2020

Empirical Study on the Effects of Acquisition Parameters for FTIR Hyperspectral Imaging of Brain Tissue

David Perezguaita
Technological University Dublin, david.perezguaita@tudublin.ie

Julia Sacharz
Monash University, Victoria, Australia

Mustafa Kasniz
Photothermal Spectroscopy Corp. 325 Chapala St, Santa Barbara, CA

See next page for additional authors

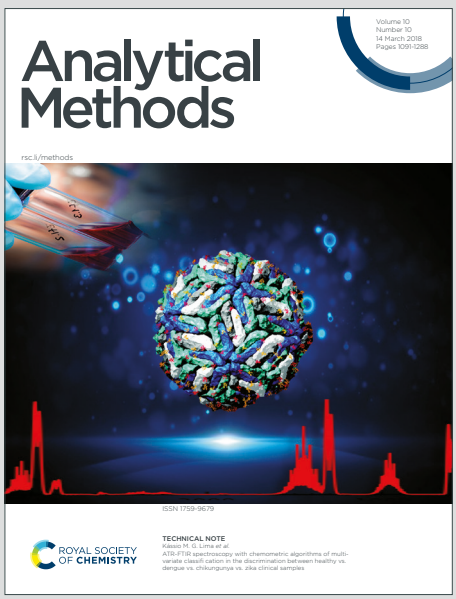
Recommended Citation

Sacharz, J. et al. (2020) Empirical study on the effects of acquisition parameters for FTIR hyperspectral imaging of brain tissue, *Anal. Methods,* 2020, DOI: 10.1039/C9AY01200A

Analytical Methods

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: J. Sacharz, D. Perez-Guaita, M. Kansiz, S. S. Nazeer, A. Weselucha-Birczynska, S. Petratos, B. R. Wood and P. Heraud, *Anal. Methods*, 2020, DOI: 10.1039/C9AY01200A.



ROYAL SOCIETY OF CHEMISTRY

rsc.li/methods

Empirical study on the effects of acquisition parameters for FTIR hyperspectral imaging of brain tissue

View Article Online
DOI: 10.1039/C9AY01200A

J. Sacharz[1,2], D. Perez-Guaita[1,3], Mustafa Kansiz[4], Shaiju S. Nazeer[1], A. Weselucha-Birczynska[2], S. Petratos,[5] B. R. Wood[1*] and P. Heraud[1,6*]

[1] Centre for Biospectroscopy and School of Chemistry, Monash University, 3800, Victoria, Australia

[2] Faculty of Chemistry, Jagiellonian University, Gronostajowa 2, 30-387, Kraków, Poland

[3]FOCAS Research Institute, Technological University Dublin, City Campus, Dublin, Ireland

[4] Photothermal Spectroscopy Corp. 325 Chapala St, Santa Barbara, CA, 93101, USA

[5] Department of Neuroscience / Central Clinical School, Monash University, Alfred Centre, 99 Commercial Rd, Prahran, 3004, Victoria, Australia

[6] Department of Microbiology and the Biomedical Discovery Institute, Faculty of Medicine, Nursing and Health Sciences, Monash University, 3800, Victoria, Australia.

Abstract

Fourier transform infrared (FTIR) spectroscopic imaging is a powerful technique for molecular imaging of pathologies associated with the nervous systems including multiple sclerosis research. However, there is no standard methodology or standardized protocol for FTIR imaging of tissue sections that maximize the ability to discriminate between the molecular, white and granular layers, which is essential in the investigation of the mechanism of demyelination process. Tissue sections are heterogeneous, complex and delicate, hence the parameters to generate high quality images in minimal time becomes essential in the modern clinical laboratory. This article presents an FTIR spectroscopic imaging study of post-mortem human brain tissue testing the effects of various measurement parameters and data analysis methods on image quality and acquisition time. Hyperspectral images acquired from the same region of a tissue using a range of the most common optical and collection parameters in different combinations were compared. These included magnification (4 × and 15 ×), number of co-added scans (1, 4, 8, 16, 32, 64 and 128 scans) and spectral resolution (4, 8 and 16 cm^{-1}). Images were compared in terms of acquisition time, signal-to-noise (S/N) ratio, and accuracy of the discrimination between three major tissue types in a section from the cerebellum

(white matter, granular and molecular layers). In the latter case, unsupervised k-means cluster (KMM) analysis was employed to generate images from the hyperspectral images, which were compared to a reference image.

The classification accuracy for tissue class discrimination was highest for the 4x magnifying objective, with 4 cm^{-1} spectral resolution and 128 co-added scans. The 15x magnifying objective gave the best accuracy for a spectral resolution of 4 cm^{-1} and 64 scans (96.3%), which was just above what was achieved using the 4x magnifying objective, with 4 cm^{-1} spectral resolution and 32 and 64 co-added scans (95.4 and 95.6 %, respectively). These findings were correlated with a decrease in S/N ratio with increasing number of scans and was generally lower for the 15x objective. However, longer scan times were required using the 15x magnifying objective, which did not justify the very small improvement in the classification of tissue types.

Keywords

Fourier Transform Infrared Spectroscopy, FTIR imaging, brain tissue, k-means cluster analysis, multiple sclerosis

Introduction

Fourier transform infrared (FTIR) spectroscopy is a powerful and non-destructive technique, which provides information about the molecular composition in the form of a unique spectrum for every sample[1-3]. The spectrum can be treated as a molecular "fingerprint", which enables investigations into the structural and macromolecular composition of complex biological materials. As a versatile tool, it has been applied to many fields of science[2-3]. In particular, biological and medical research are becoming niche fields for the technology with many papers on the detection of breast[4], skin[5], ovarian[6] tumors, heart and liver diseases[2] appearing in the literature. Advanced Vibrational spectroscopy methods and chemometric based analysis have been used also to examine brain tissues[2, 7-14] for studying glioma[2, 7-13] and neurodegenerative diseases

like epilepsy [14-17], Alzheimer's, Parkinson's and multiple sclerosis[2, 13, 18] making it an important tool in brain research[2].

View Article Online
DOI: 10.1039/C9AY01200A

The human brain is one of the most important and complex organs. It is highly specialized in every distinct heterogeneous anatomical region. Nerve cells generate electrical impulses that travel through the body based on semipermeable excitable membranes that modify permeation to small chemical molecules[19]. Any changes at the molecular level of nerve tissue can cause dysfunction to any part of the body. Brain tissue is composed of water (70–83%), protein (7.5–8.5%) and lipids (5–15%). In white matter, the concentration of lipids reaches 15% while in the gray matter only 5%[20-21]. Brain lipids can be divided into three main classes: neutral lipids, phospholipids, and sphingolipids[20-21]. However, the composition and concentration of lipids can change in very specific ways in diseased brain tissue. Hence, most studies focus on lipid concentration because it is a useful diagnostic parameter to determine for example type and grade of a brain tumor or neurodegenerative disease[2, 20]. However, the analysis of lipids in tissue by FTIR is challenging due to the standard fixation procedures that often entails the use of paraffin. Bands from paraffin obscure important regions of the spectra especially the C-H molecular vibrations of lipids. Importantly, they do not interfere with ester, acid and alcohol functional groups of lipids. Elimination of paraffin through organic solvents removes paraffin contributions, but it can also alter the lipid content of the tissue.[1]

Both the sample preparation and the method of the measurement are essential in achieving high-quality images when using FTIR spectroscopy to study the brain. Preparation of a brain tissue sample requires an experienced histologist and an appropriate procedures[2, 13, 22]. Although the FTIR method is non-destructive and quite easy to perform, measurements of large areas of tissue can take a long time. Therefore, most of the biological samples need to be safely stored for long periods of time[2, 13, 22]. In this context, the best format of the biological tissues for FTIR spectroscopic analysis are formalin-fixed paraffin embedded tissues with an appropriate thickness. To avoid absorbance saturation due to an excessive long pathlength, the thickness should be less than 10 μm [2,13] when using transmission substrates and 5 μm [2] when using transflection low emissivity (low E) slides. The sample needs to be placed on an IR-transparent material such as BaF_2 or CaF_2, for transmission measurements. This mode provides the best S/N ratio[2-3] and is free from optical

artifacts associated with transflection mode. It is also crucial to select appropriate measurement parameters, such as the type of objective (magnification), spectral resolution and number of scans[23]. The correct configuration and parameters result in optimal image acquisition speed and spectral quality. However, there is no standard procedure suitable for every case and the acquisition parameters need to be established for different tissue fixation and sectioning techniques[1-3, 23] depending on the specific aim of the experiment.

In FTIR hyperspectral imaging each FTIR spectrum is ascribed to a particular x, y spatial coordinate generating a spectral hypercube containing information on the macromolecular chemistry of the tissue section. The images can be presented as false-colour maps by using tools such as Unsupervised Hierarchical Cluster Analysis (UHCA). UHCA and *k*-means clustering (KMC) are often used to identify tissue histology and the resultant images can be directly compared to stained tissue sections to assist in identifying both anatomical and histopathological features at the molecular level[1-3].

Many recent studies have shown that the preparation of samples, experimental setup, the data pre-proceedings and data analysis plays a crucial role in the diagnosis of any type of tissues using vibrational spectroscopy [2,24-34].

Effectiveness of high magnification FTIR-FPA bench-top chemical imaging system with cluster and Principal Component Analysis (PCA) has been demonstrated for mice brain tissue. [25] The usefulness of this approach was the ability to co-localize molecular changes to different areas around the plaques and discover marker bands surrounding the plaque from the co-accumulation of molecular components indicative of inflammatory states [25]. Despite the enormous potential of FTIR spectroscopy, clinical implementation has been hampered because of practical obstacles like the speed of data acquisition and the lack of optimized computational procedures for extracting clinically relevant information. A modified Bayesian classification protocol was applied to aid in the digital identification of molecular pathology of cancer [26]. Use of a Quantum Cascade Laser (QCL) system for this study sped up the image acquisition. It has been shown that tunable QCLs provide continuous coverage of the wavenumber range of interest enabling the discrete imaging of larger sample areas in minutes and at a pixel size smaller than that generally employed by the

View Article Online
DOI: 10.1039/C9AY01200A

FTIR globar/hotwire systems. The great benefit of the QCL system is the use of a room temperature detector with no required liquid nitrogen cooling [27].

A crucial issue is the method of initial sample preparation. K-means cluster analysis was performed on spectral data from patients with prostate cancer, by monitoring the methylene hydrocarbon-associated vibrations as a function of solvent immersion time during washes of either hexane or xylene, indicated that after 5–10 minutes of immersion in the solvent the hydrocarbon-associated vibrations remain fairly consistent across the tissue. This suggests that solvent-resistant lipids remain present in formalin-fixed tissue as they are locked into protein–lipid complex matrices [28]. Advanced statistical techniques like Linear Discriminant Analysis (LDA), which is a statistical multivariate supervised method, and PCA (unsupervised method prior to the LDA analysis), achieved an ~ 85% classification accuracy among normal, polyp, and cancer groups in the tissue of the benign premalignant colonic polyps [29]. The complex analysis, performed on prostate cancer tissues using appropriate calculations for S/N ratio, reduced the data acquisition time by a factor of ca. 3 without significant degradation in classification accuracy. The classification accuracy was shown to depend on the S/N ratio of the recorded data, but improving the spectral resolution had no effect in the classification modelling [30]. A study performed on breast cancer tissue using high-performance DFIR microscopy enabled histologic imaging [31]. The optical design was combined with real-time control algorithms to reduce errors and achieve minimal-distortion on the acquired images. This enabled the use of a single-element detector plus the ability to modulate the beam, while the confocal geometry helped reduced noise within each spectral band, allowing the construction of precise classification models using a fewer number of discrete wavenumber values without extensive signal averaging or mathematical noise rejection. Enhanced spatial image quality provided detailed tissue segmentation for tumor detection [31] and similar approaches have been trailed in studies of whole organisms, like nematodes [32-34].

Research into demyelination processes in the brain such as Multiple Sclerosis (MS) requires a reliable differentiation into the three main regions of cerebral tissue. However, sometimes the discrimination is not always straightforward especially if the signal-to-noise ratio is low or if spectral bands cannot be resolved due to poor spectral resolution. In other cases, an excess of resolution and number of scans results in large acquisition times, which make it impractical for the acquisition of several samples. The aim of this work was

to verify by empirical means the influence of the data collection method and measurement parameters for tissue differentiation. We believe the approach shown here will help improve the quality and consistency of FTIR images recorded of sensitive biological material, particularly for brain sections providing a way to establish optimal collection parameters to achieve a high S/N in minimal time.

View Article Online
DOI: 10.1039/C9AY01200A

Experimental

Samples

A section of cerebellum brain tissue from a human adult female was used for our studies, obtained from post mortem material kept by the Southern Health Brain Tissue Bank, Melbourne, Victoria, Australia. All frozen human deep-cortical white matter tissues were acquired from the Victorian Brain Bank Network (VBBN) under the National Health and Medical Research Council guidelines and the Monash University Human Research Ethics Committee approval number CF13/1646-2013000831. Post-mortem interval did not exceed 56 h. All specimens were obtained from the frontal lobe deep white matter, frozen under liquid nitrogen and then stored in the Brain Biobank (-80°C until required). Brain tissue was taken post mortem and the dissected pieces (of volume of approximately 2-5 cm^3 each) dropped into liquid nitrogen. No cryo-preservatives were used. The frozen tissue pieces were fixed by placing them directly in 10% phosphate-buffered formalin for 2 weeks. The fixed tissue fragments were dehydrated by placing them three times for 2 hours in the following ethanol concentrations: 50%, 70%, and finally absolute ethanol. The tissue was then cleaned in xylene (3 changes, 2 hours in each). The cleaned tissue was placed in molten paraffin (3 times, for 2 hours in each), then embedded. The paraffin brain tissue blocks were warmed briefly (2-3 min) in a water bath at 50 ºC before cutting. Tissue sections of 10 μm thickness were mounted on the CaF_2 substrates. The samples were not deparaffinized before measurements.

FTIR hyperspectral imaging

FTIR spectra were acquired at the Centre for Biospectroscopy at Monash University in Melbourne, Australia, using an Agilent (Santa Clara, USA) Cary 620. Spectra were recorded in standard magnification, transmission mode over the region 4000–1000 cm^{-1}. The sample was imaged using two kinds of objectives:

4x and 15x with the pixel size of 19 µm × 19 µm and 5.5 µm × 5.5 µm, respectively. Single image tiles (128 x 128 of the focal plane array) covered 700 µm x 700 µm for 15x objective and 2400 µm x 2400 µm for 4x objective. For each objective, different images were obtained combining different numbers of co-added scans (1, 4, 8, 16, 32, 64 and 128 scans) and different spectral resolutions (4, 8 and 16 cm^{-1}). In total, 21 images were obtained for each objective, representing every combination of co-added scan number and spectral resolution. A background image was acquired with the same corresponding magnification and resolution in a clean area on the slide using four times the number of scans compared to the sample measurement. The region of interest in the tissue was acquired using a mosaic of 12 tiles (x=4 by y=3) for the 15x objective and 1 tile (x=1 by y=1) for the 4x objective. Special care was taken to ensure the images were always acquired at the same position on the tissue.

Data analysis

The resulting FTIR hyperspectral data cubes were processed using Matlab v2016a from Mathworks (Natic, USA) software. Data was processed using Matlab scripts and functions written in-house. Preprocessing of data was carried out using the functions available in the PLS toolbox v8.2. from Engeinvector (Manson, USA). Only the spectral regions from 1500-1800 and 1000-1350 cm^{-1} were considered for the ensuing analysis, which includes the amide region (1700–1250 cm^{-1}), the ν_{as} PO^{2-} region (1250–1200 cm^{-1}) and the ν_{as} C–O region (1060–1000 cm^{-1}). The spectral region between the 1350 cm^{-1} and 1500 cm^{-1} containing the CH bending bands from paraffin were eliminated from the spectra prior to the analysis.

KMC analysis was carried out by employing the *kmeans* function from the statistical toolbox of Matlab using the squared Euclidean distance. The following preprocessing was carried out prior to the KMC: (i) calculation of the first derivative of each spectrum using a Savitzky-Golay filter (15 points and 2nd order polynomial); (ii) standard normal variate normalization; and (iii) mean centering. The noise reduction was performed using the PCA function available in the PLS toolbox. In short, PCA was applied pixel-wise to the raw image. Data was then reconstructed considering a limited number of PCs, aiming to eliminate the spectral variation associated with random noise (see the Results section for more information).

View Article Online
DOI: 10.1039/C9AY01200A

Results & discussion

Characterization of the different types of brain tissue

A visible image of the stained tissue under investigation is shown in Figure 1a. The area included three different structures of the cerebellum: white matter; granular layer; and the molecular layer. The white matter in the cerebellum is composed mainly of myelinated axons, the granular layer contains mainly the cell bodies of neurons associated with the axons in white matter, whereas the molecular layer is comprised mainly of a mixture of myelinated axons and dendrites. Two reference maps, which contained the same number of pixels as the IR images using the 15x and 4x objectives, were created using two criteria: i) the histopathological image, and ii) preliminary results of the cluster analysis. First, each pixel was categorized as a class according to the assignment in the H&E stained image (Figure 1a and 1b), which is considered as a gold standard for differentiating between the granular, white and molecular regions. However, it should be noted that FTIR images and stained images come from adjacent tissue sections. This implies that the distribution of the tissue is very similar but not the same and edges and borderlines of tissue could be slightly modified. To improve the accuracy of the reference maps, cluster maps obtained with the highest S/N ratio, were visually compared with the reference maps and no major differences were found in the tissue sections. We marked 4 clusters corresponding to the three different tissue types: white matter (red colour), granular layer (yellow colour) and molecular layer (light blue colour) [Fig. 1b]. The violet colour on the map marked the area that included empty spaces in the sample caused by cracks in the tissue section.

Figure 1c shows the average spectra for each cluster for 1500-1800 cm^{-1} and 1100-1350 cm^{-1} obtained from the image. The spectral range between 1350-1500 cm^{-1} contained only two intense paraffin bands and was removed for analysis. In principle, one could consider the subtraction of paraffin contributions using chemometric methods. With that aim, a spectrum from pure paraffin was extracted from one of the images. The contribution of the paraffin to the spectra was calculated using the integration of the band at 2850 cm^{-1} and for each pixel this contribution was multiplied by the extracted spectrum of paraffin and subtracted from the original spectra, creating a corrected image. However, as shown in the Electronic Supplementary

View Article Online
DOI: 10.1039/C9AY01200A

Information (ESI), the corrected spectra provided similar results to the analysis excluding the 1500-1300 cm^{-1} region. Hence, it was decided to do not over treat the data and continue the analysis excluding the region instead of correcting the whole image.

The spectra in 1800-1500 cm^{-1} range were similar for all clusters. The bands at 1654 cm^{-1} and 1544 cm^{-1} correspond to the protein amide I and amide II modes, respectively. The range from 1350-1100 cm^{-1} showed more heterogeneity within the sample. The visible bands at 1345 cm^{-1} included: the CH_2 wagging vibration from phospholipids, fatty acids, triglycerides, and amino acid side chain vibrations; amide III vibrations from proteins at 1304 cm^{-1}; PO_2^- asymmetric stretch from DNA at 1242 cm^{-1}; a band at 1123 cm^{-1} assigned to the C-O stretching band of ribose ring from RNA, and a band at 1064 cm^{-1} assigned to the -CO-O-C asymmetric stretch from cholesterol esters[1,2]. Furthermore, the carbonyl band generally found at 1735 cm^{-1}, is not very intense compared to nucleic acid bands. The violet spectrum shows smaller amide bands in comparison to the other classes, and negative bands in the 1350-1100 cm^{-1} region resulting from low absorbing proteins in this region of tissue.

White matter in the cerebellum consists mostly of axons and neurons covered with myelin and is subsequently rich in lipids and proteins. The molecular layer is a mixture of the granular cell axons and Purkinje dendrites. The intensity of the band at 1345 cm^{-1} is greater for these two brain regions, which contain more lipids compared to the granular layer. The granular layer contains mostly the cell bodies dominated by nuclei. The band at 1242 cm^{-1} is the strongest in the granular layer, which indicates the presence of nucleic acids in the cell nucleus. We also observe weak DNA bands in the white matter and molecular layer, because the white matter contains several deep nuclei, like the dentate nucleus and molecular layer contains Purkinje dendrites[35-36].

Comparison of data acquisition parameters

The accuracy of tissue identification

For each image, 4 different KMC images were performed establishing 4, 5, 6 and 7 classes, respectively. In most cases, each class of the KMC analysis contained pixels associated with different tissue regions and hence the frequency of the pixels of each class in the different brain regions were computed. The whole

class was finally assigned to the region with the largest occurrence. Figure 2 shows all cluster images obtained with the 4 × and 15 × objectives from the three different spectral resolutions and 7 groups associated with different numbers of co-added scans. The best results were achieved from images acquired at 4 cm^{-1} spectral resolution and co-adding 128 scans. The 15 × objective gave better results compared to the 4 × objective based on the tissue class separation and distinct boundaries between tissues, even with the use of a smaller number of scans. By increasing the number of scans, we could achieve better images in all cases. Both spectral resolutions, 4 cm^{-1} and 8 cm^{-1} showed similar results, while resolution 16 cm^{-1} gave adequate class separation only with more scans.

The effect of the measurement parameters on the classification capability of FTIR hyperspectral imaging was quantified in Figure 3a, which shows a table with the classification accuracy of the k-means clustering of measurements from the experimental brain images acquired at different magnification, resolution and number of the scans. These values were obtained comparing the KMC images of Figure 2 with the regions of the reference image (Figure 1). In both magnifications the accuracy of tissue classification grew with the number of scans for all resolutions and the highest accuracy is obtained for resolution 4 cm^{-1} and 128 scans. From these results, it is clear the expected relation between the classification accuracy and the number of scans averaged for obtaining the spectrum, indicating that a minimum of signal-to-noise ratio is needed to detect the differences between the tissue classes. However, the measurement time grew with the number of scans and the acquisition of large image mosaics (i.e. with the 15 × objective) can be time consuming if large numbers of scans were acquired (see Table1). The image acquisition time ranged from almost two hours to less than a minute. Shortest acquisition time is important for process automation, however, the image processing time for UHCA maps is often the bottle neck in FPA imaging. The time taken for images with an accuracy above 90 % is less than 9 minutes using the 15 × objective. By comparison, the combination of 4 × objective and a larger number of scans (32 or greater) was able to achieve a 90% accuracy within 2-5 minutes. This time advantage would be further enhanced when the sample area becomes larger and there is a necessity for high throughput sample analysis. It is important to note that between 4 × and 15 × magnifications, the brain tissue presented shows clear boundaries using 4 × magnification and the tissue discrimination did not appear to require any further increase in spatial resolution. The acquisition time

also decreased when the spectral resolution is decreased because the distance the moving mirror actually moves is decreased at low spectral resolution. The shortest time of measurements was observed using 16 cm^{-1} spectral resolution, but at this resolution, the ability to discriminate the tissue regions was compromised; the tissue class discrimination accuracy was similar for all numbers of scans and complete discrimination of the tissues was only observed for 128 scans using the 4 × objective, while this could be achieved with 16 and 64 scans for the 15 × objective. Of course, the 16 cm^{-1} or even 32 cm^{-1} may be sufficient for some tissue sections depending on the anatomical or histological features. For instance, Bhargava et al.[30] found that discrimination between cell types is almost identical for 4 to 16 cm^{-1} resolution but starts degrade at 32 cm^{-1}.

This is clearly seen in figure 4, which shows the precision calculated as the percentage of true positives over the sum of true-positives and false-positives from class identification compared to the reference map. It can be seen that the differentiation of white matter was not successful in any image acquired at 16 cm^{-1}. KMC image analysis (Figure 2) indicates that the white matter was in most of the cases confused with molecular layer. This was the case using the 4 × objective with 4 and 8 cm^{-1} spectral resolutions using 1-8 scans, and with 16 cm^{-1}, spectral resolution using 1-64 scans. Acceptable discrimination of white matter was obtained at both magnifications using 4 cm^{-1} spectral resolution and greater than 16 scans. The reason for the difficulty to discriminate white matter from the molecular layer was presumably because of spectral similarities between these two classes (Figure 1c), especially in the 1100-1300 cm^{-1} region. Both contain myelinated axon and hence the spectra contain bands from lipids.

The hyperspectral matrix was reshaped as a column of spectra with every 100th spectrum being selected. For each selected spectrum, the region between 1750 and 1800 cm^{-1} was baseline corrected using the Automatic Weighted Least Squares algorithm available in the PLS toolbox package (polynomial order=1) and the standard deviation of the absorbance at this corrected region is considered as noise. The overall noise value was the mean of noise for all the selected spectra. Figure 3b shows a table that summarizes the noise obtained from the calculations. As expected, the data reveals that the noise increases with increasing the spectral resolution and decreasing with the number of scans. The latter is probably caused by the

presence of some water vapour interference, especially at 4 cm^{-1}, which is characterized by sharp narrow rotational bands that contribute to the noise based on calculating the standard deviation of the absorbance within the 1750-1800 cm^{-1} region. As seen in Figure 3b, the noise at 128 scans was less that at 64 scans, however, unexpectedly in some cases, the noise at 32 scans was less than at 64 scans. Considering that these images were obtained with a large acquisition time, the increase in the noise may be caused by changes in the environment, such as fluctuating humidity during the measurement. A large number of scans also implies a large acquisition time, which means that a longer time elapsed between the background and the sample measurement. In general, this result indicates that when selecting the number of scans, there is a tradeoff between the noise level required for the classification and the time needed for the measurement.

Effect of Principal Component Analysis (PCA) noise reduction

Results indicated that the noise level was crucial in the classification capability of the different images obtained. To determine if the impact of noise could be reduced by post processing we used a PCA noise reduction procedure. This method generates PCA over all the spectra of the image. The spectral matrix is then reconstructed using the first 20-30 principal components (PCs), with the first PCs explaining the largest amount of variance and the rest of the PCs, which are expected to contain the noise components, are rejected. Figure 5 presents the noise of the hyperspectral images before and after noise reduction using 20 and 30 PCs. PCA noise corrected images obtained with 1-16 scans show similar noise levels compared to the raw images obtained with a larger number of scans.

Our next aim was to assess whether the noise correction improved the capability of the KMC analysis to characterise the different regions of the cerebral tissue. KMC analysis was applied over the PCA noise corrected images and classification accuracy values were computed and compared with the uncorrected images. No significant differences between the accuracy values before and after PCA noise reduction were found. The paired t-test comparisons for noise between all the 21 conditions (3 resolutions and 7 number of scans) before and after PCA noise reduction (both 20 PCs and 30 PCs) gave rise to p values greater than 0.05 irrespective of the measurement parameters. Although the noise was reduced measurably, the classification capability was not improved significantly.

Conclusions

This work described the importance of measurement conditions and data processing for imaging large area samples using focal plane array IR microscopes. The classification accuracy for tissue class discrimination was highest for the 4x magnifying objective, with 4 cm^{-1} spectral resolution and 128 co-added scans. In terms of tissue classification accuracy, the 15x magnifying objective gave the best accuracy for a spectral resolution of 4 cm^{-1} and 64 scans (96.3%), which was just above that achieved using the 4x magnifying objective, with 4 cm^{-1} spectral resolution and 32 and 64 co-added scans (95.4 and 95.6 %, respectively). These findings were correlated with decreases in the noise value with increasing number of scans and was generally lower for the 15x objective. However, the longer time acquisition time using the 15x magnifying objective because of the increase in the number did not justify the very small improvement in classification accuracy (54 min for the 15x objective using 64 scans versus 2.5 min for the 4 × objective with 32 scans). We found that when using 16 cm^{-1} spectral resolution, although measurement times were shorter, it was easy to lose the information contained in the spectra necessary for tissue class discrimination for both 4x and 15x magnification.

On the basis of these observations we selected the best parameters for both magnifications. For 4x magnification the best compromise parameters are: 8 cm^{-1} spectral resolution and 128 scans. For 15x magnification the best compromise parameters are: 8 cm^{-1} spectral resolution and 32 scans. Each of the combination sets gives good results, S/N ratio and tissue specificity using cluster analysis. If the best spatial resolution is required, then the 15x would be preferred, as it has a smaller pixel projection area on the tissue plane. The fundamental and greatest advantage of the 4x objective is certainly the measurement time, which makes imaging much faster and comes into its own with large sample areas. This allows us to apply greater number of scans, which significantly improves the images and is crucial for better results. Nevertheless, the selection of the parameters depends on the accuracy needed for each clinical application and on how spectrally similar the classes are, with more spectral similarity requiring better SNR and/or better spectral resolution but it also depends on cluster size, with clusters ~>40um (roughly twice the 4x pixel size of

19um) expected to be resolved well with the 4x objective and thus benefiting from the significant time savings for spatial domains < 40um, then the 15x objective is superior.

View Article Online
DOI: 10.1039/C9AY01200A

Our work proves that FTIR hyperspectral imaging is a promising method for diagnosing anatomical and histopathological features in brain tissue and can be performed in less than 5 min providing one is judicious in the selection of the FTIR hyperspectral collection parameters. This work provides an empirical approach towards developing a Standard Operating Protocol (SOP) for FTIR hyperspectral imaging of brain tissue sections that takes into account measurement parameters including magnification, spectral resolution, and the number of scans, to achieve optimal signal-to-noise and diagnostic precision in the shortest possible time.

Acknowledgements

Julia Sacharz thanks *The Foundation of Students and Graduates of the Jagiellonian University Bratniak* (Kraków, Poland) for financial support assisting travel to and research work in Australia. DPG acknowledges support from the European Research Council MSCA grant (Spectro-metrics, 020-MSCA–IF–2017 Project ID:796287). This work is funded by an Australian Research Council Future Fellowship grant FT120100926 awarded to BW.

Electronic Supporting Information- ESI 1- Digital removal of paraffin from FTIR image sections of hamster brain

References

1 K.R. Bambery, E. Schültke, B.R. Wood, S.T. Rigley MacDonald, K. Ataelmannan, R.W. Griebel, B.H.J. Juurlink, D. McNaughton, *Biochimica et Biophysica Acta*, 2006, **1758**, 900–907

2 K. Malek, B. R. Wood, K. R. Bambery, in *Optical Spectroscopy and Computational Methods in Biology and Medicine*, ed. M. Baranska, Springer, 2014, chapter 15, 419-473, ISBN 978-94-007-7832-0
View Article Online
DOI: 10.1039/C9AY01200A

3 M. J. Baker, J. Trevisan, P. Bassan, R. Bhargava, H. J. Butler, K. M. Dorling, P. R. Fielden, S. W. Fogarty, N. J. Fullwood, K. A. Heys, C. Hughes, P. Lasch, P. L. Martin-Hirsch, B. Obinaju, G. D. Sockalingum, J. Sulé-Suso, R. J. Strong, M. J. Walsh, B. R. Wood, P. Gardner, F. L. Martin, *Nature Protocols*, 2014, **9(8)**, 1771–1791

4 H. Fabian, N. Anh Ngo Thi, M. Eiden, P. Lasch, J. Schmitt, D. Naumann, *Biochimica et Biophysica Acta*, 2006, **1758**, 874–882

5 S. M. Ali, F. Bonnier, H. Lambkin, K. Flynn, V. McDonagh, C. Healy, T.C. Lee, F.M. Lyng, H.J. Byrne, *Analytical Methods*, 2013, **5**, 2281-2291

6 R. Mehrotra, G. Tyagi, D. K. Jangir, R. Dawar, N. Gupta, *Journal of Ovarian Research* 2010, 3:27

7 K. Tanahashi, A. Natsume, F. Ohka, H. Momota, A. Kato, K. Motomura, N. Watabe, S. Muraishi, H. Nakahara, Y. Saito, I. Takeuchi, T. Wakabayashi, *BioMed Research International*, 2014, Article ID 860241

8 N. Amharref, A. Beljebbar, S. Dukic, L. Venteo, L. Schneider, M. Pluot, R. Vistelle, M. Manfait, *Biochimica et Biophysica Acta*, 2006, **1758**, 892–899

9 J. Desroches, M. Jermyn, K. Mok, C. Lemieux-Leduc, J. Mercier, K. St-Arnaud, K. Urmey, M. C. Guiot, E. Marple, K. Petrecca, F. Leblond, *Biomedical Optics Express*, 2015, **2380**, Vol. 6, No. 7

10 K. Gajjar, L. D. Heppenstall, W. Pang, K. M. Ashton, J. Trevisan, I. I. Patel, V. Llabjani, H. F. Stringfellow, P. L. Martin-Hirsch, T. Dawson, F. L. Martin, *Analytical Methods*, 2012; **5**, 89–102

11 N. Amharref, A. Beljebbar, S. Dukic, L. Venteo, L. Schneider, M. Pluot, M. Manfait, *Biochimica et Biophysica Acta*, 2007, **1768**, 2605–2615

12 Y. Zhou, C. H. Liu, Y. Sun, Y. Pu, S. Boydston-White, Y. Liu, R. R. Alfano, *Journal of Biomedical Optics*, 2012, **17(11)**, 116021

View Article Online
DOI: 10.1039/C9AY01200A

13 P. Heraud, S. Caine, N. Campanale, T. Karnezis, D. McNaughton, B. Wood, M. Tobin, C. Bernard, *Neuroimage*, 2010, **49**, 1180

14 A. Weselucha-Birczynska, J. Sacharz, J. Zieba-Palus, M. H. Lewandowski, R. Kowalski, K. Palus, L. Chrobok, M. Birczynska, A. Sozanska, *Vibrational Spectroscopy*, 2016, **85**, 48-54

15 J. Sacharz, A. Wesełucha-Birczyńska, J. Zięba-Palus, M. H. Lewandowski, R. Kowalski, K. Palus, Ł. Chrobok, P. Moskal, M. Birczyńska, A. Sozańska, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2018, **188**, 581–588

16 J. Sacharz, A. Wesełucha-Birczyńska, J. Zięba-Palus, M. H. Lewandowski, K. Palus-Chramiec, Ł. Chrobok, P. Moskal, M. Birczyńska, A. Sozańska, *Journal of Molecular Structure*, 2018, **1163**, 167–173

17 J. Zięba-Palus , A. Wesełucha-Birczyńska, J. Sacharz, M. H. Lewandowski, K. Palus, Ł. Chrobok, R. Kowalski, P. Moskal, M. Birczyńska, Agnieszka Sozańska, *Journal of Molecular Structure*, 2017, **1147**, 310–316

18 L. M. Miller, M. W. Bourassa, R. J. Smith, *Biochimica et Biophysica Acta*, 2013, **1828(10)**, 2339–2346

19 O. S. Hajjawi, *American Journal of BioScience,* 2014; **2(4)**, 122-134

20 C. Krafft, L. Neudert, T. Simat, R. Salzer*, Spectrochimica Acta Part A*, 2005, **61**, 1529–1535

21 J. S. O'Brien, E. L. Sampson, *Journal of Lipid Research*, 1965, **6**, 545-551

22 M. J. Hackett, C. J. Britz, P. G. Paterson, H. Nichol, I. J. Pickering, G. N. George, *ACS Chemical Neuroscience,* 2015, **6**, 226−238

23 P. Lasch, D. Naumann, *Biochimica et Biophysica Acta*, 2006, **1758**, 814–829

24 B. R. Wood, D. McNaughton, in *Spectrochemical Analysis using Infrared Multichannel Detectors,* ed. Rohit Bhargava, Ira W. Levin, Blackwell Publishing Ltd, 2005, chapter 10, 204–233

View Article Online
DOI: 10.1039/C9AY01200A

25 A. D. Surowka, M. Pilling, A. Henderson, H. Boutin, L. Christie, M. Szczerbowska-Boruchowska, P. Gardner, *Analyst*, 2017, **142**, 156-168

26 S. Tiwari, R. Bhargava, *Yale Journal of Biology and Medicine,* 2015, **88**, 131-143

27 P. Bassan M. J. Weida, J. Rowlette, P. Gardner, *Analyst*, 2014, **139**, 3856-3859

28 C. Hughes, L. Gaunt, M. Brown, N. W. Clarke, P. Gardner, *Anal. Methods*, 2014, **6**, 1028-1035

29 A. Zwielly, S. Mordechai, I. Sinielnikov, A. Salman, E. Bogomolny, S. Argov, *American Association of Physicists in Medicine*, *Medical Physics,* 2010, **37 (3)**, 1047-1055

30 R. Bhargava, *Anal Bioanal Chem*, 2007, **389**, 1155–1169

31 S. Mittal, K. Yeha, L. Suzanne Leslie, S. Kenkel, A. Kajdacsy-Balla, R. Bhargava, *PNAS*, 2018, **115 (25)**, E5651–E5660

32 A. Bouyanfif, S. Liyanage, E. Hequet, N. Moustaid-Moussa, N. Abidi, *Vibrational Spectroscopy,* 2018, **96**, 74–82

33 K. Laua, A. Hobro, T. Smith, T. Thurston, B. Lendl, *Vibrational Spectroscopy,* 2012, **60**, 34–42

34 A. J. Hobro, B. Lendl, *Vibrational Spectroscopy*, 2011, **57**, 213– 219

35 M. J. Turlough FitzGerald, G. Gruener, E. Mtui, in *Clinical Neuroanatomy and Neuroscience*, 6th edition, 2012, Saunders, Elsevier, chapter 6, 70-82 and 267-276, ISBN: 978-0-7020-3738-2

36 M. J. Turlough FitzGerald, G. Gruener, E. Mtui, in *Clinical Neuroanatomy and Neuroscience*, 6th edition, 2012, Saunders, Elsevier, chapter 25, 70-82 and 267-276, ISBN: 978-0-7020-3738-2

Figure legends

Figure 1. a) Image of Hematoxylin and eosin (H&E) stained section; b) Classes assigned to different regions of the brain tissue: white matter (red), granular layer (yellow) and molecular layer (light blue). The violet is assigned to the holes in the tissue; c) Standard normal variate (SNV) transformation normalized average spectra of the classes assigned to *vide* b)

Figure 1. a) Image of Hematoxylin and eosin (H&E) stained section; b) Classes assigned to different regions of the brain tissue: white matter (red), granular layer (yellow) and molecular layer (light blue). The violet is assigned to the holes in the tissue; c) Second derivatives of the average spectra of the classes assigned to *vide* b)

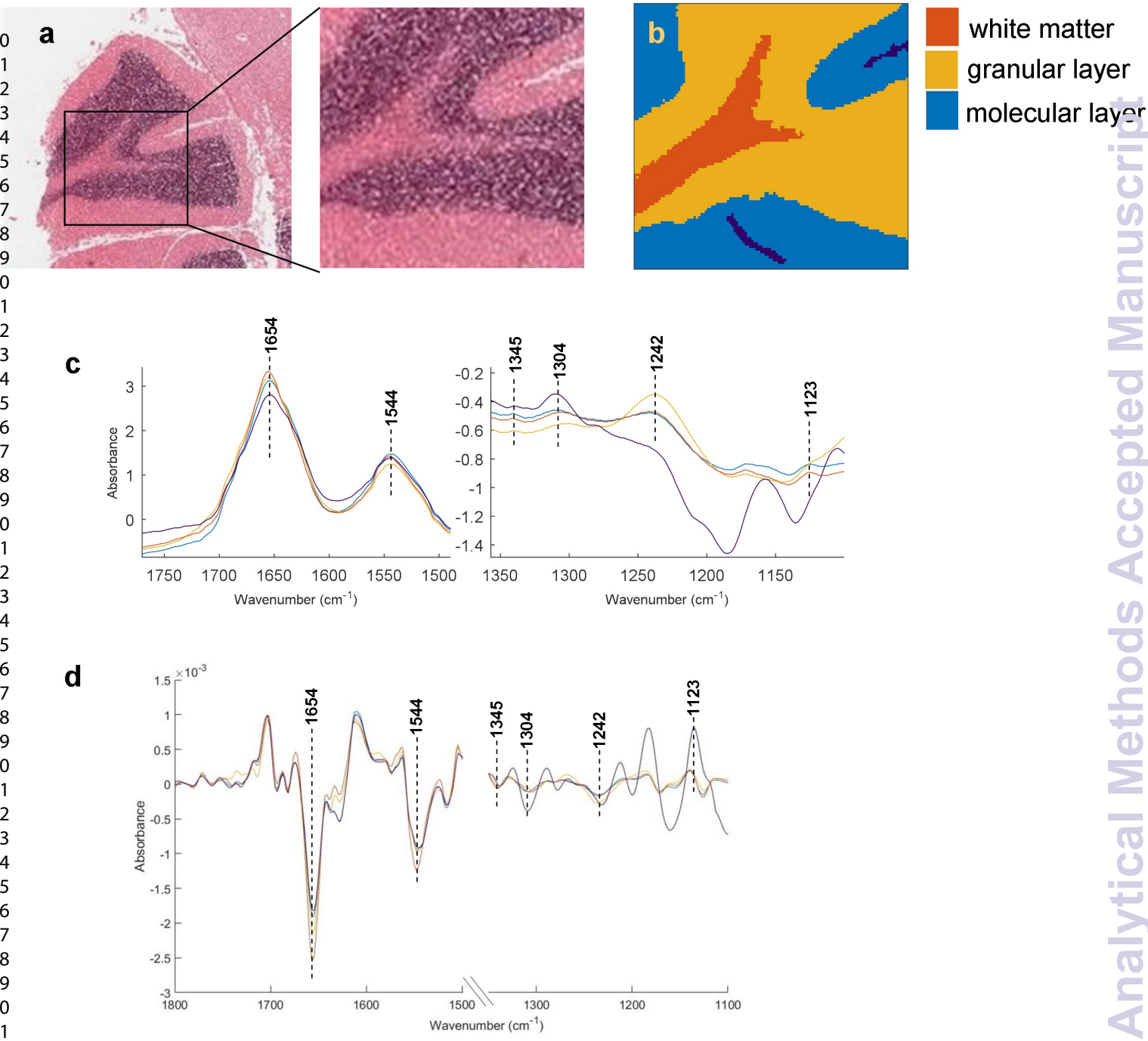
Figure 2. Cluster images obtained from the hyperspectral images acquired using objectives of different magnification, spectral resolution and number of scans. The results from the most accurate k-means cluster considering (4-7) classes is shown. See experimental section for more information.

Figure 3. Classification accuracy of the k-means clustering from the experimental brain images acquired at different magnification, resolution and number of scans.

Figure 4. Precision of the classification of the different magnification, resolution and number of scans. Precision was calculated as the percentage of true positives over the sum of true positives and false positives, in terms of correct classification of spectra to one of the four classes (white matter, molecular layer, granular layer or hole in the tissue) when compared to that position in the reference map.

Figure 5. Noise levels of the spectra of the hyperspectral images with and without PCA noise reduction. For calculating the noise, only 1% of the spectra in each image were considered. The hyperespectral matrix was reshaped as a column of spectra and ever 100^{th} spectrum was selected-. For each selected spectrum, the region between 1750 and 1800 cm-1 was baseline corrected ("Automatic Weighted Least Squares", polynomial order=1) and the standard deviation of the absorbance at this corrected region is considered as noise. The overall noise value was the mean of noise for all the selected spectra.

View Article Online
DOI: 10.1039/C9AY01200A



View Article Online
DOI: 10.1039/C9AY01200A

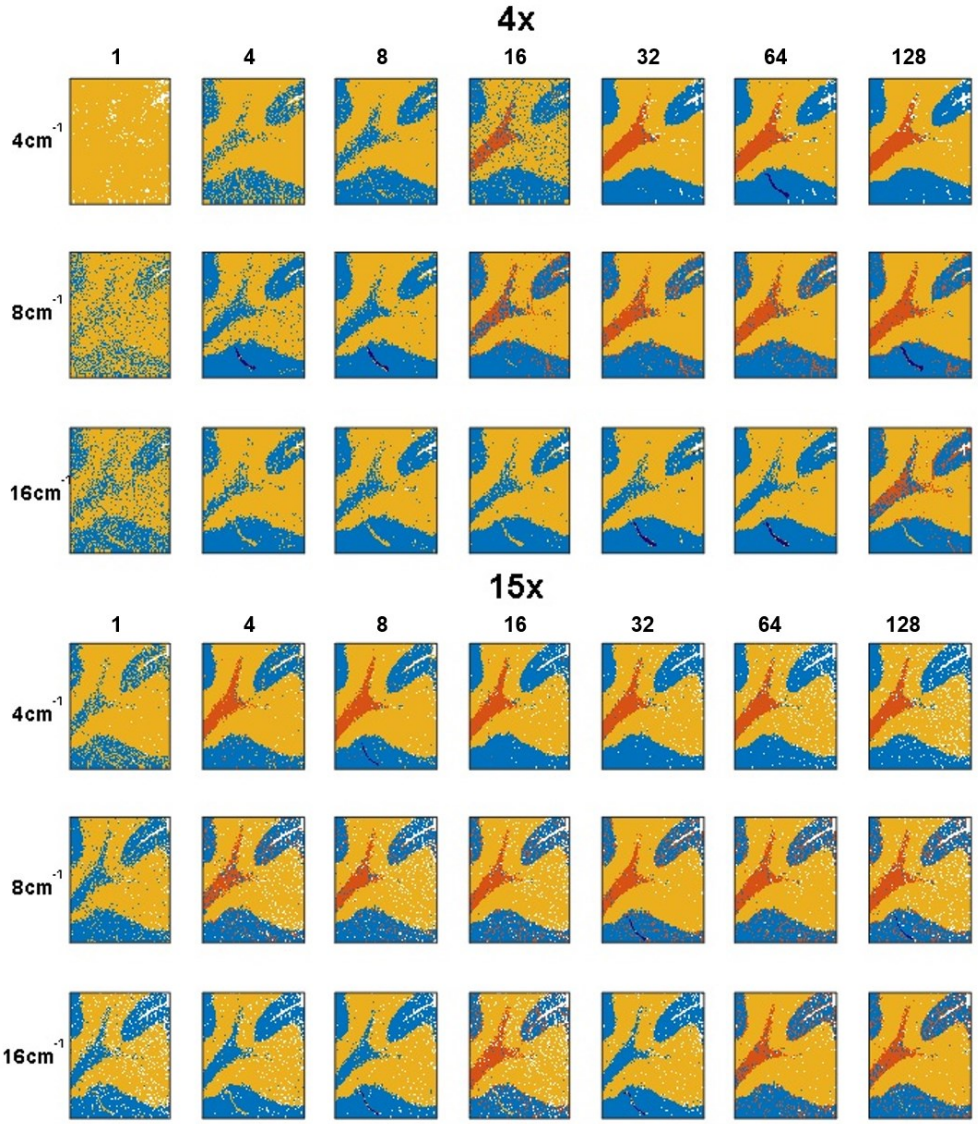


Figure 2. Cluster images obtained from the hyperspectral images acquired using objectives of different magnification, spectral resolution and number of scans. The results from the most accurate k-means cluster considering (4-7) classes is shown. See experimental section for more information.

183x217mm (150 x 150 DPI)

View Article Online
DOI: 10.1039/C9AY01200A

Figure 3.

Accuracy (%)		Numer of scans:						
Magnification	Resolution (cm^{-1})	1	4	8	16	32	64	128
4x	4	55,01	77,45	82,41	84,18	95,57	95,38	96,10
	8	63,89	82,32	84,49	85,19	87,93	87,75	91,61
	16	70,96	82,56	84,00	84,85	85,36	85,38	86,35
15x	4	77,82	91,71	94,00	94,73	95,23	96,31	95,89
	8	82,44	88,75	89,62	89,41	87,74	88,54	90,12
	16	85,95	87,56	87,69	88,76	87,98	87,74	87,74

Noise ($x10^{-3}$)		Numer of scans:						
Magnification	Resolution (cm^{-1})	1	4	8	16	32	64	128
4x	4	7,90	4,15	3,17	3,40	2,04	1,86	1,93
	8	4,87	2,56	1,77	1,79	1,45	1,38	1,52
	16	3,01	1,98	1,63	1,43	1,45	1,31	1,37
15x	4	6,85	3,76	2,87	2,46	1,93	2,82	3,14
	8	4,54	2,58	2,11	1,57	1,67	1,26	1,55
	16	2,37	1,69	1,41	1,54	1,21	1,07	1,07

Time (min)		Numer of scans:						
Magnification	Resolution (cm^{-1})	1	4	8	16	32	64	128
4x	4	<1	<1	~1	1,5	2,5	4,5	9
	8	<1	<1	~1	1	2	3	5
	16	<1	<1	~1	~1	1,5	2	3,5
15x	4	7	9	12	18	31	54	106
	8	5	6	7	10	17	31	58
	16	4	4	5	7	11	19	35

View Article Online
DOI: 10.1039/C9AY01200A

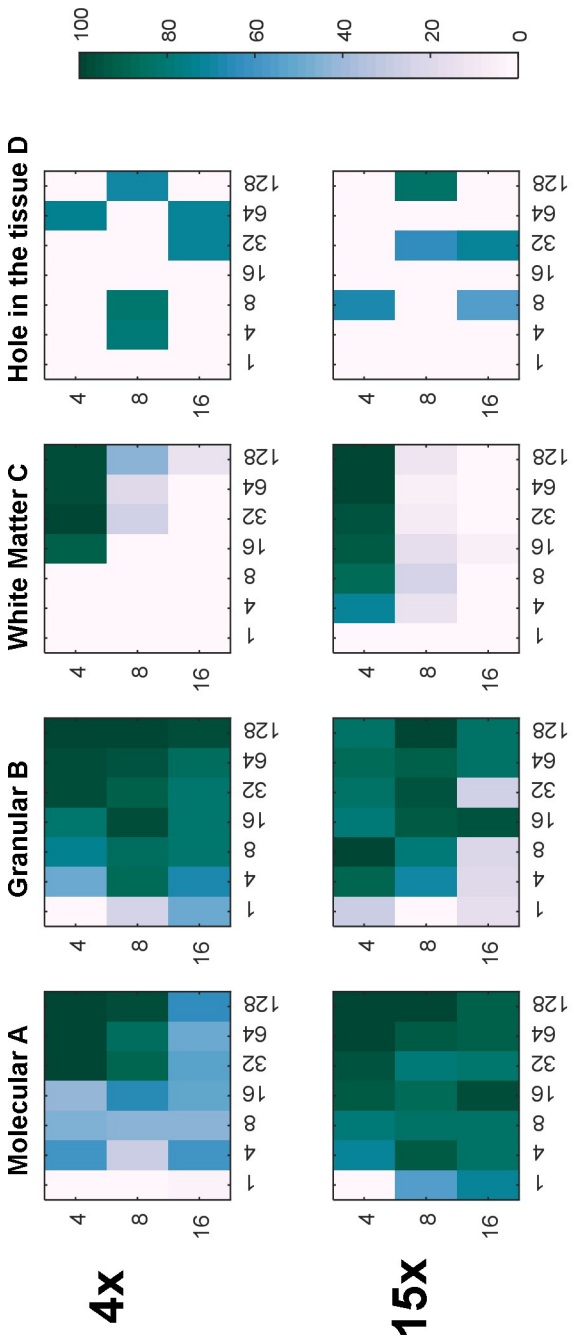
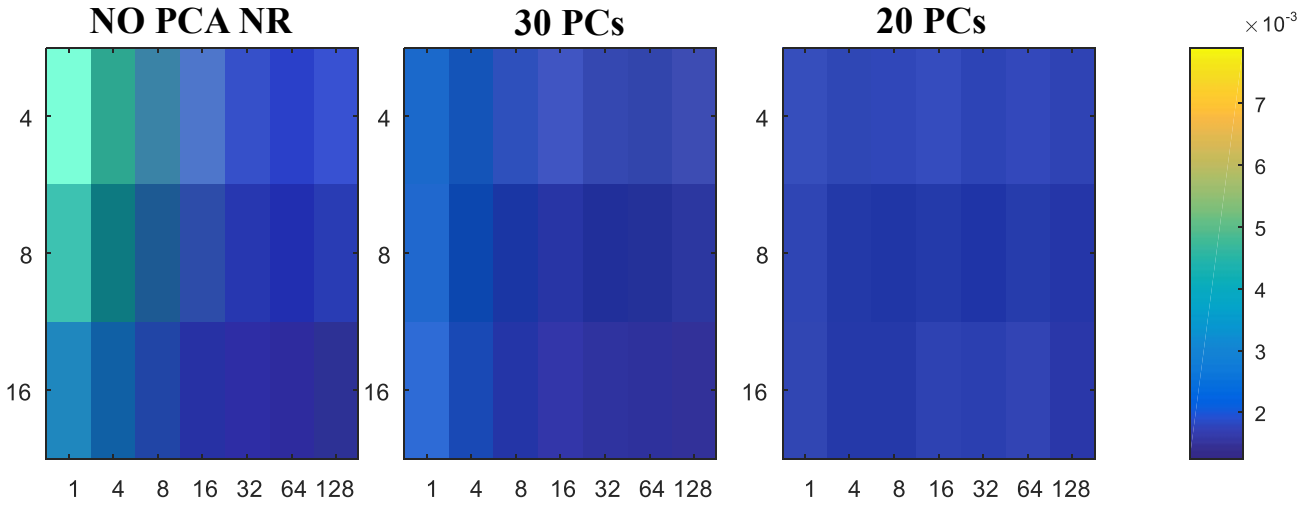
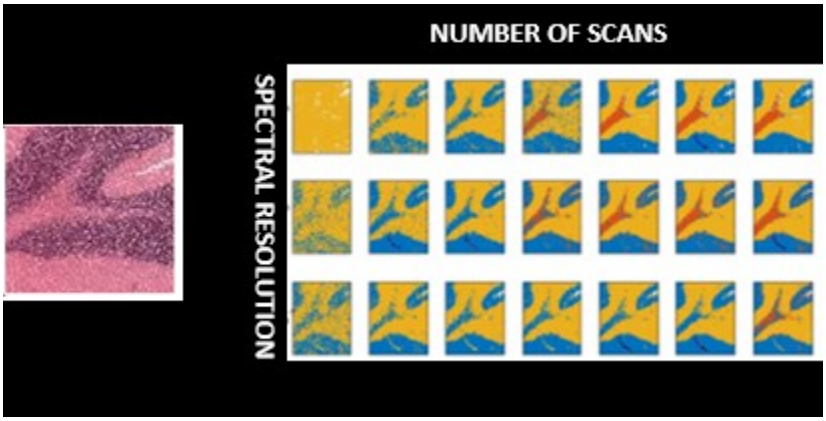


Figure 5

View Article Online
DOI: 10.1039/C9AY01200A



View Article Online
DOI: 10.1039/C9AY01200A



A method to determine the optimal measurement parameters for FTIR imaging