Faraday Discussions

Cite this: *Faraday Discuss.*, 2016, **187**, 199

PAPER

Towards quantitative molecular mapping of cells by Raman microscopy: using AFM for decoupling molecular concentration and cell topography†

Radu Boitor,[a] Faris Sinjab,[a] Stephanie Strohbuecker,[b] Virginie Sottile[b] and Ioan Notingher*[a]

Received 16th November 2015, Accepted 26th November 2015
DOI: 10.1039/c5fd00172b

Raman micro-spectroscopy (RMS) is a non-invasive technique for imaging live cells *in vitro*. However, obtaining quantitative molecular information from Raman spectra is difficult because the intensity of a Raman band is proportional to the number of molecules in the sampled volume, which depends on the local molecular concentration and the thickness of the cell. In order to understand these effects, we combined RMS with atomic force microscopy (AFM), a technique that can measure accurately the thickness profile of the cells. Solution-based calibration models for RNA and albumin were developed to create quantitative maps of RNA and proteins in individual fixed cells. The maps were built by applying the solution-based calibration models, based on partial least squares fitting (PLS), on raster-scan Raman maps, after accounting for the local cell height obtained from the AFM. We found that concentrations of RNA in the cytoplasm of mouse neuroprogenitor stem cells (NSCs) were as high as 25 ± 6 mg ml^{-1}, while proteins were distributed more uniformly and reached concentrations as high as ~50 ± 12 mg ml^{-1}. The combined AFM–Raman datasets from fixed cells were also used to investigate potential improvements for normalization of Raman spectral maps. For all Raman maps of fixed cells (n = 10), we found a linear relationship between the scores corresponding to the first component (PC1) and the cell height profile obtained by AFM. We used PC1 scores to reconstruct the relative height profiles of independent cells (n = 10), and obtained correlation coefficients with AFM maps higher than 0.99. Using this normalization method, qualitative maps of RNA and protein were used to obtain concentrations for live NSCs. While this study demonstrates the potential of using AFM and RMS for measuring concentration maps for individual NSCs *in vitro*, further studies are required to establish the robustness of the normalization method based on principal component analysis when comparing Raman spectra of cells with large morphological differences.

[a]*School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD, UK. E-mail: ioan.notingher@nottingham.ac.uk*

[b]*Wolfson STEM Centre, School of Medicine, University of Nottingham, University Park, Nottingham NG7 2RD, UK*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5fd00172b

Introduction

Cells are the fundamental units of any living organism. Their functions rely on highly orchestrated, spatially and time-dependent molecular interactions that underpin all biological processes. However, because of their small size and fragile nature, non-invasive detection and quantification of molecular interactions in living cells has always been a challenge. Various techniques for quantitative molecular analysis of cells do exist (*e.g.* mass spectrometry,[1] polymerase chain reaction (PCR)[2]), yet they require destructive procedures and usually apply to whole cell populations instead of single cells. Therefore, these techniques provide only single time-points and cannot provide insight into discrete and dynamic molecular events in living cells. Similar limitations apply to microscopy techniques based on fluorescence labeling of cells. As most molecules of interest are found inside the cells, cell fixation and membrane permeabilisation are required in order to allow the fluorescent molecules, which often are large dye molecules, to enter the cell cytoplasm and attach to the molecules of interest. To circumvent this antibody-based detection approach, transgenic strategies to express markers such as green fluorescent protein (GFP) have been developed. However, these techniques cannot be readily applied to primary cells, and genetic manipulation of cells requires laborious protocols to be developed for each cell type. Another limitation common to all fluorescence imaging methods is the difficulty in quantifying the results due to intensity variations of the fluorescence emission caused by photobleaching as well as variations in staining protocols.

Raman micro-spectroscopy (RMS) is a well-established analytical technique based on interaction of light with molecules present in the sample. Research during the last two decades showed that RMS can provide detailed molecular information on complex biological samples such as tissues and cells.[3] The main advances in this field have been recently reviewed.[4–6] An important feature of RMS is that it can measure time- and spatially resolved molecular processes in live cells maintained under physiological conditions *in vitro*, and without requiring labeling.[7–10]

Recently, attempts have been reported for quantitative molecular analysis of individual cells, as well as mapping the concentration of molecular components in live cells. The use of Raman spectral calibration models developed using samples with known concentrations (microparticles or solutions) has been reported for quantifying glycogen content in stem cell populations[11] and RNA in neuroprogenitor stem cells.[12] However, one main challenge when attempting such quantitative analysis arises from the fact that the intensity of a Raman band is proportional to the number of molecules in the sampled volume, which depends on the local concentration and the local thickness of the cell, as well as the laser focusing conditions. Furthermore, most biological processes are also accompanied by morphological changes of the cells. Therefore, it is often difficult to extract and quantify the molecular concentration in live cells because the 3D morphology of the cells is not known (Fig. 1).

When attempting to discriminate between two groups of cells based on their Raman spectra, it is common to normalize the Raman spectra, either using the intensity of certain bands that are expected to remain constant (*e.g.* 1450 cm^{-1} [13] or 1004 cm^{-1} [14]), the total area under the baseline-corrected spectrum,[15,16] or

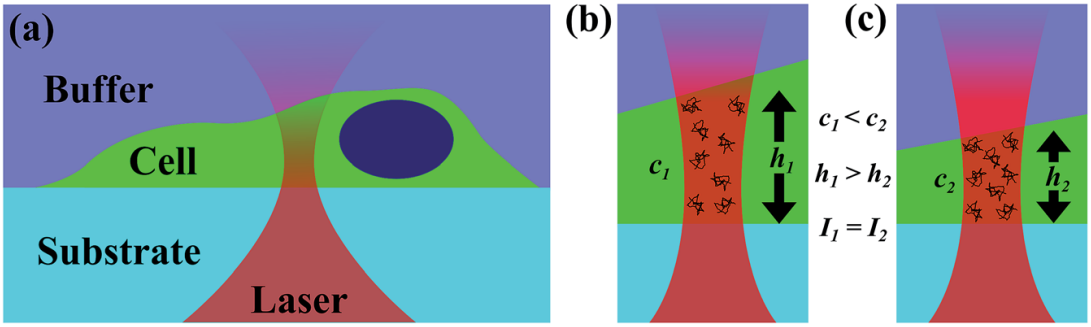

Fig. 1 Coupling between concentration and cell thickness in Raman micro-spectroscopy measurements. A schematic of the Raman sampling geometry for cells is shown in (a). The situations shown in (b) and (c) will give the same Raman scattered intensity, despite different molecular concentrations.

using the standard normal variance method.[17–19] Such normalization procedures have also been combined with solution-based calibration models to obtain molecular maps (*e.g.* RNA) in individual cells.[12] However, no studies on the validity of these normalization methods have been reported to date, thus the ability of these techniques to truly account for the morphology of the cells remains unclear.

In this paper we used confocal Raman micro-spectroscopy in conjunction with atomic force microscopy (AFM) in order to separate the effects of cell morphology and molecular concentration in the measured Raman spectra of cells. AFM was used to measure the topography maps of each individual cell, and this information was used to correct the intensity of the Raman spectra. Although in principle the topography of the cell could be obtained by recording a 3D map using a confocal Raman microscope, such measurements are time-consuming, and are prone to errors due to uncertainties related to the sampling volume of the microscope (variations in refractive index within the cell and the interfaces at the cell surface). Here, we use solution-based calibration models for proteins and nucleic acids, and combined AFM–Raman mapping to obtain quantitative molecular concentration maps for fixed cells (neuroprogenitor stem cells, NSCs). We also observed that principal component analysis (PCA) may be used for normalization of Raman spectral maps to take into account variations in cell thickness. This normalization method may be particularly useful for normalizing the Raman spectra of live cells, since AFM measurements are typically impractical because of long acquisition times for images and cell motility.

Materials and methods

Cell culture

Cell culture reagents were from Life Technologies (UK) unless otherwise stated. Mouse neural stem cells were cultured as previously described.[12] Briefly, cells were maintained in NSC medium made of DMEM/F12 and Neurobasal medium (1 : 1), N2, B27, antibiotics and supplemented with bFGF and EGF (Sigma-Aldrich, UK) at 20 ng ml^{-1} each. Cells were seeded onto low Matrigel-coated chambers (Becton Dickinson, UK) at 10^5 cells per ml using Accutase followed by a PBS wash. Cells were fixed with 4% paraformaldehyde (PFA, VWR) where indicated.

Atomic force microscopy

The cells that were fixed were washed in phosphate buffered saline (PBS) in order to remove cell debris caused by fixation with PFA. AFM images were acquired with a Nanowizard II AFM (JPK, Germany) that was mounted on top of an inverted Olympus IX-71 microscope. The measurements were performed in tapping mode to minimize the possibility of damage to the cell membrane. Custom-made sample holders were built with a thin (0.17 mm) MgF_2 window on the bottom, that would easily allow access to the microscope objective, but that would also allow space for the AFM scan head from above. Liquid-mode AFM tips (Multi75GB-G, BudgetSensors) were used for all measurements.

Raman micro-spectroscopy

The sample holders containing the cells were moved to a home-built Raman instrument, which was based on another inverted Olympus IX-71 microscope with an Olympus water immersion 60× objective (NA 1.2, a droplet of water was placed between the objective and coverslip) and with an automated stepper-motor stage (Prior Scientific) used to raster scan the sample, using custom-built software. A 785 nm Ti:Sapphire laser beam was focused to a laser spot, which provided 230 mW of power at the sample. The Raman back-scattered light was directed to a spectrometer (Andor Shamrock 303i with iDus CCD) *via* a fiber connection to obtain the Raman spectra. The spectrometer was calibrated using a polystyrene sample using peaks in the fingerprint region. Each fixed cell was scanned for 20 to 60 minutes, depending on its size. For Raman mapping, the integration time for each spectrum was 3 seconds and the step size in X and Y directions was 1 μm. For the axial scans, Raman spectra were recorded at regular 1.1 μm z intervals.

Live cell Raman measurements were performed on the same setup but with different experimental parameters. In order to facilitate live cell measurements, the setup was also equipped with an environmental enclosure (Solent Scientific), which kept the cells under physiological conditions during the measurements (37 °C, 5% CO_2). The same 785 nm laser line was used as for fixed cells, but the power was decreased to 170 mW and the acquisition time was reduced to 1 s. These parameters have been shown to have a minimal impact on cell physiology, while at the same time providing usable Raman spectra.[12]

Raman calibration models

Quantification of the two investigated cellular components was performed with the use of partial least squares regression (PLS). In order to build a PLS calibration model from Raman spectra, 14 known concentration solutions of RNA (yeast tRNA, lyophilized powder, Sigma UK) and protein (bovine serum albumin, lyophilized powder, Sigma UK) were created. All solutions were created by dissolving the two analytes in high purity deionized water (resistivity of 18.2 MΩ cm). In order to ensure the best correlation between the calibration model and the spectra of fixed cells, the same experimental parameters were used as for the acquisition of spectra from fixed cells: laser power at the sample was 230 mW and the acquisition time for each spectrum was 3 seconds. The Raman spectra of the calibration samples were measured using 6 μm-thick spacers (Omni-cell spacers from Specac).

Data processing

All data processing was carried out using custom-built MATLAB functions. Cosmic ray spikes were automatically removed from each spectrum before processing. Singular-value decomposition (SVD) was used to remove noise from the spectra whilst preserving 80–90% of the spectral information. Principal component analysis (PCA) was also used for the quick identification of various cell components and for the removal of the background (composed of contributions from the substrate, objective and the medium surrounding the cell). For the z profile data, PC1 values at different relative z positions had a linear baseline subtracted and were offset to 0 minimum value before fitting a Gaussian curve. The top-hat function representing the AFM height, used for deconvolution with the PC1 z profile had a width specified to an accuracy of 0.1 μm. The background was linearly subtracted from each spectrum used in the PLS calibration and prediction, in order to limit the spectral features to the ones that are specific to each investigated cellular component. For these spectra, the background was identified automatically with the use of k-means clustering analysis.

For AFM image processing, the background was levelled by fitting a linear plane between 3 points outside the cells and subtracting it from the dataset. In order to correlate the AFM images with the Raman maps the higher resolution AFM images were transformed and cropped in an automated routine to align the AFM image with the frame of the Raman image. The resolution of the resulting AFM image was then reduced to match that of the Raman maps, allowing the data from each to be approximately matched point-for-point.

Results and discussion

Fig. 2 presents typical examples of AFM and Raman spectral images for a fixed NSC. Fig. 2(b) shows that the selected Raman spectra at various positions inside the cell represent the molecular composition of the NSC and indicate a high chemical heterogeneity. Raman bands associated to nucleic acids include backbone vibrations at 788 cm^{-1} (O–P–O) and 1095 cm^{-1} (PO_2–), while vibrations specific to nucleotides are detected at 782 cm^{-1} (thymine, cytosine and uracil) and

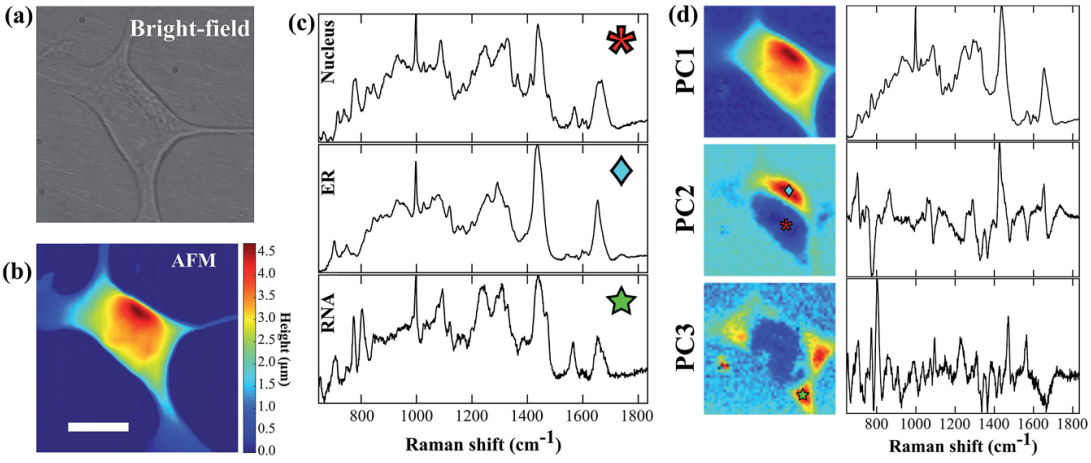

Fig. 2 Bright-field (a) and AFM (b) images of the same NSC (scale bar: 20 mm). Typical spectra from regions indicated by markers in PC score images (c) (SVD from PCs 2 : 10). (d) PCA images and loadings.

1578 cm^{-1} (guanine and adenine).[20,21] In agreement with previous reports, spectra with an intense band assigned to RNA (813 cm^{-1}, typically assigned to O–P–O vibrations in single-stranded nucleic acids) were observed at positions inside the cytoplasm of the NSC cells.[12] Raman bands corresponding to proteins are found in the ranges 1660–1670 cm^{-1} (amide I), 1450 cm^{-1} (C–H bending), 1200–1300 cm^{-1} (amide III), 1005 cm^{-1} (phenylalanine), 854 cm^{-1} (tyrosine) and 760 cm^{-1} (tryptophan).[22] All lipids are characterized by intense Raman bands at 1449 cm^{-1} (C–H bending vibrations), 1301 cm^{-1} (CH_2 twisting), 1000–1100 cm^{-1} spectral range (C–C stretching), while Raman bands characteristic to unsaturated lipids can be found at 1658 cm^{-1} (C=C stretching).[23] Membrane phospholipids also exhibit Raman bands in the 700–900 cm^{-1} spectral range assigned to different residues at the phosphate ester headgroup.[23] High-resolution Raman spectral images of cells can be created by plotting the scores obtained after principal component analysis (PCA) of all Raman spectra (Fig. 2(d)). These images can identify tentatively cellular structures, such as the nucleus (PC2 – negative scores), region rich in lipids and proteins surrounding the nucleus – potentially corresponding to the endoplasmatic reticulum (PC2 – positive scores), as well as the RNA rich regions at the edges of the cells (PC3). Indeed, bands corresponding to the molecular species found in these regions can be identified in the PCA loadings (Fig. 2(d)).

Nevertheless, PCA analysis can provide only qualitative information regarding the molecular composition of cells. To obtain quantitative information from the Raman spectra, calibration models based on water solutions of RNA and albumin (model for proteins) were developed (Raman spectra are presented in ESI Fig. S1†). To reduce the uncertainties related to the laser depth of field, the calibration solutions were placed in a sample chamber of 6 μm thickness in between two MgF_2 coverslips (Fig. 3(a)). As the thickness of the cells was 1–6 μm, the spectra recorded using this sample chamber are more representative for the Raman spectra of cells.

The PLS calibration models were built using fourteen RNA and albumin solutions with known concentrations spaced evenly over the expected concentrations in cells.[24] For the RNA solutions, the calibration model provided a good linear fit with an R^2 value of 0.961. Given the relatively low number of samples used for calibration, the prediction performance of the model was estimated by leave-one-out cross validation. When the first four components were used for the PLS model, the root mean squared error for cross-validation was RMSECV = 2.18 mg ml^{-1} (relative error 13%). Similarly, thirteen Raman spectra

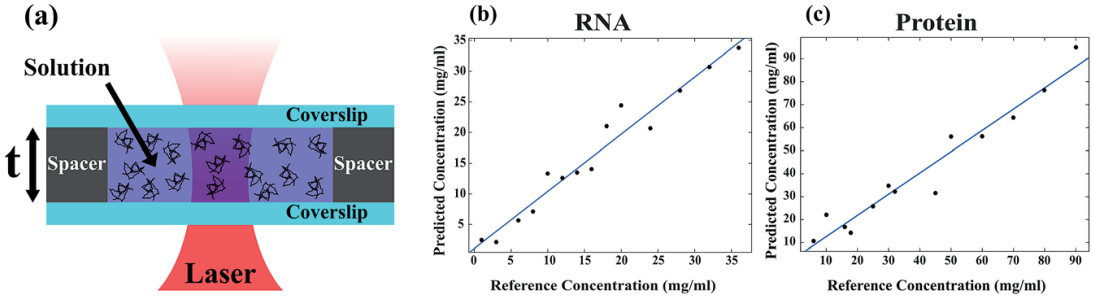

Fig. 3 PLS calibration of albumin and RNA. (a) Schematic of the experiment utilizing $t = 6$ μm spacers. Calibration curves for RNA (b) and albumin (c) in phosphate buffered saline solutions. The symbols represent the samples used for independent validation.

of albumin solutions were used for the protein model, producing an R^2 value for the calibration curve of 0.942 and a RMSECV of 6.29 mg ml^{-1} (relative error 15%).

Next, we applied the solution-based calibrations for RNA and albumin on 2D Raman maps of fixed cells in order to predict the concentration distributions of RNA and proteins. For each pixel in the Raman map (~1 μm × 1 μm), the value of cellular height obtained from the AFM map was used to calculate the local concentrations of RNA and proteins. Fig. 4 presents typical examples of quantitative maps for RNA and proteins for fixed NSCs. Individual Raman spectra from selected positions in the cells are presented in the ESI.†

At each pixel in the concentration maps in Fig. 4, the values for the RNA and protein concentrations were calculated by applying the calibration models described in Fig. 3 and then scaling using the cell thickness from the AFM map. For cell height values smaller than 1 μm in the AFM images, the corresponding Raman spectral intensity approached the limit of detection for the instrument. This resulted in the large values of concentration for the background, thus a threshold was applied based on the AFM height, allowing concentration maps only where cell spectra were detectable by the Raman instrument. When considering the accuracy of the calibration models (~15%) and errors in aligning the AFM and Raman maps, we estimate the errors in the concentration maps to be

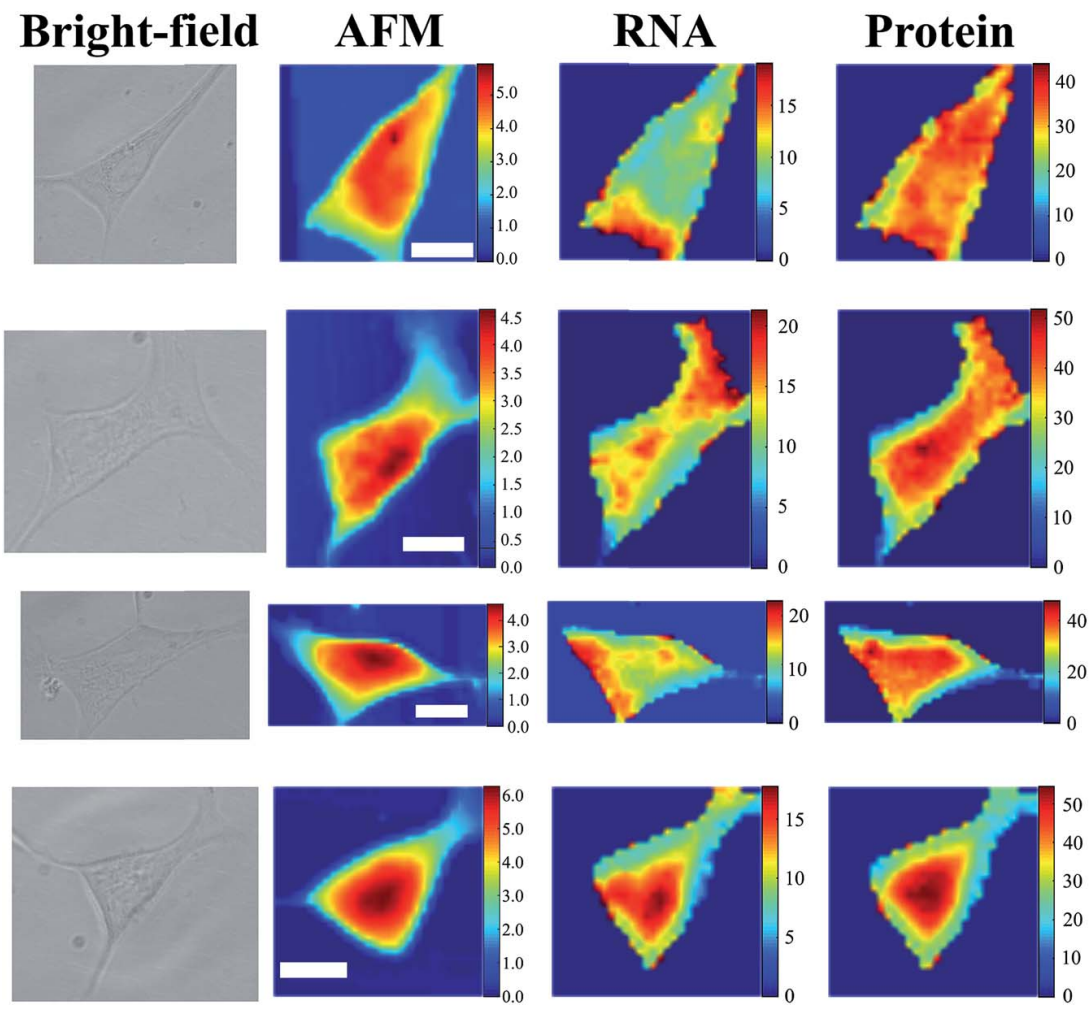

Fig. 4 Quantitative Raman mapping. Bright field images, transformed and matched AFM images (scale bar 10 μm, color scale in μm) with corresponding concentration maps for RNA and proteins (color scale in mg ml^{-1}) for three NSCs.

~25%. These errors may be reduced by increasing the number of samples in the calibration models and by using an integrated AFM–Raman instrument to reduce misalignment between AFM and Raman maps. Colocalized AFM–Raman measurements could eliminate the need for the alignment procedure in post-processing, providing more accurate results.

Fig. 4 shows that regions with high concentrations of RNA can be identified in the cytoplasm of NSCs, with maximum concentrations of typically 25 ± 6 mg ml^{-1} consistently between cells. We found that the concentration of protein can reach values as high as 50 ± 12 mg ml^{-1}, and was more widespread throughout the cell. The AFM and Raman spectral maps can also be used to quantify the total quantity of RNA and proteins in a single cell. As such, for typical NSC, the total quantity of RNA was found to be 12 pg (with values ranging from 11.1 pg to 13.1 pg), and 35 pg for proteins (values again in the range of 32.8–40.2 pg). The sensitivity level of Raman microscopy compares favorably with other single-cell analytical techniques, such as mass spectrometry and PCR. Real-time PCR (RT-PCR) is routinely used for quantification of microRNAs from cell cytoplasm, and achieves detection levels of ~15 pg of RNA.[25] Mass spectrometry (MS) is a very sensitive analysis technique that allows analysis of even single proteins.[26] However, a key challenge when analyzing cellular components is the preparation and isolation of these components. Detection of RNA was achieved by inductively coupled plasma mass spectrometry (ICP-MS), by quantifying the amount of $^{31}P^{+}$ in the investigated sample.[27] However, the reported limit of detection was ~60 pg, which is roughly 5 times larger than the expected quantity found within one cell. ICP-MS was also reported for quantification of proteins in cells, by quantifying the number of sulfur atoms.[28] However, the limit of detection was higher than 300 pg of proteins,[29] which is approximately 10 times higher than the values obtained for RMS. A major disadvantage of single-cell PCR and MS techniques is that they require cell lysis; thus, the cells are destroyed and cannot be used for subsequent analysis or imaging by other techniques.

Although the AFM allows accurate measurements of cell thickness, such information is often impractical to obtain for live cells because of the long acquisition times, during which the cell morphology can change. Thus, in the absence of AFM topography data, the absolute values of molecular concentrations cannot be obtained. However, Fig. 2 suggests that the image corresponding to the PC1 scores obtained from the PCA analysis of Raman maps resembles the height profile of the cell as measured by AFM (Fig. 2(b)). This similarity is not unexpected because the PC1 loading represents the mean Raman spectrum of the cell. Thus, at a given position in the cell, the PC1 score is expected to be proportional to the cell mass and height. However, this assumption is valid when the cell thickness is smaller than the axial resolution of the laser spot, and thus the laser excitation intensity can be considered constant across the sampled cell volume.

To investigate the relationship between PC1 scores and cell height (measured by AFM) in more depth, we analyzed the PC1 images from Raman maps of five fixed cells.

Fig. 5(a) compares the AFM height profiles of five fixed cells and the maps corresponding to PC1 scores. The calculated correlation coefficients for the AFM and PC1 maps for the individual cells was above 0.99, indicating that the PC1 scores from a Raman raster scan can provide an accurate relative height profile for the cells. Fig. 5(a) also shows that the map of PC1 scores is a more accurate

representation of the cell topography that the previously used intensity of the 1450 cm^{-1} band.[12] The use of the 1450 cm^{-1} map can be convenient to use for normalization because it corresponds mainly to CH_2 vibrations found in most biomolecules, but also it is an intense band for which the intensity can be calculated accurately. Nevertheless, Fig. 5 indicates that the images corresponding to the 1450 cm^{-1} over-represent the lipid-rich regions within the cells while under-representing the protein-rich regions. Normalization to the 1001 cm^{-1} band (Fig. 5) shows a better correlation to the AFM height map than the 1450 cm^{-1} normalization, however there is also significant noise from the background (errors in calculation of local baseline), which is avoided when using PC1. A regression analysis of the PC1 scores and AFM height at 9–11 randomly selected points within each of the five cells (Fig. 5(b)) shows that, at each sampling point, the height of the cell is proportional to the PC1 score. However, Fig. 5(b) shows that the slopes of the regression lines vary from cell to cell, therefore the maps of PC1 scores can provide only a relative profile of the cell topography.

We also investigated whether it is possible to obtain an absolute height profile of a cell by measuring the height of the cell at a single point using a Raman z scan, and use this height value to scale the height profile obtained from PC1. In order to achieve this, we first calculated the axial point spread function (PSF_z) of the Raman instrument by deconvolution of the Raman z scans on fixed cells and the cell thickness values measured by AFM (modeled as a top-hat function). Fig. 6(a) presents the 2D Raman maps corresponding to PC1, PC2 and PC3 at various values of z. These results show that when the z position along the axial direction is changed, the Raman spectral images obtained by PCA analysis change in intensity but the general profiles remain the same. This indicates that the PSF_z is significantly larger than the thickness of the cell.

Fig. 6(d) presents typical z profiles of the PC1 scores, obtained at two positions in a fixed cell. These z profiles represent the convolution of the PSF_z of the instrument and the thickness of the cell at the corresponding positions. As the heights of the cells at these positions were measured by AFM, the PSF_z of the instrument was calculated by deconvolution, obtaining a Gaussian function with FWHM of 10.7 ± 0.2 μm (data from 3 cells were used, with 13 Raman z scans for

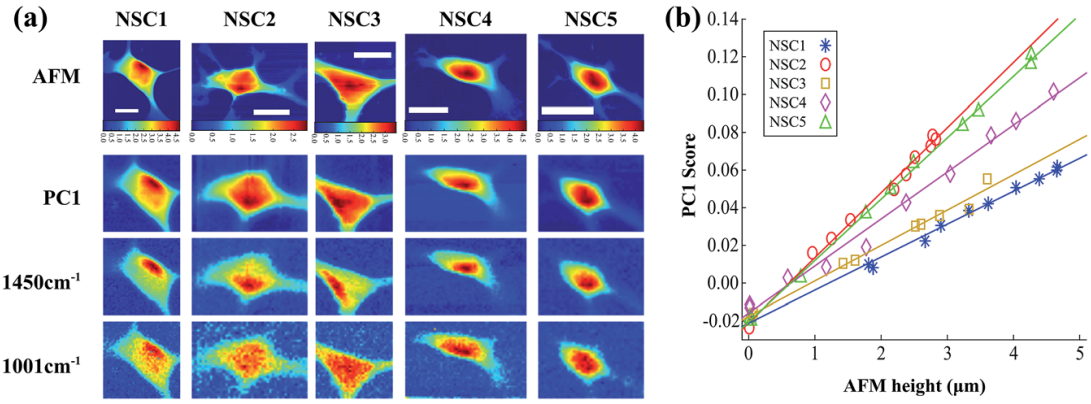

Fig. 5 Cell normalization comparison. (a) Shows AFM maps (scale bar 20 μm) for five different NSCs, with corresponding maps from Raman data, based on PC1 scores, and normalization by peak area for 1450 cm^{-1} and 1001 cm^{-1} peaks. (b) Correlation of AFM height with co-localized PC1 scores.

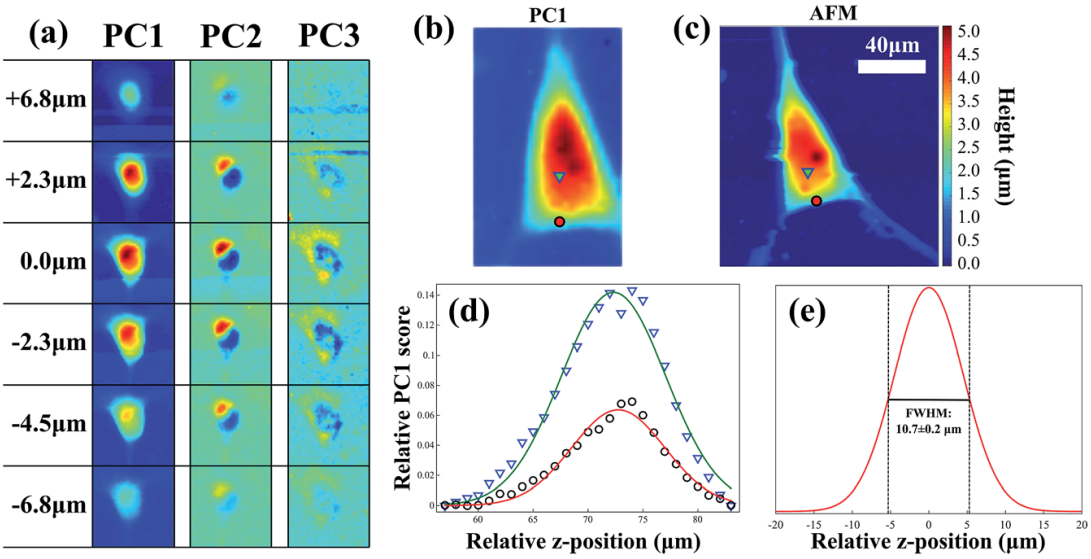

Fig. 6 (a) Raman z-profiling images. (b) PC1 map and (c) AFM for an NSC for which 25 line scans at different z positions were obtained. (d) Shows the variation in the value of the PC1 score as the sampling z position is varied at two locations on the cell of different thickness (indicated in (c) and (d)). Several PC1 z profiles were fitted to a Gaussian function, matched to a top-hat function of width given by the corresponding AFM height, and deconvolved to obtain an estimate for the PSF_z. (e) Shows the resulting average PSF_z measured from 13 deconvolution measurements.

each cell). This measured FWHM was larger than the theoretical value of 1.45 μm; discrepancies likely to be due to the use of MgF_2 rather than a glass/quartz coverslip, as well as potential differences in refractive index variations within the

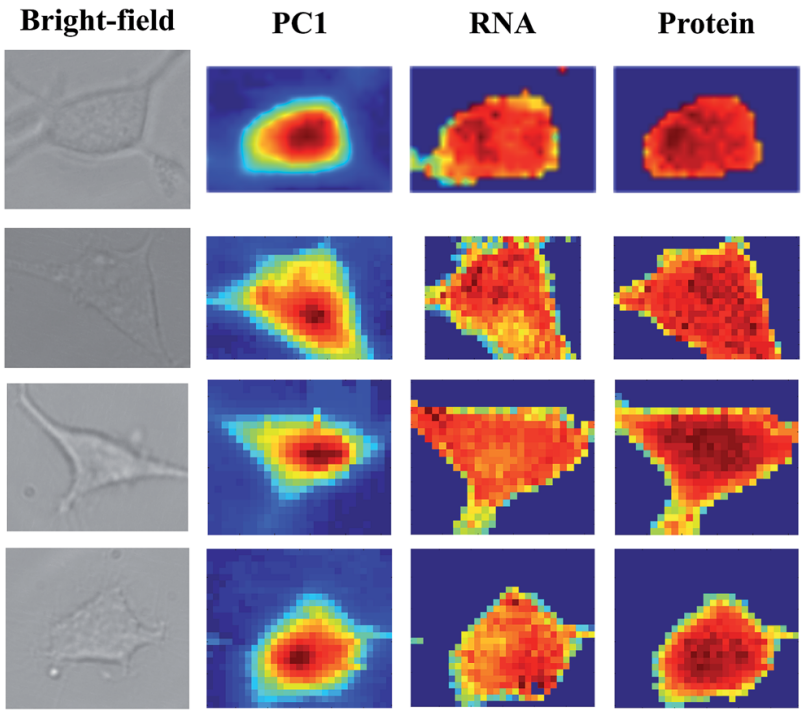

Fig. 7 Relative concentrations of RNA and proteins in live NSCs. The concentration maps were obtained by applying the calibration models described in Fig. 3 on Raman maps normalized using the scores of PC1.

cells. After the calculation of the PSF_z, Raman z scans obtained from five new cells were deconvolved in order to obtain the local height of the cells. However, comparison between the height values obtained by deconvolution and the values measured by AFM indicated typical errors of ~50%. The main factor affecting these errors is related to the fact that the FWHM of the PSF_z is much larger than the typical heights of the cells (~1–5 μm). Using a confocal Raman microscope with a narrower PSF_z may improve the measurement of the cell height. However, measuring 2D Raman maps with a narrower PSF_z would create PC1 maps that are no longer representative of cell thickness as the laser intensity in the axial direction is no longer constant throughout the cell thickness. Thus, a full 3D Raman scan would be required in order to measure the cell topography and obtain absolute molecular concentration maps of the cell.

Because the acquisition of 3D Raman maps is impractical, in particular for live cells, we measured 2D Raman maps of live NSCs and obtained relative concentration maps for RNA and proteins by applying the calibration models in Fig. 3 and normalizing the Raman spectra using the PC1 scores. Fig. 7 shows that, after normalization, the concentration of proteins in the cells is homogenous, which is in agreement with the results for fixed cells. This result is to be expected, especially considering that the acquisition time of the Raman maps was 5–15 minutes, averaging any potential short-lived heterogeneities. On the other hand, RNA distribution indicates a higher concentration of RNA in the cytoplasm than in the nucleus and the endoplasmic reticulum, in agreement with the results for fixed cells, and expected biologically.

Conclusion

RMS and AFM are non-invasive label-free techniques that can be used on cells to provide complementary information. RMS mapping allows qualitative measurements of the spatial distribution of particular biomolecules in living cells. However, when comparing cells with different morphologies, it is difficult to know whether spectral differences arise from molecular differences (*i.e.* differences in concentration) or are due to cell topography. Here, we combined RMS with AFM to understand these effects. AFM was used to map the topography of fixed cells, and thus provide a representation of cell volume. Using solution-based calibration models for RNA and albumin, the RMS and AFM data were used to create quantitative maps of RNA and proteins in individual fixed cells. Because the acquisition times of AFM and Raman measurements are too long to be compatible with live cell measurements, we also used the combined AFM–Raman dataset from fixed cells to investigate potential improvements for normalization of Raman spectral maps. We found that if the PSF of the Raman instrument is significantly wider than the thickness of the cells, the maps of Raman PC1 scores can be correlated with the AFM map, and thus the cell volume. Using this technique, qualitative maps of RNA and protein concentrations were obtained for live cells. While this study demonstrates the potential of using AFM and RMS for measuring concentration maps for individual NSCs *in vitro*, further studies are required to establish the robustness of the normalization method based on principal component analysis when comparing Raman spectra of cells with large morphological differences.

Acknowledgements

The development of the Raman micro-spectrometer was supported by the Biotechnology and Biological Sciences Research Council UK research grant BB/G010285/1. The NSC culture was available thanks to support from the Alzheimer's Society (S. S. and V. S.).

References

1 L. Mueller, H. Traub, N. Jakubowski, D. Drescher, V. I. Baranov and J. Kneipp, Trends in single-cell analysis by use of ICP-MS, *Anal. Bioanal. Chem.*, 2014, **406**(27), 6963–6977.
2 A. Ståhlberg and M. Kubista, The workflow of single-cell expression profiling using quantitative real-time PCR, *Expert Rev. Mol. Diagn.*, 2014, **14**(3), 323–331.
3 G. J. Puppels, F. F. M. de Mul, C. Otto, J. Greve, M. Robert-Nicoud, D. J. Arndt-Jovin and T. M. Jovin, Studying single living cells and chromosomes by confocal Raman microspectroscopy, *Nature*, 1990, **347**, 301–303.
4 A. F. Palonpon, M. Sodeoka and K. Fujita, Molecular imaging of live cells by Raman microscopy, *Curr. Opin. Chem. Biol.*, 2013, **17**(4), 708–715.
5 T. Huser and J. Chan, Raman spectroscopy for physiological investigations of tissues and cells, *Adv. Drug Delivery Rev.*, 2015, **89**, 57–70.
6 B. Kann, H. L. Offerhaus, M. Windbergs and C. Otto, Raman microscopy for cellular investigations—From single cell imaging to drug carrier uptake visualization, *Adv. Drug Delivery Rev.*, 2015, **89**, 71–90.
7 M. Okada, N. Isaac Smith, A. Flotildes Palonpon, H. Endo, S. Kawata, M. Sodeoka and K. Fujita, Label-free Raman observation of cytochrome c dynamics during apoptosis, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**(1), 28–32.
8 H. N. Noothalapati Venkata and S. Shigeto, Stable isotope-labeled Raman imaging reveals dynamic proteome localization to lipid droplets in single fission yeast cells, *Chem. Biol.*, 2012, **19**(11), 1373–1380.
9 F. C. Pascut, S. Kalra, V. George, N. Welch, C. Denning and I. Notingher, Non-invasive label-free monitoring the cardiac differentiation of human embryonic stem cells *in vitro* by Raman spectroscopy, *Biochim. Biophys. Acta, Gen. Subj.*, 2013, **1830**(6), 3517–3524.
10 A. Naemat, H. M. Elsheikha, A. Al-Sandaqchi, K. Kong, A. Ghita and I. Notingher, Analysis of interaction between the apicomplexan protozoan Toxoplasma gondii and host cells using label-free Raman spectroscopy, *Analyst*, 2015, **140**(3), 756–764.
11 S. O. Konorov, H. Georg Schulze, C. G. Atkins, J. M. Piret, S. A. Aparicio, R. FB Turner and M. W. Blades, Absolute quantification of intracellular glycogen content in human embryonic stem cells with Raman microspectroscopy, *Anal. Chem.*, 2011, **83**(16), 6254–6258.
12 A. Ghita, F. C. Pascut, M. Mather, V. Sottile and I. Notingher, Cytoplasmic RNA in undifferentiated neural stem cells: a potential label-free Raman spectral marker for assessing the undifferentiated status, *Anal. Chem.*, 2012, **84**(7), 3155–3162.
13 I. Notingher, S. Verrier, S. Haque, J. M. Polak and L. L. Hench, Spectroscopic study of human lung epithelial cells (A549) in culture: living cells *versus* dead cells, *Biopolymers*, 2003, **72**(4), 230–240.

14 Y. Chen, L. Su, G. N. Zhang, L. H. Wang, A. G. Shen, X. D. Zhou, X. H. Wang and J. M. Hu, *In vivo* and *in situ* monitoring of the nitric oxide stimulus response of single cancer cells by Raman spectroscopy, *Laser Phys. Lett.*, 2013, **10**(4), 045608.

15 Q. Matthews, A. Jirasek, J. Lum, X. Duan and A. G. Brolo, Variability in Raman spectra of single human tumor cells cultured *in vitro*: correlation with cell cycle and culture confluency, *Appl. Spectrosc.*, 2010, **64**(8), 871–887.

16 M. Chen, N. McReynolds, E. C. Campbell, M. Mazilu, J. Barbosa, K. Dholakia and S. J. Powis, The Use of Wavelength Modulated Raman Spectroscopy in Label-Free Identification of T Lymphocyte Subsets, Natural Killer Cells and Dendritic Cells, *PLoS One*, 2015, **10**, e0125158.

17 F. C. Pascut, H. T. Goh, N. Welch, L. D. Buttery, C. Denning and I. Notingher, Noninvasive detection and imaging of molecular markers in live cardiomyocytes derived from human embryonic stem cells, *Biophys. J.*, 2011, **100**(1), 251–259.

18 T. Tolstik, C. Marquardt, C. Matthäus, N. Bergner, C. Bielecki, C. Krafft, A. Stallmach and J. Popp, Discrimination and classification of liver cancer cells and proliferation states by Raman spectroscopic imaging, *Analyst*, 2014, **139**(22), 6036–6044.

19 S. F. El-Mashtoly, H. K. Yosef, D. Petersen, L. Mavarani, A. Maghnouj, S. Hahn, C. Kötting and K. Gerwert, Label-Free Raman Spectroscopic Imaging Monitors the Integral Physiologically Relevant Drug Responses in Cancer Cells, *Anal. Chem.*, 2015, **87**(14), 7297–7304.

20 S. C. Erfurth, E. J. Kiser and W. L. Peticolas, Determination of the backbone structure of nucleic acids and nucleic acid oligomers by laser Raman scattering, *Proc. Natl. Acad. Sci. U. S. A.*, 1972, **69**(4), 938–941.

21 J. M. Benevides and G. J. Thomas, Characterization of DNA structures by Raman spectroscopy: high-salt and low-salt forms of double helical poly(dG-dC) in H_2O and D_2O solutions and application to B, Z and A-DNA*, *Nucleic Acids Res.*, 1983, **11**(16), 5747–5761.

22 T. G. Spiro and B. P. Gaber, Laser Raman scattering as a probe of protein structure, *Annu. Rev. Biochem.*, 1977, **46**(1), 553–570.

23 C. Krafft, L. Neudert, T. Simat and R. Salzer, Near infrared Raman spectra of human brain lipids, *Spectrochim. Acta, Part A*, 2005, **61**(7), 1529–1535.

24 R. Milo, What is the total number of protein molecules per cell volume? A call to rethink some published values, *BioEssays*, 2013, **35**(12), 1050–1055.

25 C. Chen, D. A. Ridzon, A. J. Broomer, Z. Zhou, D. H. Lee, J. T. Nguyen, M. Barbisin, N. L. Xu, V. R. Mahuvakar, M. R. Andersen, K. Q. Lao, K. J. Livak and K. J. Guegler, Real-time quantification of microRNAs by stem–loop RT-PCR, *Nucleic Acids Res.*, 2005, **33**(20), e179.

26 G. A. Valaskovic, N. L. Kelleher and F. W. McLafferty, Attomole protein characterization by capillary electrophoresis-mass spectrometry, *Science*, 1996, **273**, 1199–1202.

27 S.-I. Fujii, K. Inagaki, S.-I. Miyashita, A. S. Groombridge, K. Nagasawa, K. Chiba and A. Takatsu, Separation and quantification of RNA molecules using size-exclusion chromatography hyphenated with inductively coupled plasma-mass spectrometry, *Electrophoresis*, 2014, **35**(9), 1315–1318.

28 M. Wind, A. Wegener, A. Eisenmenger, R. Kellner and W. D. Lehmann, Sulfur as the Key Element for Quantitative Protein Analysis by Capillary Liquid

Chromatography Coupled to Element Mass Spectrometry, *Angew. Chem., Int. Ed.*, 2003, **42**(29), 3425–3427.

29 N. Zinn, R. Krüger, P. Leonhard and J. Bettmer, μLC coupled to ICP-SFMS with post-column isotope dilution analysis of sulfur for absolute protein quantification, *Anal. Bioanal. Chem.*, 2008, **391**(2), 537–543.

