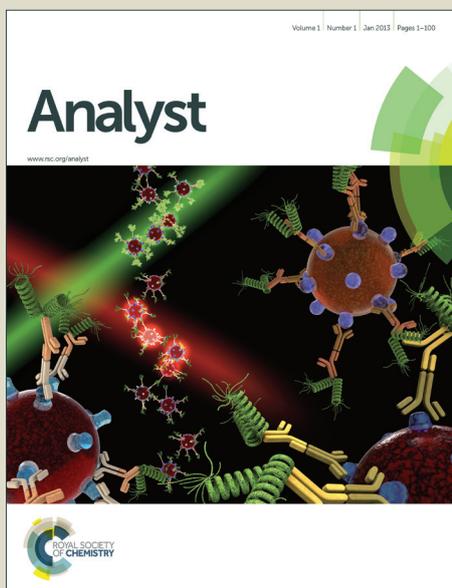


Analyst

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

Comparison of FTIR transmission and transfection substrates for canine liver cancer detection

Cite this: DOI: 10.1039/x0xx00000x

Kamila Kochan^{a,b}, Philip Heraud^{c,d}, Matti Kiupel^e, Vilma Yuzbasivan-Gurkan^f, Don McNaughton^c, Malgorzata Baranska^{a,b} and Bayden R. Wood^{*,c}.

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

FTIR spectroscopy is a widely used technique that provides insights into disease processes at the molecular level. Due to its numerous advantages it is becoming an increasingly powerful tool for the study of biological materials and has the potential to become an excellent diagnostic method, especially considering the low cost of transfection substrates. However, questions about the usefulness of the transfection measurement mode due to the complicated nature of physical processes occurring during the measurement and in particular the Electric Field Standing Wave (EFSW) effect have been raised. In this paper we present a comparison of the two most common FT-IR measurement modes: transmission and transfection using healthy and pathologically altered tissue (histiocytic sarcoma). We found the major differences between normal and cancerous tissue were associated with changes DNA and carbohydrate content. In particular we identified a band at 964 cm⁻¹ assigned to a nucleic acid phosphodiester backbone mode, which appeared more pronounced in cancerous tissue irrespective of the substrate. We applied Principal Component Analysis, Unsupervised Hierarchical Cluster Analysis and *k-means clustering* to transmission and transfection substrates and found that both measurement modes were equally capable of discrimination normal from cancerous tissue. Moreover, the differences between spectra from cancerous and normal tissue were significantly more important than the ones arising from the measurement modes.

Introduction

FTIR spectroscopy is becoming a widely used technique that provides insights into disease processes at the molecular level. It has been applied in numerous studies of various disease mechanisms including atherosclerosis¹⁻², diabetes³⁻⁴, Alzheimer's disease⁵⁻⁶, Parkinson disease and Huntington's disease⁶. However, by far the greatest potential of this technology lies in the study of cancer⁷⁻¹⁹ and numerous neoplastic entities in various organs including lungs⁷, stomach⁸, cervix⁹⁻¹¹, breast¹²⁻¹³, prostate¹⁴⁻¹⁵ and liver¹⁶⁻¹⁸ have previously been investigated.

Histopathology has been the gold standard to diagnose and prognosticate most animal and human cancers. While the morphologic phenotype allows for standardisation of nomenclature and accurate prognosis of disease progression within a population, morphologic assessment alone is often insufficient to accurately determine the biological behavior for an individual patient and/or to predict individual therapeutic response. Since most biological molecules have IR fingerprint spectral characteristics, this technology can be utilised to

identify chemical changes at a molecular level that are unique for a disease process in an individual patient.

The ability to identify chemical changes at the molecular level in formalin fixed, paraffin embedded tissue opens new opportunities for clinical studies and biomarker discovery using archival biopsy material. While formalin fixation of tissues results in rapid preservation of histological detail by creating cross-links between amino acid residues, the insoluble cross-links also interfere with efficient biomolecule analysis. Furthermore, formalin fixation fragments including nucleic acids at approximately 400 base pairs significantly limit classical RNA or DNA analysis tools²⁰. FTIR spectroscopy can be used to simultaneously detect a variety of macro-molecules in formalin fixed sections and measurements can be acquired within seconds¹¹. Due to its numerous advantages, such as the non-destructivity of the measurement, lack of any complicated sample preparation, and the possibility to obtain complex information it is becoming increasingly popular for the study of biological materials. When combined with histochemical methods, FITR may provide a rapid and relatively simple tool for the diagnosis and prognosis of neoplastic diseases.

1 Whilst the FTIR technique has numerous advantages for
2 studying biological materials there are some considerations that
3 need to be addressed before it can become a routine clinical
4 tool. In this context, the choice of substrate on which to mount
5 the tissue has become a hot topic in recent years, whilst there is
6 also debate over whether to deparaffinise tissue sections prior
7 to spectroscopy. The two most common FTIR spectroscopy
8 measurement modes are transmission and transfection – the
9 latter makes use of low cost Mirr-IR substrates (Kevley
10 Technologies, Ohio), while the former generally requires more
11 expensive infrared transparent crystalline window materials
12 including CaF₂, ZnSe, BaCl₂ and KRS-5. For diagnostic
13 applications the lower costs of transfection substrates, which
14 are the same length and width as conventional microscope
15 slides, and so can be prepared easily in a standard biomedical
16 environment have become the substrates of choice for FTIR
17 analysis of tissue sections.

18
19 In transfection mode the light beam passes through the
20 sample twice. In theory, this should result in the doubling of the
21 optical path length in comparison to a transmission
22 measurement. However, non-linearity can arise in the
23 measurement through a physical phenomenon, the electric-field
24 standing wave (EFSW)²¹⁻²³, that results from interference
25 between the incident and reflected wave. It has been suggested,
26 that due to the inhomogeneity of the electric field created by
27 a multilayer system (representing a model of a realistic sample),
28 the spectral intensities measured in transfection mode are
29 distorted non-linearly to such an extent, that their usefulness for
30 disease diagnosis has become highly questionable²³. However,
31 in a more recent study²⁴ it has been shown, that for samples of
32 certain thickness (~ 5 μm) the sample inhomogeneity along
33 with factors such as source incoherence and range of angles of
34 incidence have an averaging effect on EFSW phenomenon,
35 significantly reducing its influence on the spectrum, especially
36 in the finger-print region of the FTIR spectrum.²⁴ Since the
37 phenomena is thought to be most problematic when there are
38 differences in the thickness of compared samples and reduced
39 when using 2nd derivative spectra, the first strategy is to ensure
40 consistent thickness for all samples as done in the recent work
41 of Cao *et. al* where cell deposit thicknesses were checked using
42 AFM and 2nd derivatives were used for analysis.²⁵

43
44 Here, we present a comparison between the two most
45 common FTIR measurement modes: transmission and
46 transfection using Mirr-IR slides and CaF₂ windows,
47 respectively. In order to minimise the impact of EFSW
48 phenomena on the outcomes of analysis we have ensured a
49 consistent thickness by using the same sectioning process and
50 operator to section the material. We demonstrate that in the
51 case of a canine liver cancer (histiocytic sarcoma) for both:
52 paraffin embedded and paraffin removed tissue the diagnostic
53 utility of FTIR is not compromised by using Mirr-IR
54 transfection substrates.

Experimental

FTIR images were obtained from the livers of 7 canines with identified advanced histiocytic sarcoma. From each liver three samples were prepared from adjacent tissue sections: for FTIR measurements in (1) transmission and (2) transfection mode, as well as for (3) H&E staining. Adjacent sections of paraffin embedded liver sections were cut into slices of 4 μm and 8 μm thicknesses and placed on Mirr-IR substrates (Kevley Technologies, Ohio) and CaF₂ slides, respectively. The diseased areas were identified on the basis of H&E staining, performed on the third, neighboring section. Particular attention was paid to the imaging of the same (corresponding) fragments of the same sample on adjacent tissue sections prepared for transmission and transfection mode. FTIR images were collected using an Agilent FTIR microscope (model 600 UMA, Agilent), equipped with a 64 × 64 pixel HgCdTe Focal Plane Array (FPA) liquid nitrogen cooled detector and a 15× Cassegrain objective. Maps were collected with a pixel binning of 4. Spectra were collected in the range from 4000 – 900 cm⁻¹ with spectral resolution 8 cm⁻¹ and with 64 interferograms co-added. The size of the imaged area differed between samples due to different size of the tumours, covering a range from relatively small (4×4) to large mosaics (10×15). All spectra were collected from samples at room temperature. Additionally, in order to verify the results, the paraffin was removed by washing three times in clean xylene (3×5 min) and the measurements were performed again on the deparaffined sections, with particular attention paid to image the areas corresponding to those obtained prior to removal of the paraffin.

Cytospec (Version 2.0) and The Unscrambler (Version 10.3) software packages were used for the data analysis. *K-means* cluster (KMC) analysis was performed on single maps, after removing bad pixels and inadequate S/N spectra using a quality test (based on the intensity of amide I band). The 2nd derivatives were calculated using the Savitzky–Golay algorithm with 9 smoothing points and normalised to the amide I mode. To investigate inter-sample variability several spectra corresponding to areas of healthy and cancer regions, from each of the 7 different canines were obtained according to the following procedure. Approximately 1000 unmodified spectra were extracted from each map from samples from different animals, from areas corresponding to tumour and normal tissue. The selection of areas was determined by comparing the results with staining and only areas unambiguously defined as cancerous and normal were analyzed. To improve the quality of spectra and thereby minimise the preprocessing procedure, each of the spectra subjected for further analysis was obtained by averaging 50 single spectra extracted directly from the FTIR map. Therefore, approximately 20 spectra from each liver: 10 corresponding to normal and 10 corresponding to cancerous tissue were obtained. From these spectral averages three data sets for paraffin embedded section as well as three for deparaffined sections were created. Three data sets, each containing spectra corresponding to normal and tumour tissue,

were analyzed separately: first one included only spectra measured in transmission; second, only in transfection; and third, in both modes. Data preprocessing was performed for all spectra simultaneously and included baseline correction, normalisation, calculation of 2nd derivatives (Savitzky – Golay algorithm, 9 smoothing points) and Extended Multiplicative Scatter Correction (EMSC). Unsupervised Hierarchical Cluster Analysis (UHCA) and Principal Component Analysis (PCA) were performed in different spectral ranges, including: 1350-900 cm⁻¹ and 1680-1530 cm⁻¹ for paraffin embedded sections and 1700-900 cm⁻¹ as well as reduced regions including 1680-1530 cm⁻¹ and 1350-900 cm⁻¹ for deparaffined sections.

The ratios of intensities of amide I to amide A were calculated based on the integrated area under the bands using Opus Software (Version 7.0) for all spectra before paraffin removal.

Results and discussion

Transmission and transfection for cancer detection: paraffin embedded sections.

For all FTIR images recorded in both transmission and transfection modes efforts were made to ensure the same regions of tissue in adjacent sections were measured. Appropriate areas of tumour and normal tissue were chosen by comparing the H&E stained sections with the FTIR images. k – means cluster (KMC) analysis was performed on the 2nd derivatives (Fig 1.).

Because of the presence of strong paraffin bands at ~1450 cm⁻¹, two spectral ranges were considered: 1350-900 cm⁻¹ and 1680-1530 cm⁻¹. Since the signals from the paraffin are present in both, tumour and normal tissue areas, which result from the sample preparation rather than from the natural composition of the tissue, they are not useful to draw conclusions about the biochemical differences. KMC analysis was performed using both ranges, respectively, and each of them separately. It was found that for the paraffin containing tissue sections the range 1350-900 cm⁻¹ was more useful in identifying tumourous regions.

As shown in Fig. 1 areas corresponding to tumours are clearly distinguished (red class). The observed bands along with their assignments are tabulated in Table 1. The blue class corresponds to spectra near the border of the tissue, which is characterised by weak absorbance and light dispersion and consequently these spectra have been excluded from the ensuing analysis and are not presented in Fig.1D. The red class delineates areas with clearly visible and advanced tumour. The grey class, on the other hand, corresponds to normal tissue. The green class – not obviously distinguishable using H&E staining – exhibits spectra that bear a striking similarity to the cancerous tissue (red class).

In comparison to normal liver tissue, spectra corresponding to tumour areas (red class) can be characterised by clearly smaller intensity of bands at 930, 997, 1026, 1049, 1082, 1155 cm⁻¹. Those bands correspond to C-O stretching modes from carbohydrates (*i.a.* at 930, 997, 1049 cm⁻¹), particularly glycogen (1026, 1082, 1155 cm⁻¹). It should be noted that

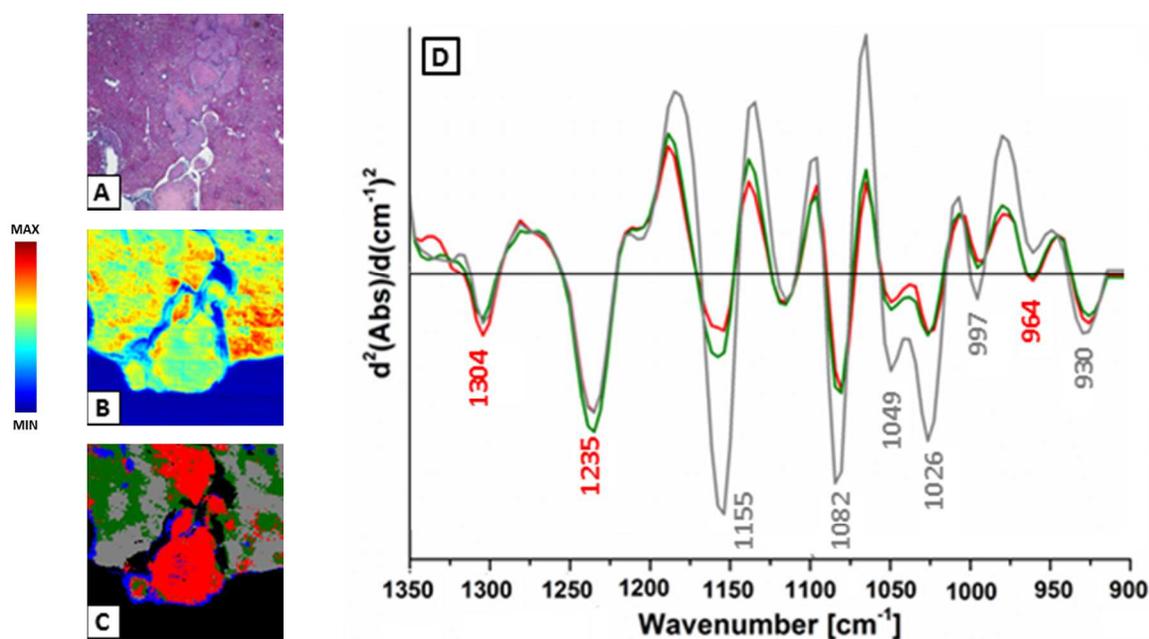


Fig. 1. An example of liver tissue section with histiocytic sarcoma measured in transfection: (A) adjacent section stained with H&E clearly showing the lighter stained cancer region, (B) map of the distribution of proteins based on the integration in the range 1680-1620 cm⁻¹ for the measured area and (C), (D) results of KMC analysis in the range 1350-900 cm⁻¹: (C) distribution of classes along with (D) corresponding spectra. Colour scale bars for the distribution map based on band integration are presented on the left side (MAX = 51.68, MIN = 0.64). More examples of cancer area identification along with examples of correlation of the measured areas in transmission and transfection are presented in supplementary (Fig. S1).

glycogen concentration varies extensively over a section of liver tissue and is not considered a clear marker for cancer. The other major difference is the higher intensity of the band at 964 cm^{-1} in tumour spectra assigned to the $\nu(\text{P-O-C})$ vibration

Table 1. Position of the most important bands observed in the IR spectra of the liver, along with their assignment.

Band position [cm^{-1}]	Assignment	Chemical compounds
930	$\nu(\text{C-O})$	Polysaccharides
964	$\nu(\text{P-O-C})$	nucleic acids; DNA&RNA
997	$\nu(\text{C-O})$	Polysaccharides
1026	$\nu(\text{C-O})$	Glycogen
1049	$\nu_s(\text{CO-O-C})$	Carbohydrates
		Glycogen
1082	$\nu_s(\text{PO}_2^-)$;	nucleic acids; DNA&RNA
1155	$\nu(\text{C-O})$	Glycogen
1235	$\nu_{as}(\text{PO}_2^-)$	nucleic acids; DNA&RNA
1304	$\delta(\text{CH}_2)$	Proteins

from the phosphodiester backbone of nucleic acids. Interestingly, this is the only nucleic acid band that seems to be more intense for tumour spectra. It is, however, important to note at this point, the complex nature of other bands corresponding to DNA and RNA. The band at 1083 cm^{-1} has contributions from both glycogen and nucleic acids because of the significantly higher glycogen content in normal tissue obscuring the nucleic acid vibrations. Similarly, in the case of the band at 1237 cm^{-1} , which in addition to nucleic acid bands, includes a contribution from phospholipids and proteins. Moreover, there were no trends in specific changes in the location of amide bands in the case of this type of cancer for paraffin embedded tissue. The amide I band for both normal and cancerous spectra was located $\sim 1650 \text{ cm}^{-1}$, which generally corresponds to the α -helical structure of proteins.

All of the differences observed between the red (cancerous tissue) and the gray (normal tissue) class are also found in the comparison between green and gray classes. Only minor differences exist between the band intensities in the green and red classes. This suggests that the green class corresponds to an area already pathologically altered, although these changes are less distinguishable with H&E staining.

Figure 1 shows an example analysis of a single map measured in transfection mode. However, similar results were obtained in the analysis of all measured maps regardless of the measurement mode. Moreover, in the case of imaging of corresponding areas in transmission and transfection from samples originating from the same individual, the distribution of classes and their spectral characteristics were consistent. An example of this correlation for the KMC analysis along with comparisons of spectra from the corresponding classes in the two different modes of measurement are shown in the Supplementary Information (Fig.S2 and S3).

To investigate inter-sample variability spectra from 7 different canines were analyzed with UHCA and PCA. Spectra were grouped in three data sets (each containing normal and tumour spectra, measured only in transmission, only in transfection and in both modes jointly). For all of the data sets processed with UHCA two ranges 1350-900 cm^{-1} and 1680-1530 cm^{-1} were analyzed both separately and also combined. It was found that the range 1680-1530 cm^{-1} along with the combined ranges (1350-900 cm^{-1} and 1680-1530 cm^{-1}) could not be used to classify spectra as cancerous or normal from the three data sets (see supplementary, Fig. S4 – S6). However, the range 1350-900 cm^{-1} enabled the unambiguous separation of normal and cancerous spectra for all three data sets (Fig.2).

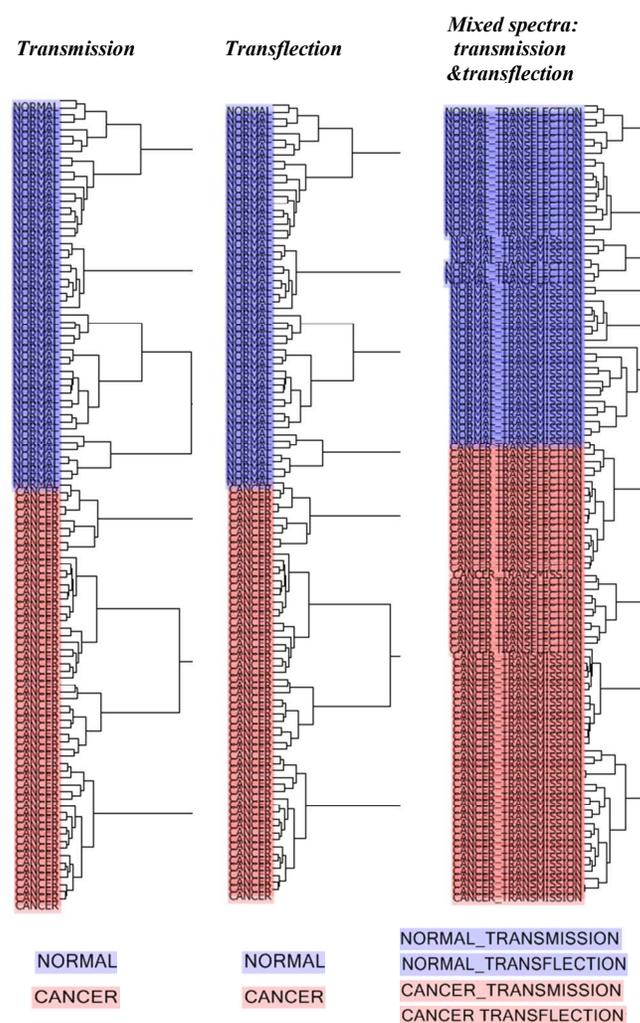


Fig.2. Unsupervised hierarchical cluster analysis (UHCA) in the range 1350-900 cm^{-1} showing the distinction between normal and cancerous spectra in three data sets: transmission (including normal and cancerous spectra measured in transmission), transfection (including normal and cancerous spectra measured in transfection) and mixed spectra (including normal and cancerous spectra measured in transmission and transfection) for paraffin embedded sections of liver tissue. Colour code for each dendrogram is provided below.

Similar results were achieved using PCA for transmission and transfection data sets (Fig.3) as well as for the mixed

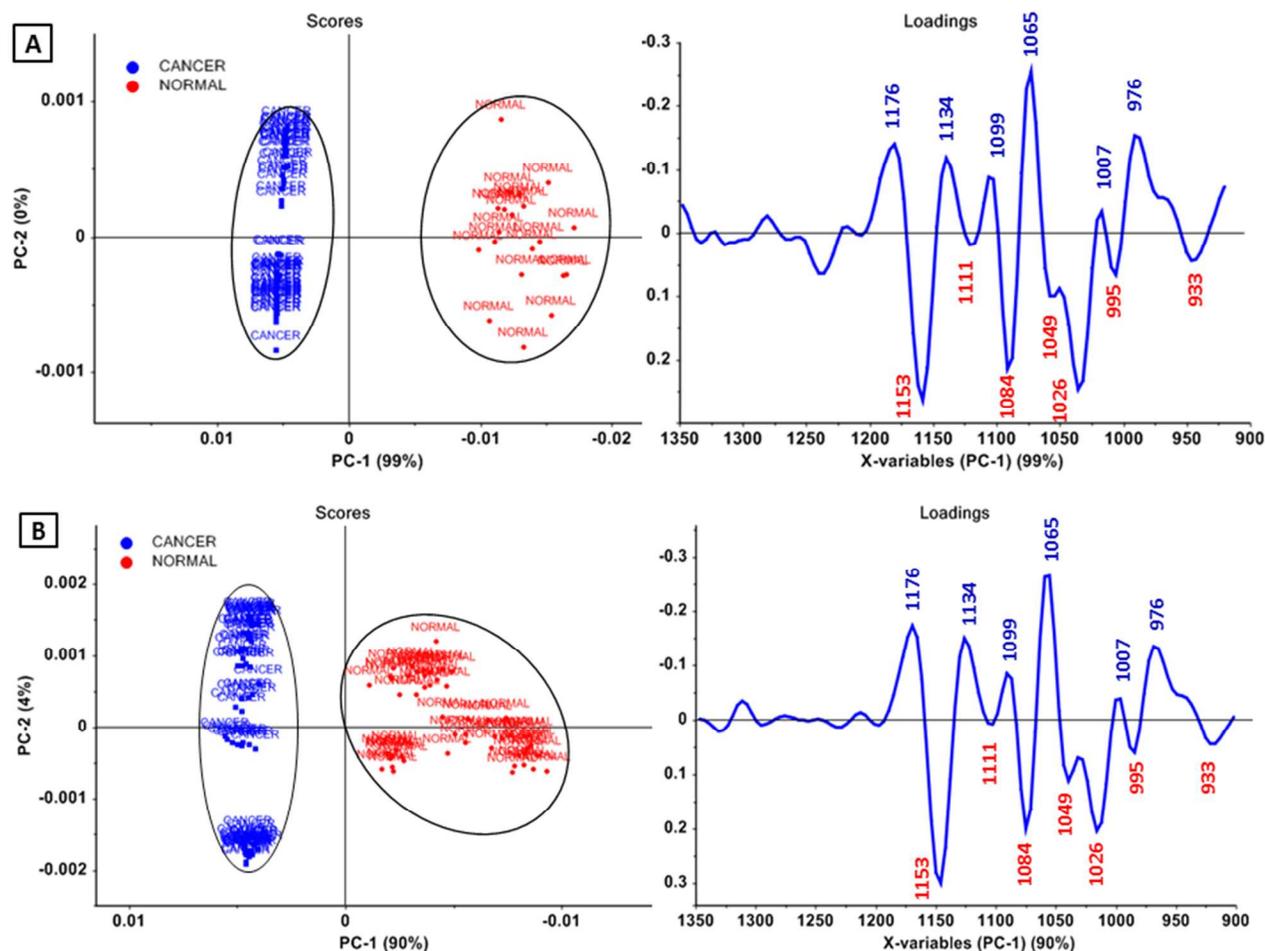


Fig.3. PCA results for spectra from healthy (marked as **NORMAL** in red) and cancer (marked as **CANCER** in blue) liver tissue, obtained in (A) transmission and (B) transfection mode: scores plot (PC1 vs PC2) and loading corresponding to PC1, in the range 1350-900 cm^{-1} . The black eclipses highlight the grouping effect (on all scores plots presented in this paper).

dataset (Fig.4). PCA, performed in the range 1350-900 cm^{-1} , allowed for a clear separation between cancerous and normal spectra. In both measurement modes (transmission and transfection – Fig.3) the separation occurred along PC1, and the corresponding loadings confirmed the previous findings. The normal areas of liver tissue can be characterised by higher content of carbohydrates (933, 995, 1049 cm^{-1}), in particular glycogen (1026, 1084, 1153 cm^{-1}), whereas tumour areas have more intense signals from nucleic acids (976 – 980, 1065 – 1068 cm^{-1}). Interestingly, in both measurement modes, the differences between tumour and normal spectra are clear and the spectral characteristics (expressed by loadings) determining the division are very similar.

In order to verify the impact of the measurement mode on the potential for the distinction of spectra, PCA in the range 1350-900 cm^{-1} was performed on all spectra simultaneously, including cancerous and normal ones, measured in both

transmission and transfection. The differences between the spectra in transmission and transfection in general can be easily seen, considering the whole measured spectral range (4000-900 cm^{-1}) by comparing the ratio of bands in the high wavenumber range (3100-2800 cm^{-1}) to the bands in fingerprint region (1350-900 cm^{-1}) (Fig.4D). For example, the ratio of amide I to amide A for the spectra measured in transmission mode is 0.35 (\pm 0.05), whereas for the spectra in transfection mode 0.21 (\pm 0.03). The mentioned ranges are separated from each other by at least 2500 cm^{-1} and consequently it shows the EFSW effect to be large when considering both ends of the spectrum together but minimal over a short wavenumber range although the high wavenumber signals are mainly paraffin based. Hence any effects are also minimised when performing classification restricted to small spectral regions (\sim 500 cm^{-1}) such as the region expanded in the figure.

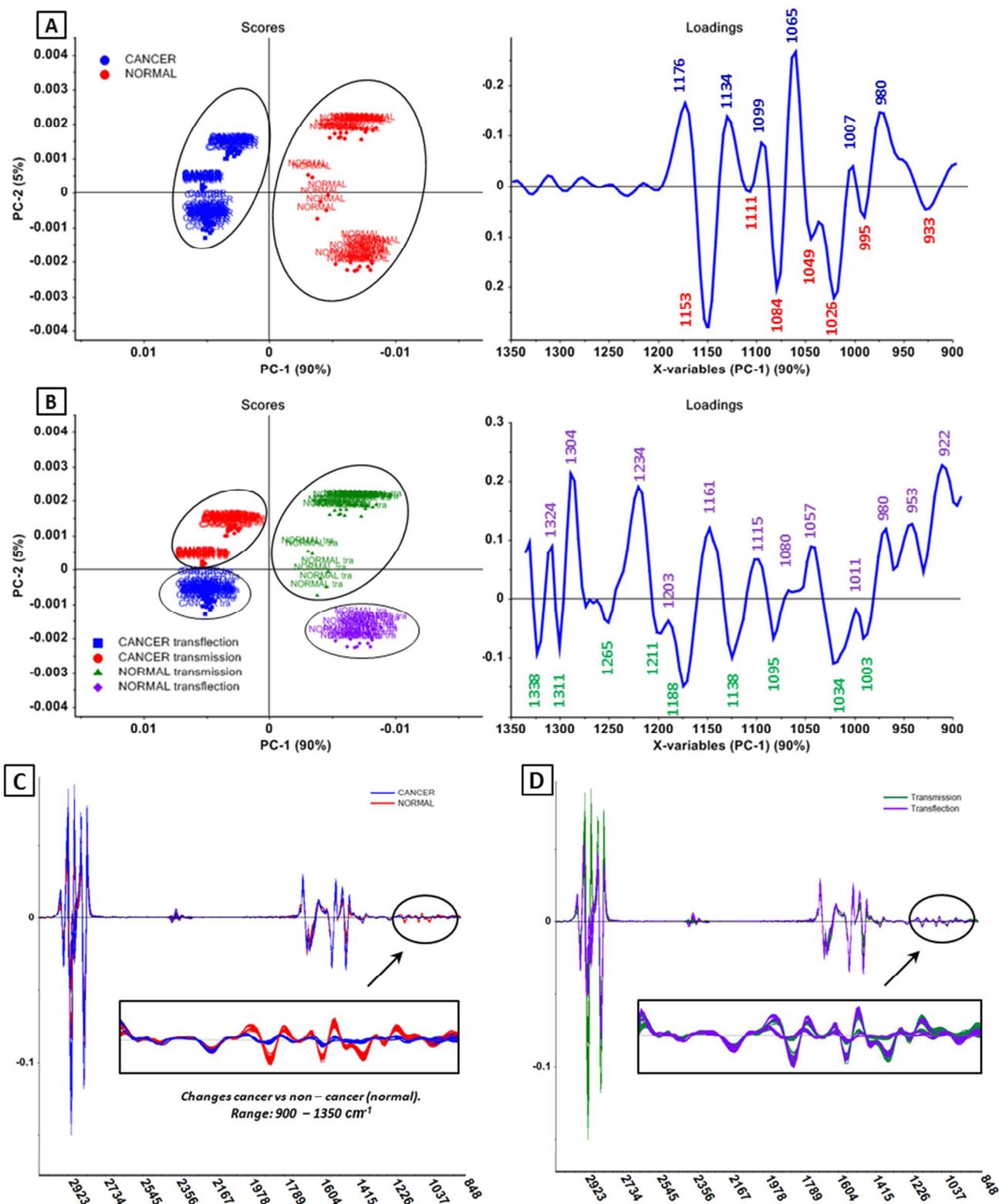


Fig.4. (A) PCA results based on spectra from healthy (marked as **NORMAL**) and cancer (marked as **CANCER**) liver tissue, obtained in transmission and transfection mode: scores plot (PC1 vs PC2) and loadings corresponding to PC1, in the range 1350-900 cm^{-1} ; (B) PCA results from the same sample set with applied grouping between measurement modes (**CANCER transfection**, **CANCER transmission**, **NORMAL transfection** and **NORMAL transmission**): scores plot (PC1 vs PC2) and loadings corresponding to PC2; all spectra from data set grouped according to: (C) area of the tissue from which they originate (**CANCER** and **NORMAL**) and (D) measurement mode (**Transmission** and **Transfection**), with highlighted area of differences.

As can be seen from the PCA results (Fig.4A, B), despite the differences in measurement modes the division between normal and cancerous spectra continues to be very clear and still occurs along PC1, explaining around 99 % of the variance. This indicates that the spectral features that change for cancer in this study are much more significant than the impact of physical effects such as EFSW related to transfection (Fig.4C,D). Just as important is the fact that the features that lead to separation remain the same, regardless of the measurement mode. Thus, both techniques are equally suitable for the detection of neoplastic change, when only a small spectral region is utilised.

However, this does not mean that the differences between measurement modes are invisible or entirely irrelevant. Although the main differentiating features (included in PC1) apply to the division between tumour and normal tissue, the subsequent ones – contained in PC2 – separate spectra according to the measuring mode. Especially the bands at 922, 1234 and 1304 cm^{-1} appear to be enhanced in the transfection mode. Similarly, in the UHCA results: although the main division occurs between classes: tumour and normal, within each of these classes spectra measured in transmission and transfection tend to be grouped together. This grouping is not perfect, but mixing of the spectra measured in the different modes occurs only to a small extent (Fig.2).

Transmission and transfection for cancer detection: sections after paraffin removal.

All of the measurements performed for the sections embedded in paraffin were repeated after paraffin removal. Once again the same regions on adjacent sections were analyzed from the different substrates. An example of a measured map corresponding to the one presented in Fig.1 along with KMC results, is shown in Fig.5.

As can be seen, KMC analysis allowed for a clear separation of tumour areas after paraffin removal. Moreover, the obtained distribution of classes corresponded to a large extent with the distribution of classes for maps measured before removing paraffin (supplementary, Fig 7S.). However, in this case, the spectral differences between cancer and non-cancerous tissue could be observed by simply integrating underneath the area of the amide I band. This reflects a different structure of the tumourous areas of the tissue, which could result from different tissue density. This change was not as visible for the paraffin embedded sections.

The same procedure of extracting and averaging spectra from cancerous and normal tissue was used to prepare sets of spectra obtained in the transmission and transfection as well as a set of combined spectra. UHCA and PCA in the range 1700-900 cm^{-1} as well as 1350-900 cm^{-1} , were performed on each of those sets. Again, the best discrimination was achieved for the 1350-900 cm^{-1} region. The results are presented on Fig. 6 and Fig.7.

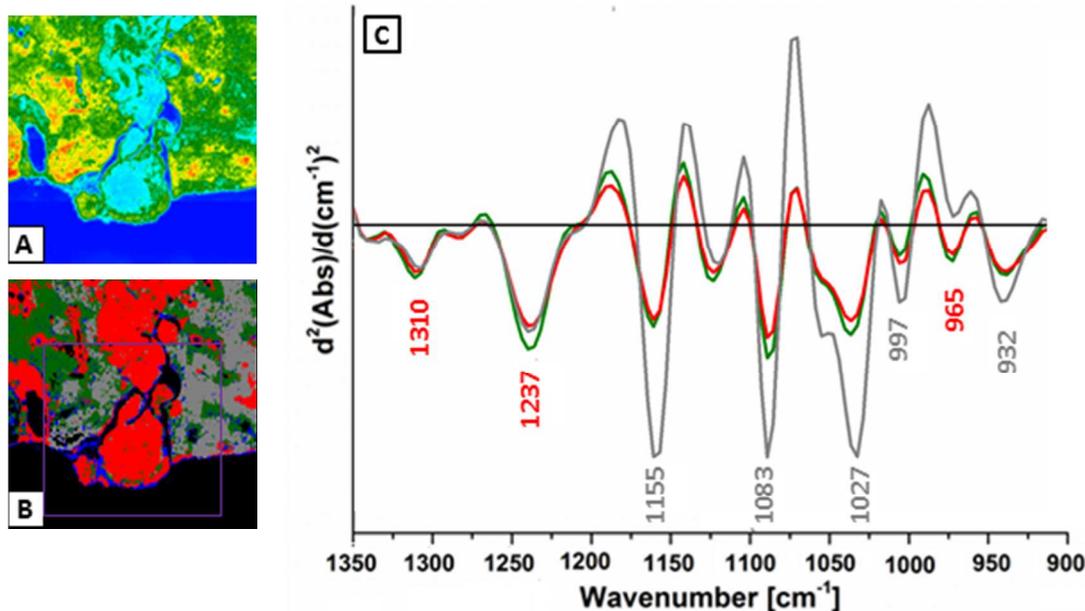


Fig. 5. An example of liver tissue section with histiocytic sarcoma, measured in transfection, after paraffin removal: (A) map of distribution of proteins based on the integrating the range 1680 – 1620 cm^{-1} (B) results of KMC in the range 1700 – 900 cm^{-1} showing the distribution of classes within the marked area measured before paraffin removal along with (C) corresponding spectra. An adjacent section stained with H&E showing a clearly distinguishable cancerous region is presented in Fig.1A. Colour scale bare for the map of distribution based on band integration is presented in Fig.1 (MAX = 77.42, MIN = 1.11) A direct comparison of two maps: before and after the removal of paraffin is presented in supplementary (Fig.7S).

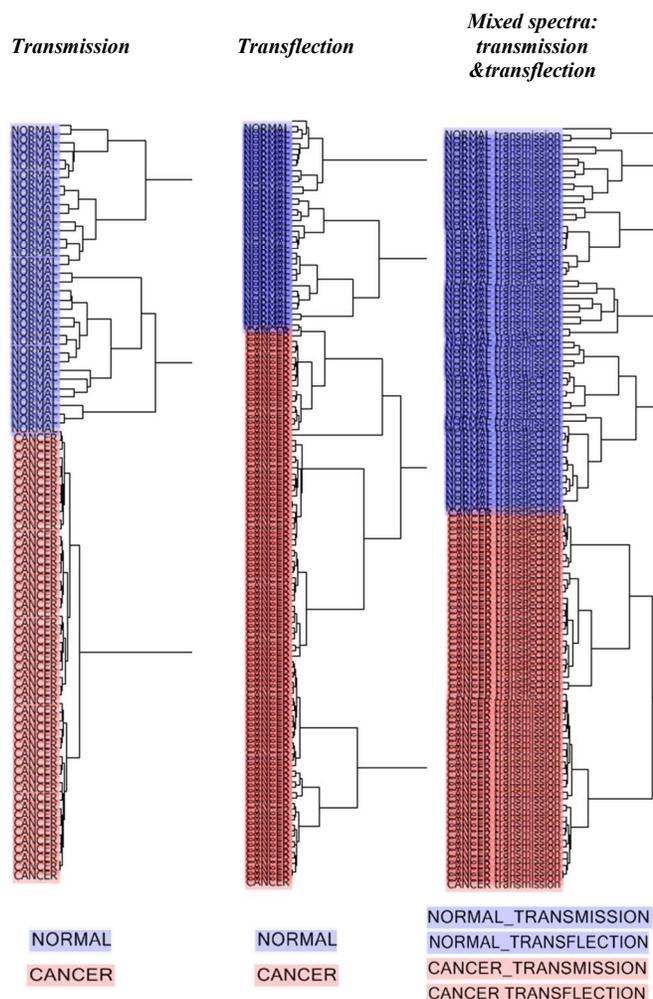


Fig.6. Unsupervised hierarchical cluster analysis (UHCA) in the range 1350-900 cm^{-1} showing the distinction between normal and cancerous spectra in three data sets: transmission (including normal and cancerous spectra measured in transmission), transflection (including normal and cancerous spectra measured in transflection) and combined spectra (including normal and cancerous spectra measured in transmission and transflection) after paraffin removal. Colour code for each dendrogram is provided below.

The UHCA performed in the range 1350-900 cm^{-1} allowed for 100% accurate discrimination of cancerous and normal spectra from all data sets. The spectral range (1700-900 cm^{-1}) failed in the case of transflection and mixed spectra, however, it

still allowed for distinction between transmission and transflection recorded spectra (supplementary, Fig. 8S). This difference between the measurement modes would seem to be a result of a clear difference in the intensity of bands from the range 1400-900 cm^{-1} relative to the bands in the range 1700-1500 cm^{-1} . For transmission measurements, the relative intensity (to amide I) of the bands in the range 1400-900 cm^{-1} is considerably higher than for transflection (supplementary, Fig. 8S). Therefore, those bands would have a greater influence on the analysis in the range 1700-900 cm^{-1} for transmission spectra most likely form the influence of the ESW effect.

Moreover, in the case of samples after paraffin removal, a shift in the amide I position in the spectra of cancerous tissue towards lower wavenumber is clearly visible. This shift occurred in both measurement modes (supplementary, Fig. 9S), from 1655 cm^{-1} (normal tissue) to 1651 cm^{-1} (cancerous tissue) in transmission, and from 1651 cm^{-1} to 1643 cm^{-1} , for transflection. For the transmission mode the shift was smaller, however, a pronounced shoulder band appeared at 1639 cm^{-1} . In the past such differences have often been attributed to changes in β -sheet structure but an alternative and more likely explanation is that the differences are the result of variation in the amount of bound water. Water has a band at \sim 1644 cm^{-1} which is largely obscured by the intense amide I band. Bound water can never be totally removed from proteins and would likely vary between cancerous and normal tissue and certainly would change after deparaffinisation. The differences in the bound water content between normal and cancerous tissue can be seen even more clearly in the O – H stretching region (3600 – 3500 cm^{-1}) (see Supporting Information).

PCA results (Fig. 7,8) confirm the previous findings for paraffin embedded sections (Fig.4). The separation between the spectra of normal and tumour tissue always occurs along PC1, in both measurement modes analyzed separately and in combination. The discriminatory spectral characteristics remain the same. In the case of the combined spectral regions the differentiation with respect to the measurement mode occurs along PC2, however, when compared to the samples in paraffin the separation is based on the band at 1036 cm^{-1} , which is a glycogen marker band, which is not an ideal marker band for cancer detection due to variation in glycogen content in liver tissue.

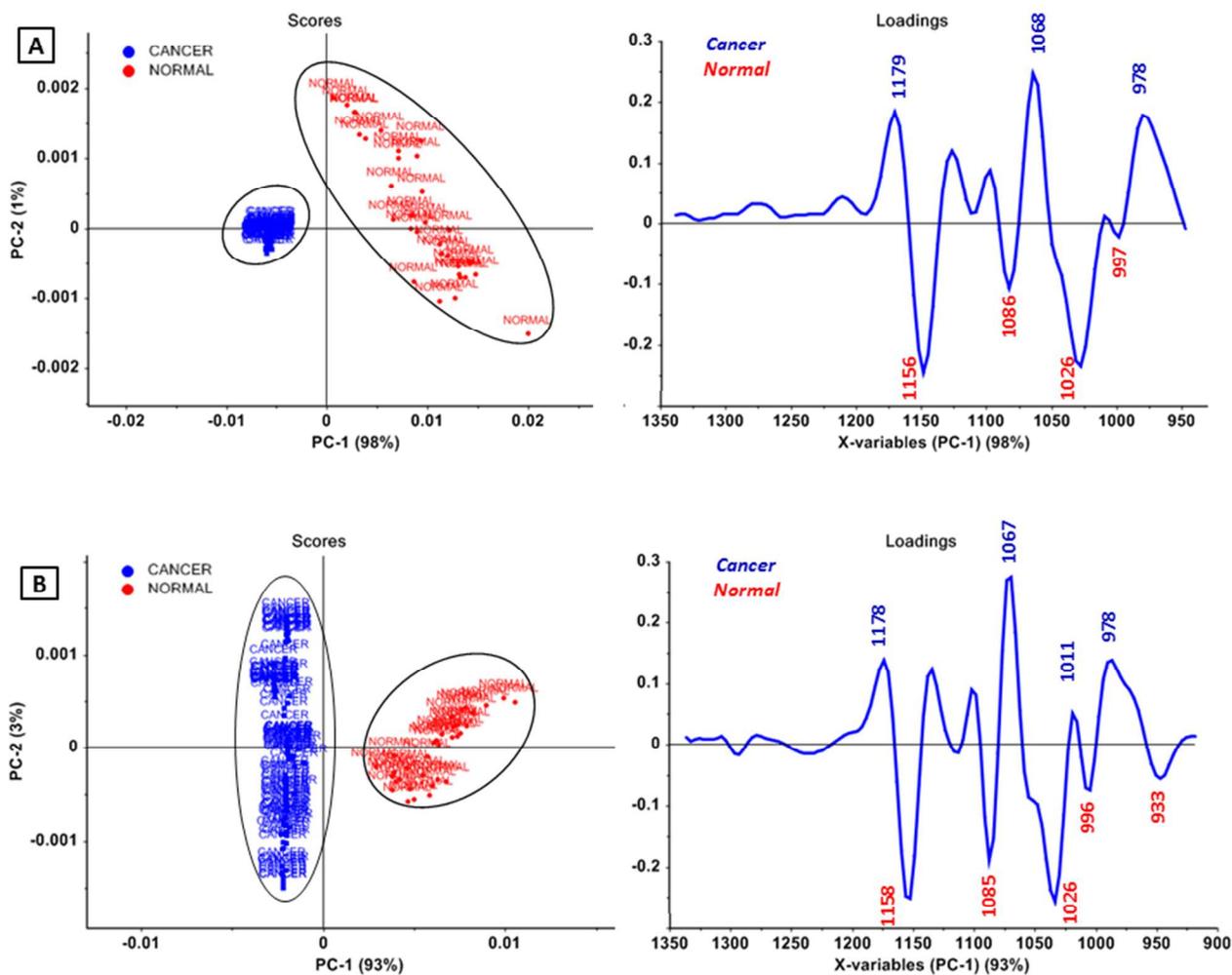


Fig.7. PCA results for spectra from healthy (marked as **NORMAL** in red) and cancer (marked as **CANCER** in blue) liver tissue, obtained for the (A) transmission and (B) transfection spectra: scores plot (PC1 vs PC2) and loading corresponding to PC1, in the range 1350-900 cm^{-1} .

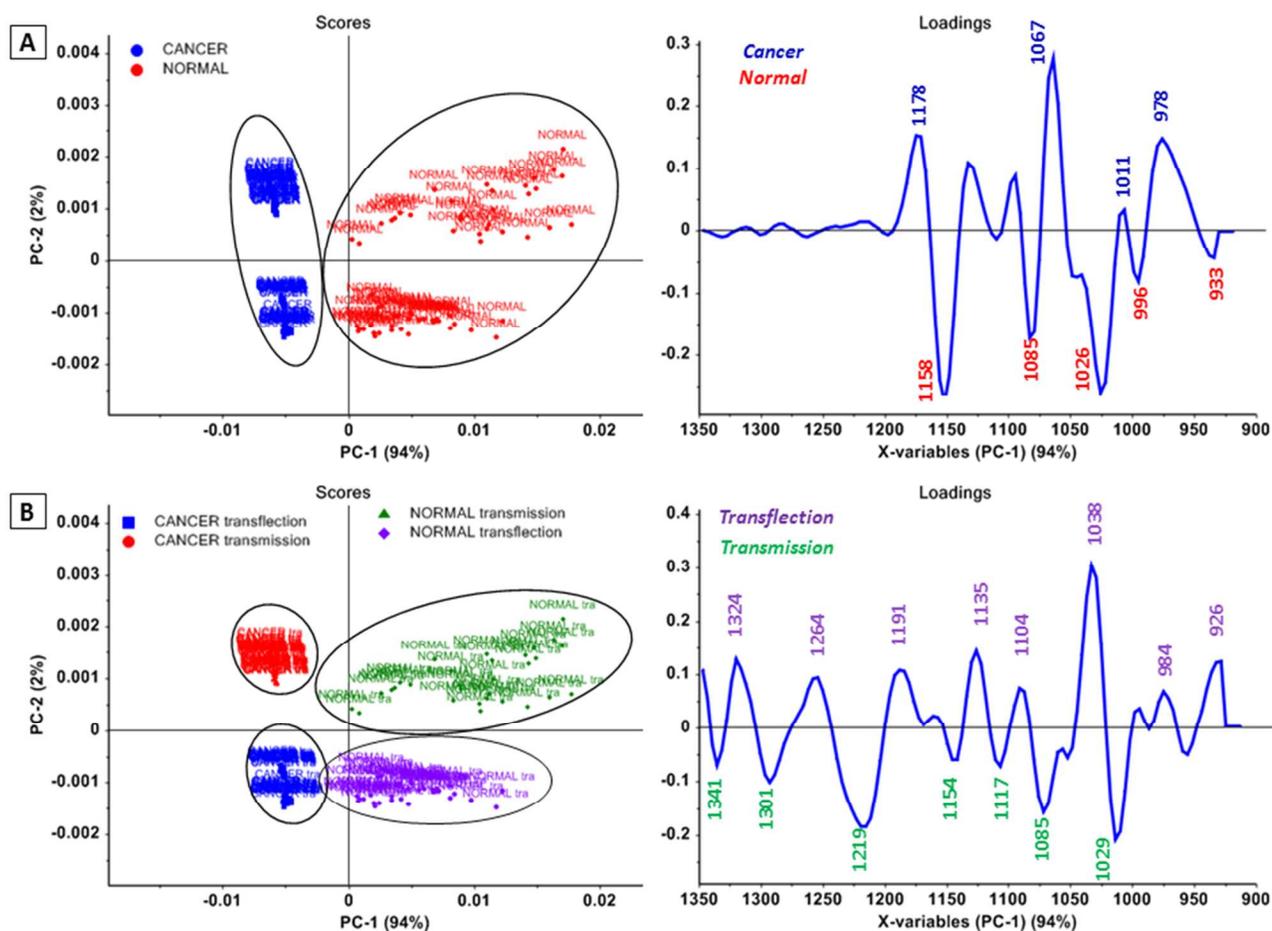


Fig.8. (A) PCA results for spectra from healthy (marked as “NORMAL” in red) and cancer (marked as “CANCER” in blue) liver tissue, obtained for the mixed spectra in the range 1350-900 cm^{-1} and (B) PCA results from the same sample showing the grouping between measurement modes (CANCER transfection, CANCER transmission, NORMAL transfection and NORMAL transmission): scores plot (PC1 vs PC2) and loading corresponding to PC2 plot (PC1 vs PC2) in the range 1350-900 cm^{-1} .

Conclusions

The results demonstrate that the diagnostic capability of FTIR imaging is not affected by the substrate, at least in the case of liver histiocytic sarcoma, when using the spectral region (1350-900 cm^{-1}) for multivariate analysis. In this case the chemical differences between normal and cancerous tissue are much greater than the contribution from the EFSW effect. Besides the differences in glycogen content between normal and cancerous tissue, which can also be attributed to nonspecific diseases, a marker band for liver sarcoma was identified at 964 cm^{-1} and assigned to a nucleic acid phosphodiester backbone mode, which appeared pronounced in cancerous tissue irrespective of the substrate. This band also appeared in regions next to the tumour albeit not as intense and thus could serve as a potential marker band to determine the tumour boundary. These differences could be observed before and after the removal of paraffin. Interestingly after deparaffinisation an improvement in distinguishing cancer from non-cancer was achieved using solely the amide I mode. It is hypothesised that this difference

relates to the amount of bound water still present in the protein after the xylene washes but this requires further experiments to verify. For liver sarcoma the best routine method to distinguish cancer from healthy tissue is to use Mirr-IR substrates with paraffin embedded tissue and do the multivariate analysis on a restricted spectral region. More work is required to ascertain whether this approach will work for other types of cancers where the spectral changes are not so large or for dysplasia where changes are usually very small but FTIR spectroscopy certainly still shows potential as a method to determine the extent of tumour penetration and possibly identify tumour boundaries not seen in conventional H&E staining using a routine clinical preparation.

Acknowledgements

This work was supported by the European Union under the European Regional Development Fund (grant coordinated by JCET-UJ, POIG.01.01.02-00-069/09) and National Science

Center (grant DEC-2013/09/N/NZ7/00626) along with an Australian Research Council Discovery Grant (DP0878464). KK acknowledges the Marian Smoluchowski Krakow Research Consortium: "Matter Energy Future" (granted the KNOW status for the 2012-2017 by the Ministry of Science and Higher Education) scholarship and financial support for international collaboration from Jagiellonian University Project, Society – Environment – Technology. We acknowledge Mr. Finlay Shanks (Monash University) for instrumental support.

Notes and references

^a Faculty of Chemistry, Jagiellonian University, Krakow, Poland, Fax: +48 12 6340515; Tel: +48 12 6632064; Email: baranska@chemia.uj.edu.pl;

^b Jagiellonian Centre for Experimental Therapeutics, Jagiellonian University, Krakow, Poland, Fax: +48 12 2974615; Tel: +48 12 6645464;

^c Centre for Biospectroscopy and School of Chemistry, Monash University, Wellington Road, Victoria 3800, Australia

^d Multiple Sclerosis Research Group, Monash Immunology and Stem Cell Laboratories, Monash University, Wellington Road, Victoria 3800, Australia

^e Diagnostic Center for Population and Animal Health, Department of Pathobiology and Diagnostic Investigation, Michigan State University, Lansing, USA

^f Department of Microbiology and Molecular Genetics, Michigan State University, Lansing, MI 48910, USA

Electronic Supplementary Information (ESI) available. See DOI: 10.1039/b000000x/

- 1 K.M. Marzec, T.P. Wrobel, A. Rygula, E. Maslak, A. Jaształ, A. Fedorowicz, S. Chlopicki, M. Baranska, *J. Biophotonics*, 2014, **3**, 1–13, DOI: 10.1002/jbio.201400014.
- 2 A. Lattermann, C. Matthäus, N. Bergner, C. Beleites, B. F. Romeike, C. Krafft, B. R. Brehm, J. Popp, *J. Biophotonics*, 2013, **6**(1), 110–121.
- 3 S. Yoshida, M. Yoshida, M. Yamamoto, J. Takeda, *Journal of Pharmaceutical and Biomedical Analysis*, 2013, **76**, 169–176.
- 4 F. Severcan, O. Bozkurt, R. Gurbanov, G. Gorgulu, *J. Biophotonics*, 2010, **3**(8-10), 621–631.
- 5 S. Kumar, Reena, S. Chaudhary, Sweetie, D. C. Jain, *Open Journal of Applied Sciences*, 2014, **4**, 103–129.
- 6 L.M. Miller, M.W. Bourassa, R.J. Smith, *Biochimica et Biophysica Acta*, 2013, **1828**, 2339–2346.
- 7 P.D. Lewis, K.E. Lewis, R. Ghosal, S. Bayliss, A.J. Lloyd, J. Wills, R. Godfrey, P. Kloer, L. A. J. Mur, *BMC Cancer*, 2010, **10**, 640.
- 8 S. C. Park, S. J. Lee, H. Namkung, H. Chung, S. H. Han, M. Y. Yoon, J. J. Park, J. H. Lee, C. H. Oh, Y. A. Woo, *Vibrational Spectroscopy*, 2007, **44**, 279–285.
- 9 S. Mordechai, R. K. Sahu, Z. Hammody, S. Mark, K. Kantarovich, H. Guterman, A. Podshyvalov, J. Goldstein, S. Argov, *J. Microsc.*, 2004, **21**, 86–91.
- 10 E. Njorge, S. R. Alty, M. R. Gani, M. Alkatib, *Conf Proc IEEE Eng Med Biol Soc.*, 2006, **1**, 5338–5341.

- 11 B.R. Wood, M. Kiupel, D. McNaughton, *Vet Pathol.*, 2014, **51**(1), 224–237.
- 12 M. Verdonck, N. Wald, J. Janssis, P. Yan, C. Meyer, A. Legat, D. E. Speiser, C. Desmedt, D. Larsimont, C. Sotiriou, E. Goormaghtigh, *Analyst*, 2013, **138**(14), 4083–4091.
- 13 W. Yang, X. Xiao, J. Tan, Q. Cai, *Vibrational Spectroscopy*, 2009, **49**, 64–67.
- 14 M.A. Mackanos, C. H. Contag, *Trends Biotechnol.*, 2009, **27**(12), 661–663.
- 15 M.J. Baker, E. Gazi, M. D. Brown, J. H. Shanks, P. Gardner, N. W. Clarke, *British Journal of Cancer*, 2008, **99**, 1859–1866.
- 16 M. Diem, L. Chiriboga, H. Yee, *Biopolymers*, 2000, **57**(5), 282–290.
- 17 J.V. Coe, Z. Chen, R. Li, R. Butke, B. Miller, C. L. Hitchcock, H. C. Allen, S.P. Provoski, E. W. Martin Jr., *Proc. of SPIE*, 2014, **Vol. 8947**, B1–6.
- 18 K. Thumanu, S. Sangrajrang, T. Khuhaprema, A. Kalalak, E. Tanthanuch, S. Pongpiachan, P. Heraud, *J. Biophotonics*, 2014, **7**(3-4), 222–231.
- 19 C. Petibois, G. Deleris, *Trends in Biotechnology*, 2006, **24**, 10, 455–462.
- 20 R.K. Maes, I. M. Langohr, A. G. Wise, R.C. Smedley, T. Thaiwong, M. Kiupel, *Vet Pathol.*, 2014, **51**(1), 238–256.
- 21 H. Brook, B. V Bronk, J.N. McCutcheon, S.L. Morgan, M.L. Myrick, *Applied spectroscopy*, 2009, **63**, 1293–1302.
- 22 J. Filik, M.D. Frogley, J.K. Pijanka, K. Wehbe, G. Cinque, *The Analyst*, 2012, **137**, 853–861.
- 23 P. Bassan, J. Lee, A. Sachdeva, J. Pissardini, K.M. Dorling, J.S. Fletcher, A. Henderson, P. Gardner, *The Analyst*, 2012, **138**, 144–157.
- 24 T.P. Wróbel, B. Wajnchold, H. J. Byrne, M. Baranska, *Vib. Spec.*, 2013, **69C**, 84–92.
- 25 J. Cao, E.S. Ng, D. McNaughton, E.G. Stanley, A.G. Elefanty, M.J. Tobin, P. Heraud, *Analyst*, 2013, **138** (14), 4147–4160.