A Quantitative Raman Spectroscopy Investigation of Parchments.

Megan Poppy Vimpany

MSc By Research

University of York

School of Physics, Engineering and Technology

January 2024

ABSTRACT

Raman spectroscopy is an innovative technique in the fields of art history and archaeology due to its ability to be non-invasive and non-destructive to its target, whilst still being able to provide a great amount of spectroscopic detail. The molecular information it can provide is a very useful tool for the purposes of artefact identification and authentication. This thesis uses these qualities of the method in an analysis on 2 parchment samples from 2 different eras. Raman profiles for parchments are well documented in literature, however there is little work out there attempting to assess the meaning of the Raman intensity values of the spectra produced. Using this data, an in depth discussion of the meanings of intensity differences in the spectra with respect to existing literature will be conducted, as well as a quantitative analysis of molecular heterogeneity differences between the samples by calculating a decay constant for the rate of data convergence, an analysis technique new to the field.

DECLARATION

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. PARCHMENT MANUSCRIPTS

Parchments are a type of writing surface thought to have been first introduced to wide use by Eumenes II of Pergamon (197-159 BC), as an alternate vessel to papyrus to replicate the lost library of Alexandria [1][2]. It was known to have been used before this, however, with basic examples dating back to the Fourth Dynasty of Ancient Egypt (late 27th- early 25th century BC)[1][2]. They are made from animal skins, often goat or sheep[3]. The two sides of a parchment have different appearances due to this, as one side was flesh, and one was external skin with hair, known as the grain side[4]. Higher quality skins produce higher quality parchments. These are often sourced from younger animals, and are called vellums[3].

The main macromolecular constituent of parchment is collagen[5][6], accounting for 95% of the molecular content of a finished parchment [7]. There are many types of collagen, however types I and III are what is mainly found in skin [6]. Typically, 80% or more of the collagen in skin is type I, with the rest being type III and possibly small amounts of type V [4][7]. The exact ratio of collagen types varies dependent on the age, species and body part of the skin source [8].

Collagen molecules are formed from 3 polypeptide chains in a triple helix formation [9]. Each chain is formed of a repeating sequence of amino acids in the pattern Gly-X-Y, with Gly representing the amino acid glycine and X and Y representing 2 other amino acids. In the majority of cases these are proline and hydroxyproline respectively[10][11], but other combinations are possible [8]. Hydrogen bonds are instrumental in stabilising the triple helix structure [8]. They are formed with water molecules found between both the polypeptide chains and the individual Gly-X-Y units [8]. In some collagen types, including I, III and V [12], adjacent polypeptide helices combine to form fibrils via covalent cross links [8]. These fibrils form collagen fibres when packed together [7]. The result of the numerous strong bonds between and within collagens structure is what give it its high tensile strength [13]

A general method for making a parchment starts by stripping the skin to remove unwanted parts of the flesh [3][4]. In older examples this was done only by mechanical scraping of the skin [3], but as parchment preparation methods advanced over time liming was introduced as

a pre-treatment to assist in the removal of non-collagenous matter [1][4] Liming involves soaking the skin in an alkaline calcium hydroxide ($Ca(OH)_2$) bath [4], breaking up the different layers of skin and allowing for easy removal of hairs, fat and other flesh components not used in a parchment with a dull blade [1]. The first documentation of this method dates back to the 8^{th} century AD [14]. After stripping all that is left of the skin is its dermal layer, comprised of mostly type I collagen [6][11].

The next step is to remove contaminants from the skin such as dirt, blood, and chemicals left over from the liming process [4]. Many of these will be removed with the hairs and fats during scraping, but to ensure as much as possible is removed the skin is also washed with cold water [1][4][. After this it must be dried. This is done under tension in order to re-orientate the collagen fibres to be parallel to the skin surface [1][4], which gives the parchment high tensile strength [4]. The parchment is then finalised to prepare the surface for writing, which includes shaving, polishing, whitening and cleaning. An example of this finishing process described by Kautek et. al. involves rehydrating the parchment with an egg wash and linseed oil [15].

Variations from this method of preparation are expected as different cultures have created their own ways of making parchments, leaving no one accepted practice for manufacturing it, and no industrial method for batch production making every piece unique[5]. An example of a variation is the use of tannins and oil in Jewish tradition for sacred documents[5]. The species of animal the original skin was sourced from can also be an indication of where the parchment was made. For example, in the middle ages, parchments from England or France were commonly made from calf or sheep skin, whereas parchments made in Italy were usually goat skin [1]. The species of animals abundant enough to be able to provide skins for parchments in a particular country can be informative of the agricultural landscape at that time.

The conditions the skin is subjected to during transformation into a parchment are harsh, and can cause damage to the molecular structure of the collagen. Bonds within the polypeptide chains are subject to hydrolytic cleavage during liming, leaving them more susceptible to further damage while aging [4][16]. The way the fibres are rearranged in the drying process also means the parchment is more susceptible to damage from outside factors [4]. There are many reasons a parchment could become damaged, especially over 2 centuries after its

creation. Exposure to light, heat or water are all ways in which the structure of the parchment, specifically the collagen fibrils, is altered.

Oxidation occurs upon reaction of the collagen fibrils with a reactive oxygen species, often a free radical [7]. This impacts the parchment fibrils by reacting with functional side groups of amino acids in the polypeptide chains, modifying their structure and increasing the acidic to basic amino acid ratio [8]. Oxidation can also cause breakdown in the polypeptide chains by cleavage of the N-C covalent bonds that connect adjacent amino acids in the chain [13]. This type of breakage is common at regions exhibiting increased polarity such as tyrosyl groups, or amino acids with a charge such as lysine. The free radicals causing the oxidation can come from many sources, including UV light (photo-oxidation) [17] or from products of the breakdown of other organic molecules such as lipids (auto-oxidation) [52].

Thermal denaturation is the damage caused upon exposure of a parchment to high temperatures. This happens when disruption of hydrogen bonds within the fibrils causes the triple helix structure of the polypeptide chains to uncoil into random conformations, which increases the intensity of the hydrogen bonds within them. This causes a change in polarity which in turn causes a change in the molecule's Raman profile [8]. This change happens progressively as a parchment heats up, with its thermal stability and the homogeneity of its structure decreasing as temperature increases until full denaturation into the random coils structure at 120˚C [13]. At this point the damage to the parchment is irreversible.

Upon expose to temperatures as high as 150˚C, the thermal stability of the collagen drops low enough to allow water molecules to react with the fragmented chains, starting the process of gelatinisation. Full conversion from collagen to gelatin is observed at 180˚C [13]. Due to its susceptibility to heat and water, humid conditions can be particularly damaging to a parchment. Temperature and humidity changes can cause shrinking and stretching of parchments [1]. Shrinkage in a new or well preserved parchment will begin at around 55˚C, whereas a more damaged manuscript will start to shrink at around 30˚C, not much higher than room temperature [8]. This proves the importance of proper parchment storage and preservation, as many countries where parchment usage would have been prominent have ambient temperatures reaching 30˚C or higher.

Parchments are particularly susceptible to these kinds of damage due to the way they are made. The process of creation involves the realignment of the collagen fibres, during which

the polypeptide chains are affected by hydrolytic breakage of covalent bonds [16], resulting in the collagen already being partially gelatinised [8].

It is important to consider both inks and parchments when examining ancient manuscript samples. Samples coming from written manuscripts may have ink traces that would be detectable, the components of which can also be identified with the Raman technique. Whilst all colours were used across history for illustrations, the literature in older works featured red headings and black text, potentially making them the most common pigments. One example of black ink is iron gall ink, made from oak galls and iron sulphate. Like parchment, it has a unique production method in different areas leading to many recipes[19]. Its insolubility and corrosiveness likely made it very permanent, making it a common choice from the middle ages until the early 20th century[19]. Red pigments were derived from crushing female cochineal or kermes insects. Other sources of natural pigments include plants such as yellow pigment created from saffron, and crushed minerals or semi-precious stones for example: lapis lazuli for an expensive blue pigment, or calcinating lead white to create the toxic, orange pigment Minium.

Parchments eventually fell out of favour to other types of paper with different manufacturing techniques, such as recycling rags by treating them and grinding them up before moulding them into paper, or using wood pulp to make paper as we know it in the current era, which was used from the 19th century.

1.2. Raman Spectroscopy

Raman spectroscopy works by irradiating a sample with laser light in the UV-visible region, and observing the scattered light. In order for scattering to occur, a fluctuating dipole moment must be induced within the molecules of the sample for analysis. This is done by using the laser to induce a fluctuating electric field of strength *E* varying with time *t* shown by,

$$E = E_0 \cos(2\pi\nu_0 t) \quad (1)$$

where E_0 is the maximum amplitude of the fields vibrations, and ν_0 is the maximum frequency of the radiation. [20] This induces a fluctuating dipole moment if a molecule is irradiated due to the release of a photon causing polarisation briefly. The dipole moment *P* is characterised by the following equation,

$$P = \alpha E = \alpha E_0 \, cos(2\pi \nu_0 t) \tag{2}$$

where α is the polarizability of the molecule, which can be expressed as a linear function of nuclear displacement q for small vibrations. At a vibrational frequency ν_m, q is defined as,

$$q = q_0 \, cos(2\pi \nu_m t) \tag{3}$$

with q_0 being the maximum amplitude of these molecular vibrations. α can therefore be written as,

$$\alpha = \alpha_0 + \left(\frac{\partial \alpha}{\partial q}\right)_0 q_0 + \cdots. \tag{4}$$

In this expression α_0 is the polarizability at the equilibrium point, and $\left(\frac{d\alpha}{\partial q}\right)_0$ is the rate of change of polarizability with respect to the change in nuclear displacement at the same point.

At the equilibrium position, $q=0$. This reduces all further terms than those shown in equation 4 to 0, and by combining this result with equation 2, the following equation is obtained,

$$P = \alpha E_0 \, cos(2\pi \nu_0 t)$$

$$= \alpha_0 E_0 \, cos(2\pi \nu_0 t) + \left(\frac{\partial \alpha}{\partial q}\right)_0 q E_0 \, cos(2\pi \nu_0 t). \tag{5}$$

By combining the result given in equation 5 with equation 3, the formula is written as follows,

$$= \alpha_0 E_0 \, cos(2\pi \nu_0 t) + \left(\frac{\partial \alpha}{\partial q}\right)_0 q_0 E_0 \, cos(2\pi \nu_0 t) \, cos \, (2\pi \nu_m t), \tag{6}$$

which can be further expanded using these known trigonometric identities,

$$cos(a + b) = cos(a) \, cos(b) - sin(a) \, sin(b) \tag{7}$$

$$cos(a - b) = cos(a) \, cos(b) + sin(a) \, sin(b). \tag{8}$$

Combining the identities in equations 7 and 8 produces the result,

$$cos(a+b)+cos(a-b)$$
$$= cos(a)\,cos(b) - sin(a)\,sin(b) + cos(a)\,cos(b) + sin(a)\,sin(b)$$

$$\cos(a+b)+\cos(a-b) = 2\cos(a)\cos(b) \quad (9)$$

and by substituting $a = 2\pi\nu_0 t$ and $b = 2\pi\nu_m t$ into equation 9, equation 6 can be rewritten as,

$$= \alpha_0 E_0 \cos(2\pi\nu_0 t) + \frac{1}{2}\left(\frac{\partial\alpha}{\partial q}\right)_0 q_0 E_0[\cos\{2\pi(\nu_0+\nu_m)t\} + cos\{2\pi(\nu_0-\nu_m)t\}]. \quad (10)$$

The form of the dipole moment equation given in equation 10 shows two terms. The first term is used in classical theory to describe a dipole radiating light at the incident frequency ν_0. This kind of scattering is called Rayleigh scattering. The second term is conditional of $\left(\frac{\partial\alpha}{\partial q}\right)_0 \neq 0$, meaning there must be a change in polarizability relative to the vibration to achieve this scattering. This term represents Raman scattering (named after C. V Raman, who first documented the phenomenon in 1928 [1][22]). When $\left(\frac{\partial\alpha}{\partial q}\right)_0 \neq 0$ a vibration is said to be Raman active [20].

Raman scattering involves inelastic scattering of the incident radiation at one of two frequencies resulting in two different kinds of Raman scattering. Raman Stokes scattering occurs when the molecule hit with the incident frequency transitions to a higher energy level, causing a lower frequency to be emitted $(\nu_0 - \nu_m)$, and Raman anti-Stokes scattering occurs when the opposite happens and the molecule is at a lower energy level after the interaction, giving out an increased frequency $(\nu_0 + \nu_m)$ [21], as described mathematically in equation 10. Raman stokes is the most favoured of these two processes because it is more probable for a photon to begin at the ground state than in an excited one [1]. However, Raman scattering only occurs approximately 0.0001% of the time a photon is scattered [1], making Rayleigh scattering the more common case.

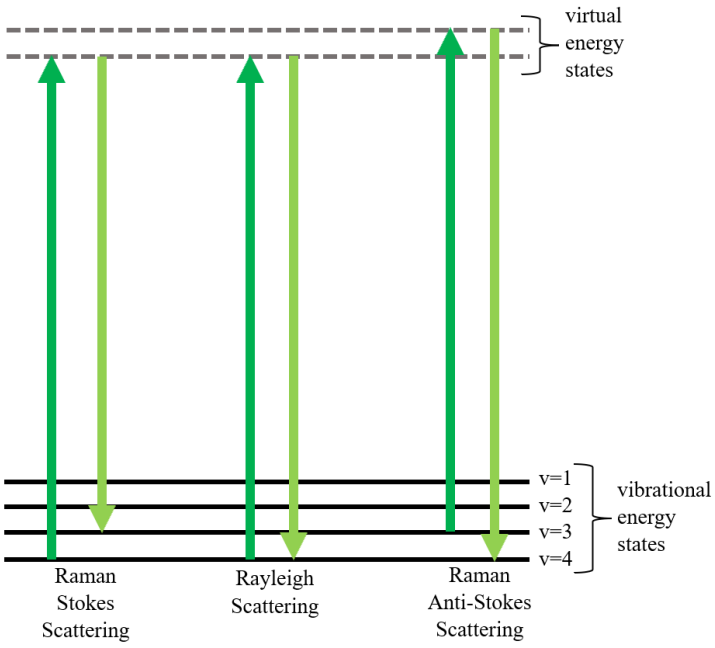


Figure 1: A diagram representing the energy state changes undergone during Rayleigh scattering (centre), Raman Stokes scattering (left) and Raman anti-Stokes scattering (right).

After interaction with the laser, the molecule enters a virtual energy state until a photon is scattered, which happens almost instantaneously. It is important to note that the molecule never stays in an excited state, as there are vibrationally excited states on the ground energy state potential surface, so there are no energy level transitions. Figure 1 illustrates the energy level changes occurring during each type of scattering.

Stokes and Anti-Stokes scatterings result in the same Raman spectra but mirrored, however the increased likelihood of Stokes scatterings lead to more intense spectra. Raman Stokes spectra are therefore used to display the Raman profile of a given substance [20]. This is demonstrated below in figure 2 with an example spectra for CCl_4.

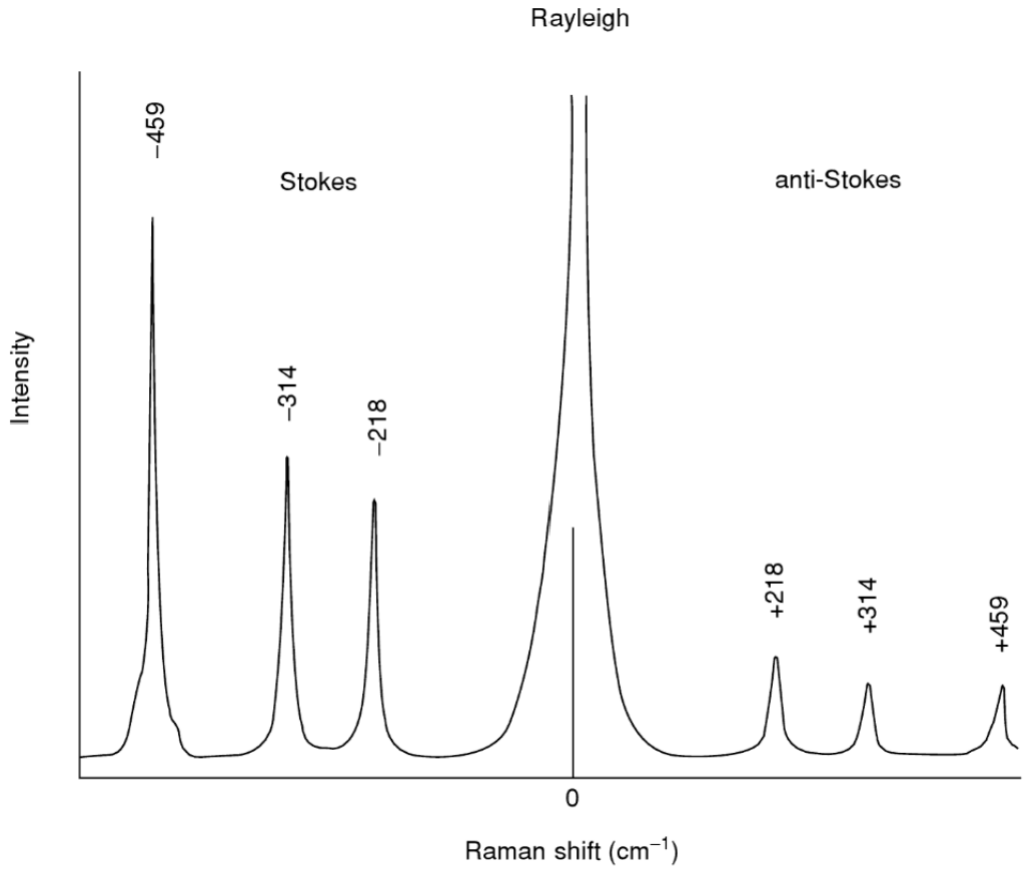


Figure 2: Depicts an example Raman profile of CCl_4 showing the Raman shift values for each of the three types of scattering, taken from Introductory Raman Spectroscopy (2nd edition) [20].

ν_M is the frequency shift from the incident to the resulting frequency, measured in wavenumbers (cm^{-1}), and is what quantises Raman scattering[21]. It is characteristic of the material under analysis which is what allows for identification using this technique. Wavenumber is defined by [22]:

$$\bar{\nu} = \frac{\nu}{c} \tag{11}$$

Where ν is frequency in s^{-1}, and c is the speed of light (accepted value $3 \times 10^8 ms^{-1}$or $3 \times 10^{10} cms^{-1}$).

Raman is often used alongside infrared spectroscopy (IR), as they provide complimentary results arising from contrasting molecular features. IR spectroscopy depends on the dipole moment of the molecule, making polarised molecules very visible, whereas in Raman the

small increase in polarisation to an already polarised molecule during vibration means they only produce weak peaks. One of the main situations in which it is useful to use both methods of analysis is when looking at hydrated molecules. The shape of water molecules makes them highly polar meaning it shows very strongly in an IR spectrum, often strong enough to prevent other molecular information from showing. Since Raman spectroscopy only increases the water molecules polarity by a fraction of the amount it is already polarised, the markers for water show much weaker, allowing the rest of the molecular information to be analysed.

1.3. Raman in Historical Conservation

Raman spectroscopy has been used in similar studies for its advantages in the field of conservation. The techniques ability to be non-destructive and non-invasive means fragile and valuable artefacts can undergo analysis without damage allowing for further expansion of knowledge, and without sacrificing any source material [11][19]. This coupled with its excellent spatial resolution and specificity make it an ideal tool for art historians and archaeologists alike for the purposes of identification and authentication [23]. One specific example includes Raman analysis in wall murals, previously difficult to examine as they cannot be moved or sampled without breaking up the art. The results of pigment analysis on one such mural was able to show different pigments used in areas of the same colour, which could point to the mural having been worked on by multiple artists [23]. and Many samples offered for this project were delicate and dated back many centuries, and therefore Raman was chosen as an analysis method that would be gentle to the sample was necessary to preserve their integrity.

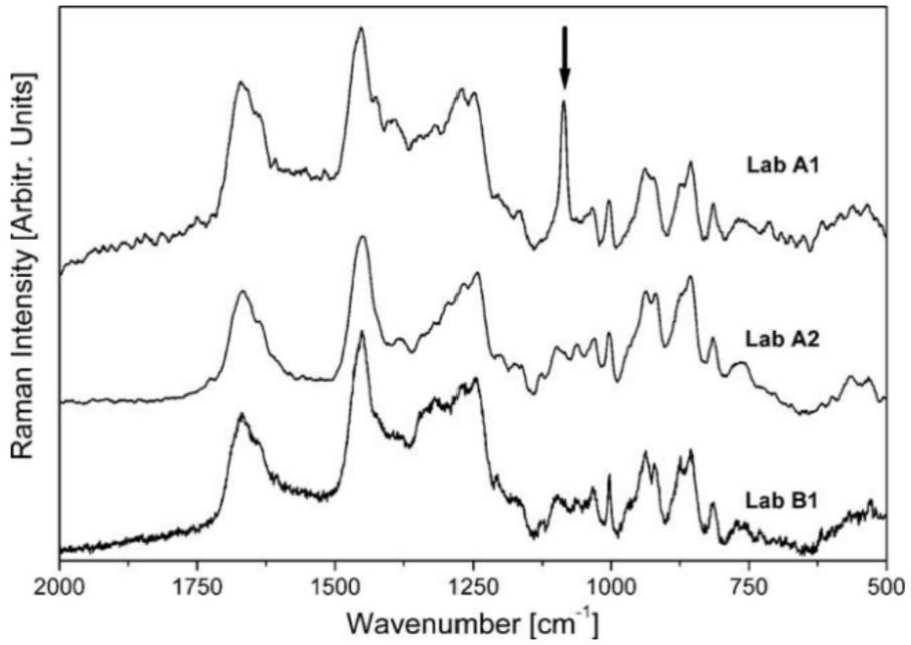


Figure 3: An example of a Raman spectra for parchments taken from Bicchieri et. al. (2011) [14]. Lab A1 represents a parchment prepared in the lab with a standard western method, Lab A2 represents Lab A1 after treatment to remove $CaCO_3$ treatment from the parchment, and Lab B1 is a parchment prepared in the lab simulating eastern development methods. The arrow points to a band assigned to $CaCO_3$.

Figure 3 shows a literature example of a Raman profile for a parchment sample from a study by Bicchieri et. al. [14]. It includes 3 spectra from parchment samples with different creation methods. Lab A1 and Lab B1 were representative of western and eastern preparation methods respectively, and differences between their spectra are apparent upon simple visual analysis. The arrow points to a band assigned to $CaCO_3$, a phase of lime used in parchment preparation as described in section 1.1. The Raman shift for this band is well documented to be $1086cm^{-1}$.

Global Raman shift values applicable to parchments have been widely documented, largely thanks to work by Talari et. al., creating a large database of known bands for biological tissues [24]. Studies examining ancient parchment degradation are of particular interest as existing quantitative analysis of this is limited.

1.4. Project Aims and Objectives

Raman Spectroscopy has been an innovative tool for identification purposes for decades due to its ability to perform analysis in vivo as well as ex vivo, as well as being non-damaging to the sample itself. For these reasons it is an ideal tool for this study, as the samples offered for this project were delicate and dated back many centuries. They therefore required an analysis method that would be gentle in order to preserve their integrity.

These qualities have also been especially useful in the medical field, as it allows investigation into materials such as cells and tissue in a way which is not invasive and able to be performed without preparing the sample in any way, therefore leaving it in the condition it would be found in the body. It is therefore widely used in cancer diagnostics [25], leading to much work having been done on tissue and skin to characterise their Raman profiles. In turn this aids the study of parchments, as they are made from animal skins and therefore the Raman shift values for their main material is already widely documented. Whilst publications in medical journals will typically use human skin samples, some studies have been done on animal tissue. Tfaili et. al. [26] performed a comparative investigation of human skin to pig skin, as it is often used in cosmetic tests for products before human consumption. Studies have also been conducted on manuscripts themselves, but often include only bond movement information and do not differentiate between parts of the spectra caused by the biological components of the parchment and those caused by contaminants added in the process of creating it.

Whilst a wide range of data has been collected investigating the Raman shift values of peaks in the collected spectra, there is a lack of quantitative information about the intensity of each peak, and what it may reveal about the sample [3]. This thesis will analyse 2 parchment samples from different eras (a modern control sample and an ancient sample dating back to 1769) and examine their respective Raman profiles, in order to determine differences in peak intensity values and their potential corelation to established biomarkers. This will be carried out over a wider range usually seen in literature, with spectra obtained starting at a Raman shift of $300cm^{-1}$ instead of the typical $500cm^{-1}$ seen in figure 3 in order to reveal more information from the highly detailed fingerprint area of the spectra. In addition, there has been minimal works of this kind that examine the high wavenumber region of the parchment Raman profile, between $2500\text{-}3200cm^{-1}$. Univariate and multivariate analysis will be conducted, in the hope of expanding the limited literature available for principal component analysis on parchment samples. Biomarkers showing statistically significant Raman

intensities across the two samples will have their heterogeneity quantified by producing decay plots for the rate of convergence of the standard error in their intensity values. This will build on a technique used by Rocha et. al. by providing a comparison between the two samples, hopefully revealing the implications of degradation of manuscripts over time [27].

2. METHODS

2.1. SPECTRA ACQUISITION

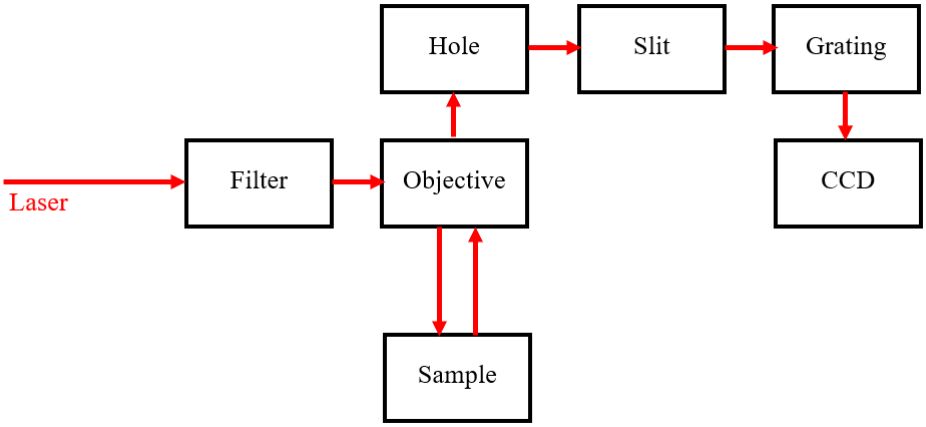


Figure 4: A diagram depicting the optical pathway of a Raman spectroscopy set up and its components, accurate to the equipment used in this thesis. The path of the laser light is indicated by the red arrows.

In this project, the Horiba XploRA instrumentation was used to acquire Raman spectra from the samples, alongside a computer running LabSpec 6 [28][29]. Figure 4 shows a basic schematic of how the spectrometer works. The incident light is first passed through a filter in order to regulate the laser power to avoid damage to the sample. The objective lens then focusses the laser on the sample, at a spot size determined by the equation,

$$x = \frac{1.22\lambda}{NA} \tag{12}$$

where x is the spot size, λ is the wavelength of the laser light and NA is the numerical aperture of the objective lens [53]. For analysis on the parchment samples in this thesis, a laser wavelength of 785nm was chosen as a balance to avoid the higher fluorescence effects of lower wavelengths but to maintain the strength of the scattering [1]. Previous works have obtained good results using this wavelength [5][30]. The objective used was a 50x magnification long working distance (LWD) lens with a numerical aperture of 0.5. Using these values in equation 12 gives a spot size value of

$$x = \frac{1.22\times 785nm}{0{\cdot}5} = 1915.4nm. \tag{13}$$

These parameters can also be used to calculate the lateral spatial resolution of the measurements, known as Rayleigh's criterion, which describes the separation nedded between two points in order for them to be distinguishable. This is done using the equation [53],

$$spatial\ resolution = \frac{0.61\lambda}{NA} \tag{14}$$

which gives a value of

$$spatial\ resolution = \frac{0.61 \times 785nm}{0{\cdot}5} = 957.7nm. \tag{15}$$

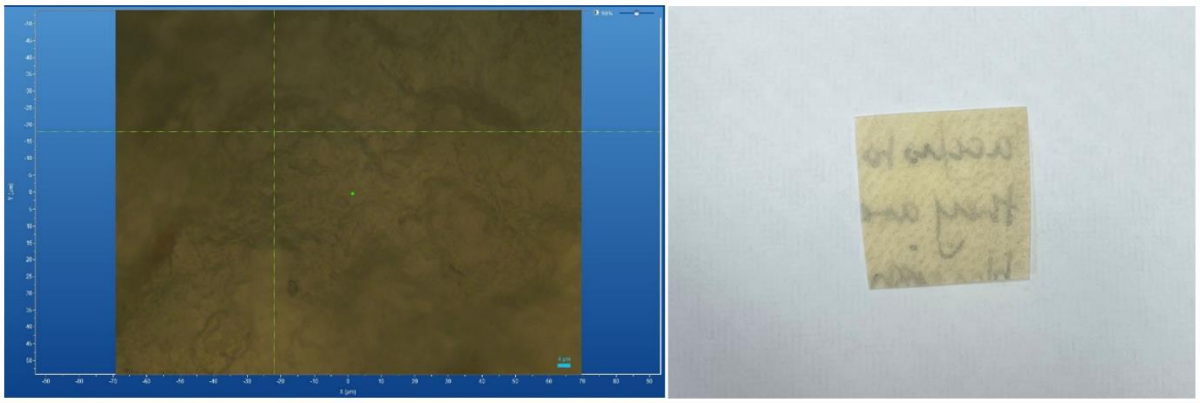
After interaction with the sample, scattered radiation passes through the confocal hole, which controls the depth of the measurements received. A small hole was used in this case (500μm) to ensure the measurements on the parchment came from a shallow depth to avoid the Raman profile from the inks on the grain side showing in the spectra. The radiation next passes through a slit which can alter its spectral resolution as desired, followed by a diffraction grating which separates the radiation into its constituent wavelengths. The parameters used were a slit aperture of 100μm, and a diffraction grating with 1200 lines per mm. A CCD interprets the dispersed wavelength information and outputs the Raman spectra.

Before any measurements could be made the instrument had to be calibrated. The material used for this was a pure silicon sample kept in the lab for this purpose, fixed to a glass slide. Experimental parameters were also decided before data acquisition, in order to ensure that the experiment was set up in such a way that the spectra produced were legible, whilst not having so long an acquisition time as to make the total time taken inefficient as a large quantity of data was needed.

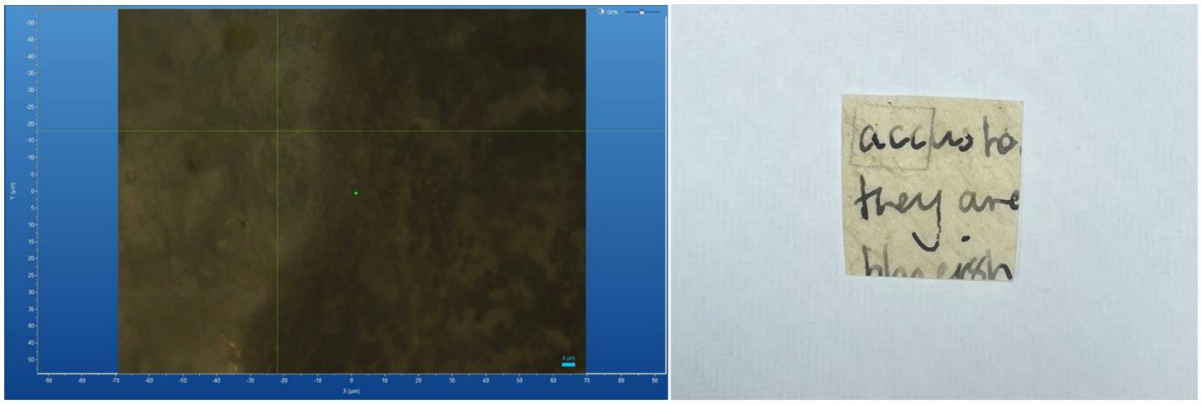
The parameters originally chosen were 20 accumulations at an acquisition time of 5s, but due to issues with the modern sample staying in focus for the duration later spectra were taken with updated parameters of 6 accumulations at an acquisition time of 6s. Steps were taken during data processing to ensure that these changes would not affect the results, discussed in section 2.3.1. The spectra were collected in two separate regions per spot measurement: the fingerprint region between 300-1800cm^{-1} and the high wavenumber region between 2500-3200cm^{-1}. This was decided via experimentation within different ranges between 0-3500cm^{-1}

to determine the largest range where signal was still apparent above the noise and fluorescence common at low wavenumbers.

2.2. SAMPLES USED



a) As seen through LabSpec camera. b) The flesh side of the sample.



c) As seen through LabSpec camera d) The grain side of the sample.

Figure 5: Four images showing a) the flesh side of the modern sample optically imaged through the LabSpec 6 software, b) the flesh side of the sample seen by eye, c) the grain side of the modern sample optically imaged through the LabSpec 6 software, d) the grain side of the sample seen by eye.

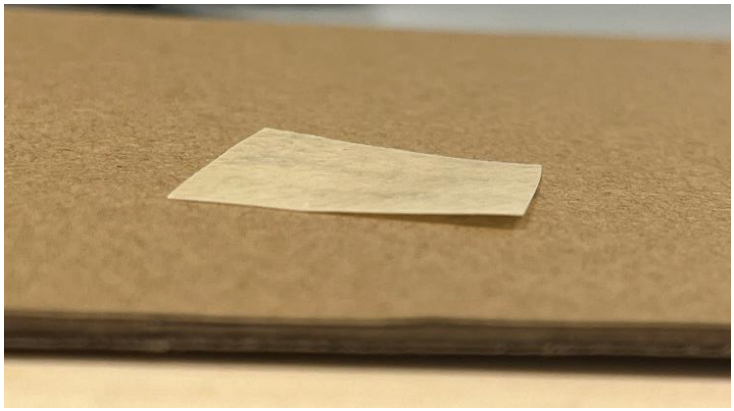
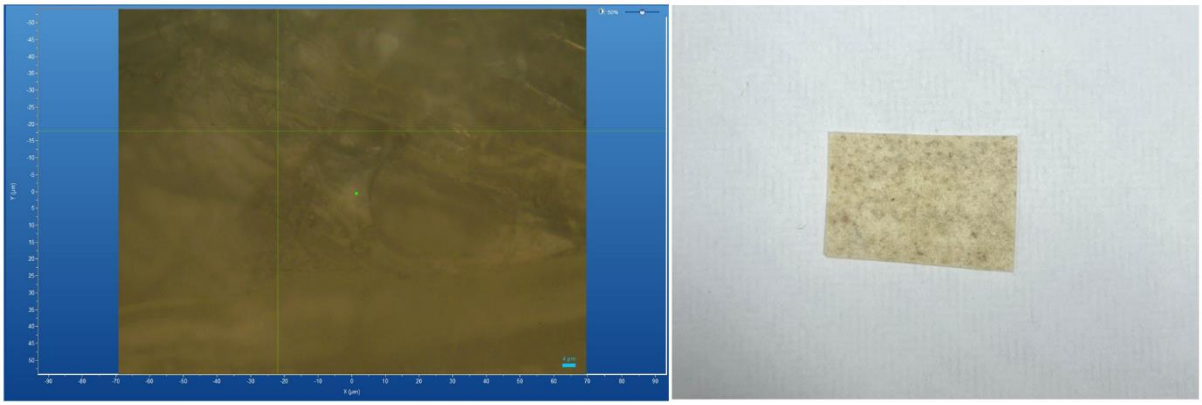
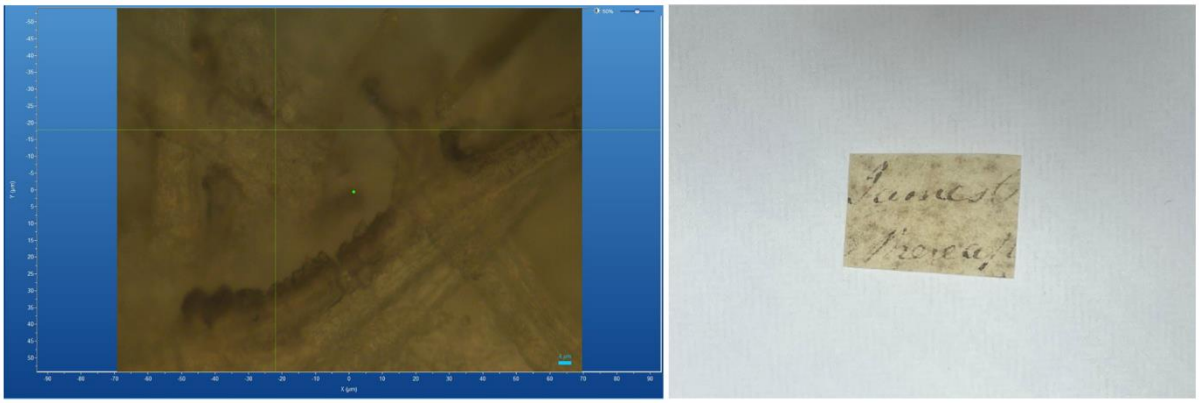


Figure 6: An image of the modern parchment sample taken from the side, showing a slight curve upwards at the front right corner.



a) As seen through LabSpec camera. b) The flesh side of the sample.



c) As seen through LabSpec camera. d) The grain side of the sample.

Figure 7: Four images showing a) the flesh side of the ancient sample optically imaged through the LabSpec 6 software, b) the flesh side of the sample seen by eye, c) the grain side of the ancient sample optically imaged through the LabSpec 6 software, d) the grain side of the sample seen by eye.

Figure 5 shows images of both the grain and flesh sides of the modern sample as seen by the eye and through the Labspec 6 software. Figure 6 is an image of the sample taken from the side, showing an elevation at the front right corner caused by a slight curl in the parchment. This proved problematic when taking data sets, as spectra taken from this region of the parchment had a higher spectral background. To limit the effects of this, the sample was stored in a flower press between data collection sessions in an attempt to flatten the parchment as much as possible. This method was referable over others attempted, such as securing the corners of the parchment to the slide with glue dots, as there was no risk of damage to the sample from adhesives or other contaminants. The ancient sample, shown in figure 7, was extremely flat and therefore no further flattening was necessary.

In parts a and c of figures 5 and 7, the green dot symbolises the spot location where Raman measurement would take place, and the scale bar shows 4μm. As stated in section 2.1, spectra were taken only on the flesh side of the parchments. This was done to lessen the chance of seeing bands produced by inks in the parchment Raman profile, allowing the focus of the thesis to remain on the parchments themselves.

Both parchment samples used for this project were part of a set acquired in 2013 courtesy of Professor Matthew Collins and their research group in the BioArCh Unit within the Department of Archaeology at the University of York. These samples in particular came from wills and indentures forming part of the Beasts2Crafts research collection [58], a project examining parchments from the perspectives of both science and the humanities in order to better understand it as a medium [59]. The modern parchment sample was made from a goat skin, while the older sample, dated from 1769, was made from a sheep skin. The samples were always handled with tweezers to avoid potential damage from contaminants on hands, and always placed into the XploRA the same way up. Analysis was done on the side of the parchment with no ink, to try and ensure results obtained were purely those of the parchments and not the pigments or other components of ink. Parchments were also placed in the spectrometer in the same orientation, due to a previous study finding amide bands to be anisotropic, having variations depending on the molecule orientation with respect to the incident radiation[9][11].

Comparing figures 5a and 7a, visual differences between the two parchments can be observed. The surface of the ancient sample appears to be less uniform, with an almost fibre like appearance. This is especially apparent in figure 7c, where the ink has highlighted the edges of the structure. This is in contrast to figure 5c, where the ink sits evenly on the surface. These differences are likely due to degradation of the parchment over time changing its structure, which will be discussed extensively in chapter 4 of this thesis.

2.3. Spectral Analysis

2.3.1. Preprocessing

The first step of data analysis was ensuring all the spectra in each data set matched a common set of wavenumbers. This was done by cutting and interpolating the data. Cutting was done first in Raman Tool Set [31]. The files were cut to the range 300.7cm^{-1} to 1799.5cm^{-1}. This range was chosen by looking at the raw data for the wavenumber measurements and eliminating variations at the start and end, creating a data set with a single range. The cut spectra were then loaded into Igor Pro 9 [32], where they could be interpolated to ensure the spectra had the same number of points. This was necessary as the movement of the spectrometer in-between the measurements for the fingerprint and high wavenumber regions can cause the wavenumber information for each spectra to be slightly different, and in order to carry out accurate multivariate analysis this cannot be the case. The interpolation was carried out with the interpolation2 procedure, sourced internally at the University of York, which was loaded in and edited to include the correct parameters. The fingerprint region was interpolated to 2342 points across each spectra. The interpolation2 procedure performs linear interpolations between points, and the large number of points present in this data set allows for this to be a good approximation.

Due to the complex nature of the data, baselining of the datas were done in a bespoke manner. The Baselines procedure file, sourced from Wavemetrics [33], was used to manually place 11 nodes on each spectra, at 319cm^{-1}, 371cm^{-1}, 508cm^{-1}, 640cm^{-1}, 721cm^{-1}, 800cm^{-1}, 1148cm^{-1}, 1224cm^{-1}, 1353cm^{-1}, 1504cm^{-1}, and 1728cm^{-1}. These were done by eye, correct to within ±3cm^{-1} to match the spectral error of the instrument. The positions were chosen to erase as much of the spectral background and noise as possible without erasing any part of the Raman profile.

Due to the changes in acquisition parameters previously mentioned in section 2.1, all data needed to be normalised and smoothed before convergence testing could take place. In Raman Tool Set, the spectra were first normalised with respect to the total spectrum area, then smoothed using a cubic spline.

The same analysis was done for the high wavelength region, except for differences with the interpolation procedure and the baselining points to make the analysis region appropriate.

The interpolation was done to 893 points in the range 2748.6-3099.8cm^{-1}, and for the baseline, 4 points at 2750.09cm^{-1}, 2841.82cm^{-1}, 3027.73cm^{-1}, and 3094.44cm^{-1} were used. This analysis was done the same way for both samples.

2.3.2. Testing for Convergence

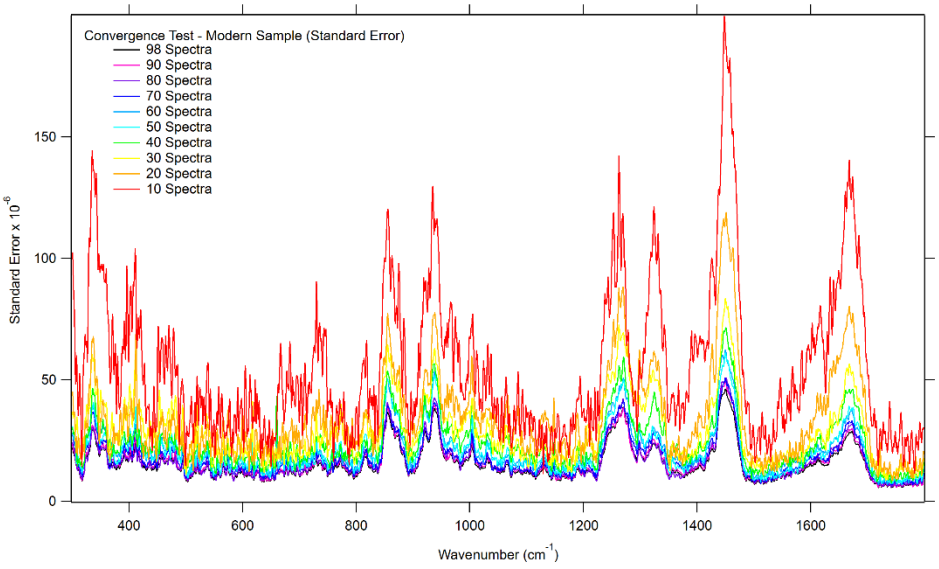


Figure 8: An example of a convergence test performed in this thesis. Conducted on the standard error in the spectral average for the fingerprint region of the modern sample, in groups with increasing numbers of spectra added to the average ranging from 10-98.

Convergence tests for the spectral average, standard error, and twice standard deviation of a group of spectra were performed in Igor Pro 9. While the convergence of the standard error is a more sensitive measure, tests were also performed on the twice standard deviation to ensure the statistical robustness of the data set and that all possible variables were converged. Figure 8 shows an example test performed on the standard error in the spectral average for the modern samples fingerprint region. Each of these values were calculated for groups of randomly selected spectra, starting with a group of 10 and increasing incrementally by 10

after that. To select these spectra, a random number generator was created in the online version of Excel 2021 [57] using the following command statement,

$$INDEX\big(UNIQUE(RANDARRAY(N\ \hat{}\ 2, 1, N_{min}, N_{max}, TRUE)), SEQUENCE(N)\big) \quad (16)$$

where N is the number of values needed in the list, and N_{min} and N_{max} are the smallest and largest numbers in the range being used to generate the list.

New groups of spectra were generated for each test, rather than adding 10 to the existing group. This helped ensure randomness in the sampling as convergence would not be achieved until any set of the given quantity of spectra produced identical plots. Convergence testing has been done in the same way in a similar study by Rocha et. al. [27] To quantify these tests, the penultimate convergence groups spectra was subtracted from the final groups, and the maximum difference between the two was taken to be the degree to which convergence was achieved.

2.3.3. Multipeak Fitting and Univariate Analysis

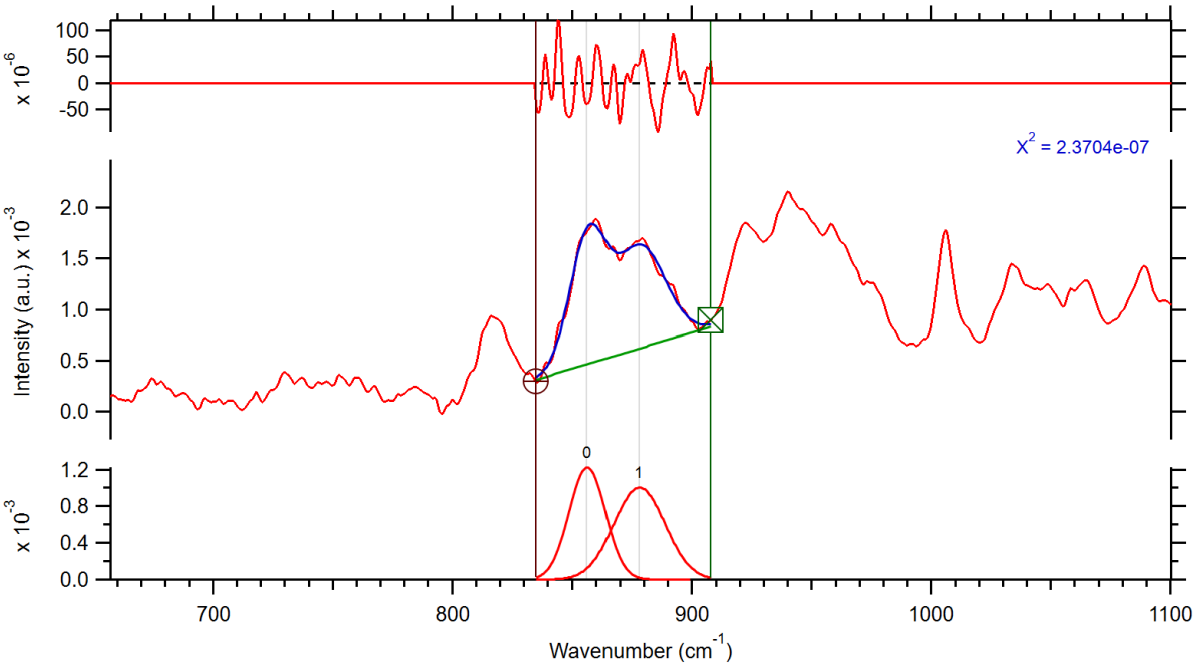


Figure 9: An example of the band fitting conducted in the fingerprint region of the ancient sample. Depicts a section of the spectra with 2 gaussian curves fit to deconvolve 2 bands, as well as a wave showing the fitting error for the region. Also given on the figure is the X^2 value, used to determine the accuracy of the fit.

[25]

The multipeak fitting was done on an average wave representative of all the spectra in a data set. The peaks of the spectra were fit with Gaussian bands using the Levenberg-Marqardt algorithm to determine both the wavenumber and intensity values[34].

Fitting was done using the Multipeak Fitting package in Igor Pro 9. For the high wavenumber region this was able to be done in one batch, fitting 3 peaks to the spectra. However, due to the larger wavenumber range and greater complexity of the spectra, the fingerprint region had to be split into 16 sections to retain the detail obtained in the data. Each section contained between 1 and 4 peaks and the regions for these were decided by eye based on the dips between peaks. The bands within these sections were chosen in a way that would minimise the fitting error. An example of this is given in figure 9. The uppermost wave in this figure represents the fitting error, and below it is the deconvolved bands and their locations. The standard error was used as the error in the fittings as opposed to the fitting error to take into account the statistical robustness of the data set.

New waves were generated in Igor Pro 9 for the average plus and minus the standard error, and peak fitting was done on these in the same way in order to calculate the standard error in the Raman intensity values. Also given in figure 9 is the value for X^2 for the region, used as a statistical measure of fitting accuracy. A low X^2 value indicates an accurate fit, and comparison of the fitting errors to the standard errors confirmed that only the standard error would be relevant.

To assess the difference between the band intensity values of the two samples in a univariate manner, a plot was made showing both average spectra, and a spectra representing the average of the modern data subtracted from the average of the ancient data. This was used to identify areas of the Raman profile that show significant differences between the two samples. A scatter plot was made to confirm these results quantitatively, including error bars representing the standard error in each intensity measurement. From this plot it can be observed if there is a significant statistical difference between the two samples at a specific biomarker by calculating the separation of the error bars at its assigned wavenumber. Detailed biomarker assignments based on existing literature will be discussed in chapter 3.

2.3.4. Principal Component Analysis (PCA)

The univariate analysis methods described in the previous section do a good job of quantifying the differences between the samples above the spectral background, however multivariate analysis is needed to take into account the full statistical robustness of this data set. This was done using principal component analysis (PCA). PCA reduces the number of dimensions present in a data set by creating new, linear variables in vector space termed principal components (PCs) [35]. A piece of code written in house at the University of York was used in Rstudio, the interface for the coding program R, was used for the analysis [36][37].

First, a summary of the importance of all components was generated, listing the principal components and the proportion of variance contained within them individually and cumulatively. Looking at these percentages and scree plots made to represent the variance in each PC graphically, it was decided that most of the necessary information was contained withing the first 6 principal components, which was confirmed by making a scree plot of the variances against the PC number to show their variances graphically.

Scatter plots were generated for all combinations of principal components 1 to 6, in order to compare variance between individual PC's. This was done by looking for visual separation between the modern and ancient data points shown on the graph. PCs deemed significant from this were pursued for further analysis.

Loadings were also made for PC's 1-6, showing each variables contribution to the principal component, which were exported and uploaded into Igor Pro 9. From here, they were labelled with the wavenumber values found from the multipeak fitting, and the values corresponding to points significantly distanced from the zero line were marked as such in the comparison table. It was worked out how many PC's each wavenumber value was significant in, and this was compared to the information found in the scatter plot and average subtraction. A table was made combining all this information, which was used to identify the biomarkers showing a significant intensity difference between the modern and ancient sample.

PCA has been used for multivariate analysis on parchment samples previously by Malissa et. al. [55], however a very limited range of wavenumbers was analysed (800-1000cm^{-1}). This thesis will expand on their work by providing PCA for more of the fingerprint region, as well

as the high wavenumber region. The combination of this work with the univariate analysis described in section 2.3.3 allows for detailed identification of significantly different biomarkers, whose heterogeneity will then be assessed.

2.3.5. Assessment of Molecular Heterogeneity

Rocha et. al. conducted a study in which heterogeneity specific to an individual biomarker was quantified through viewing the rate of convergence at its designated wavenumber as a function of the number of spectra included in each convergence group [27]. The data needed for this analysis was collected in the process of initial convergence testing detailed in section 2.3.2.

Plots were made for each wavenumber decided to be significant after univariate and multivariate analysis of the samples, and fitted with an exponential curve with equation,

$$N(x) = N_0 e^{-x/\tau} \tag{17}$$

where N(x) is the standard error of the average of x spectra, and τ is the decay constant. This decay constant can be used as a measure of heterogeneity, with a larger value meaning a slower decay and rate of convergence, therefore indicating a greater degree of heterogeneity. The exponential plots were made in Igor Pro 9, with the exponential fit function outputting values for τ^{-1} and the inverse of the error in this value. The error was calculated from its inverse using the following

$$\frac{error(\tau^{-1})}{\tau^{-1}} \times \tau = error(\tau). \tag{18}$$

The decay constants for the modern and ancient sample at each significant biomarker were tabulated alongside their error in order to quantitatively assess their heterogeneity.

3. Modern Parchment Results and Discussion

3.1. Modern Parchment Spectra

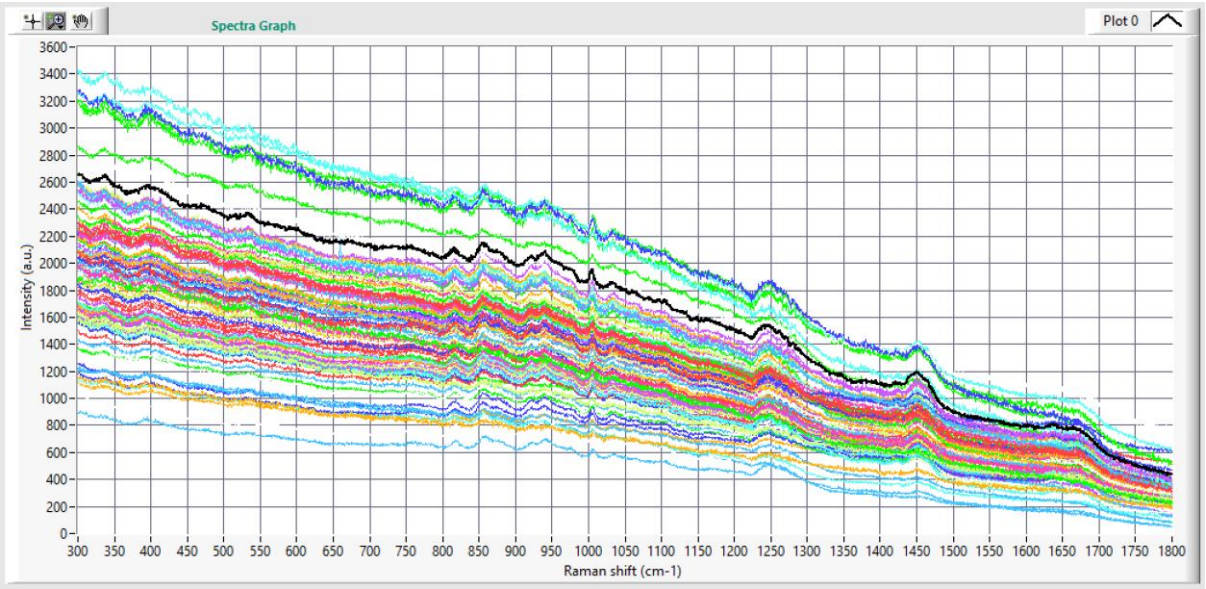


Figure 10: All the raw data collected for the fingerprint region of the modern sample on a single set of axes, displayed in Raman Tool Set. Includes 98 spectra in the range 300-1800cm^{-1}.

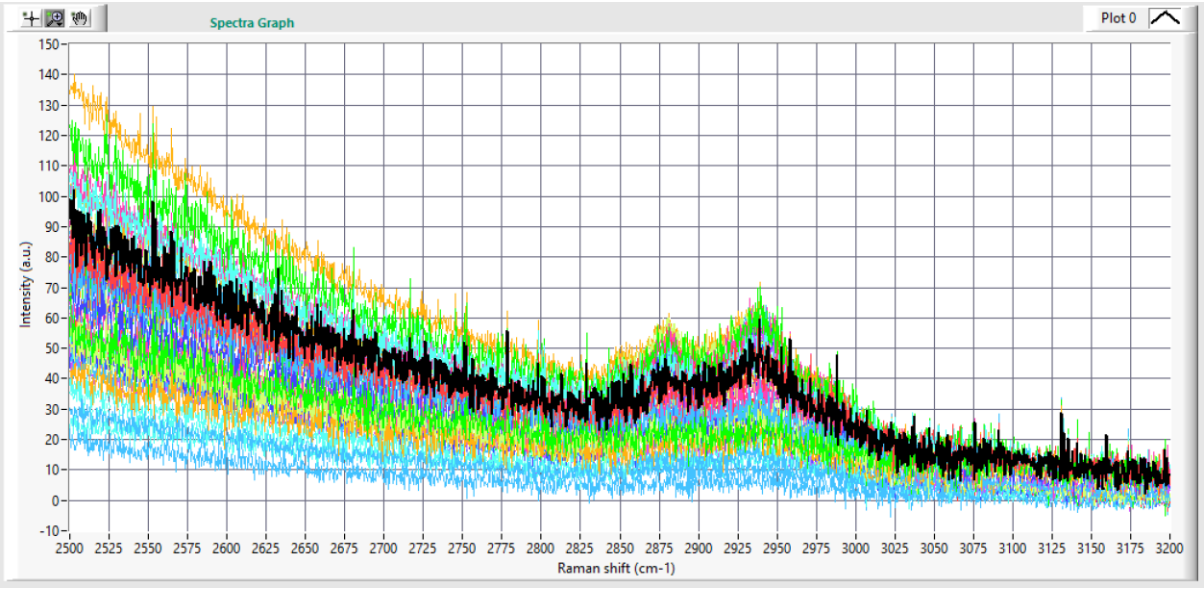


Figure 11: All the raw data collected for the high wavenumber region of the modern sample on a single set of axes, displayed in Raman Tool Set. Includes 95 spectra in the range 2500-3200cm^{-1}.

Figures 10 and 11 show all the raw data collected for analysis on the modern sample. It shows the complex nature of each spectra and the variation between them, highlighting the need for bespoke analysis. It is especially apparent at the beginning of the Raman profiles, where intensity varies the most between individual spectra. The processing done on the data does its best to eliminate this along with background and noise.

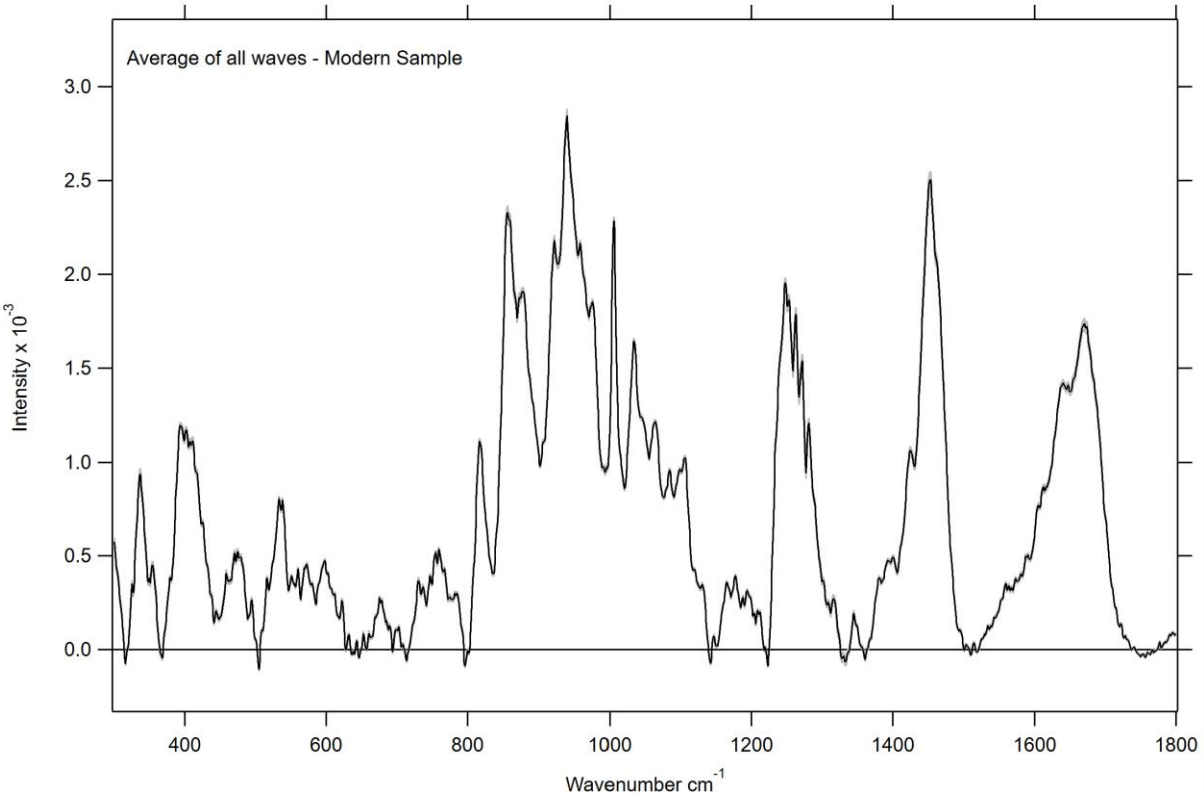


Figure 12: Shows the fully processed average of the raw data for the modern samples fingerprint region. 98 individually processed spectra are included in the average and the standard error in each value is shown with a grey envelope around the spectrum.

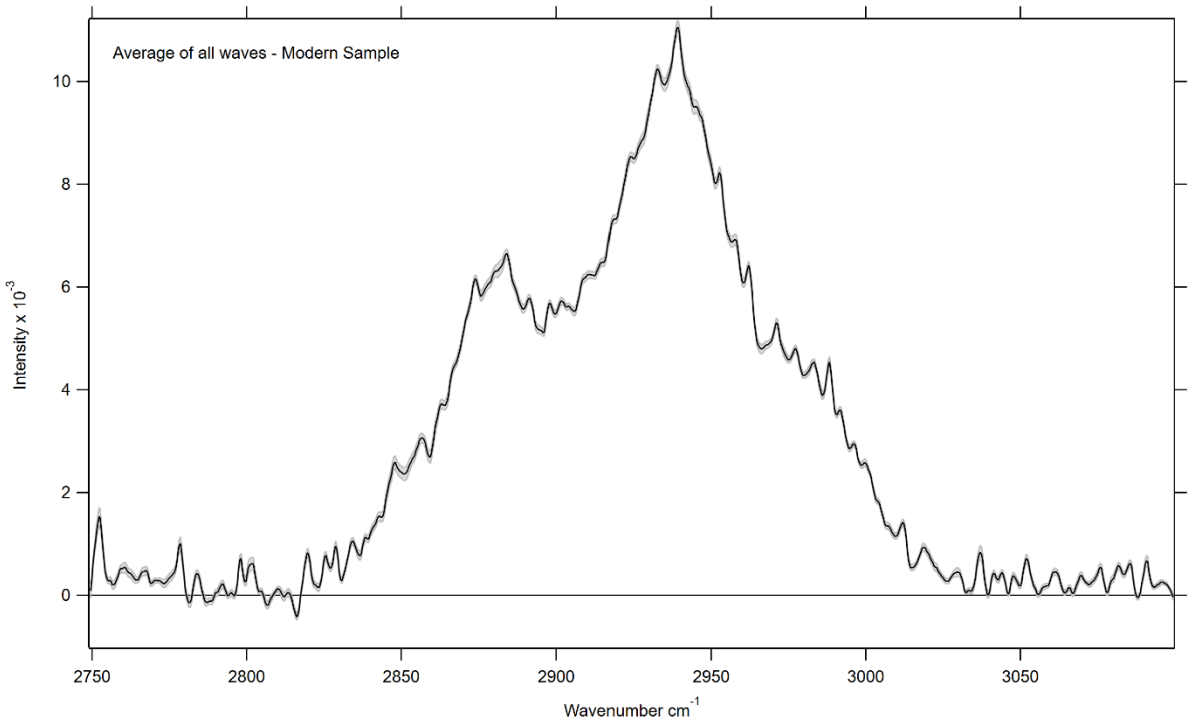


Figure 13: Shows the fully processed average of the raw data for the modern samples high wavenumber region. 95 individually processed spectra are included in the average and the standard error in each value is shown with a grey envelope around the spectrum.

Figures 13 and 14 show the fully processed spectra for both the fingerprint and high wavenumber regions of the modern sample's Raman profile. The spectra shown is an average across all the data sets taken. In the case of the fingerprint region this is an average of 98 spectra, and for the high wavenumber region it is an average of 95. As described in section 2.3.1, the pre-processing involved cutting the spectra to a common range, interpolating the wave, and baselining each spectra before normalising and smoothing them. Comparing figure 12 to figure 10 it is evident that in the processing was necessary as the spectral features are much clearer, allowing for better peak fitting for the deconvolution of the bands in the spectra. These bands are quantified below in table 1. Comparing figure 14 to figure 3 shows a lot of similarities, however there is limited literature to compare figure 13 to in order to assess its validity.

Fitting Region	Wave-number (cm^{-1})	Intensity (a.u)	±Error (a.u)	General Assignment	Detailed Assignment
FP1	324	1.99E-04	7E-06	-	-
FP1	337	9.60E-04	2E-05	-	-

FP1	355	4.60E-04	1E-05	Lime phase	$Ca(OH)_2$ [57]
FP2	404	1.262E-03	5E-06	Proteins	Skeletal deformation of collagen [44]
FP2	473	5.54E-04	9E-06	Carbohydrates, DNA	Polysaccharides, Glycogen [24][26], DNA [24]
FP3	534	7.62E-04	7E-06	Proteins, Carbohydrates, Lipids	S-S in Cysteine [26], Glucose, Cholesterol, Cholesterol Ester [24]
FP3	567	3.64E-04	3E-06	Proteins, DNA/RNA, Carbohydrates	Tryptophan [24], Cytosine, Guanine [24], Polysaccharides [26]
FP3	599	4.30E-04	4E-06	Lipids, Proteins	Amino Acids[43], Phospholipids [26]
FP4	677	2.44E-04	4E-06	DNA/RNA	Thymine and Guanine ring breathing [24] [26]
FP4	701	1.21E-04	3E-06	Proteins, Lipids	C-S Stretching: Amino Acids (methionine) [24][26], Cholesterol, Cholesterol Ester [24][26]
FP5	730	3.18E-04	5E-06	Proteins, DNA	Tryptophan [24], Proline [24], Adenine base [24]
FP5	758	5.64E-04	8E-06	Proteins	Tryptophan (symmetric and ring breathing) [24]
FP5	785	2.77E-04	4E-06	DNA/RNA	Thymine, Cytosine and Urasil ring breathing, [24] O-P-O backbone, Nucleic Acids. [24]
FP6	817	9.00E-04	2E-05	Proteins, DNA/ RNA	C-C stretching in collagen, Collagen[24], Proline, Hydroxyproline [24], Tyrosine [24][45] [26], PO_2^- stretching in nucleic acids[24].
FP7	856	1.56E-03	2E-05	Proteins, Carbohydrates	Tyrosine[24] [26], Collagen Backbone[24], Polysaccharides[26], Proline (Collagen), Hydroxyproline (Collagen[24]), C-O-C in Glycogen [24]
FP7	876	1.09E-03	3E-05	Proteins, Lipids, Carbohydrates	Hydroxyproline, Tryptophan, [24] $C\text{-}C\text{-}N^+$ symmetric stretching in Lipids[24], C-O-C ring in Carbohydrates, Collagen[24]
FP8	921	1.06E-03	2E-05	Proteins, Carbohydrates	Proline, Hydroxyproline [24], C-C stretch in Proline ring/Glucose, Praline ring

					(collagen) [24]
FP8	939	1.42E-03	4E-05	Proteins, Carbohydrates	Amino Acid chain vibrations [26], C-C skeletal stretching of Collagen backbone, Proline, Hydroxyproline [24], Glycogen[26], PPII helix in proteins[24][60]
FP8	958	1.141E-03	3E-06	Lipids	Cholesterol [24]
FP8	978	5.35E-04	3E-06	Proteins, Lipids, DNA/RNA	Phenylalanine, C-C Collagen Backbone, [24] C-C Stretching in disordered structure [24] [26][60], =CH bending in Lipids [24] Ribose [24]
FP9	1006	1.31E-03	1E-05	Proteins	Phenylalanine, [24] [26] Phenylalanine in Collagen [24]
FP10	1034	7.60E-04	2E-05	Proteins, Lipids, Carbohydrates	Collagen [24], Phenylalanine in Collagen [24] [26], C-H bending in Phenylalanine[24] [26], CH_2CH_3 bending in Collagen [26]and Phospholipids[24], C-N stretching in Proteins [24] [26], Carbohydrate residues of Collagen
FP10	1047	3.74E-04	8E-06	Proteins, Carbohydrates	Glycogen[24], Proline in Collagen[24], PO_4^{3-} symmetric stretching in Hydroxyapatite[24] [26]
FP10	1064	4.30E-04	1E-05	Proteins, Lipids, DNA/RNA	Proline in Collagen[24], C-C conformational change in Lipids [24] [26], PO_2^- stretching in DNA/RNA [24]
FP11	1085	1.97E-04	7E-06	Lipids, DNA/RNA, Lime Phase	C-C stretching[24] [26], Phosphodiester groups in nucleic acids[24] [26], CO_3^{2-}, PO_4^{3-} and C-C stretching of Lipids acyl backbone[24] [26], $CaCO_3$ [47]
FP11	1103	4.90E-04	1E-05	Proteins, Lipids	Phenylalanine[24], C-C stretching in Lipids [24]
FP12	1169	3.59E-04	6E-06	Proteins, Lipids, DNA/RNA	Tyrosine in Collagen I[24], C=C stretching and COH bending in Lipids[24],

					Cytosine, Guanine [24] [26]
FP12	1197	3.01E-04	3E-06	Proteins, DNA/RNA	Amide III[24], Nucleic acids, Phosphates[24]
FP13	1244	1.18E-03	2E-05	Proteins, Lipids, DNA/RNA	Amide III[26], CH2 wagging[24] and C-N stretching of Amide III in Collagen[24], Adenine, Cytosine, Thymine, Guanine[24] [26], Lipids[26]
FP13	1268	1.47E-03	3E-05	Proteins, Lipids	Amide III in Collagen and PPII helix proteins[24] [26][60], C-N stretching in PPII helix proteins[24][60], C-H in Lipids and Phospholipids[24] [26], C=C in unsaturated fatty acids[24]
FP13	1316	2.30E-04	2E-05	Proteins, Lipids, DNA/RNA	CH_2 and CH_3 twisting, wagging and bending modes in Collagen and Lipids, [24] Guanine ring breathing[24], Amide III in PPII helix[24][60],
FP14	1346	2.12E-04	2E-06	Proteins, Carbohydrates	CH_2 and CH_3 wagging in Collagen[24], Glucose [24]
FP15	1394	5.13E-04	3E-06	Lipids	CH_3 [24] [26], C-H rocking[24] [26],
FP15	1424	8.703E-04	8E-07	DNA/RNA	Deoxyribose[24], Adenine and Guanine ring breathing [24], NH in-plane deformation [24]
FP15	1447	1.76E-03	2E-05	Lipids, Proteins	Collagen[24], CH_2 deformation, CH_2 bending in proteins and Lipids[24], CH_2CH_3 deformation, C-H vibration[24] [26]
FP15	1464	1.61E-03	3E-05	Lipids, Proteins, DNA/RNA	CH_2CH_3 deformation in Lipids and Collagen[24], Deoxyribose[24]
FP16	1573	3.62E-04	8E-06	DNA/RNA	Guanine and Adenine ring breathing[24]
FP16	1614	6.88E-04	2E-06	Proteins, DNA/RNA	C=C mode in Tyrosine, Phenylalanine [24] [26]and Tryptophan[24], Cytosine[24] [26]
FP16	1637	6.70E-04	1E-05	Proteins	Amide I, [24]differences in Collagen content [24]
FP16	1671	1.72E-03	2E-05	Proteins,	Amide I in Collagen and

				Lipids	PPII helix[24][26][38][60], C=O stretching in disordered Amide I structure [26][38][60], C=C stretching [24] [26]
HW1	2883	5.73E-03	4E-05	Proteins, Lipids	CH_2 asymmetric stretching in Lipids and Proteins[24] [26][38], CH_3 in Lipids and Fatty Acids [24]
HW1	2938	9.66E-03	5E-05	Proteins, Lipids	CH_2 asymmetric stretching, CH_3 symmetric and asymmetric stretching in proteins[24], C-H vibrations in Lipids and Proteins[24] [26], CH_2 in Lipids and Fatty Acids[24]
HW1	2983	3.64E-03	4E-05	Lipids	CH_3 in Lipids, Fatty Acids, Cholesterol, and Cholesterol Ester[24], CH vibrations [24][38]

Table 1: All data collected from multipeak fitting done on the average spectra from the modern sample. Includes the wavenumbers of bands fit with gaussian peaks, along with their intensities and the standard error for the values. Peak assignments are also included based on an extensive literature comparison. A.u signifies arbitrary units, and the experimental error in the wavenumber values is $\pm 3cm^{-1}$ [54]

Table 1 presents a compilation of the bands fitted to the modern sample by multipeak fitting, including both their wavenumbers and intensities. Combined with this is an extensive comparison to existing literature to conclude the reasons for each band in the profile. The error in the Raman intensity of each band is the standard error of the measurement, and the error in each wavenumber presented was taken to be $\pm 3cm^{-1}$.

The literature examined has considered relevant bands from many sources including modern parchments, ancient parchments and unprocessed skins from multiple species. Edwards et al. conducted extensive analysis across samples of human, pig and goat skins and found small wavenumber variations, with the biggest differences being in the goat skin in comparison to the other two [3]. The largest difference observed is for the CH_3 asymmetric stretching peak that in table 1 has a wavenumber value of $2983cm^{-1}$. The band was deconvolved at $2975cm^{-1}$ for both human and pig skins, but at $2988cm^{-1}$ for goat skin, a difference of $13cm^{-1}$ [3**]** . A similar difference was noticed for CH_3 rocking, which was deconvolved at $958cm^{-1}$ in table 1.

Edwards et al. found that the band had a wavenumber of 970cm^{-1} for human skin, and 960cm^{-1} for pig and goat skin [3]. In table 1, both of these bands have been assigned to lipids.

Surface lipid content of skin has been shown to vary across species in a study by Nicolaides et. al. [39], who found a large difference in diester content across species, and a significantly larger ceramide and triglyceride content in human skin. The structural differences in these lipids could be the cause of the differences in wavenumbers observed, as the term lipid describes many types of molecule with varying functional groups and therefore covalent bond lengths. It is these differences in bond length that cause the bands to present at differing wavenumbers. Variations in wavenumber for lipid bands of skins from the same species could be due to the region the skin was taken from, as seen in a dermatological study by Cua et. al. [40]

Many bands attributed to phenylalanine, a component of collagen, can also be attributed to tyrosine, including 978cm^{-1}, 1006cm^{-1} and 1034cm^{-1}. This is due to tyrosine being an end product of the oxidation of phenylalanine [41]. They are similar in structure, differing only by the replacement of a hydrogen atom with an additional OH group in tyrosine, meaning they will produce similar Raman profiles. The reason for their prominence in the Raman profiles of parchments and skins is due to ring breathing, a symmetric stretching vibration within the aromatic ring in both of their structures. The 1614cm^{-1} band is attributed to tryptophan as well as phenylalanine and tyrosine, which also has an aromatic ring in its structure. Tryptophan is present in eggs, so its appearance in this table is in line with preparation methods described in section 1.1.

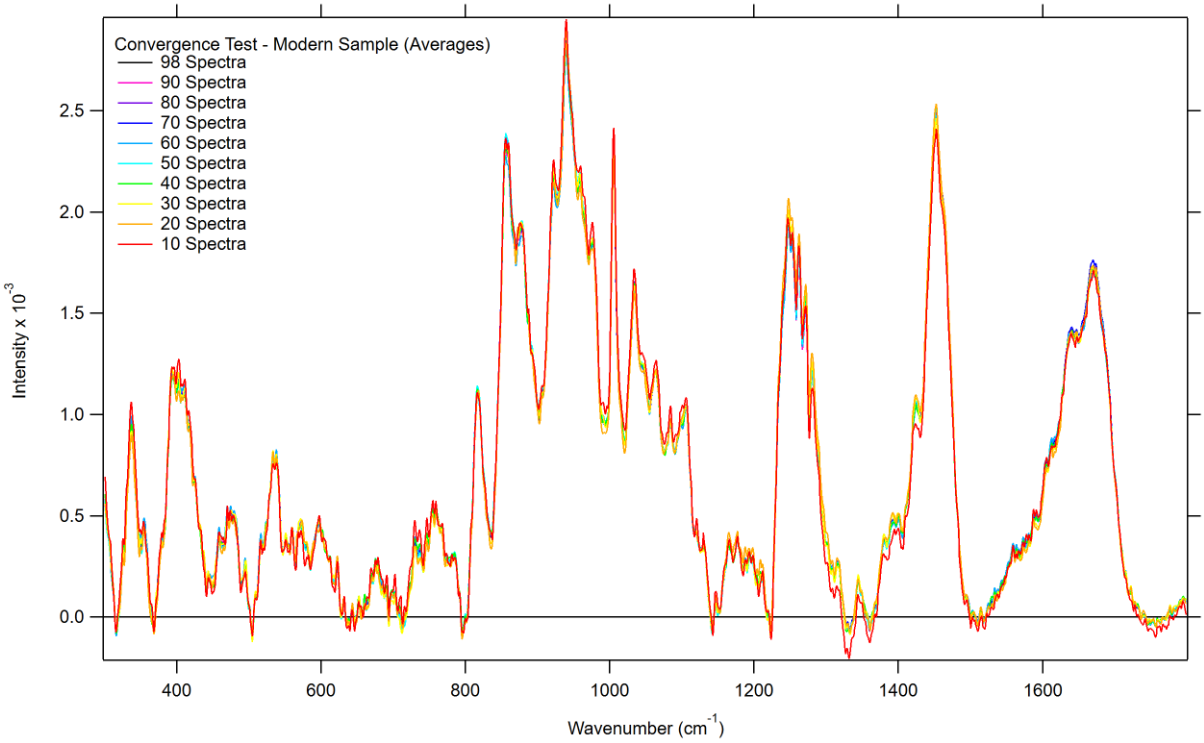
Work by Edwards et al shows the Raman spectra for cellulose, a material present in paper[3]. Paper has been known to have been used as a support for parchments, often documents were lined with it as a method of repairing the parchment surface. The bands in its spectra were quantified in a study by Agarwal, and show no overlap with any band in table 1. From this it can be confirmed that the modern sample has never been reinforced with paper, or any other cellulose based material [42].

Due to the nature of collagen structure, slight variations in wavenumber assignments in the literature can be expected [5]. The structure of the collagen chain sequence is Gly-X-Y, where X and Y are usually proline and hydroxyproline [11], however they can be any amino

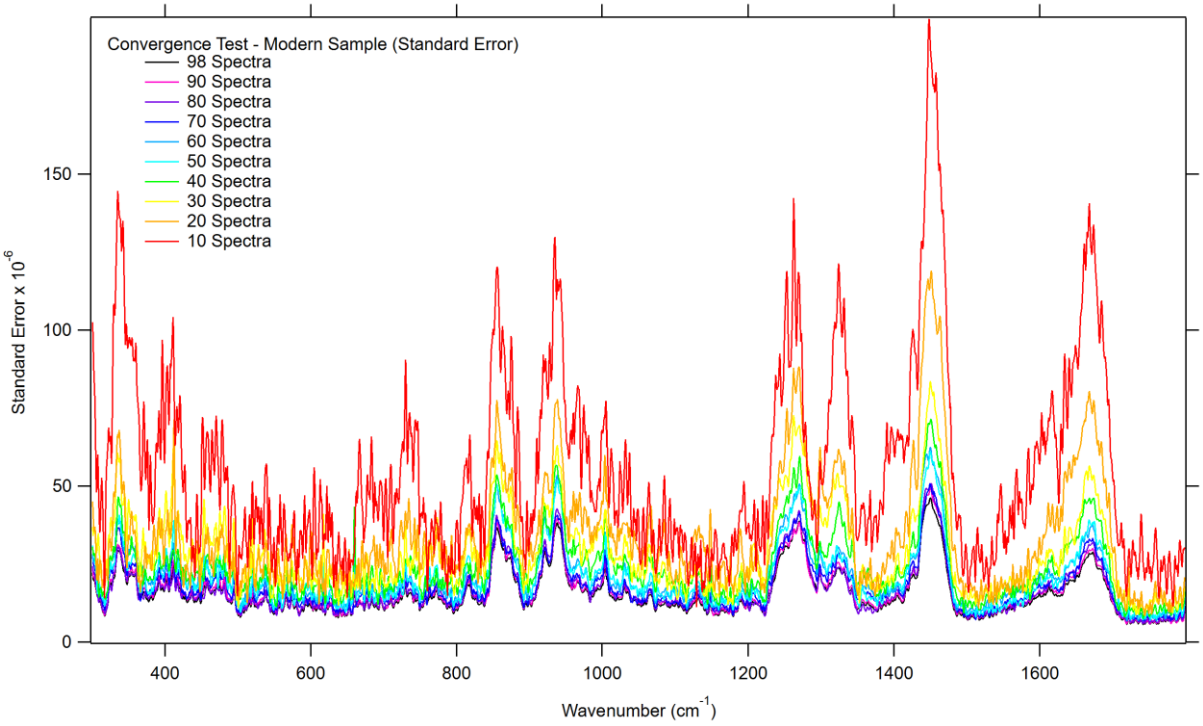
acid. Amino acids each have characteristic Raman profiles, with little variation[5]. This means any notable variations in a band assigned to collagen from existing literature is indicative of a difference in the amino acids involved in the Gly-X-Y polypeptide chains. For example, Carcamo et al noted many bands attributed to alanine[43], a potential amino acid that could be in the chain [8], not present in the data collected here, including but not limited to 1018cm^{-1}, 1147cm^{-1} and 1304cm^{-1}. The presence of many bands attributed to proline and hydroxyproline in table 1 suggests that the vast majority of the collagen in the skin sample used was composed of Gly-Pro-Hyp chains, and that any variation from this was not in a high enough quantity to be seen in the Raman profile of the parchment.

Many pieces of existing literature referring to parchments describe the structure of collagen as being an alpha helix. However, works done more specifically on collagen have instead identified these structures as a polyproline II-type helix (PPII), such as Shoulders and Raines [60]. Therefore, in table 1 and throughout, the structure of collagen has been referred to as a PPII helix.

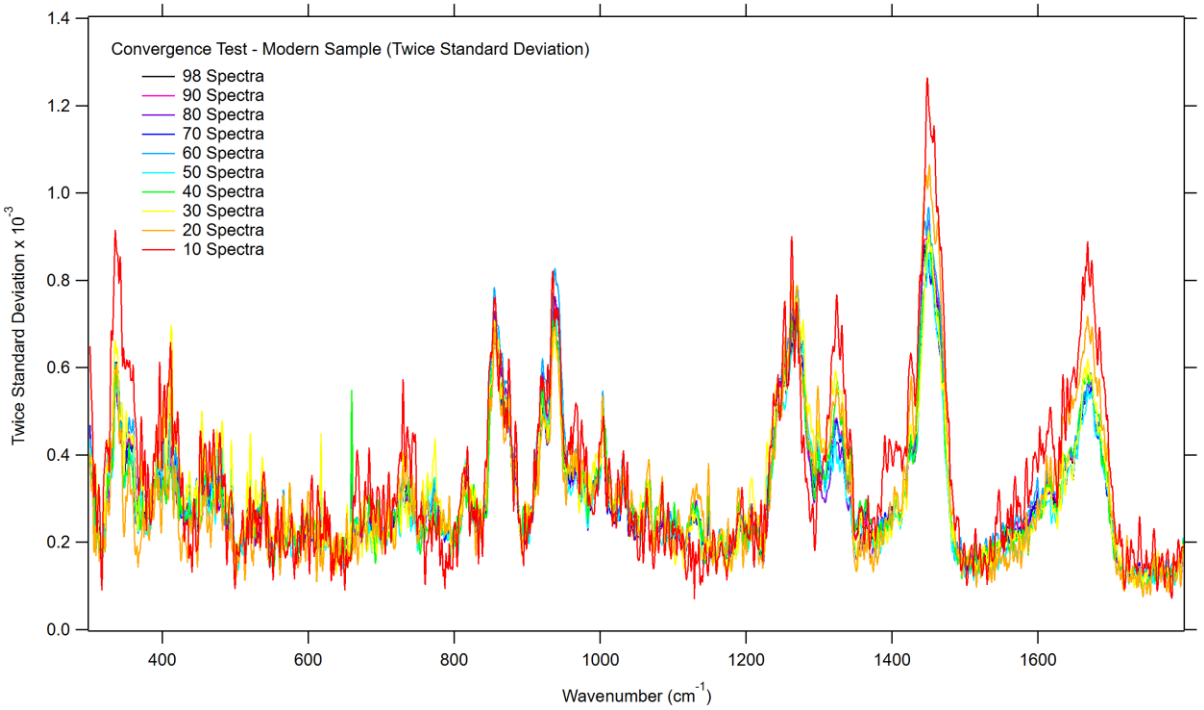
3.2. CONVERGENCE TESTS



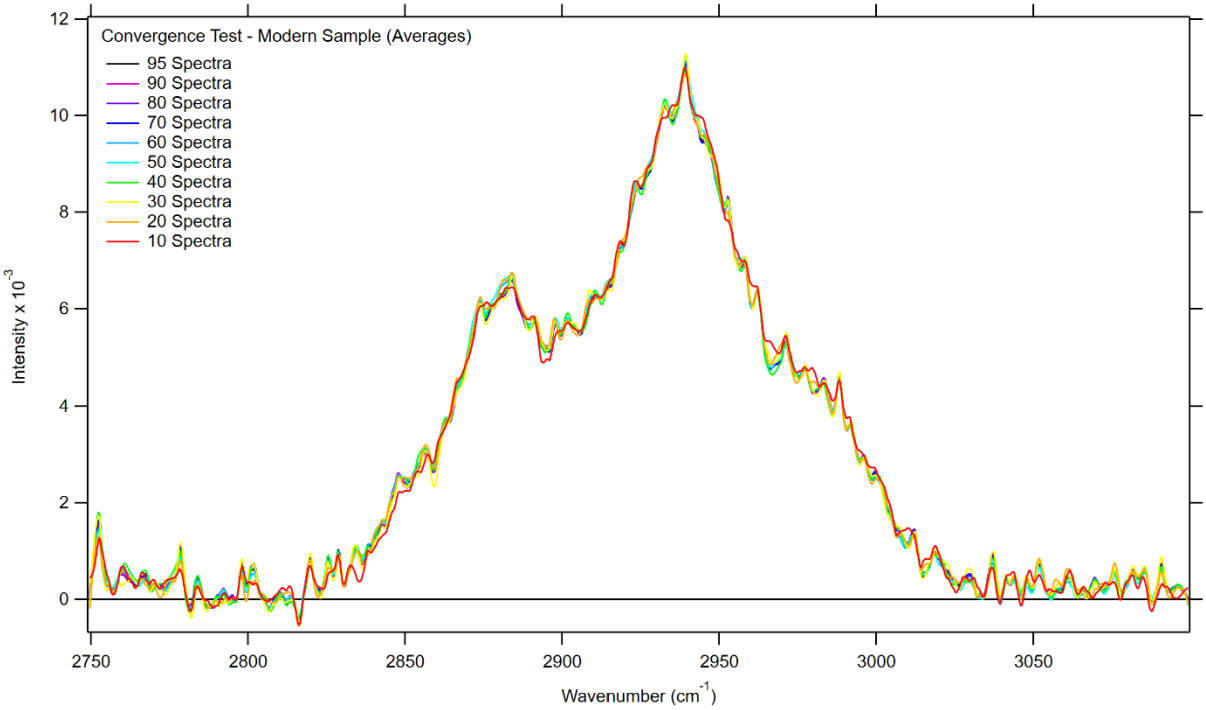
a) Convergence testing of the spectral average for the modern sample fingerprint region, across 98 spectra.



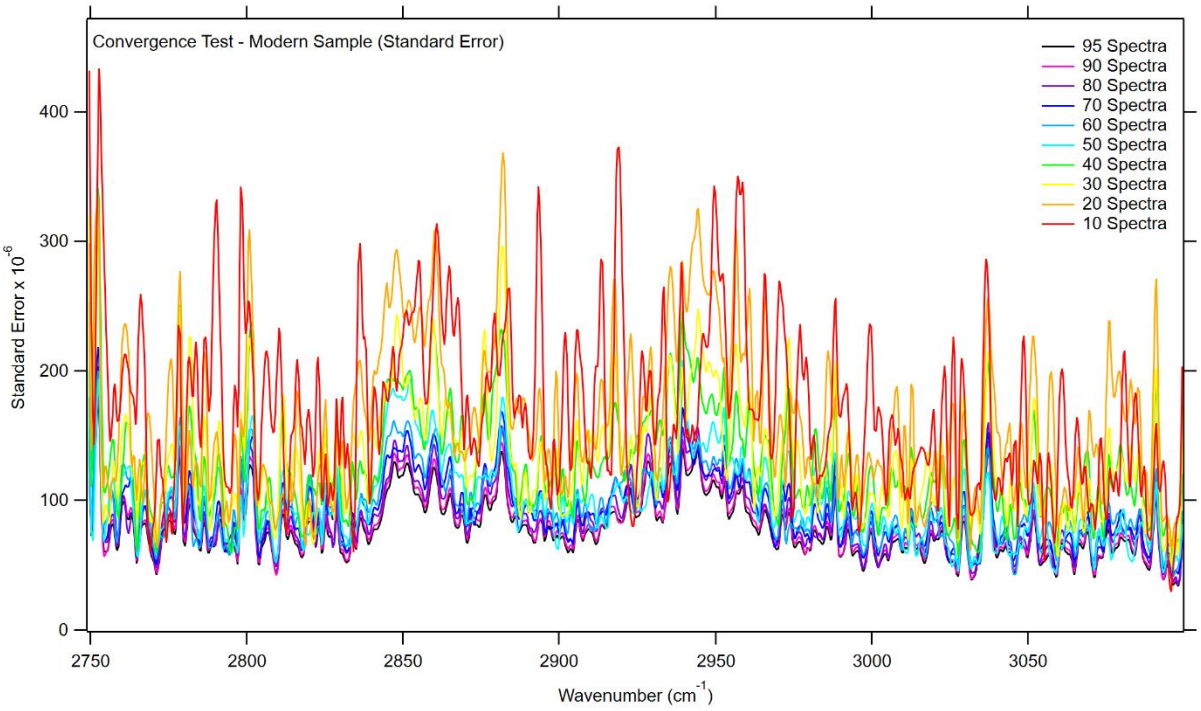
b) Convergence testing of the standard error in the spectral average for the modern sample fingerprint region, across 98 spectra.



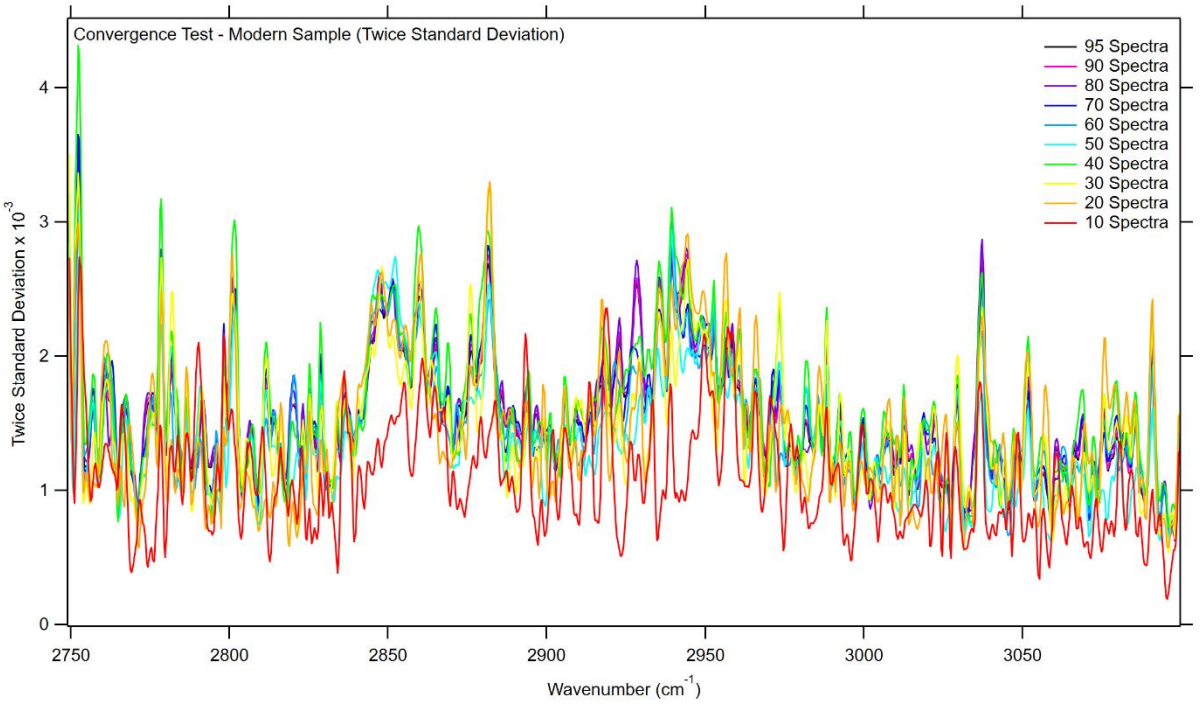
c) Convergence testing of the twice standard deviation in the spectral average for the modern sample fingerprint region, across 98 spectra.



d) Convergence testing of the spectral average for the modern sample high wavenumber region, across 95 spectra.



e) Convergence testing of the standard error in the spectral average for the modern sample high wavenumber region, across 95 spectra.



f) Convergence testing of the twice standard deviation in the spectral average for the modern sample high wavenumber region, across 95 spectra.

Figure 14: The convergence tests performed on the modern sample data. Includes convergence of the spectra averages, and the standard error and twice standard deviation in the average values for both the fingerprint and high wavelength regions.

Parts a-f of figure 14 show the convergence tests done on the modern sample, as described in section 2.3.2.

The first set of tests, in figure 14a and 14d are the convergences of the spectra averages for the fingerprint and high wavenumber regions respectively. Averages were taken in quantities of multiples of 10 until the full set was included, and plotted on the same axes to visually assess the point at which convergence was achieved. For the fingerprint region, the spectra were converged to within 2E-05 a.u., and in the high wavenumber region convergence was achieved to within 4E-05 a.u.

The same testing was done for the standard error in the band intensity values, shown in figure 14b for the fingerprint region and figure 14e for the high wavenumber region, again in increasing intervals of 10. In the fingerprint region, convergence was achieved to within 3E-06 a.u., and in the high wavenumber region the spectra were converged to within 7E-06 a.u.

The final test used in this work was convergence of the twice standard deviation of the band intensity, also tested in increasing intervals of 10. The test of the fingerprint region is seen in figure 14c, converged to within 4E-03 a.u., and the high wavenumber region is depicted in figure 14f, which is converged within 1E-04 a.u.

The rate of convergence can be used to assess the heterogeneity of a sample, with samples that converge with a smaller quantity of spectra being more homogenous [27]. The standard error values for the band intensity used in figures 14b and 14e will be used in section 4.5 to produce standard error decay plots at individual wavenumbers, in order to calculate a decay constant that could quantify heterogeneity at a molecular level. The same will be done for the ancient sample. An example of this is shown below in figure 15, at wavenumber 1637cm^{-1}, a band attributed to Amide I within collagen [24].

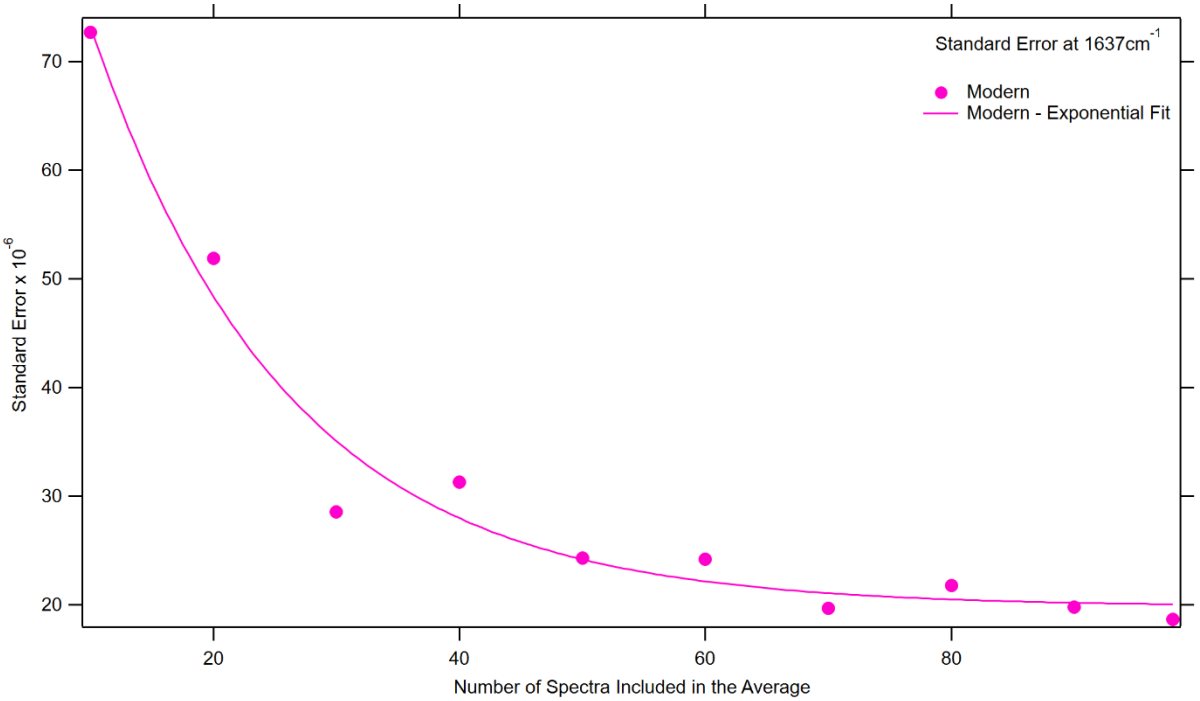


Figure 15: An example of a rate of convergence curve for the standard error at 1637cm^{-1} for the modern sample. Convergence was assessed with respect to the amount of spectra in the testing group, and the relationship shows an exponential curve with decay constant 16±2.

These tests are a new approach to assessing heterogeneity in parchments, which has been successfully carried out previously in another context by Rocha et al. [27]. A comparison will be made across the samples in the hope that it will be a useful tool for assessing the effects of degradation on parchment, and help characterise the differences between modern and ancient manuscripts.

4. Analysis of the Ancient Parchment and Comparison with the Modern Sample

4.1. Ancient Sample Spectra

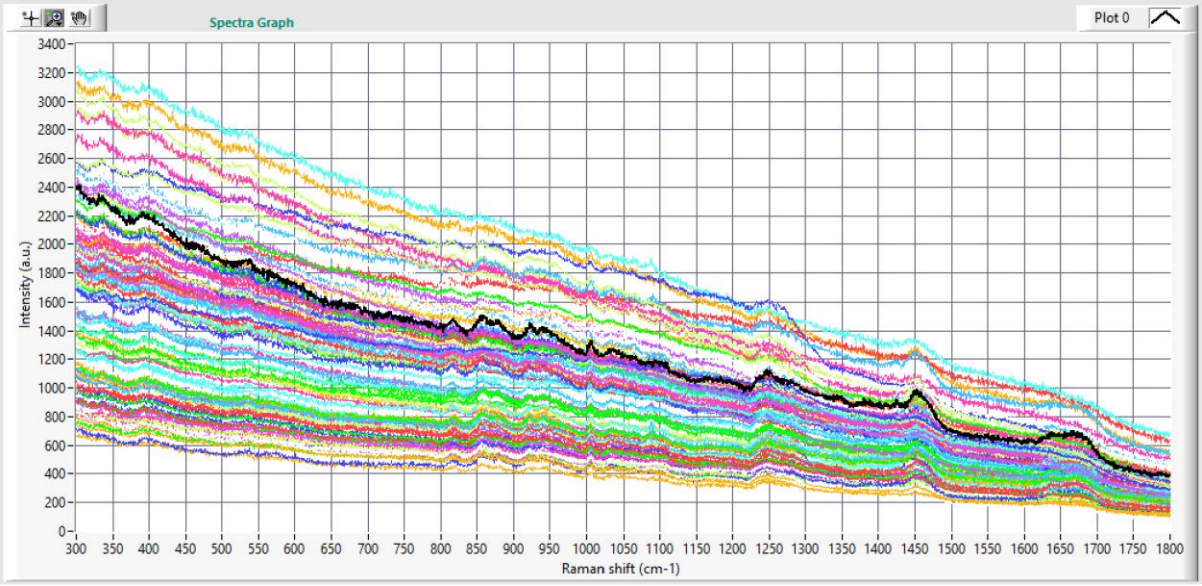


Figure 16: All the raw data collected for the fingerprint region of the ancient sample on a single set of axes, displayed in Raman Tool Set. Includes 100 spectra in the range 300-1800cm^{-1}.

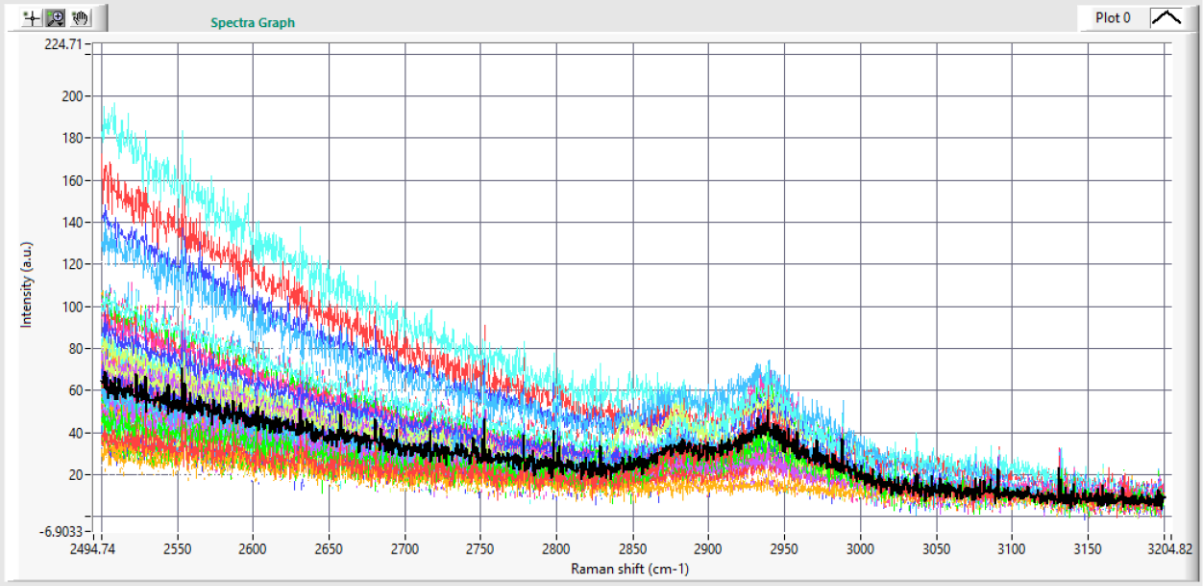


Figure 17: All the raw data collected for the fingerprint region of the modern sample on a single set of axes, displayed in Raman Tool Set. Includes 100 spectra in the range 2500-3200cm^{-1}.

Figures 16 and 17 show the full set of data collected for the analysis of the ancient sample. As with the modern sample raw data in figures 10 and 11, there is a wide variation between individual spectra. The same fluorescence is also observed. It would be expected that the ancient sample would have more fluorescence or other background radiation present in the spectra due to contaminants. However, the above data shows very little discernible difference, potentially due to increased fluorescence in the modern sample caused by the curvature seen in figure 6. Therefore, the additional processing performed on the data before analysis was especially necessary.

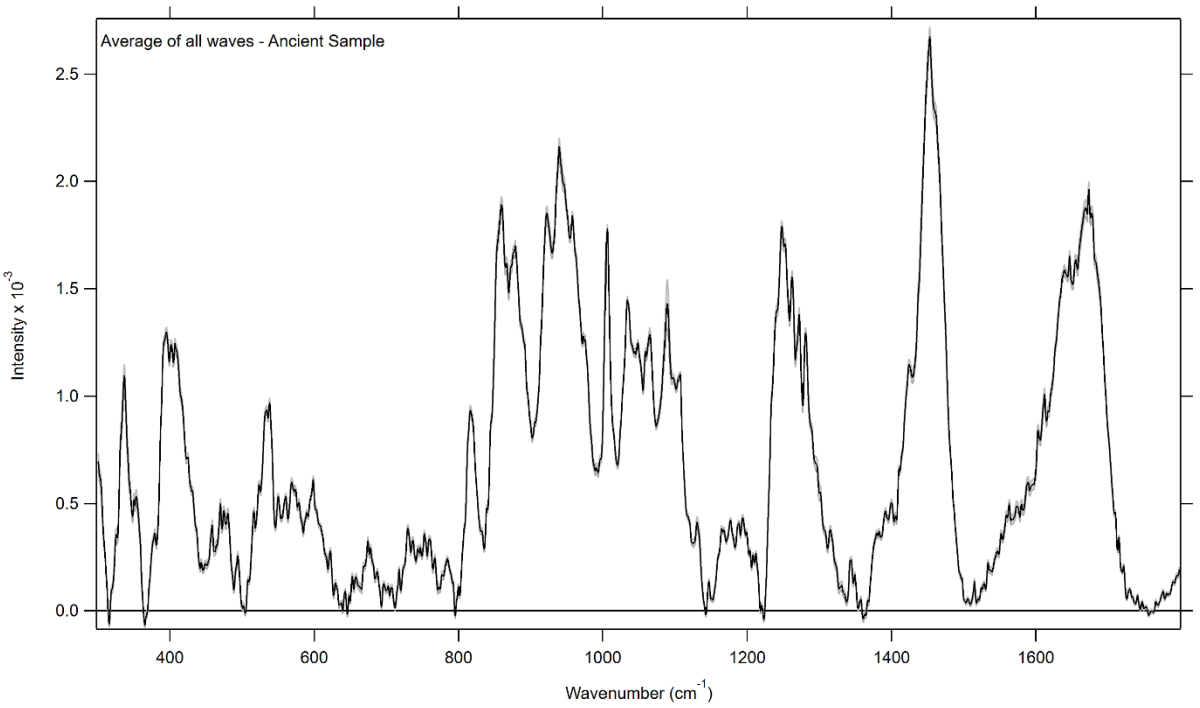


Figure 18: Shows the fully processed average of the raw data for the ancient samples fingerprint region. 100 individually processed spectra are included in the average and the standard error in each value is shown with a grey envelope around the spectrum.

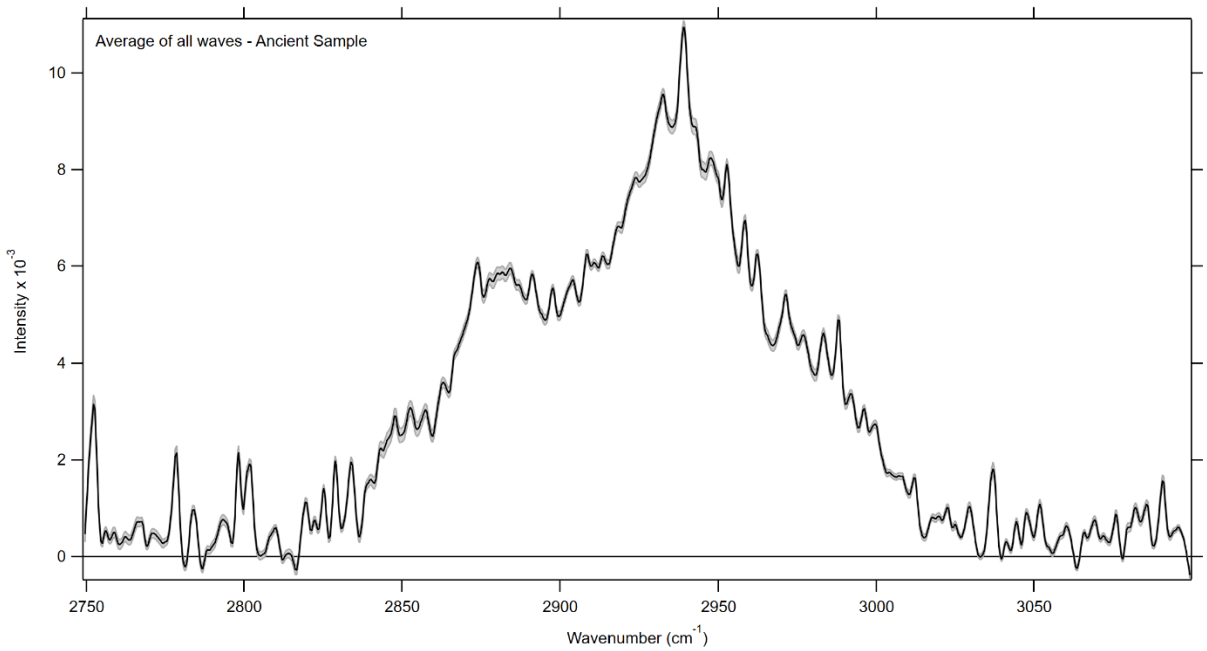
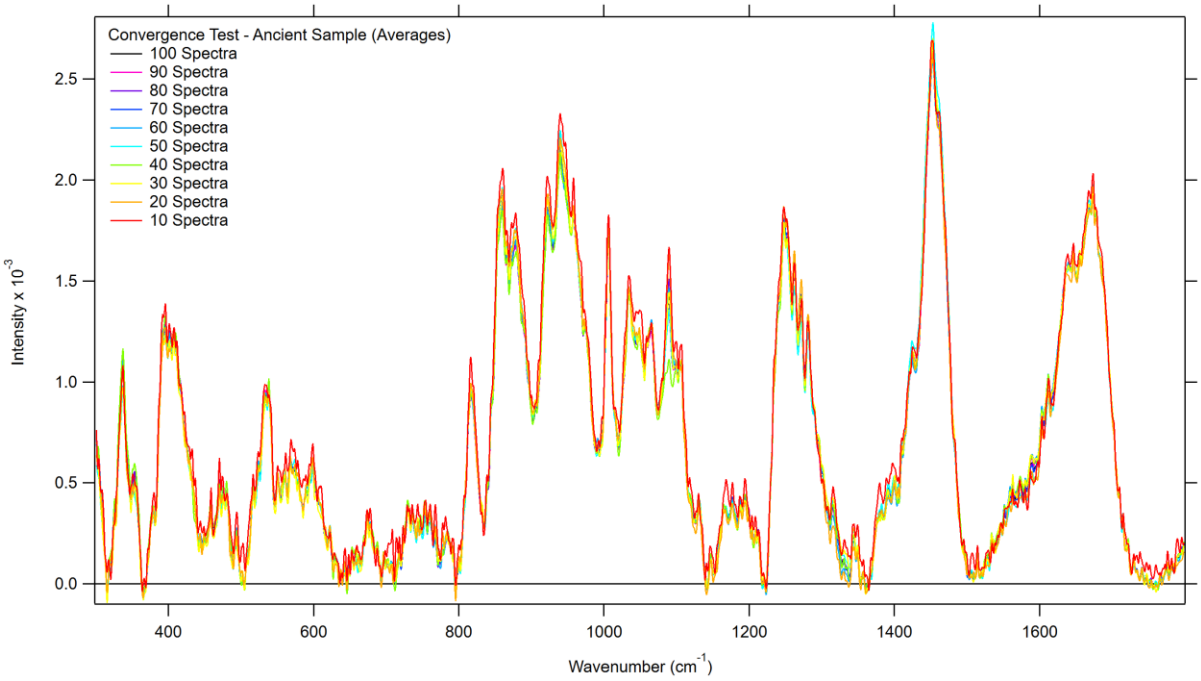


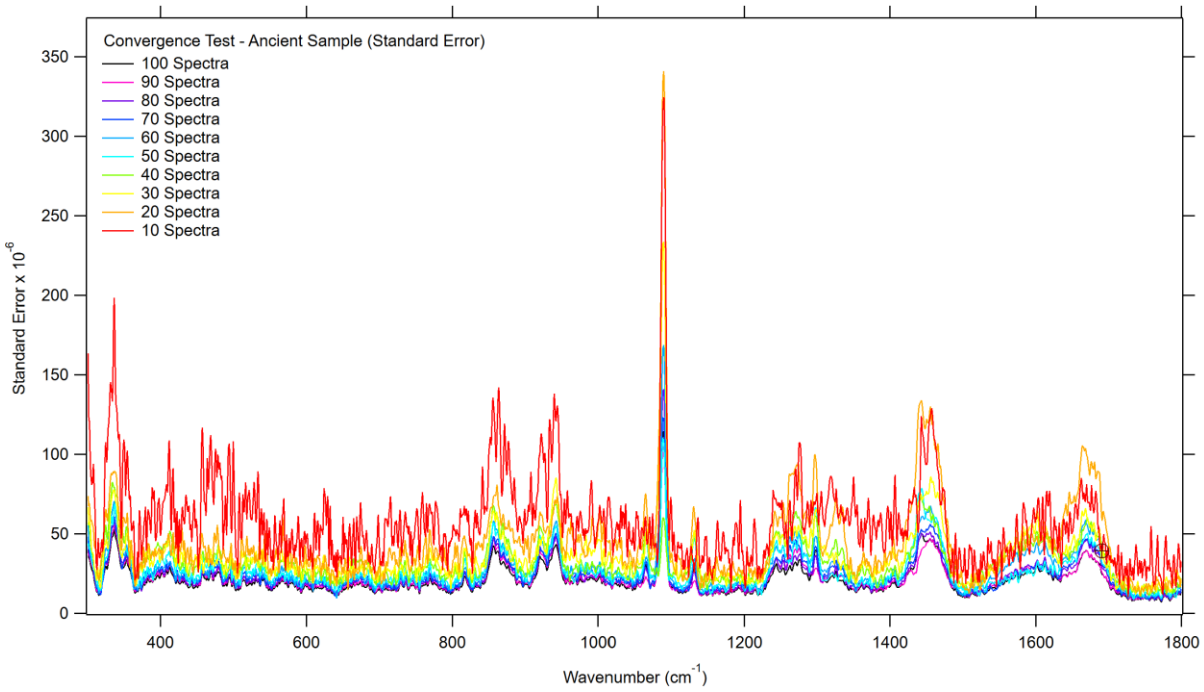
Figure 19: Shows the fully processed average of the raw data for the ancient samples fingerprint region. 100 individually processed spectra are included in the average and the standard error in each value is shown with a grey envelope around the spectrum.

Figures 18 and 19 show the fully processed data for the ancient sample, in the fingerprint and high wavenumber regions respectively. The spectra shown are averages over 100 data sets and the grey area shows the standard error in the intensity measurements.

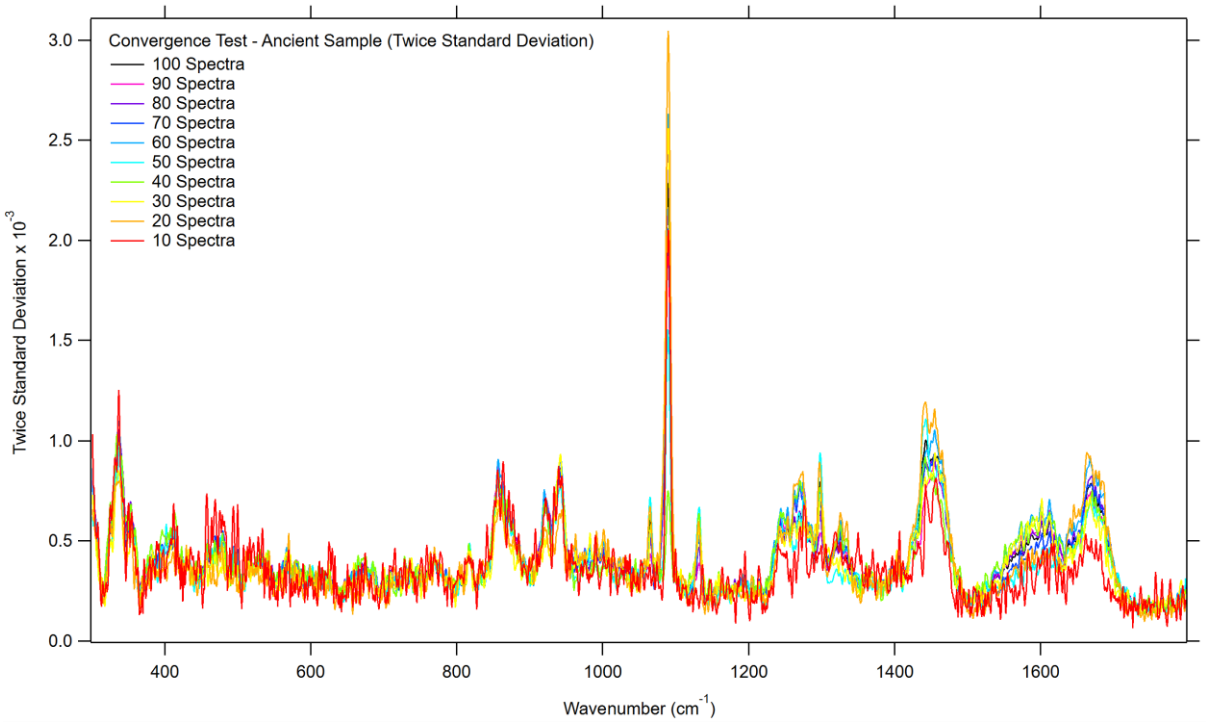
4.2. CONVERGENCE TESTS



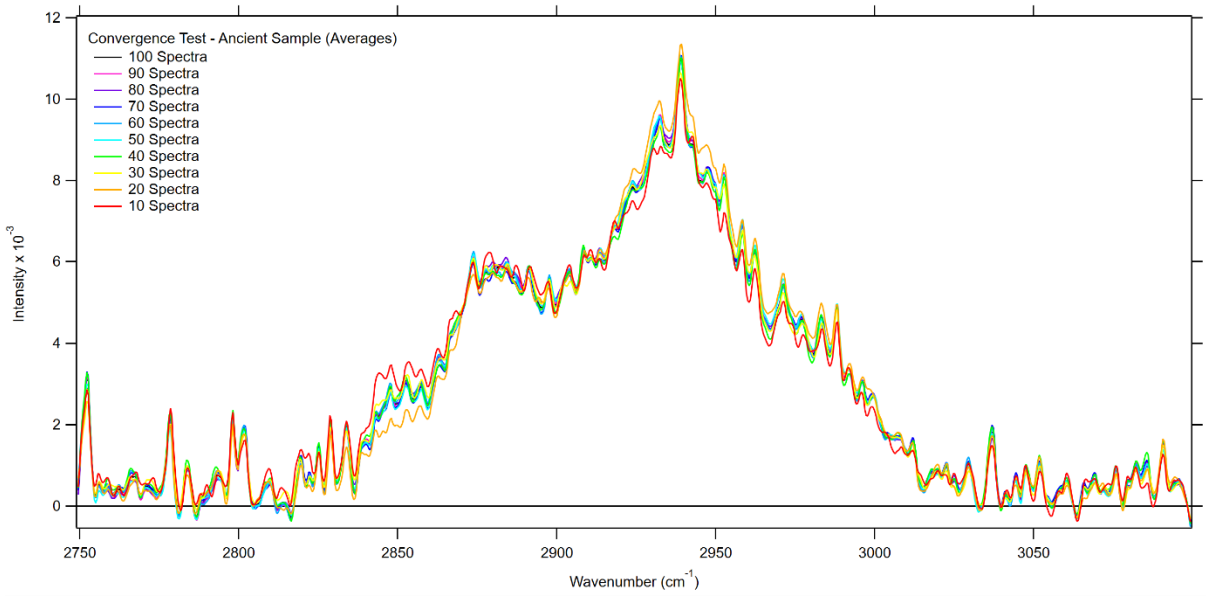
a) Convergence testing of the spectral average for the ancient sample fingerprint region, across 100 spectra.



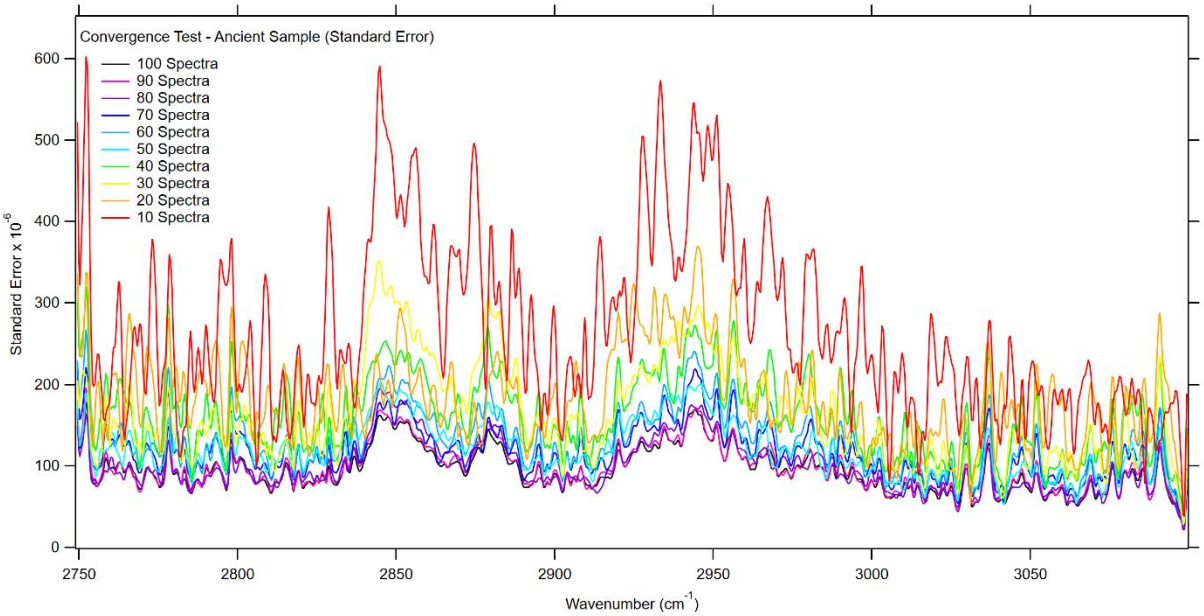
b) Convergence testing of the standard error in the spectral average for the ancient sample fingerprint region, across 100 spectra.



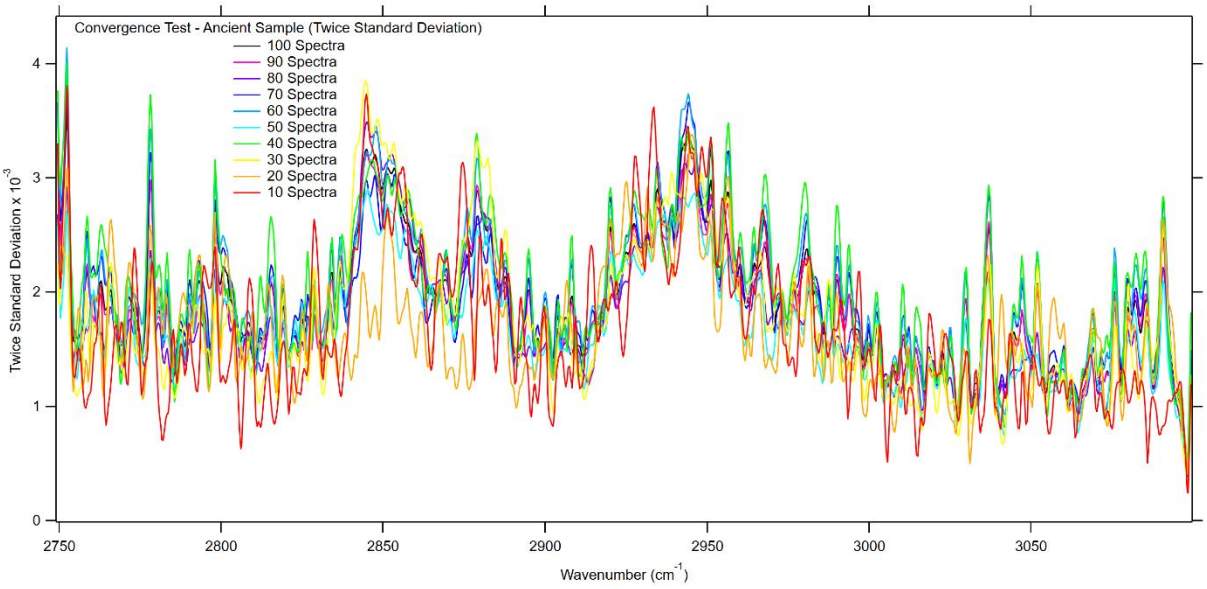
c) Convergence testing of the twice standard deviation of the spectral average for the ancient sample fingerprint region, across 100 spectra.



d) Convergence testing of the spectral average for the ancient sample high wavenumber region, across 100 spectra.



e) Convergence testing of the standard error in the spectral average for the ancient sample high wavenumber region, across 100 spectra.



f) Convergence testing of the twice standard deviation of the spectral average for the ancient sample high wavenumber region, across 100 spectra.

Figure 20: The convergence tests performed on the ancient sample data. Includes convergence of the spectral averages, and the standard error and twice standard deviation in the average values for both the fingerprint and high wavelength regions.

Figure 20 parts a-f show the convergence tests performed on the ancient sample. The tests are identical to those performed on the modern sample in section 3.2.

The first set of convergence tests on the average of different quantities of spectra is shown for the fingerprint region in figure 20a and for the high wavenumber region in figure 20d. Convergence is achieved in the fingerprint region to within 8E-05a.u, and in the high wavenumber region it was achieved within 3E-04.

The convergence of the standard error in the intensity values is shown in figures 20b and 20e. It is achieved to within 1E-05 a.u. for the fingerprint region in figure 20b, and to within 2E-05 a.u. for the high wavenumber region in figure 20e.

Figures 20c and 20f show the convergence of the twice standard deviation of the band intensities, for the fingerprint and high wavenumber region respectively. Convergence of the twice standard deviation was achieved to within 3E-04 a.u. for the fingerprint region, and within 1E-04 a.u. for the high wavenumber region.

For all but one convergence test carried out across figures 14 and 20, the modern sample achieved closer convergence. This speaks for the heterogeneity of the samples, as the better convergence of the modern sample proves fewer variations across the parchment and therefore higher homogeneity. The exception to this is the convergence of the twice standard deviation of the high wavelength region, seen in figures 14f and 20f, which are both converged within 1E-04 a.u., indicating no difference between the samples. This is the first quantitative indication of digenetic effects on the older parchment, which will be explored further in the following sections. Its heterogeneity in particular will be discussed in section 4.5, using the convergence information gathered here and in section 3.2.

4.3. Band Identification and Univariate Analysis

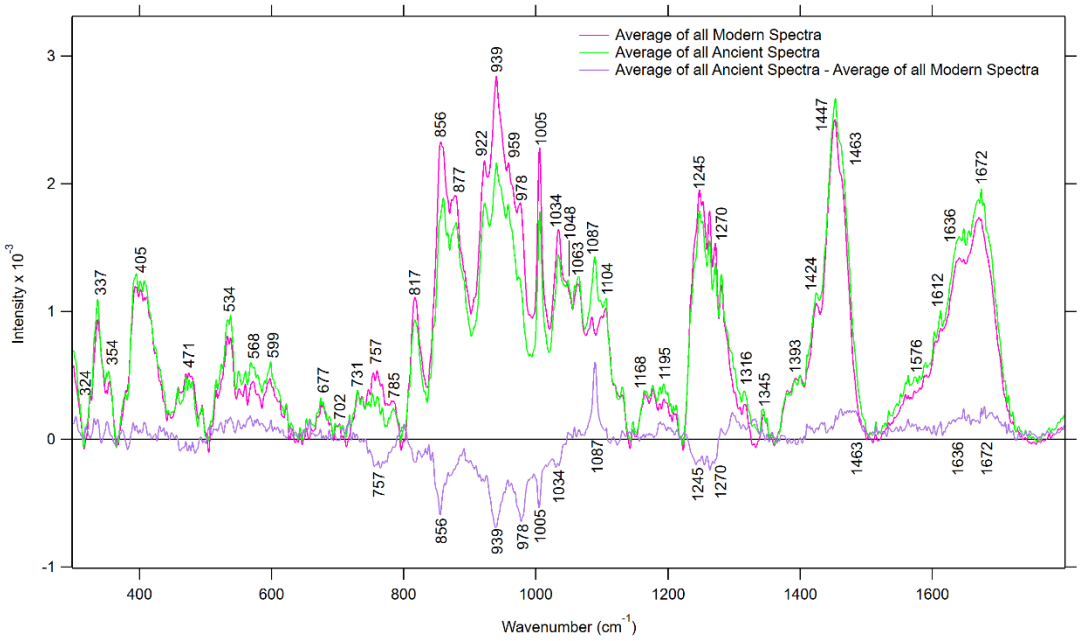


Figure 21: The average intensity plotted against wavenumber for the fingerprint region of both samples, with the wavenumbers for the peaks found during peak fitting labelled above the traces. A wave of the average intensity of the ancient sample with the modern samples average intensity subtracted from it is also included – and the places with notable differences are defined by labelling their wavenumbers below the trace.

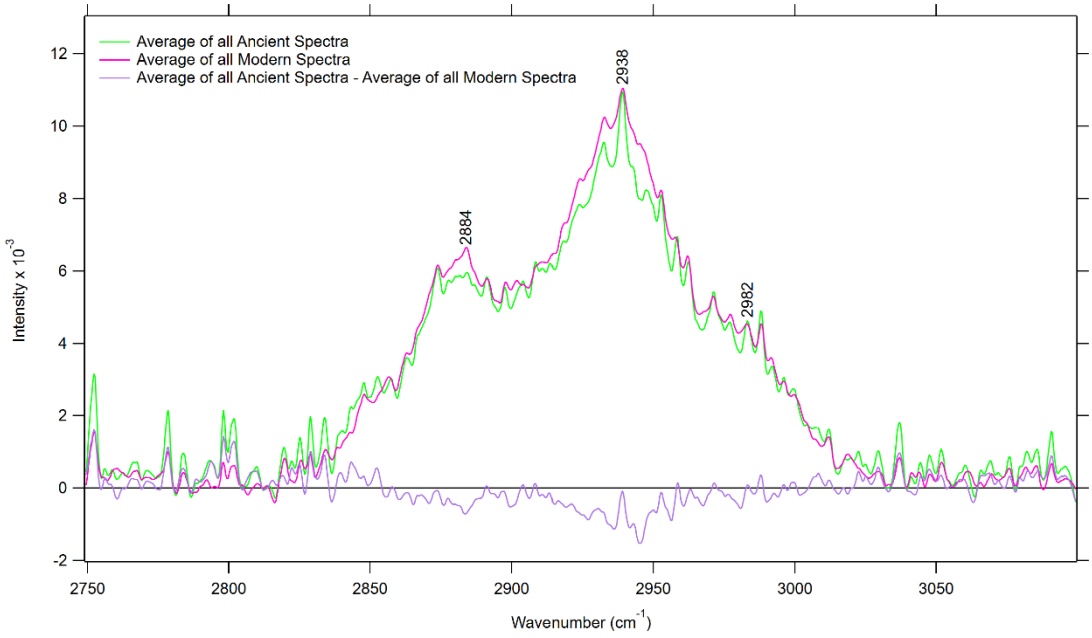


Figure 22: The average intensity plotted against wavenumber for the high wavenumber region of both samples, with the wavenumbers for the peaks found during peak fitting labelled above the traces. A wave of the average intensity of the ancient sample with the modern samples average intensity subtracted from it is also included.

WN (cm^{-1}) (Modern)	WN (cm^{-1}) (Ancient)	Av. WN (cm^{-1})	ΔAv. WN (cm^{-1})	Intensity (Modern)	Intensity SE (Modern)	Intensity (Ancient)	Intensity SE (Ancient)
324.47	323.00	324	0.7	1.99E-04	7E-06	2.20E-04	3E-05
337.42	336.51	337	0.5	9.60E-04	2E-05	1.11E-03	6E-05
355.41	353.43	354	1.0	4.60E-04	1E-05	5.90E-04	3E-05
404.84	404.54	405	0.2	1.262E-03	5E-06	1.30E-03	5E-05
472.76	470.14	471	1.0	5.54E-04	9E-06	3.90E-04	8E-05
533.79	533.95	534	0.08	7.62E-04	7E-06	8.31E-04	8E-06
567.37	567.77	568	0.2	3.64E-04	3E-06	4.67E-04	6E-06
599.09	598.88	599	0.1	4.30E-04	4E-06	4.70E-04	2E-06
677.02	676.04	677	0.5	2.44E-04	4E-06	2.20E-04	2E-05
700.94	703.44	702	1.0	1.21E-04	3E-06	7.00E-05	2E-05
730.32	730.93	731	0.3	3.18E-04	5E-06	2.70E-04	1E-05
758.42	755.87	757	1.0	5.64E-04	8E-06	3.10E-04	2E-05
784.81	785.46	785	0.3	2.77E-04	4E-06	2.30E-04	2E-05
817.09	817.21	817	0.06	9.00E-04	2E-05	7.54E-04	5E-06
855.59	856.43	856	0.4	1.56E-03	2E-05	1.22E-03	3E-05
875.93	878.06	877	1.0	1.09E-03	3E-05	1.00E-03	1E-05
921.31	921.85	922	0.3	1.06E-03	2E-05	9.90E-04	3E-05
938.83	939.65	939	0.4	1.42E-03	4E-05	1.19E-03	3E-05
958.35	959.19	959	0.4	1.141E-03	3E-06	1.019E-03	2E-06
977.89	977.47	978	0.2	5.35E-04	3E-06	3.31E-04	4E-06
1005.7	1005.9	1006	0.1	1.31E-03	1E-05	1.048E-03	2E-06
1033.8	1033.9	1034	0.05	7.60E-04	2E-05	6.80E-04	3E-05
1047.4	1048.2	1048	0.4	3.74E-04	8E-06	4.30E-04	3E-05
1063.6	1063.3	1064	0.2	4.30E-04	1E-05	4.50E-04	5E-05
1084.8	1089.2	1087	2.0	1.97E-04	7E-06	6.80E-04	8E-05
1103.4	1104.8	1104	0.7	4.90E-04	1E-05	5.50E-04	1E-05
1168.8	1166.7	1168	1.0	3.59E-04	6E-06	2.63E-04	2E-06
1197.4	1192.9	1195	2.0	3.01E-04	3E-06	3.86E-04	2E-06
1244.2	1245.2	1245	0.5	1.20E-03	2E-05	1.179E-03	1E-06
1267.9	1272.2	1270	2.0	1.47E-03	3E-05	1.16E-03	1E-05
1315.5	1317.2	1316	0.9	2.30E-04	1E-05	1.75E-04	7E-06
1345.8	1345.2	1346	0.3	2.12E-04	2E-06	1.70E-04	1E-05

1394.0	1391.1	1393	2.0	5.13E-04	3E-06	4.40E-04	5E-05
1423.8	1424.6	1424	0.4	8.703E-04	8E-07	9.70E-04	1E-05
1447.1	1447.4	1447	0.2	1.76E-03	2E-05	1.20E-03	1E-04
1464.4	1462.0	1463	1.0	1.61E-03	3E-05	1.95E-03	3E-05
1572.9	1578.8	1576	3.0	3.62E-04	8E-06	4.70E-04	1E-05
1614.1	1610.3	1612	2.0	6.88E-04	2E-06	5.10E-04	3E-05
1637.4	1635.8	1637	0.8	6.70E-04	1E-05	9.89E-04	9E-06
1671.1	1672.9	1672	0.9	1.72E-03	2E-05	1.86E-03	3E-05
2883.1	2885.3	2884	1.0	5.73E-03	4E-05	5.21E-03	4E-05
2938.3	2938.6	2938	0.2	9.66E-03	5E-05	8.17E-03	4E-05
2983.1	2981.3	2982	0.9	3.64E-03	4E-05	3.51E-03	4E-05

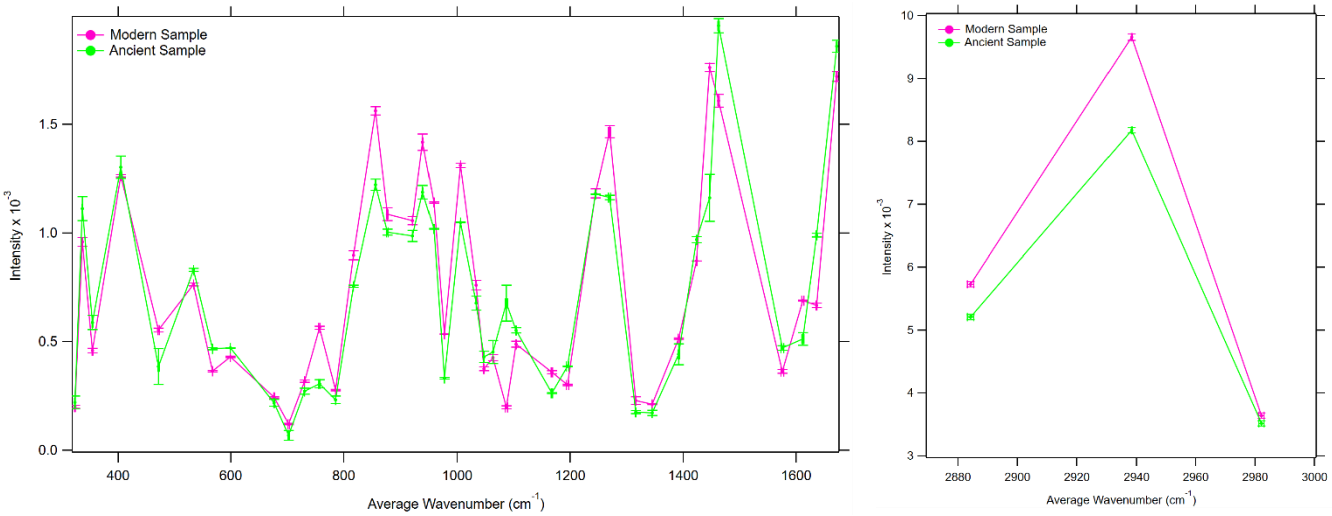
Table 2: Displays the quantitative fittings done on both samples including wavenumber and intensity values, along with errors for both. The shading indicates the sets of peaks fitted in a group.

Figures 21 and 22 show the average of all spectra taken for both samples on the same set of axes, along with a third spectra depicting the subtraction of the modern parchment Raman profile from that of the ancient parchment. This plot indicates the differences between the two samples above the spectral background. All deconvolved bands fit with the method described in section 2.3.3 are displayed in table 2. The average wavenumber given was calculated as the mean of the wavenumbers found for the two samples, presented as an integer in line with what is seen commonly in literature [24]. The Δ values for the average wavenumber were calculated by finding half the difference between the wavenumber values of the two samples. For every band fitted were lower than or equal to the spectral resolution of the instrument, $\pm 3 cm^{-1}$, so this was considered to be the error in the measurements. The Intensity values for each band are also quantified in table 2 and their standard error is given.

The wavenumbers labelled below the zero line indicate the areas with the biggest differences between the two spectra intensities that correspond to the bands fit in table 1. It is seen that all of the bands with differences seen in the fingerprint region presented in figure 21 can be attributed to collagen, or one of its structural components such as phenylalanine ($1006 cm^{-1}$) or an amide group ($1245 cm^{-1}$ and $1270 cm^{-1}$ for amide III, $1637 cm^{-1}$ and $1672 cm^{-1}$ for amide I) [24][26]. This immediately suggests an agreement with existing literature stating that collagen is susceptible to degradation while a parchment ages, and that amide groups can be characteristic indicators of this [17]. This will be further quantified in sections to follow.

In the case of the high wavelength region in figure 22, the wavenumber values found by the peak fitting method did not match the regions where the subtraction showed a significant difference between the two data sets, and no information was able to be obtained.

These two plots alone do not provide enough evidence to prove a significant difference between the two samples at any point, so further analysis was conducted. A scatter plot was made for each region to compare the average intensities for each sample and their errors, shown below in figure 23.



a) A scatter plot of peaks fitted in the fingerprint region.

b) A scatter plot of peaks fitted in the high wavenumber region.

Figure 23: Scatter plots of the intensity against the average wavenumber for both regions of both sets of data, including error bars in both the x and y directions. The information plotted is given in Table 3.

Average Wavenumber (cm^{-1})	Error Bar Separation (a.u.)	Average Wavenumber (cm^{-1})	Error Bar Separation (a.u.)
324	-1.44E-05	1048	2.17E-05
337	7.73E-05	1063	-4.28E-05
354	8.43E-05	1087	3.90E-04
405	-1.64E-05	1104	3.78E-05
471	7.67E-05	1168	8.76E-05
534	5.37E-05	1195	8.01E-05
568	9.28E-05	1245	-1.99E-05
599	3.35E-05	1270	2.63E-04
677	5.94E-06	1316	3.02E-05
702	2.61E-05	1346	2.60E-05
731	2.91E-05	1393	2.06E-05
757	2.30E-04	1424	8.46E-05
785	2.35E-05	1447	4.73E-04
817	1.16E-04	1463	2.85E-04
856	2.94E-04	1576	9.11E-05
877	3.94E-05	1612	1.46E-04
922	2.60E-05	1637	3.02E-04
939	1.60E-04	1672	8.72E-05
959	1.18E-04	2884	4.39E-04
978	1.97E-04	2938	1.39E-03
1006	2.50E-04	2982	4.36E-05
1064	2.78E-05		

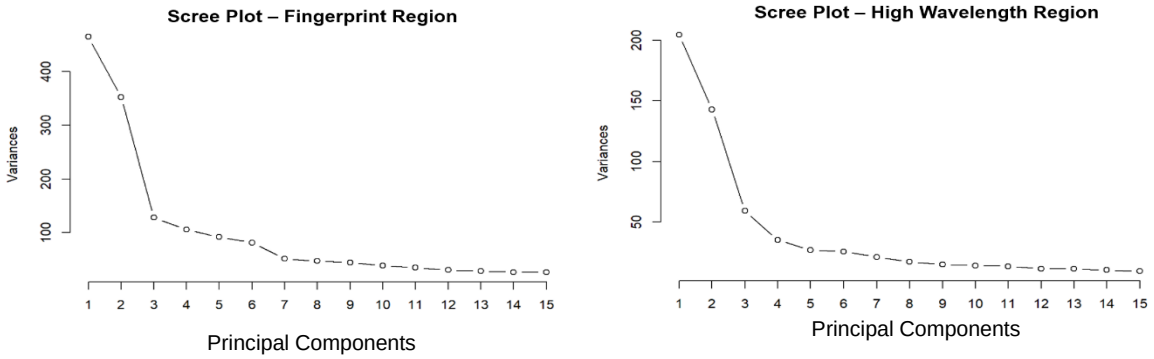
Table 3: Average wavenumber values gathered from peak fitting and their respective errors, along with the separation between the error bars shown in figures 23a and 23b. Negative values, highlighted in red, indicate an overlap in error bars and values highlighted in green were those deemed to be statistically significant.

The scatter plots in both parts of figure 23 were produced in order to visualise significant intensity differences between the two samples in a way that includes the error in the values and that is able to be easily quantified. This was done by calculating the vertical distance between error bars for the same wavenumber to determine which points produced statistically significant differences. These values are presented in table 3.

In total 15 bands were deemed significant by this method, with a wider range of assignments still including collagen constituents, but also DNA and lipids [24][26]. These bands are highlighted in green and have an intensity difference of the order of magnitude E-04 or higher

when error is taken into account. Table 3 also shows 4 wavenumber bands where the error bars overlap, highlighted in red and denoted by a negative separation value, meaning there is no statistical difference between the modern and ancient samples and are therefore not significant to the analysis. However, of these 4, 2 (1033.80cm^{-1} and 1244.70cm^{-1}) were deemed as significant by looking at the average subtraction. An additional peak seen as having a significant difference between the two samples according to the average subtraction in figure 21 at 1672.00cm^{-1} was not significant according to table 3; similarly 6 peaks decided to be significant using table 3 were not significant according to the average subtraction plot. This means that of the 18 peaks that could be significant according to the analysis so far, the two methods only agreed on 9 of them. More detailed examination was therefore needed to fully determine which areas of the spectra showed a large enough difference to be statistically significant.

4.4. Principal Component Analysis (PCA)



a)A scree plot produced for the fingerprint region.

b) A scree plot produced for the high wavenumber region.

Figure 24: Scree plots made in R of the data input for principal component analysis. Includes plots for both regions of data. The plots graphically represent the level of variance in each PC.

PCA was completed separately for the fingerprint and high wavenumber regions. Scree plots were made and used to determine how many principal components needed to be analysed to ensure enough variance is captured within them. In figure 24 parts a and b it is shown that the

variance captured in each PC after the sixth, indicating a smaller proportion of variance is represented per PC after this point. In the fingerprint region this was a drop in variance of 13% from 35% to 22%, and in the high wavenumber region it dropped from 29% to 24%, a drop of 5%. In both areas this is the largest drop in percentage variance outside of the first 5 PCs. From a total of 200 PCs, the first 6 PCs captured 52% of variance in the fingerprint region, and 55% of variance in the high wavenumber region.

PC scatter plots were produced for all combinations of PCs 1-6, and the plots which showed the most separation between PCs are included in figures 25 and 26 for the fingerprint region, and 28 and 29 for the high wavenumber region. Their accompanying loadings are included in figure 27 parts a-c and figure 30 parts a-c for the two regions, respectively.

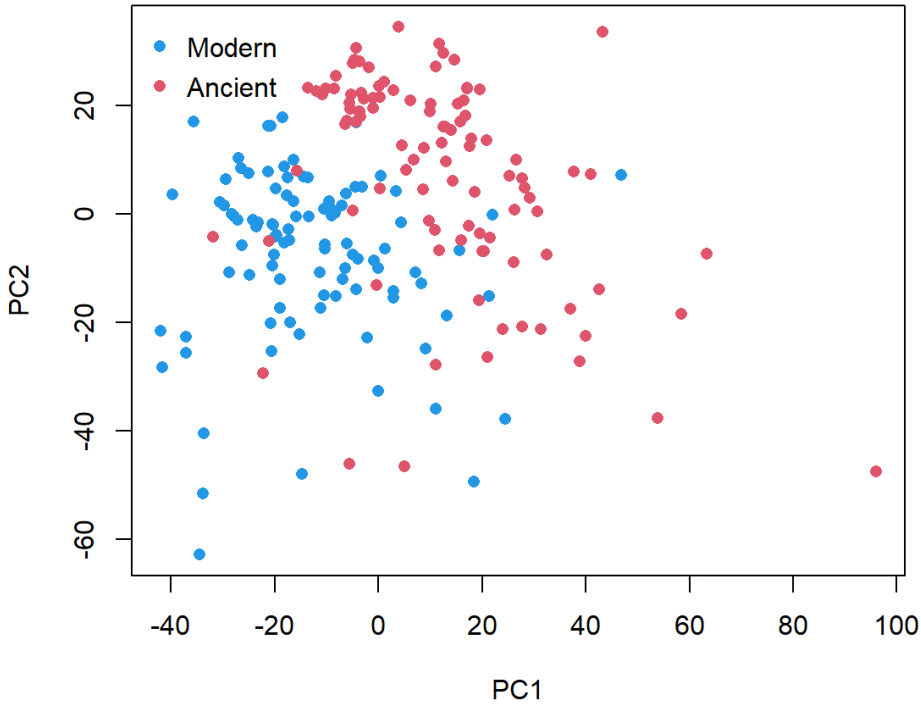


Figure 25: A scatter plot of the scores representing the relationship between PC1 and PC2 for the fingerprint region.

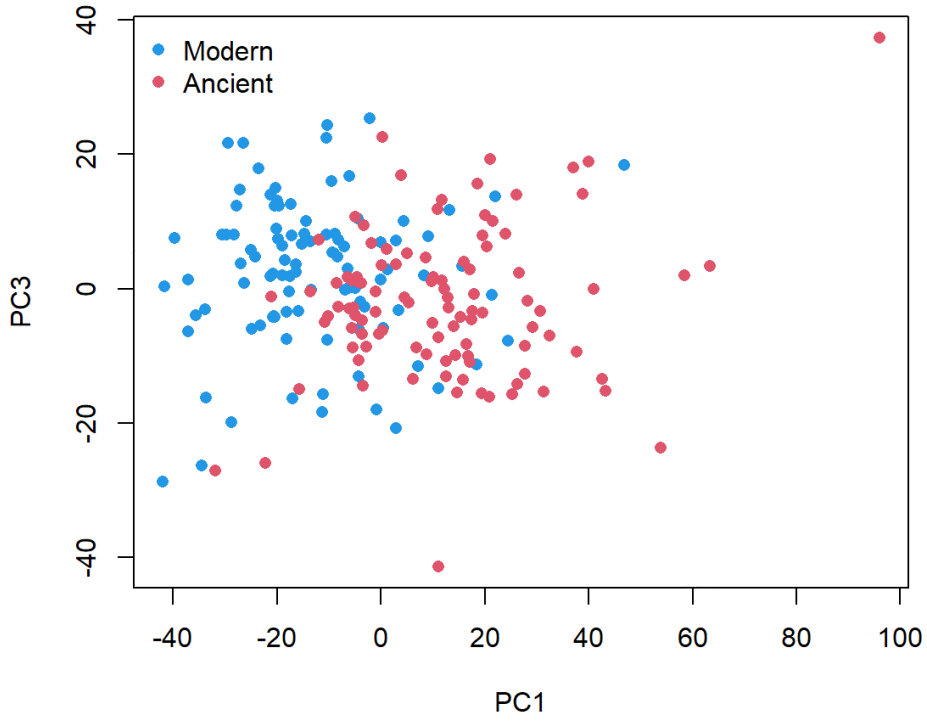
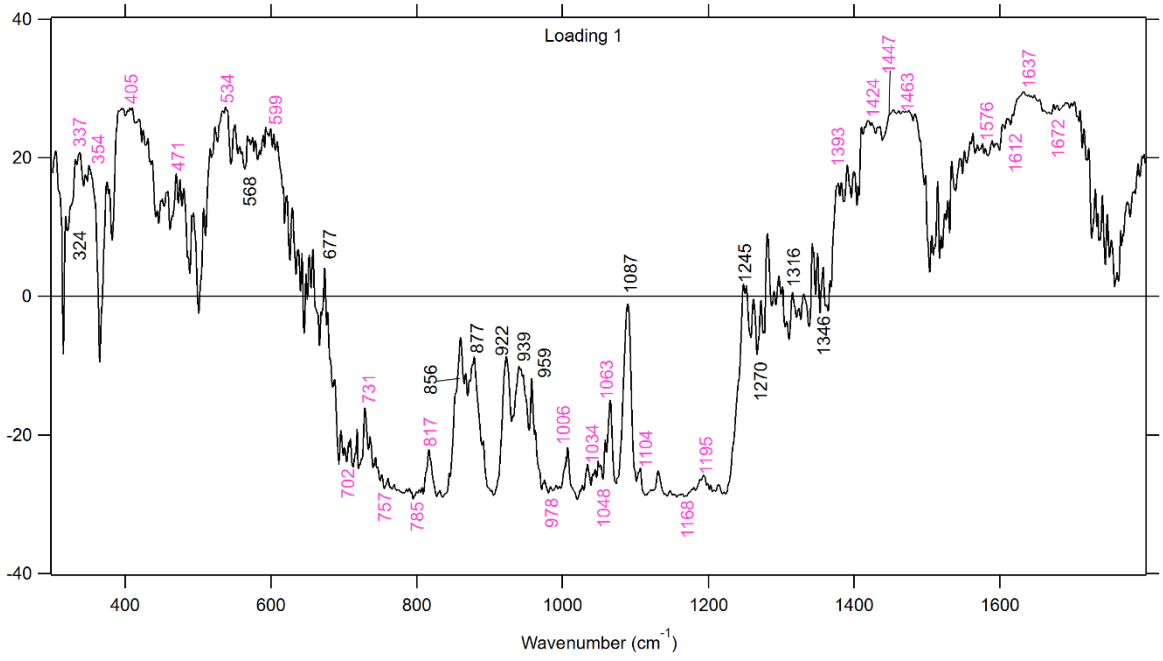
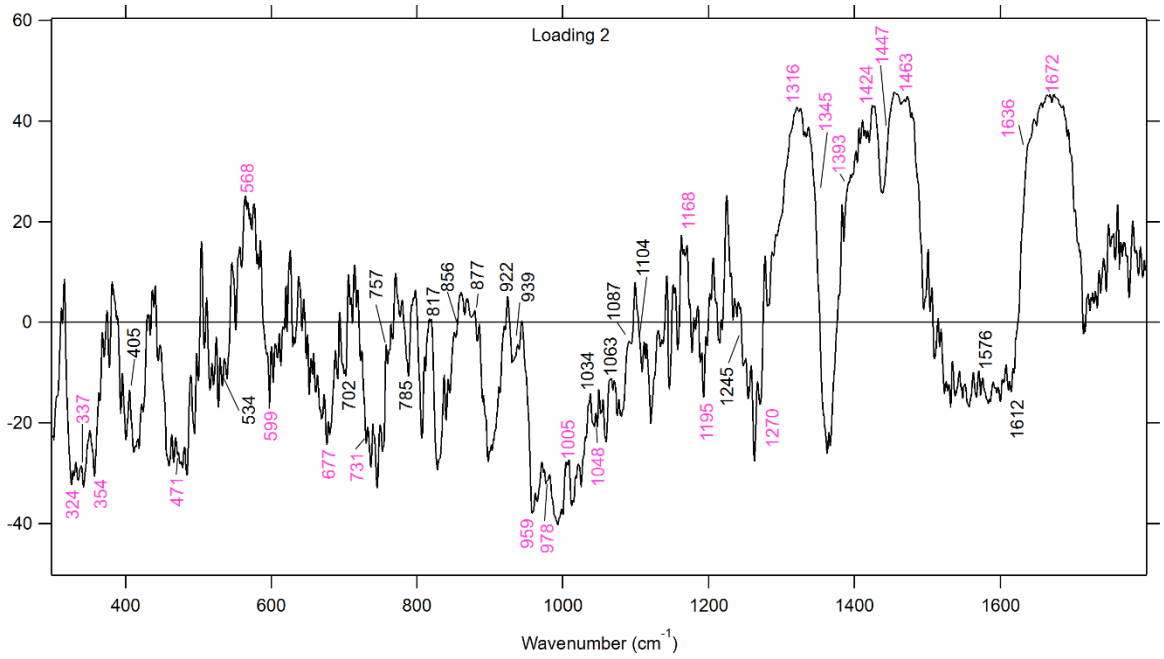


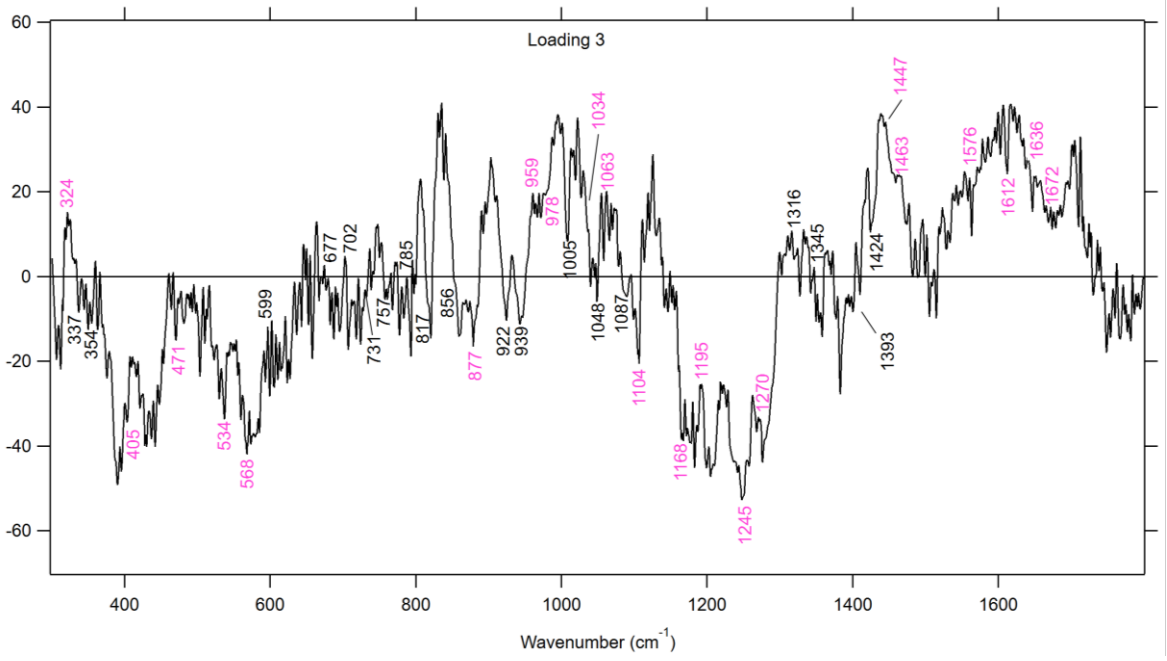
Figure 26: A scatter plot of the scores representing the relationship between PC1 and PC3 for the fingerprint region.



a) Loading 1 for the fingerprint region.



b) Loading 2 for the fingerprint region.



c) Loading 3 for the fingerprint region.

Figure 27: The loadings for PC's 1, 2 and 3, in the fingerprint region decided to be significant by assessing the separation in the score plots. All deconvolved bands are labelled with significant points indicated in pink.

Figure 27 shows the 3 most significant loadings, proved by the separation shown on the score plots in figures 25 and 26. The peak fitted wavenumber values obtained in table 2 are labelled on each figure, with the significant values indicated by pink text. Loadings highlight the contributions each variable makes to the principal component, meaning a higher magnitude signifies a more important variable. Positive loading values mean the presence of the marker is significant whereas a negative value means the markers absence is noteworthy. The same analysis was done in the high wavenumber region.

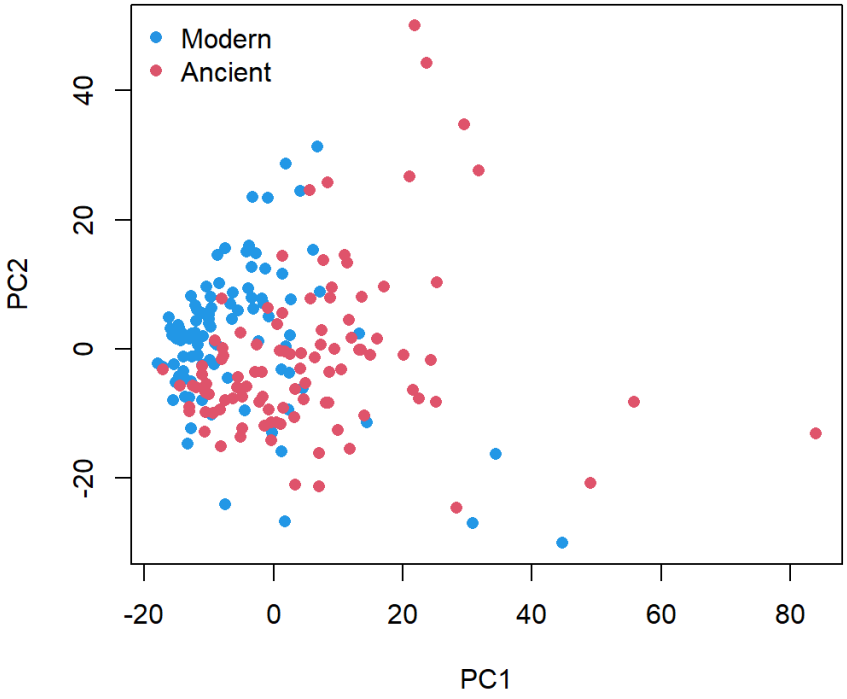


Figure 28: A scatter plot of the scores representing the relationship between PC1 and PC2 for the high wavenumber region.

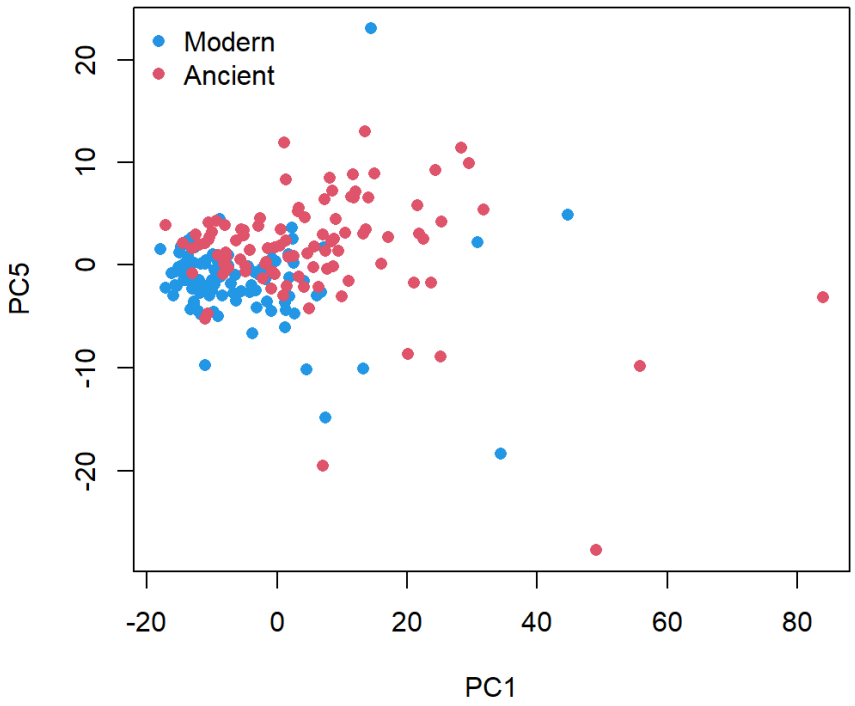
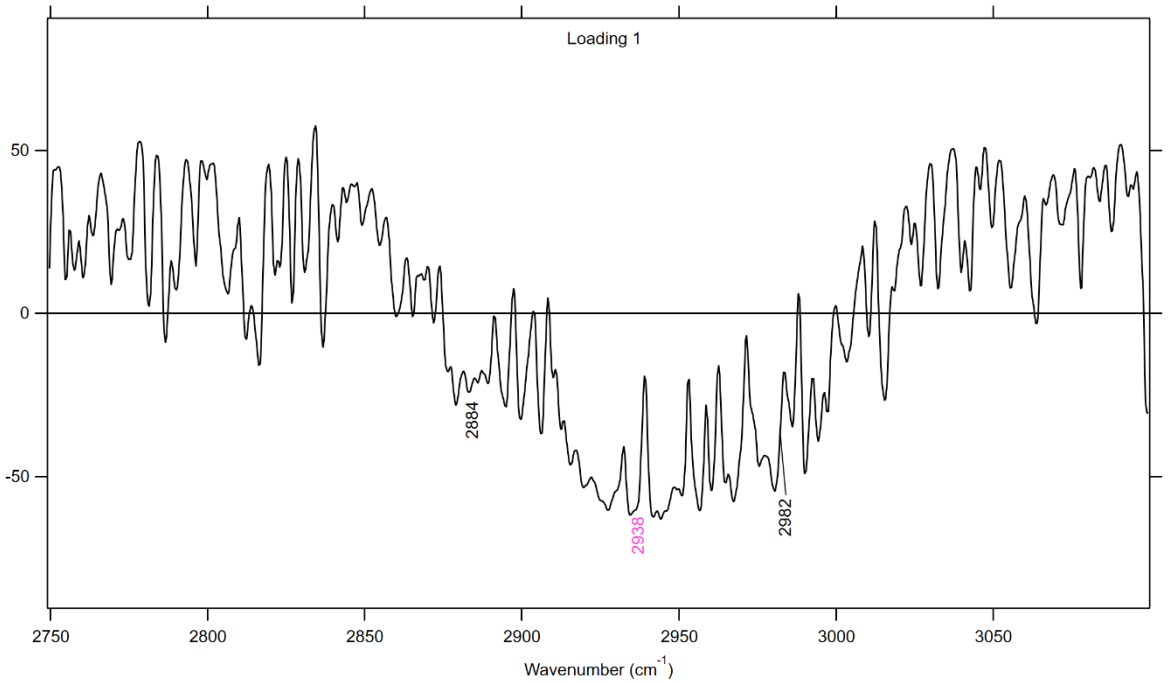
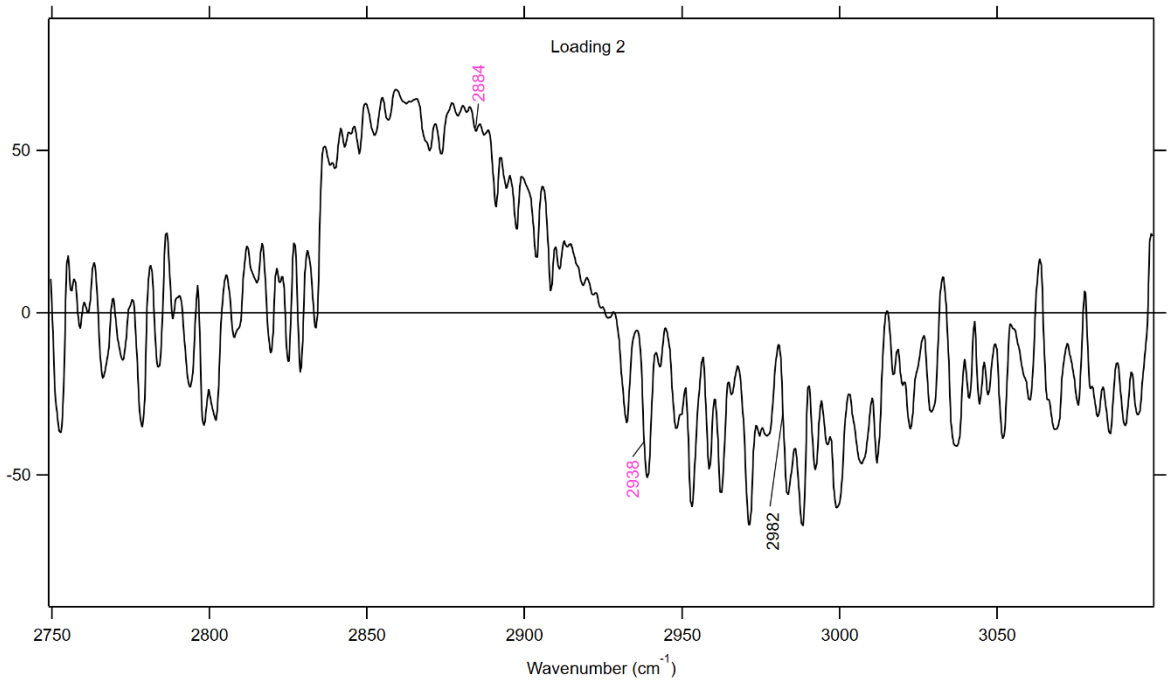


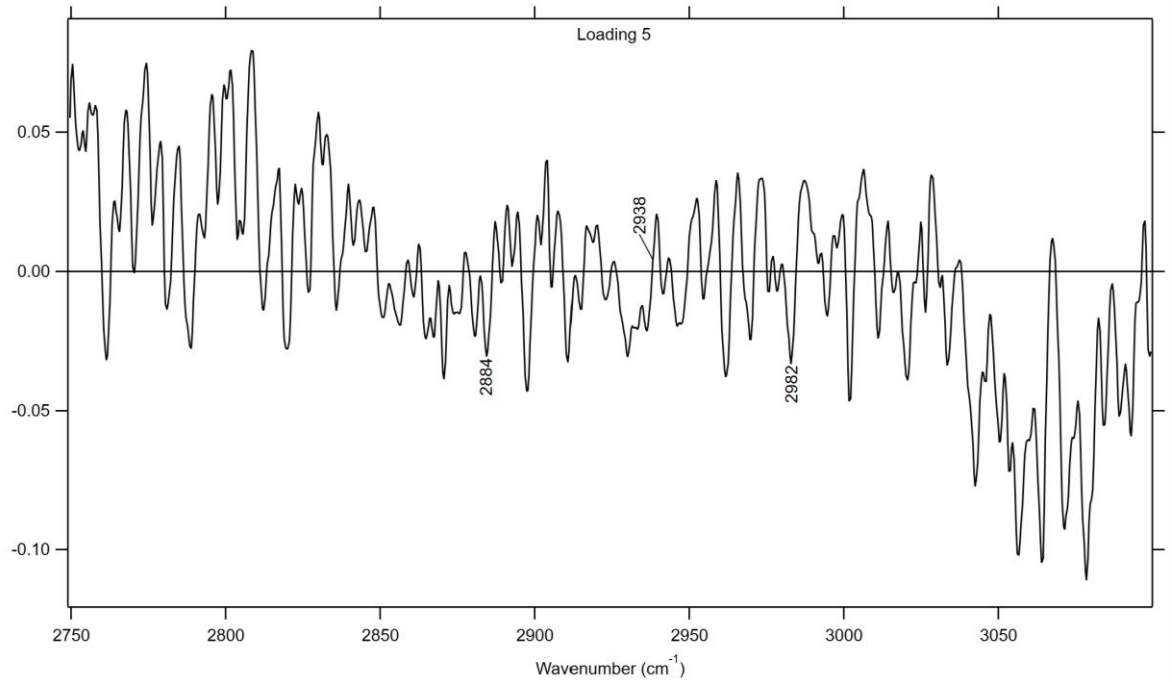
Figure 29: A scatter plot of the scores representing the relationship between PC1 and PC5 for the high wavenumber region.



a) Loading 1 for the high wavenumber region.



b) Loading 2 for the high wavenumber region.



c) Loading 5 for the high wavenumber region.

Figure 30: The loadings for PC's 1, 2 and 5, in the high wavenumber region decided to be significant by assessing the separation in the score plots. All deconvolved bands are labelled with significant points indicated in pink.

With the three analysis methods presented in figures 21 and 22, 23, 27 and 30, an assessment can be made on the significance of each biomarker in the results obtained. Table 4 below shows each biomarker and in which areas they were decided to be significant.

Wavenumber (cm^{-1})	Average Subtraction	Error Bar Separation	PC1	PC2	PC3	PC4	PC5	PC6	Overall
323.70				✓	✓			✓	
337.00			✓	✓				✓	
354.00			✓	✓				✓	
404.70			✓		✓		✓	✓	
471.00			✓	✓	✓		✓	✓	
533.87			✓		✓			✓	
567.60				✓	✓			✓	
599.00			✓	✓				✓	
676.50				✓				✓	
702.00			✓					✓	
730.60			✓	✓		✓		✓	
757.00	✓	✓	✓			✓	✓		✓

785.10			✓				✓		
817.15		✓	✓			✓			~
856.00	✓	✓				✓			✓
877.00					✓			✓	
921.60								✓	
939.20	✓	✓				✓	✓		✓
958.80		✓		✓	✓				~
977.70	✓	✓	✓	✓	✓	✓	✓		✓
1005.80	✓	✓	✓	✓			✓		✓
1033.85	✓		✓		✓				~
1047.80			✓	✓	✓	✓	✓	✓	~
1063.90			✓			✓	✓	✓	
1087.00	✓	✓				✓	✓	✓	✓
1104.10			✓		✓	✓	✓	✓	~
1168.00			✓	✓	✓	✓			
1195.00			✓	✓	✓			✓	
1244.70	✓				✓	✓	✓		~
1270.00	✓	✓		✓	✓	✓	✓	✓	✓
1316.40				✓		✓			
1345.50				✓			✓	✓	
1393.00			✓	✓		✓	✓	✓	~
1424.20			✓	✓			✓	✓	
1447.30		✓	✓	✓	✓		✓	✓	✓
1463.00	✓	✓	✓	✓	✓	✓	✓	✓	✓
1576.00			✓		✓	✓	✓		
1612.00		✓	✓		✓	✓	✓		
1636.60	✓	✓	✓	✓	✓	✓		✓	✓
1672.00	✓		✓	✓	✓	✓		✓	~
2884.00		✓		✓					
2938.50		✓	✓	✓	✓	✓		✓	✓
2982.20								✓	

Table 4: Indicates the significance of each biomarker in terms of evidence of deterioration as seen in the average subtraction (figures 21 and 22), scatter plot error bar separation (table 3), and PC loadings (figures 27 and 30). Ticks indicate significant biomarkers, and tildes represent biomarkers that may be significant, but need further work.

A more detailed investigation into the heterogeneity differences between the two samples has been completed in table 5. The biomarkers used for this analysis are the ones marked as

overall significantly different between the two samples in table 4, which was decided by considering its significance shown across the univariate and multivariate analysis performed.

Wavenumber (cm^{-1})	Intensity Difference between Samples (a.u)	Assignment	Bond Interaction
757	-2.6E-04±3E-05	Tryptophan [24][56] Ethanolamine, [24] Cytoseine [24]	Ring Breathing
856	-3.4E-04±5E-05	Tyrosine, [24][26][56] Collagen, [[24][43] Polysaccharides, [24][26] Proline, Hydroxyproline [24] Glycogen, [24] Phosphates, [24]	Ring Breathing C-O-C Skeletal Mode C-C Stretching Olefinic CHH Aromatic Bending
939	-2.3E-04±7E-05	Amino Acids, [26] Proline [24] Hydroxyproline [24] Collagen, [24][43] Protein [24] Carbohydrates [24]	C-C Stretch (11) PPII Helix C-O-C
978	-2.03E-04±6E-06	Phenylalanine [24] Protein [24]	CH wagging in CH=CH C-C Stretching Disordered Structure
1006	-2.6E-04±1E-05	Phenylalanine [24][56][38][50][3][43] Tryptophan, [24] Sulphate [3] Lime/Gypsum [38] Polysaccharides [24]	Ring Breathing C-C Stretching C-C Vibration
1087	4.8E-04±9E-05	Lime Treatment, [24][47] Collagen, [43] Phosphates (Nucleic acids) [24]	C-N Stretch C-C Stretch
1270	-3.0E-04±4E-05	Amide III, [3] [24][26] Collagen, [24][26] Lipids or Phospholipids, [24][26] Cytoseine, [24] Proline, [50]	C-H Rocking C-N Stretching
1447	-6E-04±1E-04	Lipids, Phospholipids [24] [26][50] Collagen, [24] Proline [43]	CH2 Bending CH2 Scissoring CH3 Bending C-H Deformation C-H Vibration
1463	3.5E-04±6E-05	Lipid, [24] Collagen, [24] Alanine [43] Deoxyribose [24] Disaccharides [24]	CH2 bending CH2 Wagging CH2 and CH3 deformation Fermi Interaction

1637	3.2E-04±2E-05	Amide I, [24][38][50] Collagen, [24]	C=C Stretching PPII Helix Disordered Structure C=O Stretching
2938	-1.5E-03±1E-04	Lipids, [24]	CH-CH3 Stretch CH2 Stretch CH Vibration

Table 5: Detailed band assignments along with intensity changes for the 11 bands deemed significant in table 4.

Table 5 firstly shows the intensity differences between the modern and ancient sample at the points shown to be statistically significant. These differences were calculated relative to the modern control sample, so a negative number is indicative of a lower intensity for the ancient sample. The errors given for these values are the combined standard errors for the intensity measurements on both samples. The error in the wavenumbers in all cases in table 5 was considered to be $\pm 3 cm^{-1}$. This is in agreement with the largest error calculated for an average wavenumber value in table 2, and also the instrumental error of the XploRA equipment [54].

Also given in table 5 are more detailed peak assignments from existing literature along with the specific bond interactions responsible for the presence of the band. Some peaks are very well documented and have multiple sources with the same information, whereas others do not have much existing supporting literature to aid in drawing a conclusion about their cause. An example of the former is the peak at $1006 cm^{-1}$, with 12 references within range of error in the database assembled by Talari et. al. alone [24]. There is an additional peak attributed to Phenylalanine in table 5, at 978 cm^{-1}. Studies have shown that Phenylalanine plays a part in parchment deterioration via interactions with light [6], and that these interactions change the relative intensity of the bands exhibited [56]. Edwards et al. concluded that a decrease in intensity of the Phenylalanine peak is a sign of deterioration of the parchment [3], which combined with the other works mentioned can be deduced to be via photo-oxidation.

The band at $978 cm^{-1}$ can also be assigned to C-C stretching in a disordered protein structure [24]. The presence of this is indicative of breakdown of the typical PPII helix due to hydrolysis or cleavage of the structure defining hydrogen bonds, which can happen by oxidation or thermal denaturation [13].

The product of the oxidation of Phenylalanine is Tyrosine, which is also represented in table 5 by the 856cm^{-1} band [56].The two amino acids both have an aromatic ring in their molecular structure, and previous studies have shown this to be a target location of oxidation within collagen molecules [13]. This aligns with the ring breathing bond assignment present for both the aforementioned bands [24]. This form of deterioration involves covalent bond cleavage in the amino acid chains that form collagen fibres, causing the molecule to break down. The 856cm^{-1} band can also be attributed to collagen, due to the ring structure in proline undergoing denaturation in a similar way [13].

The band at 757cm^{-1} has been attributed to tryptophan, and whilst it is an amino acid it is not one found in collagen. It is however found in other natural products such as eggs and milk [41], components of which have been known to be used as treatments to parchments. In milk, tryptophan is found specifically within Casein, the major protein component [56]. Edwards et al. reported a treatment including egg yolk for preservation [3], Bicchieri et al [14] described a smoothing paste including egg white and milk, and Kautek et al. [15] states that egg whites have been used to rehydrate parchments after undergoing the drying process. Tryptophan has been shown to be sensitive to degradation caused by many methods, including oxidation and heating [45], which is likely to be the cause of the decrease in band intensity shown in table 5. It is, however, worth noting that according to Nevin et al. [56], the changes induced by light exposure to egg whites and yolks where Tryptophan is found are sensitive enough to be induced by Raman spectroscopy, so it cannot be considered a perfect method for analysis of these particular parchment treatments.

The band at 939cm^{-1} falls in the expected range for the C-C stretching mode of collagens protein backbone, particularly the stretching of the ring structures of proline and hydroxyproline [13]

The degradation of hydroxyproline affects the parchments ability to withstand heat, as the molecules thermal stability depends on its quantity [7]. It stabilises the structure of the collagen fibres by forming hydrogen bonds between the ends of the chain and water molecules to form bridges. The decrease in intensity seen at this band implies damage to the hydroxyproline, and therefore potential heat denaturation of the collagen throughout the parchment. Work by Cappa et al. confirms that this band is sensitive to heat degradation [13]

The band at 1270cm^{-1} is characteristic of half of a doublet band for Amide III. The other portion of the band presents at 1240cm^{-1} in literature [17], which was found at average wavenumber 1245cm^{-1} in table 2. The intensity values in the same table for this value show no change at the precision shown, in disagreement with the literature. The 1270cm^{-1} section of the doublet is assigned to C-N stretching within proteins with a PPII helix structure [46], and a decrease in intensity of the band shows the breakdown of this structure, often due to hydrogen bond cleavage caused by thermal denaturation [13]. The nature of the degradation seen here is similar to that of the 939cm^{-1} peak, supporting the suspicion of widespread thermal damage to the sample.

There is extensive literature supporting the assignment of the band at 1447cm^{-1} to the bending of the C-H covalent bonds, specifically in a CH_2 formation. Cappa et al notes that a decrease in intensity is observed for this band when exposed to increasing temperatures, supporting the idea that the ancient sample has undergone thermal denaturing alongside the changes in the bands previously discussed [13]. The database assembled in work by Talari et al. points to these CH2 groups being part of a lipid [24]. Lipids have been shown to be resistant to many forms of degradation, so it follows that this band is attributed to the bending of bonds, and not the breakage of them.

Talari et. al. also attributed the band with a Raman shift of 2938cm^{-1} to lipids, and stretching of CH2 covalent bonds [24]. There has been minimal published literature examining the high wavenumber region of the parchment Raman profile, and therefore a concrete conclusion about the nature of the physical changes undergone by the parchment leading to the intensity decrease of this band is hard to draw. Given its similarities in chemical and mechanical assignments to the band at 1447cm^{-1}, an assumption can be made that this decrease in intensity is also a result of thermal denaturation of the parchment, however a more thorough analysis of this region would be necessary to determine for certain if this is the case.

The previous bands all follow the expected result was that intensity values would decrease for the ancient sample as the parchment and its components degraded over time. However, there are 3 bands in table 5 for which this is not the case.

The band at 1087cm^{-1} was deduced to be a result of lime treatment in the parchment preparation process. One reason for the increase in intensity of the band could simply be that the parchments are from different eras. As explored in section 1.1, parchment preparation

methods vary largely by time period, location and culture. The modern and ancient parchment were created multiple centuries apart, so it is highly likely that different methods were used, resulting in different quantities of leftover lime on the surface of the manuscripts, and therefore different Raman profiles. This peak is specifically caused by calcite ($CaCO_3$), but can also be caused by the reaction of calcium oxide (CaO) with atmospheric carbon dioxide (CO_2) [47]. Since the ancient parchment has been around for so much longer, it would have been exposed to more opportunity for this reaction to take place, leading to a more intense Raman peak.

The band showing an increase in intensity at 1463cm^{-1} has been attributed to lipids. Lipid molecules are resilient, and often show little to no change between samples of different ages such as the ones used in this study [48] This means the lipids detected in Raman spectra could be from the original skin, having survived both processing and the effects of degradation over time. The increase in lipid content observed here could be explained in two ways. Firstly, different areas of skin on the body are known to have different lipid contents [40]. While it is unknown which areas each sample was taken from, the samples are known to come from different animal species. As stated in section 2.2, the modern sample was made from goat skin, whereas the ancient sample was made from sheep skin. It is very likely that the intensity difference is a result of the samples not being from the same animal, as skins of different species are known to have different lipid profiles [39]. Another potential reason for the increase is that over time, the handing of these documents with bare hands could introduce additional lipids to the sample [48], which in turn could increase the Raman intensity of the band, however the large differences caused by the lipid differences between species make this difficult to deduce with any certainty.

Alternatively, Reina et al. noticed an increased intensity in bands with a Raman shift of between 1429-1517cm^{-1}, caused by one of the products of degradation of UV-irradiated proteins after their structural breakdown – free carboxylate [49]. Many bands already discussed have had intensity changes as a result of photo-degradation, however it is unlikely that the parchments have been exposed to UV light in their lifetime.

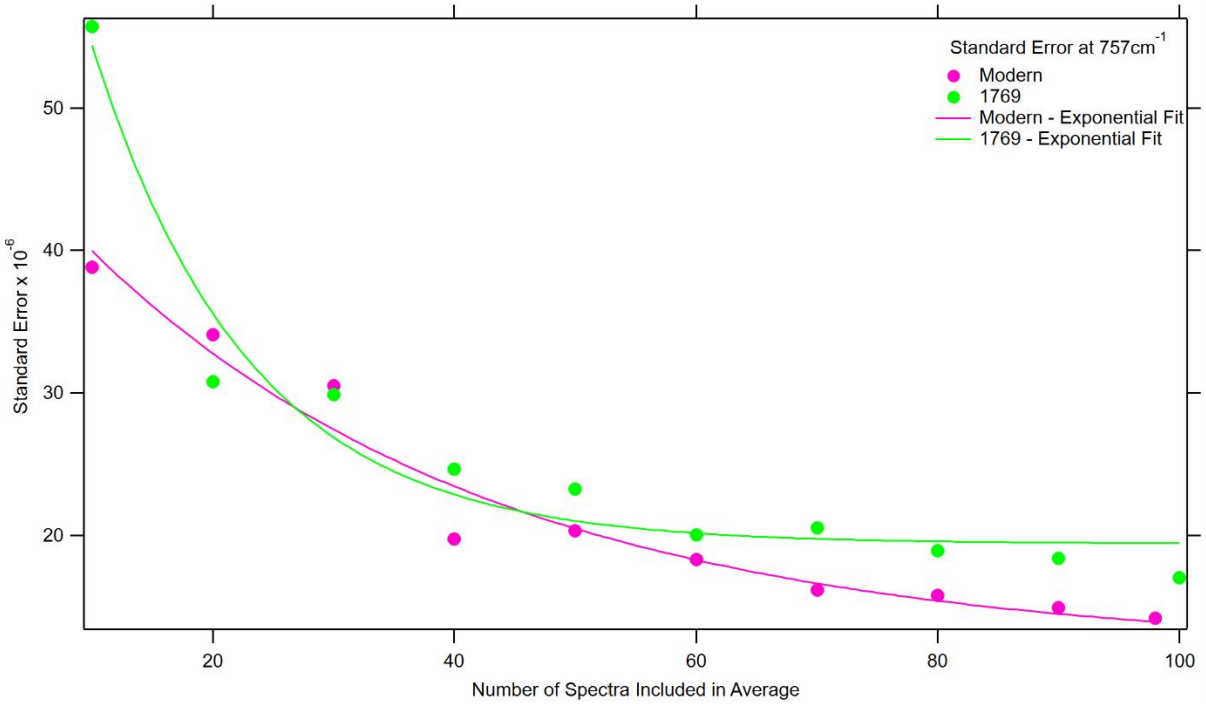
Amide I produces two bands in its Raman profile, at approximately 1650cm^{-1} and 1630cm^{-1} [17] The band deconvolved at 1637cm^{-1} present in table 5 has been attributed to Amide I by Edwards et al (04), so has been assumed to be equivalent to the 1630cm^{-1} band, which also

has slightly varying wavenumber values in the Talari et al. database [38][24]. Cappa et al. states that an increase in intensity in this region, as seen in table 5, is suggestive of the deterioration of PPII helices into disordered structure, which is supported by Pezzotti et. al. [17][50]. The band is still present and has not disappeared completely, which shows that an amount of the structure stabilising hydrogen bonds exist still, and that the PPII helix has not been completely denatured. Partial destruction of the hydrogen bonds once again points to thermal denaturation of the parchment.

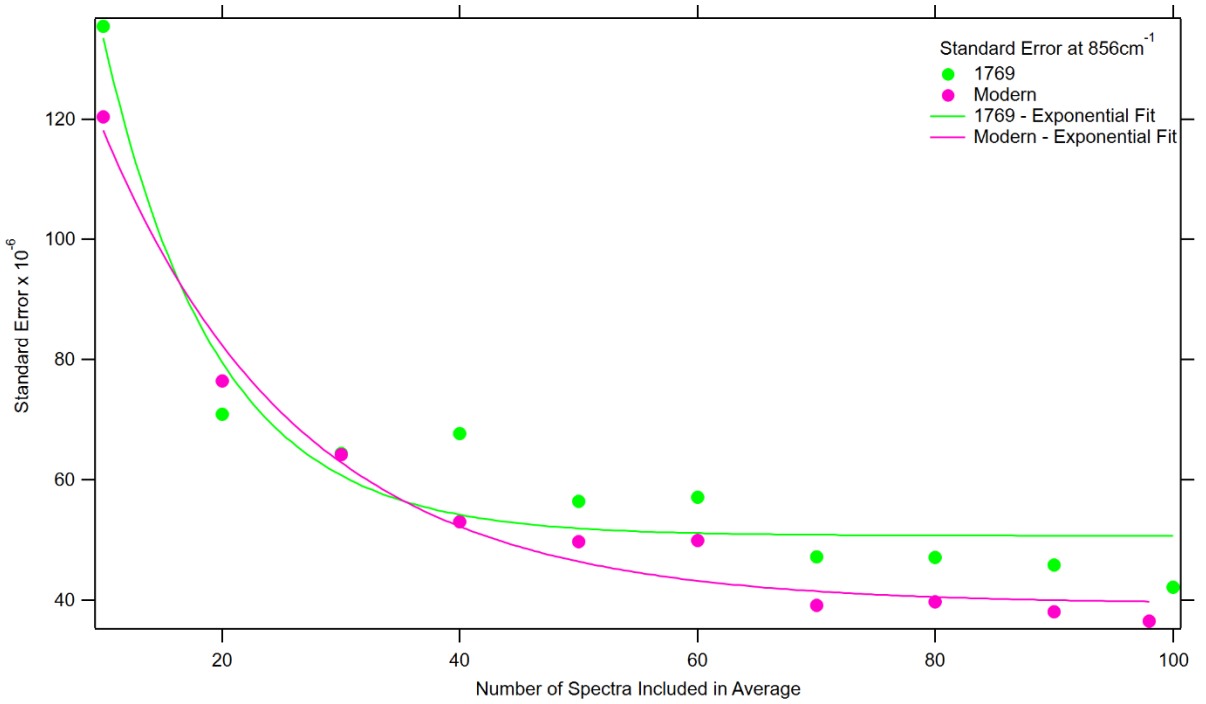
One notable peak is not present in the table, which is the Amide I peak at 1650cm^{-1}. This band has been shown as an indicator of not only if a parchment has deteriorated, but by how much [6]. A study done by Dolgin et. al. includes spectra for this peak on both sides of a modern and an aged parchment. They observed very similar Raman profiles in each case except the grain side of the ancient parchment, which had a dramatic reduction in intensity far greater than seen in table 2 [6]. The analysis in this project was carried out on the flesh side only, which could be the reason for this particular band not showing the expected reduction.

4.5. Assessing Quantified Molecular Heterogeneity Differences Between the Ancient and Modern Samples.

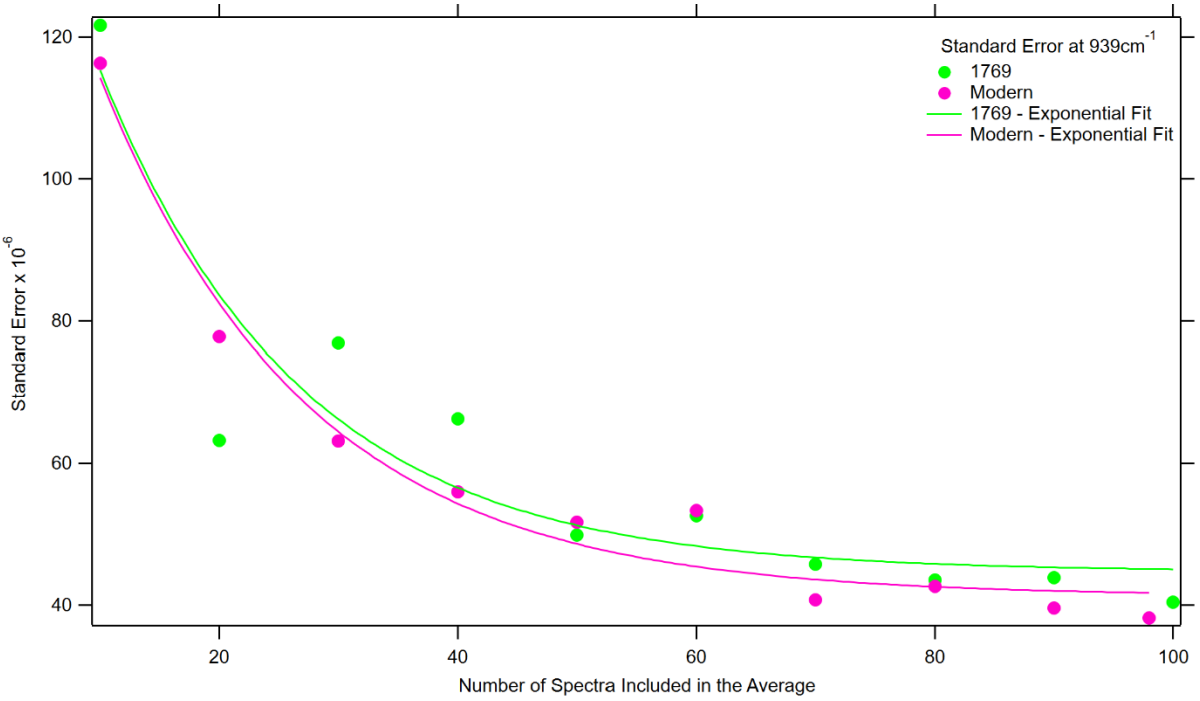
As a way to quantify heterogeneity further, standard error decay plots were produced for all the significant wavenumbers mentioned above.



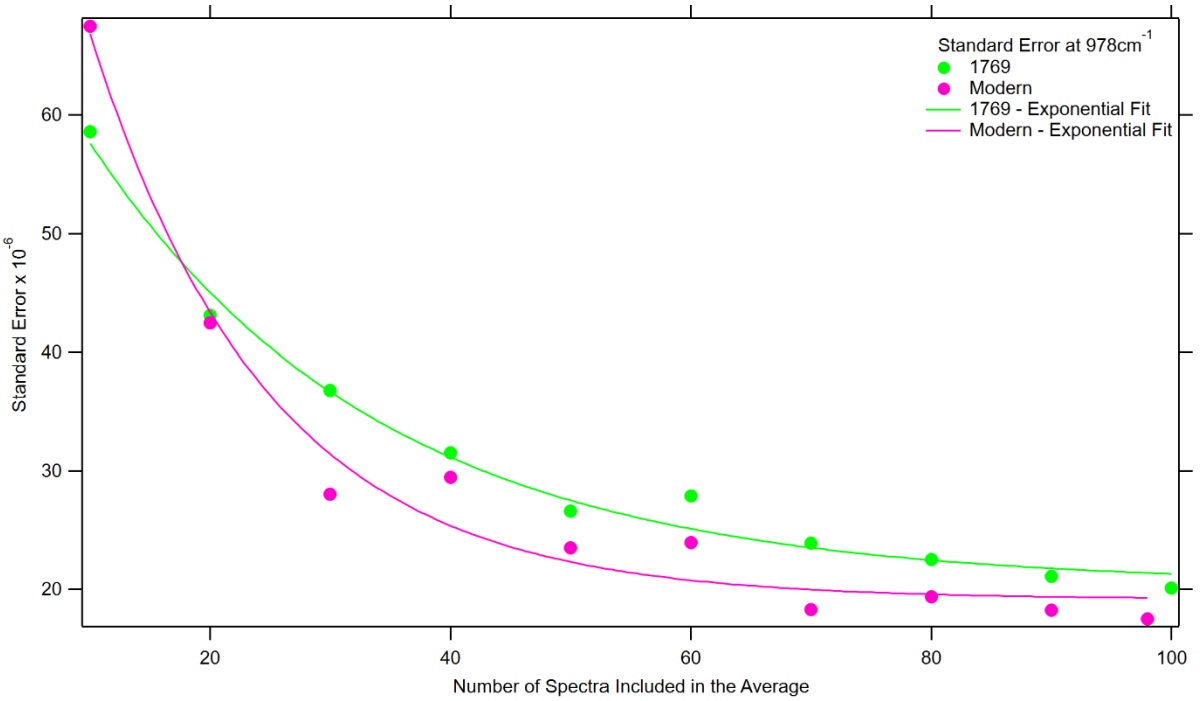
a) Standard error convergence plot for both samples at 757cm^{-1}.



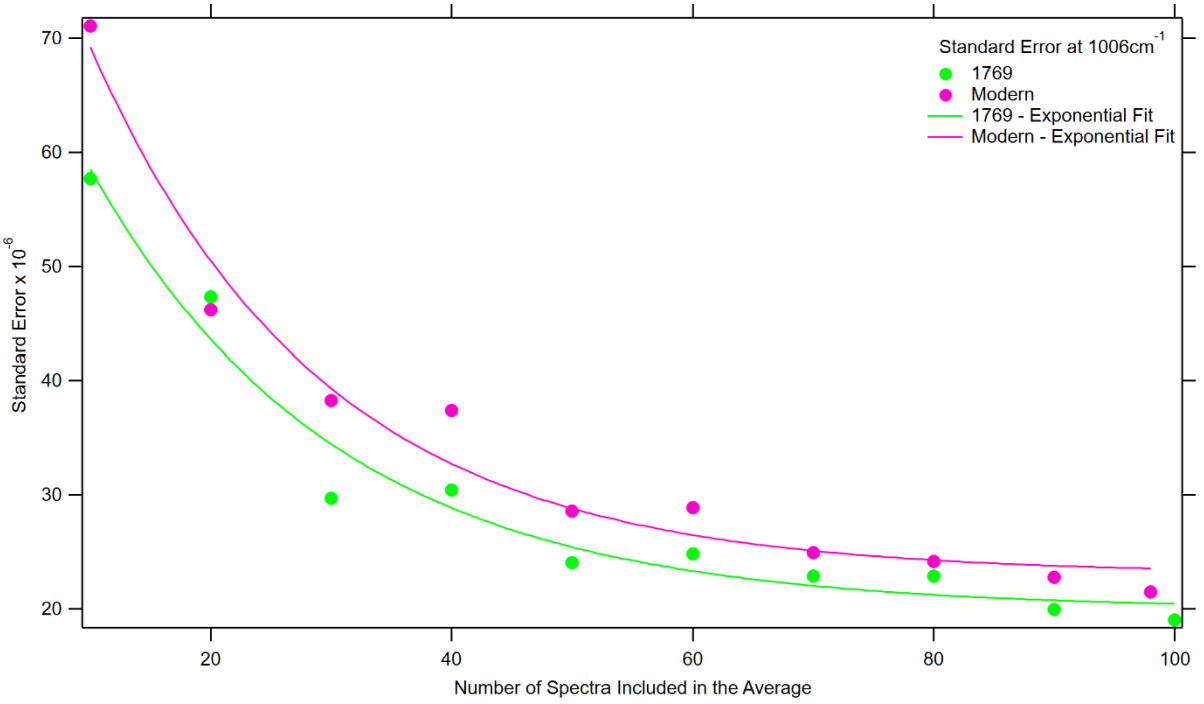
b) Standard error convergence plot for both samples at 856cm^{-1}.



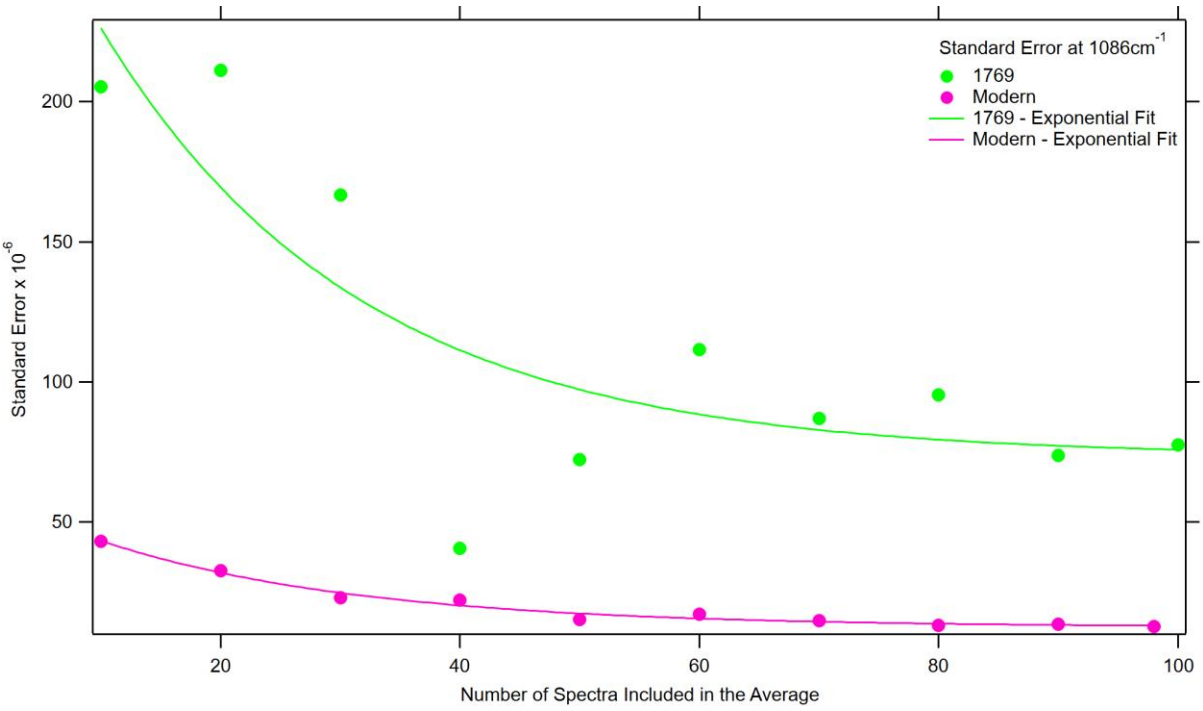
c) Standard error convergence plot for both samples at 939cm^{-1}.



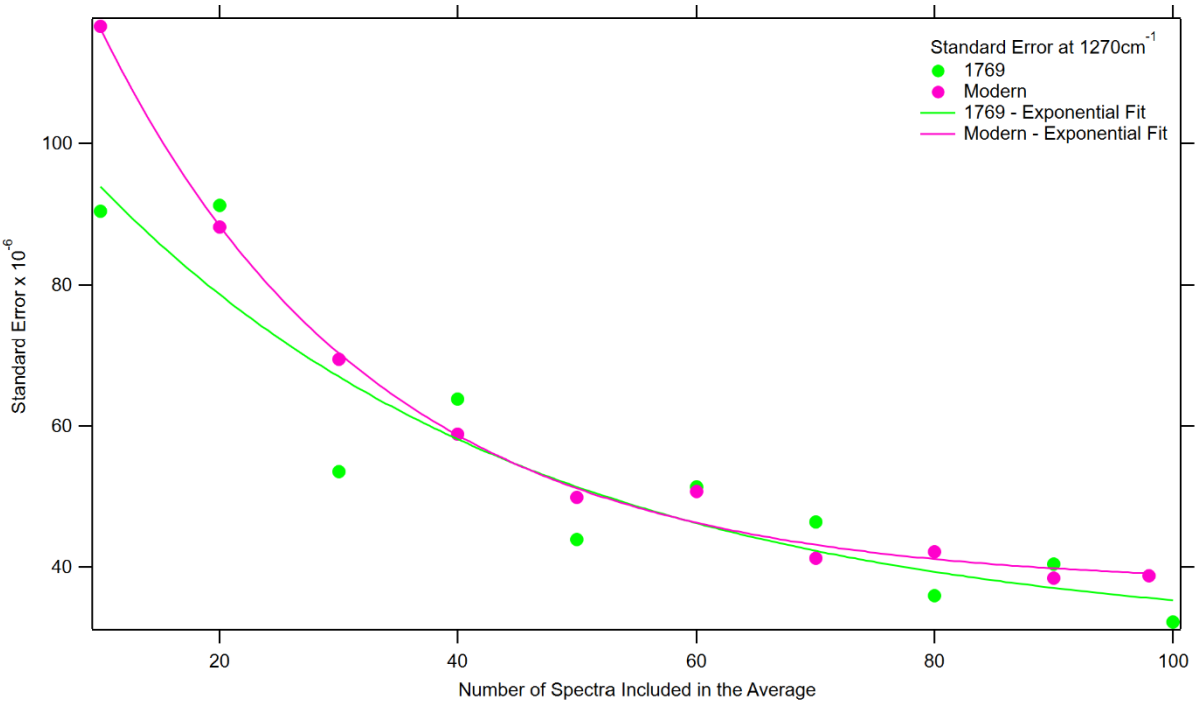
d) Standard error convergence plot for both samples at 978cm^{-1}.



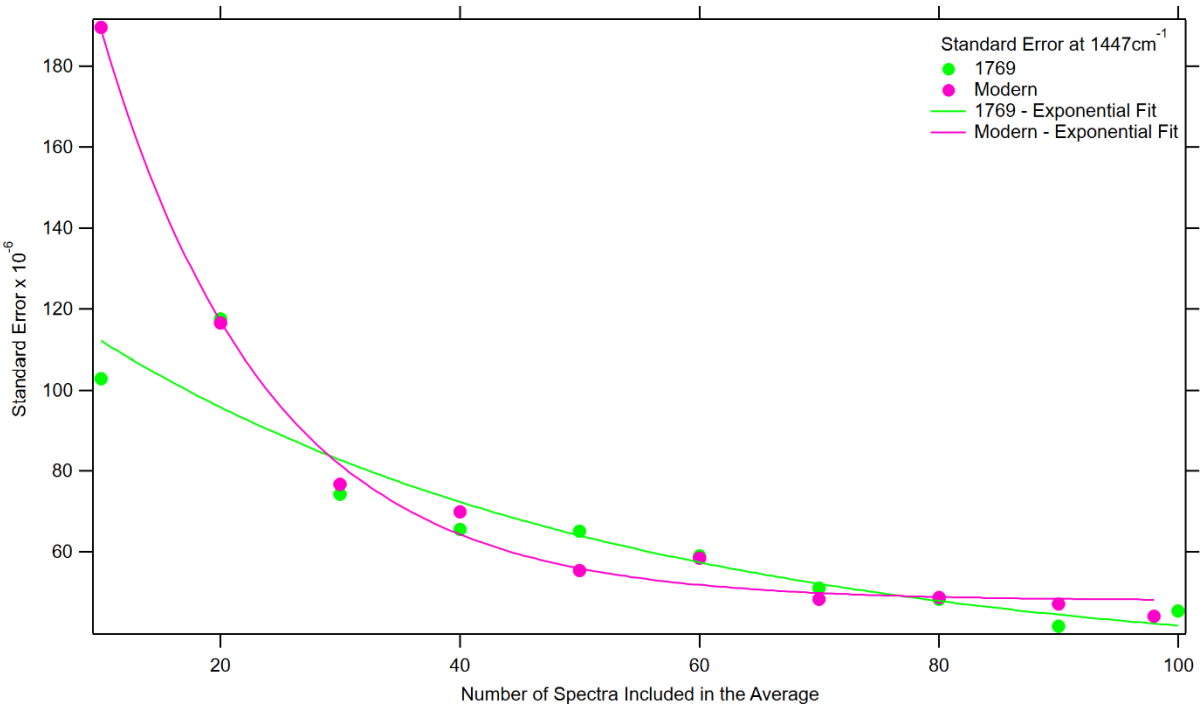
e) Standard error convergence plot for both samples at 1006cm^{-1}.



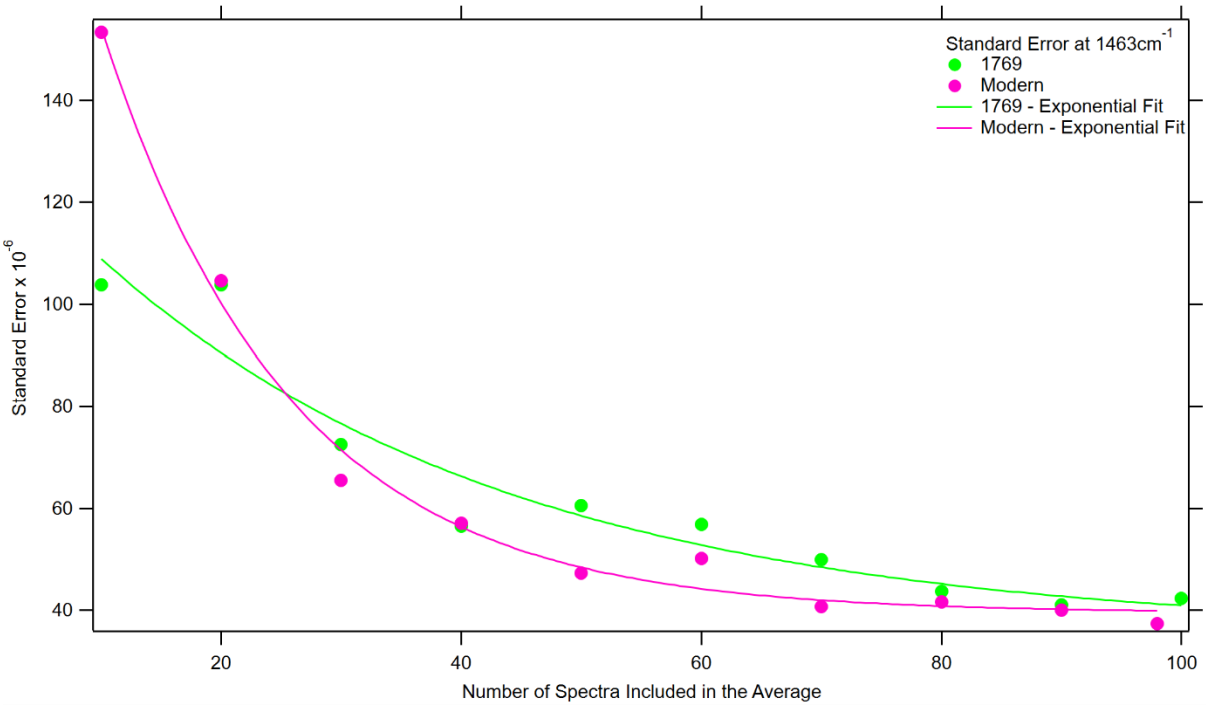
f) Standard error convergence plot for both samples at 1086cm^{-1}.



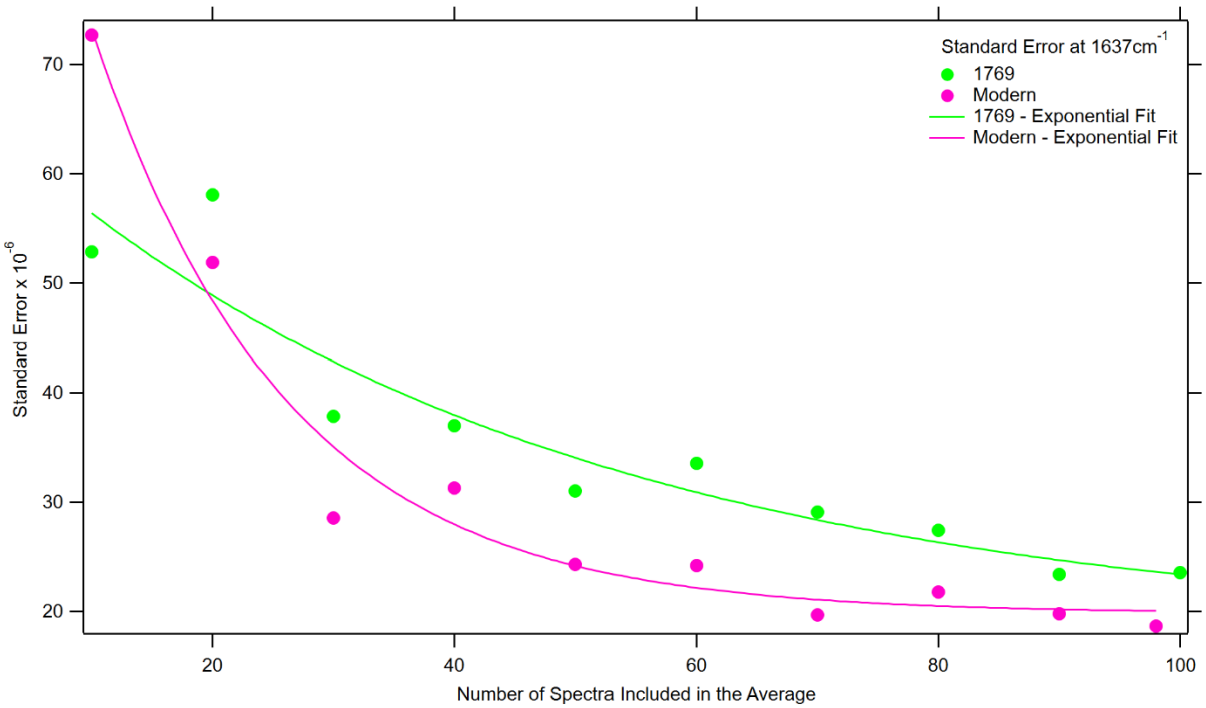
g) Standard error convergence plot for both samples at 1270cm^{-1}.



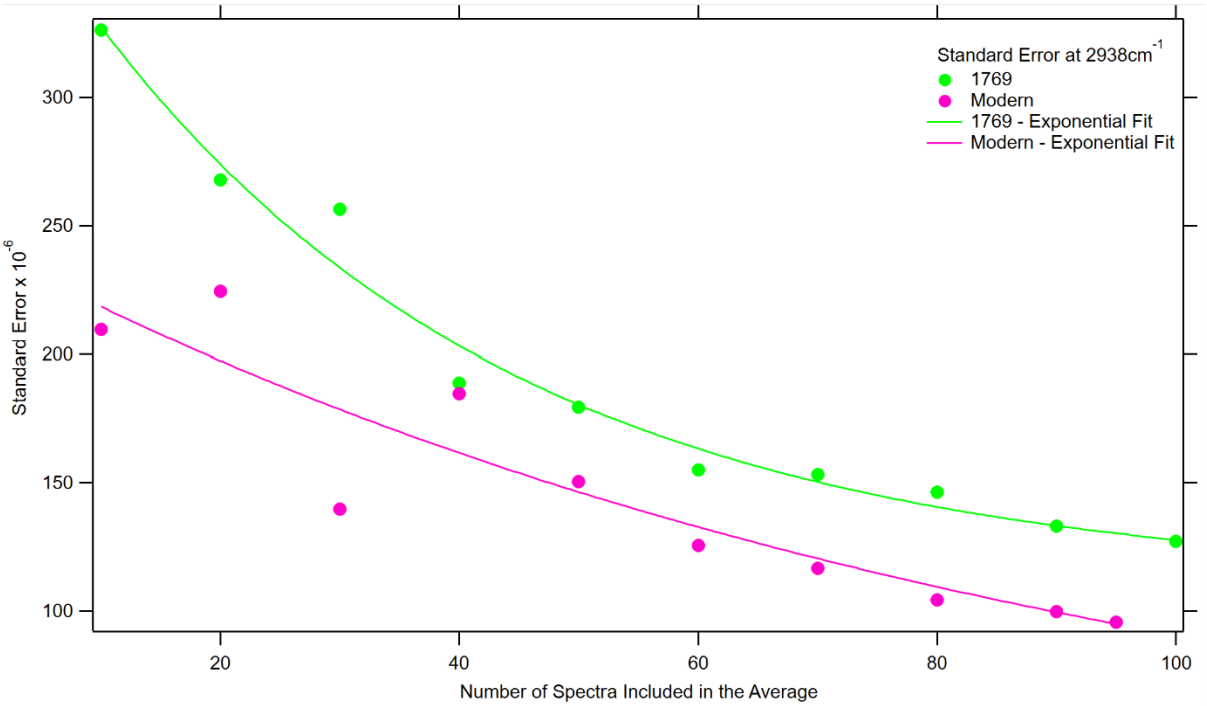
h) Standard error convergence plot for both samples at 1447cm^{-1}.



i) Standard error convergence plot for both samples at 1463cm^{-1}.



j) Standard error convergence plot for both samples at 1637cm^{-1}.



k) Standard error convergence plot for both samples at 2938cm^{-1}.

Figure 31: Standard error decay plots for all significant biomarkers defined in table 5. Includes data points and an exponential fit for each sample. Convergence was assessed I increasing intervals of 10 up to the inclusion of the full data set (100 spectra for the ancient sample, 98 for the modern sample fingerprint region and 95 for its high wavenumber region).

The plots in figure 31 parts a-k are plots of the standard error in the intensity for a spectra average for a given wavenumber against the number of spectra included in the average. One was created for each biomarker labelled as significant in table 4, and shows the error against number of spectra for both samples. The points are shown with an exponential curve fitted to them, with the goal of calculating a decay constant for each wavenumber. Similar work has been completed by Rocha et al. on stem cells, using this data to quantify heterogeneity specifically to individual biomarkers [27]. Plots for the modern and ancient samples are included on the same axes for the purpose of comparison, and the decay constants are shown in table 6.

Wavenumber (cm^{-1})	Peak Assignment	Era	Decay Constant (τ)	±Error (1sf)
757	Protein (Tryptophan – Ring Breathing)	Ancient	13	2
		Modern	34	8
856	Protein (Tyrosine, Proline – Ring Breathing)	Ancient	9	2
		Modern	17	2
939	Protein (Proline, Hydroxyproline – Ring Breathing)	Ancient	17	6
		Modern	17	3
978	Protein (disordered)	Ancient	24	3
		Modern	15	2
1006	Protein (Phenylalanine – Ring Breathing)	Ancient	20	4
		Modern	19	3
1087	Lime Phase	Ancient	21	14
		Modern	22	3
1270	Protein (Amide III – C-N Stretching)	Ancient	37	18
		Modern	23	2
1447	Lipids (C-H Bending)	Ancient	45	25
		Modern	14	1
1463	Lipids (CH_2 bending)	Ancient	35	12
		Modern	16	1
1637	Protein (Amide I – C=O in PPII helix)	Ancient	46	25
		Modern	16	2
2938	Lipids (CH_2 Stretching)	Ancient	35	7
		Modern	100	100

Table 6: The corresponding decay constants to all parts of figure 31 as integer values. Calculated according to section 2.3.5. Biomarker assignments are included.

Table 6 shows the decay constant (τ) found for all significant biomarkers. They were calculated by taking the inverse of the τ^{-1} value given by Igor Pro upon generating the graph, and the error in the value was calculated according to equation 18. The peak assignments given in table 5 are repeated here.

τ is used here as a measure of heterogeneity; a larger τ value means more data is needed to achieve convergence, meaning a higher degree of heterogeneity, as discussed in section 2.3.5. For 4 of the biomarkers in table 6, highlighted in green, the ancient sample has a higher τ value, indicating greater variations in the molecular profile of the parchment. This is the case for the markers at 978cm^{-1}, 1447cm^{-1}, 1463cm^{-1} and 1637cm^{-1}, and would be the expected result due to degradation and diagenesis effects the older parchment would have experienced during its extended lifetime.

The bands at 1006cm^{-1} for phenylalanine and 1270cm^{-1} for amide III also have higher τ values for the ancient sample, however the errors in the two values overlap so the difference between them cannot be considered statistically significant.

The biomarkers for which the modern sample has a higher heterogeneity, highlighted in pink in table 6, include 757cm^{-1} and 856cm^{-1}. This result is unexpected, as it has been shown that the ancient parchment has experienced a level of deterioration compared to the modern, and should therefore have a more disordered structure and higher heterogeneity. The increase seen here can be partially explained by the knowledge that each band is assigned to a type of ring breathing, involving no breakdown or cleavage of any bond or structure. This should, however, indicate no heterogeneity difference. Looking at table 4, both of these bands are seen as significant for PC1, but not PC's 2 and 3. PC1 is where most of the variance seen in the sample is contained, but also most of the spectral background. It is possible that the higher spectral background of the modern sample caused by its curvature seen in figure 6 has skewed the result of the standard error convergence to indicate a more disordered structure in the modern parchment, when this is not truly the case.

The band at 1087cm^{-1} for calcite has a marginally higher decay constant for the modern sample, however cannot be statistically separated from the value of the decay constant for the ancient sample. The final value in the table at 2938cm^{-1}, a biomarker for lipids, also shows a much higher τ value for the modern sample, however the error for the value is the same size as the value. This means there is a possibility of a τ value of 0, and as the decay constant was calculated using an exponential fit this would mean the fit would be a straight line. This shows that the exponential fit is breaking down at this point. The biomarker for proline at

939cm^{-1} had the same τ value for both the modern and ancient sample, again showing no differences in heterogeneity between the two samples at this wavenumber.

5. CONCLUSIONS AND SUGGESTIONS FOR FURTHER WORK

This thesis has shown that Raman spectroscopy is a viable tool for quantitatively assessing heterogeneity across parchment samples from different eras. It starts by discussing the possible forms of degradation a parchment could experience, and how these alter the parchment and its Raman profile. Parchments can be degraded via oxidation, where reactive free radicals cause covalent bond cleavage in the collagen structure, or thermal denaturation, in which the effects of added heat destabilise hydrogen bonds within the polypeptide chains holding collagens triple helix structure together, which can progress to irreversible gelatinisation in the presence of water.

Data was taken for 2 samples across a wider range typically seen in literature. A comprehensive comparison of the Raman shift values for to the modern sample to existing literature was completed, and found good correlations with existing knowledge about parchments and skins. Characteristic bands for the key structures of collagen were able to be deconvolved from an average spectra with a profile agreeing with what is seen in literature.

From here, a comparison was made to the ancient sample by assessing the Raman intensity of each band. This was assessed visually using a plot showing the subtraction of the modern average from the ancient average, identifying 12 previously deconvolved band that showed large differences in the spectra between the two samples above the spectral background. This was supported by another method of univariate analysis, the production of a scatter plot showing the peak intensities of each band with their standard errors represented with error bars. For 15 bands, there was a significant separation in the error bars of the two samples, pointing to changes in the parchment structure for the causes of these bands.

Multivariate analysis via PCA was also performed, to take into account the statistical robustness of this size of data set. As far as can be seen, little to no work has been carried out with PCA for the high wavelength region of a parchment Raman spectra to date. Analysing score plots and loadings and comparing the findings to the univariate analysis results led to the identification of 11 bands with statistically significant differences in intensity between the modern and ancient sample; these were 757cm^{-1}, for tryptophan ring breathing, 856cm^{-1} for tyrosine and proline ring breathing, 939cm^{-1}

indicative of proline and hydroxyproline C-C stretches in the collagen backbone, 978cm^{-1} for disordered protein structure, 1006cm^{-1} for phenylalanine ring breathing, 1087cm^{-1} for the calcite lime phase, 1270cm^{-1} for C-N stretching of the amide III group, 1447cm^{-1} for CH bending in lipids, 1463cm^{-1} for CH_2 bending in lipids, 1637cm^{-1} for C=O stretching of the amide I group in a proteins PPII helix structure and 2938cm^{-1} for CH_2 stretching in lipids. Potential degradation methods to support these changes were discussed, deciding that the ancient sample shows evidence of the damages caused by both oxidation and thermal denaturation.

Their heterogeneity was then quantitatively assessed using an exponential plot of rate of convergence as a function of number of spectra in each test group to result in a decay constant for each biomarker, a new technique to the study of parchment degradation. Higher decay constants indicating greater heterogeneity were seen for the ancient sample at the 978cm^{-1}, 1447cm^{-1}, 1463cm^{-1} and 1637cm^{-1} bands. It was expected due to the lengthy discussion of parchment degradation in literature that the ancient sample would exhibit greater heterogeneity throughout, and these bands agreed with that. The markers at 757cm^{-1} and 856cm^{-1} were shown to have higher decay constants and therefore greater heterogeneity for the modern sample. This did not agree with the literature, however both were attributed to ring breathing, so the complexity of the structure of the molecules that caused these bands could be a factor in the variance from the literature. Additionally, they were seen as significant in PC1 but not according to PC's 2 or 3, meaning the significance represented may have been due to variations in the spectral background.

The two bands with the lowest Raman shift, 324cm^{-1} and 337cm^{-1} were unable to be compared to existing literature, due to limited existing work on ranges covering Raman shifts below those represented in the typical fingerprint region. It would be useful for further work to be carried out at these wavenumbers, to determine if these bands can be attributed to either skin, or part of a parchment treatment.

It would be useful to expand on the work done to calculate decay constants of the rate of standard error convergence so a set of expected literature values are able to be established, which could potentially help with identification efforts by being able to

assign the decay constant of a particular biomarker to a time period for the parchments production.

REFERENCES

[1] Z. Minniti. Advanced characterization methods of organic materials of environmental and historical-artistic relevance through Micro-FTIR Mapping and Raman Microscopy. *University of Catania, International PhD Course in Chemical Science Thesis, XXIV Cycle*, 2011.

[2] R. Fuchs. The History and Biology of Parchment. *Karger Gazette* 67, 13-16, 2004.

[3] H.G.M. Edwards, D.W. Farwell, E.M. Newton, F. Rull Perez and S. Jorge Villar. Application of FT-Raman spectroscopy to the characterisation of parchment and vellum, I; novel information for paleographic and historiated manuscript studies. *Spectrochimica Acta Part A* 57 (2001) 1223-1234.

[4] C. J. Kennedy and T. J. Wess. The Structure of Collagen within Parchment – A Review. *Restaurator*, 61–80, 2003.

[5] M. Bicchieri, M. Monti, G. Piantanida, A. Sodo, M. T. Tanasi. Inside the Parchment. *9th International Conference on NDT of Art, Jerusalem Israel*, 25-30 May 2008.

[6] B. Dolgin, V. Bulatov, I. Schechter. Non-Destructive Assessment of Parchment Deterioration by Optical Methods. *Anal Bioanal Chem* (2007) 388:1885-1896.

[7] B. Dolgin, , V. Bulatov, I. Schechter. A complex analytical method for parchment characterization. *Reviews in Analytical Chemistry* Vol. 28, Nos 3-4, 2009.

[8] S. C. Boyatzis, G. Velivasaki and E. Malea. A study of the deterioration of aged parchment marked with laboratory iron gall inks using FTIR-ATR spectroscopy and micro hot table. *Herit Sci* (2016) 4:13.

[9] A. Bonifacio, V. Sergo. Effects of Sample Orientaion in Raman Microscopy of Collagen Fibres and their Impact on the Interpretation of the Amide III band. *Vibrational Spectroscopy* 53 (2010) 314-317

[10] R. Naomi, P. M. Ridzuan, H. Bahari. Current insights into collagen type I. *Polymers (Basel)* 2021 Aug; 13(16) 2642

[11] R. Schütz, I. Rabin, O. Hahn, P. Fratzi, A. Masic. In-situ Damage Assessment of Collagen within Ancient Manuscripts Written on Parchment: A Polarized Raman Spectroscopy Approach. *AIP Conf. Proc.* 1267, 319–320 (2010)

[12] J. Khoshnoodi, J. Cartailler, K. Alvares, A. Veis, B. G. Hudson. Molecular Recognition In The Assembly Of Collagens: Terminal Noncollagenous Domains Are Key Recognition Modules In The Formation Of Triple Helical Protomers. *Journal of Biological Chemistry* Volume 281, Issue 50.P38117-38121

[13] F. Cappa, I. Perganoni, C. Carsote, M. Schreiner, E. Badea. Studies on the effect of dry-heat ageing on parchment deterioration by vibrational spectroscopy and micro hot table method. *Polymer Degradation and Stability* 182 (2020) 109375

[14] M. Bicchieri, M. Monti, G. Piantanida, F. Pinzari, A. Sodo. NON-DESTRUCTIVE SPECTROSCOPIC CHARACTERIZATION OF PARCHMENT DOCUMENTS. *Vibrational Spectroscopy* 55 (2011) 267-272

[15] W. Kautek, S. Pentzien , P. Rudolph , J. Krüger, E.König. LASER INTERACTION WITH COATED COLLAGEN AND CELLULOSE FIBRE COMPOSITES: FUNDAMENTALS OF LASER CLEANING OF ANCIENT PARCHMENT MANUSCRIPTS AND PAPER. *Applied Surface Science*. Volumes 127-129, May 1998, Pages 746-754

[16] M. B. Haulică, O. Florescu V. Vasilache, I. Sandu. The Comparative Study of the State of Conservation of Two Medieval Documents on Parchment from Different Historical Periods. *Materials* 2020, 13, 4766

[17] F. Cappa, I. Perganoni, C. Carsote, E. Badea, M. Schreiner. Studies on the effects of mixed light-thermal ageing on parchment by vibrational spectroscopy and micro hot table method. *Herit Sci* (2020) 8:15.

[19] A. S. Lee, P. J. Mahon, D. C Creagh. Raman Analysis Of Iron Gall Inks On Parchment. *Vibrational Spectroscopy* 41 (2006) 170-175.

[20] J. R. Ferraro, K. Nakamoto. Introductory Raman Spectroscopy, 2nd Edition. *Academic Press*, San Diego USA, 2003.

[21] D. A. Long. Raman Spectroscopy. *The Characterization of Chemical Purity: Organic Compounds.* International Union of Pure and Applied Chemistry, Norwich, Great Britain, 1971.

[22] C. V. Raman. A New Radiation. *Indian J. Phys*. 2 387-398 (1928)

[23] G. D. Smith, R. J. H. Clark. Raman microscopy in art history and conservation science. Studies in Conservation Volume 46, 2001 - Issue sup1: Reviews in Conservation 2

[24] A. C. S. Talari, Z. Movasaghi, S. Rehman, I. Rehman. RAMAN SPECTROSCOPY OF BIOLOGICAL TISSUES. Applied Spectroscopy Reviews, Vol 50 Issue 1, 46-111. 2015

[25] X. Feng, A. J. Moy, H. T. M. Nguyen, Y. Zhang, J. Zhang, M. C. Fox, K. R. Sebastian, J. S. Reichenberg, M. K. Markey, J. W. Tunnell. Raman biophysical markers in skin cancer diagnosis. Journal of Biomedical Optics, Vol. 23, Issue 5, 057002 (May 2018).

[26] S. Tfaili, C. Gobinet, G. Josse, J. Angiboust, M. Manfait, O. Piot. Confocal Raman microspectroscopy for skin characterization: a comparative study between human skin and pig skin. Analyst, 2012, 137, 3673-3682.

[27] R. A. Rocha, J. M. Fox, P. G. Genever, Y. Hancock. Biomolecular phenotyping and heterogeneity assessment of mesenchymal stromal cells using label-free Raman spectroscopy. *Scientific Reports* volume 11, Article number: 4385 (2021).

[28] Horiba, XploRA™ PLUS Raman Spectrometer - Confocal Raman Microscope. https://www.horiba.com/gbr/scientific/products/detail/action/show/Product/xploratm-plus-1528/ Accessed online 31/01/24.

[29] Horiba, LabSpec 6 Spectroscopy Suite Software. https://www.horiba.com/gbr/scientific/products/detail/action/show/Product/labspec-6-spectroscopy-suite-software-1843/ Accessed online 31/01/2024.

[30] K. Castro, S. Pessanha, J. M. Madariaga. NONINVASIVE AND NONDESTRUCTIVE NMR, RAMAN AND XRF ANALYSIS OF A BLAEU COLOURED MAP FROM THE SEVENTEENTH CENTURY. Analytical and Bioanalytical Chemistry 391, 433-441 (2008)

[31] Raman Tool Set https://sourceforge.net/projects/ramantoolset/ Accessed online 31/01/2024.

[32] Wavemetrics, Igor Pro 9. https://www.wavemetrics.com/software/igor-pro-9 Accessed online 31/01/2024.

[33] Wavemetrics, baselines procedural file. https://www.wavemetrics.com/project/Baselines Accessed online 31/01/2024.

[34] Wavemetrics, Overview of curve fitting. https://www.wavemetrics.com/products/igorpro/dataanalysis/curvefitting Accessed online 31/01/24.

[35] A. Maćkiewicz, W. Ratajczak. Principal components analysis (PCA). Computers & Geosciences. Volume 19, Issue 3, March 1993, Pages 303-342

[36] The R Project for statistical computing https://www.r-project.org/ Accessed 31/01/2024

[37] posit, RStudio desktop https://posit.co/download/rstudio-desktop/ Accessed 31/01/2024

[38] H.G.M. Edwards and F.R. Perez. Application of Fourier transform Raman spectroscopy to the characterisation of parchment and vellum. II Effect of biodeterioration and chemical deterioration on spectral interpretation. Journal of Raman Spectroscopy 35, 754 (2004).

[39] N. Nicolaides, H. C. Fu , G. R. Rice. The Skin Surface Lipids of Man Compared with Those of Eighteen Species of Animals. Journal of Investigative Dermatology. Volume 51, Issue 2, August 1968, Pages 83-89

[40] A. B. Cua, K. P. Wilhelm, H. I. Maibach. Skin Surface Lipid and Skin Friction: Relation to Age, Sex and Anatomical Region. *Skin Pharmacology* (1995) 8 (5): 246–251.

[41] Austin Nevin, Sharon Cather, Demetrios Anglos, Costas Fotakis. Analysis of protein-based binding media found in paintings using laser induced fluorescence spectroscopy. Analytica Chimica Acta. Volumes 573–574, 28 July 2006, Pages 341-346

[42] U. P. Agarwal. Analysis of Cellulose and Lignocellulose Materials by Raman Spectroscopy: A Review of the Current Status. *Molecules* 2019, *24*(9), 1659

[43] J. J. Cárcamo, A. E. Aliaga, E. Clavijo, M. Brañes, M. M. Campos-Vallette. Raman and surface-enhanced Raman scattering in the study of human rotator cuff tissues after shock wave treatment. Journal of Raman SpectroscopyVolume 43, Issue 2 p. 248-254

[44] José J. Cárcamo, Alvaro E. Aliaga, R. Ernesto Clavijo, Manuel R. Brañes, Marcelo M. Campos-Vallette. Raman study of the shockwave effect on collagens. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. Volume 86, February 2012, Pages 360-365

[45] Stephanie Bellmaine, Alisa Schnellbaecher, Aline Zimmer. Reactivity and degradation products of tryptophan in solution and proteins. Free Radical Biology and Medicine. Volume 160, 20 November 2020, Pages 696-718

[46] Amy Robichaux Viehoever, Douglas Anderson, Duco Jansen, Anita Mahadevan-Jansen. Organotypic Raft Cultures as an Effective *In Vitro* Tool for Understanding Raman Spectral Analysis of Tissue. Photochemistry and Photobiology. Volume 78, Issue 5 p. 517-524

[47] Thomas Schmid, Petra Dariz. Shedding light onto the spectra of lime: Raman and luminescence bands of CaO, $Ca(OH)_2$ and $CaCO_3$. Journal of Raman Spectroscopy. Volume 46, Issue 1 p. 141-146

[48]C. Ghioni, J.C. Hiller , C.J. Kennedy , A.E. Aliev , M. Odlyha , M. Boulton , T.J. Wess. Evidence of a distinct lipid fraction in historical parchments: a potential role in degradation? Journal of Lipid Research Volume 46, Issue 12, December 2005, Pages 2726-2734

[49] Giacomo Reina, Silvia Orlanducci Emanuela Tamburri, Maria Letizia Terranova. Nanotechnologies for cultural heritage: Nanodiamond for conservation of papers and parchments. AIP Conf. Proc. 1603, 93–101 (2014)

[50] Giuseppe Pezzotti, Marco Boffelli, Daisuke Miyamori, Takeshi Uemura, Yoshinori Marunaka, Wenliang Zhu, Hiroshi Ikegaya. Raman spectroscopy of human skin: looking for a quantitative algorithm to reliably estimate human age. Journal of Biomedical Optics, Vol. 20, Issue 6, 065008 (June 2015)

[51] M. Bicchieri, A. Sodo, M. Montia, G. Piantanida, F. Pinzari. Critical issues in Raman characterization of parchment and leather. 5th International congress on the Application of Raman Spectroscopy in Art and Archaeology. Bilbao, Spain, 14-18 September 2009.

[52]Matija Strlič 1 *, Irena Kralj Cigić 2 , Ira Rabin3 , Jana Kolar4 , Boris Pihlar2 , May Cassar1. Autoxidation of Lipids in Parchment. Polym. Degrad. Stab. 94 (2009) 886-890

[53] Horiba, Raman Analysis. https://www.horiba.com/cze/scientific/technologies/raman-imaging-and-spectroscopy/raman-analysis/ Accessed online 31/01/2024

[54] Horiba, Spectral Resolution. https://static.horiba.com/fileadmin/Horiba/Products/Scientific/Molecular_and_Microanalysis/Raman_General/Spectral_Resolution_Tech_Note_RA-TN14.pdf Accessed online 31/01/2024.

[55] Antonia Malissa , Federica Cappa, Manfred Schreiner and Martina Marchetti-Deschmann. Spectral Features Differentiate Aging-Induced Changes in Parchment—A Combined Approach of UV/VIS, μ-ATR/FTIR and μ-Raman Spectroscopy with Multivariate Data Analysis. Molecules 2023, 28, 4584.

[56] A. Nevin et. Al. The Analysis Of Naturally And Artificially Aged Protein-Based Paint Media Using Raman Spectroscopy Combined With Principal Component Analysis. Journal of Raman Spectroscopy 2008; 39: 993-1000

[57] Microsoft 365, Microsoft Excel. https://www.microsoft.com/en-gb/microsoft-365/excel Accessed online 29/05/2024

[58]: Beasts2Craft – B2C Parchment Collection https://sites.google.com/palaeome.org/ercb2c/get-involved/b2c-parchment-collection Accessed Online 11/06/2024

[59]: Beasts 2 Craft – About B2C https://sites.google.com/palaeome.org/ercb2c/home/about-b2c Accessed Online 11/06/2024

[60]: Matthew D. Shoulders and Ronald T. Raines. Collagen Structure and Stability. Annu Rev Biochem. 2009 ; 78: 929–958. doi:10.1146/annurev.biochem.77.032207.120833

ACKNOWLEDGMENTS

I would like to firstly thank my supervisor, Dr. Y. Hancock for being able to provide a project where I am able to use my scientific knowledge to pursue my love of art and art history in a way that doesn't require me to pick up a paintbrush. She has provided academic and personal support throughout the process whenever I needed it and this would not have been possible without her guidance.

I would also like to thank my family for keeping me company and comfortable at home whilst I wrote this thesis. I am so lucky to have unconditionally supportive and encouraging people in my life.