



Cite this: *Analyst*, 2021, **146**, 1260

Self-absorption corrected non-invasive transmission Raman spectroscopy (of biological tissue)†

Benjamin Gardner, ^{a,b} Pavel Matousek ^{*c} and Nicholas Stone ^{*a,b}

The first near infrared window in biological tissue ($\lambda \sim 700\text{--}950\text{ nm}$) is of great interest for its potential to safely deliver light based diagnosis and therapeutic interventions, especially in the burgeoning field of nano-theranostics. In this context, Raman spectroscopy is increasingly being used to provide rapid non-invasive chemical molecular analysis, including bulk tissue analysis by exploiting the near infrared window, with transmission Raman spectroscopy (TRS). The disadvantage of this approach, is that when probing depths of several centimetres self-attenuation artefacts are typically exhibited, whereby TRS spectra can suffer from relative changes in the "spectral features" due to differential absorption of Raman photons by the various constituents of biological tissues. Simply put, for a homogenous substance with increasing thickness, spectral variances occur due to the optical properties of the material and not through changes in the chemical environment. This can lead to misinterpretation of data, or features of interest become obscured due to the unwanted variance. Here we demonstrate a method to correct TRS data for this effect, which estimates the pathlengths derived from peak attenuation and uses expected optical properties to transform the data. In a validation experiment, the method reduced total Raman spectral intensity variances >5 fold, and improved specific peak ratio distortions 35x. This is an important development for TRS, Spatially Offset Raman Spectroscopy (SORS) and related techniques operating at depth in the near IR window; applicable to samples where there is large sample thickness and inter- and intra-sample thickness is variable *i.e.* clinical specimens from surgical procedures such as breast cancer. This solution is expected to yield lower detection limits and larger depths in future applications such as non-invasive breast cancer diagnosis *in vivo*.

Received 28th September 2020,
Accepted 2nd December 2020

DOI: 10.1039/d0an01940b

rsc.li/analyst

Introduction

The biological near infrared (NIR) window is a spectral region within biological materials where the optical properties exhibit lower scattering (μ_s) and absorption (μ_a), as well as lower levels of auto fluorescence,^{1,2} compared with the visible spectral region. This window is often broken down into two sub regions, NIR-I (700–950 nm) and NIR-II (1000–1700 nm). The NIR windows are especially important for a number of bio-medical imaging modalities including fluorescence imaging, and spectroscopy techniques including Raman spectroscopy.

Raman spectroscopy is of particular interest as it is at a new frontier of novel theranostics,³ *i.e.* the combination of optical diagnostics and photothermal therapies mediated by nano-structures.³ These techniques aim to complement existing tomographic techniques *e.g.* MRI, CT and PET where there are either safety concerns (ionising radiation), or they yield limited temporal, spatial or chemical information.

Deep Raman spectroscopy (DRS) is a group term for closely related innovative techniques for non-invasive and non-destructive probing deep into samples using Raman spectroscopy, the principal techniques are spatially offset Raman spectroscopy (SORS)⁴ and transmission Raman spectroscopy (TRS).⁵ What these techniques share is a separation between laser illumination and collection zones on the sample surface. In SORS the two zones are generally on the same side of the sample surface, while TRS illumination and collection is separated by the sample itself, *i.e.* being on opposite sides of the sample. These approaches exploit diffuse scattering of light in the material being probed and allow a retrieval of chemical signals up to several cm in depth. There are continuously new

^aSchool of Physics and Astronomy, University of Exeter, Exeter, EX4 4QL, UK.
E-mail: N.Stone@exeter.ac.uk

^bRoyal Devon and Exeter NHS Foundation Trust, Barrack Road, Exeter, Devon, EX2 5DW, UK

^cCentral Laser Facility, Research Complex at Harwell, STFC Rutherford Appleton Laboratory, Harwell Oxford, OX11 0QX, UK. E-mail: Pavel.Matousek@stfc.ac.uk

†Electronic supplementary information (ESI) available. See DOI: 10.1039/d0an01940b



and sophisticated ways of carrying out DRS,^{6–9} yet the basic premise remains the same. DRS has seen a broad range of applications in a number of fields including pharmaceutical analysis and security,¹⁰ cultural heritage¹¹ as well as extensive biomedical applications.^{12–14}

Over time the sophistication of signal retrieval has increased from identifying a single compound,^{5,15} to multiplexed signal recovery,¹³ physical property signal recovery¹⁶ and depth localisation.^{17–21} A number of publications have demonstrated how differential optical properties in a scattering medium, *i.e.* tissue, between different Raman components can also be exploited for depth localisation. These examples have been demonstrated in well-controlled samples, wherein the sample has typically a uniform thickness in transmission geometry.^{19–21} However, in the real world there is often little control of the sample thickness, when performing deep Raman on biological samples. This variation can arise undesirably in two key ways, firstly with thickness variation within a sample during Raman mapping, secondly due to differences in thickness between different samples. Both of these phenomena can lead to differences in the Raman spectra as a function of thickness and composition, due to the optical properties experienced by Raman photons propagating through the material as a function of wavelength. For example, two Raman bands sufficiently separated in wavelength might experience significantly different optical properties therefore one can be more attenuated than another over increasingly large propagation distances. Previous DRS studies have noted this phenomena,^{22–24} having been highlighted in one TRS experiment at λ_{ex} 830 nm, with Raman spectra having shown significant loss of signal above 1250 cm^{-1} when the sample thickness increased from 20 mm to around 50 mm.²² While, other work specifically noticed the relative loss of the amide I signal^{23,24} in their spectra, and had to remove this region entirely from the subsequent multivariate data analysis. However, these works lacked extensive signal collection, *i.e.* multiple spatial locations over a sample with varying thickness. Therefore, the mechanism of how changes in thickness in samples vary with complex attenuation profiles and thus affect the collected Raman spectra was not fully realised and otherwise useful data could not be utilised.

Moreover, self-attenuation of conventional Raman signal has been observed in solution based surface enhanced Raman spectroscopy (SERS), where there was competition between extinction due to absorption of light mediated by nanoparticles and the SERS enhancement these provide.^{25,26} Previous work has also addressed some limitations of Raman spectroscopy and factors which can distort its quantitative capabilities, when applied to measurements in diffusely scattering media.^{27–31} The work of Reble *et al.*²⁹ demonstrated absolute Raman scattering coefficients were restored under back scattered Raman geometry. While the work of Barman *et al.*³⁰ also investigated distortions arising from multiple scattering events. In addition, studies using both SORS³¹ and TRS,³² discussed how the optical properties can affect sampling depth and recovered signal distributions respect-

ively. However, these studies have not addressed the influence of sample geometry or how the optical properties can distort the recovered spectra themselves.

This perturbation to the spectra can often overwhelm and easily mask many other, more subtle, changes to Raman spectral signatures of biological relevance, *e.g.* the presence or absence of disease and represent a fundamental limitation of DRS. The spectral distortions are particularly problematic for common analytical tools used in spectroscopic analysis such as Principal Component Analysis (PCA), which is often used as an important initial step to reduce data complexity prior to application of more advanced machine learning algorithms/regression techniques. In PCA for example, these distortions give rise to high ranked contributions (eigenvectors), or the distortion features are mixed with other eigenvectors that relate to chemical features of interest. This acts to complicate and confound development of robust numerical models, thereby reducing overall sensitivity and the maximal depth DRS can probe.

A useful term for describing the combined influence of the optical properties of a medium is effective transport coefficient (μ_{eff}), where $\mu_{\text{eff}} = \sqrt{3\mu_a \cdot \mu'_s}$.^{33–35} This term describes the combined influence of both absorption coefficient μ_a and the reduced scattering coefficient μ'_s , where $\mu'_s = \mu_s(1 - g)$ and g is the measure of anisotropy *i.e.* the mean cosine of scattering angles. Therefore, having an approximate knowledge of $\mu_{\text{eff}}(\lambda)$ of the material being investigated alongside the sample thickness, one can consider how significant, if at all, the spectral distortions are likely to be. Some materials will exhibit a flat profile *i.e.* uniform attenuation across a spectral window, while others contain a gradient, or worse a complex profile, which gives rise to distortions in relative peak intensities.

Here we demonstrate the effect that variation in sample thickness has on the Raman spectra measured with TRS and how this is dependent on the optical properties of the medium being probed, rather than solely the chemical composition. Moreover, the severity of the distortions, as expected, differ between common biological materials *i.e.* protein-rich, lipid-rich. Finally, we demonstrate that with limited *a priori* information *i.e.* approximate optical properties, the Raman spectrum can be restored, thereby reducing spectral variation due exclusively to material thickness and therefore enhancing the underlying differences arising in the chemistry that could otherwise be obscured. This effect can be observed, even in materials with homogenous chemistry. This work builds on previous studies that have noticed similar phenomena.^{22,25,26}

Materials and methods

Transmission Raman measurements were carried out on a custom built system, of which a similar system has been previously described.^{36,37} In summary, an 808 nm solid state laser (Innovative Photonic Solutions, Monmouth Junction, NJ, US) was coupled to illumination optics *via* an optical fibre. Two 808 nm laser clean up filters (LL01-808-25, Semrock, Rochester, New York, US) filtered the laser output prior to



focusing into a ~ 15 mm diameter beam on the sample surface. The sample stage platform consisted of a fused quartz window (2 mm thickness), sitting on a motorized stage (8MTF, Standa), providing full XY positioning control with a range of 102 mm in both directions. The collected signal was filtered by 3 edge filters, and focused into a Holospec 1.8i (Kaiser Optical Systems, Ann Arbor, Michigan, USA). The spectrometer contains a custom high dispersion grating providing a spectral range of ~ 600 – 1200 cm^{-1} and a 1 mm slit, providing an effective spectral resolution of ~ 15 cm^{-1} . The spectrometer was coupled to a deep depletion CCD detector (Andor BR-DD iDus 420) to record the Raman signals.

Sample preparation included cutting commercially available blocks of lard (pork fat) into either flat segments of a uniform thickness, or cutting these into an approximate wedge shape. Additionally, we used calcium hydroxyapatite (HAP) (Sigma Aldrich) contained in a quartz cell (Starna Scientific) as a target chemical component placed on top of lard samples, on the surface with the laser beam incident upon it. This has previously been used in studies as a substitute for *in vivo* breast calcifications.^{37,38} For all Raman measurements a laser

power of 2 W was used, delivering a power density at the sample surface of ~ 12 mW mm^{-2} . We performed TRS mapping experiments, *i.e.* spectra were gathered from multiple spatial points across the sample, under the following parameters; each spatial location was probed with 3 TRS spectra consisting each of $(0.05 \text{ s} \times 20)$ accumulations, while for static measurements 20 spectra of $(0.05 \text{ s} \times 20)$ accumulations were acquired.

All data was processed in Matlab 2017a undergoing the following pre-processing steps. Firstly a median filter was applied to the data to remove the presence of cosmic rays, and the data was then averaged to leave one mean spectrum per spatial location. A linear baseline was then subtracted from the data prior to vector normalisation.

Results and discussion

The optical properties of fat (μ_a and μ'_s) for the near-IR spectral window ~ 700 – 950 nm (Fig. 1A), show that while scattering has an inverse relationship with wavelength *i.e.* decreases monotonically as a function of increasing wavelength, the

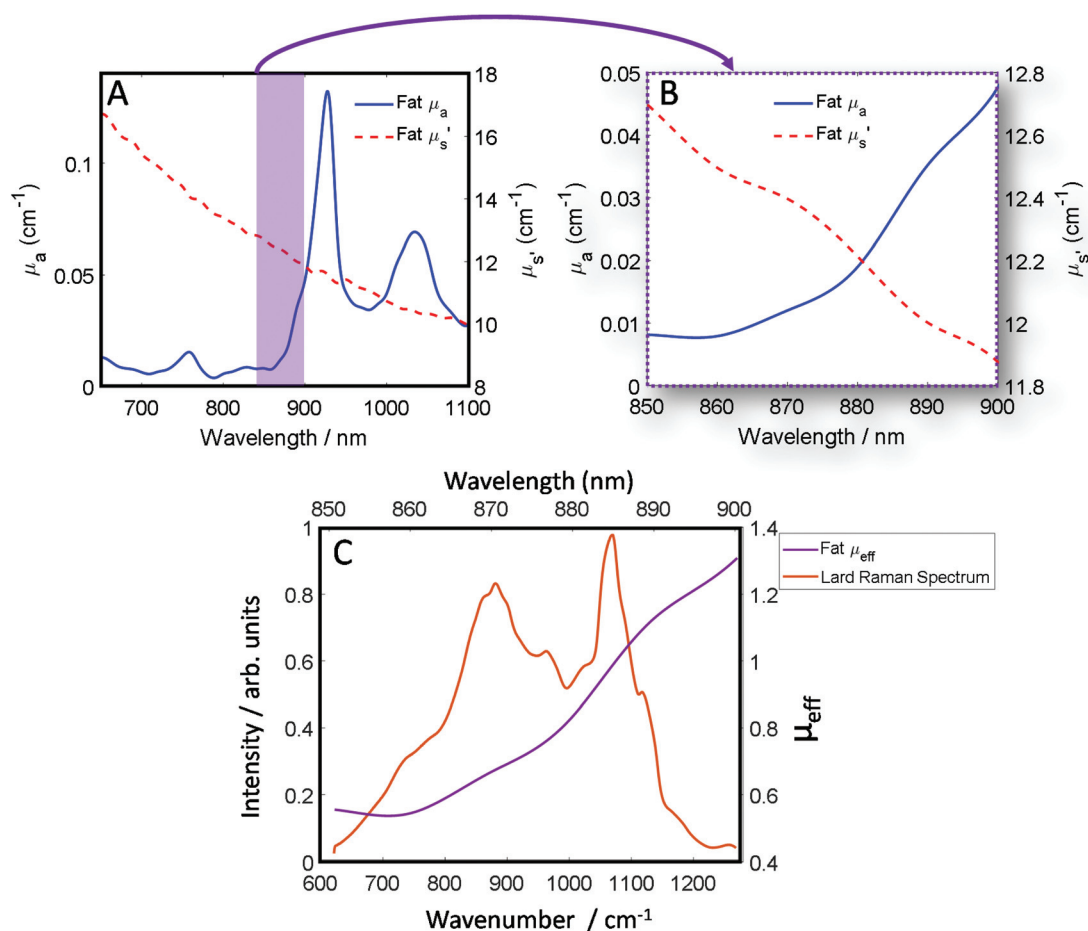


Fig. 1 (A) Optical properties (μ_a and μ'_s) of fat between 650 nm and 1100 nm. (B) Optical properties (μ_a and μ'_s) of fat between 850 nm and 900 nm, equivalent to the spectral range of 600 – 1250 cm^{-1} using an λ_{ex} 808 nm and our high dispersion Raman grating. (C) The combined optical properties contribution of μ_a and μ'_s of fat termed $\mu_{\text{effective}}$ (μ_{eff}) over the Raman spectral range (purple) while in orange a transmission Raman spectrum of lard (1 cm thick). Optical properties data (A and B) was adapted with permission from the work of Mosca *et al.*⁴⁰



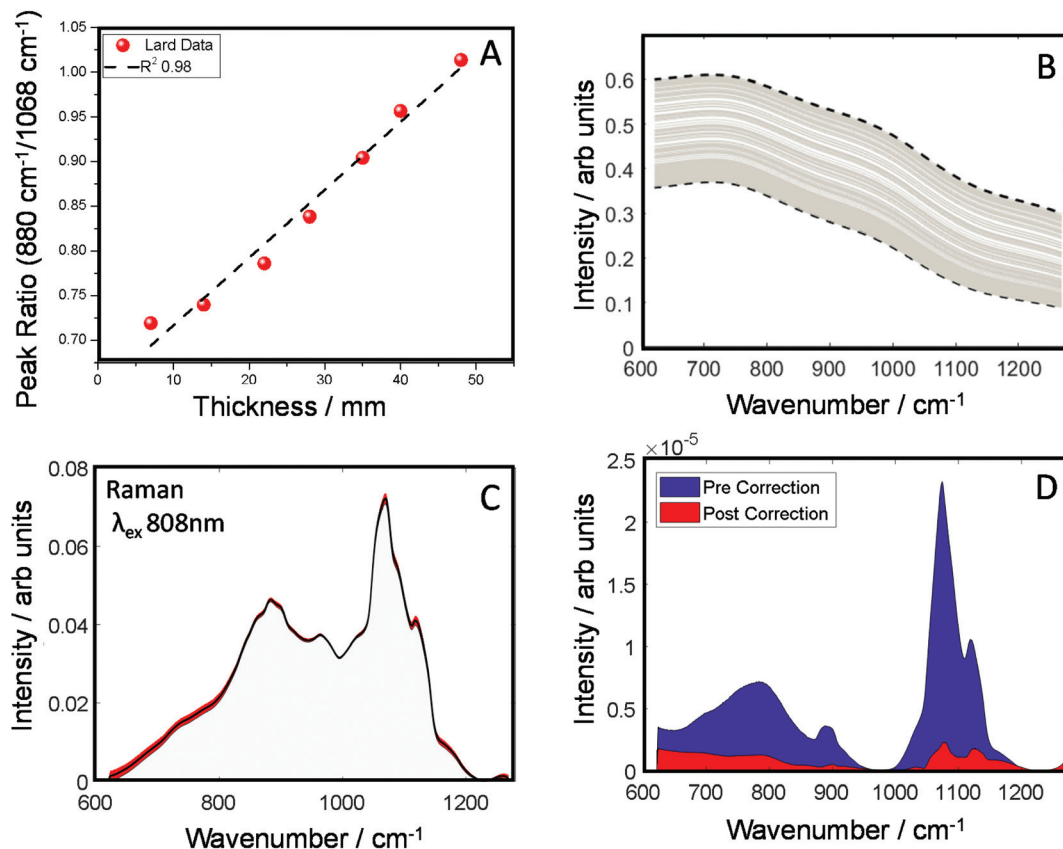


Fig. 3 (A) TRS peak intensity ratio of lard (880 cm⁻¹/1068 cm⁻¹) plotted as a function of lard thickness [~7–48 mm] with a line of best fit. (B) Correction profile as a function of predicted thickness. (C) Corrected Transmission Raman spectra (TRS), average spectrum (black), σ red, (D) spectral variance as a function of wavenumber (cm⁻¹) for TRS lard spectra pre and post correction.

a given thickness, <0.5% around the average. From this we were able to estimate sample thickness for all the TRS spectra from the mapping experiment (Fig. 2B).

Knowing the thickness for each TRS spectrum, we were then able to properly apply a correction function utilising μ_{eff} for the spectral range of the measurements (Fig. 3B) to correct for the spectral distortion. For this we used μ_{eff} as our attenuation coefficient, and pathlength ℓ , is half the sample thickness. The spectrum arising from TRS is comprised of multiple phenomena, firstly the excitation beam attenuates as it travels through the sample, usually in an exponential decay fashion. Therefore, the amount of Raman events decrease the further from the illumination point (deeper) with the sample. The Raman spectra generated are then attenuated to differing degrees depending on the total propagation distance, and the wavelength dependent optical properties have a varying influence. In short, the most Raman events occur at the sample surface, however these Raman photons are the most likely to be maximally distorted. Conversely, Raman photons generated at the exit surface are most likely to be undistorted, however there will be significantly fewer (thickness depending). However, a simplification to account for the assumed bulk transmission Raman signal origin,⁴³ can be utilised. Therefore, rearranging the transmittance equation $T = e^{-\mu_{\text{eff}}\ell}$

allowing us to create a correction function that relates the sample thickness and the optical properties.

The corrected spectra (Fig. 3C) show considerably less variance than present originally. This is further demonstrated in plots of spectral variance (Fig. 3D) in the data before (blue) and after correction (red). While some peak variance remains, as expected, it has greatly been diminished across the whole spectrum. Overall, full spectral intensity variance is reduced by a factor of ~5.5. Moreover, specifically looking at the peak ratio of interest (880/1068 cm⁻¹) this itself has been reduced 35 times. The reduction of this spectral distortion, due to physical properties of the sample, now enables a more accurate characterisation of any underlying chemical properties, provided by the Raman scattered signal, which might previously be masked by the interfering spectral distortions. This, in practical terms, is expected to enable TRS to detect with higher sensitivity more subtle spectral changes, *e.g.* those associated with a medical condition. While less variance is observed in the corrected spectra (Fig. 3C), no prior information of the pure spectrum is used in the reconstruction, therefore over or under correction can occur. Therefore, the corrected spectrum might not necessarily match that of the pure, undistorted Raman spectrum. It is noted that the goal of the correction procedure is to minimise variance in the Raman data set rather than



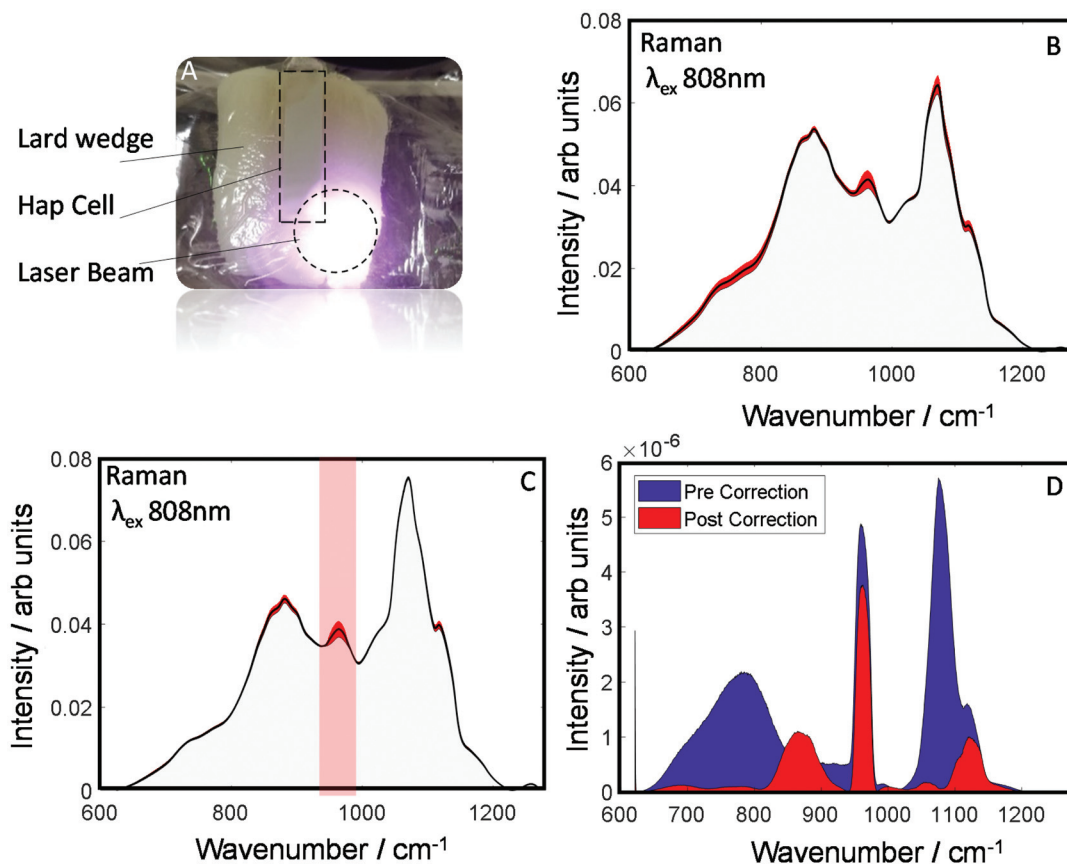


Fig. 4 (A) Image of lard with a gradual change in thickness along the length of the quartz cell containing hydroxyapatite. (B) Raman spectra of a lard/HAP phantom measured in a map configuration (TRS), average spectrum (black), σ red. (C) TRS data of lard/HAP phantom, corrected for spectral variation of optical properties. (D) Comparison of variance of pre and post corrected data.

recovering a 'perfect' undistorted Raman spectrum one would have obtained when measuring, for example, a very thin sample with negligible signal self-absorption.

To evaluate the robustness of this approach to derive more accurately the chemical information about the sample an additional dataset was investigated as is shown (Fig. 4A). This dataset was collected from a smaller wedge-shaped block of lard (with a maximal thickness of 3 cm) that had a quartz cell containing HAP placed upon it (representing a detected spectral component – e.g. calcification associated with a breast cancer lesion). TRS spectra of the maps are shown pre- and post-correction (Fig. 4B and C). As can be seen visually, post-correction acts to limit the spectral variability of the lard particularly in the region where you would expect to see the main HAP peak (PO_4 symmetric stretch), centered at $\sim 960 \text{ cm}^{-1}$. Furthermore, when comparing the variance (Fig. 4D), the pre-correction spectral features (blue) are dominated by distortions due to optical properties; while, post correction (Fig. 4D red) HAP dominates the variance plot relative to the attenuation distortions, indicating a considerably enhanced sensitivity to detect such desirable chemical signatures.

When a common analytical tool used in spectroscopy such as principal component analysis (PCA) is applied to the data-

sets, the principal component (PC) loadings (Fig. S2A†), which are weighted for the variance explained, you can see the correction decreases considerably the influence of the background distortions and increases the relative presence of HAP in PCs 1 & 2. Moreover, looking at the variance explained in the PCs, the correction increases HAP's relative importance from 16% to 43% (Fig. S2B†).

Conclusions

Here we have demonstrated within the first NIR spectroscopy optical window, a general region of lower absorption and scattering within biological tissue, TRS spectra are susceptible to distortions due to the optical properties of the bulk sample and Raman photons achieve differential transmittance across the detected spectral range. In a homogenous sample of fat over a range of 1–4 cm, large Raman spectral distortions are observed as thickness increases. By the application of a correction function that utilises the effective attenuation coefficient μ_{eff} and the approximate optical properties of fat, the spectral distortions could be significantly reduced (35 \times). This finding is vitally important for the general applicability of TRS,



- 42 K. Sowoidnich and H.-D. Kronfeldt, *ISRN Spectrosc.*, 2012, **2012**, 256326.
- 43 P. Matousek, N. Everall, D. Littlejohn, A. Nordon and M. Bloomfield, *Appl. Spectrosc.*, 2011, **65**, 724–733.
- 44 T. J. Moore, A. S. Moody, T. D. Payne, G. M. Sarabia, A. R. Daniel and B. Sharma, *Biosensors*, 2018, **8**(2), 46.
- 45 S. Pahlow, K. Weber, J. Popp, B. R. Wood, K. Kochan, A. Rütther, D. Perez-Guaita, P. Heraud, N. Stone, A. Dudgeon, B. Gardner, R. Reddy, D. Mayerich and R. Bhargava, *Appl. Spectrosc.*, 2018, **72**(S1), 52–84.
- 46 A. S. Moody and B. Sharma, *ACS Chem. Neurosci.*, 2018, **9**, 1380–1387.
- 47 R. A. Odion, P. Storbbia, B. M. Crawford and T. Vo-Dinh, *J. Raman Spectrosc.*, 2018, 1–9.
- 48 J. M. Yuen, N. C. Shah, J. T. Walsh, M. R. Glucksberg and R. P. Van Duyne, *Anal. Chem.*, 2010, **82**, 8382–8385.
- 49 M. Z. Vardaki, C. G. Atkins, H. G. Schulze, D. V. Devine, K. Serrano, M. W. Blades and R. F. B. Turner, *Analyst*, 2018, **143**, 6006–6013.

