# Three-dimensional Time-of-Flight Secondary Ion Mass Spectrometry Imaging of Primary Neuronal Cell Cultures

Sebastiaan Van Nuffel, MSc.

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

October 2016

# Abstract

TIME-OF-FLIGHT SECONDARY ION MASS SPECTROMETRY (TOF-SIMS) has proven its ability to characterise (in)organic surfaces, and is increasingly used for the characterisation of biological samples such as single cells. By combining ion imaging and molecular depth profiling it is possible to render 3D chemical images, which provides a novel, label-free way to investigate biological systems. Major challenges lie, however, in the development of data analysis tools and protocols that preserve the cell morphology. Here, we develop and employ such tools and protocols for the investigation of neuronal networks.

One of the reasons 3D ToF-SIMS imaging of cells is underused is the lack of powerful data analysis tools as 3D ToF-SIMS measurements generate very large data sets. To address this issue, we developed a method that allows the application of principal component analysis (PCA) to be expanded to large 3D images making 3D ToF-SIMS image processing of whole, intact cells and cellular networks with multivariate analysis now accessible on a routine basis. Using this method, we are able to separate cellular material from the substrate and can then correct z-offsets due to the cells' topography resulting in a more accurate surface heightmap. The method also facilitates differentiation between cellular components such as lipids and amino acids allowing the cell membrane, the cytoplasm and the extracellular matrix (ECM) to be easily distinguished from one another.

These developments permit us to investigate the intracellular localisation of specific native and non-native compounds label-free, not just in single cells but also in larger cellular networks. The visualisation of the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons and the chemical differentiation between cell types, namely primary rat cortical neurons and retinal pigment epithelium (RPE) cells, are presented as applications. Even though the dyes have distinct fragment ions in the high mass range, it was not possible to detect the fluorophores by 3D ToF-SIMS imaging of freeze-dried cells. However, it was possible to detect distinct differences in the kind of ions detected for freeze-dried primary rat cortical neurons

and RPE cells albeit in the low mass range.

To obtain meaningful results, however, it is paramount that sample preparation does not induce significant physical or chemical changes. We present the first comprehensive comparison between large 3D ToF-SIMS images of freeze-dried and frozen-hydrated cells using PCA to facilitate the data analysis of these large data sets. A higher degree of colocalisation of the $K^+$ signal with cell regions is observed for frozen-hydrated cells, which indicates a lower degree of membrane damage and migration of diffusible chemical species. Frozen-hydrated cell samples are therefore considered to best reflect the native cell state, but freeze-dried cell samples allow far easier sample handling. The mass spectrum of frozen-hydrated cellular material also has increased ion intensities for higher-mass fragments, which is an additional advantage, because the poor signal-to-noise ratio of molecular species with $m/z > 200$ is a major bottleneck in the advancement of ToF-SIMS imaging as a diagnostic tool.

Publications arising from this PhD research :

- Van Nuffel, S., Parmenter, C., Scurr, D.J., Russell, N.A. and Zelzer, M. "Multivariate analysis of 3D ToF-SIMS images: method validation and application to cultured neuronal networks." Analyst 141.1 (2016): 90-95.

# Acknowledgements

First and foremost, I would like to express my sincerest gratitude to my principal supervisor, Dr. Mischa Zelzer, for his continuous support during my PhD study and related research as well as for his patience, motivation and knowledge. His guidance was indispensable and I could not have imagined a better supervisor and mentor for my PhD study. I would also like to thank Dr. Noah Russell, joint supervisor, for his counsel and support.

Besides my supervisors, I would like to thank my examiners, Prof. Morgan Alexander and Dr. Nick Lockyear, for their insightful comments and questions.

My sincere thanks also go to Dr. David Scurr and Dr. Christopher Parmenter, who provided me with access to their laboratories and research facilities. Without their support it would not have been possible to conduct this research. I would also like to thank Dr. Emilia Moradi for providing me with ARPE-19 cells.

I would also like to thank fellow PhD students Michael Taylor and Nitzan Herzog for their assistance with cryo-SIMS experiments and cell plating, respectively.

I also thank the past and current members of the Functional Biointerfaces Group and the Neurophotonics Lab for the stimulating discussions and all the fun we have had in the last three years.

Last but not the least, I would like to thank my family and friends for supporting me throughout this PhD study.

# List of Figures

# List of Tables

# Contents

CHAPTER 1

# Introduction

Major advances in cell biology are tightly linked to innovations in microscopy. In fact, cell biology is innately tied to it as the cell was first discovered by Robert Hooke using a coarse, compound microscope and described in his book Micrographia published in 1665 under the auspices of the Royal Society. The subsequent discoveries of organelles such as the cell nucleus by Brown in 1833, mitochondria by von Kölliker in 1857 and the Golgi apparatus by Golgi in 1898 are all due to optical microscopy. The advent of the electron microscope, built by Ernst Ruska in 1931, led to the revelation of previously unresolvable organelles as well as the first visualisation of viruses by his brother Helmut Ruska. In 1941, Albert Coons labelled antibodies with FITC, thus giving birth to the field of immunofluorescence. Today, these techniques are the workhorses of cell biology and continue to lead to new discoveries every day.

## 1.1 3D Imaging of Cells

The goal of biological microscopy is to image life in its most natural form. As it happens, life occurs in three dimensions so to truly understand what is happening at a single cell level it needs to be viewed and analysed in 3D. There are several techniques available that allow 3D imaging of whole cells. The most widely used technique in the biological field is confocal fluorescence microscopy, where the cells are made fluorescent by labelling them with fluorescent stains [1]. However, many fluorescent stains are toxic to cells or require cells to be fixed and permeabilised for their entry into the cells [1]. Alternatively the cells can be genetically modified to produce a fluorescent protein. However, not all cells are readily transfected and one cannot assume that the transfected cell is entirely normal [1]. Even with fluorescent dyes that can enter cells, allowing the cells to be imaged non-destructively, fluorescence microscopy can

be challenging. If more than one cell type is present, it is necessary to separately label the different cells. Overlap of the broad emission spectra of conventional dyes limits the number of fluorescent probes that can be used for simultaneous imaging in a single experiment [2]. Furthermore, the autofluorescent properties of various extracellular matrix proteins, the dilution of fluorescent dyes following cell division and photobleaching also often hinder the imaging of cells [1]. Inherently, one is also limited to seeing what is labeled. There is therefore a need for label-free imaging technologies.

A less common, but still widely used technique for 3D imaging of cells is electron microscopy (EM) using serial section reconstructions. The development of serial block-face imaging SEM (SBF-SEM) and focused ion beam SEM (FIB-SEM) has led to a dramatic increase in 3D EM studies [3]. SBF-SEM uses an automated ultramicrotome located in the SEM chamber to remove thin sections ($\geq$ 20 nm) from the block-face and FIB-SEM uses a focused ion beam to mill away thin sections ($\geq$ 5 nm) of the sample surface [3]. Depending on the instrument, a lateral resolution of 2 nm can be achieved [1] with Z contrast information [4]. If equipped with energy-dispersive spectroscopy (EDS), it is even possible to obtain elemental maps [4]. Unfortunately, this is a terminal technique as the samples have to be imaged in an ultra-high vacuum. Such a vacuum is incompatible with biological samples such as cells and samples are therefore prepared by fixation and dehydration in room-temperature EM techniques or frozen in the case of cryo-electron microscopy.

A new technique is confocal Raman microscopy, which allows for non-invasive, label-free imaging of cells under normal physiological conditions [5, 6]. As Raman spectroscopy detects the vibrational frequencies of different functional groups, this technique provides chemical bond information, which can be used to discern different chemical compounds within the cells such as proteins, lipids and DNA [6]. Pascut et al. [7] exploited molecular markers specific to cardiomyocytes derived from human embryonic stem cells (hESC) in order to discriminate individual live cardiomyocytes within a heterogeneous population of other hESC-derived cells.

Finally, imaging mass spectrometry such as time-of-flight secondary ion mass spectrometry (ToF-SIMS) carries the greatest information depth as it is able to provide full molecular information. ToF-SIMS has proven its ability to characterise surfaces and coatings of (in)organic materials, and is increasingly used for biological applications [8]. The combination of label-free, molecular 2D imaging with a resolution comparable to a bright-field microscope image ($\approx$ 500 nm [9]) and the development of novel ion beams [9, 10] that allow beam sputtering with a depth resolution of 30 nm [11] makes it possible to render label-free 3D chemical images, which provides a novel way to

investigate single cells.

This generates exciting possibilities such as label-free chemical differentiation between intracellular components and differentiation between cell types in a culture. Label-free 3D chemical imaging is especially attractive for drug studies or toxicological risk assessment [12] and can visualise the cellular uptake of various non-native chemical species. It can also be used to characterise the extracellular matrix (ECM). The extracellular matrix tightly controls early tissue development and alterations of the extracellular matrix are responsible for (or accompany) cancer progression and the progression of many diseases associated with fibrosis [13, 14]. Characterizing the composition of extracellular matrices could therefore lead to the identification of novel biomarkers for these pathologies [14]. Once such disease-specific ECM proteins are identified, they could serve as anchors for imaging molecules or therapeutics coupled to anti-disease-specific ECM protein antibodies for example [14].

In the next section, an overview of the ToF-SIMS technique will be given for the unfamiliar reader in order to make the following discussion concerning the research on 3D TOF-SIMS imaging of single cells and its challenges (see section 1.3) easier to understand.

## 1.2 ToF-SIMS

Secondary ion mass spectrometry (SIMS) is a ultra-high vacuum (UHV) technique. In the case of SIMS, a solid sample is bombarded with a focussed ion beam of 'primary' ions (e.g. $Ar^+$, $O_2^+$, $Ga^+$, $Cs^+$, ...) accelerated with an energy between 1 and 25 keV. Desorption of sample ions is caused by a collision cascade induced by the impact of the primary ions on the sample surface (see Figure 1.1). The emitted 'secondary' ions, which can be both positively or negatively charged, are extracted to the mass analyser using a static electric field.

### 1.2.1 Analysis Modes

There are four analysis types possible in ToF-SIMS:

- mass spectrum (0D);
- depth profiling (1D);
- 2D imaging;

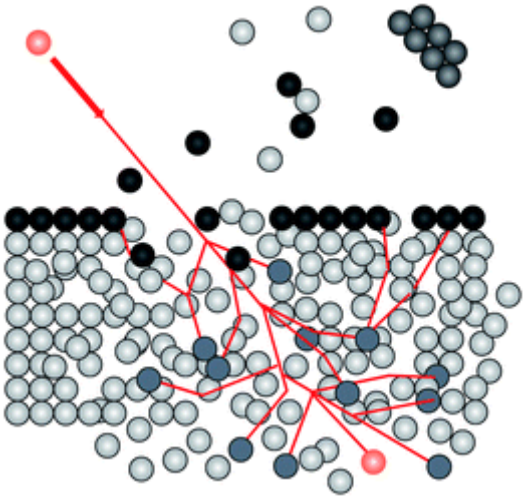


**Figure 1.1:** Collision cascade in a solid sample (grey) covered by a monolayer (black) initiated by the impact of an energetic primary ion (red). Figure obtained from [15].

- 3D imaging;

In order to generate ion images, the primary ion beam is rastered over the sample surface and a full mass spectrum is recorded for every pixel. The resolution of the images is determined by the diameter of the primary ion beam, because the pixel size needs to match the size of the beam spot. Afterward, ion images are reconstructed by showing the intensity of specific ions per pixel. The ion images visualise the distribution of components at the surface. Low secondary ion yields limit the ion intensity and therefore the contrast of the ion images.

One of the methodological problems in static SIMS is that the ion intensity is not only a function of the nature and concentration of the analyte, but also the local incidence angle of the primary ion beam and the location in the extraction field [16]. This means the detected ion intensity is dependent on the sample topography, making quantification of the surface concentration in ion images difficult.

### 1.2.2 Liquid Metal Ion Gun

A frequently used primary ion source is the liquid metal ion gun (LMIG), because of its small beam spot and high ionisation yield. A low boiling metal such as Ga for example is heated under a nozzle in a high extraction field, causing a needle shaped tip or Taylor cone to be formed and field ionisation produces the primary ions. The LMIG can be used as a cluster ion source, when for example Bi is used as a metal. The low energy per incident atom in an ion cluster limits the depth of the collision cascade and improves

both the ion yield and the depth resolution [17]. The Bi-LMIG delivers an intense beam of $Bi^+$, $Bi_2^+$, $Bi_3^+$ and $Bi_3^{2+}$. A magnetic field or Wien filter separates and selects the primary ions.

### 1.2.3 Static and Dynamic SIMS

The primary ion dose (ions/cm$^2$) is an important parameter, because it leads to two distinct modes in SIMS: static SIMS (low primary ion dose) and dynamic SIMS (high primary ion dose). The 1-25 keV energy of the primary ion, which is deposited into the sample by the collision cascade, is sufficient to break any chemical bond. The collision damage of a keV primary ion is estimated to stretch an area of several nm around its trajectory [18]. The molecular structures within this volume therefore no longer match that of the original sample [18]. By bombarding the surface with a limited number of primary ions in order to avoid hitting the same area more than once, the mass spectra retain molecular information. The so-called 'static limit' is determined empirically and is estimated at $10^{12}$ primary ions/cm$^2$ for organic samples [19]. The technique is capable of characterising the top monolayer of a solid sample with a lateral resolution of 150 nm [20]. In the case of dynamic SIMS, fast erosion of the surface occurs so depth profiling becomes a possibility, but only elemental information is obtained [21].

### 1.2.4 Gas Cluster Ion Beam

For depth profiling and 3D imaging, a second ion beam source ($C_{60}{}^+$, $Ar_n{}^+$) can be used for sputter removal of material [22]. In this case, sputtering is interleaved by 2D imaging with metal cluster beam ($Bi_n{}^+$) within the static limit. Gas phase sources are low-brightness and good spatial resolution ($\cong$ 100 nm) is only possible using small apertures. Small apertures, however, degrade the current resulting in longer acquisition times.

### 1.2.5 Time-of-Flight Mass Analyser

Time-of-flight (ToF) mass analysers are especially suited for static SIMS. Only a limited amount of secondary ions are produced so the high transmission of a ToF mass analyser is a major advantage. The inherent panoramic detection and the wide detectable mass range (from 1 to 10 kDa) are advantages as well. A ToF-MS uses the simple priciple that ions with different m/z but the same kinetic energy will separate according to their m/z in a field-free zone. An attractive quality of ToF-MS is the simple design.

Ions are accelerated in a static electric field via an electric potential difference V and all ions acquire the same kinetic energy. If the ions travel a field-free distance D to the detector , their flight time is given by:

$$t = \frac{D}{v} = D\sqrt{\frac{m}{2zeV}} \tag{1.2.1}$$

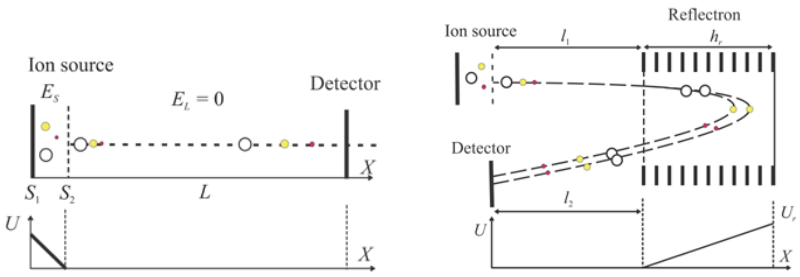
with m the mass of the ion, $v$ its velocity, z the number of charges of the ion and $e$ the electric charge of an electron. The distance between the potentials is typically about 0.5 cm and the length of the flight path D can range from 15 cm to 8 m [23]. Acceleration voltages range between 3 kV to 30 kV and the flight time amounts 100 $\mu$s during standard analysis. Time is converted to m/z by:

$$\frac{m}{z} = 2eV\left(\frac{t}{D}\right)^2 \tag{1.2.2}$$

In practice, a mass spectrum is calibrated by determining the constants a and b in the equation below based on a few known ions:

$$\frac{m}{z} = at^2 + b \tag{1.2.3}$$

A simple, linear ToF-MS has a mass resolution of about 300 to 400 [23]. Uncertainty of the time of ion formation, its place in the extraction field and the initial kinetic energy all reduce the mass resolution. To reduce the time interval of the ion formation, short ionisation pulses are required ($<$ 10 ns). In SIMS, 'passive' pulsionisation of the primary ion beam can be obtained with deflector plates. An 'active' form is bunching, where the ions of the continuous primary ion beam drift between two plates and the back reflector plate is pulsed with a repulsive potential. The ions in the vicinity of the back reflector plate are accelerated and catch up with the other ions so they arrive at the sample simultaneously. This increases the ion density in each pulse, but broadens the kinetic energy distribution, making focussing of the primary ion beam more difficult. As a result, a compromise has to be made between mass and lateral resolution. Bunching improves the mass resolution by decreasing the pulse duration without a loss in counts at the cost of the lateral resolution (1-2 $\mu$m beam spot). Sub-micrometer beam spots can be achieved in a non-bunched mode like 'burst-alignment' (300-500 nm) or the 'extreme cross-over mode' ($\approx$ 150 nm), but at the cost of mass resolution. To reduce the uncertainty of the place of ion formation in the extraction field, a flat surface is required. Surface topography therefore has a negative effect on the obtainable mass resolution [21]. To reduce the effects of a spread in initial kinetic energy, a high extraction voltage ($\approx$ 2000 V) is applied and a reflectron (ion mirror) is used (see Figure




**(a)** Linear TOF. **(b)** One-stage ion mirror.

**Figure 1.2:** Schematics of the ion optics of a linear TOF and a TOF with a one-stage ion mirror. The potentials are visualised along the ion optical axis. Figures obtained from [24].

1.2). A reflectron is a post-source compensation method for energy inhomogeneity and consists of a retarding electric field that reverses the direction of the ions. Ions with a higher kinetic energy will have a larger penetration depth so their travel distance also increases and they arrive at the same time at the detector as less energetic ions.

### 1.2.6 Flood Gun

During the primary ion bombardment not only secondary ions are emitted from the sample surface, but also an excess of electrons. This causes positive charging of dielectric samples. The change in surface potential alters the local extraction voltage, which disrupts the measurement. The electron flood gun compensates the positive charging via a beam of low energy electrons during the mass separation when the extraction voltage is absent (see Figure 1.3) [21].

### 1.2.7 Ion Formation in SIMS

In most MS techniques, ion formation occurs in the gas phase from neutrals. The transition to the gas phase and ionisation (electron ionisation, chemical ionisation, ...) are physically separated processes. SIMS, however, is a desorption-ionisation method, where ions appear to be directly emitted from solid samples through the use of a highly energetic ion beam [25]. SIMS spectra have the following characteristics:

- at high m/z, adduct ions such as $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, etc. and occasionally $M^{+\bullet}$ are found, i.e. the analytes remain intact despite the destructive energy regime;
- a large amount of fragments (depends on the used conditions, e.g. primary

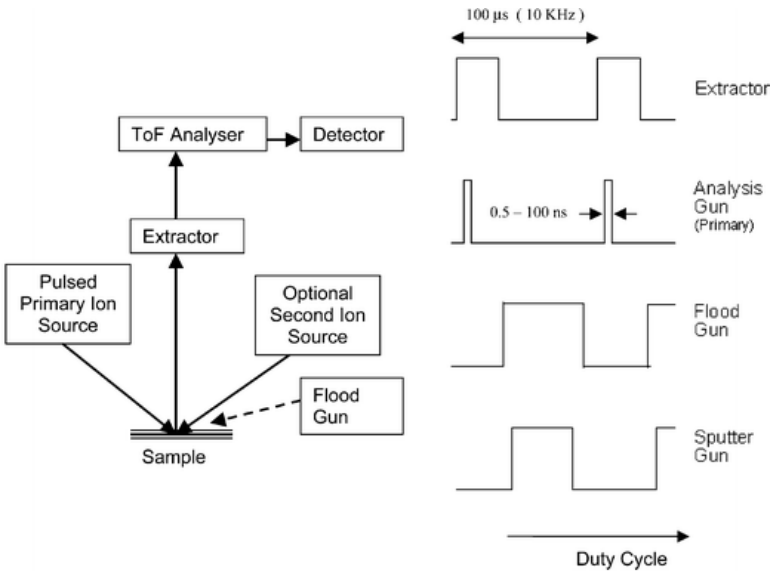



**Figure 1.3:** Schematic of the TOF-SIMS experiment. The duty cycle of the extractor and analysis gun is shown as well as those for the flood and sputter guns. Figure obtained from [21].

ions, energy, sample, ...) that are usually even-electron, but can also be uneven-electron systems;

- both organic and inorganic analytes are ionised.

The desorption-ionisation model, put forth by Cooks and Busch [26], introduces the concept of desorption by vibrational excitation. It is suggested that, regardless of the initial form of energy deposition, the energy is transformed into thermal/vibrational motion as far as the molecules are concerned. This is apparent from the similarity between spectra from S-SIMS, FAB, laser and plasma desorption. An important feature is that the desorption and ionization processes can be considered separate, i.e. there is no net creation of ions during desorption.

Alternatively, one could say the collision cascade, induced by the initial impact of the primary ion, creates a heat distribution as the kinetic energy of the atoms in the cascade can be recalculated into temperature using $E_{kin} = \frac{3}{2}Nk_BT$. In the central zone, atomisation takes place releasing radicals, element ions and electrons. Further away from the point of impact, the temperature decreases allowing fast thermal processes that gives rise to intact desorption of analytes. Some preformed ions (e.g. salts) existing at the surface prior to primary ion bombardment may be directly emitted. The majority of such ions will traverse the selvedge unperturbed and may be observed in high abundance in the spectrum; no ionization step occurs. It is suggested that neutral molecules are desorbed in high yield but to be detected must undergo an ionization step in the

gas phase of the selvedge. Even further away, thermal degradation and a long-lasting emission of for example $Na^+$ and $K^+$ occurs. In other words, the selvedge is filled with neutrals, electrons and ions.

To generate secondary ions from the neutrals, the model suggests that their desorption is followed by: i) ionization by fast ion-molecule reactions such as cationisation (i.e. adduct ionisation) in competition with ionization by the emitted secondary electrons (i.e. electron ionisation) in the gas phase of the selvedge; ii) uni-molecular fragmentation in the free vacuum, governed by the internal energy of the parent ion giving rise to the detected fragment ions. Adduct ionisation (AI) is a lower energy process than electron ionisation (EI) so most of the fragments arise from EI causing the molecular ion to disappear.

Just above the sample there is a zone of higher density (good for AI) and farther up the pressure is similar to that of the free vacuum (cf. classical gas phase EI). The density of the selvedge[1] is determined by the desorbability of the sample. In other words, a 'volatile' analyte will give rise to a larger proportion of AI.

Another important consideration is the time window of ion formation with respect to that of the used mass analyser. At first ($<$ 15 ns), a lot of $[M+H]^+$, $M^{+\bullet}$ and matching fragments are formed. Later on (0.1-1 $\mu$s), thermionic emission of $Na^+$ and $K^+$ takes off. As a TOF analyser captures the ions formed during the first 15 ns, few $[M+Na]^+$ ions are observed.

The desorption-ionisation model [26] is of course just a model and there are other models such as the nascent ion-molecule model proposed by Gerhard and Plog [27, 28]. It is also merely a qualitative model. However, it does provide an insight and helps the analyst explain the characteristics of S-SIMS mass spectra using the well-known principles of EI and AI.

## 1.3 Previous Work Regarding 3D SIMS Imaging of Single Cells

In 2004, Chandra [29] published 3D dynamic SIMS ion microscopy images of flash-frozen, fractured and freeze-dried T98G glioblastoma cells, spatially resolving elemental information in the X-, Y-, and Z-dimensions. Two years later, Kozole, et al. [30] illustrated the potential for 3D SIMS imaging of cells using model multilayer structures.

[1]The selvedge is defined as the point above which only uni-molecular fragmentation of metastable ions - generating the detected ion fragments - takes place [26].

One of the first examples of molecular 3D TOF-SIMS imaging of single cells was performed in 2007 by Fletcher, et al. [31] on freeze-dried *Xenopus laevis* oocytes ($\approx$ 1 mm in size, see Figure 1.4 a); this group has since then performed 3D TOF-SIMS imaging of freeze-dried/frozen-hydrated benign prostatic hyperplasia (BPH) cells, HeLa cells, and human cheek cells as well with their J105 - 3D *Chemical Imager* (Ionoptika Ltd., UK) [32, 33]. Also in 2007, Breitenstein, et al. [34] published 3D TOF-SIMS images of chemically fixed and air-dried normal rat kidney (NRK) cells and Nygren, et al. [35] published 3D TOF-SIMS images showing the distribution of phosphocholine, sodium and potassium ions in freeze-fractured thyroid tumor cells.

In 2008, Ghosal, et al. [36] mapped the 3D elemental distribution present within micrometre-sized, air-dried *Bacillus thuringiensis israelensis* spores with nanometer-scale spatial resolution using a NanoSIMS 50 (Cameca Instruments, Geneviers, France). In 2010, Szakal, et al. [37] investigated freeze-dried HeLa cells using a conventional TOF-SIMS instrument combined with cross-sectioning using focused ion beam milling, demonstrating in-plane resolutions of approximately 400-500 nm.

Breitenstein, et al. [38] followed with an update in 2008, where they also attempted to localise xenobiotic fluorophores. Brison, et al. [39] were able to image a non-native bromodeoxyuridine (BrdU) within chemically fixed and air-dried single HeLa cells in 2013 (see Figure 1.4 b) and Passarelli, et al. [40] similarly investigated the cellular uptake of the antiarrhythmic agent amiodarone in NR8383 cells, an immortalized cell line derived from a lung macrophage (Sprague-Dawley rat), in 2015 (see Figure 1.4 c). Graham, et al. [41] showed that ToF-SIMS 3D depth profiling is capable of localizing polymer nanoparticles within HeLa cells (see Figure 1.4 d).

Major strides have thus been made over the past decade and it has been shown that 3D ToF-SIMS imaging provides a novel way to investigate single cells. Label-free 3D chemical imaging is especially attractive for drug studies or toxicological risk assessment [12] and the research on 3D ToF-SIMS imaging of single cells has progressed to the point where the intracellular uptake and location of non-native compounds [39, 40] and nanoparticles [41] can be visualised. In this thesis, however, we strive for ToF-SIMS imaging with a large field of view containing multiple neuronal cells to eventually correlate this with other imaging techniques and multielectrode arrays (MEAs) in particular. The focus so far has been on high resolution imaging of single cells.

In spite of the progress made, there are still limitations present, which will be discussed in the following subsection.

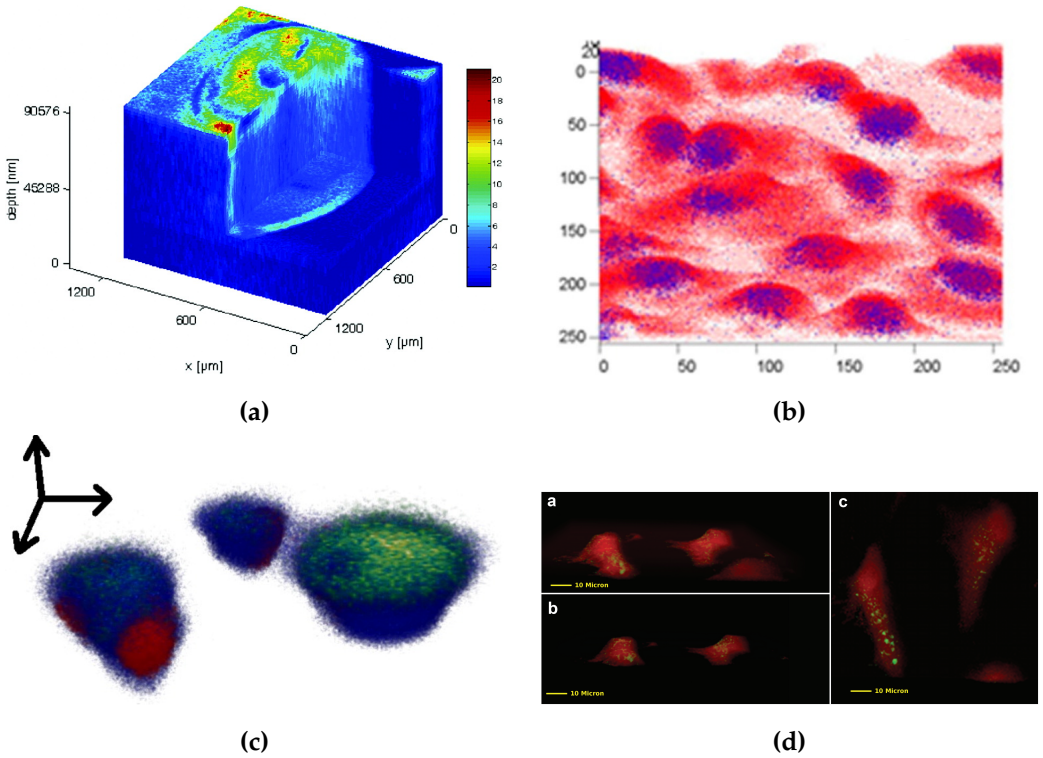
### 1.3.1 Encountered Challenges

It is observed in literature that TOF-SIMS has a few fundamental issues that challenges the data acquisition and interpretation of complex samples such as cells. One of the principal difficulties encountered is that of biological sample handling within the vacuum required for SIMS analysis [42], but this will be further discussed in section 1.4.

Furthermore, biological samples such as cells show topography. This means that 3D images created from stacked 2D images are distorted in the vertical direction. This problem has been addressed by Wagter et al. [43] who measured the topography before and after SIMS imaging using AFM and then corrected the 3D SIMS image by interpolation of the two AFM images. Breitenstein et al. [34] on the other hand vertically shifted the data points taking the substrate interface as a constant reference height and assumed a constant sputter yield in cells, demonstrating excellent agreement between AFM topography and that of the corrected 3D SIMS images. Graham et al. [44] developed a freely available MATLAB algorithm based on the latter method.

Another issue is the intrinsic incompatibility to achieve both high mass resolution and high spatial resolution. Because of the nature of the TOF mass analyser, short ($< 10$ ns) primary ion beam pulses are needed for high mass resolution. Such short pulses are achieved through bunching, which results in a less highly focused ion beam (reduced lateral resolution). Furthermore, the low duty cycle of the pulsed ion beam implicates that imaging and depth profiling experiments take a very long time.

Finally, the static limit prevents sampling more than 1 % of the surface monolayer with the analysis beam (typically a LMIG source) and the subsurface damage from the analysis beam then needs to be removed by the sputtering beam ($C_{60}^{+}$ or Ar cluster source) so most sample material is not analysed, severely limiting the sensitivity [9]. However, polyatomic ion projectiles, particularly Ar GCIBs, allow analysis at an ion fluence higher than the static limit while maintaining molecular information due to the low damage cross section and damage accumulation associated with these beams thus enabling 3D molecular imaging [45]. Additionally, ToF-SIMS data of molecular species $m/z > 500$ have a poor signal-to-noise ratio [10]. This leads to low count ion images making data interpretation even more challenging. An experimental method of increasing ionisation yields of secondary species is the current 'holy grail' in ToF-SIMS development [10]. With these limitations in mind, good data processing and analysis is of the essence [42] and will be further discussed in section 1.5.




**(a)** **(b)**

**(c)** **(d)**

**Figure 1.4:** Examples of 3D ToF-SIMS imaging of cells in the literature: a) 3D biochemical image of freeze-dried oocyte, showing the distribution of the combined signal from amino acid fragments (m/z 30, 44, 70, 120, 130, 159, 170 and 171), attributable to proteins. Reprinted (adapted) with permission from [31]. Copyright (2007) American Chemical Society. b) Z-corrected 3D overlay image of BrdU localized within cells obtained from the HMR bunched mode with the BrdU signal shown in blue and the sum of the $C_xH_yO_z$ signals in red. The image is 202 $\times$ 202 $\mu m^2$ and contains 24 slices. Reprinted (adapted) with permission from [39]. Copyright (2013) American Chemical Society. c) 3D rendering of macrophages dosed with amiodarone: amiodarone (m/z 646.0, green), lipid marker (m/z 184.1 , blue) and nuclear-marker (m/z 81.0, red). Reprinted (adapted) with permission from [40]. Copyright (2015) American Chemical Society. d) Three-dimensional reconstructed ToF-SIMS data of nanoparticles in cells. Red indicates cell and green indicates nanoparticles. Figure obtained from [41].

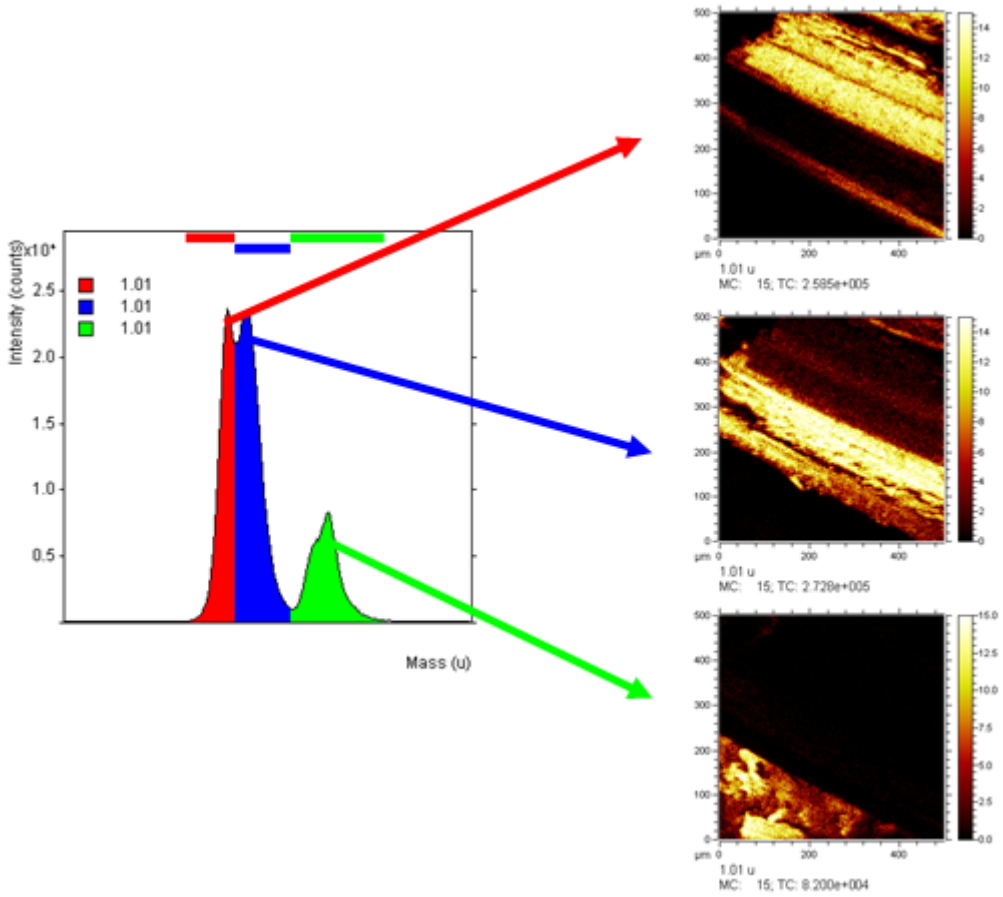



**Figure 1.5:** This ToF-SIMS data of milled, agricultural straw with extreme topography in excess of hundreds of $\mu$m illustrates the effect of topography. The $H^+$ peak is split into three separate peaks corresponding to three different heights in the sample.

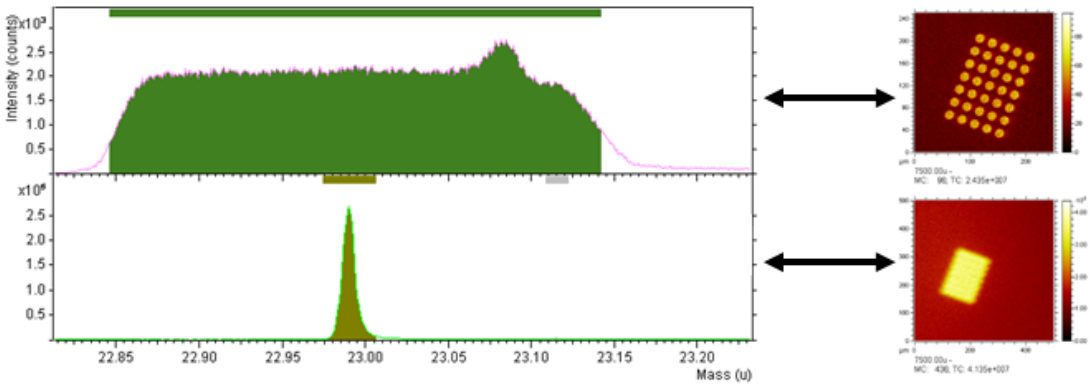



**Figure 1.6:** Trade-off between mass resolution and lateral resolution illustrated with 16 $\mu$m Au microarrays with 24 $\mu$m spacing on a glass substrate. Top: The extreme cross-over mode allows for high lateral resolution, but low mass resolution as illustrated by the broad $Na^+$ peak. Bottom: The high current bunched mode provides much better mass resolution, but sacrifices lateral resolution as illustrated by the fuzzy total ion image.

## 1.4 Sample Preparation Methods

Because ToF-SIMS is an ultra-high vacuum technique, analysis of biological materials such as cells in a close-to-native state pose a formidable challenge due to the high water content of the samples. Consequently, sample preparation is a crucial aspect for ToF-SIMS analysis of biological samples. To obtain meaningful results, it is paramount that sample preparation does not induce significant physical changes in the sample. In this area much can be learned from the field of electron microscopy (EM) that has been developing methods to measure biological samples as far back as the 1950's [46]. In addition, ToF-SIMS analysis also requires that the sample preparation procedure does not cause any significant chemical changes.

Most cell preparation techniques for ToF-SIMS (and EM) involve dehydration of the sample. Commonly used procedures are cryofixation or chemical fixation followed by freeze-drying, alcohol drying or air-drying [47–49]. A major advantage of these techniques is that the sample can be kept at room temperature, allowing convenient sample handling and the use of complementary techniques to visualize and select well-preserved cells and to measure their topography prior to analysis to allow height correction of the ToF-SIMS data.

Fixation is necessary to bring about rapid cessation of cell activity, to immobilize cellular components and to en-able the sample to withstand further processing procedures [50]. With cryofixation, it is imperative to freeze the cells rapidly enough so that they are vitrified (i.e. only contain amorphous ice) as the formation and growth of ice crystals during the freezing process may dehydrate surrounding cytoplasmic areas, displace cellular structures or even penetrate membranes [50]. Vitrification of aqueous solutions is estimated to take place at cooling rates in excess of about $3 \times 10^6$ K/s [51]. This can be achieved by plunge freezing: immersing the sample in a suitable liquid cryogen such as liquid ethane or propane [52]. The maximum sample depth where sufficiently rapid freezing occurs is only 10-15 $\mu$m. Cells below this top layer typically experience the formation of ice crystals that are large enough to distort the cell structure, essentially restricting cryofixation to thin films [50]. Chemical fixatives will penetrate at least 0.5 mm into a sample, but the fixation is much slower and can for that reason be inferior to cryofixation [50] in terms of rapid cessation of cell activity and immobilization of cellular components. A study by Malm et al. [48] compared chemical fixation and cryofixation as cell preparation techniques for ToF-SIMS and showed that the distribution of diffusible ions such as $K^+$ is not retained in cells that were chemically fixated and that they have less well preserved membrane integrity. They also showed that alcohol drying specifically reduces phosphocholine secondary ion yields and in-

ferred that the polar membrane phospholipids are extracted from the cell membrane by ethanol. Cryofixation followed by freeze-drying clearly constitutes the best dehydrated cell preparation method with regard to maintaining the cells' physical structure [48, 53].

In the aforementioned cell preparation techniques, the dehydration step induces drastic compositional changes in the sample due to the complete removal of volatile components. This subjects the cells to mechanical stresses not present in the hydrated state which can lead to undesired sample alterations such as membrane damage [48, 54–56]. These issues may be overcome by imaging the cell in its native hydrated state by vitrifying the cells and performing the ToF-SIMS measurement with a cryogenically cooled stage at a temperature below which recrystallization does not occur ($<$ -120 °C [57]). This is referred to as cryo-ToF-SIMS, a conceptually analogous approach to cryo-EM.

Cryo-ToF-SIMS can either be performed on freeze-fractured or frozen-hydrated samples. The freeze-fracture technique consists of physically breaking the frozen sample and subsequent surface analysis of the exposed fracture plane. However, the position of the fracture plane within the cell varies [58] and generally cells tend to fracture between the leaflets of the membrane bilayer [59], making intact cellular characterization difficult, especially on the cell surface.

Frozen-hydrated measurements keep the cells intact, but pose the greatest challenge for sample handling and measurement [60]. Rabbani et al. [61] showed that dried cells (both after chemical fixation and cryofixation) have a more diffuse distribution of the m/z 184 mass fragment of the phosphocholine head group compared to freeze-fractured cells. The observation that the drying process causes molecule rearrangement (seen as more diffuse spatial distributions) was also made by Lanekoff et al. [58] when 2D imaging freeze-fractured cells and subsequently freeze-drying them inside the instrument by raising the temperature in the chamber from -115 °C to -70 °C. It is clear that the risk of membrane damage and chemical redistribution as observed with drying are minimized when measuring a cell in a frozen-hydrated state. Lanekoff et al. also observed a decrease in ion intensity for certain mass fragments attributed to phosphatidylcholine (m/z 184), phosphatidylethanolamine (m/z 124 and 142) and sphingomyelin (m/z 184 and 224) after freeze-drying compared to the frozen-hydrated state. It has been hypothesized that this difference in ion yields could be due to the abundant proton sources in the matrix of frozen aqueous samples [62] as well as the decreased sample temperature [63]. In spectra from frozen-hydrated cell surfaces compared to freeze-dried ones, Robinson et al. [64] observed no differences in ion yields for most low mass ions, but an increased yield for higher mass fragments. This is consistent with

**Table 1.1:** Cells in the literature investigated using SIMS 3D imaging and their preparation.

| Cell | Preparation | Reference |
|---|---|---|
| *Xenopus laevis* oocyte | freeze-dried | [31] |
| HeLa | frozen-hydrated | [32, 33] |
| | freeze-dried | [33, 37] |
| | chemically fixed and air-dried | [39] |
| prostatic hyperplasia (BPH) cells | frozen-hydrated | [32] |
| human cheek cells | freeze-dried | [32] |
| normal rat kidney (NRK) cells | chemically fixed and air-dried | [34, 38] |
| epithelial MDCK II cells | chemically fixed and air-dried | [38] |
| thyroid tumor cells | freeze-fractured | [35] |
| | chemically fixed and air-dried | [65] |
| pig thyrocyte cells | fractured freeze-dried | [65] |
| *Bacillus thuringiensis israelensis* spore | air-dried | [36] |
| T98G glioblastoma cell | fractured freeze-dried | [29] |
| NR8383 cells | freeze-dried | [40] |
| NIH/3T3 fibroblasts | air-dried, freeze-dried and frozen-hydrated | [64] |

the reduced damage cross section mechanism produced by analysis at cryogenic temperatures, which would reduce fragmentation of larger ions. This increased yield for higher mass fragments potentially demonstrates an additional advantage of studying biological samples in a frozen-hydrated state.

Table 1.1 gives an overview of the different cell types and preparation methods investigated using SIMS 3D imaging.

## 1.5 Data Processing and Analysis

Secondary ion mass spectra are notoriously complex to interpret and informatics methods have had very little progress in the SIMS field [9]. For this reason, the SIMS community adopted multivariate analysis techniques early on, but a major issue for 3D imaging is that the data sets are too large for analysis using modern desk top computers [9].

One multivariate technique of special interest for (3D) TOF-SIMS imaging is principal component analysis (PCA). Other multivariate analysis methods that have been applied to ToF-SIMS data include partial least squares (PLS) [66], multivariate curve

resolution (MCR) [67, 68], maximum auto-correlation factors (MAF) [69, 70] and neural networks (NN) [71]. Graham et al. [72] provide an excellent summary of ToF-SIMS studies that have been carried out using at least one MVA method. PCA, however, is the most widely applied method, because of its established history and wide availability [69]. PCA also does not assume an underlying causal model and thus forms an excellent starting point prior to applying many of the other methods to process ToF-SIMS data [72]. In the following subsections the concept of PCA is explained and its applicability to (3D) ToF-SIMS image data is discussed.

### 1.5.1 Principal Component Analysis (PCA)

Because the sheer volume of multivariate data makes it difficult to see patterns and relationships, the aim of many methods of multivariate analysis is data reduction [73]. If there is some correlation between the variables, however, some of the information is redundant. Principal Component Analysis (PCA) is a technique that allows the variables in a data set to be reduced to a few, interpretable linear combinations of those variables, if there is correlation present. It was invented by Karl Pearson in 1901 [74], but its use as a multivariate analysis method was developed in the field of psychometrics in the 1930s by Harold Hotelling [75, 76]. Several decades later natural scientists caught on and PCA as a chemometric tool became widespread in the 1970s [77].

Consider a data matrix $\mathbf{X}$, where each of the n rows represents a different repetition of the experiment and each of the p columns represents a particular kind of variable that has been 'mean-centered' (see equation 1.5.1). Mean-centered means the columns of $\mathbf{X}$ have sample mean equal to zero[2].

$$\mathbf{X} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \mathbf{x_3} \\ \vdots \\ \mathbf{x_n} \end{pmatrix} = \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,p} \\ x_{3,1} & x_{3,2} & x_{3,3} & \cdots & x_{3,p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & x_{n,3} & \cdots & x_{n,p} \end{pmatrix} \tag{1.5.1}$$

PCA transforms the original multi-dimensional dataset $\mathbf{X}$ to a new dataset $\mathbf{T}$, called the principal components scores. The principal components decomposition of $\mathbf{X}$ can therefore be given as:

[2]When performing a principal components analysis by eigendecomposition of the covariance or correlation matrix, mean-centering is somewhat superfluous as the covariance matrix automatically centers the data.

$$\mathbf{T} = \mathbf{XW} \tag{1.5.2}$$

where $\mathbf{W}$ is an orthogonal p-by-p matrix, called the principal components loadings, which describe the relation between the original and new datasets.

Every p-dimensional column of weights $\mathbf{w_k}$ converts the original dataset $\mathbf{X}$ in a vector of principal component scores $\mathbf{t_k}$, given by

$$\mathbf{t_k} = \mathbf{X w_k} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \mathbf{x_3} \\ \vdots \\ \mathbf{x_n} \end{pmatrix} \begin{pmatrix} w_{1,k} \\ w_{2,k} \\ w_{3,k} \\ \vdots \\ w_{p,k} \end{pmatrix} \tag{1.5.3}$$

in such a way that these new variables of $\mathbf{T}$ successively inherit the maximum possible variance from the original dataset $\mathbf{X}$ with each loading vector w constrained to be a unit vector ($\mathbf{w_k'w_k} = 1$) and being uncorrelated with the previous loading vectors. It can be shown that the loading matrix $\mathbf{W}$ consists of the eigenvectors of $\mathbf{X^TX}$, which is proportional to the covariance matrix $\mathbf{Q}$ of the original dataset $\mathbf{X}$.

$$\mathbf{Q} \propto \mathbf{X^TX} = \mathbf{W\Lambda W^T} \tag{1.5.4}$$

with $\mathbf{\Lambda}$ the diagonal matrix of eigenvalues $\lambda_k$, which are equal to the sum of the squares over the dataset associated with each component k. The amount of variation explained by a principal component is thus defined as the eigenvalue for that component divided by the trace of the eigenvalues matrix tr($\mathbf{\Lambda}$).

There are in principle as many principal components as there are original variables and they describe the same information as the original data; the total variance is the same. However, the principal components, unlike the original variables, are not correlated with each other[3]. They are also chosen so that the first principal component (PC1) accounts for most of the variation in the data set, the second (PC2) accounts for the next largest variation and so on. Hence, when there are significant correlations between the original variables, the number of useful PCs L is much less than the number of original variables [73]. Therefore, we only need to retain the first L principal components, produced by using only the first L loading vectors. On the one hand, we want L to be as small as possible in order to obtain the simplest possible interpretation. On the other

[3]It is possible to obtain an oblique solution with correlated components, but we will only discuss orthogonal solutions [78].

hand, it is important not to lose information so the proportion of variation explained by the first L principal components needs to be as close to one as possible:

$$\frac{\lambda_1 + \ldots + \lambda_L}{tr(\mathbf{\Lambda})} \cong 1 \tag{1.5.5}$$

Alternatively, a scree plot can be used to determine the number of principal components to consider. A scree plot displays the eigenvalues in descending order versus the number k of the component. The useful number of components L is chosen at the point beyond which the remaining eigenvalues are all small and comparable.

### 1.5.2 Interpretation of the Principal Components

Interpretation of the principal components is based on finding the variables which are most strongly correlated with each component. If the coefficients given by the loading vectors have a high absolute value, there is a strong correlation between component and variable. Which values are considered sufficiently large is of course a subjective decision. Variables whose coefficients have the same sign are positively correlated and variables whose coefficients have the opposite sign are negatively correlated, but the sign itself is not important.

The values of the principal component scores of each experiment can now be calculated and will be dependent on the values of the original variables. If the score value of a sample is very high for a particular principal component, you would expect to find high values for the variables with strong correlation. The scores are usually visualised in scores plots. A scores plot is a scatter plot of the scores of two principal components and may reveal that the experiment repeats can be clustered into distinct groups, which were not readily apparent from the original data.

At this point, it is worth noting that PCA does not assume an underlying causal model in contrast to factor analysis for example [73]. It is simply a variable reduction procedure that allows the variables in a data set to be reduced to a few, interpretable linear combinations that account for most of the variance. The primary purpose of PCA is descriptive.

### 1.5.3 PCA by Eigendecomposition of the Correlation matrix

PCA by eigendecomposition of the covariance matrix will tend to give more weight to variables with higher sample variances in the first few principal components. This means that the PCA is sensitive to the scaling of the variables. This also means that

PCA is somewhat arbitrary, when variables have different units of measure. Scaling the data so as to give all variables unit variance can be done by standardising the data or simply by carrying out the PCA by eigendecomposition of the correlation matrix instead of the covariance matrix. For standardised data, the total variance is equal to the number of variables: $\mathrm{tr}(\mathbf{\Lambda}) = p$. Principal component analysis should only be used with non-scaled data, if all variables have the same unit and only if you wish to give variables with higher variances more weight.

### 1.5.4 Applying PCA to ToF-SIMS Data

Because principal component analysis is a powerful tool for reducing the number of observed variables in a data set into a smaller number of principal components that account for most of the variance, PCA has been applied to a very large number of analytical methods and problems in recent years. In the case of ToF-SIMS images, the mass peaks can be regarded as variables and each pixel of the image can be considered an individual observation or sample. If we perform PCA on ToF-SIMS images, we could reduce the hundreds of ion images generated from each mass peak to a few, interpretable images of the principal component scores. However, we must first consider the applicability of PCA for (3D) TOF-SIMS image data and review if the properties of (3D) TOF-SIMS images comply with the assumptions underlying PCA.

Because a principal component analysis is performed on a matrix of Pearson correlation coefficients, the data should satisfy the assumptions for this statistic [78]:

- Interval-level measurement:
  All analyzed variables should be assessed on an interval or ratio level of measurement [78]. The numerous mass peaks in a ToF-SIMS spectrum (approx. 300 different $m/z$) are of course (continuous) ratio data.

- Random sampling:
  Each sample will contribute one score on each observed variable and should represent a random sample drawn from the population of interest [78]. Principal component analysis is also a large-sample procedure so there is a question of sampling adequacy. As a rule of thumb [79], it is suggested that there should be 10 samples for each variable and 50 should be added to that to ensure sufficient sample size for small sets of variables[4] : $n \geq (10 \times p) + 50$. ToF-SIMS images luckily are 'hyperspectral' and each pixel equates to a full mass spectrum. A $256 \times$

[4]This rule regarding the number of subjects per variable constitutes a lower bound, however, and larger samples may be required under less optimal conditions.

**Table 1.2:** Maximum number of m/z a mass spectrum can contain for 2D ToF-SIMS image sizes to be analysed using PCA, given that the sample minimum for PCA n $\geq$ (10 $\times$ p) + 50 with p the number of variables (m/z).

| Image Size (pixels) | Max. Variables |
|---|---|
| 8 x 8 | 1 |
| 16 x 16 | 20 |
| 32 x 32 | 97 |
| 64 x 64 | 404 |
| 128 x 128 | 1633 |
| 256 x 256 | 6548 |
| 500 x 500 | 24995 |

256 $\times$ 150 voxel 3D ToF-SIMS image for example represent 9,830,400 individual mass spectra (samples). In table 1.2 the maximum number of m/z a mass spectrum can contain to be analysed using PCA is given for different 2D ToF-SIMS image sizes.

Given that a typical mass spectrum does not exceed more than 1633 different m/z, it is save to assume that any 3D ToF-SIMS image with an xy-size $\geq$ 128 x 128 can be analysed using PCA.

- Linearity:
  Obviously, the variables can't be uncorrelated or PCA would not be useful, but the relationship between all observed variables should also be linear [78]. Linearity is PCA's one stringent assumption[5]. The mass peak intensities in ToF-SIMS are correlated with each other, of course, but they are not always perfectly linear. Detector saturation can result in non-linear intensity variations [72]. Therefore, care should be taken to avoid detector saturation during data collection. Dead time corrections (using Poisson statistics) can also be used to reduce such non-linearity issues [81]. A more fundamental issue is the fact that the emission probability of any secondary ion in SIMS depends on its environment. This is known as the matrix effect and complicates quantitative analysis in static SIMS [82–84]. It is still poorly understood and has defied any description beyond phenomenology for many years although some progress has been made using good reference materials [85].

[5]Methods for performing a nonlinear form of PCA have been explored. This has been termed kernel PCA [80].

- Normality:
  Each observed variable should be normally distributed[6] and each pair of observed variables should display a bivariate normal distribution, i.e. they should form an elliptical scattergram when plotted [78]. Because ToF-SIMS counts ions, its data will be governed by 'counting statistics' for which the Poisson probability distribution is the appropriate model [72, 87]. For large numbers of events $\lambda$, Poisson probabilities approach a normal distribution thanks to the central limit theorem. In other words, when the ion count is high, Poisson noise or 'shot noise' becomes indistinguishable from Gaussian noise. This is unfortunately not often the case, especially when considering individual pixels in ToF-SIMS images, so shot noise is typically dominant. However, the Pearson correlation coefficient is robust against violations of the normality assumption, when the sample size is large [78].

When performing PCA using the covariance matrix there is no explicit assumption about the distribution of the data as the sample covariance matrix is unbiased. There is, however, an implicit assumption of normality, because the only zero-mean probability distribution that is fully described by the variance ('sufficient statistics') is the normal distribution [88]. For multivariate normal distributions, zero covariance between components implies independence. PCA therefore works extremely well with normally distributed data. But if the data is not normal, higher order statistics beyond variance are not taken into account by PCA and orthogonal components are not necessarily independent[7]. However, PCA is robust to marginal deviations from normality [88].

In summary, PCA is applicable to (3D) ToF-SIMS image data. The variables (mass peaks) are continuous and by regarding each pixel/voxel as an individual mass spectrum we readily achieve adequate sampling. The Poisson nature of ToF-SIMS data is not an issue as long as the sample size is large. Provided there are no significant matrix effects, the relationship between all observed variables should also be sufficiently linear.

In the literature, it is quite common to find examples of PCA performed on 2D TOF-SIMS images, but so far only Fletcher [33] performed PCA on a 3D TOF-SIMS image (with a size of $256 \times 256 \times 10 = 655{,}360$ voxels). One reason for this is the fact that the

[6] Variables that demonstrate marked skewness or kurtosis may be transformed to better approximate normality [86].

[7] The class of algorithms that attempt to find the components that satisfy this statistical constraint are termed independent component analysis [89].

eigendecomposition needed for PCA is computationally heavy [9].

## 1.6 Other Techniques

In the following subsections some of the other instrumental techniques that are used in this thesis are discussed.

### 1.6.1 Cryoscopic Osmometer

A cryoscopic osmometer measures the freezing point of a solution to determine the total osmolality (Osm/kg). The osmolarity (Osm/L) of a solution can be calculated as follows:

$$\text{Osm/L} = \sum_i \phi_i n_i C_i \tag{1.6.1}$$

with $\phi$ the osmotic coefficient of the solute (between 0 and 1 where 1 indicates 100% dissociation), $n$ the number of ions into which the solute dissociates, $C$ the molar concentration of the solute and $i$ the identity of a particular solute. For example, a solution of 1 mol/L NaCl corresponds to an osmolarity of 2 Osm/L.

Adding a solute to a solvent decreases the freezing point of the solvent. A cryoscopic osmometer measures the freezing point of an (aqueous) solution and then uses a calibration value to convert it to the total osmolality (Osm/kg). Before the total osmolality of sample solutions can be measured, the osmometer must be calibrated, e.g. with distilled water and a calibration solution. The normal osmolality of extracellular fluid is 280-295 mOsm/kg.

### 1.6.2 Freeze-drying

Freeze-drying or lyophilisation is a dehydration process where the hydrated material is brought around the triple point of water (T = 273.16 K and P = 611.73 Pa, see Figure 1.7) by freezing it and then reducing the surrounding pressure below that of the triple point. This allows the solid ice to sublimate directly into water vapour, avoiding the liquid-gas transition. Freeze-drying generally causes less deformation than other dehydration methods and chemical compounds remain unchanged. One of the major advantages is that the samples can be stored and handled at room temperature. The TOF-SIMS measurements can also be executed without the need for a cryostage. However, other volatile compounds beside water are removed as well.

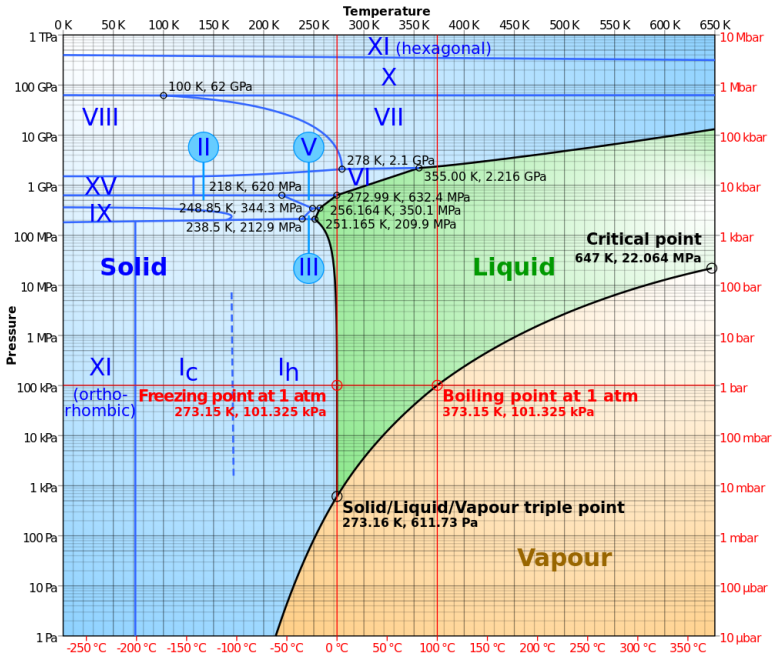



**Figure 1.7:** Phase diagram of water as a log-lin chart with pressure from 1 Pa to 1 TPa and temperature from 0 K to 650 K. Figure obtained from Cmglee (`https://commons.wikimedia.org/wiki/File:Phase_diagram_of_water.svg`).

Before freeze-drying, lyoprotectants can be added to protect the freeze-dried material from forced hydrogen bonding with other molecules after the loss of water [90]. Lyoprotectants are typically sugars such as trehalose and sucrose.

In the freezing step, it is important to cool the material well below the temperature of the triple point to avoid melt-back before the drying step. In the case of cells, it is also important to use a high cooling rate to avoid the formation of large ice crystals, which will cause damage to the cell walls [91]. This flash freezing is done by submerging the sample in a cryogenic liquid such as liquid nitrogen or liquid ethane. Liquid nitrogen has the advantage of being inflammable and readily available, but the Leidenfrost[8] effect reduces the mean cooling rate to $0.5 \times 10^3$ K$^{-1}$, whereas the mean cooling rate of liquid ethane equals $13 - 15 \times 10^3$ Ks$^{-1}$ [92]. To avoid the formation of microscopic ice crystals, the cooling rate should be at least $10 \times 10^3$ Ks$^{-1}$ [92].

Next, the pressure is lowered below that of the triple point, and enough heat is supplied (either from the environment or controlled via a heating element) to the material for the solid ice to sublime. Higher temperatures and lower pressures increase the sublimation rate, which can be described by the Hertz-Knudsen equation [93]. Apart from a vacuum system and (optional) heating element, a freeze-dryer will have a con-

[8]The Leidenfrost effect is the physical phenomenon where a liquid in contact with a surface significantly warmer than the liquid's boiling point, produces an insulating vapour layer.

denser plate (at a temperature of approximately -50 to -60 °C) for the water vapour to re-solidify on, preventing it from reaching the vacuum pump.

### 1.6.3 Coherence Scanning Interferometry

Coherence scanning interferometry is an optical method for areal surface topography measurements that relies on visible-wavelength light (white light). It is extensively used for surface metrology.

Interferometry is based on the wave superposition principle, which states that when two or more propagating waves with the same frequency traverse the same space, the resultant amplitude at each point is equal to the vector sum of the amplitudes of the individual waves. When two waves are in phase, i.e. the phase difference between the waves is a multiple of $2\pi$, they will undergo constructive interference. When two waves are out of phase, i.e. the phase difference is an odd multiple of $\pi$, they will undergo destructive interference. In an interferometer using a single light source, a beam splitter is typically used to split a wave into two parts, which travel different paths. The two waves are then combined to be interfered. The Michelson interferometer and the Mirau interferometer are two examples of such amplitude-division systems that are commonly used in designs of interferometer objective lenses used for a coherence scanning interferometer [94]. The main difference between them is the position of the reference mirror.

A coherence scanning interferometer is constructed as follows (see Figure 1.8). Light is emitted by a broadband white light source. A beam splitter, integrated inside the interferometric objective, splits the light beam into a reference and a measurement beam. The reference beam is reflected by the reference mirror and the measurement beam is reflected at the sample surface (or at an internal interface in the case of transparent layers). The two returning beams are relayed by the beam splitter to a charge coupled device (CCD) image sensor, and form an interference pattern that is spatially sampled by the individual CCD pixels [94]. Due to the incoherent nature of the white light, the optical path length to the sample and the reference must be almost identical for interference to be observed [94]. This results in an image showing interference fringes following the equal-height lines on the sample. As the interferometric objective is actuated in the vertical direction, a change of intensity due to interference will be observed for each pixel when the path length of the two beams are the same. Therefore, the z-value for the point on the surface imaged by this pixel corresponds to the z-value of the interferometric objective, when the modulation of the correlogram (interference signal) is greatest. With this information a matrix with the height values of the object surface is

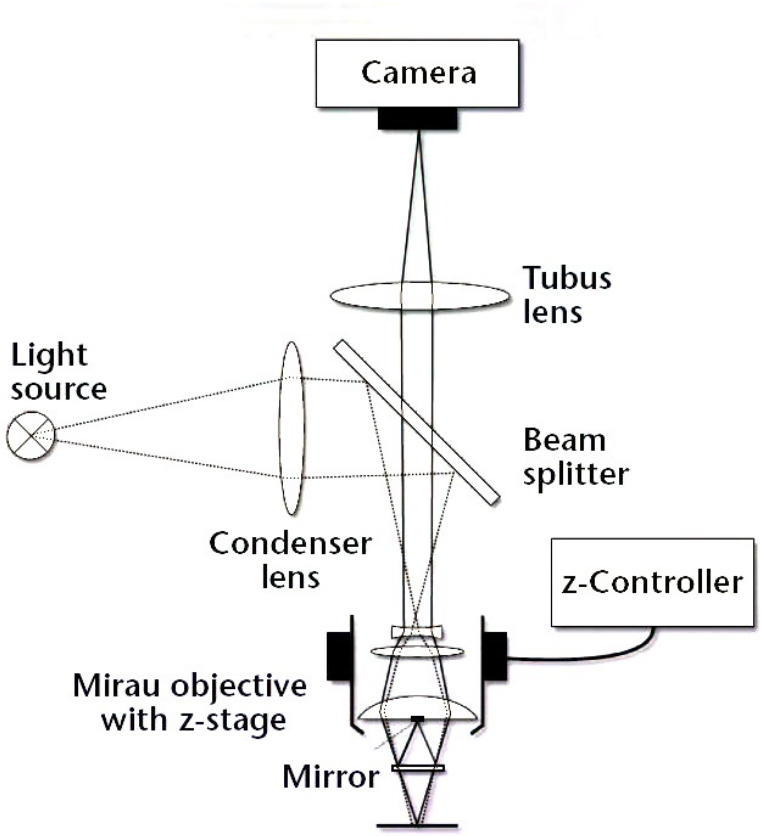



**Figure 1.8:** Schematic layout of an interference microscope with Mirau objective. Figure obtained from Polytec GmbH (`https://commons.wikimedia.org/wiki/File:Interferenzmikroskop_Aufbau_sw.jpg`).

produced, which is then converted into a three-dimensional map of the sample surface with a vertical resolution down to 0.1 nm.

A quantitative description of the areal height deviation of the surface topography is often needed to better understand the morphology of a surface. The amplitude parameters are the principal parameters in characterizing the surface topography and the most common amplitude parameters are the average roughness ($R_a$), the root mean square roughness ($R_q$) and the total roughness ($R_t$) [94, 95].

The average roughness ($R_a$) is the arithmetic mean of the absolute height, and is given by:

$$R_a = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z}) \tag{1.6.2}$$

with $n$ the number of pixels and $z_i$ the $i^{th}$ pixel's height value.

The root mean square roughness ($R_q$) is the root mean square value of the ordinate values, and is given by:

$$R_q = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})^2} \tag{1.6.3}$$

The total roughness ($R_t$) is defined as the sum of the maximum peak height ($R_p$) and maximum valley depth ($R_v$) values:

$$R_t = R_p + R_v = max(z) - min(z) \tag{1.6.4}$$

These parameters can be useful independently, but describe surface topography more comprehensively when used in conjunction. For example, the $R_q$ value and the $R_t$ value are both needed to determine whether apparent roughness is due to isolated features or the overall surface roughness [94, 95].

## 1.7 Aims and Objectives

Proof-of-concept experiments have clearly demonstrated the unique capabilities of Time-of-flight secondary ion mass spectrometry (ToF-SIMS) for biological analysis [8]. By combining ion imaging and molecular depth profiling it is possible to render 3D chemical images, which provides a novel way to investigate biological samples such as single cells. An investigation of the literature revealed that there still lie major challenges in the development of data analysis tools and protocols that preserve the cell morphology [42].

The aim of this work is to assess PCA as a data analysis tool for 3D TOF-SIMS images and to establish a cell preparation protocol for TOF-SIMS that preserves the cell morphology.

The first is done in chapter 3, where we expand the application of PCA to large 3D images without requiring any computing resources beyond a desk top computer. We also present a method to construct TOF-SIMS image data consisting of multiple chemistries with spatial distributions of their own to identify the effects of different preprocessing procedures.

In chapter 4, the effectiveness of this approach to PCA is tested on 3D ToF-SIMS images of the more complex, biological sample consisting of primary rat cortical neurons. The sample preparation consists of cryofixation followed by freeze-drying (cf. reference [48]). The technique must also prove itself in answering biological questions and, as a first application, we attempt to visualise the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons. In second application, we

attempt to differentiate between two different cell types: primary rat cortical neurons and retinal pigment epithelium (RPE) cells. RPE cells are densely packed with melanin granules. It might therefore be possible to differentiate RPE cells from neuronal cells based on the presence of melanin mass fragments.

To obtain meaningful results, however, it is paramount that sample preparation does not induce significant physical or chemical changes. In chapter 5, the first comprehensive comparison between large 3D ToF-SIMS images of freeze-dried and frozen-hydrated cells is presented using PCA to facilitate the data analysis of these large data sets.

CHAPTER 2

# Experimental Details

## 2.1 Cell Substrate Preparation

Borosilicate glass coverslips (631-1570 - VWR) were solvent cleaned with methanol (HPLC), acetone (LRG) and isopropanol (HPLC). All solvents were purchased from Fisher Scientific and used as obtained. After each solvent rinse, the coverslips were blow-dried with compressed air. Then, adhesive polyester microscope slide-grids with 1 mm line spacing (Z688533 - Sigma) were cut to fit the glass coverslips and stuck onto the glass in order to facilitate locating the same cells with different techniques. The coverslips were then placed in a vacuum oven (Thermo Scientific Vacutherm 6025) at 60 °C and 50 mbar for 1 h and rinsed with isopropanol (HPLC) again before being stored in a vacuum desiccator. Before the cell plating, the side of the coverslips without the grid is aseptically coated with 0.01 % (w/v) poly-L-lysine (P4707 - Sigma) in water. After 40 minutes, the poly-L-lysine solution is removed by aspiration and thoroughly rinsed with sterile tissue culture grade water (W3500 - Sigma). The coverslips were allowed to dry for at least 12 hours before introducing cells and medium.

## 2.2 Cell Culture

Low-density cultures of disassociated embryonic rat cortical neurons were prepared and cultured in serum-free media on the poly-L-lysine coated glass coverslips. All experiments involving animals were performed according to the UK Home Office regulations in accordance with the Animals (Scientific Procedures) Act 1986 and were approved by the University of Nottingham's animal welfare committee. Cortices were dissected from rat embryos at day E18 of gestation and incubated in 2 ml of HEPES-buffered Hanks' balanced salt solution (HBSS), containing 1 % (w/v) trypsin (T1426

- Sigma) and 0.005 % (w/v) DNAse (D5025 - Sigma) for 20 min at 37 °C.[1] After that, 0.05% (w/v) trypsin inhibitor (T9003 - Sigma) was added to de-activate the trypsin. The cortices were then rinsed with 1.5 mL of Neurobasal medium and gently triturated using a fire-polished glass Pasteur pipette in the presence of 0.005 % (w/v) DNAse. The resulting cell suspension was then centrifuged and the pellet re-suspended in Neurobasal media containing B27 supplement (17504-044 - Invitrogen), 0.5 mM L-glutamine (G7513 - Sigma) and 1 % (v/v) of Penicillin-Streptomycin (P0781 - Sigma). The prepared coverslips were placed in 6-well tissue culture plates and plated with a total of 150,000 disassociated cells per coverslip. After incubating for 30 min at 37 °C to allow cell adhesion to occur, the wells were flooded to a total volume of 2 mL of Neurobasal/B27 media. After 24 h, the media was replaced with fresh media and the cultures were maintained at 37 °C in a humidified 5 % $CO_2$ atmosphere for the next 9 days. N.B. The (primary) embryonic rat cortical neurons are non-dividing.

### 2.2.1 Cell Staining

For fluorescence staining, 100 $\mu$L of a 2 $\mu$M calcein-AM and 4 $\mu$M propidium iodide solution in PBS was added to each well (04511 - Sigma) and incubated for 15 minutes before viewing the labelled cells under the fluorescence microscope. As soon as possible (less than 40 minutes) after adding the working solution, labelled cells were cryofixated and freeze-dried.

### 2.2.2 ARPE-19

ARPE-19 cells (the retinal pigment epithelium) was purchased from American Type Culture Collection (ATCC, USA) and was cultured in Dulbecco's modified Eagle's medium (DMEM) with F12 supplemented with penicillin (100 U/mL), streptomycin (0.1 mg/mL), amphotericin (0.25 $\mu$g/mL), glutamin (200mM) and Fetal Bovine Serum (FBS, 10% v/v, Sigma-Aldrich, UK).

ARPE-19 cells were then seeded on PLL-coated coverslips at $10^5$ cells/cm$^2$ in Neurobasal/B27 media.

[1] Trypsin is necessary to dissociate the dissected cells and obtain a single cell suspension. However, cell surface proteins are often cleaved due to the proteolytic activity of trypsin and there is some evidence that that this can lead to dysregulation of the cell functions [96].

## 2.3 Cell Preparation

Before being flash-frozen by plunging the coverslips into 20-30 mL of liquid ethane for 20 s, the cell covered coverslips were dipped three times with a pair of tweezers in an ammonium formate (516961 - Aldrich) solution that matched the osmolality of the cell culture media as determined with a cryoscopic osmometer (Osmomat 030). The ethane was liquefied by letting a stream of ethane gas condensate into a 50 mL Falcon tube held in a bath of liquid nitrogen. The samples were stored in liquid nitrogen until the freeze-drying process ($<$ 1h) or the ToF-SIMS measurement in the case of a frozen-hydrated cell measurement ($<$ 12h). To obtain freeze-dried cells, cold samples were transferred to the freeze-dryer (VirTis Benchtop Freeze Dryer) in falcon tubes and the chamber was immediately pumped down to approximately 10 Pa. When the freeze-drying process was complete, the samples were stored in a vacuum desiccator. Well-preserved cells for sub-sequent ToF-SIMS analysis were identified with light microscopy.

## 2.4 Optical Microscopy

Optical microscopy images were captured on a Nikon Eclipse Ti inverted microscope with a 20x/0.50 objective and an ORCA-Flash4.0 (C11440 - Hamamatsu) digital CMOS camera (2048 $\times$ 2048, 6.5 $\mu$m pixels). The resulting images are analysed and processed using ImageJ.

## 2.5 Interferometry

In order to assess the samples' surface topography, a 3D mapping was achieved using a vertical scanning interferometer (Fogale Nanotech Photomap 3D; 763 x 573, 9.4 $\mu$m pixels) with a 5x/0.12 objective. The total roughness $R_t$ and root-mean-square roughness $R_q$ are calculated from the data using Matlab (Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States). More details can be found in section .

## 2.6 ToF-SIMS

The ToF-SIMS analyses were executed using a ION-TOF TOF-SIMS IV instrument (Muenster, Germany), equipped with a Bi liquid metal ion gun (LMIG) and Ar gas cluster ion beam (GCIB). The primary ion beam is directed at the sample under an angle of 45 $^\circ$ in relation to the normal and has a beam spot of 1-2 $\mu$m in the high-current bunched

mode. 25 keV $Bi_3^+$ primary ions were used in all measurements. Charging of the sample is compensated with the low-energetic electrons of the flood gun. 20 keV $Ar_{5000}^+$ ion clusters were used for sputtering. Ion images were recorded in the high current bunched mode, which allows for higher mass resolution.

### 2.6.1 Freeze-dried cells

Unless stated otherwise, 250 × 250 $\mu m^2$ raster scans with 256 × 256 pixels and 5 frames per scan (1 shot per pixel) were analysed. With a target current of 0.3 pA in AC mode, this resulted in an ion dose of $1.02 \times 10^{13}$ primary ions $cm^{-2}$, which approximates the static limit. The 20 keV $Ar_{5000}^+$ GCIB sputtered a region of 500 x 500 $\mu m^2$ with a target current of 1.2 nA for 5 s (= 6 C) between the 160 scans or $1.50 \times 10^{22}$ primary ions $cm^{-2}$, unless stated otherwise. The positive ion mass spectra were calibrated with m/z 1 ($H^+$), 15 ($CH_3^+$), 29 ($C_2H_5{}^+$), 43 ($C_3H_7{}^+$) and 57 ($C_4H_9{}^+$). The negative ion mass spectra were calibrated with m/z 1 ($H^-$), 12 ($C^-$), 13 ($CH^-$), 16 ($O^-$) and 17 ($OH^-$).

### 2.6.2 Frozen-hydrated cells

For ToF-SIMS analysis of frozen-hydrated cells, cell samples were mounted one at a time onto an ION-TOF cryostage immersed in liquid nitrogen (cf. reference [97]). The sample stage was removed from the liquid nitrogen and placed onto the precooled sample transfer arm in the entry chamber under nitrogen flow. The entry chamber was immediately pumped down to $2 \times 10^{-5}$ mbar. When this pressure was reached, the entry door to the main chamber was opened to transfer the cryostage. Analysis was performed at -120 °C.

250 × 250 $\mu m^2$ raster scans with 256 × 256 pixels and 5 frames per scan (1 shot per pixel) were analysed. With a target current of 0.3 pA in AC mode, this resulted in a ion dose of $1.02 \times 10^{13}$ primary ions $cm^{-2}$, which approximates the static limit. 36 C of 20 keV $Ar_{5000}{}^+$ clusters (GCIB) were used for sputtering frozen-hydrated samples over a region of 500 x 500 $\mu m^2$ between scans or $8.99 \times 10^{22}$ primary ions $cm^{-2}$, because of the lower sputter yield of an ice matrix [98, 99]. The positive ion mass spectra were calibrated with m/z 1 ($H^+$), 15 ($CH_3^+$), 29 ($C_2H_5{}^+$), 43 ($C_3H_7{}^+$) and 57 ($C_4H_9{}^+$). The negative ion mass spectra were calibrated with m/z 1 ($H^-$), 12 ($C^-$), 13 ($CH^-$), 16 ($O^-$) and 17 ($OH^-$).

## 2.7 Data-processing

A peak search was performed to locate relevant mass peaks with a minimum of 15,000 counts, a minimum of 3.0 SNR and a maximum of 0.8 background (unless stated otherwise) and secondary ion images were reconstructed from the raw data files with the commercial ION-TOF software (Surfacelab 6). These values allowed all major peaks in the spectra to be selected and were determined empirically for these specific data sets. Images for the integrated peaks were then exported in an ASCII file format for further data processing in Matlab (Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States). All further data processing was performed using in-house generated Matlab scripts. All calculations were performed on a 64-bit Windows 10 platform with 8GB of RAM, using an Intel Core i3, 1.8GHz processor.

CHAPTER 3

# Principal Component Analysis of Large ToF-SIMS Datasets[1]

Despite the increasing capabilities of ToF-SIMS instruments thanks to the the development of novel ion beams [9, 10], typical ToF-SIMS measurements have a number of fundamental limitations that make data acquisition and interpretation challenging [9]. Chief among these is the intrinsic trade-off between high mass resolution and high spatial resolution as explained in subsection 1.2.5. Analysis in the static regime limits the signal-to-noise ratio as no more than 1% of the surface can be bombarded with primary ions in order to avoid hitting sites damaged by the analysis beam, which means only a very small fraction of the sample is used for analysis. The low duty cycle of the pulsed ion beam leads to long depth profiling experiments, which frequently causes samples to be analysed well below the static limit as well, in order to save time. Additionally, the ion images of high-mass molecular species often have a poor signal-to-noise ratio due to the low ion count per pixel [10]. There are also complications involving the secondary ion yield, when the sample material has a curvature or a surface topography in excess of several tens of $\mu$m [101]. With these limitations in mind, powerful data analysis is of the essence, which is why the SIMS community has embraced multivariate analysis (MVA) methods such as PCA [69]. In section 1.5.1, the theoretical background for PCA is discussed as wells as its applicability to ToF-SIMS data.

While it is important to select an appropriate MVA technique for ToF-SIMS data, it is equally important to select appropriate data pretreatment. In section 3.1, a method to construct TOF-SIMS image data consisting of multiple chemistries with spatial distributions of their own is proposed. With this method, it is possible to identify the

[1] This chapter has been partly published in Analyst (RSC publishing) as 'Multivariate analysis of 3D ToF-SIMS images: method validation and application to cultured neuronal networks' [100]. Some passages are quoted verbatim from this source.

effects of different preprocessing procedures such as mean-centering, standardisation (autoscaling) and normalisation.

PCA already proved useful for 2D ToF-SIMS image analysis, but 3D ToF-SIMS data sets are typically very large and therefore often unsuitable for MVA using the processing power of standard desk top computers [9]. As a result, up until now the only published application of PCA on a 3D ToF-SIMS dataset was reported by Fletcher et al. [33] on a relatively small 3D ToF-SIMS image with a size of 256 x 256 x 10 pixels. Given the developments in 3D ToF-SIMS imaging, there is need for a method of facilitating PCA of large ToF-SIMS data sets as their size will only increase. In section 3.2, we demonstrate that it is possible to expand the application of PCA to large (128 x 128 x 622) 3D images without requiring any computing resources beyond a desk top computer by using a small training subset comprising 6.1% of the total amount of pixels to determine the PCA loadings. We have validated our method using an established data set with known composition and distribution that was previously published [102].

## 3.1 Evaluation of Data Preprocessing Effects using Mock Data

Data pretreatment is specific to the instrumental technique used. In the case of ToF-SIMS, data preprocessing begins with the selection of peaks from the spectra and can also include normalisation, centering, scaling and non-linear transformations [69, 83]. In subsection 3.1.1, an overview of typical TOF-SIMS image data preprocessing steps for PCA is given.

When choosing a given preprocessing procedure, there is a tendency to choose the procedure which gives the best looking results [72]. The problem is of course that there is no way to determine which procedure is 'correct'. To identify the effects of different preprocessing procedures, it is necessary to know what the nature and spatial distribution of the sample components are and if there is convolution in the data due to topography and/or matrix effects. In other words, we would already need to know the answer to understand the effect of different preprocessing steps. In subsection 3.1.2, a method to construct TOF-SIMS image data consisting of multiple chemistries with spatial distributions of their own is proposed. The method also allows to include matrix effects, topography and noise.

### 3.1.1 Preprocessing

There are a variety of preprocessing procedures, a number of them - specifically related to ToF-SIMS data - will be discussed below.

The first preprocessing step of ToF-SIMS data typically involves **peak selection**. One can manually select known peaks or select all peaks that fit certain criteria such as a minimum number of counts, a minimum SNR or maximum background [83]. It is also common to remove contaminant peaks from the data set [83]. Peak selection can significantly impact the results and the assumptions made about the data must be carefully considered.

**Normalization** is typically done to remove variance due to sample charging, instrumental conditions, topography or matrix effects [69, 72]. It should be noted however that normalization removes some information from the spectra.[83] and can accentuate noise in ToF-SIMS images due to the low count rates often found within the data [72]. Therefore, care should be taken when using normalization of ToF-SIMS images[72].

In our description of PCA in subsection 1.5.1, **mean-centering** is necessary to ensure that the first principal component is proportional to the maximum variance. With different algorithms, however, mean-centering is optional, when performing PCA. Avoiding to center the data matrix results in a more general first component that mainly reflects the mean of the full data dataset.

**Scaling** refers to dividing each variable by a constant. Subsection 1.5.3 already mentioned that PCA is sensitive to the scaling of the variables and that the correlation matrix can be used as the basis for the PCA instead of the covariance matrix used for standard (i.e. mean-centered) PCA. Because the correlation coefficients are obtained by dividing the covariance of the variables by the product of their standard deviations, a correlation matrix is a covariance matrix of the standardized data. Therefore, principal component analysis using the correlation matrix is equivalent to standard (i.e. mean-centered) principal component analysis using the standardized data. Standardised means the variables are rescaled to have a mean of zero and a standard deviation of one by subtracting the mean from each variable and then dividing the difference by the standard deviation of the variable.

There is a consensus that data should be scaled in a manner that is consistent with the noise structure of the data [69, 72]. As stated in 1.5.4, ToF-SIMS data often follows a Poisson distribution and scaling methods that account for Poisson noise (such as scaling by the square root of the mean) work well for ToF-SIMS data [69, 72]. On the other hand, while the (corrected) sample standard deviation is a biased estimator for the

population standard deviation, its bias declines with 1/N as sample size increases. If one were to generate a random sample set of observations from a Poisson distribution with different parameters $\lambda$ that are consistent with the number of counts per pixel observed in ToF-SIMS images and different sample sizes, it is easy to see that the difference between the sample standard deviation and the known value $\sqrt{\lambda}$ becomes very small for sample sizes consistent with the large number of pixels typically present in 2D ToF-SIMS images[2].

To illustrate this decrease in bias as sample size increases, a graph of the bias in the sample standard deviation and square root of the mean vs. $\lambda$ and sample number $n$ is presented in Figure 3.1. Random numbers from the Poisson distribution with mean parameter $\lambda$ = 1, 10 and 100 were generated with sample sizes $n$ = 10, 100, 1000, 10000 and 100000, which are consistent with the ion counts/pixel and image sizes for (2D) ToF-SIMS images, respectively. The sample standard deviations and square of the means were calculated for each vector of random numbers and their bias was expressed as the percentage of the true value $\sqrt{\lambda}$. The average bias of 1000 repeats is presented in the graph. At n = 1000, the bias of the sample standard deviation and the square of the mean start to converge. For even higher sample numbers, the bias becomes negligible with less than 0.7 % for n = 10,000 ($\equiv$ 100 $\times$ 100 pixels) and less than 0.2 % for n = 100,000 ($\approx$ 316 $\times$ 316 pixels). This illustrates that standard deviation scaling is equivalent to Poisson scaling for large sample sizes.

It is also important to note that standardisation makes no assumptions regarding the distribution of the data. When the assumption of a Poisson distribution is badly violated, which occurs when data is normalized [69], root mean scaling will not provide a good estimate of the measurement uncertainty. The sample standard deviation, on the other hand, will be consistent with the noise structure of the data, provided the sample number is sufficiently high. One must be careful with standardisation, however, as it has a tendency to amplify noise peaks relative to peaks which show image contrast [69] in which case root mean scaling might provide better results.

**Transformation** refers to transforming the data with a function such as the logarithm or square root [72].

[2]In the case of a parametric family of distributions, the standard deviation can be expressed in terms of the parameters. The normal family is parametrized by $\theta = (\mu,\sigma)$, whereas the Poisson family of distributions is parametrized by a single number $\lambda$. The interesting thing about the Poisson distribution is that its expectation and its variance are both equal to its parameter $\lambda$.

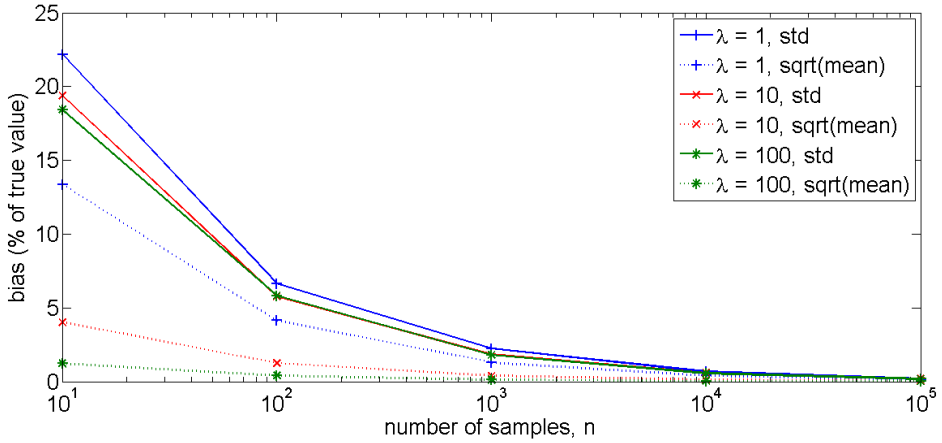



**Figure 3.1:** Graph of the bias in the sample standard deviation and square root of the mean vs. sample number n for different $\lambda$. Random numbers from the Poisson distribution with mean parameter $\lambda$ = 1, 10 and 100 were generated with sample sizes n = 10, 100, 1000, 10000 and 100000. The sample standard deviations and square of the means for each were calculated and their bias expressed as the percentage of the true value $\sqrt{\lambda}$. The average bias of 1000 repeats is presented in the graph.

### 3.1.2 Creating Mock Data

To identify the effects of different preprocessing procedures, a method to construct TOF-SIMS image data consisting of multiple chemistries with spatial distributions of their own is proposed. The method also allows to include matrix effects, topography and noise.

Ion intensities for compounds can be obtained from databases (see Table A.1). The spatial distribution of an analyte can be supplied from a monochrome image with the same size as the TOF-SIMS image one wants to form. An 8 bit grayscale image of 256 × 256 pixels, for example, can be imported into Matlab as a matrix and then reshaped into a vector P containing the pixels' intensities[3]. In the case of 8 bit grayscale images, the pixel intensity will vary between 0 and 255. This pixel intensity vector **P** can then be multiplied with a vector **I** containing the ion intensities, creating a matrix **C** matching SIMS image data:

$$C = P \times I = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_{65536} \end{pmatrix} \begin{pmatrix} i_{\text{total}} & i_{\text{sum of rest}} & i_{\text{m/z 1}} & \cdots & i_{\text{m/z 150}} \end{pmatrix} = \begin{pmatrix} c_{\text{total},1} & c_{\text{sum of rest},1} & c_{\text{m/z 1},1} & \cdots & c_{\text{m/z 150},1} \\ c_{\text{total},2} & c_{\text{sum of rest},2} & c_{\text{m/z 1},2} & \cdots & c_{\text{m/z 150},2} \\ c_{\text{total},3} & c_{\text{sum of rest},3} & c_{\text{m/z 1},3} & \cdots & c_{\text{m/z 150},3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{\text{total},65536} & c_{\text{sum of rest},65536} & c_{\text{m/z 1},65536} & \cdots & c_{\text{m/z 150},65536} \end{pmatrix} \quad (3.1.1)$$

[3] In this subsection, the simulation of 2D images will be outlined but the formation of 3D images follows easily.

If the ion intensities are normalised, the total ion count $i_{total}$ equals 1. The sum of rest $i_{sum\ of\ rest}$ is the sum of the intensities that fall outside of the mass range or mass increments. If the ion intensities are normalised, the maximum number of counts (MC) per pixel will be $\leq 255$ for each chemistry, because $c_{total,i} \leq 255$. Normalisation also implies that all compounds are given the same ionisability. This can of course be changed by scalar multiplication.

Several chemistries can be added to the TOF-SIMS image through linear combination of such matrices using the same or different images. The variance of each ion is then given by the sum of all elements in the covariance matrix of the components.

$$Var\left(\sum_{i=1}^{n} a_i P_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(P_i, P_j) \tag{3.1.2}$$

If the variables are uncorrelated, i.e. $Cov(P_i, P_j) = 0\ \forall (i \neq j)$, the variance of their sum is equal to the sum of their variances (Bienaymé formula):

$$Var\left(\sum_{i=1}^{n} a_i P_i\right) = \sum_{i=1}^{n} a_i^2 Var(P_i) \tag{3.1.3}$$

This means that when the spatial distributions (images) are linearly independent, which is roughly the case when different images are used, the coefficient of variation is the same for all ions.

**Matrix effects** can for example be included by element-wise multiplication of the image vector with a (scaled) image vector associated with a different chemistry. The spatial distributions are then of course no longer independent.

**Topography** can be incorporated in the TOF-SIMS image by element-wise multiplication with a (scaled) image vector that serves as a 'topography vector'. If the variables are independent, the variance of their product is given by:

$$Var(aTP) = a^2\left(\mu_T^2 Var(P) + \mu_P^2 Var(T) + Var(T)Var(P)\right) \tag{3.1.4}$$

Multiplicative **noise** can be added to the TOF-SIMS image I by using the equation $I + nI$, where n is random noise of a particular distribution with mean 0.

### 3.1.3 Example: Coumarin Distribution in Acrylamide

To illustrate the effects of standardisation, normalisation and the combination of normalisation and standardisation, the PCA results are calculated for a mock coumar-

in/acrylamide sample. One could think of this as the distribution of a coumarin dye in an acrylamide gel after gel electrophoresis.

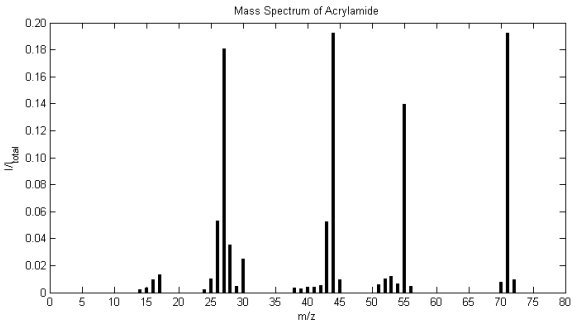
The ion intensities for acrylamide and coumarin are easily obtained from a database (see Table A.1). In order to create a chemical image, the image vector of a detail of The Blind Beggar (1853) by Jozef Laurent Dyckmans is multiplied with a vector containing the normalised intensities of acrylamide (see Figure 3.2) and the image vector of The Starry Night (1889) by Vincent van Gogh is multiplied with a vector containing the normalised intensities of coumarin (see Figure 3.3). The resulting two matrices, are added together to form a $256 \times 256$ TOF-SIMS image. Noise is added to the TOF-SIMS image by element-wise multiplicating with a matrix containing uniformly distributed random numbers with mean 1 and variance 0.04. The variance of an ion is then given by Var(aNP) with a the normalised intensity of that ion, N the noise vector of that ion and P the sum of the two image vectors $P_1 + P_2$. It is important to note that, because the mean and variance of every noise vector is set at 1 and 0.04 respectively, Var(NP) is approximately the same for all ions, which in turn means that the coefficient of variation is roughly the same for all ions.

Successful PCA ought to separate the chemistries so that the loadings resemble the mass spectrum and the scores images match the spatial distribution. One way of quantifying this is to look at the peak signal-to-noise ratio (PSNR) between the scores images and the images used for coumarin and acrylamide's spatial distribution. The benchmark for useful PCA is given by the PSNR of the ion images of the ions with the highest intensity that are specific for coumarin (m/z 118) and acrylamide (m/z 71). The PSNR of m/z 71 with the reference being the 8 bit grayscale image of The Blind Beggar is 22.1515 dB, and the PSNR of m/z 118 with the reference being the 8 bit grayscale image of The Starry Night is 15.1411 dB (see Figure 3.4). These values are the benchmark for useful PCA, i.e. the scores images need to have a higher PSNR than the specific ions.

The colour scheme of ion images (see Figure 3.4) is the same as the default hot colour scheme used in Surfacelab, the commercial ION-TOF software. This colour scheme has been used in this thesis, when presenting positive valued image data such as (single) ion images and interferometry heightmaps. Race et al. [103] recommends the use of perceptually linear colour schemes. The lightness values of the hot colour scheme more or less monotonically increases apart from a small kink. The hot colour scheme therefore approximately satisfies this criterion.

With no data preparation 95 % of the variance can be explained by the first five PCs. N.B. the PCA function centers the data by default. PC1's scores image corresponds to

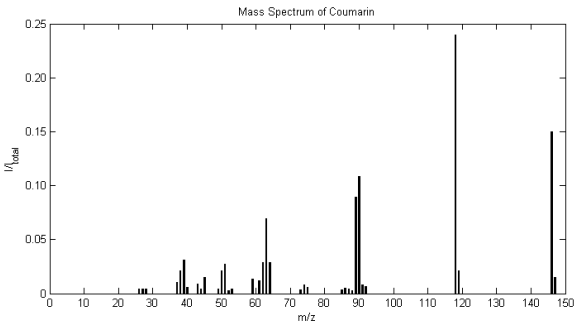
**(a)** The Blind Beggar (1853) by Jozef Laurent Dyckmans.




**(b)** Normalised mass spectrum of acrylamide.

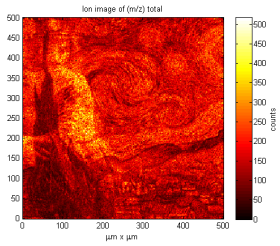
**Figure 3.2:** Spatial distribution and ion intensities for acrylamide.

**(a)** The Starry Night (1889) by Vincent van Gogh.

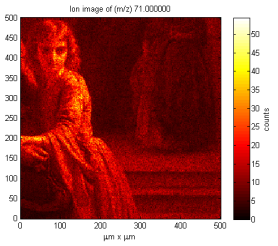



**(b)** Normalised mass spectrum of coumarin.

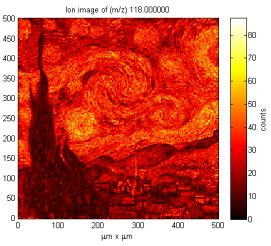
**Figure 3.3:** Spatial distribution and ion intensities for coumarin.




**(a)** MC 517, TC 1.0522e+07




**(b)** MC 54, TC 6.3295e+05




**(c)** MC 87, TC 1.7310e+06

**Figure 3.4:** Simulated ion images. (a) Total ion image (b) m/z 71 which is specific for acrylamide (PSNR = 22.1515 dB) and (c) m/z 118 which is specific for coumarin (PSNR = 15.1411 dB).

coumarin's spatial distribution and its loadings resemble coumarin's mass spectrum (see Figure 3.5), whereas PC2's scores image corresponds to acrylamide's spatial distribution and its loadings resemble acrylamide's mass spectrum (see Figure 3.6). The PSNR of the scores images are higher than those of the specific ions.

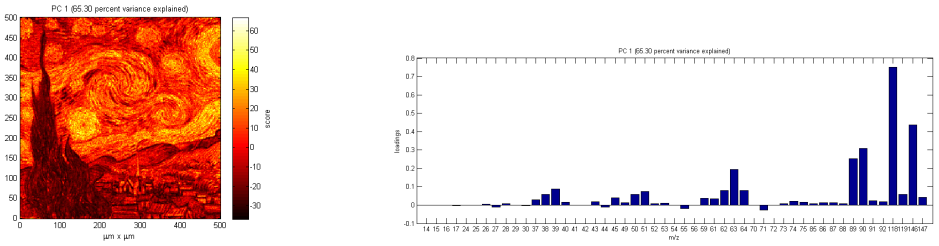
With standardisation the first 35 PCs are needed to explain 95 % of the variation. Again, PC1's scores image corresponds to coumarin's spatial distribution and its loadings resemble coumarin's mass spectrum (see Figure 3.7), whereas PC2's scores image corresponds to acrylamide's spatial distribution and its loadings resemble acrylamide's mass spectrum (see Figure 3.8). The scores images have a higher PSNR than that of the specific ions and the PCA with no data preprocessing. The loading plots, however, shows that all variables are given the same weight regardless of how abundant they are.

With normalisation of the data before PCA, 95 % of the variance is explained by the first 6 PCs. PC1's loadings plot contains both mass spectra: the positive loadings resemble coumarin's spectrum, whereas the negative loadings resemble acrylamide spectrum (see Figure 3.9). Correspondingly, PC1's scores image contains both chemistries, where a positive score is related to coumarin and a negative one to acrylamide. This makes for a poor PSNR compared to either specific ion image.

With normalisation and standardisation of the data before PCA, 95 % of the variance is explained by the first 34 PCs. PC1's loadings plot again contains both mass spectra: the positive loadings resemble acrylamide's spectrum, whereas the negative loadings resemble coumarin's spectrum (see Figure 3.10). Correspondingly, PC1's scores images contains both chemistries, where a positive score is related to acrylamide and a negative one to coumarin. The PSNR compared to either specific ion image is poorer and this time the variables are given the same weight due to the standardisation.

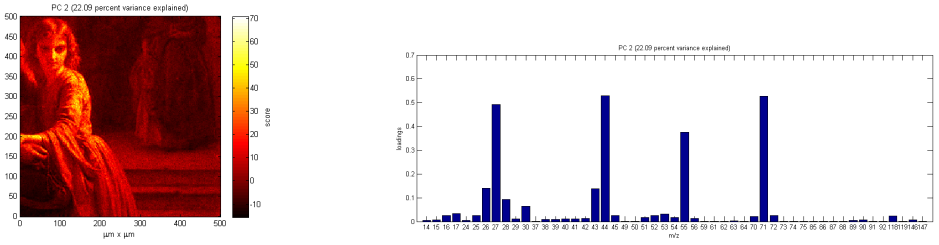
It is thus possible to simulate chemical images and assess the use of PCA and the various preprocessing options. This simulation responds best to standardisation as far as the scores images are concerned (highest PSNR). Without data preprocessing variables with the highest sample variances tend to be emphasized, whereas standarisation gives all variables the same weight in the loading plots. Normalisation tends to contain both chemistries in the first PC, leading to a merged scores image with poor PSNR compared to the specific ion images. This can, however, be circumvented by only using the positive or negative coefficients to form scores images.

The PCA scores images this section have been presented in the same hot colour scheme as the ion images. However, it is apparent there is a drawback using this colour scheme as it is not immediately evident which regions of the image have a positive value and



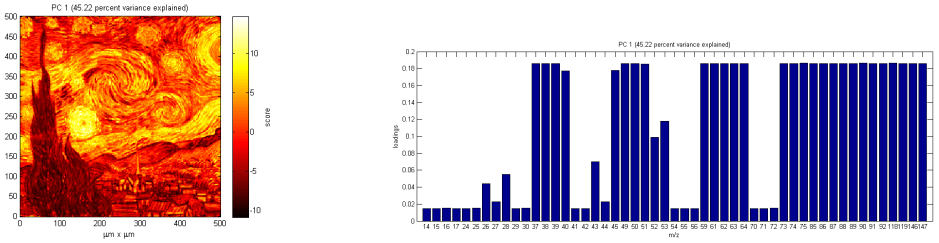

**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.5:** No data preparation: PC1 corresponds to coumarin. The scores image has a PSNR equal to 18.8553.



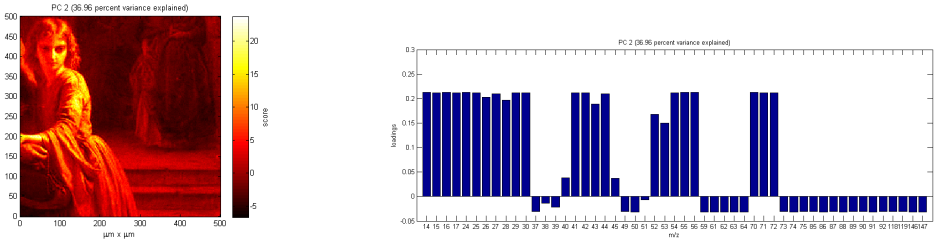

**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.6:** No data preparation: PC2 corresponds to acrylamide. The scores image has a PSNR equal to 25.9409.



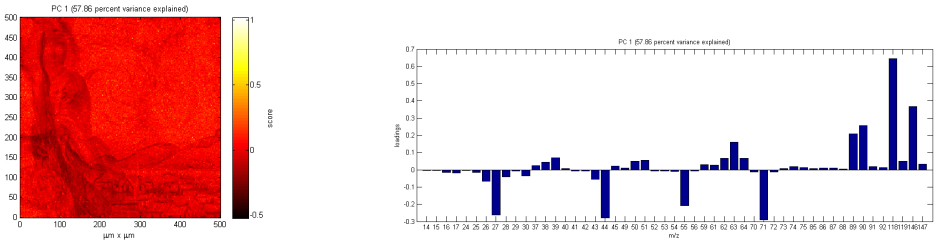

**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.7:** Standardisation: PC1 corresponds to coumarin. The scores image has a PSNR equal to 26.7833.



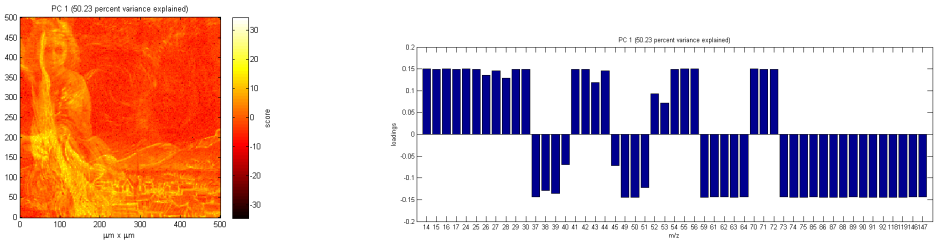

**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.8:** Standardisation: PC2 corresponds to acrylamide. The scores image has a PSNR equal to 34.9276.




**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.9:** Normalisation: PC1 corresponds to both coumarin and acrylamide. The scores image has a PSNR of 13.9185 for coumarin and 13.4932 for acrylamide.




**(a)** Scores image. **(b)** Loadings plot.

**Figure 3.10:** Normalisation and standardisation: PC1 corresponds to both coumarin and acrylamide. The scores image has a PSNR of 11.4606 for coumarin and 10.6283 for acrylamide.

which have a negative value. This improves when using the Matlab's default jet colour scheme, which has been used in sections 4.3 and 4.4. However, it becomes a colour-matching exercise to determine which values are positive and which are negative. In the case of PCA scores images, Race et al. [103] recommends the use of a diverging colour scheme centred around zero such as the red and blue scheme centred around black used in sections 3.2 and 4.2.

In summary, the method described here allows the effect of preprocessing steps to be evaluated on completely known data sets. It can be particularly useful, if one has prior knowledge of the composition or distribution of a real sample. These mock data sets can also be used during the development of new preprocessing methods to test their performance.

## 3.2 Training Set Method for Large ToF-SIMS Data-sets

While PCA already proved useful for 2D ToF-SIMS image analysis, 3D ToF-SIMS data sets are typically very large and unsuitable for MVA using the processing power of standard desk top computers [9]. As a result, up until now the only published applica-

tion of PCA on a 3D ToF-SIMS dataset was reported by Fletcher et al. [33] on a relatively small 3D ToF-SIMS image with a size of 256 x 256 x 10 pixels. Recently, Cumpson et al. [104] developed faster algorithms that allowed PCA to be performed on large 2D data sets. In this section, we demonstrate that it is possible to expand the application of PCA to large (128 x 128 x 622) 3D images under 30 minutes without requiring any computing resources beyond a desk top computer [100]. We used a small training subset comprising 6.1% of the total amount of pixels, which were randomly selected from the full 3D image, to determine the PCA loadings (i.e. linear combinations of the original mass peaks accounting for amounts of variance). These loadings were then applied to the full data set. We have validated our method using an established data set with known composition and distribution that was previously published [102].

### 3.2.1 Development of Training Set Method

A simplified schematic of our data processing method for large 3D ToF-SIMS images is shown in Figure 3.11. After a peak search to identify the relevant mass peaks, the respective secondary ion images are imported into Matlab (Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States) and reshaped into (scan resolved) matrices, where the rows represent pixels (or samples) and the columns represent mass peaks (or variables).

Normalisation with the total ion count per pixel or the sum of the selected peaks is an option at this point in case one would like to minimise variations in the secondary ion signal due to differences in topography, sample charging or instrumental conditions such as variations in primary ion current or detector efficiency [83, 105]. All data sets presented here are normalised prior to analysis.

Because the eigendecomposition involved is computationally intensive, a smaller training subset of randomly selected pixels is created to calculate the principal component coefficients (i.e. the loadings). Depending on whether the covariance or the correlation matrix of the training set is decomposed, the data is either mean-centered or standardised (auto-scaled) respectively. Because the correlation coefficients are obtained by dividing the covariance of the variables by the product of their standard deviations, the correlation matrix is equal to the covariance matrix of the standardized data. When mean-centering, PCA will give more weight to variables that have higher variances, which tend to be the variables with higher means. If the variables are standardised, all variables will be weighted equally regardless of how abundant they are. It is important to note that standardisation has a tendency to amplify noise peaks relative to peaks which show image contrast [69] and is therefore not generally recommended

over scaling methods that are consistent with the structure of the noise (see also 3.1.1). Root mean scaling, derived from the assumption that the image noise is Poisson in nature, often yields better results. However, this is not the case when the Poisson assumption is badly violated, which occurs when the data is normalized [69]. While the (corrected) sample standard deviation is a biased estimator for the population standard deviation, its bias drops off as 1/N as sample size increases (see also 3.1.1). Given the size of the training sets used and the fact that the sample standard deviation makes no assumptions regarding the distribution, scaling using the sample standard deviation was chosen. The training set data presented here has always been standardised.

The full data matrix then needs to be standardised and multiplied with the loadings in order to calculate the scores for every pixel in the image. This can be done efficiently one scan at a time (block processing).

The Matlab code is presented in appendix A.

### 3.2.2 Method Validation using a Model 3D ToF-SIMS Data-set

To validate the method, we are using a previously published [102] model 3D ToF-SIMS data set of a spin-cast multilayer sample comprising ten well-defined, alternating layers of 50 nm polystyrene (PS) and 200 nm polyvinylpyrrolidone (PVP) on a silicon wafer substrate. Mass calibration, peak search and image reconstruction are performed with the commercial ION-TOF software (SurfaceLab 6) and all further data processing is performed with Matlab. The test image consists of 128 x 128 x 622 pixels with 258 relevant mass peaks each ($\approx 2.6 \times 10^9$ data points). First, the data is normalised to the total number of ion counts per pixel to account for the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling. In the work of Bailey et al. [102], the specific layers of PS, PVP and the silicon wafer are identified using the $C_7H_7^+$ (m/z 91), $C_6H_{10}NO^+$ (m/z 112) and $Si^+$ (m/z 28) ions respectively. For the z-scaling of the data the silicon wafer interface first needs to be established. A Gaussian function ($R^2_{adj}$ = 0.92) is fitted to the gradient of the average $Si^+$ intensity of each XY plane in the z-direction (see Figure 3.12 A) to identify the position of the interface. The position of the centre of the peak is considered to be the interface and was set to z = 0 nm. The sputter yields for PVP and PS under the used experimental conditions were previously determined [102] and equal (0.654 $\pm$ 0.006) nm/s and (0.828 $\pm$ 0.01) nm/s respectively. The PS-PVP interfaces are similarly determined by Gaussian fits and their sputter times are converted into layer thicknesses. Ion images for m/z 91 and m/z 112 are presented in Figures 3.13 A and B to show the PS and PVP layers, respectively. Their SNR is calculated as the ratio between the

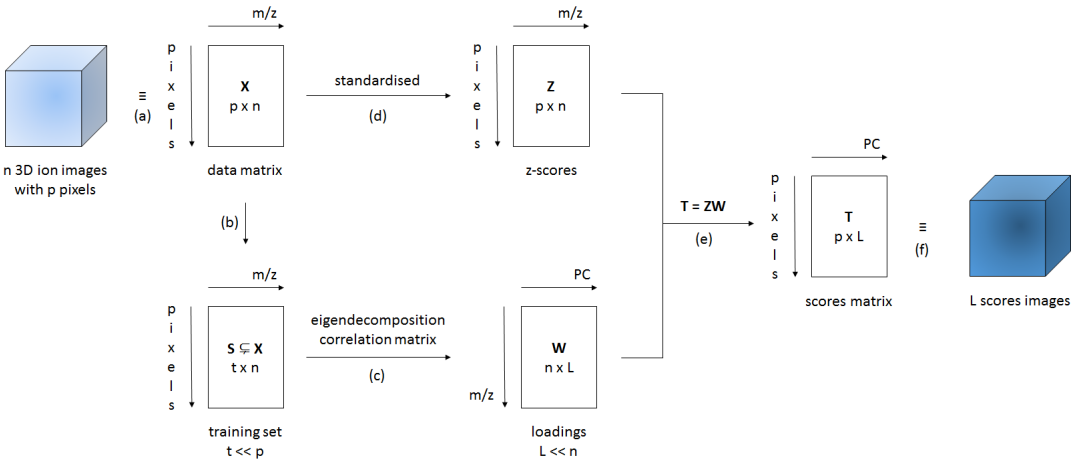



**Figure 3.11:** Simplified schematic of the data processing method used. (a) The n different 3D (normalised) ion images for every m/z can be presented as a data matrix **X** with n columns (one for every m/z) and p rows (one for every xyz pixel). (b) In order to calculate the loadings matrix **W** of the L ($\ll$ n) principal components, a smaller training data set **S** with t ($\ll$ p) randomly selected pixels is created; the training set **S** (t x n) is a subset of the data matrix **X** (p x n). (c) Eigendecomposition of the correlation matrix of **S**, provides the loadings matrix **W** with L columns (one for every PC) and n rows (one for every m/z). (d) Because the training set **S** was standardised for the calculation of the loadings **W**, the data matrix **X** has to be standardised as well using the mean and standard deviation for each column n of the training set **S** generating the z-scores matrix **Z** (p x n). (e) The scores matrix **T** with L columns and p rows is calculated as the matrix product of **Z** (p x n) and **W** (n x L). (f) The scores matrix **T** can now be presented in the form of L ($\ll$ n) interpretable 3D scores images.

mean and standard deviation ($\mu_{sig}/\sigma_{sig}$) of the average ion intensity of each plane in the z-direction and equals 0.99 for the ion image for m/z 91 and 0.79 for the ion image for m/z 112 (see Figures 3.14 A and B). The low SNR is a consequence of the low ionisability of organic samples; the maximum count per pixel equals only 1. The range of intensities seen in the ion images is solely due to the pixel to pixel variation in the total ion signal. The depth resolution $\Delta z$ for the various interfaces is calculated by fitting Gaussian functions to the gradient of the average intensity of the specific ions at m/z 91 and 112 in the z-direction (see Figures 3.15 A and B) and using the definition that the depth resolution $\Delta z = 2\sigma$ where $\sigma$ is the standard deviation of the Gaussian [106]. The average $\Delta z = (4.2 \pm 0.7)$ nm (n = 9).

PCA is performed by regarding the mass peaks in the spectra as variables and each pixel as an individual sample. As the eigendecomposition involved is computationally

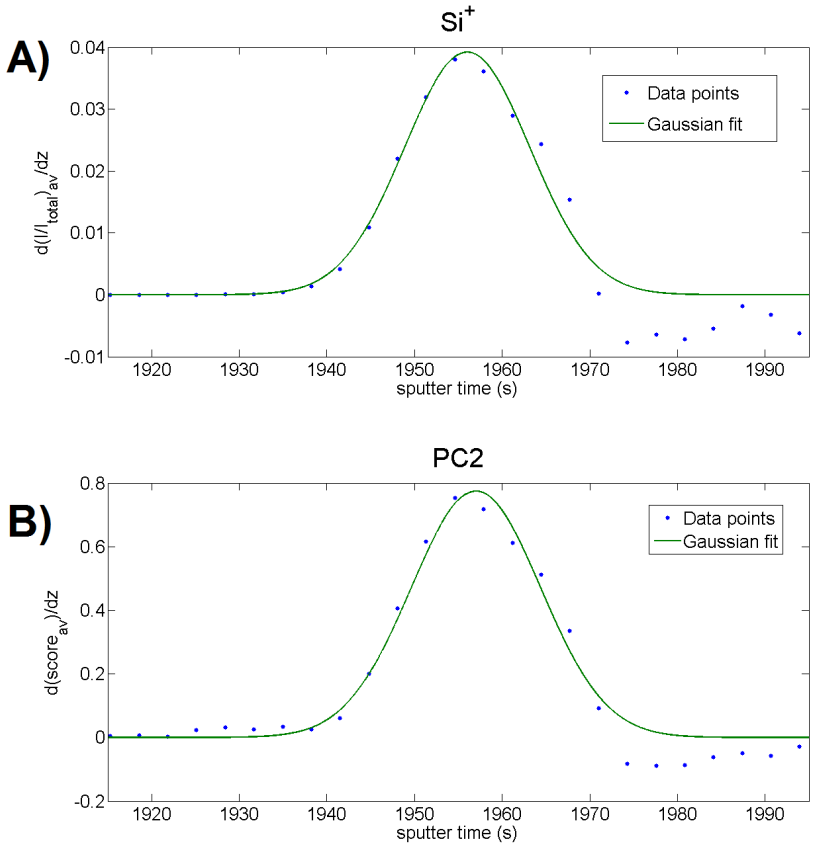



**Figure 3.12:** Z-calibration for the PS-PVP multilayer sample using the Si wafer interface. A) Gaussian curve fit ($ae^{-\left(\frac{x-b}{c}\right)^2}$ with a = 0.03921, b = 1956 and c = 10.13) to the gradient of the average $Si^+$ intensity in the z-direction. This is used to scale the ion images. N.B. only the data points in the vicinity of the interface are shown. B) Fitting a Gaussian curve ($ae^{-\left(\frac{x-b}{c}\right)^2}$ with a = 0.7744, b = 1957 and c = 10.48) to the gradient of the average PC2 scores in the z-direction. This is used to scale the scores images. N.B. only the data points in the vicinity of the interface are shown.

intensive, the PCA is executed on a training set created by randomly selecting a thousand pixels from each z-plane (see 3.2.1); the training set thus consists of 622,000 pixels (i.e. mass spectra) or 6.1 % of the total number of pixels. Prior to PCA we tested if the normalised variables follow a Poisson distribution. The variance-to-mean ratio (VMR) was calculated and a chi-square goodness of fit test for a Poisson distribution was performed for each variable. If the variables are truly Poisson distributed, the VMR of the variables ought to equal 1; they average to $0.07 \pm 0.09$ (n = 258) for our data. The goodness of fit test yielded p-values $< 0.0001$. Both tests indicate that our data does not follow a Poisson distribution. Therefore, the loadings are generated for standardised variables (mass peaks) in the training set. Processing times and memory usage can be found in Table 3.1.

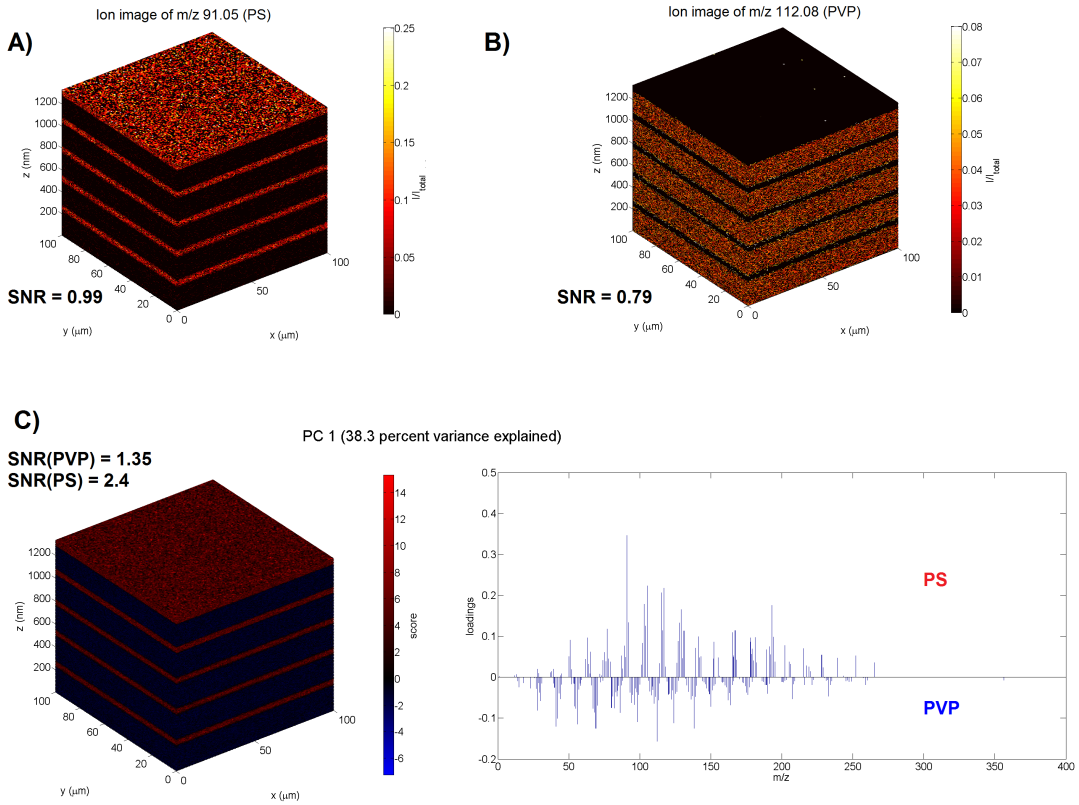



**Figure 3.13:** PCA of the PS-PVP multilayer sample. A) Normalised and scaled ion image of the specific ion for PS (m/z = 91). B) Normalised and scaled ion image of the specific ion for PVP (m/z = 112). C) PC1 explains 38.3% of the variance: 3D scores image (left) and loadings plot (right). The scores clearly visualise the alternating PS-PVP layers. The positive loadings of PC1 correspond to the mass spectrum of PS and the specific ion at m/z 91 is the one with the highest weight. The negative loadings of PC1 correspond to the mass spectrum of PVP and the specific ion at m/z 112 is the one with the highest weight. The silicon substrate has a score of approximately zero, i.e. the loadings do not apply.

In order to assess whether this random pixel selection is representative of the entire data set, the PCA is repeated ten times to determine if the amount of variance explained and loadings remain the same (see Table 3.2). The coefficient of variation (CV) of the different principal components was found to be smaller than 0.19 % (n = 10) indicating that the pixel selection is indeed representative. It should be noted that the sign of the loadings varies during these repeats, however, this does not alter their interpretation.

A direct comparison of the results of the training set approach with those of a PCA performed on the full data set is not possible due to memory limitations for a typical PC setup, but can be performed on a high performance system as kindly provided by

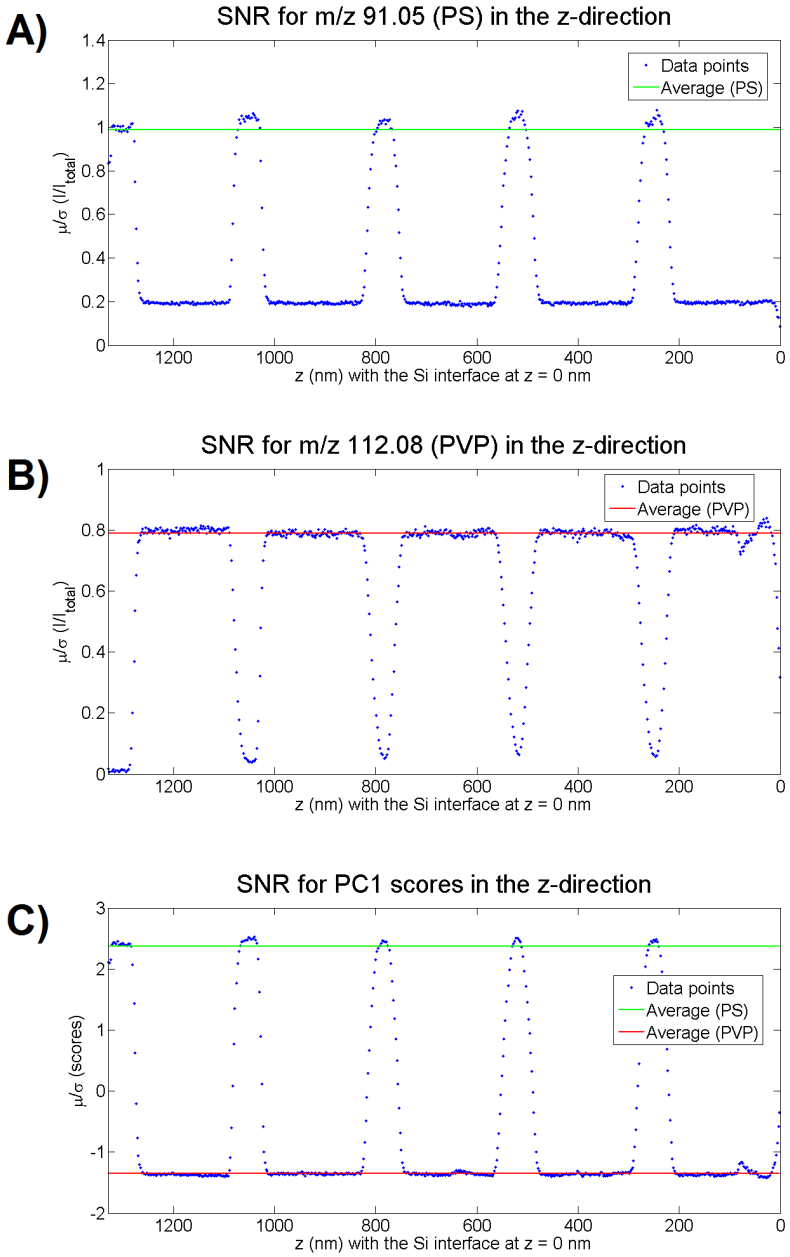



**Figure 3.14:** Calculation of the SNR for the PS-PVP multilayer sample ion and scores images. The SNR is calculated as $\mu_{sig}/\sigma_{sig}$ to allow direct comparison between ion images and scores images. A) The SNR for m/z = 91.05 (PS) equals 0.99. B) The SNR for m/z = 112.08 (PVP) equals 0.79. C) The SNR for the positive scores (PS) of PC1 equals 2.4 and the SNR for the negative scores (PVP) of PC1 equals 1.35.

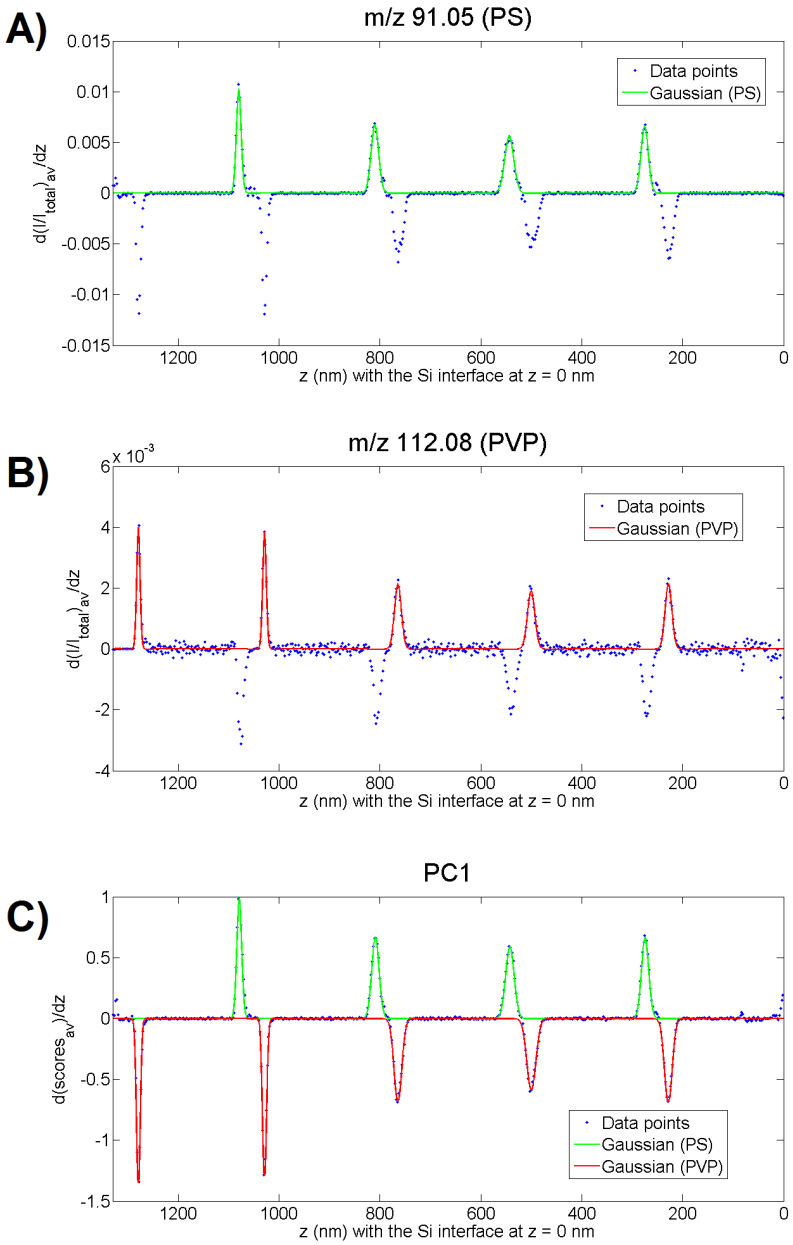



**Figure 3.15:** Calculation of the depth resolutions $\Delta z$ for the PS-PVP interfaces in the ion and scores images. The depth resolutions $\Delta z = 2\sigma$ are calculated by fitting Gaussian curves ($a_i e^{-((x_i - b_i)/(c_i))^2}$) to allow direct comparison between ion images and scores images. A) Fitting Gaussian curves to the gradient of the average m/z = 91.05 (PS) intensity in the z-direction. B) Fitting Gaussian curves to the gradient of the average m/z = 112.08 (PVP) intensity in the z-direction. C) Fitting Gaussian curves to the gradient of the average scores for PC1 in the z-direction. The parameters of the Gaussian functions can be found in Table A.2.

**Table 3.1:** Processing times and memory usage of the PCA performed on the validation data set. All calculations were performed on a 64-bit Windows 10 platform with 8GB of RAM, using an Intel Core i3, 1.8GHz processor.

| Process | Time (s) | Input (Mb) | Output (Mb) | Peak memory (Kb) |
|---|---|---|---|---|
| Training set formation | 44.3 | 1386.3 | 1283.8 | 1255792 |
| Eigendecomposition | 8.5 | 1283.9 | 0.012432 | 1256176 |
| Scores formation | 1667.4 | 2670.0 | 489.58 | 1256300 |

Dr. Alex Henderson[4]. The calculation of the eigendecomposition was performed on a 64-bit Windows 7 Ultimate platform in 296.7 seconds and required 33 GB of memory. Now, the amount of variance explained (%) for the first six PCs of the full PCA can be compared to those obtained in subsampling PCA and the values appear to be very similar. A two-sided one-sample t-test was performed for the null hypothesis that the amount of variance explained in subsampling PCA equals that of the full PCA (P values given in Table 3.2). No significant difference is observed for PC3 and PC5. A significant different with 95 % confidence was observed for PC1 and PC4, a significant difference with 99 % confidence was observed for PC2 and a significant difference with 99.99 % confidence was observed for PC6. Overall, it appears the training set method assigns slightly less variance to the first two PCs and slightly more variance to the subsequent PCs. The more important question is of course if the loadings are similar to those obtained from full PCA. A visual comparison of the loadings obtained from full PCA and those obtained from a PCA with the training set method is given in Figure A.1. Based on the loadings, it can be surmised that only PC1 and PC2 contain relevant information regarding our sample.

In order to get a quantitative measure of the similarity between the loadings, their cosine similarity $\cos\theta$ is calculated. Cosine similarity (also called the spectral contrast angle) is a measure of similarity between two vectors, loadings in our case, and is often used as a mass spectral similarity measure [107]. The value of $\cos\theta$ ranges from -1 to 1 with 0 indicating orthogonality (decorrelation), 1 meaning that the two spectra are exactly the same and -1 meaning they are exactly opposite. If two loadings are presented as vectors of the coefficients, $\mathbf{A}$ and $\mathbf{B}$, their cosine similarity, $\cos\theta$, is calculated by dividing their dot product by the product of their magnitudes:

[4]Surface Analysis Research Centre, School of Chemical Engineering and Analytical Science, University of Manchester, United Kingdom.

**Table 3.2:** Comparison between the results of a full PCA with the training set method. The amount of variance explained (%) is given for the first six PCs as well as the standard deviations for 10 repeats in the case of the training set method. Also included is the P value of a two-sided one-sample t-test testing the null hypothesis that the amount of variance explained in subsampling PCA equals that of the full PCA. The cosine similarity $\cos\theta$ with the loadings of the full PCA is calculated for all 10 repeats of the subsampling PCA.

| PC | Variance Explained (%) | | P value | $\cos\theta$ (n = 10) |
|---|---|---|---|---|
| | Full PCA | Subsampling PCA (n = 10) | | |
| 1 | 38.33 | $38.29 \pm 0.04$ | $0.0162^{*}$ | $0.99986 \pm 0.00001$ |
| 2 | 14.55 | $14.52 \pm 0.03$ | $0.0042^{**}$ | $0.9987 \pm 0.0002$ |
| 3 | 12.22 | $12.23 \pm 0.02$ | 0.4898 | $0.9918 \pm 0.0006$ |
| 4 | 11.82 | $11.84 \pm 0.02$ | $0.0215^{*}$ | $0.982 \pm 0.003$ |
| 5 | 11.65 | $11.66 \pm 0.02$ | 0.0544 | $0.969 \pm 0.006$ |
| 6 | 11.44 | $11.47 \pm 0.01$ | $0.0001^{****}$ | $0.94 \pm 0.04$ |

$$\cos\theta = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\,\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{3.2.1}$$

$A_i$ and $B_i$ are the scalar components (coefficients) of vectors (loadings) **A** and **B** with n different mass peaks. The cosine similarity $\cos\theta$ with the loadings of the full PCA is calculated for the loadings of all 10 repeats of the subsampling PCA (see Table 3.2). It should be noted that due to occasional sign inversion of the loadings during the repeats, the averages of the absolute values are presented. It can be observed that all PCs show a great degree of similarity ($> 0.9$). The first three PCs are near identical, whereas increasing dissimilarity can be observed for subsequent PCs. This also coincides with increasing coefficients of variation (CVs).

The first two principal components elucidate the three different chemistries of the sample (53.8 % variance explained), where the positive loadings of PC1 (see Figure 2C) correspond to the mass spectrum of PS, the negative loadings of PC1 correspond to the mass spectrum of PVP and the positive loadings of PC2 with ions as $Si^+$, $SiH^+$, $SiO^+$, $SiOH^+$, $Si_2O^+$ and $Si_2OH^+$ correspond to the mass spectrum of the silicon wafer. Next, the loadings are applied to the whole data set, which was first standardised with the mean and standard deviation of the training set, to generate scores for every pixel in the 3D image. The scores images were then z-scaled (with the silicon wafer interface

set at z = 0 nm). The silicon interface is established by fitting a Gaussian ($R^2_{adj}$ = 0.93) to the gradient of the average scores for PC2 in the z-direction (cf. the z-calibration with $Si^+$, as shown in Figure 3.12 B. The PS-PVP interfaces are similarly determined by Gaussian fits and their sputter times are converted into layer thicknesses. The scaled scores image for PC1 is presented in Figure 3.13 C. The SNR for the PS (2.4) and PVP signal (1.35) is calculated as $\mu_{sig}/\sigma_{sig}$ of the positive and negative scores of PC1, respectively (see Figure 3.14 C). The SNR has clearly improved, specifically the SNR is 2.4 times higher for PS and 1.7 times higher for PVP. Similarly, the depth resolution for the various interfaces is calculated by fitting a Gaussian to the gradient of the average scores of PC1 in the z-direction (see Figure 3.15 C) and are not significantly different from those calculated with the ion images as shown by a pairwise t-test (P = 0.31). The average $\Delta z = (4.3 \pm 0.7)$ nm (n = 9).

## 3.3 Conclusions

A method to construct TOF-SIMS image data consisting of multiple chemistries with spatial distributions of their own is presented. This method makes it is possible to identify the effects of different preprocessing procedures such as mean-centering, standardisation (autoscaling) and normalisation.

Next, we have demonstrated that it is possible to expand the application of PCA to large 3D images without requiring any computing resources beyond a desk top computer. The method reported here presents the first time PCA has been performed on a large scale (128 x 128 x 622 pixels) 3D ToF-SIMS image. This was made possible by first calculating the PCA loadings using a smaller subset of randomly selected pixels as a training set that could then be applied to the full data set to generate the scores images. The method has been validated using a well-defined 3D ToF-SIMS data set of a PS-PVP multilayer system. The results clearly show that PCA separates the different chemistries in its loadings and provides information on spatial chemical distribution via the scores. Furthermore, the scores images have a 1.7-2.4 times better signal-to-noise ratio than can be obtained with single ions. The depth resolution of the scores images does not differ from that of the single ion images. Given the developments in 3D ToF-SIMS imaging, this method of facilitating PCA of large ToF-SIMS data addresses an important need as the size of the data will only continue to increase.

CHAPTER 4

# 3D ToF-SIMS Imaging of Neuronal Networks: Freeze-Dried Cell Preparation and Data Analysis[1]

The development of novel ion beams [9, 10, 45, 108] has provided Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS) with the basis to be established as a routine tool for label-free 3D chemical imaging of single cells. Label-free 3D chemical imaging is especially attractive for drug studies or toxicological risk assessment [12] and the research on 3D ToF-SIMS imaging of single cells has progressed to the point where the intracellular uptake and location of non-native compounds [39, 40] and nanoparticles [41] can be visualised.

Having developed and validated an approach to PCA of 3D ToF-SIMS images using a well-defined test data set in chapter 3, it is time to test its effectiveness on a more complex, biological sample consisting of primary rat cortical neurons that were cultured on poly-L-lysine coated glass slides. First, we have to develop a cell preparation protocol for 3D ToF-SIMS measurements, however. The proposed sample preparation consists of cryofixation followed by freeze-drying (cf. reference [48]).

As an application, we attempt to visualise the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons. In another application, we attempt to differentiate between two different cell types: primary rat cortical neurons and retinal pigment epithelium (RPE) cells.

[1]This chapter has been partly published in Analyst (RSC publishing) as 'Multivariate analysis of 3D ToF-SIMS images: method validation and application to cultured neuronal networks' [100]. Some passages are quoted verbatim from this source.

## 4.1 Primary Rat Cortical Neurons

### 4.1.1 Cell Morphology and Size

When designing a ToF-SIMS measurement protocol for cells, it is necessary to first establish the size and morphology of the cells to make sure they can be resolved with ToF-SIMS imaging. For this reason, phase contrast microscopy images – obtained by Dr. Christopher Towlson – of primary rat cortical neurons cultured for 2, 5, 6, 8 and 14 days in vitro (DIV) were analysed (see Figure 4.1). All five images were recorded with $10\times$ magnification.

After creating a thresholded binary image (see Figure 4.1), it is possible to measure the area, perimeter, circularity, Feret diameter, and minimum Feret diameter of the cells using particle analysis algorithms. Circularity is defined by equation 4.1.1. A circularity value of 1 indicates a perfect circle. As the value approaches 0, it indicates an increasingly elongated polygon. The Feret diameter is the longest distance between any two points along the selection boundary, i.e. maximum caliper. The minimum Feret diameter is the minimum caliper diameter.

$$\text{circularity} = 4\pi \left( \frac{\text{area}}{\text{perimeter}^2} \right) \tag{4.1.1}$$

Histograms of the area were made using a bin width of 10 $\mu$m$^2$ and a 0-1000 $\mu$m$^2$ range. The first few bins have a high frequency ($> 100$) attributed to non-cellular debris (see Figure 4.2). Between 80 and 400 $\mu$m$^2$, the histograms approach a normal distribution representing counts of single cells. The normality of the 80-400 $\mu$m$^2$ distribution was checked using normal quantile plots yielding coefficients of determination between 0.88 and 0.98 for the various images (see Figure 4.2). The higher areas are assigned to aggregated cells or larger debris.

After a visual check to determine whether or not the 80-400 $\mu$m$^2$ distribution truly corresponds to the cell dimensions by overlaying images of the particle outlines and the original images, the means and standard deviations of the different parameters are calculated. The fact that the standard deviations of the individual parameters overlap for all the different images, indicates that there is no significant difference between the different images. Therefore, all particles are pooled together to generate the overall means given in Table 4.1.

This image processing is rather laborious. Therefore, only a limited number of images has been processed. The large amount of user input also induces bias and some cells are

**Table 4.1:** Means and standard deviations of the different parameters with all particles pooled together (n = 851).

| Parameter | Mean ± StDev |
|---|---|
| Area | (195 ± 69) $\mu$m$^2$ |
| Perimeter | (75 ± 32) $\mu$m |
| Circularity | 0.53 ± 0.22 |
| Feret | (29 ± 13) $\mu$m |
| MinFeret | (13.2 ± 3.5) $\mu$m |

left out by only considering a certain range of results. However, this small investigation does provide a sense of the average cell size and morphology, which are needed to check if they can be resolved with ToF-SIMS imaging.

It is clear the cells are large enough to be investigated using ToF-SIMS. As stated in section 1.2, a 1-2 $\mu$m beam spot can be achieved in the high-current bunched mode of the Bi LMIG. A cell area of 195 ± 69 $\mu$m$^2$ (n = 851) should translate to 49 to 195 pixels in a ToF-SIMS image obtained in the high-current bunched mode. Sub-micrometer beam spots can be achieved in a non-bunched modes like burst-alignment (300-500 nm) or the extreme cross-over mode ($\approx$ 150 nm), but at the cost of mass resolution. Because of their inherent complexity and the close chemical similarities of most of the compounds of interest (proteins, lipids and carbohydrates), biological samples such as cells require a high mass resolution.

### 4.1.2 Cell Substrate Preparation

The next thing to be determined is the substrate on which to grow cells. Silicon wafers are sometimes used as the substrate material, because they are atomically flat [109]. In our case, transparent wafers made of glass slides are better suited as they allow the use of optical techniques.

It is also very important the cells are able to adhere to the surface. Cell adhesion is preceded by protein adsorption, and cell-surface interaction is actually an interaction between cells and surface bound proteins, e.g. fibronectin, or other biomolecules [110]. After a surface has been placed in a biological milieu containing cells, water molecules are the first to reach the surface. The properties of this surface water shell, which in turn depend on the surface properties, influence the proteins that arrive later and can determine their orientation, if they denature or not, surface coverage, etc. Finally, the

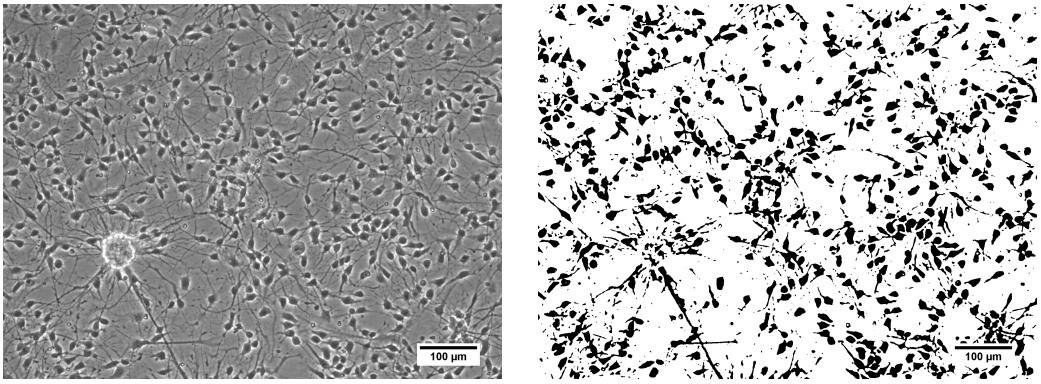



**Figure 4.1:** Microscopy (left) and thresholded binary image (right) of primary rat cortical neurons 5 days in vitro (DIV). The microscopy image was taken by Dr. Christopher Towlson.

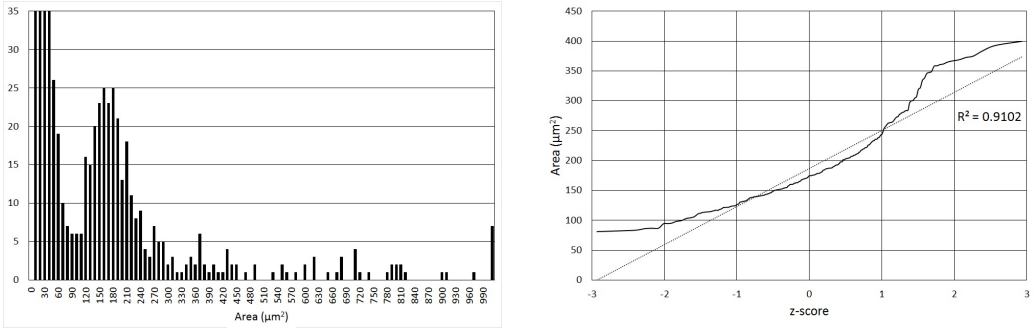



**Figure 4.2:** Histogram of the analysed particles' areas (i.e. cells and debris) for the 5 DIV image (left) and the normal quantile plot of the 80-400 $\mu$m$^2$ area distribution (right).

cells arrive at a protein-covered surface with properties that are ultimately determined by the surface properties [110]. It is therefore necessary to functionalise the glass surface in order to make it cytophilic. The glass surface is turned cytophilic by depositing a poly(L-lysine) (PLL) polymer layer on top of it.

To make correlative imaging between optical microscopy, white light interferometry and ToF-SIMS possible, a marker system is needed in order to be able to image the same cells with all three techniques. Therefore, adhesive polyester microscope slide-grids with 1 mm line spacing are stuck onto the back of the glass coverslips.

Full experimental details can be found in section 2.1.

### 4.1.3 Cell Culture

Low-density cultures of disassociated embryonic rat cortical neurons were prepared and cultured in serum-free media on the poly-L-lysine coated glass coverslips. The prepared coverslips were placed in 6-well tissue culture plates and plated with a total of 150,000 disassociated cells per coverslip. After incubating for 30 min at 37 °C to allow cell adhesion to occur, the wells were flooded to a total volume of 2 mL of Neurobasal/B27 media. After 24 h, the media was replaced with fresh media and the cultures were maintained at 37 °C in a humidified 5 % $CO_2$ atmosphere for the next 9 days.

Full experimental details can be found in section 2.2.

## 4.2 3D ToF-SIMS Images of Freeze-dried Neuronal Networks

### 4.2.1 Cell Preparation for ToF-SIMS

Before being flash-frozen by plunging the coverslips into 20-30 mL of liquid ethane for 20 s, the cell covered coverslips were dipped three times with a pair of tweezers in an ammonium formate solution that matched the measured osmolality of the cell culture media in order to prevent cytolysis. The osmolality of the cell culture media was determined with a cryoscopic osmometer and equalled 237.3 $\pm$ 4.7 mOsm/L (n = 6), which equates to 119 mM ammonium formate solution.

During the first experimental run, the temperature and pressure of the freeze-dryer chamber was monitored as it is important that the sample stays frozen until the pressure drops below that of triple point (273.16 K, 611.73 Pa). Figure 4.3 shows the evolution of the temperature and pressure of the freeze-dryer chamber over time. It can be seen that the pressure of the freeze dryer gets below 611.73 Pa after 30 minutes before the temperature rises over 0 °C. Without samples the system was able to reach 5.3 Pa after 20 min. The pressure in the system continues to drop and remains relatively constant at c. 10 Pa after t $>$ 3000 min, indicating drying is complete. The temperature in the sample chamber drops to a minimum of -3 °C after the transfer of the falcon tubes, and reaches an equilibrium (c. 12 °C) after 125 minutes. The temperature of the condensor plate was constant within 0.5 °C: (-59.29 $\pm$ 0.49) °C (n = 20).

In order to assess the success of freeze-drying process, optical microscopy images were taken before and after freeze-drying. The microscopy grid attached to the bottom of the glass coverslips allowed the same regions on the coverslips to be imaged. The before and after images are imported into Matlab and aligned using an in-house Matlab code presented in section B.2. The optical microscopy images presented in Figures 4.4 and B.1 consist of nine individual 20x images that were stitched together using ImageJ's stitching plugin [111]. In the images prior to freeze-drying, several neurons can be seen as well as their axons and dendrites. After cryofixation and freeze-drying, optical microscopy shows that the morphology of the cells is largely preserved: the same neurons can be seen as well as their axons and dendrites. However, not all neurons are preserved and there are visual signs of damage. This raises questions about how representative the samples are of the native cellular state after freeze-drying and whether this also induces chemical changes such as chemical compound diffusion.

In order to assess the surface topography of cells after the freeze-drying process, interferometry images were recorded after freeze-drying. The raw data is imported into Matlab and processed using Matlab code presented in section B.1. The microscopy

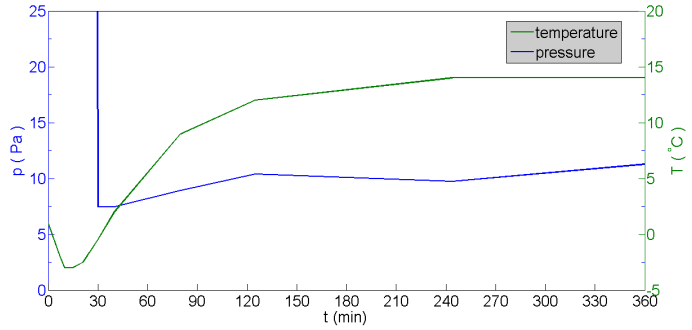



**Figure 4.3:** Temperature and pressure in the freeze-dryer chamber during the freeze-drying of cellular samples.

grid attached to the bottom of the glass coverslips allowed the same regions on the coverslips to be imaged. The interferometry images were aligned with the microscopy images using Matlab code presented in section B.2. The interferometry images of the freeze-dried cells (see Figures 4.4 and B.1) show a good correlation with their corresponding optical microscopy images. It can be seen that the dry cellular material has a topography in the order of several $\mu$m. Based on 16 individual interferometry images the average $R_{t,av}$ and $R_q$ were determined to be respectively 6.7 $\pm$ 0.7 $\mu$m and 0.57 $\pm$ 0.08 $\mu$m. The $R_{t,av}$ of an interferometry image can be considered a measure of the maximum cell height. It is known that there are complications involving the secondary ion yield in ToF-SIMS, when the sample material has a curvature or a surface topography in excess of several tens of $\mu$m [101]. Based on the interferometry data, the cells appear to show curvature and surface topography to the extent, where it could affect the secondary ion yield.

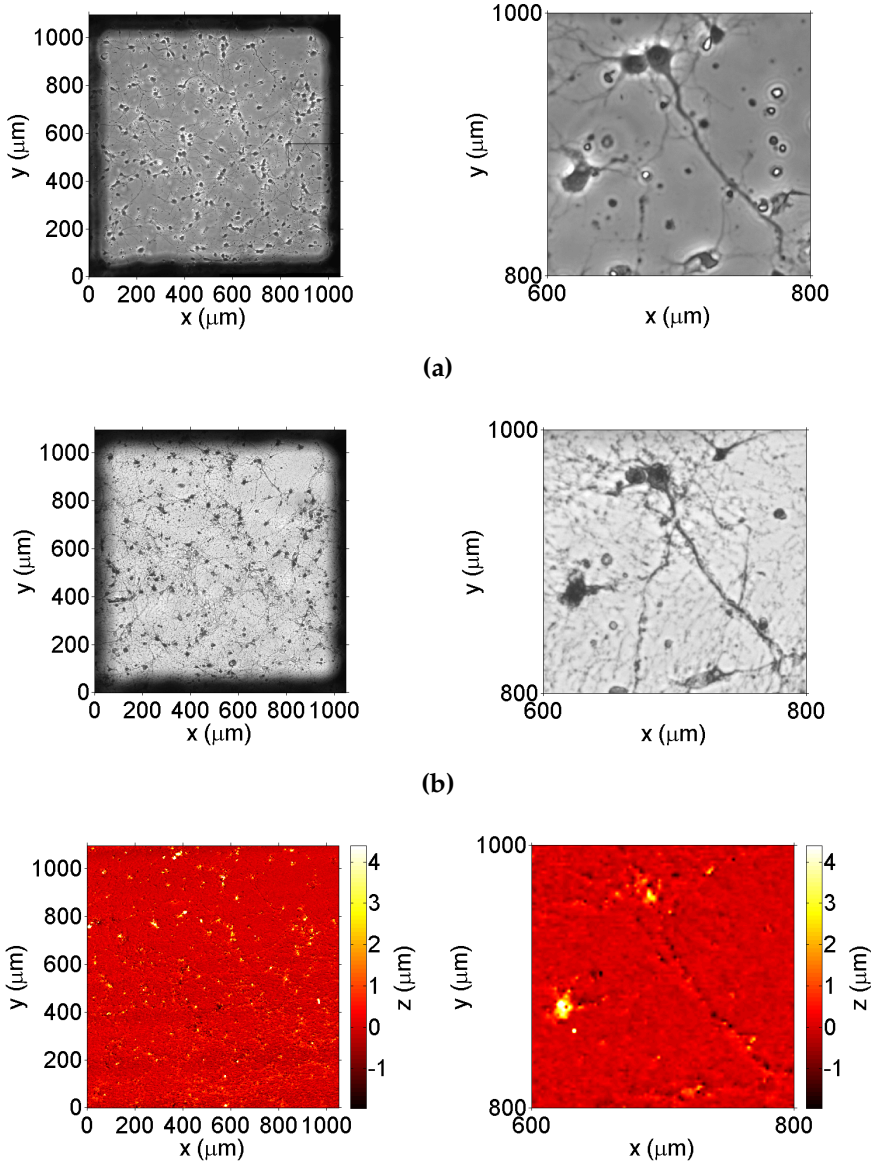



**Figure 4.4:** Correlative imaging of neuronal cell networks before and after freeze-drying. a) Differential interference contrast (DIC) microscopy image (20x) before freeze-drying (7 days in vitro). b) Bright-field microscopy image (20x) after freeze-drying. Several neurons (grey ovals) can be seen as well as their axons and dendrites. The black edges on the optical microscopy images originate from the grid fixed to the back of the microscope slide, which were used to locate cells during interferometry and TOF-SIMS measurements. c) White light interferometry heightmap ($R_{t,av}$ = 6.36 $\mu$m, $R_q$ = 0.54 $\mu$m) of the freeze-dried neuronal network.

### 4.2.2 Data Analysis using PCA

We performed a 3D ToF-SIMS imaging experiment on the samples described above (experimental details can be found in chapter 2). After mass calibration, a peak search and image reconstruction, the raw TOF-SIMS data is again imported into Matlab for data processing and analysis. The image has a size of 256 $\times$ 256$\times$ 160 pixels and the peak search extracted 173 mass peaks ($\approx 1.8 \times 10^9$ data points). The data is normalised to the total number of ion counts per pixel particularly to account for variations in the secondary ion signal due to the topography of the cell sample as well as the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling.

The variance-to-mean ratio (VMR) of the variables averages 0.023 $\pm$ 0.02 (n = 169) and all chi-square goodness of fit tests yielded p-values < 0.0001 indicating again that the variables do not follow a Poisson distribution. Prior to PCA the $Na^+$ and $K^+$ ion intensities, because of their dominance, are removed as contaminant peaks (in accordance with other studies [83]) that likely originated from the cell culture medium [48]. PCA is performed using the training set method described in section 3.2.1. The training set is formed by randomly selecting 4000 pixels per z-plane; the training set thus consists of 640,000 pixels (i.e. mass spectra) or 6.1 % of the total amount of pixels (i.e. the same relative amount of pixels as for the multilayer sample in section 3.2). The first two principal components explain 64.3% of the variance. The positive loadings of PC1 (48.8% variance explained, see Figure B.2) contain organic and higher-mass ions, whereas the negative loadings contain inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27) and $Si^+$ (m/z 28).

Biological samples such as the cells imaged here have a surface topography, which means that the 3D image created from the stacked 2D images is distorted in the vertical direction. Because PC1 differentiates between the borosilicate glass substrate and cellular material, its indication of the substrate interface (where the scores equal zero) can be utilised to apply the necessary z-offset correction to account for the surface topography of the cells. Note that this assumes a constant sputter yield through the cellular material. The Matlab code is presented in section B.3.

This computational transformation is then calibrated against interferometry data that shows a $R_{t,av}$ of 2.0 $\mu$m (see Figure 4.5 B), giving each pixel a height of 12.6 nm in the z-direction. This approach to account for topography has previously been demonstrated by Fletcher et al. [33] and is very similar to the method employed by Breitenstein et al. [34] and Robinson et al. [44] who vertically shift data points using a single ion as a

substrate marker. However, using a linear combination of ion intensities (i.e. the PCA loadings) instead of a single ion has the advantage of increased SNR (see Figure B.3), especially given the fact that each XY line has to be z-corrected individually, leading to an improved z-correction (see Figure 4.5).

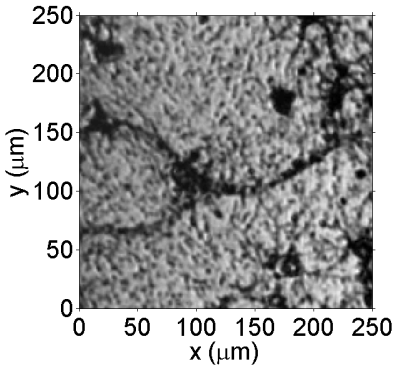
The positive loadings of PC2 (see Figure 4.6) contain a strong correlation with the ion at m/z 184, which is specific for phosphocholine-containing phospholipids and a common marker for cell membranes in ToF-SIMS analysis [33, 40]. Its fragment ions at m/z 166, 104, 86 and 58 are also present in the loadings [33]. The negative loadings of PC2 contain peaks that are commonly associated with amino acids [112] such as m/z 84 (Lys), 100 (Arg), 110 (His), 120 (Phe) and 130 (Trp).

Based on the loadings, it appears that PC2 distinguishes between the cell membrane and the cytoplasm. This supposition is strengthened by the scores plots in Figure 4.7 that show positive scores at the top of the cells (2nd analysis layer) and negative scores inside the cell material (15th analysis layer). The presence of the ion at m/z 184 only persists in the top two analysis layers, indicating that, with the given depth resolution, they originate from a 13-26 nm layer on the surface of the cell, which corresponds, with an order of magnitude, to the 8-10 nm thickness of a neuronal cell membrane [113]. In contrast, ion fragments associated with amino acids can be detected over all subsequent analysis layers in areas coinciding with the location of cells, indicating that they originate from the cytoplasm. The negative scores of the background (areas not occupied by cells) in the scores plot of analysis layer 2 are attributed to the extracellular matrix, which is supported by the disappearance of these fragments from the surrounding material in deeper analysis layers that subsequently display a score of zero, because neither lipids nor amino acids are present in the glass substrate. Notably, if single ions such as m/z 184 or m/z 130 are used instead of the principal components, the cell features are not clearly visible due to the low SNR.

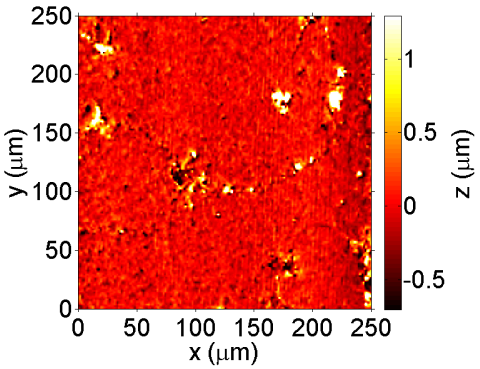
The results clearly show that PCA separates the different chemistries in its loadings and provides information on spatial chemical distribution via the scores. In addition, the PCA scores can be used to correct z-offsets due to the cells' topography. Importantly, this approach now makes 3D SIMS image processing of biological samples with multivariate analysis accessible on a routine basis and considerably facilitates data analysis.

### 4.2.3 Determining the sputter yield of freeze-dried cells

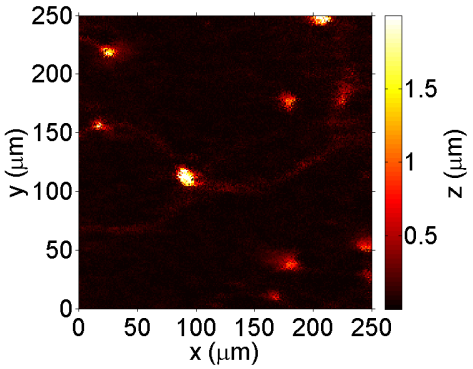
Biological samples such as the cells imaged here have a surface topography, which means that the 3D image created from the stacked 2D images is distorted in the vertical



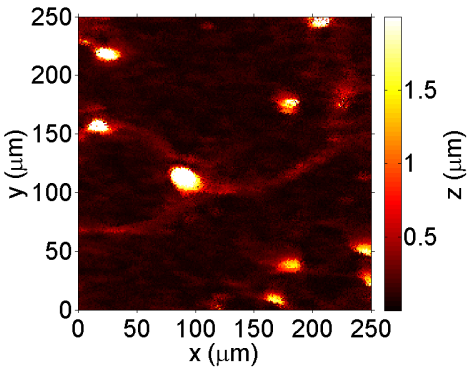

**(a)** Optical image obtained from the interferometer.




**(b)** White light interferometry heightmap ($R_{t,av}$ = 1.99 $\mu$m, $R_q$ = 0.28 $\mu$m).

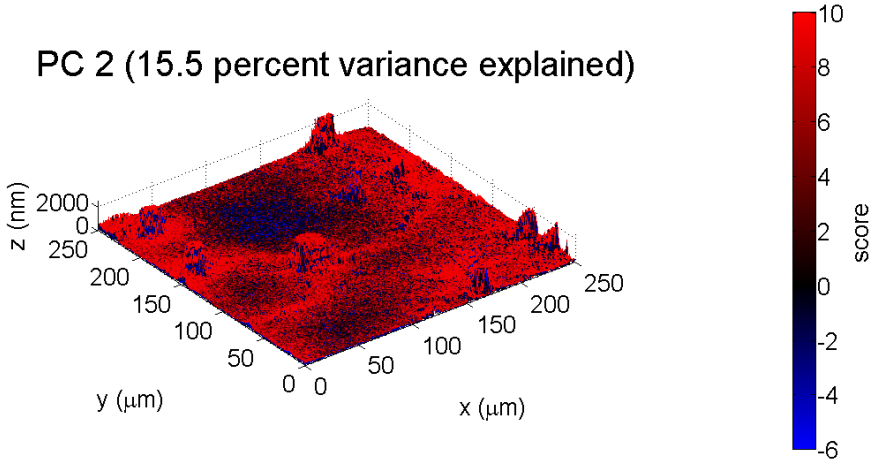



**(c)** Heightmap based on the ion intensity of $Si^+$ ($R_{t,av}$ = 0.72 $\mu$m, $R_q$ = 0.14 $\mu$m).

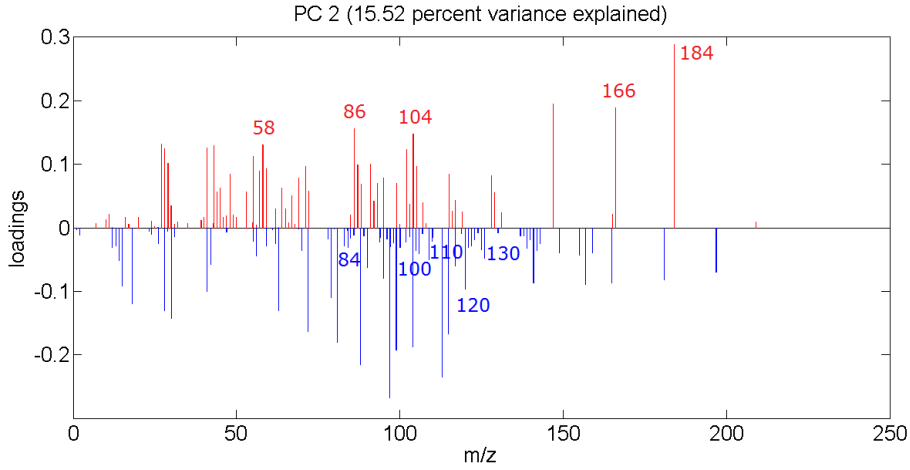



**(d)** Heightmap based on the scores for PC1 ($R_{t,av}$ = 1.17 $\mu$m, $R_q$ = 0.25 $\mu$m).

**Figure 4.5:** Illustration of the improved z-offset correction of the ToF-SIMS image of freeze-dried neurons when using a principal component instead of a single ion. For the heightmap based on the ion intensity of $Si^+$ (c), the interface is defined as $I_{normalised}$ = 0.05 and taken as a reference for the substrate plane. Scaling is performed using a maximum height of 2.0 $\mu$m based on the interferometry data (b) and assuming a constant sputter yield. In the case of the heightmap based on the scores for PC1 (d), the interface is defined as scores = 0 and taken as a reference for the substrate plane. This heightmap has more features in common with the optical image (a).




**(a)** 3D scores image.




**(b)** Loadings plot.

**Figure 4.6:** PCA of the ToF-SIMS data of the freeze-dried neuronal cell network: PC2 explains 15.5% of the variance. The positive loadings of PC2 correspond to fragments associated with lipids and the negative loadings correspond to fragments associated with amino acids.

direction. As described above, the PCA scores can be used to correct z-offsets due to the cells' topography [100]. If a component differentiates the borosilicate glass substrate from cellular material, its indication of the substrate interface can be utilised to apply the necessary z-offset correction to account for the surface topography of the cells. This computational transformation is then calibrated against interferometry data, assuming a constant sputter yield through the cellular material. It is possible to do the latter more accurately by fitting the scores based z-offset image to the interferometry heightmap.

First, the interferometry images are transformed to obtain an aligned (i.e. rotated and scaled) image with the same size as the scores based z-offset image using feature matching. Next, the pixels of the scores based z-offset image, which values represent the number of scans needed to reach the substrate interface, are fitted to the interferometry image's height values using a linear function where the slope represents the sputter yield (nm/scan) and the intercept represents the layer thickness (nm) of the lowest point in the interferometry heightmap (see Figure 4.8). Using this method the sputter yield for an individual image can be calculated accurately, resulting in a good correspondence between the interferometry heightmap and the scores based heightmap (see Figure 4.8). The Matlab code is presented in Section B.4.

Based on three separate 3D ToF-SIMS images of freeze-dried cell samples with corresponding interferometry measurements, the average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material equals (100 $\pm$ 50) $nm^3$/ion. The large variation in sputter yield between different samples might be due to instrumental variation, biological variability, or perhaps varying degrees of cell collapse during the freeze-drying process [114] resulting in different material densities. Because similar overall sputter times were observed for the different freeze-dried cell samples in combination with varying sample heights, the latter explanation seems likely.

Importantly, knowing the average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material makes complementary interferometry images redundant. It is now possible to calibrate z-offset corrections using only this known sputter yield. Although it is of course recommended to obtain complementary interferometry images for the best accuracy.

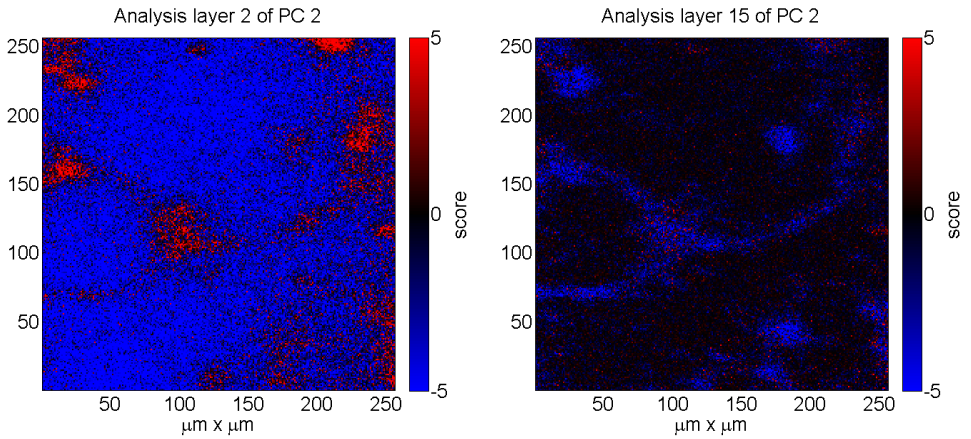



**Figure 4.7:** PC2 scores plots of analysis layers 2 and 15 of the ToF-SIMS image of the freeze-dried neuronal cell network. The PC2 scores plot for analysis layer 2 (left) shows red pixels with positive scores (lipids) in areas where cells are present and blue pixels with negative scores (amino acids) in areas without cells. The PC2 scores plot for analysis layer 15 (right) shows blue pixels with negative scores (amino acids) in areas where cells are present and black pixels with scores equal to zero (substrate) in areas without cells.

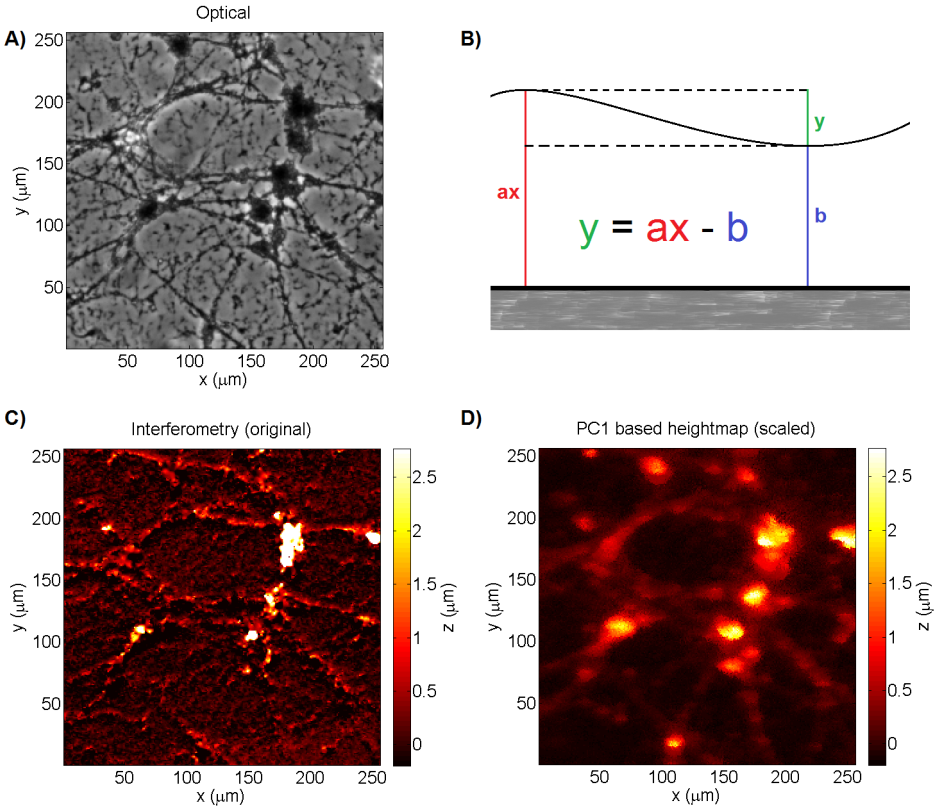



**Figure 4.8:** Scaling of z-corrected ToF-SIMS images using interferometry. A) Bright-field microscopy image (10x) of freeze-dried neuronal cells investigated with ToF-SIMS. Several neurons can be seen as well as their axons and dendrites. B) Diagram illustrating the linear relation between the scores based z-offset image and the interferometry heightmap with y the height in nm, x the number of scans, a the sputter yield in nm/scan and b the layer thickness at the lowest point in the heightmap in nm. C) White light interferometry heightmap ($R_t$ = 18 $\mu$m, $R_q$ = 0.79 $\mu$m) of the part of the neuronal network investigated with ToF-SIMS. D) A heightmap based on the scores of principal component 1, which differentiates the borosilicate glass substrate from cellular material, by fitting the z-offset image to the interferometry heightmap (a = $7.0 \pm 2 \cdot 10^{-6}$ nm/scan and b = $206.2 \pm 1 \cdot 10^{-4}$ nm). The result is a heightmap very similar to that obtained with the interferometer so that the peak signal-to-noise ratio (PSNR) equals 62 dB.

## 4.3 Application I : Visualising the Cellular Uptake of Fluorescent Dyes

Having shown that the PCA method developed in Chapter 3 is applicable to cell samples, we now attempt to visualise the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons. The dyes of interest are calcein-AM (acetoxymethyl ester of calcein) and propidium iodide (PI).

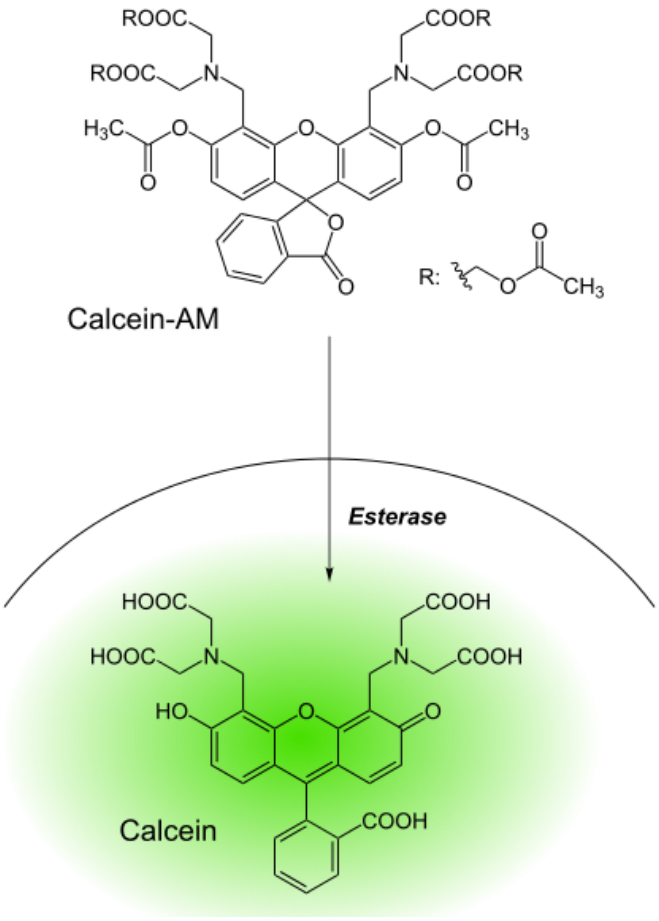
Calcein-AM, the non-fluorescent acetomethoxy derivate of calcein, can be transported through the cellular membrane into live cells. After transport through the cell membrane, intracellular esterases remove the acetomethoxy groups (see Figure 4.9). The hydrolysis of the acetomethoxy groups result in a larger conjugated ring system and a strong green fluorescence. The hydrolysis of the esters on the amine side groups also increase hydrophilicity and enable chelatation with intracellular $Ca^{2+}$, $Mg^{2+}$, $Zn^{2+}$ and other ions, causing the dye to stay inside the cell as it becomes more hydrophilic. As dead cells lack active esterases, Calcein-AM only stains viable cells.

Propidium iodide (PI) is a fluorescent molecule and an intercalating agent that is used as a DNA stain (see Figure 4.9). PI is membrane impermeant and generally excluded from viable cells. If a cell's nucleus and other DNA-containing organelles are stained, it indicates cytolysis or membrane leakage. PI is thus commonly used as a counterstain for calcein-AM in so-called viability assays.

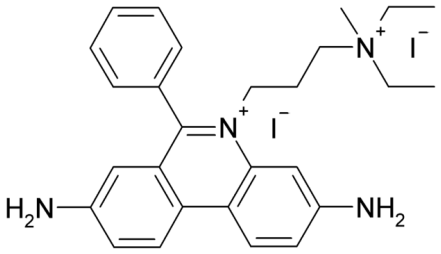
First, we need to collect reference mass spectra of the pure dyes to establish potential diagnostic ions. Next, primary rat cortical neurons will be prepared as described above, but will be stained with calcein AM and propidium iodide prior to their cryofixation. The goal is to be able to observe diagnostic ions for calcein and propidium iodide in the ToF-SIMS 3D data sets of stained primary rat cortical neurons.

### 4.3.1 Reference Spectra for Calcein AM and Propidium Iodide

The molecular propidium cation can be observed in the positive mass spectrum of propidium iodide at m/z 207 ($C_{27}H_{34}N_4{}^{2+}$) (see Figure B.4). The mass fragments at m/z 454, 385, 362, 256, 250, 164, 87, 86 and 72 in the positive ion mass spectrum of propidium iodide can be explained as logical mass fragments from the intact propidium cation initiated by one of three possible inductive cleavages of the C-N bonds (see Figure 4.10). One inductive cleavage leads to the formation of $[C_{22}H_{21}N_3.I]^+$ (m/z 454) and $[C_{22}H_{21}N_3.Cl]^+$ (m/z 362). The anion exchange from $I^-$ to $Cl^-$ could have occurred in situ or in the selvedge. The $C_{16}H_{16}N_3{}^+$ ion at m/z 250 is then formed from m/z 454




**(a)** Calcein AM is converted to calcein.




**(b)** Molecular structure of propidium iodide.

**Figure 4.9:** The dyes used for cell staining.

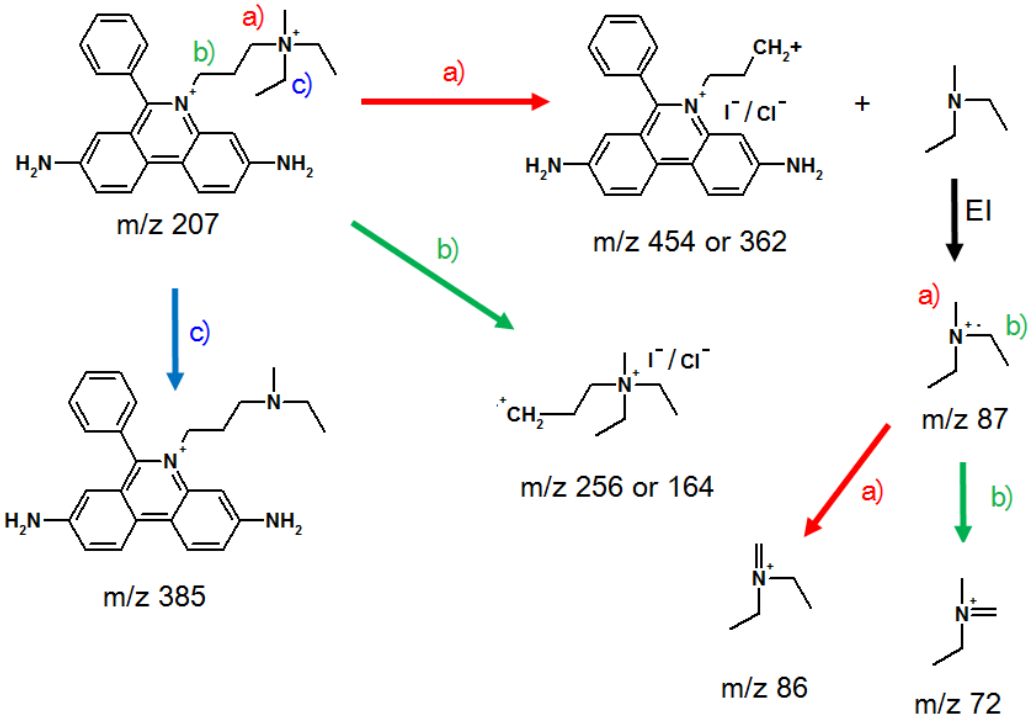



**Figure 4.10:** Fragmentation scheme of the ions at m/z 454, 385, 362, 256, 164, 87, 86 and 72 in the positive ion mass spectrum of propidium iodide.

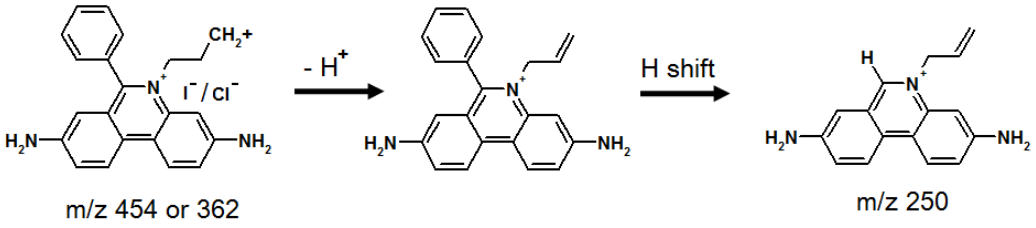



**Figure 4.11:** Fragmentation scheme of the ion at m/z 250 in the positive ion mass spectrum of propidium iodide.

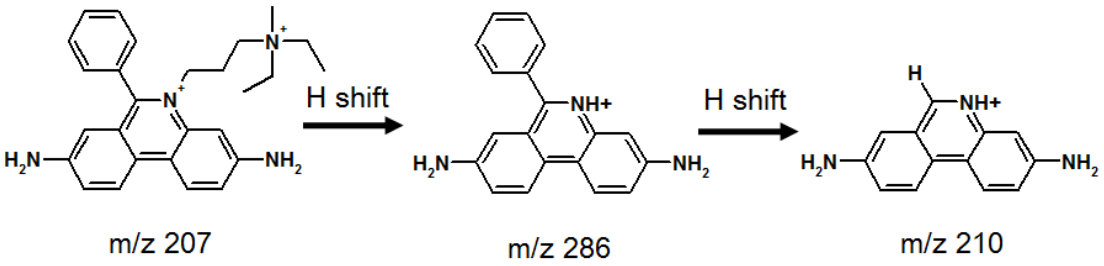



**Figure 4.12:** Fragmentation scheme of the ions at m/z 286 and 210 in the positive ion mass spectrum of propidium iodide.

and 362 by deprotonation followed by H rearrangement (see Figure 4.11). The neutral

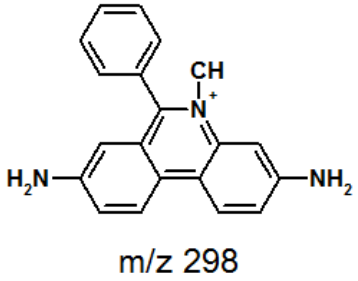



**Figure 4.13:** Molecular structure assignment for the ion at m/z 289 in the positive ion mass spectrum of propidium iodide.

fragment formed by this first possible inductive cleavage is ionised to $C_5H_{13}N^{+\cdot}$ (m/z 87), which then fragments to $C_5H_{12}N^+$ (m/z 86) and $C_4H_{10}N^+$ (m/z 72) by $\alpha$-cleavage of the first C-H and C-C bond respectively (see Figure 4.10). The second possible inductive cleavage leads to the formation of $[C_8H_{19}N.I]^+$ (m/z 256) and $[C_8H_{19}N.Cl]^+$ (m/z 164). The third possible inductive cleavage leads to the formation of $C_{25}H_{29}N_4{}^+$ at m/z 385. The ions at m/z 286 and 210 can be explained as logical mass fragments from the intact propidium ion by consecutive H rearrangements (see Figure 4.12). The molecular structure assigned to m/z 298 is shown in Figure 4.13. The $C_{20}H_{17}N_3{}^+$ ion at m/z 298 could be formed by $\alpha$-cleavage followed by a H· loss, but this would violate the even-electron rule.

The main peaks observed in the negative mass spectrum of propidium iodide (see Supplementary Figure B.5) are inorganic (cluster) ions such as $Cl^-$ (m/z 35), $I^-$ (m/z 127), $HCNI^-$ (m/z 154), $NaCNI^-$ (m/z 176), $I_2{}^-$ (m/z 254), $HI_2{}^-$ (m/z 255), $NaI_2{}^-$ (m/z 277), $KI_2{}^-$ (m/z 293), $I_3{}^-$ (m/z 381), $NaI_3{}^-$ (m/z 404) and $Na_2I_3{}^-$ (m/z 427). The negative mass spectrum of propidium iodide also exhibits the presence of hydrocarbon peaks such as $C^-$ (m/z 12), $CH^-$ (m/z 13), $C_2{}^-$ (m/z 24), $C_2H^-$ (m/z 25), $C_3{}^-$ (m/z 36), $C_3H^-$ (m/z 37), $C_4{}^-$ (m/z 48) and $C_4H^-$ (m/z 49), $C_5{}^-$ (m/z 60) and $C_5H^-$ (m/z 60), $C_6{}^-$ (m/z 72) and $C_6H^-$ (m/z 72) and nitrogen containing ions such as $CN^-$ (m/z 26), $CNO^-$ (m/z 42), $C_3N^-$ (m/z 50) and $C_6N^-$ (m/z 74). Some high mass organic peaks are observed (such as m/z 282) but with very low intensity.

Calcein-AM has a low ionisation yield and not many m/z with high intensity are observed in the positive mass spectrum. The molecular ion can be observed in the positive mass spectrum of calcein-AM at m/z 994 ($C_{46}H_{46}N_2O_{23}{}^{+\cdot}$), but with very low intensity. Many mass fragments are present in the low mass range (see Supplementary Figure B.6). The positive spectrum of calcein-AM shows very few high mass organic peaks apart from m/z 357, 358 and 359 for which the molecular structures assigned are

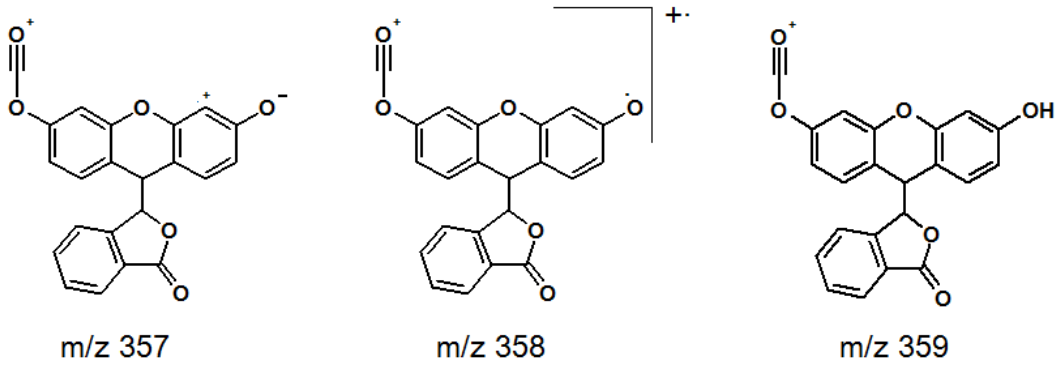



**Figure 4.14:** Molecular structure assignment for the ions at m/z 357, 358 and 359 in the positive ion mass spectrum of calcein-AM.

shown in Figure 4.14.

Calcein-AM again has a low ionisation yield and not many fragments with high intensity are observed in the negative mass spectrum. The molecular ion can be observed in the negative mass spectrum of calcein-AM at m/z 994 ($C_{46}H_{46}N_2O_{23}^{-\cdot}$). Many fragments are present in the low mass range (see Supplementary Figure B.7). The negative spectrum of calcein-AM shows very few high mass organic peaks apart from m/z 255.

In conclusion, the reference spectra show that the propidium cation has several high mass fragments in its positive spectrum that could be used as diagnostic ions. The negative mass spectrum for propidium iodide is dominated by $I^-$ containing anorganic cluster ions, which are not indicative of the propidium cation itself, which is the fluorescent, intercalating agent. The positive spectrum of calcein-AM shows very few high mass fragments and the only ions that could be used as diagnostic ions are the molecular ion at m/z 994 and m/z 357. Similarly, the negative spectrum of calcein-AM shows very few high mass fragments and the only ion that could be used as a diagnostic ion is the molecular ion at m/z 994. However, these ions can not be formed by calcein, which is the actual fluorophore to be observed in cells. If calcein displays the same fragmentation behaviour as its acetomethoxy derivative, the molecular ion ($C_{30}H_{26}N_2O_{13}^{+\cdot}$) at m/z 622 is the one to look out for.

### 4.3.2 Cell Preparation and Fluoresence Microscopy

Low-density cultures of disassociated embryonic rat cortical neurons were prepared and cultured in serum-free media on the poly-L-lysine coated glass coverslips. The prepared coverslips were placed in 6-well tissue culture plates and plated with a total of 150,000 disassociated cells per coverslip. After incubating for 30 min at 37 °C

to allow cell adhesion to occur, the wells were flooded to a total volume of 2 mL of Neurobasal/B27 media. After 24 h, the media was replaced with fresh media and the cultures were maintained at 37 °C in a humidified 5 % $CO_2$ atmosphere for the next 9 days. Full experimental details can be found in section 2.2.

For fluorescence staining, 100 $\mu$L of a 2 $\mu$M calcein-AM and 4 $\mu$M propidium iodide solution in PBS was added to each well and incubated for 15 minutes before viewing the labelled cells under the fluorescence microscope. As soon as possible (less than 40 minutes) after adding the working solution, labelled cells were cryofixated and freeze-dried.

Before being flash-frozen, the cell covered coverslips were dipped three times with a pair of tweezers in an ammonium formate solution that matched the measured osmolality of the cell culture media. The osmolality of the cell culture media was determined with a cryoscopic osmometer and equalled 242.2 $\pm$ 5.2 mOsm/L (n = 6), which equates to 121 mM ammonium formate solution.

In order to assess the success of the freeze-drying process and the surface topography of the cells, optical microscopy and interferometry images were taken after freeze-drying.

Calcein and propidium iodide remained fluorogenic after the freeze-drying process. Calcein's fluorescence microscopy image after freeze-drying indicates that the cells were viable prior to freeze-drying (see Figures 4.16 and 4.18). Importantly, these images allow us to match ToF-SIMS ion images with the known spatial distribution of both dyes.

### 4.3.3 3D ToF-SIMS Imaging

After mass calibration, a peak search and image reconstruction, the raw TOF-SIMS data is imported into Matlab for data processing and analysis. The 3D ToF-SIMS image in the positive polarity has a size of 256 $\times$ 256 $\times$ 410 pixels and the peak search extracted 153 mass peaks ($\approx 4.1 \times 10^9$ data points). The data is normalised to the total number of ion counts per pixel particularly to account for variations in the secondary ion signal due to the topography of the cell sample as well as the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling. Because of the dominance of the $Na^+$ and $K^+$ ion intensities, these ions are removed as contaminant peaks (in accordance with other studies [83]). The training set is formed by randomly selecting 1500 pixels per z-plane; the training set thus consists of 615.000 pixels (i.e. mass spectra) or 2.29% of the total amount of pixels.

The first three principal components explain 59.48% of the variance. The positive load-

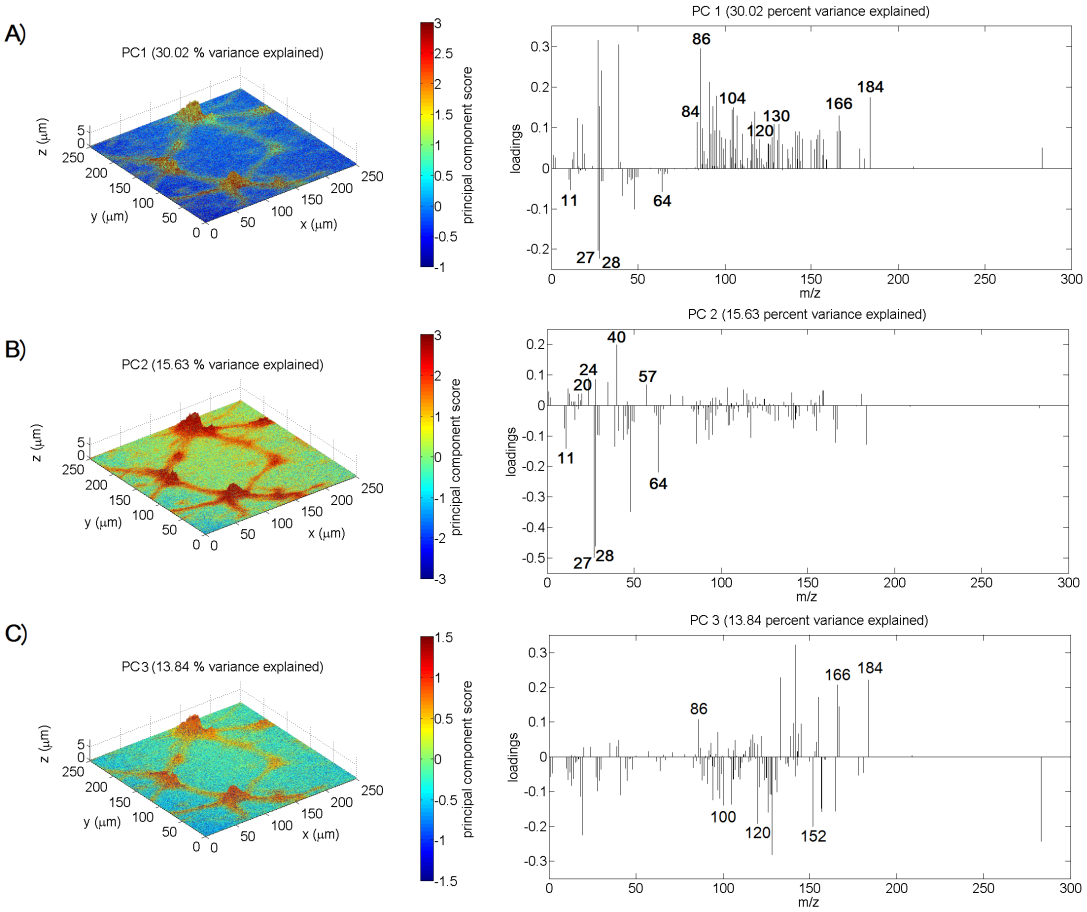


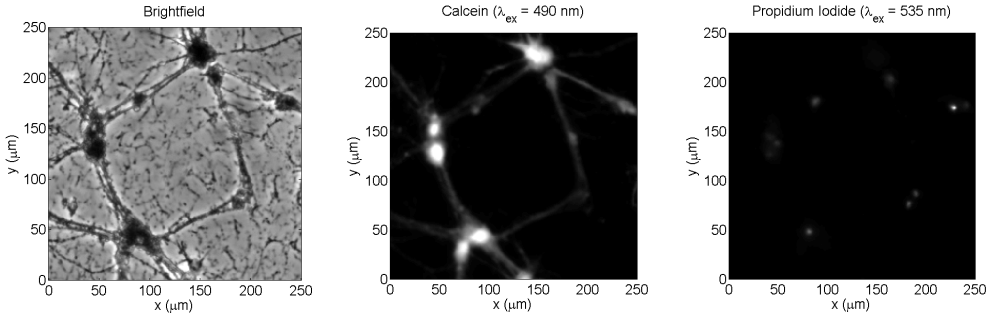

**Figure 4.15:** PCA of the freeze-dried neuronal cells stained with calcein-AM and propidium iodide and investigated with ToF-SIMS in the positive polarity after removing $Na^+$ and $K^+$ peaks from the mass spectrum; 3D scores images (left) and loadings plots (right). A) PC1 explains 30.02% of the variance. The positive loadings correspond to fragments associated with bio-organic material and the negative loadings correspond to fragments associated with the borosilicate glass. B) PC2 explains 15.63% of the variance. The positive loadings of PC2 are a measure of intracellular calcium and magnesium. C) PC3 explains 13.84% of the variance. The positive loadings correspond to fragments associated with phospholipids and the negative loadings appear to be associated with amino acids and nucleic material.

ings of PC1 (30.02% variance explained, see Figure 4.15 A) contain organic and higher-mass ions that can be attributed to bio-organic material such as phosphatidylcholine and amino acids. A common marker for phosphocholine-containing phospholipids is the ion at m/z 184 and its fragment ions at m/z 166, 104 and 86 [33, 40]. Peaks that are commonly associated with amino acids are for example m/z 84 (Lys), 120 (Phe) and 130 (Trp) [33]. The negative loadings of PC1 contain inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27), $Si^+$ (m/z 28) and $Zn^+$ (m/z 64).

Because PC1 differentiates between the borosilicate glass substrate and cellular material, its indication of the substrate interface (where the scores equal zero) is utilised to apply the necessary z-offset correction to account for the surface topography of the cells. Note that this assumes a constant sputter yield through the cellular material. This computational transformation is then calibrated against interferometry data. Figure B.8 C shows the heightmap based on the scores of principal component 1.

PC2 explains 15.63% of the variance (see Figure 4.15 B). The positive loadings correspond to $Ca^{2+}$ (m/z 20), $Mg^+$ (m/z 24), $Ca^+$ (m/z 40) and $CaOH^+$ (m/z 57) as well as other inorganic salt clusters such as ${K_2}^+$ (m/z 78), $K_2CN^+$ (m/z 104), $K_2Cl^+$ (m/z 113), $K_2{}^{37}Cl^+$ (m/z 115), $NaKPO_3{}^+$ (m/z 141), $K_2PO_3{}^+$(m/z 157) and $NaP_2O_6{}^+$ (m/z 181). The positive loadings also contain some high mass organic ions such as m/z 100, 118, 120, 130, 136 and 152. $C_5H_6N_5{}^+$ (m/z 136, A) and $C_5H_6N_5O^+$ (m/z 152, G) are DNA peaks [115] whereas $C_4H_{10}N_3{}^+$ (m/z 100, Arg), $C_8H_{10}N^+$ (m/z 120, Phe), $C_9H_8N^+$ (m/z 130, Trp) and $C_8H_{10}NO^+$ (m/z 136, Tyr) are amino acid peaks [112]. The negative loadings of PC2 resemble those of PC1 with high coefficients for inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27), $Si^+$ (m/z 28) and $Zn^+$ (m/z 64). The negative loadings also contain high mass organic ions, notably phospholipid peaks such as $C_5H_{12}N^+$ (m/z 86), $C_5H_{14}NO^+$ (m/z 104), $C_5H_{13}NPO_3{}^+$ (m/z 166) and $C_5H_{15}NPO_4{}^+$ (m/z 184). The positive loadings of PC2 are therefore a measure of intracellular calcium and magnesium. Perhaps unsurprising, the summed 2D image of the $Ca^{2+}$ (m/z 20), $Mg^+$ (m/z 24), $Ca^+$ (m/z 40) and $CaOH^+$ (m/z 57) ions without the substrate ($z > 0$ $\mu$m) resembles the calcein fluorescence microscopy image (PSNR = 41.47, see Figure 4.17). Calcein after all chelates intracellular $Ca^{2+}$ and $Mg^{2+}$. And perhaps surprising, not a single ion that can be attributed to calcein or cannot be explained as a cellular mass peak is detected in the positive loadings of PC2.

PC3 explains 13.84% of the variance (see Figure 4.15 C). The positive loadings correspond to phospholipid peaks such as $C_5H_{12}N^+$ (m/z 86), $C_5H_{13}NPO_3{}^+$ (m/z 166) and $C_5H_{15}NPO_4{}^+$ (m/z 184). The negative loadings correspond to amino acid peaks such




**(a)** Bright-field. **(b)** Calcein. **(c)** Propidium.

**Figure 4.16:** Live/dead fluorescence staining of neuronal cells investigated with ToF-SIMS in the positive polarity. A) Bright-field microscopy image (10×) of freeze-dried neuronal cells stained with calcein-AM and propidium iodide. Several neurons can be seen as well as their axons and dendrites. B) Fluorescence microscopy image (10×, $\lambda_{ex}$ = 490 nm, $\lambda_{em}$ = 515 nm) for calcein indicating viable cells. C) Fluorescence microscopy image (10×, $\lambda_{ex}$ = 535 nm, $\lambda_{em}$ = 617 nm) for propidium iodide indicating dead cells.

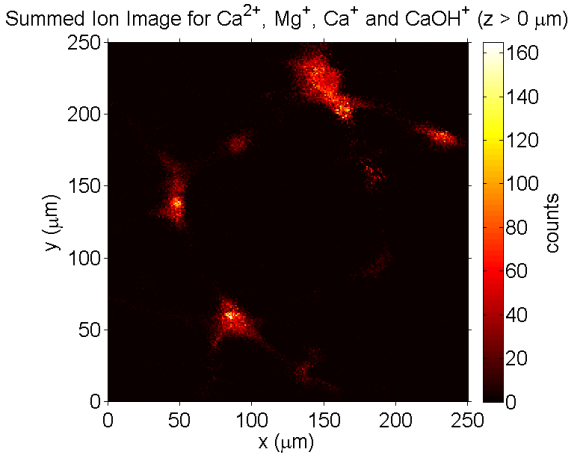



**Figure 4.17:** The summed 2D image of the $Ca^{2+}$ (m/z 20), $Mg^{+}$ (m/z 24), $Ca^{+}$ (m/z 40) and $CaOH^{+}$ (m/z 57) ions without the substrate ($z > 0$ $\mu$m) for neuronal cells investigated with ToF-SIMS in the positive polarity.

as $C_4H_{10}N_3^+$ (m/z 100, Arg), $C_5H_8N_3^+$ (m/z 110 Arg/His), $C_8H_{10}N^+$ (m/z 120, Phe), $C_9H_8N^+$ (m/z 130, Trp) and $C_8H_{10}NO^+$ (m/z 136, Tyr) as well as DNA peaks such as $C_5H_7N_2O_2^+$ (m/z 127, T) $C_5H_6N_5^+$ (m/z 136, A) and $C_5H_6N_5O^+$ (m/z 152, G). PC3 thus distinguishes lipids from amino acids and nucleic material. Again, not a single ion that can be attributed to calcein or propidium is detected in the loadings.

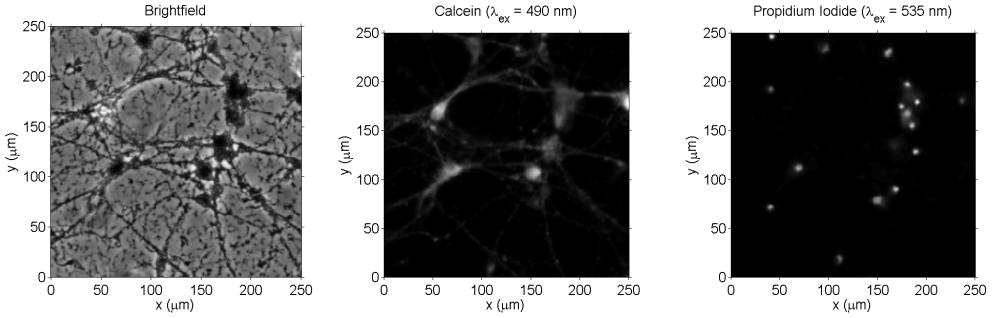
The 3D ToF-SIMS image in the negative polarity has a size of 256 $\times$ 256 $\times$ 425 pixels and the peak search extracted 288 mass peaks ($\approx$ 8.0 $\times$ $10^9$ data points). The data is normalised to the total number of ion counts per pixel particularly to account for variations in the secondary ion signal due to the topography of the cell sample as well as the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling. The training set is formed by randomly selecting 1000 pixels per z-plane; the training set thus consists of 425.000 pixels (i.e. mass spectra) or 1.53% of the total amount of pixels.

The first two principal components explain 61.65% of the variance. The positive loadings of PC1 (49.99% variance explained, see Figure 4.19 A) contain ions that can be attributed to bio-organic material [40] such as $CN^-$ (m/z 26), $PO_2^-$ (m/z 63), $PO_3^-$ (m/z 79). The negative loadings of PC1 correspond to inorganic mass peaks that can easily be attributed to the silica ($SiO_2$) and alumina ($Al_2O_3$) [116, 117] present in the borosilicate glass such as $O^-$ (m/z 16), $AlO^-$ (m/z 43), $SiO_2^-$ (m/z 60), $SiO_3^-$ (m/z 76), $Al_2O_4H^-$ (m/z 119), $Si_2O_5^-$ (m/z 136) and $Al_3O_6H_2^-$ (m/z 179).

Because PC1 differentiates between the borosilicate glass substrate and cellular material, its indication of the substrate interface (where the scores equal zero) is utilised to apply the necessary z-offset correction to account for the surface topography of the cells. Note that this assumes a constant sputter yield through the cellular material. This computational transformation is then calibrated against interferometry data. Figure B.9 C shows the heightmap based on the scores of principal component 1.

PC2 explains 11.66% of the variance (see Figure 4.19 B). The positive loadings of PC2 are somewhat similar to the negative loadings of PC1 and correspond to OH- (m/z 17), CN- (m/z 26), Cl- (m/z 35), $SiO_3H^-$ (m/z 77), $Si_2O_5H^-$ (m/z 137), $Si_3O_7H^-$ (m/z 197). The negative loadings of PC2 contain DNA peaks such as $C_4N_3^-$ (m/z 90, A), $C_4H_3N_2O_2^-$ (m/z 111,T) $C_5H_4N_5^-$ (m/z 134, A) [115] and a fatty acid peak (16:0) at m/z 255 ($C_{16}H_{31}O_2^-$) [118] but also inorganic ions such as $NaCl_2^-$ (m/z 93). Overall, PC2 appears to make a distinction between the substrate or more general organic material versus nuclear material.

Unfortunately, no ions in the negative spectrum can be attributed to calcein or propidium iodide in a conclusive manner.




**(a)** Bright-field. **(b)** Calcein. **(c)** Propidium.

**Figure 4.18:** Live/dead fluorescence staining of neuronal cells investigated with ToF-SIMS in the negative polarity. A) Bright-field microscopy image (10×) of freeze-dried neuronal cells stained with calcein-AM and propidium iodide. Several neurons can be seen as well as their axons and dendrites. B) Fluorescence microscopy image (10×, $\lambda_{ex}$ = 490 nm, $\lambda_{em}$ = 515 nm) for calcein indicating viable cells. C) Fluorescence microscopy image (10×, $\lambda_{ex}$ = 535 nm, $\lambda_{em}$ = 617 nm) for propidium iodide indicating dead cells..

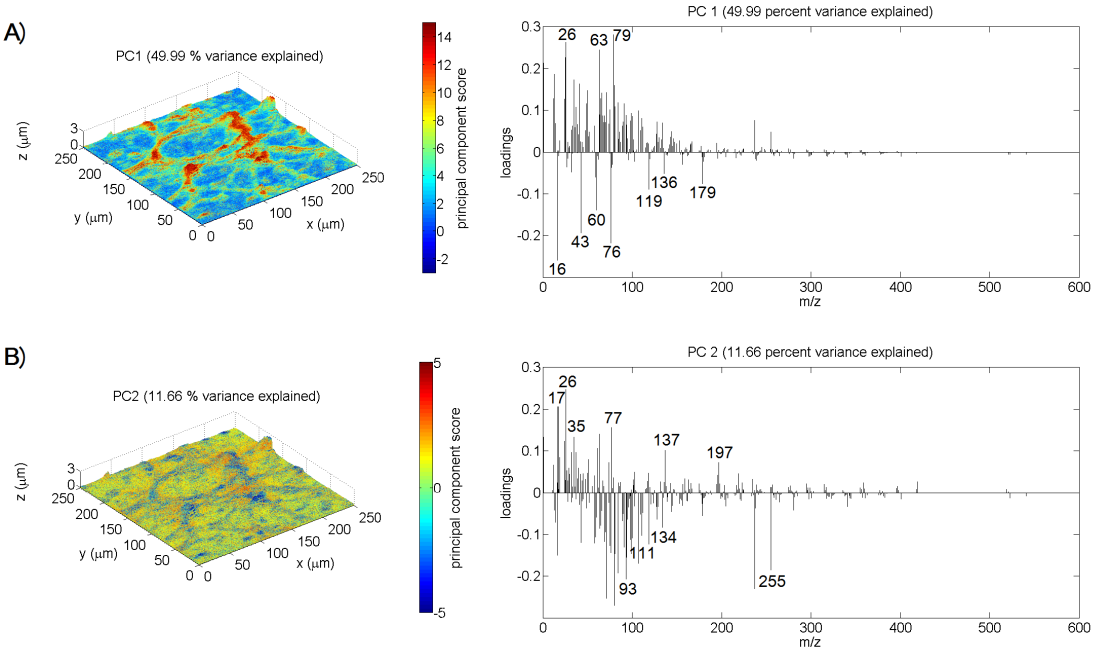



**Figure 4.19:** PCA of the freeze-dried neuronal cells stained with calcein-AM and propidium iodide and investigated with ToF-SIMS in the negative polarity; 3D scores images (left) and loadings plots (right). A) PC1 explains 49.99% of the variance. The positive loadings correspond to fragments associated with bio-organic material and the negative loadings correspond to fragments associated with the borosilicate glass. B) PC2 explains 11.66% of the variance. PC2 appears to make a distinction between the substrate or more general organic material versus nuclear material.

## 4.4 Application II : Differentiation between Neurons and RPE cells

In a second application, we attempt to differentiate between two different cell types: primary rat cortical neurons and retinal pigment epithelium (RPE) cells.

Retinal pigment epithelium (RPE) is a monolayer of highly polarized cells which forms the outer layer of the retina, they are responsible for the transport of nutrients from the vascular choroid, the formation of the blood-retinal barrier and absorption of scattered light. They are also in charge of secreting some factors with trophic effect. In order to absorb scattered light on the retina and thus diminish the photo-oxidative stress as well as improve the quality of the optical system, the cells are densely packed with melanin granules. Eumelanin and pheomelanin have distinct fragment ions in a mass range between m/z = 150 and m/z= 800 [119]. It might therefore be possible to differentiate RPE cells from neuronal cells based on the presence of melanin mass fragments.

### 4.4.1 ARPE-19

The sample consists of ARPE-19, which is an in-vitro model of native RPE. ARPE-19 cells are well pigmented and are arranged in a regular hexagonal mosaic.

It was first determined whether or not primary rat cortical neurons and ARPE-19 cells could survive in coculture. The ARPE-19 cells are prepared as described in the appendix (see subsection 2.2.2) and then seeded on the poly-L-lysine coated glass slides used for neuronal cells in the Neurobasal/B27 media used for neuronal cells without and in the presence of primary rat cortical neurons. Figure 4.20 B shows that ARPE-19 is able to survive in Neurobasal medium/B27 supplement and figure 4.20 C shows that ARPE-19 is able to survive in co-culture with primary rat cortical neurons as well. ARPE-19 are 30-40 $\mu$m in size and have a morphology distinct from that of neurons.

ARPE-19 samples in Neurobasal medium/B27 supplement were then flash-frozen and freeze-dried in an identical manner as the neurons.

### 4.4.2 3D ToF-SIMS Imaging of RPE Cells

Figure 4.21 A shows a bright-field microscopy images of the freeze-dried RPE cells investigated with ToF-SIMS in the positive polarity. An out of focus bright-field microscopy image (see Figure 4.21 B) indicates that the cells are elevated and therefore intact.

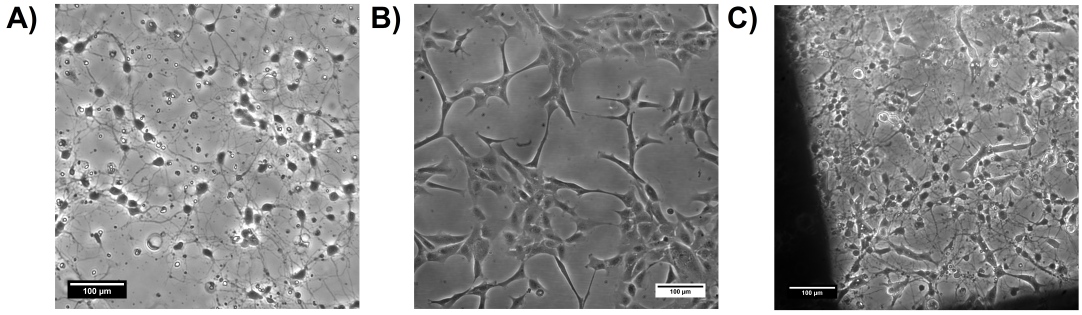



**Figure 4.20:** Morphology of primary rat cortical neurons and ARPE-19 cells. A) Bright-field microscopy image (20x) of primary rat cortical neurons. B) Bright-field microscopy image (20x) of ARPE-19 cells. C) Bright-field microscopy image (20x) of a coculture of primary rat cortical neurons and ARPE-19 cells. The scale bars are 100 $\mu$m in size.

After mass calibration, a peak search and image reconstruction, the raw TOF-SIMS data is imported into Matlab for data processing and analysis. The image has a size of 256 $\times$ 256 $\times$ 170 pixels and the peak search extracted 163 mass peaks ($\approx 1.8 \times 10^9$ data points). The data is normalised to the total number of ion counts per pixel particularly to account for variations in the secondary ion signal due to the topography of the cell sample as well as the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling. Because of the dominance of the $Na^+$ and $K^+$ ion intensities, they are removed as contaminant peaks (in accordance with other studies [83]). The training set is formed by randomly selecting 1000 pixels per z-plane; the training set thus consists of 170.000 pixels (i.e. mass spectra) or 1.53% of the total amount of pixels.

The first three principal components explain 73.42% of the variance. The positive loadings of PC1 (51.24% variance explained, see Figure 4.22 A) contain organic and higher-mass ions that can be attributed to bio-organic material such as phosphatidylcholine and amino acids. A common marker for phosphocholine-containing phospholipids is the ion at m/z 184 and its fragment ions at m/z 166, 104, 86 and 58 [33, 40]. Peaks that are commonly associated with amino acids are for example m/z 84 (Lys), 100 (Arg), 110 (His), 120 (Phe) and 130 (Trp) [33]. The negative loadings of PC1 contain inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27), $Si^+$ (m/z 28), $SiH^+$ (m/z 29), $CaH^+$ (m/z 41), $SiO^+$ (m/z 44), $SiOH^+$ (m/z 45) and $Zn^+$ (m/z 64).

Biological samples such as the cells imaged here have a surface topography, which means that the 3D image created from the stacked 2D images is distorted in the vertical direction. Because PC1 differentiates between the borosilicate glass substrate and cel-

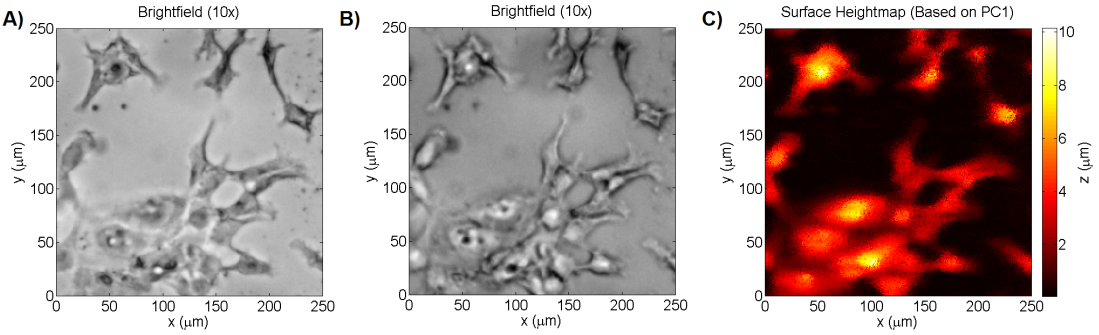



**Figure 4.21:** Freeze-dried RPE cells investigated with ToF-SIMS in the positive polarity. A) Bright-field microscopy image (10x) of freeze-dried RPE cells investigated with ToF-SIMS. B) Out of focus bright-field microscopy image (10x) of freeze-dried RPE cells investigated with ToF-SIMS, indicating elevated and intact cells. C) A heightmap based on the scores of principal component 1, which differentiates the borosilicate glass substrate from cellular material. The z-offset was calibrated to the previously calculated average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material, which equals $(100 \pm 50)$ nm$^3$/ion.

lular material, its indication of the substrate interface (where the scores equal zero) can be utilised to apply the necessary z-offset correction to account for the surface topography of the cells. Note that this assumes a constant sputter yield through the cellular material. This computational transformation is then calibrated using the previously calculated average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material, which equals $(100 \pm 50)$ nm$^3$/ion (based on three separate 3D ToF-SIMS images of freeze-dried cell samples with corresponding interferometry measurements, see section 4.2.3). Knowing the average sputter yield in nm$^3$/ion and that the $Ar^+_{5000}$ beam was rastered over a $250 \times 250$ $\mu$m$^2$ area with a target current of 1.2 nA for 5 s between analysis scans, we can calculate the sputter yield in nm/scan:

$$\frac{(100\ \text{nm}^3/\text{ion}) \times (1.2\ \text{nA}/e)}{(250 \times 250)\mu\text{m}^2} \times \frac{5\text{s}}{\text{scan}} = (60 \pm 30)\ \text{nm/scan} \tag{4.4.1}$$

and convert the number of scans to a height in nm. This gives each voxel a height of 60 nm in the z-direction. Figure 4.21 C shows the heightmap based on the scores of principal component 1.

PC2 explains 13.11% of the variance (see Figure 4.22 B). The positive loadings correspond to cluster ions generated from salts (see Table 4.2) and the negative loadings are associated with fragments associated with bio-organic material. The positive loadings of PC2 are therefore a measure of the presence of salts in the cell bodies.

PC3 explains 9.07% of the variance. The positive loadings correspond to fragments as-

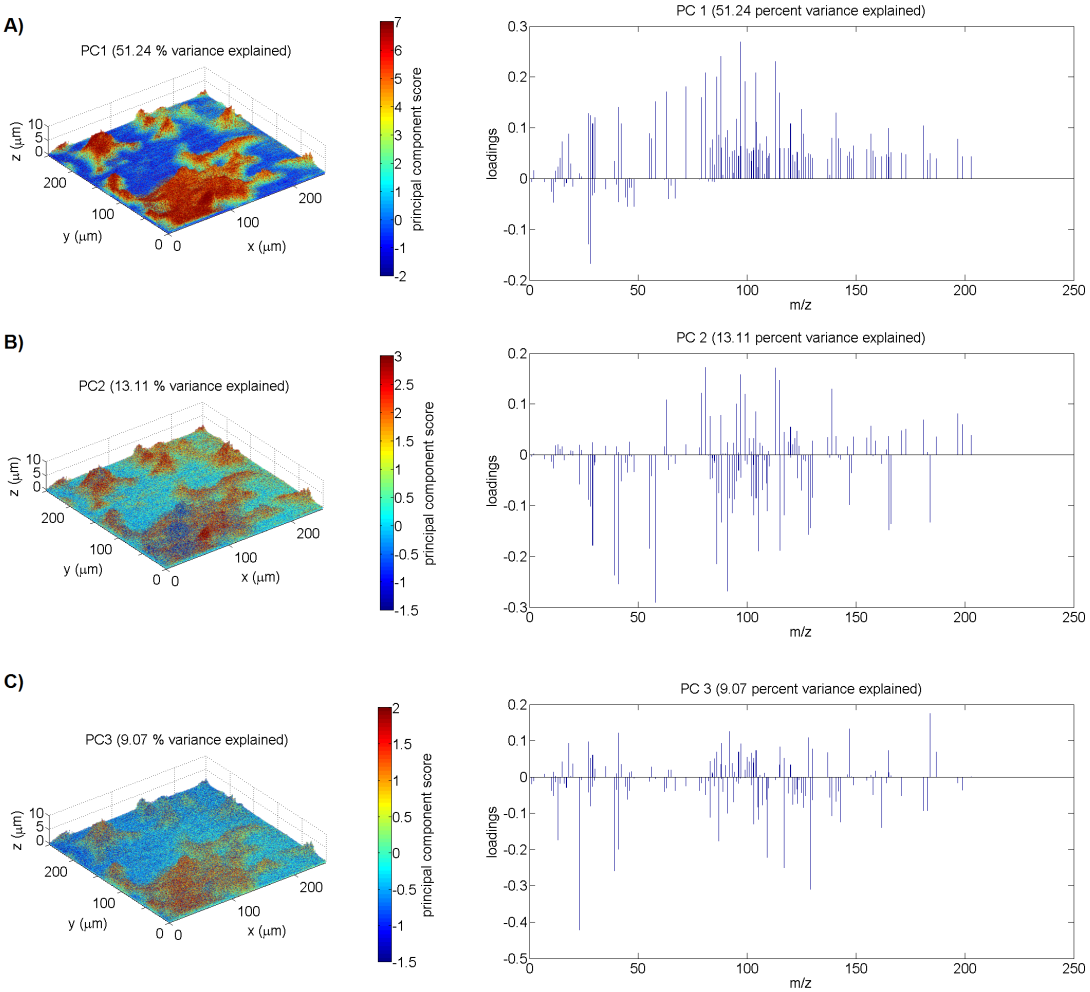



**Figure 4.22:** PCA of the freeze-dried RPE cells investigated with ToF-SIMS in the positive polarity after removing $Na^+$ and $K^+$ peaks from the mass spectrum; 3D scores images (left) and loadings plots (right). A) PC1 explains 51.24% of the variance. The positive loadings correspond to fragments associated with bio-organic material and the negative loadings correspond to fragments associated with the borosilicate glass. B) PC2 explains 13.11% of the variance. The positive loadings correspond to cluster ions generated from salts and the negative loadings are associated with fragments associated with bio-organic material. C) PC3 explains 9.07% of the variance. The positive loadings correspond to fragments associated with bio-organic material and the negative loadings appear to be associated with material surrounding the cells.

**Table 4.2:** Positive ions indicative of cluster ions generated from salts.

| Nominal Mass | Molecular Formula | Exact Mass |
|---|---|---|
| 46 | $Na_2^+$ | 45.979540 |
| 63 | $Na_2OH^+$ | 62.982280 |
| 72 | $Na_2CN^+$ | 71.982614 |
| 78 | $K_2^+$ | 77.927416 |
| 79 | $NaKOH^+$ | 78.956218 |
| 81 | $Na_2Cl^+$ | 80.948393 |
| 83 | $Na_2{}^{37}Cl^+$ | 82.945443 |
| 88 | $NaKCN^+$ | 87.956551 |
| 95 | $K_2OH^+$ | 94.930156 |
| 97 | $NaKCl^+$ | 96.922330 |
| 99 | $NaK^{37}Cl^+$ | 98.91938 |
| 103 | $Na_3(OH)_2^+$ | 102.974790 |
| 104 | $K_2CN^+$ | 103.930489 |
| 113 | $K_2Cl^+$ | 112.896268 |
| 115 | $K_2{}^{37}Cl^+$ | 114.893318 |
| 119 | $Na_2K(OH)_2^+$ | 118.948728 |
| 137 | $Na_2K(CN)_2^+$ | 136.949395 |
| 139 | $Na_3Cl_2^+$ | 138.907016 |
| 141 | $Na_3{}^{37}ClCl^+$ | 140.904066 |
| | $NaKPO_3^+$ | 140.911986 |
| 155 | $Na_2KCl_2^+$ | 154.880954 |
| 157 | $Na_2K^{37}ClCl^+$ | 156.878004 |
| | $K_2PO_3^+$ | 156.885924 |
| 171 | $NaK_2Cl_2^+$ | 170.854891 |
| 173 | $NaK_2{}^{37}ClCl^+$ | 172.851941 |
| 181 | $NaP_2O_6^+$ | 180.906786 |
| 197 | $KP_2O_6^+$ | 196.880724 |

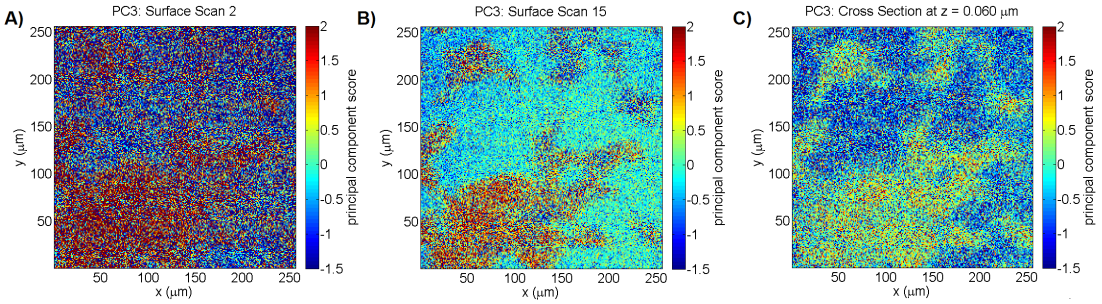



**Figure 4.23:** PC3 scores plots of the freeze-dried RPE cells investigated with ToF-SIMS in the positive polarity after removing $Na^+$ and $K^+$ peaks from the mass spectrum. PC3 explains 9.07% of the variance. A) The PC3 scores plot for analysis layer 2 shows red pixels with positive scores in areas where cells are present and blue pixels with negative scores in areas without cells. B) The PC3 scores plot for analysis layer 15 shows red pixels with positive scores in areas where cells are present and pixels with scores equal to zero (substrate) in areas without cells. C) The PC3 scores plot for the cross section at z = 0.060 $\mu$m shows pixels with positive scores in areas where cells are present and blue pixels with negative scores in areas without cells.

sociated with bio-organic material and the negative loadings appear to be associated with material surrounding the cells. This supposition is strengthened by the scores plot of surface scan number 2 (see Figure 4.23 A) that shows positive scores at regions containing cells and negative scores outside of the cell material. The scores of the background (areas not occupied by cells) subsequently display a score of zero in surface scan number 15 (see Figure 4.23 B), because neither lipids nor amino acids are present in the glass substrate. Figure 4.23 C is the cross section at z = 0.060 $\mu$m, which again shows positive scores at regions containing cells and negative scores outside of the cell material. These observations support the conclusion that the negative loadings of PC3 relate to a component that is present in the extracellular matrix. Several organic mass peaks such as m/z 129, 123, 117, 109, 105, 103, 87 and 39 are present in the negative loadings. Some of these such as m/z 129 ($C_{10}H_9^+$), 91 ($C_7H_7^+$), 55 ($C_4H_7^+$) and 39 ($C_3H_3^+$) could point towards the presence of (polycyclic) aromatic hydrocarbons [120]. Mass peaks at m/z 109 ($C_8H_{13}^+$) and 95 ($C_7H_{11}^+$) are known fragments of cholesterol [121, 122]. Mass peaks at m/z 117, 105 and 91 can also be attributed to cholesterol [121]. However, the identification of individual compounds such as cholesterol is difficult, because only low mass fragment ions are observed. Only if a (near) molecular ion such as $[M\text{-}H]^+$ at m/z 385.3 ($C_{27}H_{46}O^+$) or $[M\text{-}OH]^+$ at m/z 369.3 ($C_{27}H_{45}^+$) can be resolved, unambiguous identification is possible.

Figure 4.24 A shows a bright-field microscopy images of the freeze-dried RPE cells

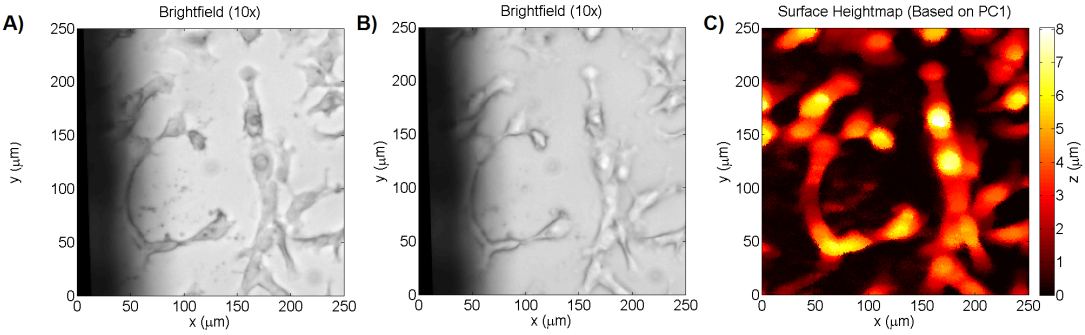



**Figure 4.24:** Freeze-dried RPE cells investigated with ToF-SIMS in the negative polarity. A) Bright-field microscopy image (10x) of freeze-dried RPE cells investigated with ToF-SIMS. The dark edge on the left belongs to the 1 mm spaced sample grid to aid in locating specific cells. B) Out of focus bright-field microscopy image (10x) of freeze-dried RPE cells investigated with ToF-SIMS, indicating elevated and intact cells. C) A heightmap based on the scores of principal component 1, which differentiates the borosilicate glass substrate from cellular material. The z-offset was calibrated to the previously calculated average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material, which equals $(100 \pm 50)$ nm$^3$/ion.

investigated with ToF-SIMS in the negative polarity. An out of focus bright-field microscopy image (see Figure 4.24 B) indicates that the cells are elevated and therefore intact.

After mass calibration, a peak search and image reconstruction, the raw TOF-SIMS data is imported into Matlab for data processing and analysis. The image has a size of 256 $\times$ 256 $\times$ 135 pixels and the peak search extracted 367 mass peaks ($\approx$ 3.2 $\times$ $10^9$ data points). The data is normalised to the total number of ion counts per pixel particularly to account for variations in the secondary ion signal due to the topography of the cell sample as well as the decrease of the ion yield in the initial transient region and fluctuations in the secondary ion signals during depth profiling. The training set is formed by randomly selecting 1000 pixels per z-plane; the training set thus consists of 135.000 pixels (i.e. mass spectra) or 1.53% of the total amount of pixels (i.e. the same relative amount of pixels as for the positive polarity).

The first two principal components explain 75.04% of the variance. The positive loadings of PC1 (68.24% variance explained, see Figure 4.25 A) contain ions that can be attributed to bio-organic (and more specifically nuclear) material [40] such as $PO_2^-$ (m/z 63), $PO_3^-$ (m/z 79), $C_5H_4N_5^-$ (m/z 134, A), $HP_2O_6^-$ (m/z 159) and $NaP_2O_6^-$ (m/z 181). The negative loadings of PC1 correspond to inorganic mass peaks that can easily be attributed to the silica ($SiO_2$) and alumina ($Al_2O_3$) [116, 117] present in the borosilicate glass, see Table 4.3.

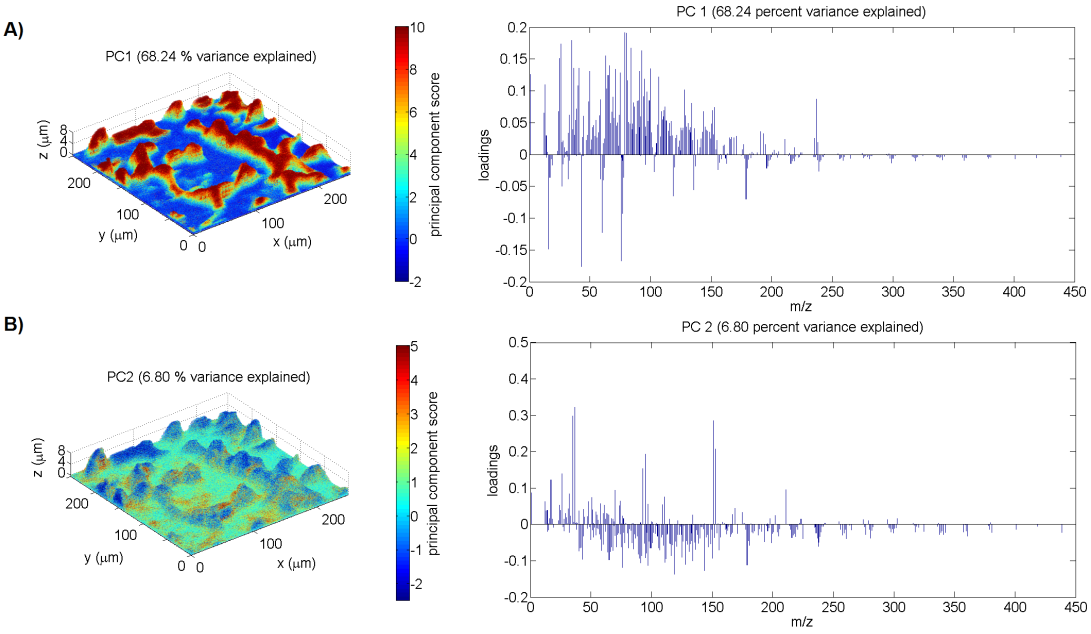



**Figure 4.25:** PCA of the freeze-dried RPE cells investigated with ToF-SIMS in the negative polarity; 3D scores images (left) and loadings plots (right). A) PC1 explains 68.24% of the variance. The positive loadings correspond to fragments associated with bio-organic material and the negative loadings correspond to fragments associated with the borosilicate glass. B) PC2 explains 6.80% of the variance. The positive loadings correspond to fragments associated with salts and the negative loadings correspond to fragments associated with bio-organic material.

**Table 4.3:** Negative ions indicative of the borosilicate glass substrate.

| Nominal Mass | Molecular Formula | Exact Mass |
|---|---|---|
| 16 | $O^-$ | 15.994915 |
| 17 | $OH^-$ | 17.002740 |
| 27 | $Al^-$ | 26.981541 |
| 32 | $O_2^-$ | 31.989830 |
| 43 | $AlO^-$ | 42.976456 |
| 59 | $AlO_2^-$ | 58.971371 |
| 60 | $SiO_2^-$ | 59.966758 |
| 61 | $SiO_2H^-$ | 60.974583 |
| 76 | $SiO_3^-$ | 75.961673 |
| 77 | $SiO_3H^-$ | 76.969498 |
| 103 | $Al_2O_3H^-$ | 102.955652 |
| 119 | $Al_2O_4H^-$ | 118.950567 |
| 136 | $Si_2O_5^-$ | 135.928431 |
| 137 | $Si_2O_5H^-$ | 136.936256 |
| 179 | $Al_3O_6H_2^-$ | 178.929763 |
| 196 | $Si_3O_7^-$ | 195.895189 |
| 197 | $Si_3O_7H^-$ | 196.903013 |
| 239 | $Al_4O_8H_3^-$ | 238.908959 |

Because PC1 differentiates between the borosilicate glass substrate and cellular material, its indication of the substrate interface (where the scores equal zero) can be utilised to apply the necessary z-offset correction to account for the surface topography of the cells. Note that this assumes a constant sputter yield through the cellular material. This computational transformation is then calibrated using the previously calculated average 20 keV $Ar^+_{5000}$ sputter yield for dry cellular material, which equals (100 $\pm$ 50) $nm^3$/ion (based on three separate 3D ToF-SIMS images of freeze-dried cell samples with corresponding interferometry measurements). This gives each voxel a height of 60 nm in the z-direction. Figure 4.24 C shows the heightmap based on the scores of principal component 1.

PC2 explains 6.80% of the variance. The positive loadings of PC2 correspond to inorganic mass peaks such as $Cl^-$ (m/z 35), $NaCl_2^-$ (m/z 93), $KCl_2^-$ (m/z 109), $Na_2Cl_3^-$ (m/z 151), $NaKCl_3^-$ (m/z 167) and their $^{37}Cl$ isotopes. The positive loadings PC2 are therefore a measure of the presence of salts. The negative loadings correspond to fragments associated with bio-organic material. The 3D scores image (see Figure 4.25 B) shows positive scores values surrounding the cellular material.

PC2 for both positive and negative polarity describes the distribution of cluster ions generated from salts, but interestingly, in the positive polarity PC2 (see Figure 4.22 B) is a measure of the presence of salts in the cell bodies, whereas in the negative polarity PC2 (see Figure 4.25 B) is a measure of the presence of salts surrounding the cellular material. An explanation might lie in the cluster ions observed in the loadings. In the positive polarity, the ions are clusters formed from $Na^+$ and $K^+$ cations and $OH^-$, $CN^-$, $Cl^-$, $PO_3{}^{3-}$ and $P_2O_6{}^{2-}$ anions (see Table 4.2). $CN^-$ and $PO_3{}^{3-}$ are known to originate from cellular and nuclear material respectively [123]. In the negative polarity, the ions are clusters formed from $Na^+$ and $K^+$ cations and $Cl^-$ anions and their $^{37}Cl$ isotopes.

### 4.4.3 Comparison to Neurons

Having performed PCA on the ARPE-19 positive and negative ToF-SIMS 3D image data, which revealed major components of the samples and allowed us to perform a z-offset correction, we now focus on determining whether there are differences in cell-related ion signals and their intensities between ARPE-19 and primary rat cortical neurons in a freeze-dried sample state.

It is essential that only the mass signals obtained from voxels corresponding to cellular material are compared in order to exclude contributions from the substrate to mass

signals and the total ion intensity. Selecting voxels corresponding to cellular material can be readily done by using the voxels' scores values. Any voxels for which PC1 $> 0$ (not substrate) are considered cellular material voxels for freeze-dried ion images. The mass signals from voxels corresponding to cellular material are summed to produce a mass spectrum specific to the cellular material in our samples. These (to the voxel's total ion count) normalized ion intensity summations, however, need to be divided by the total number of voxels to obtain the average normalized ion intensity signal per voxel before the mass spectra of different cells can be compared.

Finally, we are not very interested in inorganic mass peaks. Therefore, only organic mass peaks are selected for comparison. Figures 4.26 and 4.27 show the average normalized ion intensities of organic mass peaks between ARPE-19 (red) and primary rat cortical neurons (blue) cellular material in positive and negative 3D ToF-SIMS images in a freeze-dried sample state.

In order to get a quantitative measure of the similarity between the mass spectra of ARPE-19 and neuronal reference mass spectra, their cosine similarity is calculated. Cosine similarity (also called the spectral contrast angle) is a measure of similarity between two vectors, mass spectra in our case, and is often used as a mass spectral similarity measure [107]. The value of $\cos\theta$ ranges from 0 to 1 in positive space with 0 indicating orthogonality (decorrelation) and 1 meaning that the two spectra are exactly the same. If two mass spectra are presented as vectors of mass peaks, **A** and **B**, their cosine similarity, $\cos\theta$, is calculated by dividing their dot product by the product of their magnitudes:

$$\cos\theta = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\,\|\mathbf{B}\|} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{4.4.2}$$

$A_i$ and $B_i$ are the scalar components (mass peak intensities) of vectors (spectra) **A** and **B** with n different mass peaks.

The average $\cos\theta$ between the organic mass spectrum of RPE cells and three separate reference mass spectra of neuronal cells in the positive polarity equals $0.77 \pm 0.02$ (n = 3). The $\cos\theta$ between the three reference mass spectra of neuronal cells equals $0.89 \pm 0.07$ (n = 3), which is significantly different according to a two-tailed t-test (P = 0.0463). This indicates that the organic mass peaks and their intensities are quite similar for both freeze-dried RPE and freeze-dried neuronal cells although there are some differences which are not merely measurement variation.

Organic mass peaks which are unique for RPE cells include m/z 125, 88, 83, 82 and 39. One of these can be explained as an amino acid peak, $C_5H_7O^+$ (m/z 83, Val) [112], or a DNA peak, $C_3H_5N_3{}^+$ (m/z 83, G) [115]. Unfortunately, none of these ions can be attributed to melanin, a molecule expected to be present in RPE cells, in a conclusive manner. Most of the potential diagnostic peaks for melanin appear above m/z 200 such as $C_{25}H_{17}N_3O_8{}^+$ (m/z 487), $C_{24}H_{17}N_3O_6{}^+$ (m/z 443), $C_{18}H_{11}N_2O_9{}^+$ (m/z 399) and $C_{16}H_{11}N_2O_5{}^+$ (m/z 311) for eumelanin and $C_{13}H_{12}N_2O_{10}S^+$ (m/z 388) and $C_{11}H_{12}N_2O_5S^+$ (m/z 284) for pheomelanin [119].

Organic mass peaks which are unique for neuronal cells include m/z 147, 131, 119, 116, 98, 72, 71, 70, 69, 59, 57, 53 and 43. Some of these can be explained as amino acid peaks such as $C_9H_8O^+$ (m/z 131, Phe), $C_4H_{10}N^+$ (m/z 72, Val) and $C_4H_8N^+$ (m/z 70, Pro) [112], or phospholipid peaks such as $C_5H_8NO^+$ (m/z 98), $C_5H_{11}{}^+$ (m/z 71), $C_5H_9{}^+$ (m/z 69), $C_3H_9N^+$ (m/z 59), $C_4H_9{}^+$ (m/z 57), $C_4H_5{}^+$ (m/z 53) and $C_3H_7{}^+$ (m/z 43) [124]. A PDMS contamination could explain m/z 147 ($C_5H_{15}OSi_2{}^+$).

The $\cos\theta$ between the organic mass spectrum of RPE cells and a reference mass spectrum of neuronal cells in the negative polarity equals 0.96. This indicates that the organic mass peaks and their intensities are very similar for both freeze-dried RPE and freeze-dried neuronal cells. Some common organic peaks, i.e. observed in the mass spectra of both cell types, can be explained as DNA peaks such as $C_5H_4N_5O^-$ (m/z 150, G), $C_5H_4N_5{}^-$ (m/z 134, A), $C_4H_3N_2O_2{}^-$ (m/z 111, T) and $C_4N_3{}^-$ (m/z 90, A) [115].

There are also several organic mass peaks which are unique for the RPE cells' mass spectrum such as $C_5H_3N_5{}^-$ (m/z 133, G), $C_5H_5N_2O_2{}^-$ (m/z 125, T), $C_4H_4N_3O^-$ (m/z 110, C) and $C_4H_2N_3{}^-$ (m/z 92, C) [115], but this is likely due to measurement variation (better SNR for the RPE ToF-SIMS data) given the high $\cos\theta$ value. There are no unique ions present in the negative RPE spectrum that are indicative of melanin such as $C_9N^-$ (m/z 122), $C_5N^-$ (m/z 74), $C_3NO^-$ (m/z 66) and $C_3N^-$ (m/z 50) [125].

A single organic mass peak which is only found in the reference spectrum for neuronal cells is a fatty acid peak (16:0) at m/z 255 ($C_{16}H_{31}O_2{}^-$) [118] so this might be distinctive for neuronal cells.

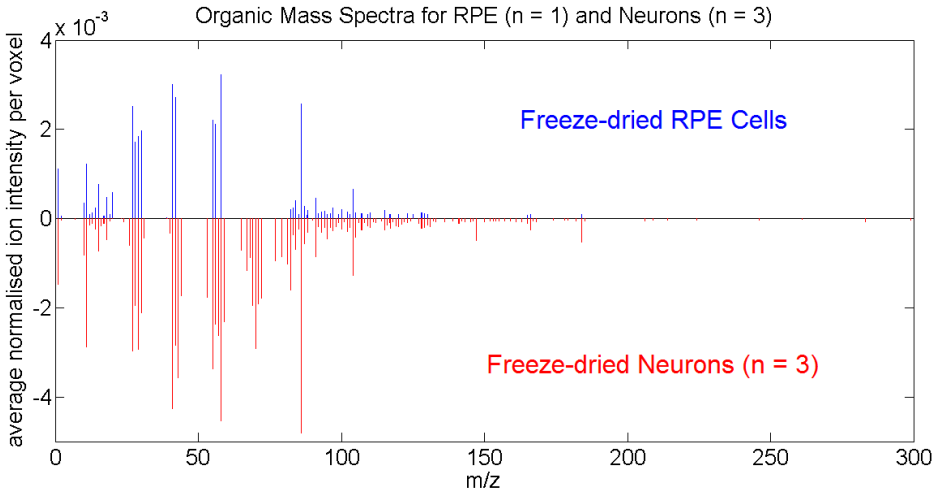



**Figure 4.26:** Comparison of the average normalized ion intensity per voxel of organic mass peaks between neuronal (red) and RPE (blue) cellular material in positive polarity 3D ToF-SIMS images of freeze-dried cells (10-15 cells per image).

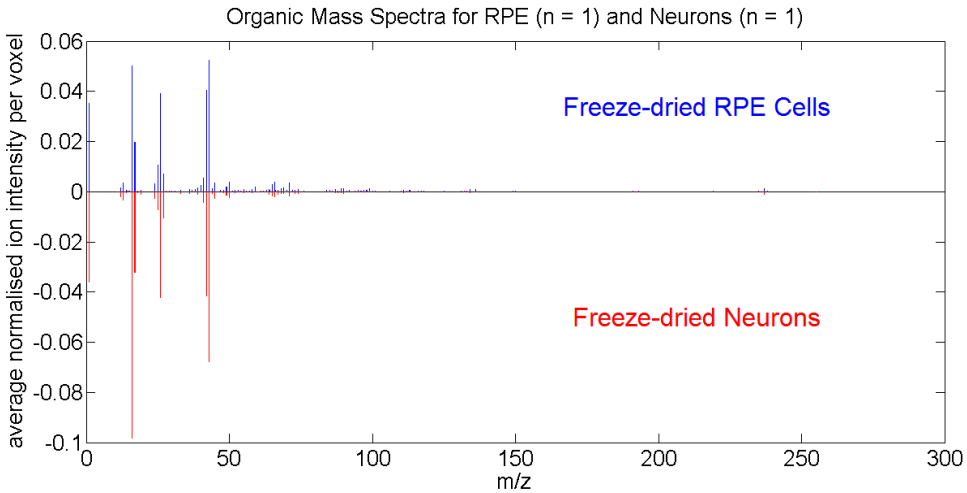



**Figure 4.27:** Comparison of the average normalized ion intensity per voxel of organic mass peaks between neuronal (red) and RPE (blue) cellular material in negative polarity 3D ToF-SIMS images of freeze-dried cells (10-15 cells per image).

## 4.5 Conclusions

In this chapter, we have developed a cell preparation protocol for 3D ToF-SIMS measurements. Primary rat cortical neurons were cultured on poly-L-lysine coated glass slides for 9 days in vitro. The sample preparation consists of cryofixation followed by freeze-drying (cf. reference [48]). Microscopy and interferometry measurements show that the morphology of the cells is more or less preserved. However, not all neurons are preserved and there are visual signs of damage. This raises questions about how representative of the native cellular state the samples are after freeze-drying and whether this also induces chemical changes such as chemical compound diffusion.

Having developed and validated an approach to PCA of 3D ToF-SIMS images using a well-defined test data set in chapter 3, the method was subsequently applied to 3D ToF-SIMS data obtained from freeze-dried neuronal cell samples to test its effectiveness on this more complex, biological sample. The results clearly show that PCA separates the different chemistries in its loadings and provides information on spatial chemical distribution via the scores. In addition, the PCA scores can be used to correct z-offsets due to the cells' topography. This in combination with calibration using corresponding interferometry images allows the average sputter yield in freeze-dried cellular material to be calculated accurately. Importantly, this approach now makes 3D SIMS image processing of biological samples with multivariate analysis accessible on a routine basis and considerably facilitates data analysis.

As an application, we attempted to visualise the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons. Even though propidium iodide and calcein to a lesser extent have distinct fragment ions in the high mass range, it was not possible to detect the fluorophores by 3D ToF-SIMS imaging of the cells, while their presence was confirmed by fluorescence microscopy. Breitenstein et al. [38] conducted a similar experiment using ethidium homodimer, DiI, fuchsine, calcein and LysoTracker Green, where they reached similar conclusions as well. One of the reasons cited is the fact that '3D ToF-SIMS data of molecular species $m/z > 200$ have shown up to now a comparatively poor signal-to-noise ratio which makes a scientifically proper interpretation challenging' [38].

In another application, we attempted to differentiate between two different cell types: primary rat cortical neurons and retinal pigment epithelium (RPE) cells. It is possible to detect differences in the kind of ions detected, but again only low mass fragment ions are observed. This makes the identification of individual compounds (such as melanin or cholesterol) difficult. Only if a (near) molecular ion or diagnostic fragment ($m/z$

$>$ 200) can be resolved, unambiguous identification is possible. Some caution is also required as to whether these differences in the kind of ions detected are due to inherent differences in cell chemistry or caused by the trypsinisation of the primary neuronal cells (see 2.2).

We have to conclude that the poor signal-to-noise ratio of molecular species with m/z $>$ 200 is a major bottleneck in the advancement of ToF-SIMS imaging as a diagnostic tool for the uptake of non-native compounds in cells or the differentiation of different cell types. Gilmore et al. [9] argues that increasing the ionisation yield is the current 'holy grail' in ToF-SIMS development. The ionisation efficiency in SIMS is low and it has been established that the majority of sputtered atoms and molecules are neutral [9]. Post-ionisation of the sputtered neutral molecules could therefore deliver huge increases in ion signal [126]. Fartmann et al. [127] have demonstrated the application of laser post-ionisation in the analysis of cultured cell lines. However, this approach has yet to be widely adopted, because of the added complexity of the experiments and the expense of the additional equipment required [9].

CHAPTER 5

# 3D ToF-SIMS Imaging of Frozen-Hydrated Neuronal Networks

Time-of-Flight Secondary Ion Mass Spectrometry (ToF-SIMS) has the capacity to become a routine tool for label-free 3D chemical imaging of biological samples such as cellular networks. To obtain meaningful results, it is paramount that sample preparation does not induce significant physical or chemical changes. In the previous chapter, we have seen that freeze-drying does not preserve all neurons and that there are visual signs of damage. This raises questions about how representative of the native cellular state the samples are after freeze-drying and whether this also induces chemical changes such as chemical compound diffusion. We have also observed a poor signal-to-noise ratio of molecular species with $m/z > 200$. This poses a major bottleneck in the advancement of ToF-SIMS imaging as a diagnostic tool for the uptake of non-native compounds in cells or the differentiation of different cell types.

As discussed in section 1.4, the literature tells us that cryofixation may be regarded as the most suitable fixation method currently available. The drying step, however, has shown to cause membrane damage as well as molecule rearrangement and a decrease in certain ion signals compared to ToF-SIMS measurements of samples in a frozen-hydrated state. However, these observations were made by surface analysis of either intact cells or fracture planes within cells. Furthermore, to our knowledge nearly all examples of freeze-fractured and frozen-hydrated cell imaging experiments in the literature are executed in 2D. Only Fletcher et al. [33] and Angerer et al. [128] presented 3D ToF-SIMS images ($256 \times 256 \times 10$ and $128 \times 128 \times 13$ voxels in size respectively) of a whole, intact frozen-hydrated cell. Angerer et al. also compared 2D ion images of air-

dried cells at room temperature and cells in a frozen-hydrated state after removing the outer membrane by etching with $C_{60}^{+}$, and found that the ions of the latter had a greater degree of localisation, which led them to infer that the drying process causes internal structures to collapse. However, they did not investigate the differences between the different sample preparations beyond the degree of localization of a few select ions.

A comprehensive comparison between 3D ToF-SIMS images of freeze-dried and frozen-hydrated cells is therefore still lacking. A few studies that included frozen-hydrated ToF-SIMS surface analysis of cells, when comparing various cell preparation techniques, also raised the temperature of the sample holder well above the recrystallization temperature in order to sublimate water from the cell surfaces prior to imaging [64, 129]. This step ultimately provides in-situ freeze-dried samples and raises the question to what extent the sample can still be considered frozen-hydrated and if recrystallization deformed the cell structure. Here, we present 3D ToF-SIMS image data of whole, intact cells in a frozen-hydrated state. The cells have been kept well below the recrystallization temperature prior and during ToF-SIMS analysis and the 3D ToF-SIMS image data shows the cells completely embedded in an ice matrix clearly demonstrating the cells are in a frozen-hydrated state. A comprehensive comparison between large 3D ToF-SIMS images of freeze-dried and frozen-hydrated cells is made using principal component analysis to facilitate the data analysis of these large data sets.

## 5.1 Cell Preparation and ToF-SIMS

In this chapter, primary rat cortical neurons are investigated in a frozen-hydrated state with 3D ToF-SIMS imaging. $Ar_{5000}^{+}$ gas clusters (GCIB) are used to etch through the vitreous ice matrix to obtain 3D ToF-SIMS data of intact cells, while the stage temperature is kept below the recrystallization temperature. This data is compared to 3D ToF-SIMS image data of freeze-dried cells, obtained at room temperature, in order to ascertain any quality differences in terms of cell preservation and chemical information. In order to make a valid comparison, the same primary ion beam configuration and other instrumental settings are used. Next, because of the sample's topography and the fluctuations in the secondary ion signals during depth pro-filing (such as the decrease of the ion yield in the surface transient region), the images are normalised to their total ion count per voxel. This normalisation also serves to minimise differences in the secondary ion signal between 3D ToF-SIMS images due to differences in sample topography, sample charging or instrumental conditions such as variations in primary ion current or detector efficiency, making them more comparable.

Primary rat cortical neurons were cultured on poly-L-lysine coated glass slides for 9 days in vitro. Details on the cell substrate preparation and cell culture protocol can be found in appendix 2.

For ToF-SIMS analysis of frozen-hydrated cells, cell samples were mounted one at a time onto an ION-TOF cryostage immersed in liquid nitrogen (cf. reference [97]). The sample stage was removed from the liquid nitrogen and placed onto the precooled sample transfer arm in the entry chamber under nitrogen flow. The entry chamber was immediately pumped down to $2 \times 10^{-5}$ mbar. When this pressure was reached, the entry door to the main chamber was opened to transfer the cryostage. Analysis was performed at -120 °C as measured by a thermocouple.

## 5.2 Data Analysis using PCA

There are three main sample components in 3D ToF-SIMS images of frozen-hydrated cells: bio-organic material, vitreous ice and glass substrate. 3D ToF-SIMS images of freeze-dried cells on the other hand only consist of two main sample components: bio-organic material and glass substrate. By performing principal component analysis (PCA) on our 3D ToF-SIMS data sets of freeze-dried and frozen-hydrated cells, we are able to distinguish the main components of our samples and assign one major component to each voxel. The image containing frozen-hydrated cells has a size of $256 \times 256 \times 205$ voxels and the peak search extracted 714 mass peaks ($\approx 9.6 \times 10^9$ data points). The training set is formed by randomly selecting 1000 voxels per z-plane; the training set thus consists of 205,000 voxels (i.e. mass spectra) or 1.5% of the total number of voxels. The correlation matrix of the training set is then decomposed to generate the loadings for standardised variables (mass peaks). These loadings are then applied to the full data set via block processing to generate 3D scores images.

The first two principal components elucidate the three different chemistries of the frozen-hydrated sample (75.5% variance explained). The positive loadings of PC1 (65.7% variance explained, see Figure 5.1 A) contain protonated water clusters $H(H_2O)_n^+$ (n = 5-11) that are indicative of ice [97] as well as organic ions indicative of cellular material, whereas the negative loadings contain inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27) and $Si^+$ (m/z 28). Because PC1 differentiates between the borosilicate glass substrate and ice/cellular material, its indication of the substrate interface (where the scores equal zero) can be utilised to apply the necessary z-offset correction to account for surface topography of the sample [100]. Note that this assumes a constant sputter yield through the sample matrix.

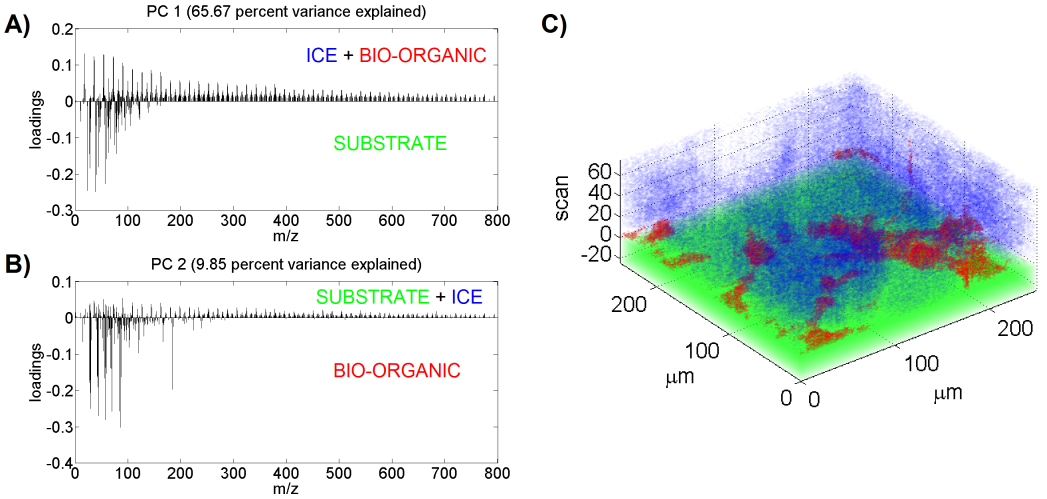



**Figure 5.1:** Elucidating the three main sample components in the 3D ToF-SIMS image of frozen-hydrated primary rat cortical neurons using principal component analysis (PCA). A) The loadings of PC1 (65.7% variance explained) differentiate between the borosilicate glass substrate (negative loadings) and ice/cellular material (positive loadings). B) The loadings of PC2 (9.8% variance explained) distinguish cellular material (negative loadings) from its surroundings be it ice or substrate (positive load-ings). C) Voxel classification in the z-corrected 3D ToF-SIMS image based on their scores values for PC1 and PC2. Voxels for which PC1 $< 0$ are considered substrate voxels (green). Voxels for which PC1 $> 0$ and PC2 $< 0$ are considered cellular material (organic) voxels (red). Voxels for which PC1 $> 0$ and PC2 $> 0$ are considered ice voxels (blue). N.B. Because neither the thickness nor the topography of the ice layer could be determined, it is only possible to perform the z-offset correction (without z-calibration and calculation of the sputter yield).

The negative loadings of PC2 (9.8% variance explained, see Figure 5.1 B) contain a strong correlation with common markers for phosphocholine-containing phospholipids such as the $C_5H_{15}NPO_4^+$ ion at m/z 184 [124] and peaks commonly associated with amino acids such as the $C_8H_{10}NO^+$ ion (Tyr) at m/z 136, the $C_9H_8N^+$ ion (Trp) at m/z 130 and the $C_8H_{10}N^+$ ion (Phe) at m/z 120 [112]. The positive loadings of PC2 contain both protonated water clusters and inorganic ions specific for the borosilicate glass substrate, but with low coefficients indicating limited correlation. PC2 therefore distinguishes cellular material from its surroundings be it ice or substrate.

Using these PCA assignments, each voxel in the 3D ToF-SIMS image can be classified based on its scores values for PC1 and PC2 (see Figure 5.1 C). Voxels for which PC1 $< 0$ are considered substrate voxels, voxels for which PC1 $> 0$ and PC2 $< 0$ are considered cellular material voxels and voxels for which PC1 $> 0$ and PC2 $> 0$ are considered ice voxels.

Similarly, we can perform PCA on the freeze-dried cells' 3D ToF-SIMS data to distinguish the main components of our samples and identify to which component each voxel belongs (see Figure 5.2). In this case, PC1 differentiates between the borosilicate glass substrate and organic material. In the case of freeze-dried cells, the z-offset correction to account for the surface topography can also be calibrated against interferometry data. Based on three separate 3D ToF-SIMS images of freeze-dried cell samples with corresponding interferometry measurements, the average 20 keV $Ar_{5000}^+$ sputter yield for dry cellular material equals (100 $\pm$ 50) $nm^3$/ion (n = 3). The calculation of the average 20 keV $Ar_{5000}^+$ sputter yield for dry cellular material makes it possible to calibrate z-offset corrections without complementary interferometry data. Although it is of course recommended to obtain complementary interferometry images for the most accurate result given the average sputter yield's large coefficient of variation.

Because it was not possible to conduct interferometry measurements on frozen samples, it was not possible to determine their sputter yield directly. However, based on the average cell height determined by interferometry on dried samples (6.7 $\pm$ 0.7 $\mu$m, see 4.2.1), we estimate the sputter yield in frozen samples to be six-fold that of dry samples.

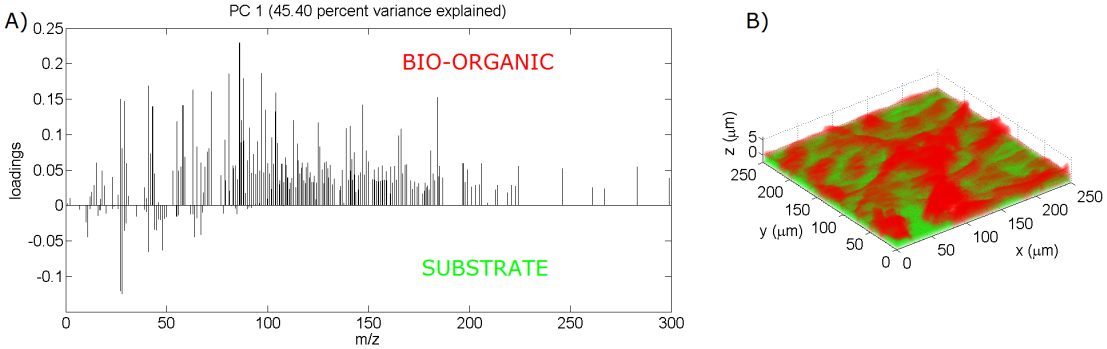



**Figure 5.2:** Elucidating the two main sample components in the 3D ToF-SIMS image of freeze-dried primary rat cortical neurons using principal component analysis (PCA). A) The loadings of PC1 (45.4% variance explained) differentiate between the borosilicate glass substrate (negative loadings) and cellular material (positive loadings). B) Voxel classification in the z-corrected 3D ToF-SIMS image based on their scores values for PC1. Voxels for which PC1 $< 0$ are considered substrate voxels (green). Voxels for which PC1 $> 0$ are considered cellular material (organic) voxels (red). The z-corrected 3D ToF-SIMS image is calibrated using the average 20 keV $Ar_{5000}^{+}$ sputter yield determined for freeze-dried cellular material.

## 5.3 Improved Cell Membrane Preservation

Preservation of cell integrity is a critical aspect of any biological sample preparation method. Complementary imaging techniques such as optical or electron microscopy could provide this information before ToF-SIMS analysis for freeze-dried samples, but for frozen-hydrated samples this is not readily achievable as the samples need to be kept below the recrystallization temperature and the cells are embedded in an ice matrix. Instead, antilocalization of $Na^+$ and $K^+$ in ToF-SIMS images is used as an indicator of instantaneous fixation of the cell in the living state [48] and membrane preservation [58, 130]. Due to active transport of sodium and potassium ions by the $Na^+/K^+$-ATPase enzyme found in the plasma membrane of all animal cells, mammalian cells contain a relatively high concentration of potassium ions but low concentrations of sodium ions whereas the extracellular environment has significantly higher levels of sodium than potassium ions.

As the freeze-dried cells are not embedded in a matrix like the frozen-hydrated cells and the intracellular $Na^+$ and $K^+$ intensities are to be compared to their extracellular intensities, slice planes in the z-axis close to the substrate interface that contain both cells and their surrounding extracellular matrix are chosen for this comparison. A scores image of PC1 at a slightly higher slice plane (see Figure 5.3 A) is used to determine intracellular (PC1 $> 0$) regions. The freeze-dried cell images show similar intracellular $K^+$ and $Na^+$ intensities (see Figures 5.3 B and C). The average normalised $K^+$ and $Na^+$ ion intensities per pixel inside the cell regions equal 0.087 $\pm$ 0.018 (n = 7606 pixels) and 0.088 $\pm$ 0.017 (n = 7606 pixels) respectively. In regions corresponding to the extracellular matrix the $Na^+$ intensities are higher than within the cells, whereas the $K^+$ intensities are slightly lower. The average normalised $K^+$ and $Na^+$ ion intensities per pixel outside the cell regions equal 0.080 $\pm$ 0.048 (n = 57371 pixels) and 0.113 $\pm$ 0.047 (n = 57371 pixels) respectively. Two-tailed t-tests show that all ion intensities are significantly different from one another with 95% confidence. The overlap of the $K^+$ signal with the cell regions is nonetheless limited. To obtain a quantitative measure of the overlap between $K^+$ and cell regions, the PC1 and $K^+$ images are converted to binary images and their peak signal-to-noise ratio (PSNR) is calculated (see Figures 5.4 A and B). The PSNR between the binary images equals only 3.30 dB. It is also shown that the freeze-dried cells were viable before the freeze-drying step. The cells had been stained with calcein-AM (acetoxymethyl ester of calcein) and propidium iodide (PI) prior to their cryofixation (see Figure 4.16). Calcein-AM only stains viable cells so its fluorescence microscopy image after freeze-drying indicates that there was no membrane leakage prior to freeze-drying. (N.B. Only the freeze-dried cells presented in this

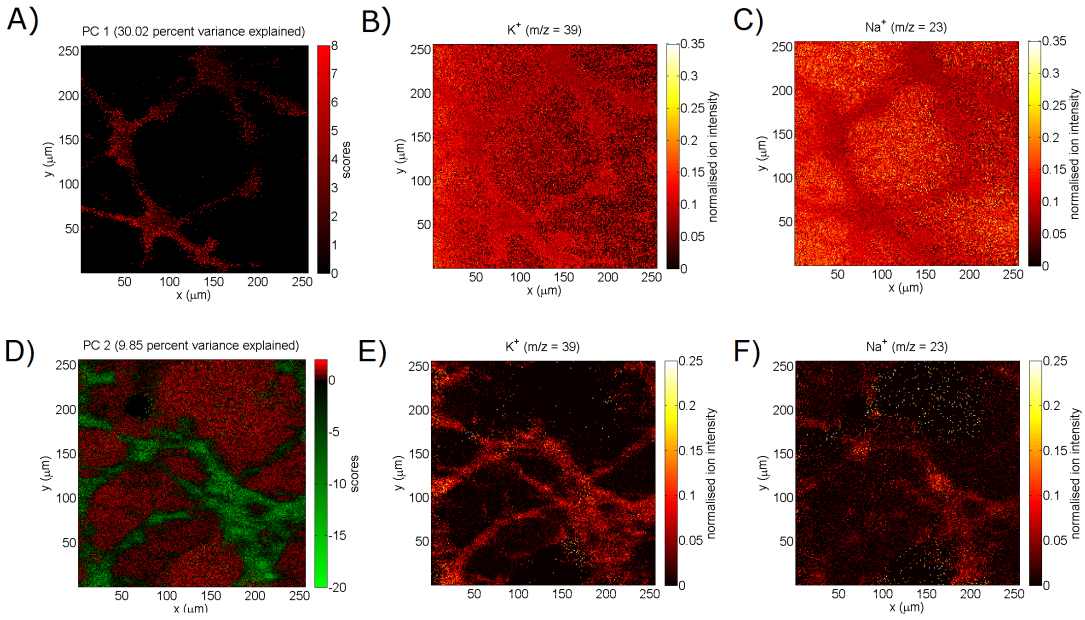



**Figure 5.3:** $Na^+/K^+$ antilocalisation in z-corrected ToF-SIMS image slices of freeze-dried (top row) and frozen-hydrated (bottom row) neuronal cells. A) Scores image for PC1 of a freeze-dried cell sample at z = 189 nm (10 scans) from the substrate interface. The red regions (positive scores) correspond to cellular signals. No extracellular regions are visible due to the short distance from the substrate interface. B) Normalised ion image for $K^+$ (m/z = 39) of the freeze-dried cell sample at z = 21 nm (2 scans) from the substrate interface. The average normalised $K^+$ intensities per pixel inside and outside (i.e. the extracellular matrix) the cell regions equal 0.087 $\pm$ 0.018 (n = 7606 pixels) and 0.080 $\pm$ 0.048 (n = 57371 pixels) respectively. C) Normalised ion image for $Na^+$ (m/z = 23) of the freeze-dried cell sample at z = 21 nm (2 scans) from the substrate interface. The average normalised $Na^+$ intensities per pixel inside and outside (i.e. the extracellular matrix) the cell regions equal 0.088 $\pm$ 0.017 (n = 7606 pixels) and 0.113 $\pm$ 0.047 (n = 57371 pixels) respectively. D) Scores image for PC2 of a frozen-hydrated cell sample at z = 5 scans from the substrate interface. The green regions (negative scores) correspond to cellular signals and the red regions (positive scores) correspond to signals originating from the ice matrix. E) Normalised ion image for $K^+$ (m/z = 39) of the frozen-hydrated cell sample at z = 5 scans from the substrate interface. The average normalised $K^+$ intensities inside and outside the cell regions equal 0.023 $\pm$ 0.032 (n = 33923 pixels) and 0.005 $\pm$ 0.018 (n = 29608 pixels) respectively. F) Normalised ion image for $Na^+$ (m/z = 23) of the frozen-hydrated cell sample at z = 5 scans from the substrate interface. The average normalised $Na^+$ intensities inside and outside the cell regions equal 0.015 $\pm$ 0.026 (n = 33923 pixels) and 0.007 $\pm$ 0.026 (n = 29608 pixels) respectively.

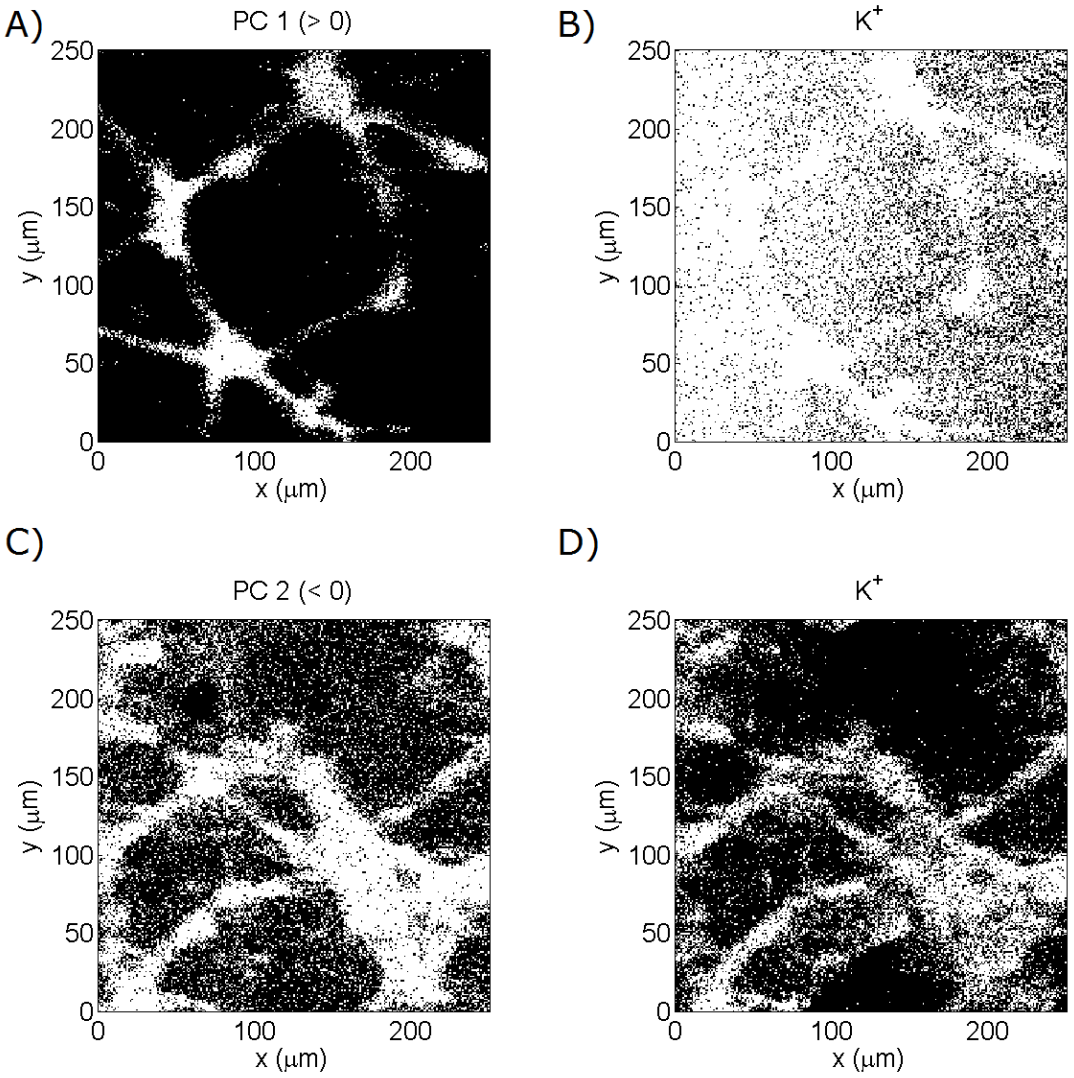



**Figure 5.4:** Overlap of the $K^+$ distribution with cellular regions in z-corrected ToF-SIMS image slices of freeze-dried (top row) and frozen-hydrated (bottom row) neuronal cells. A) Binary scores image for PC1 $> 0$ of the freeze-dried cell sample at z = 189 nm (10 scans) from the substrate interface. B) Binary normalised ion image for $K^+$ (m/z = 39) of the freeze-dried cell sample at z = 21 nm (2 scans) from the substrate interface. The peak signal-to-noise ratio (PSNR) between the $K^+$ image and the PC1 image equals 3.30 dB. C) Binary scores image for PC2 $< 0$ of the frozen-hydrated cell sample at z = 5 scans from the substrate interface. D) Binary normalised ion image for $K^+$ (m/z = 39) of the frozen-hydrated cell sample at z = 5 scans from the substrate interface. The PSNR between the $K^+$ image and the PC1 image equals 12.64 dB.

section were stained. Other cell samples used for the comparison between freeze-dried and frozen-hydrated cells were not.) These observations indicate that the dehydration step induces some amount of membrane damage and/or migration of diffusible chemical species.

A slice plane in the z-axis close to the substrate interface is also chosen for the frozen-hydrated cells. Because it is not readily achievable to perform any complementary imaging on frozen-hydrated samples, the scores image of PC2 (see Figure 5.3 D) is used to determine intracellular (PC2 $< 0$) and extracellular (PC2 $> 0$) regions. The frozen-hydrated ToF-SIMS images of the neuronal networks show cells embedded in a vitreous ice matrix (see Figure 5.3 D). The $K^+$ intensities observed inside the cell regions are higher than those of $Na^+$ (see Figures 5.3 E and F). The average normalised $K^+$ and $Na^+$ ion intensities per pixel inside the cell regions equal $0.023 \pm 0.032$ (n = 33923 pixels) and $0.015 \pm 0.026$ (n = 33923 pixels) respectively. The $K^+$ ion image of the frozen-hydrated cells also overlaps better with the cell regions than that of the freeze-dried cells (see Figures 5.4 C and D). The PSNR between the binary PC2 and $K^+$ images equals 12.64 dB in this case. The vitreous ice regions contain lower amounts of both $Na^+$ and $K^+$ compared to cells, but the $Na^+$ intensity is higher in these regions than that of $K^+$. The average normalised $K^+$ and $Na^+$ ion intensities per pixel outside the cell regions equal $0.005 \pm 0.018$ (n = 29608 pixels) and $0.007 \pm 0.026$ (n = 29608 pixels) respectively. Two-tailed t-tests again show that all ion intensities are significantly different from one another with 95% confidence. The observations confirm rapid fixation and a good degree of membrane preservation as well as successful removal of salts during the ammonium formate washing step for the frozen hydrated sample preparation. The greater degree of antilocalization of $Na^+$ and $K^+$ in the frozen-hydrated cell sample's 3D ToF-SIMS image data means that the frozen-hydrated cell samples reflect the native cell state better than the freeze-dried ones, which is in agreement with the literature [58, 61, 128].

## 5.4 Increased Intensities for High-Mass Fragments

Having established which method preserves the cell structure best physically, we then focused on determining whether there are differences in cell-related ion signals and their intensities between the frozen-hydrated and the freeze-dried samples. It is essential that only the mass signals obtained from voxels corresponding to cellular material are compared in order to exclude contributions from the substrate or ice matrix to mass signals and the total ion intensity. Selecting voxels corresponding to cellular material

can be readily done by using the voxels' scores values (see Figure 5.1 C). Any voxels for which PC1 $>$ 0 (not sub-strate) and PC2 $<$ 0 (not ice) are considered cellular material voxels for frozen-hydrated ion images. Any voxels for which PC1 $>$ 0 (not substrate) are considered cellular material voxels for freeze-dried ion images (see Figure 5.2 B). The mass signals from voxels corresponding to cellular material are summed to produce a mass spectrum specific to the cellular material in our samples. These (to the voxel's total ion count) normalized ion intensity summations, however, need to be divided by the total number of voxels to obtain the average normalized ion intensity signal per voxel before the mass spectra of freeze-dried and frozen-hydrated cellular material can be compared. Finally, the cellular material of frozen-hydrated cells still contains vitrified ice and will produce a mass spectrum containing ion signals for both organic material and ice. The cell preparation method can also influence the organic ion signals produced. Therefore, a list of known mass peaks for cellular material (lipids and amino acids) from literature [112, 124, 131] was used to select relevant mass peaks for the comparison (see Table C.1).

Figure 5.5 shows the average normalized ion intensities of identified lipid and amino acid mass peaks between freeze-dried (red) and frozen-hydrated (blue) cellular material in 3D ToF-SIMS images of neurons. In order to get a quantitative measure of the similarity between the (known fragments) mass spectra of frozen-hydrated and freeze-dried cellular material their cosine similarity is calculated. Cosine similarity ($\cos\theta$) is a measure of similarity between two vectors, mass spectra in our case, and is often used as a mass spectral similarity measure [107]. The value of $\cos\theta$ ranges from 0 to 1 in positive space with 0 indicating orthogonality (decorrelation) and 1 meaning that the two spectra are exactly the same. Comparing the mass spectrum of frozen-hydrated cellular material separately to three mass spectra of freeze-dried cellular material, the average $\cos\theta$ equals $0.30 \pm 0.02$ (n = 3) indicating intermediate dissimilarity. The average $\cos\theta$ between the three mass spectra of freeze-dried cellular material equals $0.95 \pm 0.03$ (n = 3) indicating a high degree of similarity. The cell preparation method clearly has an influence on the mass peaks obtained.

Many mass peaks of the frozen-hydrated cellular material are higher in intensity than those of the freeze dried samples, especially those at higher masses. The $C_7H_7^+$ ion at m/z 91 shows a fifteen-fold increase and the $C_5H_{15}NPO_4^+$ ion at m/z 184 shows a four-fold increase, for example. A two-sided t-test shows that the average of the identified mass peak intensities of the frozen-hydrated cellular material is significantly larger (P = 0.02) in the higher mass range (m/z $>$ 100) than that of those of the freeze-dried cellular material. However, when limited to the low mass range up to m/z 100, no

significant difference (P = 0.32) is observed between the average identified mass peak intensities of the frozen-hydrated cellular material and those of the freeze-dried cellular material. Notably, several fragments at higher mass attributed to phosphatidylcholine such as $C_7H_{17}NPO_5^+$ (m/z 226), $C_6H_{17}NPO_4^+$ (m/z 198) and $C_5H_{13}NPO_4^+$ (m/z 182) are only observed in the mass spectrum of the frozen-hydrated cellular material [124]. Equally, several low-mass fragments attributed to phosphatidylcholine such as $C_5H_{14}NO^+$ (m/z 104), $C_5H_{12}NO^+$ (m/z 102), $C_7H_{13}^+$ (m/z 97), $C_7H_{11}^+$ (m/z 95), $C_5H_{14}N^+$ (m/z 88), $C_6H_{13}^+$ (m/z 85) and $C_6H_{11}^+$ (m/z 83) are only observed in the mass spectrum of the freeze-dried cellular material [124]. Mass fragments attributed to amino acids are found in the mass spectra of both freeze-dried and frozen-hydrated cellular material, but there are certain differences. $C_3H_8NO^+$ (m/z 74, Thr), $C_3H_6NO^+$ (m/z 72, Gly), $C_2H_5S^+$ (m/z 61, Met) and $C_2H_6NO^+$ (m/z 60, Ser) for example are only observed in the frozen-hydrated spectrum and $C_4H_7N_2O_2^+$ (m/z 115, Gly), $C_5H_8N_3^+$ (m/z 110, Arg/His), $C_4H_{10}N_3^+$ (m/z 100, Arg), $CH_5N_3^+$ (m/z 59, Arg) and $C_3H_3O^+$ (m/z 55, Tyr) are only observed in the freeze-dried spectrum [112, 131]. There are several more ions in the low mass range that are only found in one of the two spectra, but these can be attributed to both phosphatidylcholine and amino acids or proteins in general so no clear conclusions can be drawn from them.

Fragmentation products and their ratios are the result of competing and consecutive unimolecular reactions, whose reaction rates are dependent on the internal energy of the precursor ion [132]. The increase in higher-mass fragments for phosphatidylcholine indicates that the internal energy ($E_{int}$) of the phosphatidylcholine precursor ions is lower in the frozen-hydrated state. Differences in internal energy would also explain the observed differences in fragments attributed to amino acids. The internal energy content of molecular ions derives from the thermal energy of the molecule before ionization and the amount of energy deposited in the ionization process [132]. On the one hand, the low temperature of the liquid nitrogen cooled frozen-hydrated cell measurements in ToF-SIMS could affect the internal energy of an ion via a decrease of its thermal energy before ionization. On the other hand, the energy deposition of the $Bi_3^+$ projectiles is likely influenced by the ice matrix.

To assess the effect of the sample's temperature a simple experiment comparing ToF-SIMS mass spectra of freeze-dried cells at room temperature with those of freeze-dried cells cooled to the same temperature as the frozen-hydrated cells was conducted. Two mass spectra of freeze-dried cellular material at T $<$ -120 °C are compared with the three mass spectra of freeze-dried cellular material at room temperature by calculating the cosine similarity between them. The average $\cos\theta$ equals $0.96 \pm 0.03$ (n = 6) indicat-

ing a high degree of similarity. In other words, it would seem the temperature has little effect on the mass fragments obtained suggesting that not the lower temperature but the ice matrix is responsible for the increase in higher-mass fragments in the frozen-hydrated state. We hypothesize that the ice matrix as a source of protons increases the likelihood of 'soft' chemical ionization over 'hard' electron ionization of neutrals in the relatively high-pressure region close to the surface (selvedge) and that its relative volatility increases the particle density in the selvedge increasing the likelihood of collision induced energy-exchange [18, 133].

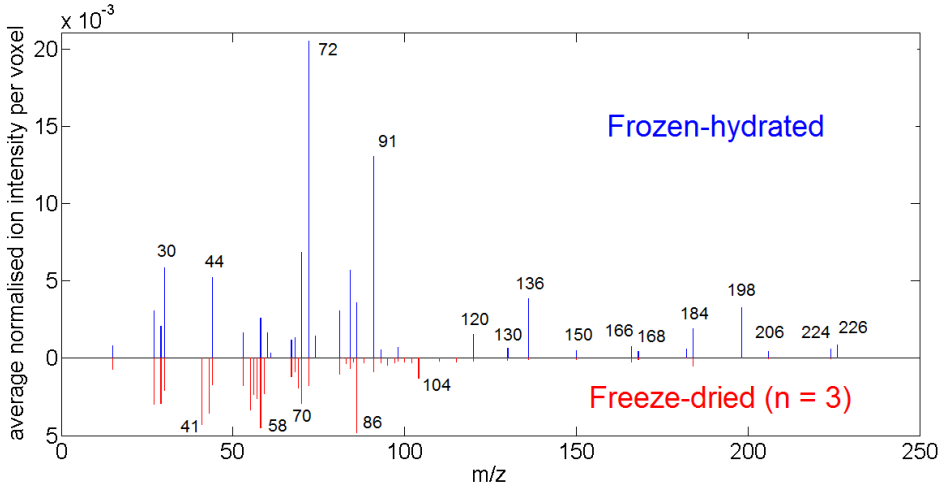



**Figure 5.5:** Comparison of the average normalized ion intensity per voxel of selected mass peaks between freeze-dried (red) and frozen-hydrated (blue) cellular material in 3D ToF-SIMS images of neurons. Only mass peaks identified as lipid and amino acid mass peaks are shown (see Table C.1 for assignments). The frozen-hydrated cell material has higher average normalized ion intensities for identified lipid and amino acid mass peaks in the higher mass range ($m/z > 100$).

## 5.5 Conclusions

$Ar_{5000}^{+}$ gas clusters (GCIB) were used to etch through the vitreous ice matrix to obtain large 3D ToF-SIMS image data ($256 \times 256 \times 205$ voxels) of whole, intact primary rat cortical neurons (a small section of a neuronal network) in a frozen-hydrated state. The cells have been kept well below the recrystallization temperature prior and during ToF-SIMS analysis and the 3D ToF-SIMS image shows the cells completely embedded in an ice matrix. This enabled, for the first time, a comprehensive comparison with 3D ToF-SIMS images of freeze-dried cells. Previous ToF-SIMS imaging studies of cells generally employed a freeze-fracture methodology or sublimated the surface ice to expose the cellular material for 2D analysis.

The dehydration step for the freeze-dried samples induces some amount of membrane damage and/or migration of diffusible chemical species based on a comparison of the spatial distributions of $Na^{+}$ and $K^{+}$ in the ToF-SIMS images. This effect was significantly reduced in frozen-hydrated samples indicating that these samples reflect the native cell state better than the freeze-dried ones. Furthermore, significant differences between the mass spectra of frozen-hydrated and freeze-dried cellular material were found by calculating their cosine similarity. Notably, an increased ion intensity for higher-mass fragments ($m/z > 100$) is observed in the mass spectrum of frozen-hydrated cellular material. This increased ion intensity for higher mass fragments in the mass spectrum of frozen-hydrated cellular material is an additional advantage of studying biological samples in a frozen-hydrated state. Many chemical species present in cells will generate similar low mass fragments and differ in their higher mass fragments so unambiguous chemical identification often relies on the presence of the high mass fragments in the spectrum. It is likely that the ice matrix is responsible for the increase in higher-mass fragments in the frozen-hydrated state rather than the low temperature used for the cryogenic analysis conditions.

In spite of the quality improvements in terms of cell preservation and chemical information, there are draw-backs to this approach. The increased technical difficulty, the inability to assess a sample location beforehand using complementary imaging techniques and the increased measurement time severely limit sample throughput. If there is need for a high number of measurements and the application does not require the cell to be studied in its most native state and higher-mass fragments are not of interest, one may consider freeze-dried cell preparation for convenience. The best results in terms of cell preservation and secondary ion yields, however, are obtained with cell samples in a frozen-hydrated state.

CHAPTER 6

# Conclusions

In this thesis we set out to to assess PCA as a data analysis tool for 3D TOF-SIMS images and to establish cell preparation protocols for TOF-SIMS that preserves the cell morphology.

In chapter 3, we first presented a method to construct ToF-SIMS image data consisting of multiple chemistries with spatial distributions of their own to identify the effects of different preprocessing procedures. We were also able to expand the application of principal component analysis (PCA) to large 3D ToF-SIMS images without requiring any computing resources beyond a desk top computer. The method reported here presents the first time PCA has been performed on a large scale (128 x 128 x 622 pixels) 3D ToF-SIMS image. This was made possible by first calculating the PCA loadings using a smaller subset of randomly selected pixels as a training set that could then be applied to the full data set to generate the scores images. The method has been validated using a well-defined 3D ToF-SIMS data set of a PS-PVP multilayer system. The results clearly show that PCA separates the different chemistries in its loadings and provides information on spatial chemical distribution via the scores. Furthermore, the scores images have a 1.7-2.4 times better signal-to-noise ratio than can be obtained with single ions. The depth resolution of the scores images does not differ from that of the single ion images.

Having developed and validated an approach to PCA of 3D ToF-SIMS images using a well-defined test data set in chapter 3, the method was subsequently applied to 3D ToF-SIMS data obtained from freeze-dried neuronal cell samples to test its effectiveness on this more complex, biological sample in chapter 4. The results clearly show that PCA separates the different chemistries in its loadings and provides information on spatial chemical distribution via the scores. The method facilitates differentiation between cellular components such as lipids and amino acids, which allowed the cell membrane,

the cytoplasm and the extracellular matrix (ECM) to be easily distinguished from one another. In addition, the PCA scores can be used to correct z-offsets due to the cells' topography. This in combination with calibration using corresponding interferometry images allows the average sputter yield in freeze-dried cellular material to be calculated accurately. Importantly, this approach now makes 3D SIMS image processing of biological samples with multivariate analysis accessible on a routine basis and considerably facilitates data analysis.

As an application, we attempted to visualise the cellular uptake of non-native compounds, namely fluorescent dyes, in primary rat cortical neurons. Even though propidium iodide and calcein to a lesser extent have distinct fragment ions in the high mass range, it was not possible to detect the fluorophores by 3D ToF-SIMS imaging of the cells, while their presence was confirmed by fluorescence microscopy. We have to conclude that the poor signal-to-noise ratio of molecular species with $m/z > 200$ is a major bottleneck in the advancement of ToF-SIMS imaging as a diagnostic tool for the uptake of non-native compounds in cells.

In another application, we attempted to differentiate between two different cell types: primary rat cortical neurons and retinal pigment epithelium (RPE) cells. It is possible to detect differences in the kind of ions detected, but again only low mass fragment ions are observed. This makes the identification of individual compounds (such as melanin or cholesterol) difficult. Only if a (near) molecular ion ($m/z > 200$) can be resolved, unambiguous identification is possible.

In chapter 4, microscopy and interferometry measurements show that the morphology of freeze-dried cells is more or less preserved. However, not all neurons are preserved and there are visual signs of damage. This raised questions about how representative of the native cellular state the samples are after freeze-drying and whether this also induces chemical changes such as chemical compound diffusion. In chapter 5, we presented a comprehensive comparison with 3D ToF-SIMS images of freeze-dried cells. Previous ToF-SIMS imaging studies of cells generally employed a freeze-fracture methodology or sublimated the surface ice to expose the cellular material for 2D analysis. The dehydration step for the freeze-dried samples was shown to induce some amount of membrane damage and/or migration of diffusible chemical species based on a comparison of the spatial distributions of $Na^+$ and $K^+$ in the ToF-SIMS images. This effect was significantly reduced in frozen-hydrated samples indicating that these samples reflect the native cell state better than the freeze-dried ones.

Furthermore, significant differences between the mass spectra of frozen-hydrated and freeze-dried cellular material were found by calculating their cosine similarity. No-

tably, an increased ion intensity for higher-mass fragments ($m/z > 100$) is observed in the mass spectrum of frozen-hydrated cellular material. This increased ion intensity for higher mass fragments in the mass spectrum of frozen-hydrated cellular material is an additional advantage of studying biological samples in a frozen-hydrated state. Many chemical species present in cells will generate similar low mass fragments and differ in their higher mass fragments so unambiguous chemical identification often relies on the presence of the high mass fragments in the spectrum. It is likely that the ice matrix is responsible for the increase in higher-mass fragments in the frozen-hydrated state rather than the low temperature used for the cryogenic analysis conditions.

## 6.1 Future Work and Outlook

While we may not have found melanin related mass fragments in the mass spectra of retinal pigment epithelium (RPE) cells, it was possible to detect quite a few distinct differences in the kind of ions detected for primary rat cortical neurons and RPE cells albeit in the low mass range. We need to determine if it is possible use these mass peaks to identify cell types in unknown samples (e.g. a co-culture of neuronal and RPE cells) by 'fingerprinting'.

Even though propidium iodide and calcein to a lesser extent have distinct fragment ions in the high mass range, it was not possible to detect the fluorophores by 3D ToF-SIMS imaging of freeze-dried cells, while their presence was confirmed by fluorescence microscopy. Like Breitenstein et al. [38], we conclude that this is due to the poor signal-to-noise ratio of molecular species with $m/z > 200$. Gilmore et al. [9] argues that increasing the ionisation yield is the current 'holy grail' in ToF-SIMS development. In chapter 5, we showed that the mass spectrum of frozen-hydrated cellular material has increased ion intensities for higher-mass fragments compared to freeze-dried cellular material. We thus need to follow up on this and see if it is possible to detect the fluorophores by 3D ToF-SIMS imaging of frozen-hydrated cells.

Similarly, while differences are observed in the kind of ions detected for primary rat cortical neurons and retinal pigment epithelium (RPE) cells, only low mass fragment ions are observed, which makes the identification of individual compounds difficult. If frozen-hydrated cell measurements resolve (near) molecular ions ($m/z > 200$), it will be possible to unambiguously identify specific compounds such as melanin or cholesterol.

In spite of the quality improvements in terms of cell preservation and chemical information, there are draw-backs to measuring cells in a frozen-hydrated state. The

increased technical difficulty, the inability to assess a sample location beforehand using complementary imaging techniques and the increased measurement time severely limit sample throughput. This needs to be addressed, because biological samples (due to inherent biological variability) require high sample numbers before the results are considered reliable.

Alternatively, matrix-enhanced SIMS (ME-SIMS) [134, 135] can be used to enhance molecular yields. However, these all in all incremental ion yield enhancements may not be sufficient and then new technological approaches will be need to explored. For example, developments in laser post ionisation of the neutral portion of the secondary species ($>$ 99% of ejected material) have the potential to deliver huge increases in signal if the problem of photon induced fragmentation can be overcome [9].

# References


[1] Louise E Smith, Rod Smallwood, and Sheila Macneil. A comparison of imaging methodologies for 3d tissue engineering. Microscopy research and technique, 73(12):1123–1133, 2010.

[2] Sarwat B Rizvi, Shirin Ghaderi, Mo Keshtgar, and Alexander M Seifalian. Semiconductor quantum dots as fluorescent probes for in vitro and in vivo biomolecular and cellular imaging. Nano Reviews & Experiments, 1, 2010.

[3] Anna Kremer, Saskia Lippens, Sona Bartunkova, Bob Asselbergh, C Blanpain, Matyas Fendrych, Alain Goossens, Matthew Holt, Sophie Janssens, Michiel Krols, et al. Developing 3d sem in a broad biological context. Journal of microscopy, 259(2):80–96, 2015.

[4] Marcel Menezes Lyra da Cunha, Sylvain Trepout, Cédric Messaoudi, Ting-Di Wu, Richard Ortega, Jean-Luc Guerquin-Kern, and Sergio Marco. Overview of chemical imaging methods to address biological questions. Micron, 84:23–36, 2016.

[5] Katarzyna Majzner, Agnieszka Kaczor, Neli Kachamakova-Trojanowska, Andrzej Fedorowicz, Stefan Chlopicki, and Malgorzata Baranska. 3d confocal raman imaging of endothelial cells and vascular wall: perspectives in analytical spectroscopy of biomedical research. Analyst, 138(2):603–610, 2013.

[6] Rachael Smith, Karen L Wright, and Lorna Ashton. Raman spectroscopy: an evolving technique for live cell studies. Analyst, 2016.

[7] Flavius C Pascut, Huey T Goh, Nathan Welch, Lee D Buttery, Chris Denning, and Ioan Notingher. Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells. Biophysical journal, 100(1):251–259, 2011.

[8] Alain Brunelle, David Touboul, and Olivier Laprévote. Biological tissue imag-

ing with time-of-flight secondary ion mass spectrometry and cluster ion sources. Journal of Mass Spectrometry, 40(8):985–999, 2005.

[9] Ian S Gilmore. Sims of organics: Advances in 2d and 3d imaging and future outlook. Journal of Vacuum Science & Technology A, 31(5):050819, 2013.

[10] John S Fletcher, John C Vickerman, and Nicholas Winograd. Label free biochemical 2d and 3d imaging using secondary ion mass spectrometry. Current opinion in chemical biology, 15(5):733–740, 2011.

[11] MS Wagner. Molecular depth profiling of multilayer polymer films using time-of-flight secondary ion mass spectrometry. Analytical chemistry, 77(3):911–922, 2005.

[12] Harald Jungnickel, Peter Laux, and Andreas Luch. Time-of-flight secondary ion mass spectrometry (tof-sims): A new tool for the analysis of toxicological effects on single cell level. Toxics, 4(1):5, 2016.

[13] Kristopher E Kubow, Radmila Vukmirovic, Lin Zhe, Enrico Klotzsch, Michael L Smith, Delphine Gourdon, Sheila Luna, and Viola Vogel. Mechanical forces regulate the interactions of fibronectin and collagen i in extracellular matrix. Nature communications, 6, 2015.

[14] Alexandra Naba, Karl R Clauser, Huiming Ding, Charles A Whittaker, Steven A Carr, and Richard O Hynes. The extracellular matrix: Tools and insights for the omics era. Matrix Biology, 49:10–24, 2016.

[15] Mathias Senoner and Wolfgang ES Unger. Sims imaging of the nanoworld: applications in science and technology. Journal of Analytical Atomic Spectrometry, 27(7):1050–1068, 2012.

[16] Srinath Rangarajan and Bonnie J Tyler. Topography in secondary ion mass spectroscopy images. Journal of Vacuum Science & Technology A, 24(5):1730–1736, 2006.

[17] Greg Gillen and Sonya Roberson. Preliminary evaluation of an sf5+ polyatomic primary ion beam for analysis of organic thin films by secondary ion mass spectrometry. Rapid communications in mass spectrometry, 12(19):1303–1312, 1998.

[18] Luc Van Vaeck, Annemie Adriaens, and Renaat Gijbels. Static secondary ion mass spectrometry (s-sims) part 1: methodology and structural interpretation. Mass Spectrometry Reviews, 18(1):1–47, 1999.

[19] A Benninghoven, D Jaspers, and W Sichtermann. Secondary-ion emission of amino acids. Applied physics, 11(1):35–39, 1976.

[20] A Benninghoven. Surface analysis by secondary ion mass spectrometry (sims). Surface Science, 299:246–260, 1994.

[21] Rana NS Sodhi. Time-of-flight secondary ion mass spectrometry (tof-sims): versatility in chemical and imaging surface analysis. Analyst, 129(6):483–487, 2004.

[22] John S Fletcher, Nicholas P Lockyer, and John C Vickerman. C60, buckminsterfullerene: its impact on biological tof-sims analysis. Surface and interface analysis, 38(11):1393–1400, 2006.

[23] Robert J Cotter. Time-of-flight mass spectrometry: basic principles and current state. In Time-of-Flight Mass Spectrometry, page 16, 1994.

[24] Anna Radionova, Igor Filippov, and Peter J Derrick. In pursuit of resolution in time-of-flight mass spectrometry: A historical perspective. Mass spectrometry reviews, 2015.

[25] Kenneth L Busch et al. Desorption ionization mass spectrometry. Journal of Mass Spectrometry, 30(2):233–240, 1995.

[26] RG Cooks and KL Busch. Matrix effects, internal energies and ms/ms spectra of molecular ions sputtered from surfaces. International Journal of Mass Spectrometry and Ion Physics, 53:111–124, 1983.

[27] W Gerhard and C Plog. Secondary ion emission by nonadiabatic dissociation of nascent ion molecules with energies depending on solid composition. Zeitschrift für Physik B Condensed Matter, 54(1):59–70, 1983.

[28] C Plog and W Gerhard. Secondary ion emission by nonadiabatic dissociation of nascent ion molecules with energies depending on solid composition. Zeitschrift für Physik B Condensed Matter, 54(1):71–86, 1983.

[29] Subhash Chandra. 3d subcellular sims imaging in cryogenically prepared single cells. Applied surface science, 231:467–469, 2004.

[30] Joseph Kozole, Christopher Szakal, Michael Kurczy, and Nicholas Winograd. Model multilayer structures for three-dimensional cell imaging. Applied surface science, 252(19):6789–6792, 2006.

[31] John S Fletcher, Nicholas P Lockyer, Seetharaman Vaidyanathan, and John C Vickerman. Tof-sims 3d biomolecular imaging of xenopus laevis oocytes using buckminsterfullerene (c60) primary ions. Analytical Chemistry, 79(6):2199–2206, 2007.

[32] John S Fletcher, Sadia Rabbani, Alex Henderson, Paul Blenkinsopp, Steve P Thompson, Nicholas P Lockyer, and John C Vickerman. A new dynamic in mass spectral imaging of single biological cells. Analytical chemistry, 80(23):9058–9064, 2008.

[33] John S Fletcher, Sadia Rabbani, Alex Henderson, Nicholas P Lockyer, and John C Vickerman. Three-dimensional mass spectral imaging of hela-m cells–sample preparation, data interpretation and visualisation. Rapid Communications in Mass Spectrometry, 25(7):925–932, 2011.

[34] Daniel Breitenstein, Christina E Rommel, Rudolf Möllers, Joachim Wegener, and Birgit Hagenhoff. The chemical composition of animal cells and their intracellular compartments reconstructed from 3d mass spectrometry. Angewandte Chemie International Edition, 46(28):5332–5335, 2007.

[35] Håkan Nygren, Birgit Hagenhoff, Per Malmberg, Mikael Nilsson, and Katrin Richter. Bioimaging tof-sims: High resolution 3d imaging of single cells. Microscopy research and technique, 70(11):969–974, 2007.

[36] Sutapa Ghosal, Stewart J Fallon, Terrance J Leighton, Katherine E Wheeler, Michael J Kristo, Ian D Hutcheon, and Peter K Weber. Imaging and 3d elemental characterization of intact bacterial spores by high-resolution secondary ion mass spectrometry. Analytical chemistry, 80(15):5986–5992, 2008.

[37] Christopher Szakal, Kedar Narayan, Jing Fu, Jonathan Lefman, and Sriram Subramaniam. Compositional mapping of the surface and interior of mammalian cells at submicrometer resolution. Analytical chemistry, 83(4):1207–1213, 2011.

[38] D Breitenstein, CE Rommel, J Stolwijk, Joachim Wegener, and B Hagenhoff. The chemical composition of animal cells reconstructed from 2d and 3d tof-sims analysis. Applied Surface Science, 255(4):1249–1256, 2008.

[39] Jeremy Brison, Michael A Robinson, Danielle SW Benoit, Shin Muramoto, Patrick S Stayton, and David G Castner. Tof-sims 3d imaging of native and non-native species within hela cells. Analytical chemistry, 85(22):10869–10877, 2013.

[40] Melissa Kathleen Passarelli, Carla J Newman, Peter S Marshall, Andrew West, Ian S Gilmore, Josephine Bunch, Morgan Russell Alexander, and Colin T Dollery. Single-cell analysis: Visualizing pharmaceutical and metabolite uptake in cells with label-free 3d mass spectrometry imaging. Analytical Chemistry, 2015.

[41] Daniel J Graham, John T Wilson, James J Lai, Patrick S Stayton, and David G Castner. Three-dimensional localization of polymer nanoparticles in cells using tof-sims. Biointerphases, 11(2):02A304, 2016.

[42] Amy V Walker. Why is sims underused in chemical and biological analysis? challenges and opportunities. Analytical chemistry, 80(23):8865–8870, 2008.

[43] ML Wagter, AH Clarke, KF Taylor, PAW Van Der Heide, and NS McIntyre. Topographic correction of 3d sims images. Surface and interface analysis, 25(10):788–789, 1997.

[44] Michael A Robinson, Daniel J Graham, and David G Castner. Tof-sims depth profiling of cells: z-correction, 3d imaging, and sputter rate of individual nih/3t3 fibroblasts. Analytical chemistry, 84(11):4880–4885, 2012.

[45] Sadia Rabbani, Andrew M Barber, John S Fletcher, Nicholas P Lockyer, and John C Vickerman. Tof-sims with argon gas cluster ion beams: a comparison with c60+. Analytical chemistry, 83(10):3793–3800, 2011.

[46] Russell L Steere. Electron microscopy of structural detail in frozen biological specimens. The Journal of biophysical and biochemical cytology, 3(1):45–60, 1957.

[47] Nicholas Winograd and Anna Bloom. Sample preparation for 3d sims chemical imaging of cells. In Mass Spectrometry Imaging of Small Molecules, pages 9–19. Springer, 2015.

[48] Jakob Malm, Dimitrios Giannaras, Mathis O Riehle, Nikolaj Gadegaard, and Peter SjoÌĹvall. Fixation and drying protocols for the preparation of cell samples for time-of-flight secondary ion mass spectrometry analysis. Analytical chemistry, 81(17):7197–7205, 2009.

[49] Kaija Schaepe, Julia Kokesch-Himmelreich, Marcus Rohnke, Alena-Svenja Wagner, Thimo Schaaf, Anja Henss, Sabine Wenisch, and Jürgen Janek. Storage of cell samples for tof-sims experimentsâĂŤhow to maintain sample integrity. Biointerphases, 11(2):02A313, 2016.

[50] Michael J Dykstra and Laura E Reuss. Biological electron microscopy: theory, techniques, and troubleshooting. Springer Science & Business Media, 2011.

[51] WB Bald. On crystal size and cooling rate. Journal of microscopy, 143(1):89–102, 1986.

[52] Marc Adrian, Jacques Dubochet, Jean Lepault, and Alasdair W McDowall. Cryo-electron microscopy of viruses. 1984.

[53] Jeremy Brison, Danielle SW Benoit, Shin Muramoto, Michael Robinson, Patrick S Stayton, and David G Castner. Tof-sims imaging and depth profiling of hela cells treated with bromodeoxyuridine. Surface and Interface Analysis, 43(1-2):354–357, 2011.

[54] Joe Wolfe and Gary Bryant. Freezing, drying, and/or vitrification of membrane–solute–water systems. Cryobiology, 39(2):103–129, 1999.

[55] HP Castro, PM Teixeira, and R Kirby. Evidence of membrane damage in lactobacillus bulgaricus following freeze drying. Journal of Applied Microbiology, 82(1):87–94, 1997.

[56] Merry Brennan, Bahijah Wanismail, MC Johnson, and Bibek Ray. Cellular damage in dried lactobacillus acidophilus. Journal of Food Protection®, 49(1):47–53, 1986.

[57] JG Linner, SA Livesey, DS Harrison, and AL Steiner. A new technique for removal of amorphous phase tissue water without ice crystal damage: a preparative method for ultrastructural analysis and immunoelectron microscopy. Journal of Histochemistry & Cytochemistry, 34(9):1123–1135, 1986.

[58] Ingela Lanekoff, Michael E Kurczy, Kelly L Adams, Jakob Malm, Roger Karlsson, Peter Sjövall, and Andrew G Ewing. An in situ fracture device to image lipids in single cells using tof-sims. Surface and Interface Analysis, 43(1-2):257–260, 2011.

[59] Elena SF Berman, Susan L Fortson, Kyle D Checchi, Ligang Wu, James S Felton, Kuang Jen J Wu, and Kristen S Kulp. Preparation of single cells for imaging/profiling mass spectrometry. Journal of the American Society for Mass Spectrometry, 19(8):1230–1236, 2008.

[60] Alan M Piwowar, Selda Keskin, Melissa Ortiz Delgado, Kan Shen, Jonathan J Hue, Ingela Lanekoff, Andrew G Ewing, and Nicholas Winograd. C60-tof sims imaging of frozen hydrated hela cells. Surface and Interface Analysis, 45(1):302–304, 2013.

[61] S Rabbani, JS Fletcher, NP Lockyer, and JC Vickerman. Exploring subcellular imaging on the buncher-tof j105 3d chemical imager. Surface and Interface Analysis, 43(1-2):380–384, 2011.

[62] Thomas P Roddy, Donald M Cannon, Sara G Ostrowski, Andrew G Ewing, and Nicholas Winograd. Proton transfer in time-of-flight secondary ion mass spectrometry studies of frozen-hydrated dipalmitoylphosphatidylcholine. Analytical chemistry, 75(16):4087–4094, 2003.

[63] Alan M Piwowar, John S Fletcher, Jeanette Kordys, Nicholas P Lockyer, Nicholas Winograd, and John C Vickerman. Effects of cryogenic sample analysis on molecular depth profiles with tof-secondary ion mass spectrometry. Analytical chemistry, 82(19):8291–8299, 2010.

[64] Michael A Robinson and David G Castner. Characterization of sample preparation methods of nih/3t3 fibroblasts for tof-sims analysis. Biointerphases, 8(1):15, 2013.

[65] Per Malmberg, Christian Kriegeskotte, Heinrich F Arlinghaus, Birgit Hagenhoff, Jan Holmgren, Mikael Nilsson, and Håkan Nygren. Depth profiling of cells and tissues by using c 60+ and sf 5+ as sputter ions. Applied Surface Science, 255(4):926–928, 2008.

[66] Michael Taylor, Andrew J Urquhart, Daniel G Anderson, Robert Langer, Martyn C Davies, and Morgan R Alexander. Partial least squares regression as a powerful tool for investigating large combinatorial polymer libraries. Surface and Interface Analysis, 41(2):127–135, 2009.

[67] VS Smentkowski, SG Ostrowski, and MR Keenan. A comparison of multivariate statistical analysis protocols for tof-sims spectral images. Surface and Interface Analysis, 41(2):88–96, 2009.

[68] MS Wagner, DJ Graham, and DG Castner. Simplifying the interpretation of tof-sims spectra and images using careful application of multivariate analysis. Applied surface science, 252(19):6575–6581, 2006.

[69] Bonnie J Tyler, Gaurav Rayal, and David G Castner. Multivariate analysis strategies for processing tof-sims images of biomaterials. Biomaterials, 28(15):2412–2423, 2007.

[70] Alex Henderson, John S Fletcher, and John C Vickerman. A comparison of

pca and maf for tof-sims image interpretation. Surface and Interface Analysis, 41(8):666–674, 2009.

[71] OD Sanni, MS Wagner, D Briggs, DG Castner, and JC Vickerman. Classification of adsorbed protein static tof-sims spectra by principal component analysis and neural networks. Surface and Interface Analysis, 33(9):715–728, 2002.

[72] Daniel J Graham and David G Castner. Multivariate analysis of tof-sims data from multicomponent systems: the why, when, and how. Biointerphases, 7(1):49, 2012.

[73] James N Miller and Jane Charlotte Miller. Statistics and chemometrics for analytical chemistry. Pearson Education, 2005.

[74] Karl Pearson. Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11):559–572, 1901.

[75] Harold Hotelling. Analysis of a complex of statistical variables into principal components. Journal of educational psychology, 24(6):417, 1933.

[76] Harold Hotelling. Relations between two sets of variates. Biometrika, 28(3/4):321–377, 1936.

[77] Richard G Brereton. Applied chemometrics for scientists. John Wiley & Sons, 2007.

[78] Norm O'Rourke, Larry Hatcher, and Edward J Stepanski. A step-by-step approach to using SAS for univariate & multivariate statistics. SAS Institute, 2005.

[79] Robert M Thorndike. Correlational procedures for research. Wiley, 1976.

[80] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In International Conference on Artificial Neural Networks, pages 583–588. Springer, 1997.

[81] Thomas Stephan. Tof-sims in cosmochemistry. Planetary and Space Science, 49(9):859–906, 2001.

[82] Sandro Ferrari and Buddy D Ratner. Tof-sims quantification of albumin adsorbed on plasma-deposited fluoropolymers by partial least-squares regression. Surface and interface analysis, 29(12):837–844, 2000.

[83] MS Wagner, DJ Graham, BD Ratner, and David G Castner. Maximizing information obtained from secondary ion mass spectra of organic thin films using multivariate analysis. Surface Science, 570(1):78–97, 2004.

[84] Vincent S Smentkowski. Quantitative imaging of trace b in si and sio2 using tof-sims. Journal of Vacuum Science & Technology A, 33(5):05E107, 2015.

[85] Alexander G Shard, Rasmus Havelund, Steve J Spencer, Ian S Gilmore, Morgan R Alexander, Tina B Angerer, Satoka Aoyagi, Jean-Paul Barnes, Anass Benayad, Andrzej Bernasik, et al. Measuring compositions in organic depth profiling: Results from a vamas interlaboratory study. The Journal of Physical Chemistry B, 119(33):10784–10797, 2015.

[86] Rudolf J Rummel. Applied factor analysis. Northwestern University Press, 1988.

[87] Michael R Keenan and Paul G Kotula. Accounting for poisson noise in the multivariate analysis of tof-sims spectrum images. Surface and interface analysis, 36(3):203–212, 2004.

[88] Jon Shlens. A tutorial on principal component analysis: Derivation, discussion and singular value decomposition, 2003.

[89] Pierre Comon. Independent component analysis, a new concept? Signal processing, 36(3):287–314, 1994.

[90] Willem F Wolkers, Fern Tablin, and John H Crowe. From anhydrobiosis to freeze-drying of eukaryotic cells. Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology, 131(3):535–543, 2002.

[91] Hans Moor. Die gefrier-fixation lebender zellen und ihre anwendung in der elektronenmikroskopie. Zeitschrift für Zellforschung und Mikroskopische Anatomie, 62(4):546–580, 1964.

[92] Patrick Echlin. Low-temperature microscopy and analysis. Springer Science & Business Media, 2013.

[93] Konrad J Kossacki and Jacek Leliwa-Kopystynski. Temperature dependence of the sublimation rate of water ice: Influence of impurities. Icarus, 233:101–105, 2014.

[94] Richard Leach, L Brown, X Jiang, R Blunt, M Conroy, and D Mauger. Guide to the measurement of smooth surface topography using coherence scanning interferometry. Measurement good practice guide, 108:17, 2008.

[95] M Raposo, Q Ferreira, and PA Ribeiro. A guide for atomic force microscopy analysis of soft-condensed matter. Modern research and educational topics in microscopy, 1:758–769, 2007.

[96] Hsiang-Ling Huang, Hsiang-Wei Hsing, Tzu-Chia Lai, Yi-Wen Chen, Tian-Ren Lee, Hsin-Tsu Chan, Ping-Chiang Lyu, Chieh-Lin Wu, Ying-Chieh Lu, Szu-Ting Lin, et al. Trypsin-induced proteome alteration during cell subculture in mammalian cells. Journal of biomedical science, 17(1):1, 2010.

[97] Michael Taylor, David Scurr, Matthias Lutolf, Lee Buttery, Mischa Zelzer, and Morgan Alexander. 3d chemical characterization of frozen hydrated hydrogels using tof-sims with argon cluster sputter depth profiling. Biointerphases, 11(2):02A301, 2016.

[98] Christopher Szakal, Joseph Kozole, and Nicholas Winograd. Fundamental studies of the cluster ion bombardment of water ice. Applied surface science, 252(19):6526–6528, 2006.

[99] AG Shard, PJ Brewer, FM Green, and IS Gilmore. Measurement of sputtering yields and damage in c60 sims depth profiling of model organic materials. Surface and interface analysis, 39(4):294–298, 2007.

[100] Sebastiaan Van Nuffel, Christopher Parmenter, David J Scurr, Noah A Russell, and Mischa Zelzer. Multivariate analysis of 3d tof-sims images: method validation and application to cultured neuronal networks. Analyst, 141(1):90–95, 2016.

[101] Yannick Vercammen, Nicola Moons, Sebastiaan Van Nuffel, Linda Beenaerts, and Luc Van Vaeck. Experimental study of the organic ion intensity distribution in the ion imaging of coated polymer fibres with s-sims. Applied Surface Science, 284:354–365, 2013.

[102] James Bailey, Rasmus Havelund, Alexander G Shard, Ian S Gilmore, Morgan R Alexander, James S Sharp, and David J Scurr. 3d tof-sims imaging of polymer multilayer films using argon cluster sputter depth profiling. ACS applied materials & interfaces, 7(4):2654–2659, 2015.

[103] Alan M Race and Josephine Bunch. Optimisation of colour schemes to accurately display mass spectrometry imaging data based on human colour perception. Analytical and bioanalytical chemistry, 407(8):2047–2054, 2015.

[104] Peter J Cumpson, Naoko Sano, Ian W Fletcher, Jose F Portoles, Mariela Bravo-Sanchez, and Anders J Barlow. Multivariate analysis of extremely large tof-

sims imaging datasets by a rapid pca method. Surface and Interface Analysis, 47(10):986–993, 2015.

[105] Daniel J Graham, Matthew S Wagner, and David G Castner. Information from complexity: Challenges of tof-sims data interpretation. Applied surface science, 252(19):6860–6868, 2006.

[106] Siegfried Hofmann. From depth resolution to depth resolution function: refinement of the concept for delta layers, single layers and multilayers. Surface and interface analysis, 27(9):825–834, 1999.

[107] Katty X Wan, Ilan Vidavsky, and Michael L Gross. Comparing similar spectra: from similarity index to spectral contrast angle. Journal of the American Society for Mass Spectrometry, 13(1):85–88, 2002.

[108] John S Fletcher, Sadia Rabbani, Andrew M Barber, Nicholas P Lockyer, and John C Vickerman. Comparison of c60 and gcib primary ion beams for the analysis of cancer cells and tumour sections. Surface and Interface Analysis, 45(1):273–276, 2013.

[109] Didier Falconnet, Gabor Csucs, H Michelle Grandin, and Marcus Textor. Surface engineering approaches to micropattern surfaces for cell-based assays. Biomaterials, 27(16):3044–3063, 2006.

[110] Bengt Kasemo. Biological surface science. Surface Science, 500(1):656–677, 2002.

[111] Stephan Preibisch, Stephan Saalfeld, and Pavel Tomancak. Globally optimal stitching of tiled 3d microscopic image acquisitions. Bioinformatics, 25(11):1463–1465, 2009.

[112] MS Wagner and David G Castner. Characterization of adsorbed protein films by time-of-flight secondary ion mass spectrometry with principal component analysis. Langmuir, 17(15):4649–4660, 2001.

[113] Luc J Gentet, Greg J Stuart, and John D Clements. Direct measurement of specific membrane capacitance in neurons. Biophysical Journal, 79(1):314–320, 2000.

[114] J Kistler and E Kellenberger. Collapse phenomena in freeze-drying. Journal of ultrastructure research, 59(1):70–75, 1977.

[115] Collin J May, Heather E Canavan, and David G Castner. Quantitative x-ray photoelectron spectroscopy and time-of-flight secondary ion mass spectrometry characterization of the components in dna. Analytical chemistry, 76(4):1114–1122, 2004.

[116] Kyoko Shimizu, Marie-Laure Abel, and John F Watts. Evaluation of the interaction and adsorption of $\gamma$-glycidoxy propyl trimethoxy silane with grit-blasted aluminium: A tof-sims and xps study. The Journal of Adhesion, 84(8):725–741, 2008.

[117] Heidy Visbal, Kohji Nagashima, and Kazuyuki Hirao. Surface analysis of three aluminum foils and relation to hydrogen generation capability. Korean Journal of Chemical Engineering, 33(4):1255–1260, 2016.

[118] Melissa K Passarelli and Nicholas Winograd. Lipid imaging with time-of-flight secondary ion mass spectrometry (tof-sims). Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 1811(11):976–990, 2011.

[119] Suet YiáLiu et al. Elucidation of the chemical composition of avian melanin. RSC Advances, 4(76):40396–40399, 2014.

[120] Jan Toporski and Andrew Steele. Characterization of purified biomarker compounds using time of flight-secondary ion mass spectrometry (tof-sims). Organic Geochemistry, 35(7):793–811, 2004.

[121] Paul D Piehowski, Anthony J Carado, Michael E Kurczy, Sara G Ostrowski, Michael L Heien, Nicholas Winograd, and Andrew G Ewing. Ms/ms methodology to improve subcellular mapping of cholesterol using tof-sims. Analytical chemistry, 80(22):8662–8667, 2008.

[122] Melissa Kathleen Passarelli. Lipid Characterization with Time-of-flight Secondary Ion Mass Spectrometry (ToF-SIMS). PhD thesis, The Pennsylvania State University, 2011.

[123] Daniel J Graham and David G Castner. Image and spectral processing for tof-sims analysis of biological materials. Mass Spectrometry, 2(Special_Issue):S0014–S0014, 2013.

[124] Christopher R Anderton, Bita Vaezian, Kaiyan Lou, Jessica F Frisz, and Mary L Kraft. Identification of a lipid-related peak set to enhance the interpretation of tof-sims data from model and cellular membranes. Surface and Interface Analysis, 44(3):322–333, 2012.

[125] Maria E McNamara, Bart E Dongen, Nick P Lockyer, Ian D Bull, and Patrick J Orr. Fossilization of melanosomes via sulfurization. Palaeontology, 59(3):337–350, 2016.

[126] Nicholas P Lockyer. 14 laser post-ionisation for elemental and molecular surface analysis. TOF-SIMS: Materials Analysis by Mass Spectrometry, page 361, 2013.

[127] M Fartmann, S Dambach, C Kriegeskotte, HP Wiesmann, A Wittig, W Sauerwein, D Lipinsky, and HF Arlinghaus. Characterization of cell cultures with tof-sims and laser-snms. Surface and interface analysis, 34(1):63–66, 2002.

[128] Tina B Angerer and John S Fletcher. 3d imaging of tio2 nanoparticle accumulation in tetrahymena pyriformis. Surface and Interface Analysis, 46(S1):198–203, 2014.

[129] Kaija Schaepe, Julia Kokesch-Himmelreich, Marcus Rohnke, Alena-Svenja Wagner, Thimo Schaaf, Sabine Wenisch, and Jürgen Janek. Assessment of different sample preparation routes for mass spectrometric monitoring and imaging of lipids in bone cells via tof-sims. Biointerphases, 10(1):019016, 2015.

[130] Shawn Parry and Nicholas Winograd. High-resolution tof-sims imaging of eukaryotic cells preserved in a trehalose matrix. Analytical chemistry, 77(24):7950–7957, 2005.

[131] Heather E Canavan, Daniel J Graham, Xuanhong Cheng, Buddy D Ratner, and David G Castner. Comparison of native extracellular matrix with adsorbed protein films using secondary ion mass spectrometry. Langmuir, 23(1):50–56, 2007.

[132] Károly Vékey. Internal energy effects in mass spectrometry. Journal of mass spectrometry, 31(5):445–463, 1996.

[133] Xavier A Conlan, Nicholas P Lockyer, and John C Vickerman. Is proton cationization promoted by polyatomic primary ion bombardment during time-of-flight secondary ion mass spectrometry analysis of frozen aqueous solutions? Rapid communications in mass spectrometry, 20(8):1327–1334, 2006.

[134] Tina B Angerer, Masoumeh Dowlatshahi Pour, Per Malmberg, and John S Fletcher. Improved molecular imaging in rodent brain with time-of-flight-secondary ion mass spectrometry using gas cluster ion beams and reactive vapor exposure. Analytical chemistry, 87(8):4305–4313, 2015.

[135] Masoumeh Dowlatshahi Pour, Per Malmberg, and Andrew Ewing. An investigation on the mechanism of sublimed dhb matrix on molecular ion yields in sims imaging of brain tissue. Analytical and bioanalytical chemistry, 408(12):3071–3081, 2016.

# Appendices

APPENDIX A

# Supplementary Data to Chapter 3

## A.1 Data Used for Mock TOF-SIMS Data Formation

**Table A.1:** Normalised ion intensities for acrylamide and coumarin, obtained from SDBSWeb : http://sdbs.db.aist.go.jp (National Institute of Advanced Industrial Science and Technology, December 2014). The sum of rest is given by the sum of the intensities that fall outside of the mass range or mass increments. The underlined m/z are those used as specific ions.

| m/z | acrylamide | coumarin |
|---|---|---|
| total | 1.0000 | 1.0000 |
| sum of rest | 0.0000 | 0.0077 |
| 14 | 0.0023 | 0.0000 |
| 15 | 0.0031 | 0.0000 |
| 16 | 0.0096 | 0.0000 |
| 17 | 0.0131 | 0.0000 |
| 24 | 0.0019 | 0.0000 |
| 25 | 0.0102 | 0.0000 |
| 26 | 0.0529 | 0.0045 |
| 27 | 0.1804 | 0.0043 |
| 28 | 0.0354 | 0.0043 |
| 29 | 0.0046 | 0.0000 |
| 30 | 0.0248 | 0.0000 |
| 37 | 0.0000 | 0.0103 |
| 38 | 0.0033 | 0.0213 |
| 39 | 0.0027 | 0.0313 |
| 40 | 0.0038 | 0.0060 |
| 41 | 0.0040 | 0.0000 |

| | | |
|---|---|---|
| 42 | 0.0052 | 0.0000 |
| 43 | 0.0523 | 0.0089 |
| 44 | 0.1923 | 0.0043 |
| 45 | 0.0094 | 0.0151 |
| 49 | 0.0000 | 0.0045 |
| 50 | 0.0000 | 0.0213 |
| 51 | 0.0058 | 0.0273 |
| 52 | 0.0100 | 0.0029 |
| 53 | 0.0119 | 0.0045 |
| 54 | 0.0067 | 0.0000 |
| 55 | 0.1398 | 0.0000 |
| 56 | 0.0048 | 0.0000 |
| 59 | 0.0000 | 0.0134 |
| 61 | 0.0000 | 0.0120 |
| 62 | 0.0000 | 0.0285 |
| 63 | 0.0000 | 0.0692 |
| 64 | 0.0000 | 0.0287 |
| 70 | 0.0079 | 0.0000 |
| 71 | 0.1922 | 0.0000 |
| 72 | 0.0094 | 0.0000 |
| 73 | 0.0000 | 0.0031 |
| 74 | 0.0000 | 0.0079 |
| 75 | 0.0000 | 0.0060 |
| 85 | 0.0000 | 0.0031 |
| 86 | 0.0000 | 0.0048 |
| 87 | 0.0000 | 0.0045 |
| 88 | 0.0000 | 0.0024 |
| 89 | 0.0000 | 0.0893 |
| 90 | 0.0000 | 0.1084 |
| 91 | 0.0000 | 0.0084 |
| 92 | 0.0000 | 0.0062 |
| 118 | 0.0000 | 0.2393 |
| 119 | 0.0000 | 0.0213 |
| 146 | 0.0000 | 0.1503 |
| 147 | 0.0000 | 0.0148 |

## A.2 Matlab Code to Import ToF-SIMS Data

```
function [ FileList, N, IMAGECHAR, id, alpha ] = im_info
%im_info gathers FileList, number of files N, FOV, size, name, id
% and alpha

clearvars

% imports files from directory we're in (all image files together)
FileList = dir('*.txt');

% number of files ( = total + sum of rest + ions )
N = size(FileList,1);

% load first file and set delimiter
filename = FileList(1).name;
delimiter = ' ';

% retrieve FOV

startRow = 3;
endRow = 3;
formatSpec = '%*s%*s%*s%*s%f%*s%f%*s%*s%[^\n\r]';
fileID = fopen(filename,'r');
textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
dataArray = textscan(fileID, formatSpec, endRow-startRow+1,
'Delimiter', delimiter, 'MultipleDelimsAsOne', true, 'ReturnOnError',
false);
fclose(fileID);

FOV = [dataArray{1:end-1}];

clearvars startRow endRow formatSpec fileID dataArray ans;

% retrieve SIZE

startRow = 9;
endRow = 9;
formatSpec = '%*s%*s%*s%f%*s%f%*s%f%*s%[^\n\r]';
fileID = fopen(filename,'r');
textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
dataArray = textscan(fileID, formatSpec, endRow-startRow+1,
'Delimiter', delimiter, 'MultipleDelimsAsOne', true, 'ReturnOnError',
false);
```

```
fclose(fileID);

SIZE = [dataArray{1:end-1}];

clearvars startRow endRow formatSpec fileID dataArray ans;

% retrieve NAME

endRow = 1;
formatSpec = '%*s%*s%*s%s%*s%*s%*s%*s%[^\n\r]';
fileID = fopen(filename,'r');
dataArray = textscan(fileID, formatSpec, endRow, 'Delimiter',
delimiter, 'MultipleDelimsAsOne', true, 'ReturnOnError', false);
fclose(fileID);
NAME = dataArray{:, 1};

clearvars filename delimiter startRow endRow formatSpec
fileID dataArray ans;

IMAGECHAR{1,1} = NAME;
IMAGECHAR{2,1} = FOV;
IMAGECHAR{3,1} = SIZE;

clearvars NAME FOV SIZE;

alpha = cell(1,N-2);
id = ones(1,N-2);

for k = 3:N;

    % loads individual txt file
    filename = FileList(k).name;
    delimiter = ' ';

    % ID

    startRow = 2;
    endRow = 2;
    formatSpec = '%*s%*s%*s%*s%*s%*f%*s%*s%[^\n\r]';
    fileID = fopen(filename,'r');
    textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
    dataArray = textscan(fileID, formatSpec, endRow-startRow+1,
        'Delimiter', delimiter, 'MultipleDelimsAsOne', true,
        'ReturnOnError', false);
    fclose(fileID);

```

```
    ID = [dataArray{1:end-1}];

    id(1,k-2)=ID;

    % ions
    startRow = 2;
    endRow = 2;
    formatSpec = '%*s%*s%*s%s%*s%*s%*s%*s%[^\n\r]';
    fileID = fopen(filename,'r');
    textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
    dataArray = textscan(fileID, formatSpec, endRow-startRow+1,
        'Delimiter', delimiter, 'MultipleDelimsAsOne', true,
        'ReturnOnError', false);
    fclose(fileID);

    IND_ION = dataArray{:, 1};

    alpha{1,ID-1} = IND_ION;

end

clearvars k filename delimiter startRow endRow formatSpec
fileID dataArray ID IND_ION ans;


end
```

```
function [ temp ] = im_ionint2( FileList, IMAGECHAR, k )
%UNTITLED3 Summary of this function goes here
%   Detailed explanation goes here

temp = cell(1,1);

filename = FileList(k).name;
delimiter = ' ';

startRow = 10;
%formatSpec = '%*s%*s%*s%f%*s%*s%*s%*s%*s%[^\n\r]';
formatSpec = '%*s%*s%*s%f%*s%*s%*s%*s%*s%*s%[^\n\r]';
fileID = fopen(filename,'r');
textscan(fileID, '%[^\n\r]', startRow-1, 'ReturnOnError', false);
dataArray = textscan(fileID, formatSpec, 'Delimiter',
delimiter, 'MultipleDelimsAsOne', true, 'EmptyValue' ,0.0,
'ReturnOnError', false);
```

```
fclose(fileID);

clearvars filename delimiter startRow formatSpec fileID ;

temp{1,1} = sparse( [dataArray{1:end-1}] );

clearvars dataArray ;

end
```

```
clearvars

str1 = sprintf('%s - BEGIN PROSSEX', datestr(now));
disp(str1)
clearvars str1

% gather image information

[FileList,N,IMAGECHAR,id,alpha] = im_info;

save im_info.mat FileList N IMAGECHAR id alpha -v7.3;

clearvars alpha

% gather ion intensities (excl. total + sum of rest) in num1

num1 = cell(1,N-2);

for k = 3:N

    [ temp ] = im_ionint2( FileList, IMAGECHAR, k );

    num1(1,id(1,k-2)-1) = temp;

    clearvars temp

    str1 = sprintf('%s - FILE %d OF %d COMPLETE', datestr(now),k-2,N-2);
    disp(str1)
    clearvars str1

end

clearvars k FileList N IMAGECHAR id

```

```
save num1.mat num1 -v7.3;

clearvars num1

str1 = sprintf('%s - OUTPUT SAVED', datestr(now));
disp(str1)
clearvars str1
```

```
clearvars

str1 = sprintf('%s - BEGIN PROSSEX', datestr(now));
disp(str1)
clearvars str1

% gather image information

load im_info.mat
clearvars N alpha id

% gather total ion intensities
[ total1 ] = im_ionint2( FileList, IMAGECHAR,1 );

save total1.mat total1 -v7.3;

clearvars

str1 = sprintf('%s - OUTPUT SAVED', datestr(now));
disp(str1)
clearvars str1
```

## A.3 Matlab Code to Reshape Imported Data

```
% needed: IMAGECHAR, N, num1

str1 = sprintf('%s - BEGIN PROSSEX', datestr(now));
disp(str1)
clearvars str1

load im_info.mat
clearvars FileList alpha id
load num1.mat
```

```

num=cell(IMAGECHAR{3,1}(1,3),N-2);

A=linspace(0,prod(IMAGECHAR{3,1}),IMAGECHAR{3,1}(1,3)+1);

for k = 1:N-2;

for l= 1:IMAGECHAR{3,1}(1,3) ;

    num{l,k} = rot90( reshape( num1{1,k}(A(l)+1:A(l+1),1),
        IMAGECHAR{3,1}(1,1),IMAGECHAR{3,1}(1,2) ) );

end

    str1 = sprintf('%s - ION %d OF %d COMPLETE', datestr(now),k,N-2);
    disp(str1)
    clearvars str1

end

clearvars A k l num1 IMAGECHAR N

save num.mat num -v7.3;

clearvars num

str1 = sprintf('%s - END PROSSEX', datestr(now));
disp(str1)
clearvars str1
```

```
% needed: IMAGECHAR, total

str1 = sprintf('%s - BEGIN PROSSEX', datestr(now));
disp(str1)
clearvars str1

load im_info.mat
clearvars FileList N alpha id
load total1.mat

% reshape
total=cell(IMAGECHAR{3,1}(1,3),1);

A=linspace(0,prod(IMAGECHAR{3,1}),IMAGECHAR{3,1}(1,3)+1);
```

```
15
16 for l= 1:IMAGECHAR{3,1}(1,3) ;
17
18     total{1,1} = rot90( reshape( total1{1,1}(A(l)+1:A(l+1),1),
19         IMAGECHAR{3,1}(1,1),IMAGECHAR{3,1}(1,2) ) );
20
21 end
22
23 clearvars IMAGECHAR total1 A l
24
25 save total.mat total -v7.3
26
27 clearvars total
28
29 str1 = sprintf('%s - END PROSSEX', datestr(now));
30 disp(str1)
31 clearvars str1
```

## A.4 Matlab Code for the Normalisation of the Image Data

```
1  clearvars
2
3  str1 = sprintf('%s - BEGIN PROSSEX', datestr(now));
4  disp(str1)
5  clearvars str1
6
7  %load variables
8  load num.mat
9  load total.mat
10
11 % create new cell for normalised intensities
12 normalised = cell(size(num));
13
14 save normalised_v2.mat normalised -v7.3
15
16 clearvars normalised
17
18 % load matfile
19
20 m = matfile('normalised_v2.mat','Writable',true);
21
22 for k = 1:size(num,2)
```

```
23
24      temp = cell(size(num,1),1);
25
26      for l=1:size(num,1)
27
28      temp{l,1} = num{l,k}./total{l,1};
29
30      temp{l,1}(isnan(temp{l,1})) = 0; % clears nan when total = 0
31
32      end
33
34      m.normalised(:,k) = temp;
35
36      str1 = sprintf('%s - ION %d OF %d COMPLETE', datestr(now),k,
37          size(num,2));
38      disp(str1)
39      clearvars str1 l temp
40
41
42  end
43
44  clearvars k m total num
45
46  str1 = sprintf('%s - OUTPUT SAVED', datestr(now));
47  disp(str1)
48  clearvars str1
```

## A.5 Matlab Code to Calculate the PCA Loadings

```
1  % create a training set
2  % needs normalised, IMAGECHAR, amount of random pixels per plane
3
4  training=ones(1000*IMAGECHAR{3,1}(1,3),size(normalised,2));
5  L=linspace(0,1000*IMAGECHAR{3,1}(1,3),IMAGECHAR{3,1}(1,3)+1);
6
7  for l = 1:IMAGECHAR{3,1}(1,3);
8
9      A=randperm(IMAGECHAR{3,1}(1,1)*IMAGECHAR{3,1}(1,2),1000);
10     B=unique(A);
11     clearvars A
12
13     for k = 1:size(normalised,2);
```

```
14
15          training(L(l)+1:L(l+1),k)=normalised{l,k}(B);
16
17      end
18
19      clearvars B
20
21  end
22
23  clearvars L l k
24
25  save analysis_normalised.mat training -v7.3
26
27  str1 = sprintf('TRAINING SET FORMED');
28  disp(str1);
29  clearvars str1
30
31  % calculate PCA coefficients
32
33  % [coeff,~,~,~,explained] =
34  % pca(training,'Algorithm','eig','Centered',false,'NumComponents',6);
35  % not centered pca
36
37  % [coeff,D] = eigs(cov(training)); % equivalent to centered pca (cov)
38
39  [coeff,D] = eigs(corr(training)); % equivalent to standardised pca (corr)
40
41  explained = (diag(D)/trace(D))*100; % variance as a percentage
42  clearvars D
43
44  save analysis_normalised.mat coeff explained -v7.3 -append
45
46  str1 = sprintf('PCA COEFF CALCULATED');
47  disp(str1);
48  clearvars str1
49
50  % transform alpha from cell to double, which can be used in bar
51  % needs alpha
52
53  ALPHA = ones(size(alpha));
54
55  for k = 1:size(alpha,2);
56
57      A = cell2mat(alpha{k});
58      A = str2double(A);
59      ALPHA(k)=A;
```

```
60
61      clearvars A
62
63  end
64
65  clearvars k
66
67  % loadings plots
68  % needs ALPHA, coeff, explained
69
70  for k = 1:6 % chart for each PC
71
72  h = figure; % create figure
73  bar(ALPHA, coeff(:,k)) % bar chart for PC k
74
75  str1=sprintf('PC %d (%1.2f percent variance explained)', k, explained(k));
76  title(str1); % puts PC number and its explained variance in title
77
78  xlabel('m/z');
79  ylabel('loadings');
80
81  str2=sprintf('Loadings_normalised_PC%d', k);
82  saveas(h,str2,'fig') % save figure
83
84  close(h) % close figure
85
86  end
87
88  clearvars k h str1 str2
89
90  str1 = sprintf('LOADINGS PLOTS FORMED');
91  disp(str1);
92  clearvars str
```

## A.6 Matlab Code to Calculate PCA Scores

```
1  % formation of scores
2  % needs normalised, IMAGECHAR, coeff, training
3
4   A=mean(training);
5   B=std(training);
6  %
```

```
% clearvars training

normalised_SCORES = cell(IMAGECHAR{3,1}(1,3),6);

for l = 1:IMAGECHAR{3,1}(1,3);

    data=ones(IMAGECHAR{3,1}(1,1)*IMAGECHAR{3,1}(1,2),size(normalised,2));

    for k = 1:size(normalised,2);

        data(:,k)=reshape(normalised{l,k},IMAGECHAR{3,1}(1,1)
                *IMAGECHAR{3,1}(1,2),1);

    end

    data_stand =(data - (ones(size(data,1),1)*A))
        ./(ones(size(data,1),1)*B);

    scores = data_stand*coeff;

    for k = 1:6

        normalised_SCORES{l,k}=reshape( scores(:,k),IMAGECHAR{3,1}(1,1),
                IMAGECHAR{3,1}(1,2) );

    end

end

clearvars k l data data_stand scores

save scores.mat normalised_SCORES -v7.3

str1 = sprintf('SCORES FORMED');
disp(str1);
clearvars str1
```

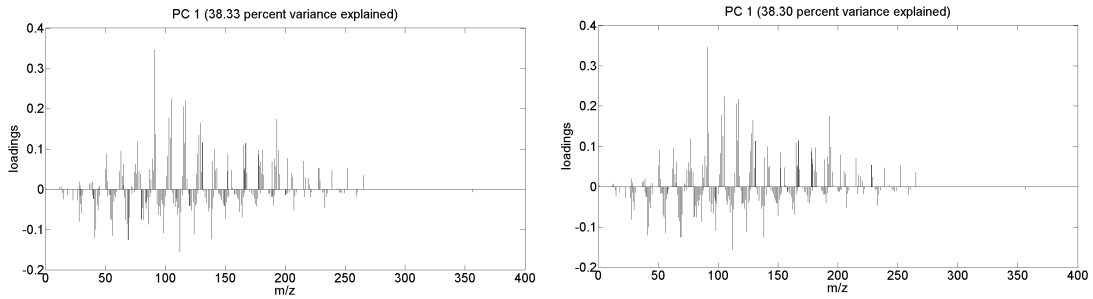
## A.7 Matlab Code to Form 3D Scores Images

```
datestr(now)

V = cell(max(Zr{1,1}(:)) - min(Zr{end,1}(:)) + 1,1);
```

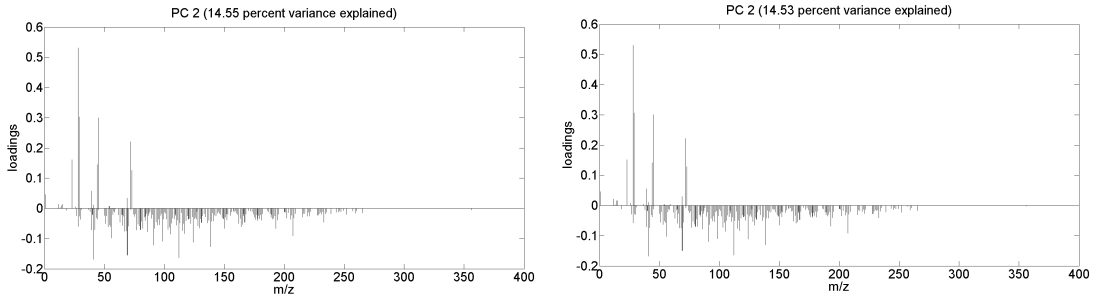
```
for l = 1:size(V,1)
V{l,1}=NaN(256,256);
end
clearvars l

%[I,J] = ind2sub(size(Zr{k,1}),find(Zr{k,1}==l));

MaxZR = max(Zr{1,1}(:));
MinZR = min(Zr{end,1}(:));


LS=linspace(MinZR,MaxZR,size(V,1));

for k = 1:size(Zr,1)

    if max(Zr{k,1}(:)) > -11;

    for l = 1:size(V,1);

        A=find(Zr{k,1}==LS(l));

        V{l,1}(A)=normalised_SCORES{k,3}(A);

        clearvars A
    end


    end

    clearvars k
%
% A=find(Zr{k,1}==l)
%
% V{1,1}(B)=normalised_SCORES{1,1}(B);
%
%
% V{l+1,1}=normalised_SCORES{k,1}(I,J);

end

datestr(now)

%
%
%
for l = 1:size(V,1)
```

```
50
51      Vmat(:,:,l) = V{l,1};
52
53  end
54
55  clearvars l
56
57  for l = 1:size(V,1)
58
59      Z(:,:,l) = (ones([256,256])*(l-MaxZR+1)*0.060)-0;
60      X(:,:,l) = ones([256,1])*linspace(0, 250,256);
61      Y(:,:,l) = linspace(0, 250,256)'*ones([1,256]);
62
63  end
64
65  figure;
66  k=slice(X,Y,Z,Vmat,[],[],[-0.6:0.060:10.140]);
67  set(k,'edgecolor','none');
68  axis xy tight vis3d;
69  xlabel('x (\mum)','FontSize', 20)
70  ylabel('y (\mum)','FontSize', 20)
71  ylabel(colorbar,'principal component score','FontSize', 20)
72  zlabel('z (\mum)','FontSize', 20)
73  % title('PC1 (68.24 % variance explained)','FontSize', 20)
74  %caxis([0,10]);
75  daspect([4,4,1]);
76
77  for n=1:length(k)
78
79      set(k(n),'alphadata',get(k(n),'cdata'),'facealpha',0.25)
80
81  end
```

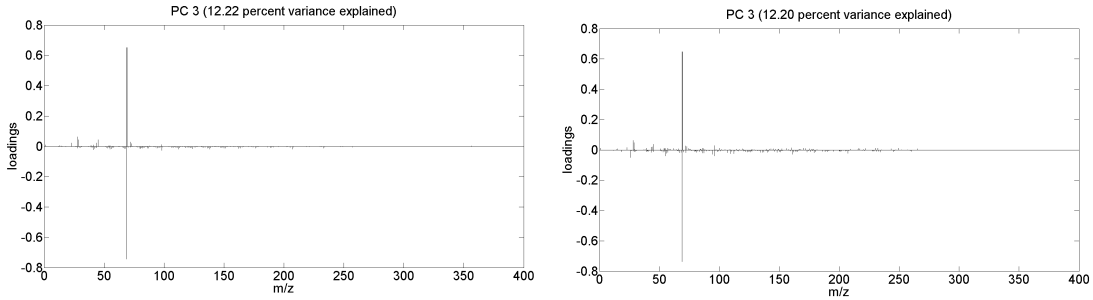
## A.8 Supplementary Figures




**(a)** PC1: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals 0.99987.

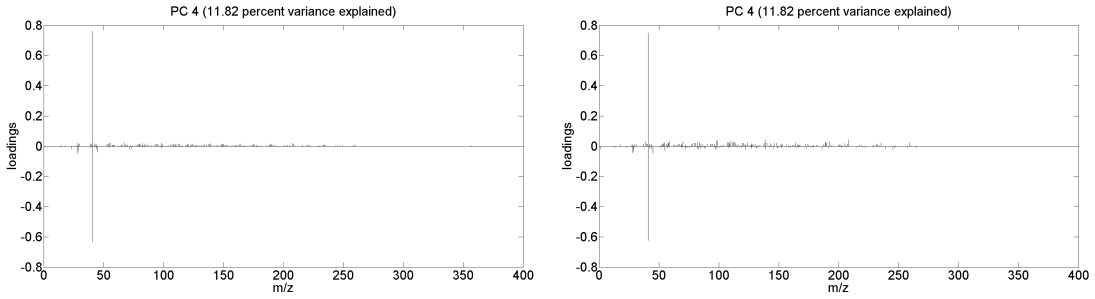



**(b)** PC2: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals 0.9989.

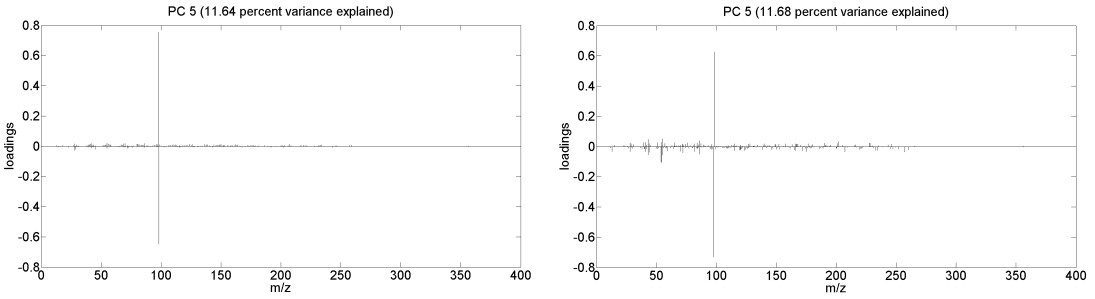



**(c)** PC3: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals 0.9907.

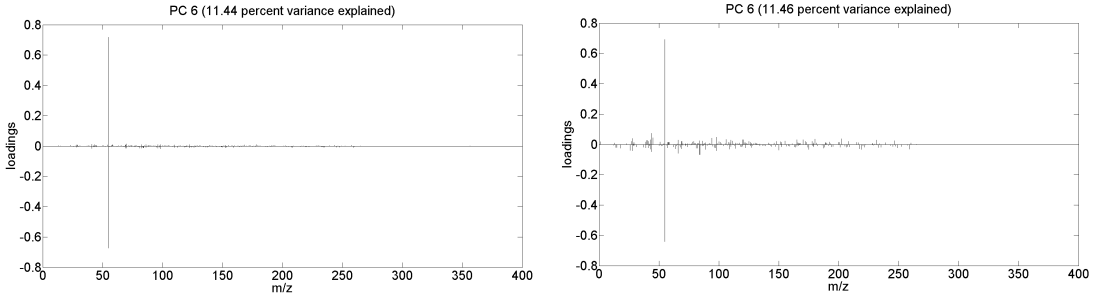
**Figure A.1:** A visual comparison of the loadings obtained from full PCA and those obtained from a PCA with the training set method.




**(d)** PC4: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals 0.984.




**(e)** PC5: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals -0.965.




**(f)** PC6: full PCA result (left) and subsampled PCA result (right). The cosine similarity $\cos\theta$ between the loadings equals 0.96.

**Figure A.1:** A visual comparison of the loadings obtained from full PCA and those obtained from a PCA with the training set method. (cont.)

## A.9 Supplementary Tables

**Table A.2:** Parameters of the Gaussian curves ($a_i e^{-((x_i-b_i)/(c_i))^2}$) fitted to the gradient of the average m/z = 91.05 (PS) intensity, m/z = 112.08 (PVP) intensity and scores for PC1 in the z-direction, seen in Figure 3.15, to allow direct comparison between ion images and scores images. The depth resolutions Δz are calculated as $2\sigma$ and there is no significant difference between those calculated with the ion images and those calculated with the scores as shown by a pairwise t-test (P = 0.31).

| Interface | m/z 91.05 (PS) or m/z 112.08 (PVP) | | | | PC1 scores | | | |
|---|---|---|---|---|---|---|---|---|
| | $a_i$ | $b_i$ (nm) | $c_i$ (nm$^2$) | Δz (nm) | $a_i$ | $b_i$ (nm) | $c_i$ (nm$^2$) | Δz (nm) |
| 1 (PVP) | 0.003996 | 1279 | 5.083 | 3.19 | -1.346 | 1279 | 5.132 | 3.20 |
| 2 (PS) | 0.010280 | 1080 | 7.087 | 3.76 | 0.9903 | 1079 | 7.181 | 3.79 |
| 3 (PVP) | 0.003839 | 1029 | 5.400 | 3.29 | -1.290 | 1029 | 5.540 | 3.33 |
| 4 (PS) | 0.006753 | 810.3 | 10.90 | 4.67 | 0.6732 | 809.1 | 10.56 | 4.60 |
| 5 (PVP) | 0.002134 | 764.6 | 9.882 | 4.45 | -0.6745 | 764.6 | 10.69 | 4.62 |
| 6 (PS) | 0.005647 | 544.0 | 13.08 | 5.11 | 0.5918 | 542.6 | 12.20 | 4.94 |
| 7 (PVP) | 0.001923 | 500.5 | 10.65 | 4.62 | -0.5838 | 500.1 | 12.38 | 4.98 |
| 8 (PS) | 0.006661 | 275.6 | 10.58 | 4.60 | 0.6717 | 274.6 | 10.48 | 4.58 |
| 9 (PVP) | 0.002146 | 228.3 | 9.854 | 4.44 | -0.6844 | 228.6 | 10.58 | 4.60 |

APPENDIX B

# Supplementary Data to Chapter 4

## B.1 Matlab Code for Interferometry Image Data Processing

Below is the Matlab code written for the processing of interferometry image data obtained from the Fogale Nanotech Photomap 3D, a coherence scanning interferometer (see section 1.6.3) with a 5x/0.12 objective and a camera size of 763 $\times$ 573 (9.4 $\mu$m pixels). The raw data files were exported in a ASCII text file containing XYZ triplet data (in $\mu$m). The data is imported into Matlab (Release 2013a, The MathWorks, Inc., Natick, Massachusetts, United States) as three [437199 $\times$ 1] columns.

The image is leveled using a baseline correction that consists of fitting a polynomial equation which is then subtracted from the image. Next, the total roughness $R_t$ and the root-mean-square roughness $R_q$ are calculated (see section 1.6.3). Finally, the array is reshaped as a matrix. The $R_{t,av}$ of the image is calculated as the average of the $R_t$ values for each individual y-profile line in the image. Finally, a surface heightmap is plotted.

```
% Surface plot of XYZ triplets

% Input data
disp('Surface plot of XYZ triplets');
X =  input('Enter X (n x 1) : ');
Y =  input('Enter Y (m x 1) : ');
Z =  input('Enter Z (nm x 1) : ');
filename =  input('Enter filename : ');

% Baseline correction
f = fit([X,Y], Z,'poly33','Exclude', Z > 1);
Zn=Z-f([X,Y]);

% Roughness parameters
```

```
d = Zn-mean(Zn);
Rv = min(d(:));
Rp = max(d(:));
Rt = Rp - Rv
Rq = sqrt(mean(d.^2))

clearvars d Rv Rp

% Reshape data
Zr=rot90(reshape(Zn,763,573));

% Average Rt per y profile
for k = 1:size(Zr,2)
Rp(k)=max(Zr(:,k));
Rv(k)=min(Zr(:,k));
end
Rp_av = mean(Rp);
Rv_av = mean(Rv);
Rt_av = Rp_av - Rv_av

clearvars Rp Rv k

% Create figure
h=figure;

% Plot
clims = [Rv_av, Rp_av];
imagesc([min(X), max(X)],[min(Y), max(Y)],Zr,clims)

colormap(hot)
colorbar EastOutside
axis xy image

xlabel('x (\mum)')
ylabel('y (\mum)')
ylabel(colorbar,'z (\mum)')

str=sprintf('R_{t} = %f and R_{q} = %f', Rt,Rq );
title(str);

% Save figure
str1=sprintf('IF_%s', filename);
str2 = strrep(str1, '.', '_');
saveas(h,str2,'fig')

% Close figure
```

```
close(h)

clearvars X Y Z Zn f filename h str str1 str2 clims ans Rp_av Rv_av
```

## B.2 Matlab Code for Aligning Images

This code aligns two images by rotation and scale change after manually picking corresponding points. It is used to align correlated microscopy, interferometry and ToF-SIMS images.

The original and distorted images are first imported as grayscale images or transformed to grayscale images (mat2gray). Matlab's Control Point Selection Tool (cpselect) is then used to pick pairs of control points, which are saved to the workspace. A similarity transformation is fitted to the control points. Next, the scale and angle are recovered from the geometric transformation. The distorted image is then transformed to align with the original image. The mean-squared error between the two images is calculated to provide a quantitative measure of alignment quality.

```
% show point matches
figure;
showMatchedFeatures(original,distorted,base_points,input_points);
title('Putatively matched points (including outliers)');

% estimate transformation
[tform, inlierDistorted, inlierOriginal] =
estimateGeometricTransform(input_points,base_points, 'similarity');

% display matching point pairs used in the computation of the
% transformation matrix
figure;
showMatchedFeatures(original,distorted, inlierOriginal, inlierDistorted);
title('Matching points (inliers only)');
legend('ptsOriginal','ptsDistorted');

% solve for Scale and Angle
Tinv  = tform.invert.T;
ss = Tinv(2,1);
sc = Tinv(1,1);
scale_recovered = sqrt(ss*ss + sc*sc)
theta_recovered = atan2(ss,sc)*180/pi

% recover the original image by transforming the distorted image
outputView = imref2d(size(original));
recovered  = imwarp(distorted,tform,'OutputView',outputView);

% compare recovered to original
figure, imshowpair(original,recovered,'montage')
```

```
30
31 % image MSE
32 MSE = mean((original(:)-recovered(:)).^2)
```

## B.3 Matlab Code for Z-Offset Correction

Based on a vector containing the scores of a principal component differentiating between substrate and the material on the substrate, it is possible to calculate the z-offset correction for ToF-SIMS images with topography. First, the scores matrix is converted into a matrix [pixels $\times$ scans] where each column is a xy-pixel's z-profile of the scores' values. Next, the first row number (i.e. scan number) for which the score value $< 0$ is determined for each xy-pixel's z-profile. If there is none such row, the maximum number of scans is assigned instead. Then, a matrix with new scan numbers, where the interface is defined as scan 0, is created and reshaped into scan resolved images. Finally, a heightmap of the surface topography is produced with z the number of scans removed from the interface.

```
% z-offset correction based on PC scores

% total number of scans
SN = size(normalised_SCORES,1);

% create matrix where each column is a xy-pixel's z-profile
for k = 1:SN
    pp(:,k) = normalised_SCORES{k,1}(:);
end

clearvars k

% find row number of first score value < 0 for each pixel
% if there is none, assign max number of scans instead
% create matrix with new scan numbers (interface = 0)
for k = 1:size(pp,1)
    zero = find(pp(k,:) < 0,1,'first');
    TF = isempty(zero);
    if TF == 1
       zero = SN;
    end
    Z(k,:)= linspace(zero-1,-SN+zero,SN);
end

clearvars k zero TF

% reshape Z-matrix into scan resolved images
Zr=cell(SN,1);

```

```
for l= 1:SN
    Zr{l,1} = reshape(Z(:,l),size(normalised_SCORES{1,1},1),
                              size(normalised_SCORES{1,1},2));
end

clearvars l Z SN pp

% create heightmap
h = figure;
imagesc([0, IMAGECHAR{2,1}(1)],[0, IMAGECHAR{2,1}(2)],Zr{1,1});
colormap(hot);
colorbar;
axis xy image;
xlabel('x (\mum)');
ylabel('y (\mum)');
ylabel(colorbar, 'z (number of scans removed from the interface)');
title('Surface Heightmap in Number of Scans');

% save z-file
save zfile_pc1.mat Zr -v7.3
```

## B.4 Matlab Code for Determining the Sputter Yield

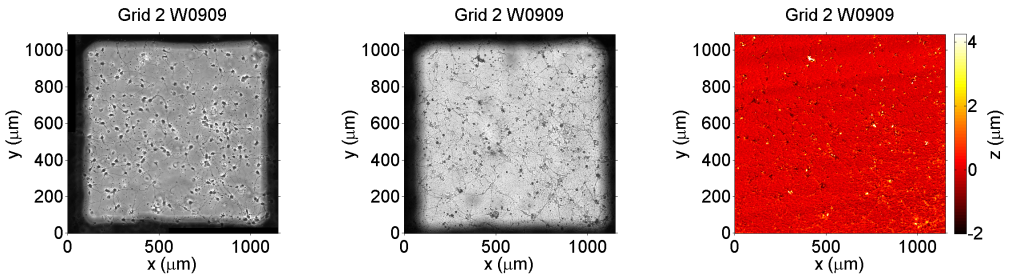
If one possesses a scores based z-offset image and a corresponding interferometry heightmap, the z-offset correction can be calibrated by fitting the scores based z-offset image to the interferometry heightmap using a linear function. The pixel values of the scores based z-offset image represent the number of scans needed to reach the substrate interface. Therefore, the slope of the linear function represents the sputter yield (nm/scan) and the intercept represents the layer thickness (nm) of the lowest point in the interferometry heightmap. It is important to note this assumes a constant sputter yield through the cellular material. If necessary, the interferometry image is first transformed to obtain an aligned (i.e. rotated and scaled) image with the same size as the scores based z-offset image using feature matching (see Section B.2).

The interferometry and z-offset corrected heightmap are transformed into vectors before applying a first degree fit. Then the standard deviation of the coefficients and the $R^2$ of the linear fit are calculated. The scaled z-offset correction is calculated and reshaped into a matrix. The mean squared error (MSE) is then calculated in order to calculate the peak signal-to-noise ratio (PSNR) between the scaled z-offset heightmap and the interferometry heightmap. Finally, both images are plotted next to each other for comparison.

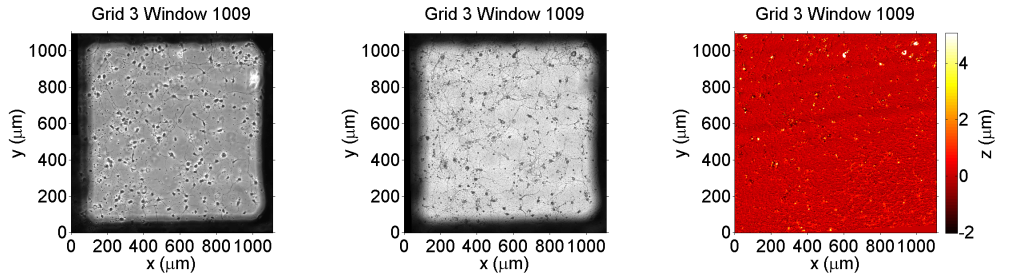
```
% fitting sputterspeed and offset

%load data
H = IF_trans(:);
N = Zr_PC1{1,1}(:);

% least-squares polynomial fit of degree 1
[p,s] = polyfit(N,H,1);

% calculate standard deviation of coefficients and R2
SD = sqrt(diag(inv(s.R)*inv(s.R'))./s.normr.^2./s.df);
R2=1 - s.normr^2 / norm(H-mean(H))^2;

clearvars s

% form scaled image
E1 = (N*p(1)) +p(2);
E2 = reshape(E1,256,256);

% calculate MSE between images
```

```
21 MSE = mean((E1-H).^2);
22
23 % calculate PSNR between images
24 PSNR = log ((max(IF_trans(:))^2) / (MSE))*10;
25
26 clearvars E1 H N
27
28 % compare scaled to original heightmap
29 h=figure;
30 h1 = subplot(1,2,1);
31 imagesc(IF_trans)
32 axis xy image
33 colormap(hot)
34 colorbar
35 caxis([min(E2(:)) max(E2(:))])
36 xlabel('x (\mum)','FontSize',20)
37 ylabel('y (\mum)','FontSize',20)
38 ylabel(colorbar,'z (\mum)','FontSize',20)
39 title('Interferometry (original)','FontSize',20)
40 h2 = subplot(1,2,2);
41 imagesc(E2)
42 axis xy image
43 colormap(hot)
44 colorbar
45 xlabel('x (\mum)','FontSize',20)
46 ylabel('y (\mum)','FontSize',20)
47 ylabel(colorbar,'z (\mum)','FontSize',20)
48 title('Z-corrected PC1 (scaled)','FontSize',20)
49 linkaxes([h2,h1])
50
51 saveas(h,'scaling.fig')
52
53 clearvars h h1 h2 E2
```

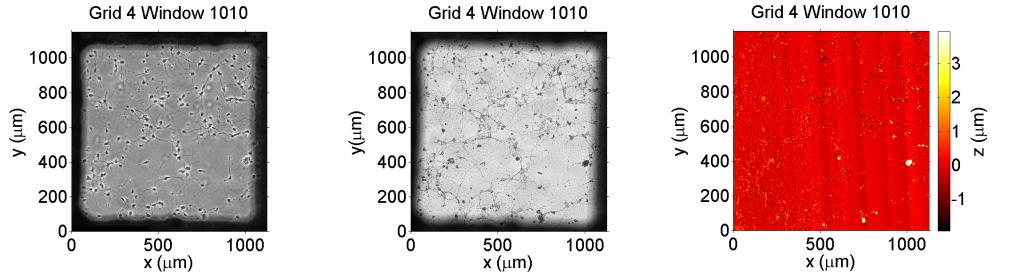
## B.5 Supplementary Figures




**(a)** DIC microscopy image (20x) of neuronal cells. **(b)** Bright-field microscopy image (20x) of freeze-dried neuronal cells. **(c)** White light interferometry heightmap of the freeze-dried neuronal network ($R_{t,av}$ = 6.25 $\mu$m, $R_q$ = 0.52 $\mu$m).

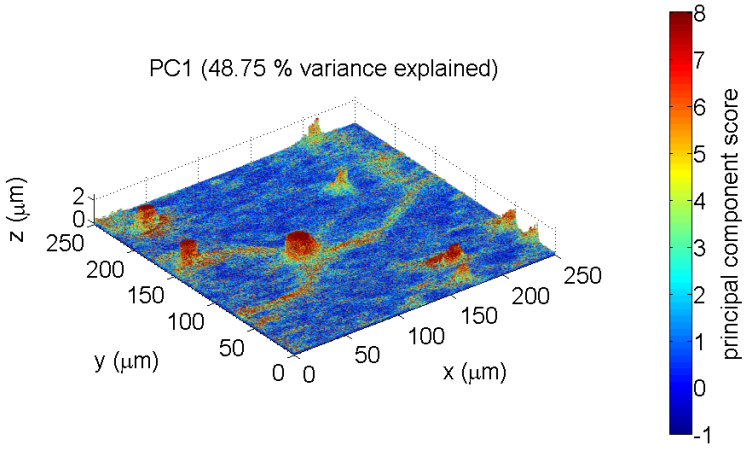



**(d)** DIC microscopy image (20x) of neuronal cells. **(e)** Bright-field microscopy image (20x) of freeze-dried neuronal cells. **(f)** White light interferometry heightmap of the freeze-dried neuronal network ($R_{t,av}$ = 7.10 $\mu$m, $R_q$ = 0.58 $\mu$m).

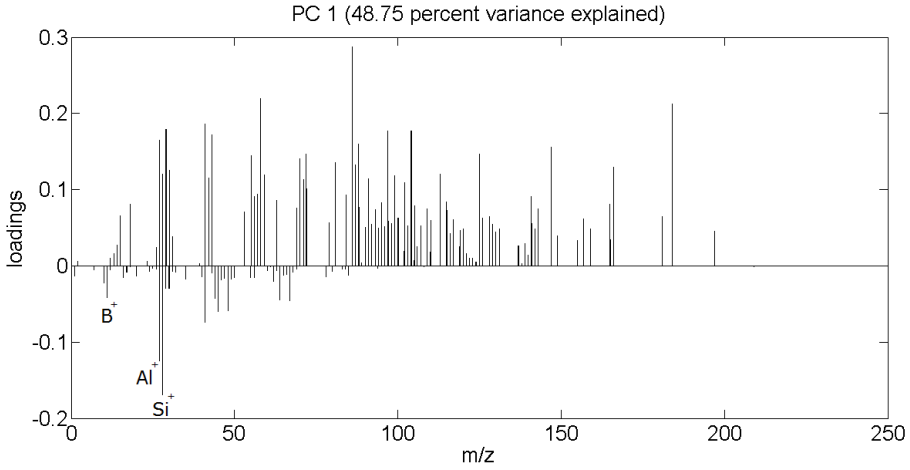



**(g)** DIC microscopy image (20x) of neuronal cells. **(h)** Bright-field microscopy image (20x) of freeze-dried neuronal cells. **(i)** White light interferometry heightmap of the freeze-dried neuronal network ($R_{t,av}$ = 5.88 $\mu$m, $R_q$ = 0.51 $\mu$m).

**Figure B.1:** Correlative imaging of neuronal cell networks before and after freeze-drying.

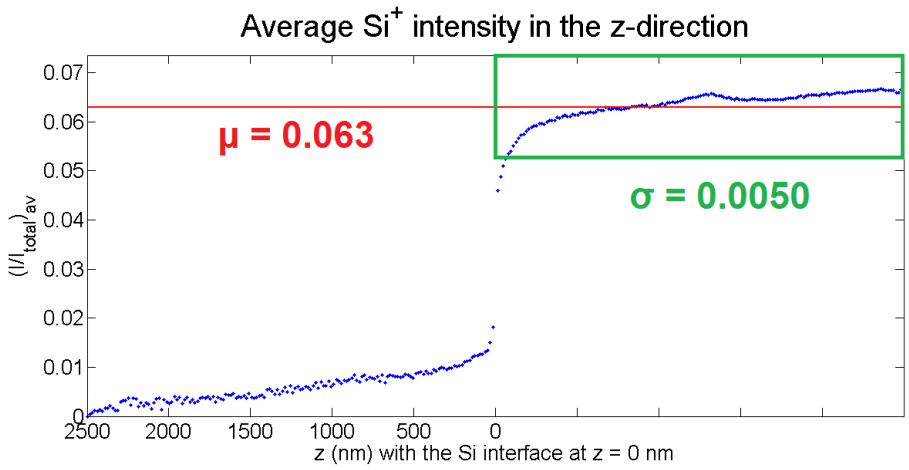



**(a)** 3D scores image.

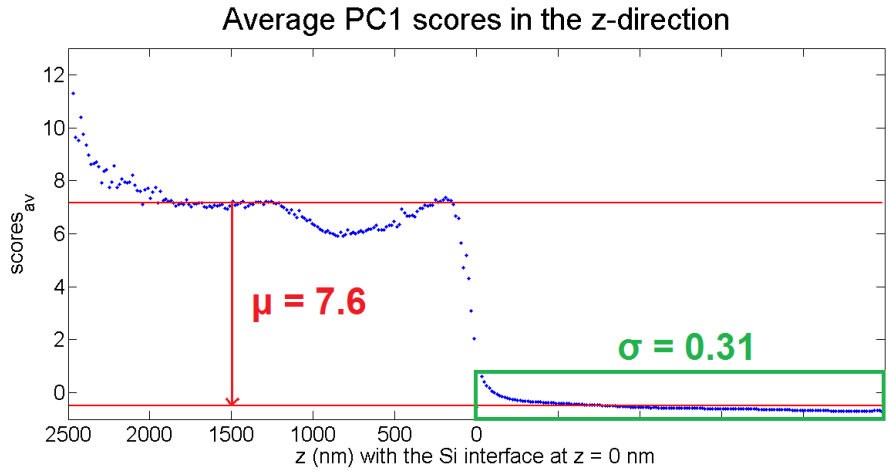



**(b)** Loadings plot.

**Figure B.2:** PCA of the neuronal cell network: PC1 explains 48.8% of the variance. The positive loadings of PC1 correspond to fragments associated with bio-organic material and the negative loadings correspond to inorganic ions specific for the borosilicate glass substrate such as $B^+$ (m/z 11), $Al^+$ (m/z 27) and $Si^+$ (m/z 28).




**(a)** Calculation of the SNR for the average $Si^+$ intensity in the z-direction. The SNR equals 13.




**(b)** Calculation of the SNR for the average PC1 scores in the z-direction. The SNR equals 25, which is 1.9 times that of the average $Si^+$ intensity.

**Figure B.3:** Using a linear combination of ion intensities (i.e. the PCA loadings) instead of a single ion has the advantage of increased SNR, leading to an improved z-correction. As a result, a heightmap that resembles the optical image more closely is obtained (see Figure 4.5).

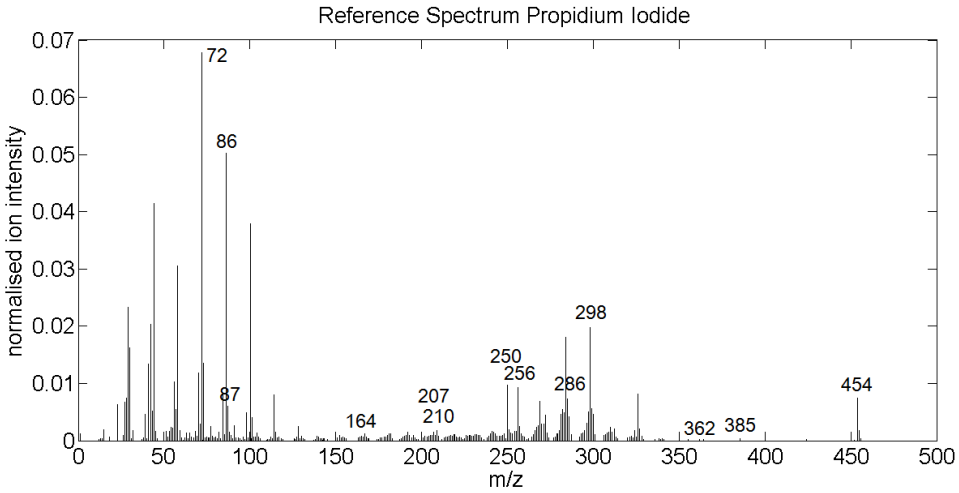



**Figure B.4:** The normalised mass spectrum in the positive polarity of propidium iodide. The m/z labelled peaks correspond to $C_4H_{10}N^+$ (m/z 72), $C_5H_{12}N^+$ (m/z 86), $C_5H_{13}N^{+\cdot}$ (m/z 87), $[C_8H_{19}N.Cl]^+$ (m/z 164), $C_{27}H_{34}N_4{}^{2+}$ (m/z 207), $C_{13}H_{12}N_3{}^+$ (m/z 210), $C_{16}H_{16}N_3{}^+$ (m/z 250), $[C_8H_{19}N.I]^+$ (m/z 256), $C_{19}H_{16}N_3{}^+$ (m/z 286), $C_{20}H_{17}N_3{}^+$ (m/z 298), $[C_{22}H_{21}N_3.Cl]^+$ (m/z 362), $C_{25}H_{29}N_4{}^+$ (m/z 385) and $[C_{22}H_{21}N_3.I]^+$ (m/z 454).

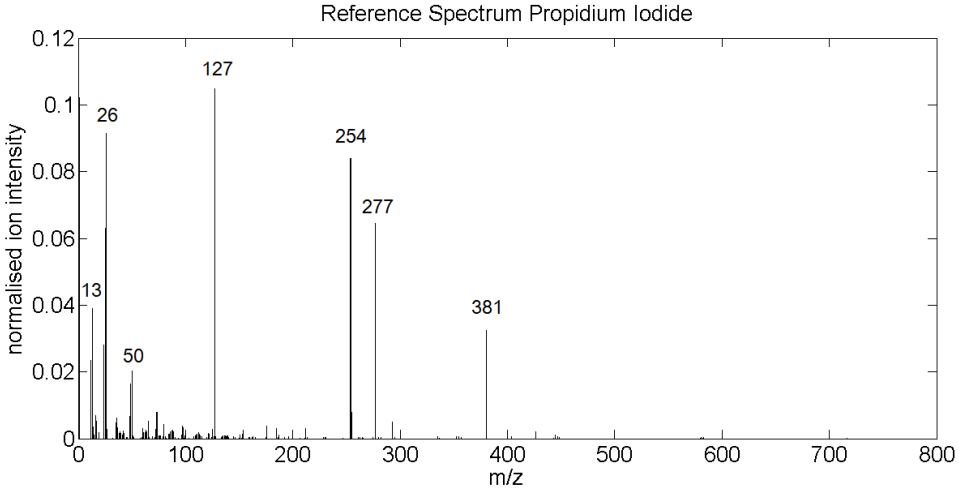



**Figure B.5:** The normalised mass spectrum in the negative polarity of propidium iodide. The m/z labelled peaks correspond to $CH^-$ (m/z 13), $CN^-$ (m/z 26), $C_3N^-$ (m/z 50), $I^-$ (m/z 127), $I_2{}^-$ (m/z 254), $NaI_2{}^-$ (m/z 277), and $I_3{}^-$ (m/z 381).

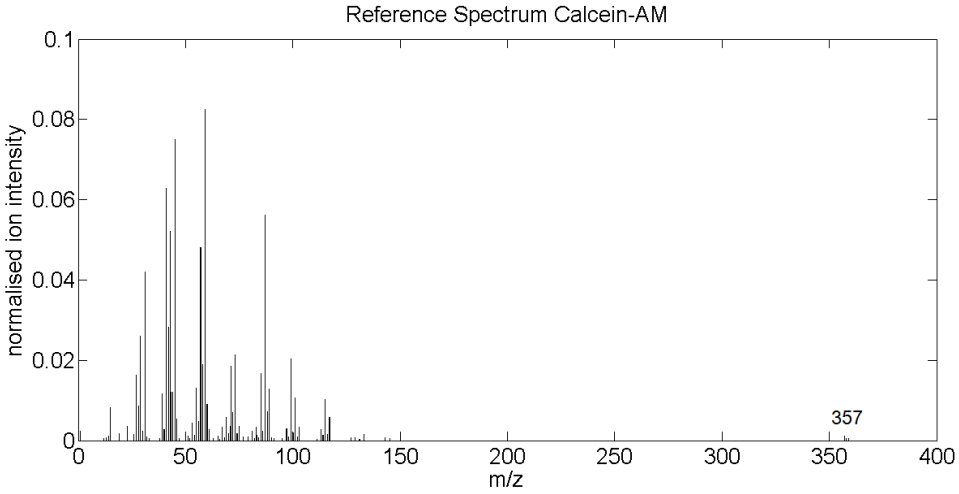



**Figure B.6:** The normalised mass spectrum in the positive polarity of calcein-AM. The m/z labeled peak corresponds to $C_{21}H_9O_6^+$ (m/z 357).

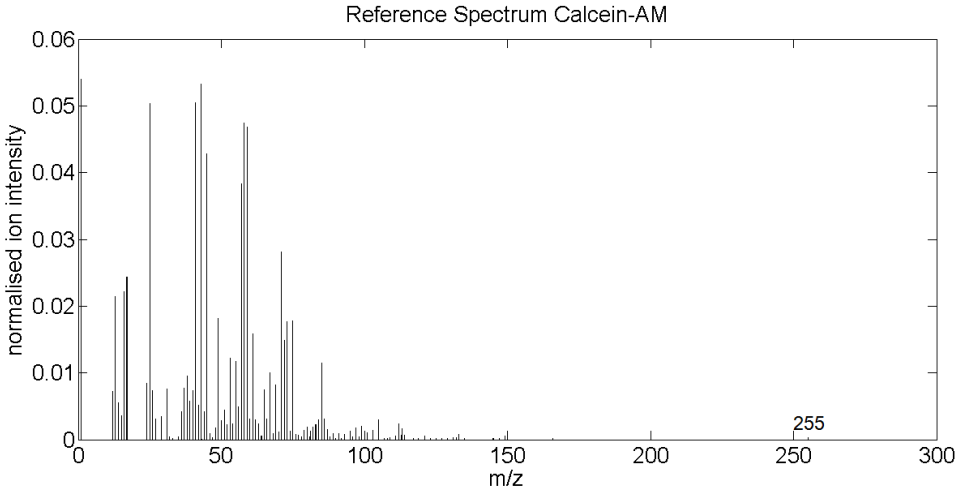
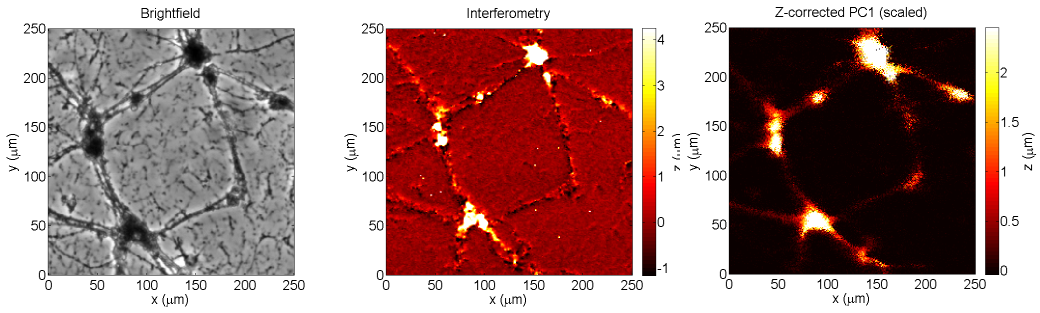
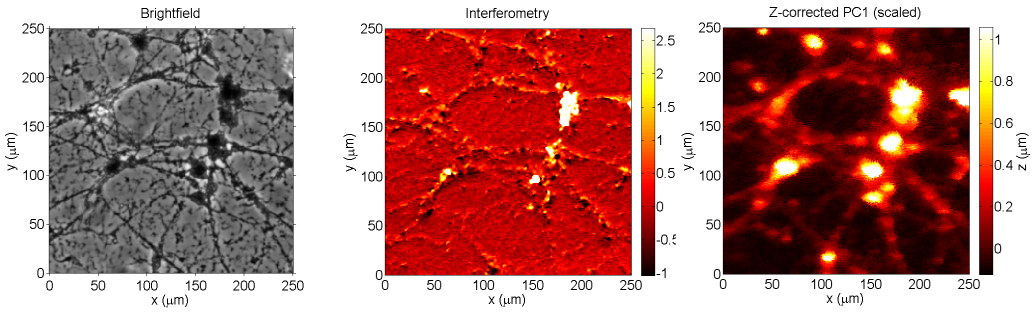



**Figure B.7:** The normalised mass spectrum in the negative polarity of calcein-AM.




**(a)** Bright-field. **(b)** Interferometry. **(c)** Heightmap based on PC1.

**Figure B.8:** Scaling of the z-offset corrected heightmap based on PC1 of stained, neuronal cells investigated with ToF-SIMS in the positive polarity. A) Bright-field microscopy image (10×) of freeze-dried neuronal cells stained with calcein-AM and propidium iodide. Several neurons can be seen as well as their axons and dendrites. B) White light interferometry heightmap ($R_q$ = 1.05, $R_{t,av}$ = 5.4) of the part of the neuronal network investigated with ToF-SIMS. C) Scaled z-offset corrected heightmap ($R_q$ = 0.64, $R_{t,av}$ = 2.50) based on the scores of principal component 1, which differentiates the borosilicate glass substrate from cellular material, by fitting the z-offset image to the interferometry heightmap. The result is a heightmap very similar to that obtained with the interferometer so that the peak signal-to-noise ratio (PSNR) equals 63 dB. The calculated sputter yield equals 21.2 nm/scan and the layer thickness of the lowest point in the interferometry heightmap equals 44.1 nm.




**(a)** Bright-field. **(b)** Interferometry. **(c)** Heightmap based on PC1.

**Figure B.9:** Scaling of the z-offset corrected heightmap based on PC1 of stained, neuronal cells investigated with ToF-SIMS in the negative polarity. A) Bright-field microscopy image (10×) of freeze-dried neuronal cells stained with calcein-AM and propidium iodide. Several neurons can be seen as well as their axons and dendrites. B) White light interferometry heightmap ($R_q$ = 0.69, $R_{t,av}$ = 3.72) of the part of the neuronal network investigated with ToF-SIMS. C) Scaled z-offset corrected heightmap ($R_q$ = 0.27, $R_{t,av}$ = 1.18) based on the scores of principal component 1, which differentiates the borosilicate glass substrate from cellular material, by fitting the z-offset image to the interferometry heightmap. The result is a heightmap very similar to that obtained with the interferometer so that the peak signal-to-noise ratio (PSNR) equals 64 dB. The calculated sputter yield equals 7.0 nm/scan and the layer thickness of the lowest point in the interferometry heightmap equals 206.2 nm.

APPENDIX C

# Supplementary Data to Chapter 5

## C.1 Supplementary Tables

**Table C.1:** Identified mass peaks for cellular material (lipids and amino acids) used to select relevant mass peaks for comparison between the mass spectra of freeze-dried and frozen-hydrated cellular material. The nominal masses (nom. m/z), molecular formulas and their assignments were obtained from literature [112, 124, 131]. Three-letter abbreviations are used for amino acids and PC indicates phosphatidylcholine. Furthermore, the observed m/z (obs. m/z), mass accuracy (mass acc.) and normalised ion intensity per pixel (norm. ion int. per pixel) are given for both freeze-dried and frozen-hydrated cellular material. The normalised ion intensities per pixel are the ones graphically presented in Figure 5.5.

| nom. m/z | Formula | Assignment | Frozen-hydrated (PC1 > 0, PC2 < 0) | | | Freeze-dried (PC1 > 0) | | |
|---|---|---|---|---|---|---|---|---|
| | | | obs. m/z | mass acc. (ppm) | norm. ion int. per pixel (x $10^5$) | obs. m/z | mass acc. (ppm) | norm. ion int. per pixel (x $10^5$) (n = 3) |
| 15 | $CH_3^+$ | PC | 15.02 | -193 | 0.81 | 15.02 | -193 | 0.73 ± 0.23 |
| 27 | $C_2H_3^+$ | PC/proteins | 27.02 | -107 | 3.04 | 27.02 | -107 | 2.96 ± 0.87 |
| 29 | $C_2H_5^+$ | PC | 29.04 | 48 | 2.06 | 29.04 | 48 | 2.94 ± 0.77 |
| 30 | $CH_4N^+$ | PC/Gly | 30.03 | -127 | 5.86 | 30.04 | 206 | 2.11 ± 0.81 |
| 41 | $C_3H_5^+$ | PC/proteins | | | | 41.04 | 34 | 4.27 ± 1.54 |
| 43 | $C_3H_7^+$ | PC/proteins | | | | 43.05 | -98 | 3.57 ± 1.00 |
| 44 | $C_2H_6N^+$ | PC/Ala/Cys | 44.04 | -216 | 5.23 | 44.05 | 11 | 1.73 ± 1.65 |
| 53 | $C_4H_5^+$ | PC | 53.03 | -162 | 1.66 | 53.04 | 26 | 1.76 ± 0.39 |
| 55 | $C_3H_3O^+$ | Tyr | | | | 55.02 | 40 | 0.44 ± 0.77 |
| 55 | $C_4H_7^+$ | PC/proteins | | | | 55.06 | 105 | 3.36 ± 0.96 |
| 56 | $C_3H_6N^+$ | PC/proteins | | | | 56.05 | 9 | 2.37 ± 0.58 |
| 57 | $C_4H_9^+$ | PC/proteins | | | | 57.07 | 2 | 2.63 ± 0.74 |
| 58 | $C_3H_8N^+$ | PC | 58.06 | -88 | 2.59 | 58.07 | 84 | 4.53 ± 1.69 |
| 59 | $CH_5N_3^+$ | Arg | | | | 59.04 | -132 | 0.30 ± 0.52 |
| 59 | $C_3H_9N^+$ | PC | | | | 59.07 | -51 | 2.31 ± 0.91 |
| 60 | $C_2H_6NO^+$ | Ser | 60.05 | 93 | 1.63 | | | |
| 61 | $C_2H_5S^+$ | Met | 61.01 | -10 | 0.34 | | | |
| 67 | $C_5H_7^+$ | PC | 67.04 | -212 | 1.20 | 67.06 | 86 | 1.17 ± 1.07 |
| 68 | $C_4H_6N^+$ | PC/Pro | 68.04 | -140 | 1.35 | 68.05 | 7 | 0.88 ± 0.80 |
| 69 | $C_5H_9^+$ | PC/proteins | | | | 69.07 | 1 | 1.95 ± 0.83 |
| 70 | $C_4H_8N^+$ | PC/Arg/Leu/Pro | 70.06 | -73 | 6.86 | 70.07 | 70 | 2.92 ± 1.04 |
| 72 | $C_3H_6NO^+$ | Gly | 72.05 | 78 | 20.51 | | | |
| 72 | $C_4H_{10}N^+$ | PC/Val | | | | 72.08 | -11 | 1.78 ± 0.46 |
| 74 | $C_3H_8NO^+$ | Thr | 74.05 | -135 | 1.42 | | | |
| 81 | $C_6H_9^+$ | PC | 81.05 | -245 | 3.04 | 81.07 | 1 | 1.01 ± 0.92 |
| 83 | $C_6H_{11}^+$ | PC | | | | 83.09 | 54 | 0.36 ± 0.35 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **84** | $C_5H_{10}N^+$ | PC/Lys | 84.07 | -128 | 5.69 | 84.09 | 109 | 0.69 ± 0.34 |
| **85** | $C_6H_{13}^+$ | PC | | | | 85.09 | -132 | 0.24 ± 0.24 |
| **86** | $C_5H_{12}N^+$ | PC/Leu/Ile | 86.1 | 42 | 3.56 | 86.11 | 158 | 4.81 ± 2.97 |
| **88** | $C_5H_{14}N^+$ | PC | | | | 88.11 | -24 | 0.31 ± 0.20 |
| **91** | $C_7H_7^+$ | PC/proteins | 91.04 | -156 | 13.04 | 91.05 | -46 | 0.85 ± 0.38 |
| **93** | $C_7H_9^+$ | PC | 93.05 | -214 | 0.56 | 93.07 | 1 | 0.31 ± 0.14 |
| **95** | $C_7H_{11}^+$ | PC | | | | 95.09 | 47 | 0.46 ± 0.27 |
| **97** | $C_7H_{13}^+$ | PC | | | | 97.08 | -218 | 0.27 ± 0.13 |
| **98** | $C_5H_8NO^+$ | PC | 98.08 | 204 | 0.71 | 98.08 | 204 | 0.18 ± 0.10 |
| **100** | $C_4H_{10}N_3^+$ | Arg | | | | 100.09 | 25 | 0.23 ± 0.12 |
| **102** | $C_5H_{12}NO^+$ | PC | | | | 102.09 | -13 | 0.30 ± 0.16 |
| **104** | $C_5H_{14}NO^+$ | PC | | | | 104.11 | 29 | 1.28 ± 0.96 |
| **110** | $C_5H_8N_3^+$ | Arg/His | | | | 110.08 | 79 | 0.20 ± 0.12 |
| **115** | $C_4H_7N_2O_2^+$ | Gly | | | | 115.04 | -89 | 0.26 ± 0.14 |
| **120** | $C_8H_{10}N^+$ | Phe | 120.06 | -173 | 1.52 | 120.08 | -7 | 0.18 ± 0.11 |
| **130** | $C_9H_8N^+$ | Trp | 130.06 | -39 | 0.65 | 130.07 | 38 | 0.15 ± 0.09 |
| **136** | $C_8H_{10}NO^+$ | Tyr | 136.06 | -115 | 3.84 | 136.09 | 105 | 0.06 ± 0.07 |
| **150** | $C_5H_{13}NPO_2^+$ | PC | 150.04 | -185 | 0.49 | 150.08 | 81 | 0.07 ± 0.08 |
| **166** | $C_5H_{13}NPO_3^+$ | PC | 166.06 | -17 | 0.73 | 166.08 | 104 | 0.25 ± 0.11 |
| **168** | $C_5H_{15}NPO_3^+$ | PC | 168.03 | -288 | 0.42 | 168.07 | -50 | 0.08 ± 0.08 |
| **182** | $C_5H_{13}NPO_4^+$ | PC | 182.05 | -42 | 0.59 | | | |
| **184** | $C_5H_{15}NPO_4^+$ | PC | 184.1 | 145 | 1.92 | 184.12 | 254 | 0.53 ± 0.21 |
| **198** | $C_6H_{17}NPO_4^+$ | PC | 198.09 | 5 | 3.29 | | | |
| **206** | $C_5H_{14}NPO_4Na^+$ | PC | 206.07 | 71 | 0.42 | 206.11 | 265 | 0.05 ± 0.05 |
| **224** | $C_8H_{19}NPO_4^+$ | PC | 224.08 | -110 | 0.59 | 224.14 | 158 | 0.04 ± 0.04 |
| **226** | $C_7H_{17}NPO_5^+$ | PC | 226.08 | -17 | 0.86 | | | |