Application of Hyperspectral Imaging in Pharmaceutical Analysis

A thesis submitted in part fulfilment of the requirements for the degree of Doctor of Philosophy

by

Joseph Muench

Strathclyde Institute of Pharmacy and Biomedical Sciences

University of Strathclyde

2013

Author's Declaration

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Signed:

Date:

Abstract

This thesis describes the application of hyperspectral imaging (HSI) as a novel technique for the analysis of spectral data derived from image analysis of tablet breakdown during dissolution. Whilst defining the rate of release is the common output from traditional dissolution experiments, the intention of the research presented in this thesis was also to describe the physical changes occurring within the disintegrated mass during the dissolution process.

The initial stages of the investigation focused on determining which wavelength ranges were most discriminatory in separating out similar polymer samples; the two wavelength ranges investigated were the visible (400-860 nm) and the near infrared (835-1650 nm). As the polymer samples were similar, comparison of the spectra for distinguishing features was unsuccessful and principal component analysis was used to separate the spectral signals. The near infrared (nIR) was shown to be the most effective wavelength range at separating signals due to an increased number of peaks for comparative analysis. The examination of the dissolution of paracetamol tablets produced by several different manufacturers was described in this thesis. In particular, the rate of expansion of tablet material and the identification of caffeine containing regions during tablet dissolution. The initial results tracked the expansion of the tablet within the flow cell and determined the cause of the signal attenuation affecting the results. The final experiments combined the spatial and time resolutions of the previous experiments. This system used a minimised path length to reduce signal attenuation and give spectra with enhanced spectral features. This system was able to show caffeine rich regions in the tablet during the dissolution in a caffeinated paracetamol brand, the rate of caffeine loss also being calculated. These results were then corroborated using conventional analytical techniques to establish whether HSI had robustly measured the release of a known compound during dissolution.

Acknowledgement

Firstly I would like to thank my many supervisors for their invaluable advice, knowledge and guidance throughout my time in Glasgow, especially Dr Andrew Urquhart, Professor Stephen Marshal and Professor Clive Wilson. Thanks also to AstraZeneca for funding the PhD and to my industrial supervisor Stephen Wren who was always on hand to answer any pharmacy questions and who also helped with proofing the thesis. I would also like to acknowledge Tim Kelman and Dr Paul Murray from the Hyperspectral Imaging Centre for their help. You were always around to help me when I had a little panic or when MATLAB wouldn't work, which was quite often.

I would also like to thank the people at Gilden Photonics for the help throughout this project, their knowledge and expertise was invaluable in both the use of the equipment and the design of new experiments. I would like to thanks Suvitha Karthick, Neil Cairns and Dr John Gilchrist in particular for their help and assistance.

On a personal note I would like to thank my family especially my mum and dad whose support along the way was invaluable. A huge thanks also goes to all the friends I made while I was up in sunny Scotland. Your good cheer, unique banter and shared appreciation of fermented fruit drinks were a great help throughout my time in Glasgow.

Finally I would also like to thank Helen for all the love and moral support.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 Hyperspectral imaging

1.1.1 Introduction

Hyperspectral imaging (HSI) is an imaging technique which allows the user to capture large amounts of spectral and spatial data in a very short space of time. In a standard digital camera at each pixel the spectral information is only represented as intensities of set wavelengths to give the coloured images [1] (red, green, blue etc.) Hyperspectral imaging however gives the spectral information of the image taken with a predetermined set of wavelengths (e.g. 300-1000 nm). At each pixel in the hyperspectral image a spectrum, of a specific range of wavelengths, is recorded with a spectral resolution as high as 3-10 nm [2]. Until recently HSI was an expensive technique to implement due the high costs of the optics and electronics required. The expensive nature of these systems meant that the only groups to use hyperspectral imaging for many years where those with large budgets such as NASA and the military. Recent advances in detectors have reduced the cost of production resulting in cheaper hyperspectral cameras, which has resulted in the technique becoming more utilised in the larger scientific community. It is quickly gaining a reputation as a fast reliable imaging system for the gathering of large amounts of spatial and spectral information.

The data received from the hyperspectral imaging is stored so as to give the user an image in the (x,y) plane where each individual pixel contains the spectral information in the z plane and this data is referred to as a 'hypercube [3]. Figure 1-1 is an example of a hyperspectral image, the image on the left shows two pieces of tablet on a glass slide and the plot on the right is a spectrum generated from a pixel in the tablet. In a hyperspectral image the x and y axis have the spatial information while the z axis shows the spectral information. In this image the colour on the z axis indicates intensity of the spectrum at each wavelength for the edge pixels and it is clear that the tablets give a strong spectral signal by the deep red intensity. An image of this size can be captured within a few seconds and contains thousands of individual spectra.

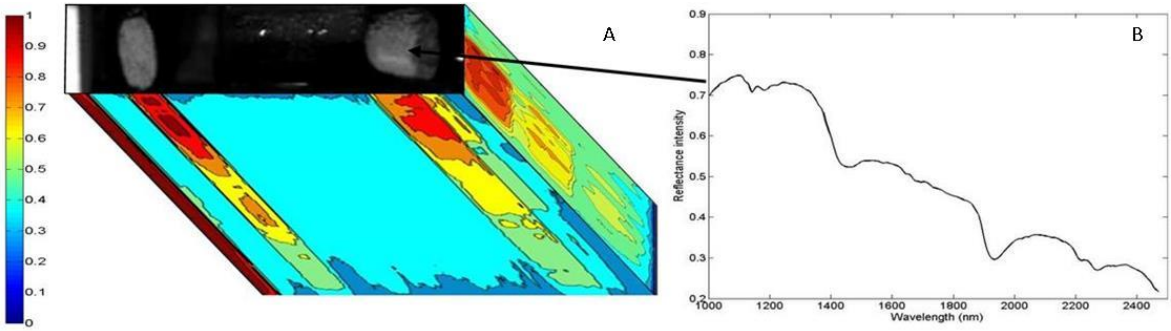


Figure 1-1 An example of the data within a hyperspectral image which is stored as a hypercube. Figure 1-1A is of two tablet pieces and the edges of the cube show the intensities of the spectral wavelengths for each spectra contained within the edge pixels. Figure 1-1B is a spectrum generated from a single pixel in the image.

A hyperspectral camera is an imaging device which can capture huge amounts of both spatial and spectral information in just a few seconds. There are two main components to a HSI system, the first is a charge-coupled detector (CCD) which stores and translates the electrical impulse received into spectral data. The second component of the system is where the light is manipulated or filtered in some way; this component is a liquid crystal tuneable filter (LCTF) [4,5] , an acoutso-optical tuneable filter (AOTF) [6,7] or a spectrograph [8,9].

1.1.2 Building a hyperspectral image

There are three ways to create a hyperspectral image point and these are known as point, plane or line scanning. The detector for any HSI camera is determined by the wavelength range used and not the approach to generating the hyperspectral image. The theory of how the detectors work and the composition of the detectors are discussed in detail in Chapter 2.1.1. The three approaches to creating a hyperspectral image differ in how the image is created and each use a different system for separating wavelengths to be captured by the system.

1.1.2.1 Point scan hyperspectral imaging

The point scan approach to hyperspectral imaging is where the image is built up one pixel at a time with the system collecting a full spectrum at each pixel. This is the process used to create Raman [10,11] or Infrared (IR) [12,13] mapping which, while not often discussed as such, are types of hyperspectral image. The major advantage using a point scan is that the very high spectral resolution as the spectroscopy is performed using normal Raman or IR equipment; however a mobile stage and specific software for performing the analysis is

required. Disadvantages include the target being motionless and an extensive time required for data capture. These disadvantages make this technique a poor choice for dynamic systems or large target areas. However, when the sample is stationary, stable and a small target area is sufficient then point scan HSI can be very powerful.

1.1.2.2 Plane scan hyperspectral imaging

Plane scan hyperspectral imaging is a technique that uses a filter which only allows a specific wavelength of light to pass through. However, these filters are not fixed and the wavelength filtered can be altered in a controlled manner. The two main filters used are the Acousto-optic filter and the liquid crystal filter. Therefore with this system the target is imaged at the first wavelength, the filter changed and the target re-imaged. By proceeding through the different wavelengths a hypercube is created with the full spectral data at each point. Plane scanning is very sensitive to movement as the images are built up over time and so if the target is undergoing a physical change then the image will be distorted by this change. The filters are also sensitive to temperature changes and so are often cooled by bulky cooling systems to prevent any fluctuations from the ideal operational temperatures the system is calibrated for. However, as they don't require additional setup other than the exposure time for the system they are excellent at quickly imaging static objects such as a tablet surface [14].

1.1.2.3 Line scan hyperspectral imaging

The third system for generating hyperspectral images is to create the image one line at a time and is known as push broom or line scan hyperspectral imaging. These systems collect a narrow beam of photons from the target and require the movement of either the camera or the system to generate the picture. A spectrometer is used to split incoming light, separating out the wavelengths for detection. The theory behind these systems is described in greater detail in Chapter 2.1.2. These systems are more robust than the other HSI systems, without the need for cooling or the sensitivity to movement that some of those systems suffer from. Due to the requirement for either the sample or the camera to move these systems have been employed extensively by the remote sensing community[15] where large areas of land are imaged using planes or satellites.

1.1.3 Remote sensing applications of hyperspectral imaging

Hyperspectral imaging was first developed for the remote sensing community as a way to image large areas with sufficient spectral information to investigate processes or changes which cannot be easily recorded on foot. Differentiating between similar plant types in coastal regions or large tree canopies [16,17], detecting mineral deposits and separation out similar ores across mountainous areas [18,19], water quality in coastal areas [20] and classification of building and infrastructure in urban areas [21,22] are just a few examples of how HSI is used in remote sensing. HSI is also used by the military, while many uses are not published for obvious reasons there are some declassified documents which have been shown in the literature. These detail the use of the technique in identification of target and target areas such as using satellite data to find landmines in a minefield [23]. The use of HSI data has also been used to detecting potential military targets such as planes [24] with a high degree of success in an automated system which reduces the time to spot these targets.

1.1.4 Industrial applications of hyperspectral imaging

One of the largest fields of hyperspectral imaging outside of the military and remote sensing imaging is its use in industrial quality control systems, especially the food industry. The large volume of spatial, spectral and temporal information allows the user to check the quality of a large area very quickly whether it's looking for diseased or bruised apples[25] or checking the uniform distribution of a polymer blend [26].The use of HSI to investigate polymer uniformity [26] is of particular interest; here Gosselin *et al.* used a mixture of chemometric methods such as multivariate image analysis and partial least squares regression to show the average distributions of the polymer blend. This can be done using current techniques but only by removing the film from the process and testing it in the laboratory. Hyperspectral imaging however, can be done while the polymer films are still being made to give real time accurate feedback about polymer distributions in the polymer blend.

In the case of checking for bruising of apples[25] the technique is used to penetrate just under the skin on the apple where the earliest signs of bruising are shown, these signs are easy to detect by looking at the data using specific wavelengths which show a spectral difference between bruised and whole apples. By viewing the full image at a single wavelength a greyscale depiction of the intensities within the image at that wavelength is created, which shows the areas of bruising under the surface of the apples. This system allows for any heavily bruised or damaged apples to be removed from the packaging process giving

higher quality results in the shipping of this product which reduces returns and boost profits.

There are numerous different publications showing the use of HSI to detect disease [27,28], bruising [29,30] or contamination[31,32] in food products by using techniques similar to those above. The analysis of these images can be performed quickly and could potentially be used in an automated system. These systems can be used to save companies money by reducing waste; one such example of this is in paper mills.

There are also examples in the literature where hyperspectral imaging has been tested to see if it can be used to separate better quality products which could be sold at a higher price. One such study looked at the tenderness of beef steaks[33]. The tenderness of a steak is related to key biochemical properties (cytoskeletal proteins etc.) and the muscle structure, both of which can be measured using hyperspectral imaging. By investigating the spectra of these areas it was possible to distinguish between the lean and fat areas of the beef using simple spectral comparisons. For example the reflectance in a pixel of fat starts at a much higher value (approximately 0.4 compared to 0.25) and the position of certain peaks are shifted. These differences could be used in computer software for an automated system of detecting the quality of the meat as a quantitative value of fat percentage in a piece of meat. This is just a small selection of the wide breadth of uses hyperspectral imaging has in industrial applications.

1.1.5 Laboratory applications of hyperspectral imaging

The remote sensing and industrial uses of HSI are often performed on a large scale but there are many applications of HSI in the microscopic scale; mainly it is used for the detection and identification of cells [34-36]. As Hyperspectral imaging requires an illumination source to produce an image the only thing necessary to view microscopic samples is simply an amplification device such as an adapted microscope [34]. This uses the microscope as a zoom lens with the reflected or transmitted photons entering into the hyperspectral camera rather than an eye piece. This is the same principle as is seen in Raman and near infrared (nIR) microscopy systems. In the microscopic scale however most measurements are done using fluorescence and so require more advanced equipment to detect the data received along with a different microscope set up.

The identification of specific bacterial growths in a group of bacteria is just one example where microscopic hyperspectral imaging has been implemented previously [34]; this was done by the detection of specific wavelength signatures unique to the bacteria in question. These wavelengths signatures are specific differences in the spectra of the bacteria which can be used to distinguish between them. A false colour image can be created by using these wavelengths which show clearly the different species of bacteria in the image which may be otherwise difficult to distinguish. There are limitations to this technique as with most hyperspectral imaging of this kind in that overlapping bacteria (or samples of any kind) can lead to confused signals which are a combination of the two sources. These should be easily detected, especially when using false colouring, and removed from the sample or solved using multivariate statistics.

Hyperspectral imaging has also been used to investigate areas of interest on or near the surface of an animal [37] or even internally using an endoscopic approach to HSI [38]. The detection of skin tumours or melanomas can be difficult in the early stages as they can leave little visible marking on the skin or indeed just a distortion of the skin rather than a discolouration. Using hyperspectral imaging researchers have been able to correctly detect these tumours [39,40] by looking for the change in certain specific peaks which indicates the presence of the growth. The fluorescence data from these systems clearly shows that the spectra for the malignant tumour has a much lower intensity than normal tissue up to approximately 520 nm after which the intensity of the fluorescence from the tumour is higher than that of the normal tissue. These changes in the spectra between normal skin and a tumour are due to the changes caused by the tumours such as an increased melanin concentration and a change in the blood flow through the affected area.

Hyperspectral imaging has also been used to research levels of certain compounds in the blood of patients[41,42]. These papers where looking at the levels of haemoglobin in the blood of test patients with one paper specifically investigating patients with sickle cell anaemia to discover whether certain treatments help oxygen saturation. Using HSI was an ideal way to detect levels of haemoglobin using a non-invasive technique which can be used over a period of time to show how the levels of haemoglobin change both before and after the release of a pressure cuff. The hyperspectral imaging allows for a very detailed analysis of the concentrations of the target compounds with minimal discomfort for the patient as the imaging is fast and non-invasive.

1.1.6 Hyperspectral imaging in pharmaceutical analysis

There are two 'styles' of hyperspectral imaging in pharmaceutical analysis, the oldest and most established is chemical imaging. Chemical imaging is traditionally a style of point scan imaging used with near infrared, infrared and Raman spectroscopy. This style sacrifices speed to spectral resolution with larger images often taking hours to complete but with a spectral resolution much higher than is normally attributed to hyperspectral imaging systems. The use of hyperspectral imaging systems, specifically those with a liquid crystal tuneable filter (LCTF) or acoutso-optical tuneable filter (AOTF) filter or plane scanning spectrograph, is a more recent addition to the field of pharmaceutical analysis but is quickly becoming more common in the literature. However, it should be noted that both of these techniques are hyperspectral in nature and the distinctions only exist in so far as how the data has been collected previously and the nomenclatures that were created for them. Chemical imaging and hyperspectral imaging are becoming interchangeable terms [14] and so for ease the term hyperspectral imaging will be used throughout this work to prevent confusion.

Near Infrared spectroscopy is common in pharmaceutical analysis and most HSI systems in this field detect in the near infrared. A traditional spectrometer will average the signal across the target area giving a general description of the bulk properties of the sample. Pharmaceutical products are often complex systems which are designed to release the drug contained within at the intended source at a predetermined rate. The greater the degree of information which can be discovered on these formulations the greater the level of design which goes back into creating better formulations. The use of tradition vibrational spectroscopy techniques reduces the volume of information on the sample relative to a hyperspectral approach and this is why these techniques are being investigated. Hyperspectral imaging does not average over a large area but gives spatial resolved spectral information throughout a system.

Tablets are designed to be homogenous with respect to the excipients and active pharmaceutical ingredients contained within. This prevents any unusual or unexpected release properties which may be caused by clustering of the components. Investigations into the homogeneity of tablets have been performed using HSI in an attempt to monitor the concentration difference throughout [43,44]. In these papers different approaches have been taken to the same goal of identifying the major components in a tablet using a nIR HSI system. The first paper by Lopes *et al* [43] uses a style of post image analysis called

simples identification via split augmented Lagrangian (SISAL) which is a form of spectral unmixing algorithm. This algorithm assumes that the data within a pixel is not pure, that is that more than one component is present and causing the spectral signal to be a composite of these different compounds. The algorithm then calculates the abundance of each component in each pixel which shows the efficiency of the mixing. The approach of Ravn *et al* [44] was to test three separate forms of data analysis to determine which is the best for spectral unmixing of the data. These were a single wavenumber method, classical least squares regression and partial least squares regression. It was found that the partial least squares gave the best results with the highest degree of accuracy however, this method does require a calibration set which may not always be available. Both of the papers detail how HSI can be used, in different ways, to ascertain the relative concentration of components from a tablet in a spatial resolved manner. This shows the ability of HSI as a tool for investigating the mixing of tablet materials. Another example is the use of hyperspectral imaging to test the concentrations of acetylsalicylic in commercially available aspirin tablets [14,45]. Here a near IR spectral range was used to scan the surface of the aspirin and the data received along with chemometric approaches like those above was converted into concentrations of active pharmaceutical ingredient (API) and then the image of the tablet recreated as a relative concentration map.

There is also interest in the use of HSI to identify counterfeit drugs [46,47]. Hyperspectral imaging can be used to test large batches of drugs rapidly which allows for quicker validation of suspect drugs. The work by Lopez *et al* [46] showed that using HSI it is possible to take a large sample of tablets of unknown content (although the intended API was known) and discover which of these tablets were counterfeit. Furthermore the use of principal component analysis (PCA) showed that the counterfeit tablets could be grouped into 13 separate groups. Within these groupings it was also possible to determine which of the counterfeit tablets were placebos and which contained either too little API or a substitute API. The exact composition of the counterfeit tablets was not investigated but it was suggested that using further analysis it may be possible to determine which APIs and excipients were used. These results show how HSI can be used to quickly distinguish between tablets which by visual inspection may seem similar but are in fact very different. This was performed quickly, with multiple tablets being imaged at the same time and with minimal post imaging analysis needed to determine the counterfeit drugs.

1.2 Tablets

Tablets are a common approach to the delivery of drugs into patients; they are a simple way to deliver a controlled dosage to the patient over time with good stability.

A tablet contains a number of compounds other than the API and these other compounds are used to enhance the physical properties of the tablets to improve its efficiency. These other compounds are known as excipients and can be used to create a range of different physical properties in the tablets. A large proportion of tablet bulk is often a filler excipient, such as cellulose or starch, which is used to bulk up the tablet. This is used for easier processing of the tablets and to make the drug a suitable size for patient consumption as many tablets contain only a few micrograms of API. A tablet containing only this small volume of API would be too small to handle easily and effectively. Other excipients, such as magnesium stearate, are lubricants which help the tableting process during the expulsion of the tablet from the tableting machine. Compounds which aid the disintegration of the tablet by promoting expansion are also common and lead to improved release properties for the tablets. Binders are excipients that hold the tablets together after the tableting process and some such as starch and cellulose are also used as disintegrants. Preservatives and sweeteners are also used in tablets for a longer shelf life and to mask the unpleasant taste of the tablets. Coating tablets is commonly used to disguise the taste of some drugs making them easier to swallow. They are also used on tablets which are designed to release their API further along the gastrointestinal (GI) tract than the stomach. By coating the tablet in a compound resistant to the stomach acids but susceptible to the conditions later in the GI tract the tablets release the API in the desired location.

Tablets are created as a thoroughly mixed blend of excipients and API which are then pressed into a tablet using compression of the loose material into a die. The mixing of the tablet ingredients is not always possible using pure dry powders as these may cluster together due to electrostatic charged created by small particle sizes. To allow for even mixing throughout the tablet a process of granulation is performed. Granulation is a process of combining the desired API and excipients to form controlled particle sizes using either a liquid binder or light compaction and low pressures. By using this process the mixing of the API and excipient is assured and tablets can be made which show a homogenous distribution of API and excipients throughout.

Before any tablet can be approved for general use it must undergo a number of tests to determine the quality of the tablet. For example tablets must be tested to show that the ingredients are evenly mixed and within specific quantity limits. Two techniques used to test tablets are dissolution testing to investigate the release rate of the drugs and vibrational spectroscopy to investigate the mixing, purity and surface quality of the tablets.

1.3 Pharmaceutical analysis

1.3.1 Introduction

Pharmaceutical analysis is a tightly controlled rigorous discipline with strict guidelines and procedures controlling the analysis of pharmaceutical products before they can be sold. These procedures are detailed within the pharmacopoeias, with the United States Pharmacopeia [48] containing many of the most common techniques and procedures for testing pharmaceutical products of all types.

1.3.2 Dissolution testing

Dissolution testing is the standard experimental procedure for investigating the release of the drug or active pharmaceutical ingredient (API) from a pharmaceutical product. It uses ultraviolet (UV) spectroscopy to detect the increasing intensities of spectral peaks in the UV from specific drug absorptions. Then using the Beer-Lambert law and known standards the change in spectral absorbance can be attributed to a specific change in the amount of API released into the system. The dissolution properties of a compound are important as the drugs cannot be absorbed through the gut if they are not dissolved into the gastrointestinal liquids. The specific protocols used in dissolution testing are detailed within the USP and a number of different types are available for different experimental needs [48].

The USP contains four different dissolution apparatus which are designed for different experiment needs and pharmaceutical products and may even give different release rates for certain drugs [49,50]. Each system is different but there are core similarities between them, each uses a 900 ml solution, often acidic buffer, into which the pharmaceutical sample will dissolve. The use of the acidic buffer is to provide a basic model of the environment within the stomach although the stomach acids are more complex, containing numerous enzymes to aid in the breakdown of food. Each procedure uses a UV spectrometer to measures the

solution and follow the release of API into the system, each requires heating of the solution to 37°C and each system thoroughly mixes the solution.

There are a number of different areas of research within dissolution testing other than the testing of drugs to show release rate of a new formulations or for testing batch to batch variability. Investigations into complimentary drugs or excipients which improve the absorption rate of the target drug have long been an area of great interest. An increase in the rate of gastrointestinal absorption by the addition of a secondary compound to the target drugs was shown by Goldberg *et al* [51] in 1966. Improving the rate of uptake makes your drug formulation more effective and thus more desirable.

Before the USP dissolution procedures were agreed as a standard there were a number of different dissolution approaches. Mattok *et al*[52] investigated three different dissolution procedures for the release of acetaminophen (paracetamol) and found each to give different results. During the experiment it was also shown that older tablets may have a slower release than recently created tablets and no tablet showed release comparable to the physiological availability of acetaminophen estimated from blood and urine. This paper shows why it is important to have a rigorous set of standards, that some tablets degrade over time and that the release of API from a tablet in a reaction vessel is different to the release and uptake inside living organism. There are some good reviews on where USP dissolution apparatus is currently in the field of pharmaceutical analysis and the changes that may be needed in the future [53].

Van der Weerd *et al* [54] looked into combining FTIR imaging and dissolution testing to investigate the release of API from a tablet with the image to show the spatial locations where this occurs. The outcome of this was to show how the release of a niacinamide drug from a HPMC tablet occurs homogenously throughout the region of tablet imaged. As the dissolution progressed a gel state is seen to form around the HPMC which is the major excipient in the tablet. This is an example of how spatial information can show physical changes during the dissolution which are not seen with the traditional UV dissolution apparatus alone. There are however problems with the work as the tablet was placed under pressure constantly from the golden gate holding it close to the IR beam and a very restricted flow space for the water to dissolve the tablet in. It is promising to see the potential of imaging and spectral analysis merged in this way.

A new vista in dissolution testing is the combination of the UV spectrometry and imaging to create an image of the dissolution of samples in a time resolved manner. The images are captured using a band pass filter to select the desired wavelength and then shining monochromatic light through the flow cell containing the target sample. These are not hyperspectral images as they are only captured a single wavelength and not the full spectrum. These imaging systems show great promise for pharmaceutical analysis of the in situ release of API from different pharmaceuticals.

The use of this technique on transdermal patches was performed to investigate patch release profiles and also to highlight the ability of this technique for analysis [55]. The patches contained nicotine as the active ingredient with no other UV active compound so determining the wavelength for imaging was relatively simple. One of the complications with these imaging systems is when multiple UV active compounds are present as it may be difficult to calculate the concentration in the system with only a single wavelength. The imaging clearly showed the drug being released from the patch with release rates being calculated from the data and corroborated against standard UV dissolution tests. This work clearly showed how UV imaging is able to better show the release properties of a simple drug system and the potential of the technique for more complex systems or analysis. One major problem with this paper is that the spatial information is not used fully and the data is used only to show the change in the concentration of drug over time. There is no discussion of the changes in the shape of the patch or rate of change in the size of the patch which are simple analyses which could lead to more information on the physical changes occurring during the dissolution.

The UV imaging of the dissolution of a pre-made tablet of pure Amlodipine [56] does contain more analysis on the spatial information contained within the data. Much of this work mirrors that done previously, using the data to calculate the release of the API. However, this work also investigates the images and the physical effects which are occurring that can be detected. The images showed that the release of one formulation of Amlodipine tablet was much quicker the others. Also by investigation of the images it is possible to see the uneven distribution of the tablet as the flow causes the preferential release from one side of the tablet. This is visible as a slight skew in the tablet and as a thin line of dissolved material at the side of the tablet. This work shows how it is important to utilise the spatial information in this work as by only considering the spectral information the imaging becomes less powerful.

Dissolution testing is a powerful tool for analysing the release properties of a drug system, it now has a firm place in pharmaceutical analysis with tightly controlled procedures. The data collected when a UV spectrometer is used to detect the release of the API is an average of the total dissolution in the system and is dependent on sufficient mixing and diffusion throughout the system to prevent erroneous results. It also shows no information on the physical properties of the tablets during the dissolution such as rate of swelling or any unusual tablet properties. However, UV imaging is a novel method for analysing dissolution which gives these physical changes alongside better, spatially resolved information on the release of the API from the system.

1.3.3 Vibrational spectroscopy

1.3.3.1 nIR spectroscopy

Near-infrared (nIR) spectroscopy is a fast, non-destructive technique which is a common tool for pharmaceutical analysis. The nIR region is between the visible and the mid infrared and is defined as being between the wavelengths of 780 – 2526 nm. The region was first discovered by Herschel[57] in 1800 when he separated the electromagnetic spectrum using a prism and noted that the temperature increased greatly past the red end of the spectrum. This region in which the temperature rose is the near infrared. Experiments have been performed using nIR since the early 1920s but more advanced statistical analysis alongside reduced production costs were needed before major advances could occur with the technique. In the 1960s Karl Norris from the U.S. Department of Agriculture recognised the potential of nIR and introduced it for the analysis of agricultural and food products [58].

The use of nIR spectroscopy for pharmaceutical analysis is now widespread with the technique being used for quality testing of materials before, after and during production. The increased interest in the technique is due to a number of factors including the fast nature of the data collection, the lack of any pre-treatment steps, the non-destructive nature of the technique and advances in chemometrics.

The bands seen in nIR spectroscopy are from the combinations and overtones of the fundamental vibrations of the -CH, -NH and -OH groups with some contributions from other similar functional groups. These vibrations are much rarer than the fundamental vibrations seen in the mid IR and so are much weaker in intensity, they are also strongly affected by dipoles and hydrogen bonding which slightly changes the energetics of the

systems. This creates broad absorptions which are the result of a number of similar but not identical absorptions which are so close that they cannot be distinguished by the measurements. There is also constant background irradiation from thermal energy, which cannot be removed that causes a broadening of the absorption peaks by exciting all molecules and causing some degree of vibration. Thus a nIR spectrum contains peaks which are broader and weaker than their corresponding mid IR absorptions. However, there is an advantage to the weaker nIR signal as the signal penetration of a nIR photon is much greater than that of the mid IR as they are less readily absorbed. This is caused by two factors, firstly that the energy transitions caused by near IR irradiation are less likely to occur and so a nIR photon can, on average, pass further through an absorbing medium before interacting than mid IR photons. The second reason is that the scattering of light, as explained by Rayleigh, is proportional to λ^{-4} where λ is the wavelength. This means that the longer wavelengths of the mid IR photons will be scattered more than near IR photons by the same particles, which impedes signal penetration. The theory of vibrational spectroscopy is covered in greater detail in Chapter 2.21.

This increase in signal penetration gives rise to a greater path length through the sample and a greater sensitivity to the physical changes in a sample, such as particle size, than the mid IR. Changing particle sizes causes variability in the degree of scattering in a sample which with chemometrics and reference data can be detected and tracked throughout a sample.

As these peaks are broad they contain less spectral definition and similar peaks from two or more distinct environments may overlap. This data is multivariate in nature and so the spectrum requires more complex statistical analysis than its mid IR counterpart. This analysis is multivariate and uses the large quantity of information within the peaks to better distinguish the small changes between and within spectra which are of interest. One of these techniques is principal component analysis and this is explained in more detail in section 1.4

The use of near infrared spectroscopy for pharmaceutical analysis has become much more prevalent since the 1990s due to improvements in both the hardware and software used to capture and analyse the spectra. The technique is now used to investigate a wide range of different pharmaceutical forms and processes from tablets to polymer implants [59]. One major use of nIR spectroscopy is for the identification of raw materials before processing or in production lines as it can be used as a non-destructive technique to track and identify

materials in real time without the need to remove samples to an external location for testing. There are a number of papers which discuss the use of nIR for the identification of both active pharmaceutical ingredients (API) and excipients [60,61]. NIR spectroscopy is also used to look at the copolymer ratios in some common pharmaceutical copolymers [62] and even how the characteristics of certain excipients affect the tablet properties [63]. Other physical attributes of raw materials have also been analysed using nIR such as particle size determination [64-67], polymorphic forms [68,69], the ratio of amorphous to crystalline forms [70-73] and moisture content [74-79].

The determination of moisture content within either excipients or the finished pharmaceutical product is of great importance as both the excipients and APIs may be largely affected by the presence of water, with potentially lower efficacy. While there are a number of techniques used to analyse this property, such as Karl Fischer titration [80], nIR spectroscopy is a powerful approach due to its sensitivity to the OH vibrational stretching frequencies observed in the water molecules. The O-H bands of water are intense and include 5 separate absorption peaks at 760, 970, 1190, 1450, and 1940 nm. These peaks are from different overtones and combinations of the fundamental vibrational frequency and so each has a different intensity and the strength of the hydrogen bonding in the sample structure will affect the frequency. It is this characteristic of near infrared spectroscopy which has caused it to be a powerful tool in moisture content analysis and there are a great number of papers on this subject.

One of the most common pharmaceutical dosage delivery forms are tablets as they are small, provide an easy form of dose regulation and stable. However there are a number of different parameters for which nIR spectroscopy is used to detect, track or identify such as tablet hardness[81-85], thickness [86,87], dissolution rate [88,89], coating thickness [87,90,91], product degradation [92,93] and identification of tablet components [47,94-98]. The deep penetration of nIR spectroscopy also allows for the identification of tablets inside sealed blister packs [99,100]. These are just a few of the papers published on nIR analysis of tablets which shows the breath of information which can be determined using this technique.

1.3.3.2 Raman spectroscopy

Raman scattering was discovered by Sir C. V. Raman in 1928 [101] and Raman spectroscopy is a form of vibrational spectroscopy which involves the excitation from a vibrational energy level to a virtual energy level and the subsequent relaxation of the bond to a lower

vibrational energy level with the emission of a photon. It is these photons which are detected during the spectroscopy. The theory of Raman spectroscopy is covered in greater detail in Chapter 2.21. Near IR and mid IR spectroscopy detects the vibration of a bond which undergoes a change in the diploe moment of that bond during vibration. Raman spectroscopy detects vibrations caused by a change in polarisation of that bond and can detect non-polar bonds such as carbon-carbon backbones. Due to this difference between IR and Raman spectroscopy an intense absorption from a functional group in the IR does not necessarily appear in the Raman and vice versa. One of the most important differences for pharmaceutical analysis is the absorption by water, in the IR this absorption is strong and causes issue for any spectral definition in this region. However the water absorption spectrum in the Raman is very weak and so it is possible to investigate compounds and samples which are in or contain a large quantity of water.

Raman spectroscopy is well suited to pharmaceutical analysis and in particular the investigation of solid state forms of pharmaceutical products such as tablets[11,12,102-106]. The technique is specific and can be used to detect very low concentrations of a target compound. Widjaja *et al* [102] showed that the pure spectra of a compound which has a concentration of 0.2% by weight can be calculated using a multivariate approach. With this analysis the spectra of a tablet were captured and run through an advanced multivariate statistical algorithm to generate the pure spectra of each component in the tablet with a high degree of accuracy. This is just one example of using Raman spectroscopy to see very low concentrations of a target molecule without a calibration set of spectra. The use of multivariate statistics is very common in Raman, especially Raman mapping as there is a large volume of high resolution spectral data which is often comprised of multiple signals combining. [107-110] Raman mapping is a common technique which creates a hyperspectral image of the target [10] and is used to study the concentrations of the surface and near surface of a target. This process is often lengthy as each spectrum is collected individually and a Raman spectrum takes longer the capture than a standard IR measurement.

1.4 Multivariate analysis of data

Multivariate analysis is a set of techniques which are used to analyse large sets of data which contain substantial amount of variables and observations[111]. It is commonly used in the fields of nIR spectroscopy and chemical imaging, Raman mapping, HSI and image analysis where there is a large volume of data which cannot be easily analysed using more basic statistical analysis or plotting of the data. With only two or three variables it is

relatively simple task to plot the data and see trends but this is insufficient for analysing data sets with large numbers of variables. Principal component analysis, partial least squares, classical least squares and multivariate regression analysis are just a small number of the different approached to multivariate analysis which can each be used to determine different properties of the data such as isolating groups of similar signals or identifying specific signals. As with other chemometric systems these analyses are mathematically complex and thus only started to become more common with the improvement in computers and the software packages to perform the analysis.

1.4.1 Principal component analysis

Principal component analysis (PCA) is one of the most common forms of multivariate analysis and is used to identify clusters of signals which are similar within large data sets which contain shared variables. The techniques used today for PCA were first developed by Hotelling in 1933 [112,113] although various forms of PCA were discovered prior to this with the first recorded paper being published in 1901 by Karl Pearson [114]. The theory of PCA is detailed in Chapter 2.5.1 but for the ease of the reader a brief description will be included here. Principal component analysis is a way to reduce down the dimensionality of a large data set with numerous variables to a small number of principle components. These components show the largest variation within the data set with each subsequent component showing less of the total sample variation than the one preceding it. In this way a set of vectors or axis are created which can show the majority of the variation from a system containing many variables in as few as two to four new principal component 'variables'. This allows for similarities between samples in a complex system to be determined such as spectra which share a specific peak or regions in an image which contains the same material. These principle components contain two factors which are the scores and loadings. The scores are a measure of the variation shown by each sample for that component and the loadings detail which variables are causing the variation which is being described in the data.

Principal component analysis is often employed in hyperspectral imaging as a way to determine which areas within an image which have a similar spectral response. Remote sensing imaging is one of the oldest forms of the HSI and involves the imaging of large areas of land either by plane or satellite often to look for specific vegetation [115,116] or mineral deposits [18]. In remote sensing it is common to use the PCA as way of identifying aspects for further analysis especially the loadings which can show the wavelengths which

contain the greatest degree of variation in the images. This can reduce down 200+ wavelengths of interest to 5-10 wavelengths which is a much smaller sample set for comparison [115]. The use of PCA in HSI, as either a standalone analysis or a midpoint for more complex analysis, is also common in laboratory experiments. PCA has been used to analyse a wide range of systems such as bacterial colonies[117-119], demineralisation of enamel [120] biological materials such as lung cells [121] and even determining food quality [122,123].

Principal component analysis is also very commonly used in the fields of vibrational spectroscopy especially the near infrared where peaks are broad and spectral difference slight. The analysis is often performed to distinguish between numerous spectra and to group similar spectra together as a way of identifying them. There are numerous different examples of this in the literature such as identifying tablets within sealed blister packs [99], identifying similar cellulose tablet excipients [60], analysis of cellulose and lactose content in ecstasy tablets [124] and determining the degree of water absorption onto microcrystalline cellulose[125]. There are also a number of reviews which cover the use of PCA within the areas of nIR spectroscopy in greater detail [126-129].

1.4.2 Other multivariate approaches

There are a number of other multivariate analytical techniques which are used in the analysis of vibrational spectroscopy and hyperspectral images. Partial least squares (PLS) analysis is used as a way of deconvolving composite signals to generate the original spectral data along with relative intensities of this data. If the image pixel is sufficiently large then the spectrum collected for that pixel will be a composite created from the addition of multiple different spectral sources. The resolution of an image pixel is dependent on what magnification has been used and the distance from the camera to the sample. The pharmaceutical compounds can have an effective diameter of less than 10 μm[130] and image pixels can represent an area much larger than this and so multiple compounds are contained within a single image pixel. For example a pixel of tablet data will contain API and excipients which all add to create the outputted spectral signal. By using the calibration data PLS is able to ascertain which spectra are contained within the data and the relative contributions from each source. PLS is very common in nIR spectroscopy of known mixtures [131-134] where it is used to show the concentrations of the compounds being investigated.

Support vector machines (SVM) are a form of supervised learning model [135] which has been shown to be effective at separation out similar signals for the qualification and quantification of samples [136,137]. An SVM works by using a set of data which includes known values and their labels and then assigning each individual data point in the sample to one of those labels. So in an image containing a mixture of different compounds, if the pure spectra of these compounds are known, then each pixel in the image can be labelled as belonging to a specific compound. SVMs are also used in machine learning as a way of training the data inputted to identify new target genes [138] or drugs [139]. These can be used on their own to produce a reduced data set or as a precursor to further computational analysis such as artificial neural networks.

1.5 Hyperspectral dissolution imaging

The dissolution and release of active pharmaceutical ingredient from a tablet can be tracked using traditional UV spectroscopy and dissolution apparatus to gather information on the release of drugs from the system. These traditional dissolution experiments however, do not show the effects occurring within tablet only how much drug has been released into the solution [53]. New imaging approaches have been shown which image the dissolution using a UV system at a single wavelength to see the change in the tablet with the increased information that the spatial dimension can give. The disadvantage to these systems is that they can only track a small number of wavelengths simultaneously and so any complex drug release system which requires multiple wavelengths to track affectively may be unsuitable for this sort of analysis. An IR based system which tracks the release of the drug using a FT-IR imaging system has also been shown [54] which is able to track the release from an area on the surface of the tablet during the dissolution. The information on the release from the target sample during the experiment was good with spatial effects visible, however the total area imaged is small and there is no way to confirm that the data shown in this region is consistent throughout the tablet.

Hyperspectral imaging, especially using tuneable filters or spectrographs, has shown itself as a powerful tool in various different research areas including pharmaceutical analysis[140]. The point scan imaging systems have been in use in pharmaceutical analysis for some time where it is most commonly referred to as chemical imaging [44]. These new systems which use the tuneable filters and spectrographs for line scanning are much faster than the point scanning which leads to quicker data acquisition and the ability to investigate time resolved and dynamic systems such as chemical reactions or the dissolution of a tablet. The

line scan systems are also implemented in production lines where constant analysis of product quality is needed which is what HSI can provide.

It should be possible to combine the analytical power of hyperspectral imaging with dissolution testing to create a system able to track the changes in the whole tablet throughout the dissolution process. By combining HSI and dissolution testing it should be possible to show the release of APIs or excipients out of the tablet over time with spatial information on which areas release more efficiently. However by investigating the entire tablet and surrounding areas physical qualities such as the rate and manner of the tablet swelling and disintegration as well as the movement of tablet particles can all be investigated. This is a large volume of information which can be discovered using a single technique.

1.6 Aims and objectives

The aim of this work is to use a hyperspectral imaging system as a novel way to investigate the dissolution of known tablet systems and to show the physical changed which are occurring during the process using both the spatial information in the image and the spectral information at each pixel.

The key objectives are as follows:

- The investigation of the most efficient HSI systems for analysis of pharmaceutical compounds.

- The investigation of the physical process which occur during the dissolution of a tablet, specifically the expansion and disintegration of tablet material.

- To investigate different ways to use the HSI system to achieve the best information possible from the data received such as improvement of time resolution.

- To use multivariate statistical analysis on both individual spectra and the complete images to determine which is best able to show the changes occurring within the tablet over time

- To show the loss of API or excipient from the tablet as a function of time

Chapter 2: Instrumentation and theory

2.1 Hyperspectral imaging

Hyperspectral imaging (HSI) is a spectroscopic technique which enables the user to obtain an image with both spatial and spectral information [141]. There are three main ways to collect hyperspectral images which are single point[12], plane scan[134] and push broom scanning[36]. The systems used in this work all use plane scanning HSI, with the exact experimental setup of the camera changing between the different experimental chapters. However, the necessary component pieces remain the same throughout and they are the detector, the spectrograph, the stage and the light source. Figure 2-1 shows the experimental setup in a schematic diagram, the light source illuminates the target which is on a stage and the scattered photons from this target then enter into the spectrometer. A series of slits inside the spectrometer narrow down the incoming photons to a narrow beam which then enters into the diffraction grating, the different energy photons are split and then hit the detector which records the spectral information.

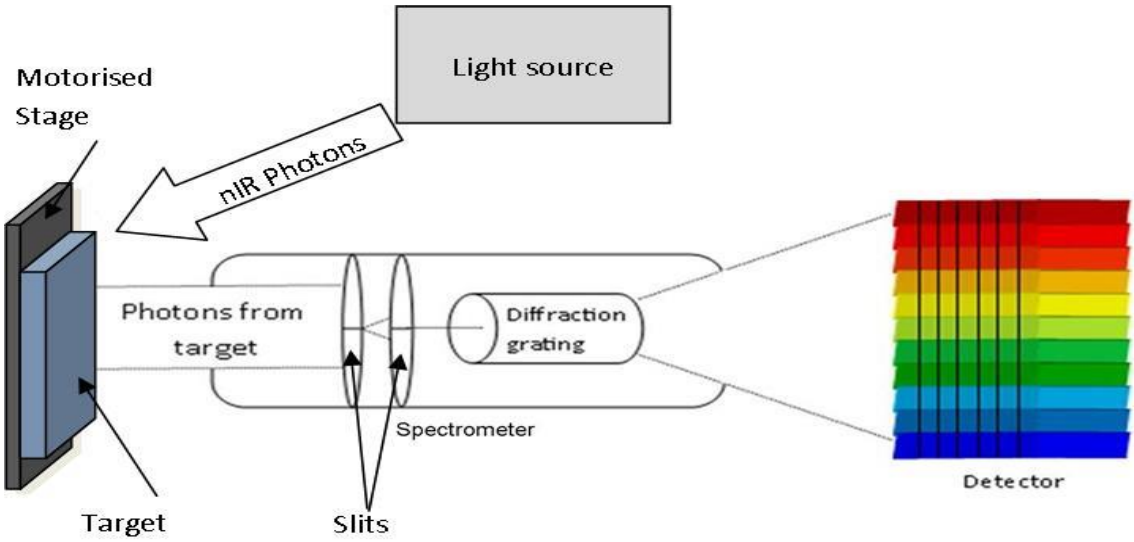


Figure 2-1 Schematic representation of the hyperspectral imaging setup

2.1.1 The detector

The detector is a device which stores and translates the electrical impulses received into spectral data. It is located behind the spectrometer either as a separate unit, in the case of the visible and near infrared (IR) systems, or enclosed with the spectrometer which is a

style implemented by the SWIR camera used in Chapter 6. The general method of action for any detector is the same with incoming photons hitting the detector and causing an electrical charge to be created by the excitation of an electron. This excitation is of the free electrons in the valence band electronic structure to the conduction band and the energy difference between this is known as the band gap. The valence band in is created by the overlapping of the highest occupied energy levels from all of the atoms in a metallic structure. As each atom has a similar but not identical energy level to the surrounding atoms a band of energies if formed and it is from this band that the electron is excited from to create an electrical charge. Each charge is counted and the magnitude of the total counts gives the relative intensity of the signal at that position in the detector. The detectors are attuned to the wavelength range that is being investigated as they are made with specific alloys which respond to photons within a narrow band of energies. By modifying the alloys used to create the detector it is possible to change the band gap and thus the sensitivity range of the detector. One such alloy is mercury cadmium telluride which is used in MCT detectors which are used in this work and is a composite of two alloys, mercury telluride and cadmium telluride. The cadmium telluride has a band gap of approximately 1.5 eV and mercury telluride is a semi metal and has a band gap of 0. By combining these it is possible to create a MCT detector with a band gap anywhere between 0 and 1.5 eV. The energy of a photon at a set wavelength is specific and so by manipulating the band gap you can manipulate which photons will cause excitation of electrons and thus be detected.

For this work three separate wavelength ranges were investigated and three different detectors were used. An electron-multiplying charge-coupled device (EMCCD) for the visible wavelengths of 350 – 860 nm, a InGaAs CCD for the very near infrared region (780 – 1860 nm) and an MCT detector for the near infrared (also defined as the short wave infrared) region of 970 – 2500 nm. The EMCCD has a detector made of silicon, the band gap of which allows for photons in the visible wavelengths to be detected. For the near infrared region the detector uses an alloy of indium gallium arsenide (InGaAs) which is effective up to the 1860 nm. An InGaAs system is insufficient with the short wave IR as this goes as far as 2500 nm and this is too small a band gap energy for an InGaAs system to sufficiently detect the photons. For this range an MCT detector was used which contains a detector made from a mercury cadmium telluride alloy.

Noise in a spectrum is often as a consequence of noise in the detector, the biggest cause of this is thermal. Even room temperature heating will give erroneous signals as it causes excitation of the electrons in the detector without any external irradiation of photons. To

minimise this noise the detectors are cooled with inbuilt cooling systems. Using cooling also reduces the dark current which is present in detectors at times when they are not exposed to incoming light. The pushing of electron charge through the detector can also cause some error, where every step adds a small degree of noise into the signal, although this is often much smaller than the thermal noise.

2.1.2 The spectrometer

A spectrometer is a device commonly used in a number of different spectroscopic techniques; its primary function is to separate out photons into different energies for detection. In hyperspectral imaging the image is built up from a series of line scans which are collated together by moving either target (or the camera) during the imaging. In a hyperspectral imaging system the spectrometer consists of three major components which are the objective lens, a series of slits and a diffraction grating or prism. The photons which enter the device, in these experiments, are scattered from the target surface. Figure 2-1 shows a schematic of how these pieces fit together and work. An objective lens is often used to focus the incoming beam of photons onto the detector, once the reflected light has entered the spectrometer it passes through a series of slits which reduce down the beam into a single narrow beam of photons. This line of photons then enters into a diffraction grating which causes the different energy photons to be separated out such that a spectrum of different energies hits the detector.

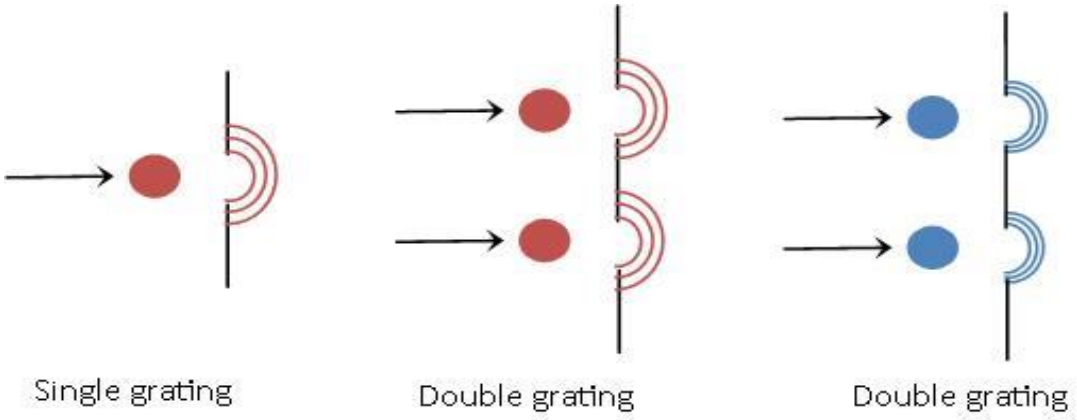


Figure 2-2 Diagram of how photons react with a diffraction grating, the different colours denote a photon of red light and one of blue

The diffraction grating contains a series of very narrow slits through which the photons of light may pass. Figure 2-2 is a diagram of the diffraction grating and shows how the photons react upon reaching the grating. As a photon (which for ease can be thought of

here as a particle) hits the diffraction grating it is diffracted outwards in all directions, this is known as a diffraction pattern and is shown in the single grating example. If two gratings are present then a series of diffraction patterns will occur in close proximity and when these diffraction patterns meet the waves will interact. If the waves are in phase then constructive interference will occur however, if the waves are not in phase destructive interference will occur. Thus only a small amount of the diffract pattern reaches the detector. The wavelength of the photon determines the size of the pattern, the larger the wavelength the larger the wave which passes through the grating which is also shown in Figure 2-2. A blue photon will have a wavelength which is shorter than a red photon which creates a diffraction pattern with tighter banding. The relative positions in space where these two diffraction patterns will react to create constructive interference is different due to the different wavelengths of the photons as they enter into the grating. The grating equation in Equation 2-1 gives the angle of diffraction θ_m at which the diffracted light will be at a maximum. The other terms in the equation are *a* which is the distance between the mid-point of two adjacent gratings, *m* is an integer which specifies the order of the various principal maxima and I is the wavelength.

$$\boldsymbol{a\,sin\theta_m = m\lambda}$$

Equation 2-1 The grating equation

For more in depth theory on diffraction and diffraction grating readers are advised to read Optics by Hecht[142]

As the exact wavelength range the light is split into depends on the diffraction gratings used each spectrometer can only be used for one set wavelength range. As with the detector three separate spectrographs were used for the three different wavelength ranges investigated. The visible used a Specim V8E, the near IR used a Specim V17E and the near IR used an ImSpector N25E spectrograph.

The spatial resolution of the pixels in the image are determined as a function of the distance from the camera to the target and whatever magnification optics are attached to the spectrometer. During these experiments the camera was positioned as close as possible to the target while still remaining in focus and no magnification was used which gave a pixel resolution of 219 x 219 μm^2.

2.1.3 The stage

The spatial image shows the physical representation of the target in a manner similar to a normal digital camera. However in plane scanning HSI the image is created one line at a time; these lines contain all of the spectral data for each pixel which are then collated together to from the full hyperspectral image as shown in Figure 2-3. The top image in the figure shows one line being captured on the detector, which then correlated to a single line in the spatial image. The bottom image shows how when multiple lines are collected the image begins to build up, with each new line of imaging adding another spatial line in the y dimension.

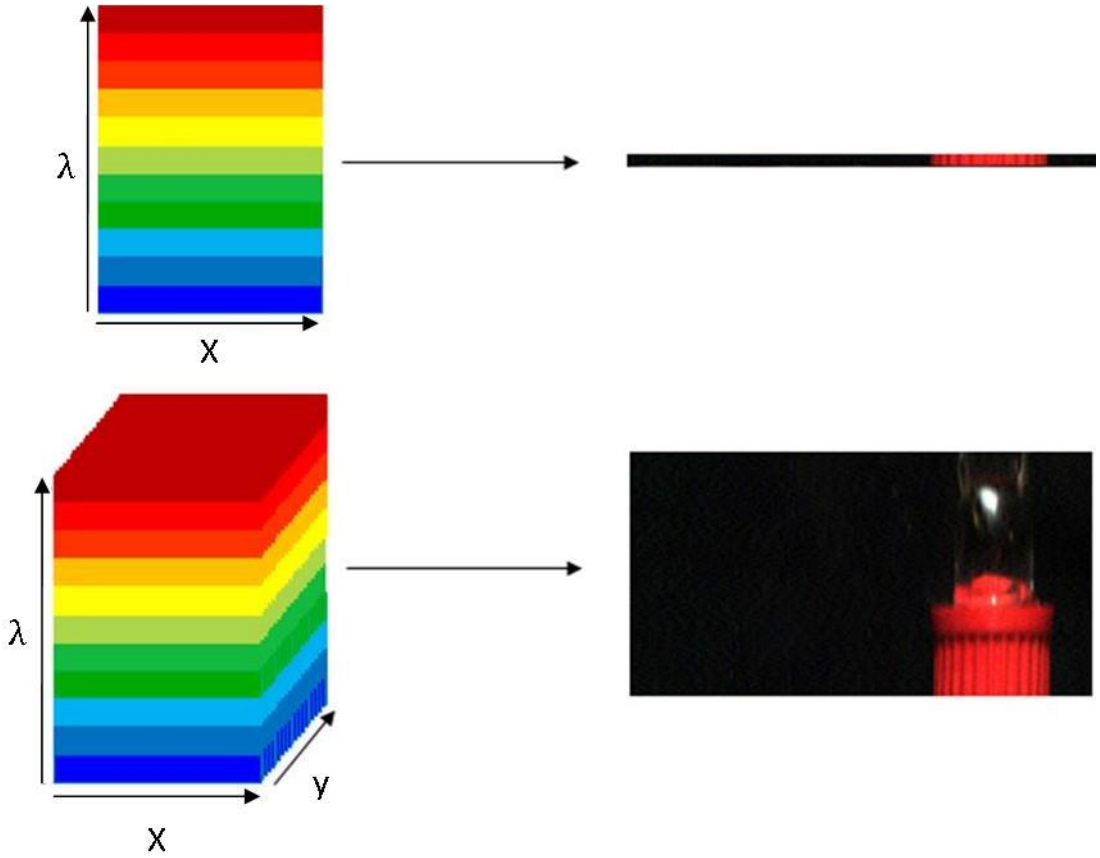


Figure 2-3 A diagram showing how HSI with a spectrograph builds up an image with the image being created one line at a time

Movement of either the camera or the sample is essential to build up a 2 dimensional spatial image of the target when imaging with a push broom system. Running the experiment without this movement the same spatial location would be imaged repeatedly which can be a powerful technique for monitoring a specific location but loses spatial information. The speed of the stage depends on the exposure time used in the imaging and the spatial resolution of the image pixels. If the speed of the stage is correct then pixels should be square, if the speed is too slow the pixels become elongated and if the speed if

too fast the pixels become short and data is lost. Once the exposure time and frame rate had been set (which is discussed in 2.1.4) the correct speed of the stage could calculated. The speed was determined by imaging a checkerboard pattern using the experimental setup for the experiment. If the speed of the camera was too fast then the squares in the pattern become truncated and the speed was slowed, and when the speed was too slow the squares were elongated into rectangles. By a manual iterative process the ideal speed was determined and incorporated into the experimental setup for the imaging as once calculated the speed is always correct as long as exposure time, frame rate and camera distance remain constant. A precise motorised stage was used to ensure the smooth, continuous motion of the sample. Incredibly precise distances can be scanned at a constant speed with no jerking of the sample. The same stage was used throughout all experiments, which was a Zolix KSA 11-200 s4N with an IMS Mdrive 17+ motor attached.

2.1.4 Exposure times

The exposure time is how long the detector receives photons to create the data in the image. Too long and the detector will saturate and be inaccurate and too short will cause weak spectral signals. The exposure time was determined using a piece of the PTFE calibration tile which was used to calculate reflectance values (section 2.3.2). By imaging this tile and inspecting the raw counts values the exposure time can be manipulated so that spectral signals were intense but no saturation of the signal would occur.

2.1.5 Light sources

The light source used in these experiments is important, if the light source is variable then inter-image variation may become a substantial source of variation within the data. The visible HSI system used an Armley 150W Halogen lamp however, this was later found to be unsuitable for imaging in the near IR with fast frame rates such as were used in Chapters 4, 5 and 6. Polytetrafluoroethylene (PTFE) is used as a standard in these experiments for 100% reflectance and is used to calculate the reflectance of the data, this process is shown in section 2.3.2. This material should have a value of 1.0 at all points throughout the image however, when imaged using the Armley lamp a pattern of variation was noted in the values of the reflectance. It was concluded that the AC power of the halogen lamp was causing the flicker in the reflectance intensity values in the image of the PTFE piece. A direct current halogen lamp was implemented for all experiments from Chapter 4 onwards to remove this inter image variation.

2.2 Pharmaceutical analysis

2.2.1 Vibrational spectroscopy

This information on vibrational spectroscopy is based on the information given in 'Spectroscopy of Polymers' by J. Koenig [143] and further details on the theory of vibrational spectroscopy can be found there. There are three separate and distinct techniques within vibrational spectroscopy; these are Raman, near Infrared (nIR) and mid Infrared (IR) spectroscopy. These three techniques are used to show different and distinct information about the chemical structure of a sample [144]. However, the cause of the absorption bands in each of these techniques is the same. Vibrational spectroscopy revolves around the energy changes within a compound due to the molecular vibrations of its atoms and bonds. The simplest form of molecular vibrations is explained by the diatomic oscillation model shown in Figure 2-4, in which two different atoms are vibrating back and forth along a bond.

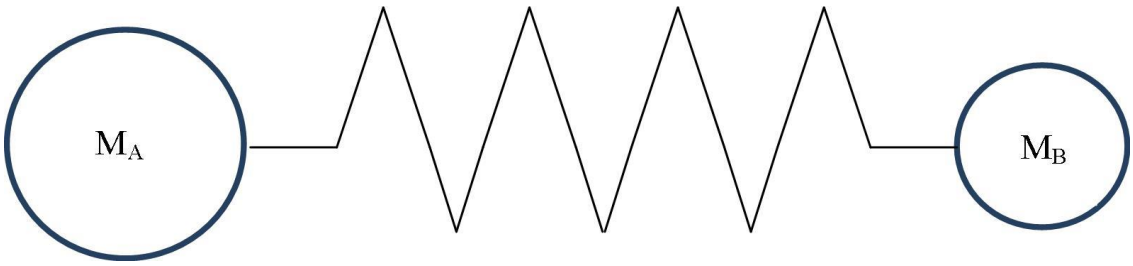


Figure 2-4 Schematic representation of the stretching of a simple heterogeneous diatomic bond

The vibrational frequency ν is a measure of the energy required to make a bond vibrate, moreover it can be calculated using the harmonic oscillator approximation in Equation 2-2

$$\boldsymbol{\nu} = \frac{\mathbf{1}}{\mathbf{2}\boldsymbol{\pi}} \sqrt{\frac{\boldsymbol{f}}{\boldsymbol{\mu}}}$$

Equation 2-2 The harmonic oscillator approximation

In this equation f is the force constant of the bond and is proportional to the length and valency of the bond and the electronegativity of the atoms in the bond. The reduced mass μ is used to account for the masses of the two atoms and is calculated using Equation 2-3

$$\mu = \frac{M_A M_B}{M_A + M_B}$$

Equation 2-3 Equation for calculating the reduced mass

These two factors show that the vibrational frequency of a bond is directly linked to the atoms contained within a vibrating bond and the strength of that bond. This means that changing the atoms incorporated in the bonds or changing the valency of the bond (single, double or triple bond) can have a significant impact on the value of ν. The environment around the bond will also affect the vibrational frequency of the bond, a carbonyl group attached to a benzene ring will have a different frequency to a ketone and a terminal amino group will have a different vibrational frequency to one in the centre of long carbon chain. The vibrational frequencies are also distinct and discrete, no two bonds have the same vibrational energy and the energy levels are quantised so the energy required for a carbonyl asymmetrical stretch is always the same.

It is this feature of the molecular vibration of bonds which allows vibrational spectroscopy to be used to determine what bonds are within the target molecule. It is quite common with vibrational spectroscopy that the bonds detected are not a simple diatomic bond but a polyatomic bond which contains a number of different atoms bonded to a single central atom. These bonds vibrate in unison to create more complex vibrational structure such as the bonds in a methyl group which has 4 bonds which are all be involved in vibration.

A vibrating bond containing N atoms can exhibit 3N-6 (3N-5 in a linear system) different molecular vibrations (also called vibrational states) where the minus 6 vibrations are those related to translation and rotation of the bond in the x, y and z spatial co-ordinates. Each unique bond vibration is known as a vibrational mode, for example a H_2O molecule has three vibrational modes (3x3-6). These modes are scissoring, antisymmetrical stretching and symmetrical stretching vibrational modes and the movement of the bond is detailed in Figure 2-5.

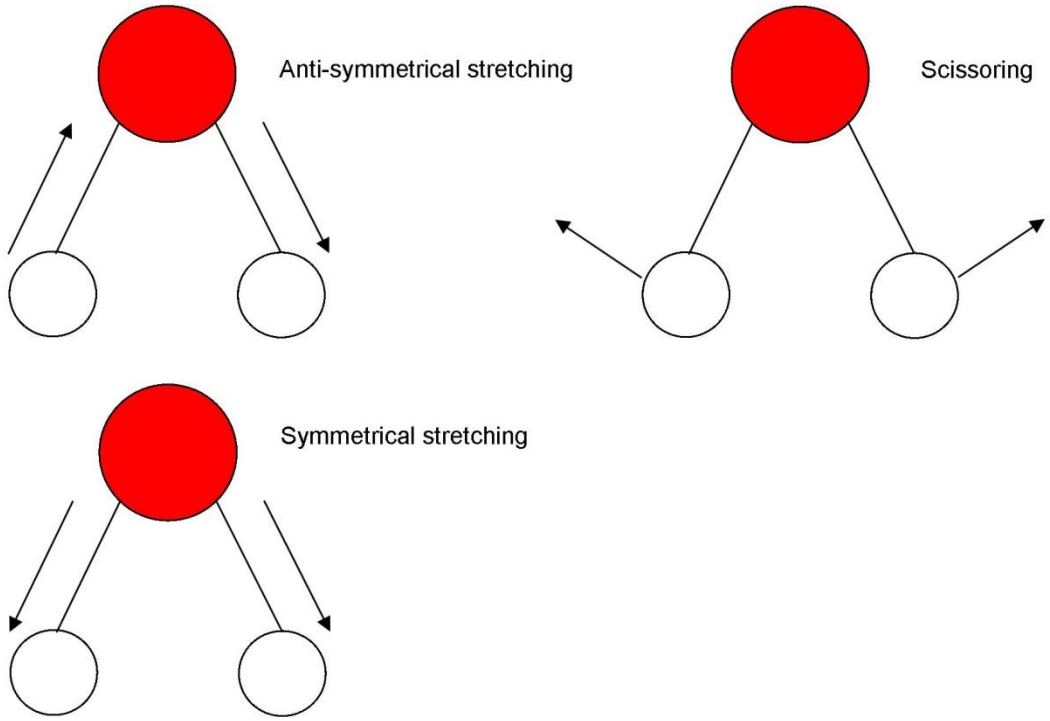


Figure 2-5 The three different vibrational modes shown by a water molecule

Each possible vibration within a compound will occur when a photon of the correct energy hits the molecule however, it should be noted that there are always some vibrations from populated vibrational states which get there energy from background thermal radiation. Molecules will always exhibit some vibrational modes unless they are close 0°K and have no access to background thermal energy. The energy absorbed from the IR photons allows the vibrational energy level of the bonds to increase, causing the corresponding vibration to occur. Each individual molecular vibration has its own distinct energy and so each bond can be distinguished by the relative degree of absorption at these energies.

The vibrational frequency is related to the vibrational energy level, the higher the energy level the higher the frequency of the vibration. Again it is important to note that the energy levels are different and distinct between the vibrational modes exhibited in a single bond and also different to any other bond. The energy required for the scissoring of a water molecule is not the same as the symmetrical stretching of the same water molecule and the symmetrical stretching of water is different than the symmetrical stretching of hydrogen sulphide, which has a similar structure. The vibrational frequency is affected by the atoms involved, the valency of the bond and also potentially the environment close to the vibrating bond. Therefore every functional group requires a unique and distinct quantity of energy and every vibration within a molecule occurs at a different distinct energy.

The photons required to cause each vibrational mode are of a set wavelength as the energy, frequency and wavelength are all related by the Planck relation shown in Equation 2-4:

$$E = \frac{hc}{\lambda}$$

Equation 2-4 The Planck relation

Where E is the energy of a photon, *h* is Planck's constant, *c* is the speed of light and λ is the wavelength of the photon. The frequency of the photon is given by *c*/ λ.

This means that vibrational spectroscopy is able to not only show what types of bond exist within a structure but also the environment of the bonds such as adjoining bonds which can have a notable effect on the vibration frequency and the specific types of vibration that are occurring. It is this ability of vibrational spectroscopy which has resulted in it becoming one of the most common techniques for chemical identification and structure elucidation.

Near and mid IR spectra are both measured as a function of the absorption of photons which either pass through or are reflected off the surface of the samples. The photons are streamed from a source to a detector via the sample and the difference in the starting intensity and the intensity which reaches the detector gives the absorption (or transmission) of the target sample. These absorptions are quantised as the energy of any bond vibration is specific and discrete. Raman spectroscopy is uses high energy lasers to cause excitation and relaxation of the vibrational quantum states. The relaxation of the bond emits a photon with a different energy to the incident photons if the vibration is a Raman active vibration. Raman is a scattering technique with the emission of photons being random in all direction and the different types of Raman scattering are shown in section 2.2.1.3. The data collected in Raman spectroscopy is not a measure of the wavelength at which a vibration occurred but rather the energy difference between the emitted photon and the incident photon.

Mid IR spectroscopy details the absorbance of photons from ~200cm^{-1} to 4000cm^{-1} while the nIR photons are found in the range of 12400 to 4000 cm^{-1}. Raman is often in the range of ~100cm^{-1} to 4000cm^{-1} but not as absorbed photons but the energy difference between incident and emitted photons. The intensity of the techniques can depend on a number of factors such as the surface roughness, quantity of sample and the physical form of the target substance (liquid etc.). Infrared spectroscopy is often quite intense however, Raman and nIR are both much weaker. In Raman scattering the absorption of a photon is more

common than the scattering of that photon as the photon needs to interact with the electrons in a bond in the correct orientation in space for the scattering to occur. This reduces the efficiency of this technique so that 1000 photons will not create 1000 scattered photons to detect but a much lower number, to counter this issue a huge number of photons are used in Raman spectroscopy and the power of the laser can sometimes damage the sample if used for too long. Raman spectroscopy is detecting the scattered photons which are emitted from the relaxation of the quantum energy level of the bond. This process is random and the directionality of the photons cannot be controlled, which results in a much weaker (but still sufficiently strong) spectral signal which is also a reason that more photons are used than may be necessary in IR of the same compound. The molecular vibrations involved in nIR spectroscopy require a change in quantum energy level of greater than +1, these vibration are much less common that those in the IR region. As the vibrations are reduced in number so too is their relative intensity, but the signal is still sufficiently strong accurate chemical analysis of the target.

2.2.1.1 IR spectroscopy

Infrared spectroscopy (IR) is the most common form of vibrational spectroscopy, largely due to its speed, ease to analyse and none destructive nature. The range used in a standard IR spectrometer is 2500 – 100000 nm and this portion of the electromagnetic spectrum is often referred to as the mid IR.

For a bond to be visible and significant in an IR spectrum it must be 'IR active' which means that the vibrational of the bond must obey certain selection rules. The selection rules are that the vibrational energy level must increase by only 1 and that the bond vibration causes a change in its dipole moment. An energy level diagram is shown in Figure 2-6 which details allowed energy transitions in IR spectroscopy. The absorption of a photon must causes an increase in the vibrational energy level ν of 1 but the starting energy level is not a factor in the selection rule. The excitation from higher energy levels is less common however, as the higher the energy level the lower the population with most bond vibrations being in the lowest vibrational energy level. The dipole moment of a bond is the relative distribution of electron density across a bond moreover when one atom in a bond has a greater electronegativity, electron affinity, then that atom will contain the larger portion of the electron density. This causes a dipole across the bond with a positive and negative end and the degree of the polarity is the dipole moment. Vibration of a polar bond will cause the dipole moment to change as the poles are moved closer together and further

away and this change is necessary for a bond to be 'IR active'. The stretching of a bond such as in the C-C chain of a polymer backbone will not show up in an IR spectra, unless one of the carbons has a unique group attached which affects the dipole moment (a C=O for example).

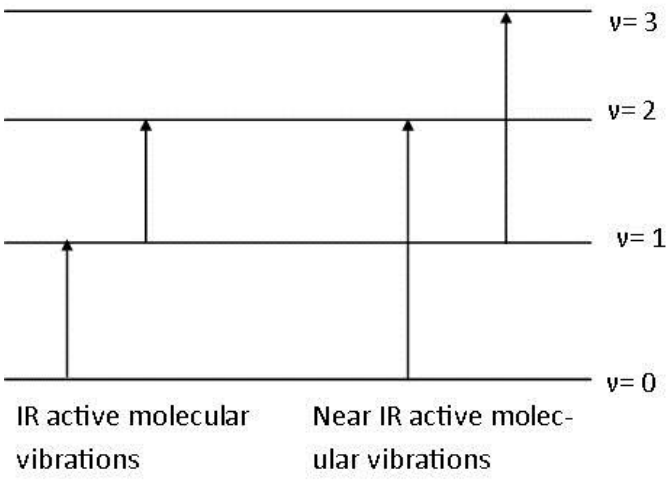


Figure 2-6 A vibrational energy level diagram showing the IR and near IR active energy transitions

2.2.1.2 Near IR

The near infrared region (nIR) of the electromagnetic spectrum is approximately 850 – 2500 nm and contains very different but analogous information to the IR region. The peaks in nIR spectroscopy come from two sources, combinations of mid IR vibrational and from an increase in vibrational energy level (ν) of +2 or more which is shown in Figure 2-6. As with the IR vibrational excitations the initial energy level is not important to if a vibration occurs however, they are much less common and are referred to as overtone vibrations.

IR spectroscopy and nIR spectroscopy both show the vibrations of polar molecules where a change in the dipole moment of the bond occurs during the vibration. This means that the bonds seen in the nIR are from heterogeneous bonds with X-H being the most commonly seen bonds in the nIR. The peaks seen in the near IR are associated with peaks in the IR spectrum a carbonyl stretch in the IR has a complimentary peak in the nIR from the overtone vibration.

The overtone and combination vibrational bands seen in the nIR are broader and weaker than there corresponding IR bands. These bands can be 10-100 times weaker than in their

corresponding IR bands [127] and while most IR absorptions are narrow (with the exception of hydrogen bonding O-H bonds) the nIR peaks can be much broader, up to 10 - 100 of nm wide or are seen as sharper peaks within a much broader peak. This is especially true when water is within the sample as this has strong absorption bands. The broad nature of these peaks means that it is very common procedure to use more advanced analytical technique than are normally applied in analysis of mid IR data. Multivariate analysis is the style most often applied to the data using techniques such as partial least squares, principal component analysis and artificial neural networks to de-convolute the data and extract valuable information.

The path length of nIR photons through the target is not a simple penetration and reflection but instead the photons show multiple deflections before being reflected out of the substrate. The penetration of the nIR photons is greater, typically, than that of the mid IR photons. This is due to the reduce absorption coefficient of near IR photons and also that the scattering of the photons is proportional to λ^{-4} thus the mid IR undergoes considerably fewer scattering events per unit length than the near IR, though will be scattered from a greater depth in the sample. Traditionally IR and nIR spectrometers used a monochromator along with a multichannel detector to build up a spectrum. This process is time consuming as it requires each wavenumber to be scanned individually however, most modern IR spectrometers do not use this kind of setup and instead use a Fourier transform system to allow for rapid data acquisition. In a Fourier transform setup (FTIR) an interferometer is used to allow the collection of all the data at once giving an interferogram. This interferogram is then put through a Fourier transform and the spectrum of the sample is output.

The convention for both IR and nIR spectra is to measure the transmission of the beam from the source to the detector. This gives a spectrum where the 'peaks' are decreases in the transmission energy of the beam due to the absorbance by the chemical bonds within the sample. The x axis in an IR spectrum is in the units of wavenumber which are $1/\lambda$ where λ is the wavelength of the IR beam and the units of wavenumber are cm^{-1}. The wavelength range used in a typical mid-IR spectrometer is 2500 - 100000 nm, however, the use of wavenumbers reduces this to 100 - 4000 cm^{-1}. A wavenumber is a measure of energy, Equation 2-4 shows that the relationship between energy and wavelength (λ) is proportional to λ^{-1} therefore a change in the value of the wavenumber (λ^{-1}) is a direct change to the energy of the related photon. The difference between vibrational modes can also be expressed as a form of energy change using wavenumber. There is no clear

nomenclature for the near IR as to the units of the energy specific to the peaks, it is acceptable to use either wavelength or wavenumber although the wavelength is more commonly seen. In this work the wavelength is used as it is easiest with the hyperspectral systems being calibrated to use wavelength.

2.2.1.3 Raman spectroscopy

Raman spectroscopy is similar to infrared spectroscopy in that it is also sensitive to vibrational changes within a molecule. Raman spectroscopy however, deals in a different type of excitation than the near and mid IR spectroscopy. Infrared spectroscopy deals with a simple increase in the vibrational state of +1 (or more in the case of nIR) but in Raman spectroscopy there is an excitation and emission back to a vibration level of ±1 from the initial vibrational energy level. The vibration must also be accompanied by a change in the polarisation of the bond for it to be Raman active. These vibrations are most common with non-polar groups which are why Raman spectroscopy is very good for identification of the carbon-carbon backbones as this is a non-polar structure which cannot be detected easily in the IR.

The quantum model of the excitation exhibited during the Raman scattering is shown in Figure 2-7 which depicts the way in which the incoming photons affect the target molecule. A laser light is shone onto the sample which causes the excitation of the vibrational state to a virtual energy state; the vibrational state quickly relaxes to the lower energy and release a photon in the process. This photon is the particle detected in Raman spectroscopy which is why the technique is a disperse scattering technique and not a tightly controlled absorption such as is the case with the IR. There are three types of scattering observed from this process which are shown in Figure 2-7, they are Rayleigh scattering, Stokes Raman scattering and anti-Stoke Raman scattering. Rayleigh scattering involves the excitation of a bond to a virtual energy (excited) state which then relaxes to the same vibrational energy level causing no change in overall vibrational energy level. Rayleigh scattering is the most common of the scattering processes created in Raman spectroscopy. An incoming photon can also excite the vibrational state of a bond which then relaxes to a vibrational energy level which is higher than original ground state of the bond and this kind of energy change is known as Stokes Raman scattering. Anti-stokes Raman scattering is the third and least common kind of scattering process seen in Raman spectroscopy. In anti-Stokes scattering a bond absorbs the energy of a photon causing it from go from an excited ground state to a virtual energy level. The bond then relaxes by expelling a photon

and the vibrational level lowers from the virtual excited state back down to the ground state in a lower energy level than it originated at before the initial Raman excitation. Raman scattering requires a change in the ground state energy level of ±1, this is the quantum selection rule for Raman spectroscopy.

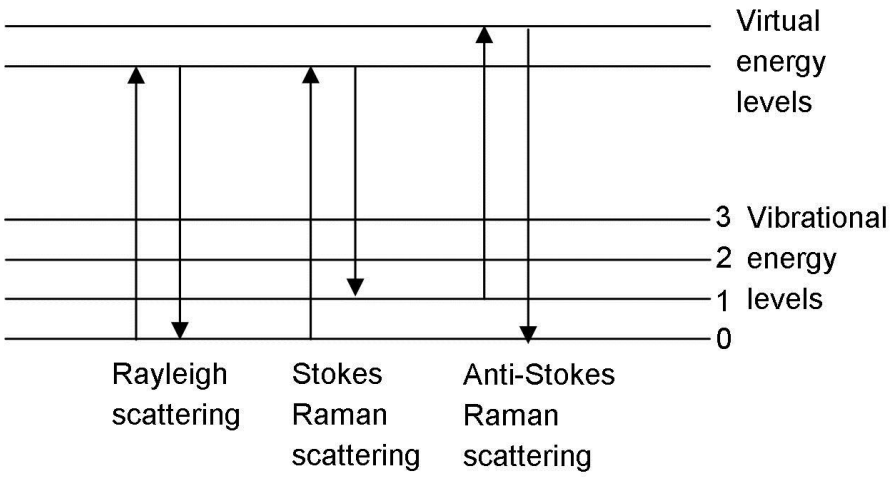


Figure 2-7 A quantum energy level diagram which shows the three different possible scattering which can occur during Raman spectroscopy.

2.2.2 Dissolution testing and apparatus

Dissolution testing is a way of detecting and following the breakdown and release of tablets by using spectrometry to track spectral changes in the system over time, leading to a quantification of the release from the target sample [52]. Dissolution testing apparatus vary in their exact setup but must contain a few key components, a vessel which holds 900 ml of liquid, often hydrochloric acid or simulated stomach/intestinal acid and somewhere for the tablet to be placed so that it is submerged in the water. Some way to keep the water moving to facilitate the tablets dissolution and a UV spectrometer which constantly takes reading of the UV absorption of the liquid in the system are also necessary components. The exact apparatus and how they are assembled can be varied and a number of different 'allowed' dissolution tests are contained within the United States Pharmacopeia (USP) [48]. The protocols contained within the USP are rigid and must be adhered to when running pharmaceutical analysis of any drugs. As the tablet swells and eventually begins to break down there is a release of the API from within the tablet. Chemicals become dissolved into the solution and the UV spectrometer detects a steadily increasing absorption of UV photons by these compounds. In this way the UV spectrometer can track the release from the tablet and once the release has finished a rate of release can be calculated.

There are a number of other factors such as the flow rate, pH concentration, buffer composition, temperature and stirring rate which must also be considered when setting up a dissolution experiment. Many of these factors are often controlled by the experimental protocol, such as stirring rate and temperature but others such as flow rate and buffer composition are normally altered for the specifics of what is being dissolved. If a tablet is designed to release in the intestine then the pH of the buffer solution needs to be much higher, not the pH1 or 2 that might be used for absorption in the stomach.

The dissolution apparatus used briefly in this work was a USP type IV vessel which was chosen as it was the system which best allows for imaging of the tablet during the experiment if this was required. In the other USP apparatus the tablet is suspending in a large volume (900 ml) of aqueous solution which would cause severe imaging artefacts due to the refractive nature of the solution. In the type IV apparatus however, the tablet is placed inside a flow cell which contains only a portion of the total volume of solution thus minimising these effects. A schematic of this vessel is shown in Figure 2-8. The dissolution apparatus contains a flow cell which has buffer solution pumped through constantly during the experiment. The buffer enters the vessel from the bottom and passes through a series of glass beads, these beads help to remove the pulsed nature of the flow from a peristaltic pump and promote a constant steady flow. The buffer then passes over the tablet and goes back into the buffer reservoir. A second flow system is also used which pumps the solution in the buffer reservoir through a UV spectrometer which is set to take readings of the solution as often as possible.

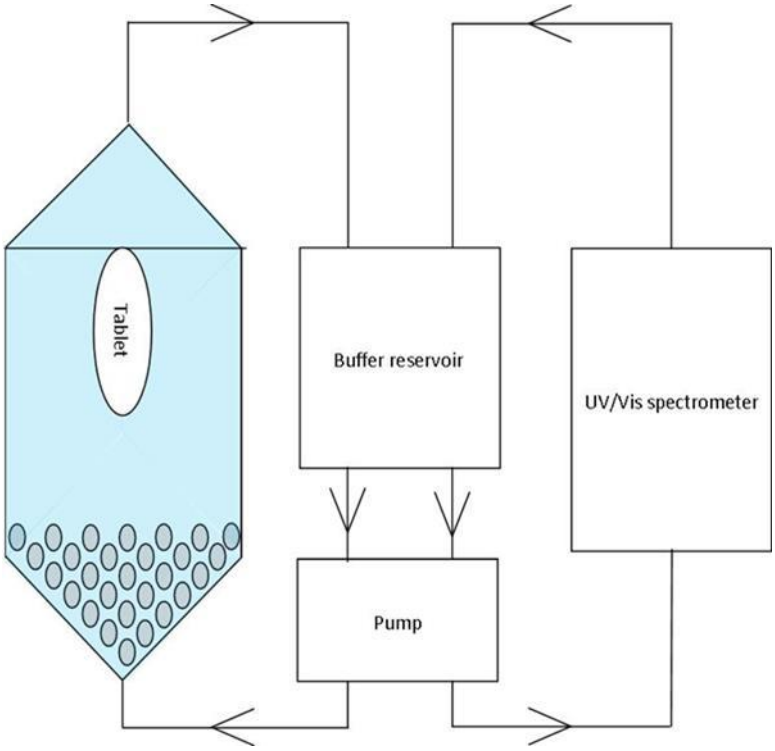


Figure 2-8 a schematic representation of a USP type IV dissolution apparatus

The exact protocol used in the hyperspectral imaging changes for each experiment and is explained in detail in the individual chapter however, the flow cell used doesn't change. A picture of the flow cell is shown in Figure 2-9; the cell is a modified 10 mm quartz UV cell with a screw cap and in and out flow ports for the tubing. The picture was taken using a digital camera is the same relative position as the HSI camera which is placed above the flow cell during these experiments. This system uses a flat vessel to minimise imaging problems caused by the imaging of, or through, curved surfaces and the short path length is also advantageous as it lowers signal loss or distortion from photons passing through the solution. The flow of solution through this cell is achieved by using a peristaltic pump with a flow rate of 10 ml/min for all experiments unless otherwise stated. The peristaltic pump created a flow from the solution from a central reservoir through the flow cell and back to the reservoir. The tubing for this can be seen at the bottom of the image with the inflow at the right hand side of the image.

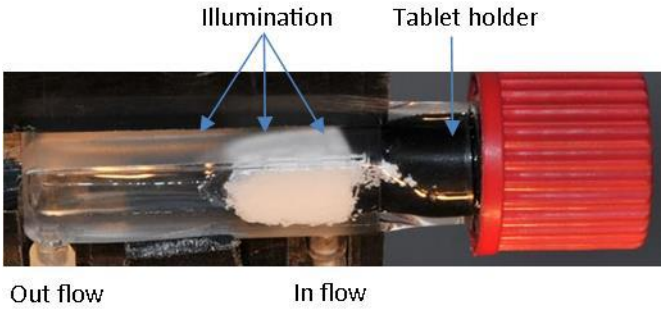


Figure 2-9 A picture of the flow cell used in the hyperspectral dissolution experiments

2.3 Data calibration and correction

2.3.1 Data calibration

The spectrograph was recalibrated before use if the system has been moved or disturbed in any way so that the correct wavelengths can be attributed to the correct pixel locations. This is achieved by using a line scan to image the emissions from a mercury lamp, these emission are tightly defined and have known wavelengths [145].

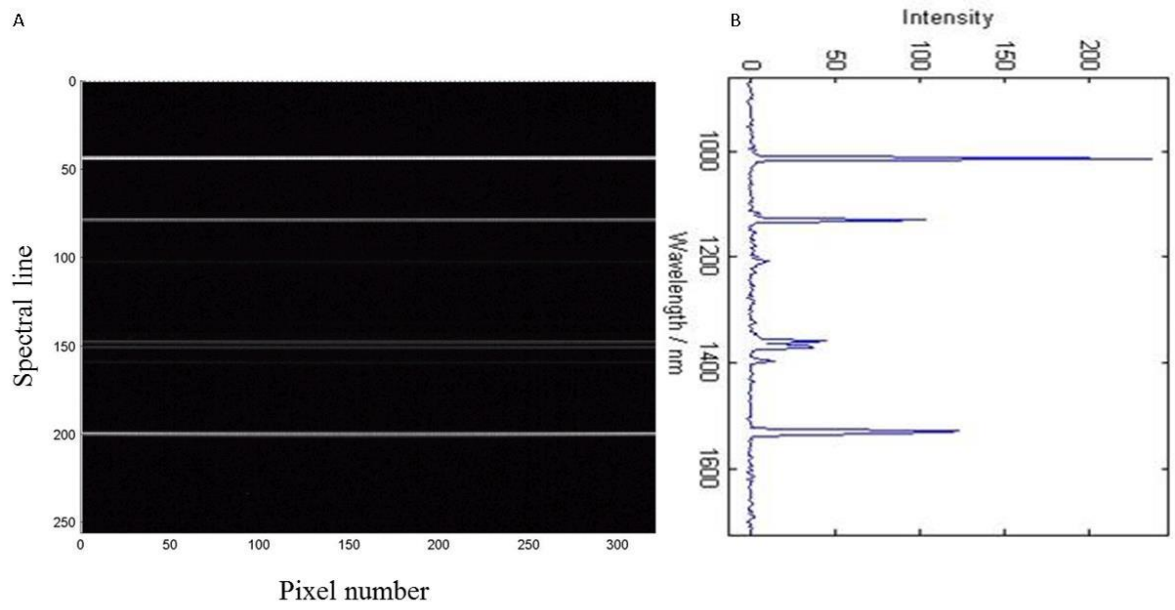


Figure 2-10 The calibration mercury spectrum with A) the line scan data from a mercury lamp and B) the spectra from one pixel of the line scan

Figure 2-10 A shows a single frame which was captured with the mercury emission used for the calibration. The x axis is the spatial location across the image and the spectral data

is shown in the y axis. There are 5 clear and distinct lines in this image, these are assigned to known Hg emissions so that the exact wavelengths of the spectral pixels can be determined. By knowing the location of two or more well defined, intense peaks it is possible to calculate which spectral pixels are which wavelengths. Figure 2-10B is the spectrum from a single pixel in the image of Figure 2-10A and shows the spectrum of the mercury emissions. Viewing this spectrum shows which are the intense spectral signals which give the bright intensities in Figure 2-10A and also can be used to check that the calibration was performed correctly by matching other spectral peaks with known mercury emissions.

2.3.2 Reflectance calculations

The data captured in its raw format is a function of the counts of detected photons at each wavelength at each pixel in the image. Any variations in lighting throughout the sample can cause significant variations with the raw data and this is problematic. To correct for this the data is transformed into reflectance values by calculating a 100% reflectance value. This is created using a piece of white PTFE tile which will completely reflect all wavelengths of light evenly from the visible to the near infrared.

Once the image is captured a set of calibration images are created using this white PTFE tile for the 100% reflectance. A 0% reflectance image is also recorded by switching off the lights and placing the lens cap onto the camera. These two values are used to convert the raw data into the reflectance of the radiation from the sample.

The calculation for determining the reflectance of any point is shown in Equation 2-5 below

$$R_{s\lambda} = \frac{I_{s\lambda} - D_{\lambda}}{I_{o\lambda} - D_{\lambda}}$$

Equation 2-5 The Reflectance correction equation

Where $I_{0\lambda}$ and $I_{s\,\lambda}$ are the intensity (in counts) of the 100% reflectance standard and the target pixel respectively at a set wavelength λ. D is the intensity of the dark or 0% reflectance standard and $R_{s\,\lambda}$ is the reflectance value for any specific pixel in the image at a set wavelength.

Any lighting deviations across the image should be removed with this procedure as the white tile will have the reflectance at that point increase by the same factor as the sample, thus the data will scale correctly.

2.4 Reagents

2.4.1 Chemicals

All chemicals and solvents were purchased from Sigma Aldrich unless otherwise stated specifically.

Chemicals used were:

- Hydrochloric acid, 1M
- Acetone
- Polyethylene Glycol
- Methoxypolyethylene glycol
- ε-Capralactone
- Tin(2) ethylhexanoate
- Toluene
- Dichloromethane
- Hexane

2.4.2 Tablets

2.4.2.1 Panadol Actifast

Ingredients: Paracetamol (500mg), Sodium Bicarbonate, Starch Pre-gelatinised, Povidone, Maize Starch, Potassium Sorbate (E202), Microcrystalline Cellulose, Magnesium Stearate, Carnauba Wax, Titanium Dioxide (E121), Polydextrose, Hypromellose, Glycerol Triacetate and Polyethylene Glycol.
Produced by: GlaxoSmithKline
Expiry Date: 02/2012
Batch no: 100578A

2.4.2.2 Panadol Advance

Ingredients: Paracetamol (500mg), Pregleatanised Starch, Calcium Carbonate, Alginic Acid, Crospovidone, Povidone, Magnesium Stearate, Colloidal Anhydrous Silica and Sodium Ethyl (E215), Sodium Methyl (E219) and Sodium Propyl (E217) Parahydroxybenzoates.
Produced by: GlaxoSmithKline
Expiry Date: 05/2013
Batch no: 100654A

2.4.2.3 Panadol Extra

Ingredients: Paracetamol (500 mg), Caffeine (65 mg), Starch pre-gelatinised, Maize Starch, Polyvinyl Pyrrolidone, Potassium Sorbate (E202), Purified Talc, Stearic Acid, Croscarmellose Sodium, Hypermellose and Glycerol Triacetin.
Produced by: GlaxoSmithKline
Expiry Date: 05/2015
Batch no: 100696A

2.4.2.4 Tesco Paracetmol

Ingredients: Paracetamol (500 mg), Pre-gelatinised Maize Starch, Sodium Metabisulphate (E223) and Magnesium Stearate.
Produced by: Galpharm Internation Ltd.
Expiry Date: 04/2013
Batch no: PAK52E01

2.4.2.5 Tesco Paracetamol Extra

Ingredients: Paracetamol (500 mg), Caffeine (65 mg), Starch Pre-gelatinised, Povidone K-30, Talc, Stearic Acid, Magnesium Stearate.
Produced by: Aspar pharmaceuticals Ltd
Expiry Date: 06/2014
Batch no: 10409

2.4.2.6 Boots Paracetamol Extra

Ingredients: Paracetamol (500 mg), Caffeine (65 mg), Maize Starch, Methycellulose Povidone, Purified Water, Talc, Calcium Stearate, Hypromellose and Polyethylene Glycol.
Produced by: Wafton laboratories Limited
Expiry Date: 06/2013
Batch no: 56108A

2.5 Statistics

2.5.1 Principal component analysis

Principal component analysis (PCA) is a form of multivariate statistical analysis which is used to show grouping in datasets with a large number of variables which was first invented in 1901 by Karl Pearson [114], although the modern version of PCA was first developed by Hotelling [112,113].

If you have an experiment with two or three variable it is easy to plot the data and to see groups or trends appearing in the plot. However, once the number of variable begins to escalate it becomes more difficult to accurately investigate the data in a simple graphical manner without creating a multitude of plots. PCA is a way to take data sets like these and to quickly analyse the data to discover these same trends and groupings as can be seen in the basic plots [119,146]. PCA plots the data in n-dimensional space so where n is the number of variables in the dataset. A vector is created which passes through the greatest variation in this data and this vector is known as the first principal component. Each subsequent component is created to account for the next highest variation in the data on a plane, orthogonal to the previous components. Figure 2-11 shows this idea in a simple plot where the first component is the blue line which cuts through the data so the greatest possible degree of variation is transect by this vector. The second component must be orthogonal to the first and must account for the second largest variation in the complete data set and in Figure 2-11 this is shown as the green line. When two vectors are orthogonal then there is no relation or interaction between these vectors, in terms of principal components this means that the each components is independent of all others.

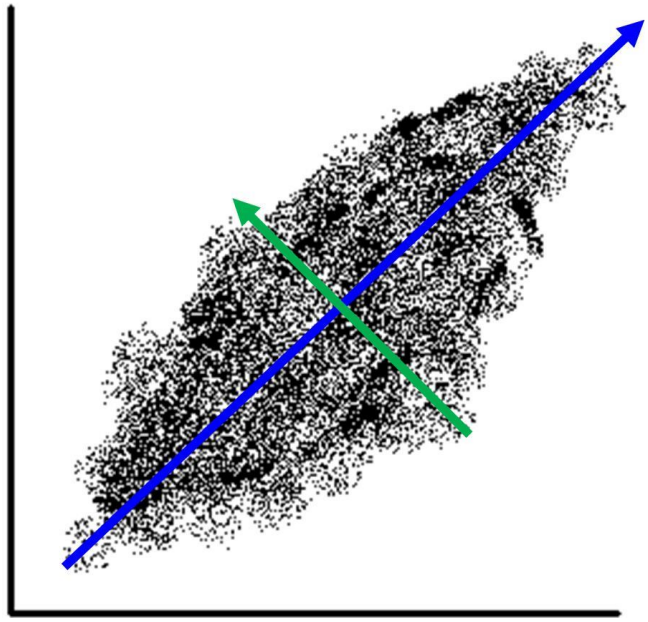


Figure 2-11 A diagram showing the how the principal components are defined. The blue line is the first principal component and the green line denotes the second principal component

The principal component is also known as the eigenvector and the degree of the total variation described by a eigenvector is the eigenvalue. The principal component on its own is of little use as it only shows variation in a non-real space and so the data must be transformed to a value which can be used to show the groupings and trends which are of importance. Two new sets of values are created to show the information within the principal components and they are known as the loadings and the scores. The loadings show to what degree each variable effect a principal component and thus which variable are responsible for the variation which the component describes. The scores detail how much of the variation described in the loadings can be attributed to the individual samples analysed and by plotting theses scores as a 2 or 3 dimensional bi-plot then grouping of scores should occur, these groupings are samples which have a similarity in the variable described in the loadings.

It is often necessary to apply some form of pre-processing to the data to prevent any bias occurring in the results. This can happen when the data contains within it a bias which is not due to a real difference between the data. This can be caused by effects such as lighting variation from the lamps or emission spectra which are based on a random event such as the angle of emission being consistent with the detector location. Any data like this which is not pre-processed in some way will lead to poor results and potential the PCA failing to discover true variations, such as a peak shift, which are masked by these ‘false’ variations in the data. To remove the bias in the data it is normalised in some way, there are a number

of different approaches to normalisation and they are tailored to remove bias in specific ways. One of the most effective with HSI is mean centring, a process in which the mean spectrum is subtracted from each sample and allows the PCA to better show variations from the mean along. A second pre-processing is normalisation by total intensity which is a way to remove topographical changes in the sample surface by subtracting each sample value by the total sample intensity. Normalisation by total peak intensity was used in Chapter 3 to reduce noise and improves spectral definition while also overlapping spectra, it was also used as it is effective with both near IR and visible spectra. In later chapters mean centering is used to improve spectral definition, this is done by subtracting the mean value for an individual spectrum from each wavelength in that spectrum. The resulting spectra then have a mean of 0 which reduces variations such as intensity change from variability in light intensity across or between samples. The second derivative spectra are used in Chapter 6 to improve the results of PCA by highlighting small changes in the gradient of the peaks in the spectra. This is a common approach to PCA of nIR spectra where the broad peaks can mask small spectral changes.

2.6 Software

The data processing of this work is often difficult and involves complex mathematical manipulation which is not possible using most common software. A number of specialist programs were used throughout the data acquisition, processing and analysing.

2.6.1 SpectraSENS

SpectraSENS® is software developed by Gilden Photonics for use with their hyperspectral imaging systems. This software has two main functions, the first is the setup and running of the hyperspectral camera. The software can be used to set a number of factors such as the stage speed, the exposure time, the pixel binning and distance imaged. The second function of the software is to perform the corrections onto the data to generate the reflectance values.

2.6.2 ENVI

Exelisvis's ENVI® software is widely used for investigating hyperspectral data. It allows the user to open the hypercube, make false colour images and even perform some statistical analysis. While ENVI can perform analysis like principal component analysis or

support vector machines there are restrictions to what can be analysed due to the constraints of the software. This software was heavily used for the first few experiments where the results required less interpretation but was phased out towards the ending chapters in favour of MATLAB.

2.6.3 MATLAB

The MATLAB® software is published by Mathworks® and is used to create code for performing a wide range of mathematical functions on data sets. The statistical toolbox and basic inbuilt MATLAB functions were employed in this work. There are very few restriction on what data analysis can be performed and it is easy to reduce down or concatenate data sets for improved analysis. This software was used heavily in the last two chapters due to the ability to do the majority of the analysis in the one program. The software is not without its issues however, allocation of RAM often meant that when the data sets where too large then no analysis could be performed. The meant that a powerful PC was needed for some analysis while for others a subset of the total data was used to reduce down the memory constraints.

2.6.4 SIMCA – P 11

Principal component analysis can be performed by a number of different statistical programs, once such program is SIMCA – P 11®. This was used as it is able to produce clear scores plots which can be easily tailored, while also giving a lot of customisation to how the analysis is performed. One of the major disadvantages to this software is that it cannot be used to generate score images easily, only biplots. It is used for predominantly for the first few chapters alongside ENVI.

2.6.5 Origin Pro

Origin Pro® is a common statistics and graphing package which is used for its competent handling of large data sets and the ability to apply a range of statistical or graphical options onto the data.

Chapter 3: A comparative study investigating two different hyperspectral imaging systems to determine which wavelength range, the visible or the near IR, is better for the differentiation of polymer sample

3.1 Introduction

The aims of this chapter are to investigate the use of hyperspectral imaging as a novel approach to distinguish between similar polymers samples and to investigate which wavelength range is best suited to this task. The polymers samples being investigated are simple monomeric polymers and block co-polymers which contain the monomeric polymer. A visible HSI system and a near infrared (nIR) HSI system is used to differentiate between polymer samples to show which is the most efficient system. The monomeric polymers used for this analysis are biocompatible and are a valid comparison to pharmaceutical materials. The wavelength range best able to distinguish between these polymers samples should be the wavelength range most effective for pharmaceutical analysis.

3.1.1 Overview of the technique

Hyperspectral imaging is restricted to specific wavelength range by the internal optics used to capture the spectrum. The two most common wavelength ranges used are the visible spectral region of 300 – 800 nm and the near Infrared (nIR) 835 – 1700 nm. In current HSI applications, the visible spectral region of 300-800 nm is widely used, for example, in geospatial imaging of vegetation [147]. The use of nIR HSI is less common due to its price and also that many uses of HSI are for investigating colour changes such as with the geospatial imaging, which has been a common use for HSI for many years. However, the nIR is becoming more common in laboratory situations where samples often have the same colouration and it is more difficult to use the visible systems to distinguish samples.

3.1.2 Pharmaceutical analysis

There are very few examples of HSI systems which use a scanner or tuneable filter for fast acquisition of spectral data pertinent to pharmaceutical analysis. There are many examples of imaging within pharmaceutical analysis where either single wavelengths are captured or the data is captured one pixel at a time. These systems are often known as chemical

imaging and are most common with nIR and Raman spectroscopies. The majority of the available HSI literature on pharmaceutical analysis centres on using a HSI system to ensure even blending of drugs in solid dosage forms. This is useful to verify the correct packaging and labelling of drugs and also to check their authenticity [134]. This research is often carried out using HSI systems which operate in the near-infrared (nIR) (835-1650 nm) region of the electromagnetic spectrum. This is due to the speed, ease of use and non-destructive nature of the nIR region and because standard nIR spectroscopy is common in pharmaceutical analysis. A detailed discussion of the use of nIR in pharmaceutical applications is beyond the scope of this chapter, and there are a great many comprehensive review papers which cover the subject [148,149]. Outside of pharmaceutical research, extensive comparisons have been made between nIR and visible HSI in order to establish the most effective wavelength range for the imaging application in question [131]. Here we present, to the authors' knowledge, the first study of the relative benefits of nIR and visible HSI in the field of pharmaceutical analysis.

3.1.3 Polymer materials

The monomeric polymers used in this chapter are polyethylene glycol (PEG) and methoxypolyethylene glycol (MPEG) and the block copolymers are polyethylene glycol-polycaprolactone- polyethylene glycol (PEG-PCL-PEG) and methoxypolyethylene glycol-polycaprolactone (MPEG-PCL). The structure of both the monomeric and block copolymers is shown in Figure 3-1

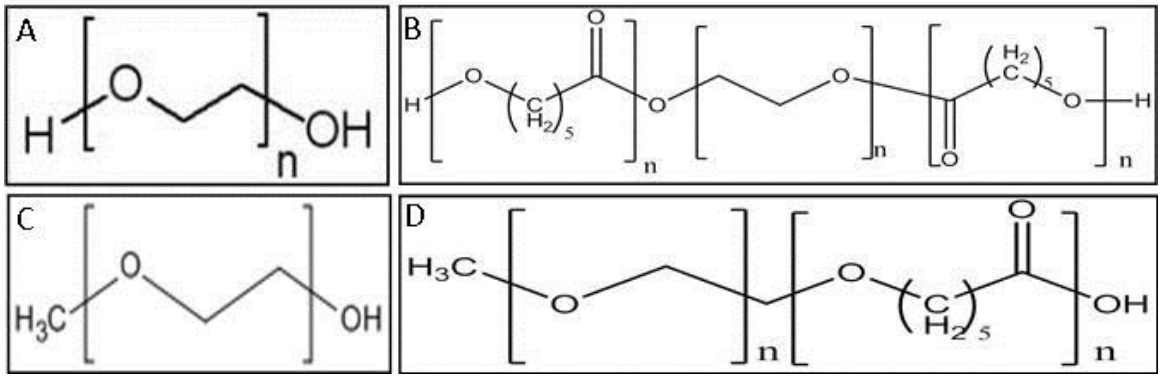


Figure 3-1 Structures of polymers investigated in this chapter A) PEG, B) PEG-PCL-PEG copolymer C) MPEG and D) MPEG-PCL copolymer

3.2 Materials and Methods

3.2.1 Experimental procedure

Two similar co-polymers, PEG-PCL-PEG and MPEG-PCL were synthesised using similar thermal reaction methods as described in [150]. Polymer starting materials (PEG and MPEG respectively) were reacted with the monomer ε-caprolactone in 5 mL of toluene with tin (2) ethylhexanoate as a catalyst. The reaction mixtures were heated and stirred and the resulting polymer dissolved in dichloromethane and extracted using cold hexane. The block copolymers have a marked difference to the form of the original monomer and monomeric polymers, the synthesised polymers were off white, large thin sheets while the starting materials were white powders. The reactions were deemed successful without performing additional qualitative analysis as the aim of this work was to separate like signals not quantifiably show that degree of polymerisation in the samples. Polymer samples and starting materials were placed onto a microscope slide and fixed in place. The slide was then placed onto a motorised stage beneath the hyperspectral camera. The camera was positioned 135 mm above the stage with an exposure time of 50 ms, a scanner speed of 4mm/s and the pixel resolution was 219 x 219 μm^2. The images were captured using these setting and the SpectraSENS software, the data was also corrected into reflectance values using the software with and 100% and 0% standard.

Two HSI cameras were used: the first imaged the 280-860 nm UV/visible spectrum region with a spectral resolution of 2.36 nm and the second imaged the 835-1650 nm nIR region with a spectral resolution of 3.28 nm. Data in the visible spectrum region was collected using an Andor Luca EMCCD camera with a Specim V8E spectrograph. The nIR region was imaged using a Xenics Xeva CCD camera coupled to a Specim N178 Spectrograph. Samples were mounted onto a Zolix KSA 11-200S4N motorised stage and illuminated with an Armley 150W Halogen lamp. No magnification was used in this setup to allow for all of the polymer to be imaged. Hyperspectral data was manipulated using both the ITTVis ENVI software package and MATLAB.

3.2.2 Image analysis techniques

Principal component analysis (PCA) was used on the spectral data to better separate the spectra and to determine which peaks are most efficient at separating out the two similar polymer samples. A 10 x 10 area that showed a large degree of homogeneity throughout was chosen for the PCA analysis and each pixel in that area was used. This created a set of

100 data points per polymer in the PCA analysis. All PCA was performed using the SIMCA-P 11 software from Umetrics.

Pre-processing is a common procedure in data analysis to prevent any bias or skew in the data. In this work the topographical changes in the polymer surfaces and variations in light intensity caused the greatest variation within the images. The variations caused by the changes in light intensity are corrected by calculating reflectance values and normalisation can reduce the effects of the topographical changes present in the samples.

$$X_{i,\lambda} = \frac{x_{i,\lambda}}{\left(\sum_{k}^{N} x_{i,\lambda}\right)}$$

Equation 3-1 The equation for calculation of normalisation by total peak intensity

The pre-processing used in this work was normalisation by total peak intensity, the equation for this is shown in Equation 3-1. X is the normalised value derived using x the raw data, i is the sample number (pixel number) and λ denotes the variable number (in this case wavelength). This normalisation divides each point in a spectrum by the integrated intensity of that spectrum, which helps to enhance spectral features while also bringing pixel intensities closer together,

3.2.3 Data correction

Natural variations in light intensity between and within images can cause substantial difference between images and so must be corrected for. This is done by calculating the reflectance value of the data and using this value rather than the raw ‘counts’ value. The calculation for determining the reflectance of any point is shown in the Chapter 2 and all data is changed into this format. The correction was performed using the imaging software.

3.3 Results and Discussion

The images were captured using the HSI systems and the data analysed; for the analysis a 10 × 10 area of pixels was chosen. A greyscale image of the PEG-PCL-PEG co-polymer formed using the data from the spectral band at 501 nm is shown in Figure 3-2. This band was chosen as it showed strong intensities and gave a clear image. The red square shows the 10 × 10 region selected for spectral analysis; the area chosen is relatively homogenous with both minimal curvature of the surface and no notable shadowing effects. Greyscale images of the other polymers are shown in the appendix for reference.

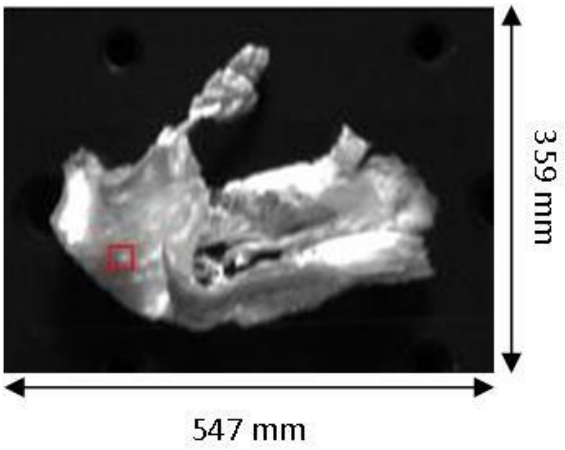


Figure 3-2 grey scale image of the PEG-PCl-PEG co-polymer using intensities at 501 nm

The mean values of these pixels were used to produce the spectra shown in Figure 3-3. This sampling method resulted in data sets of sufficient size to accurately analyse the results without having to use the whole image. Thus it saves on computer processing, allowing for faster analysis.

3.3.1 Results from the visible light HSI Camera

3.3.1.1 Analysis of sample spectra collected from HSI images

The polymers were individually placed under the camera and imaged using the setup in section 3.2.1. The spectra of the data gathered using the 280-860 nm visible camera are shown in Figure 3-3 with the PEG-PCL-PEG polymer and monomer in Figure 3-3A and the MPEG-PCL polymer in Figure 3-3B. The 280-450 nm region of the spectra was omitted from the plots due to a large degree of noise dominating the spectra in this region. The region of polymer imaged for the analysis in Figure 3-3 was chosen as it showed little to no changes in light intensity as well as minimal topographical variation across the

sample. It is clear from the graphs of Figure 3-3A that the two different materials have different reflective properties which are giving vastly different intensities. The starting material, PEG, is a fine white powder with a strong constant reflectance in the visible wavelength while the synthesised copolymer is a thin film with a much weaker reflector of the light across every wavelength.

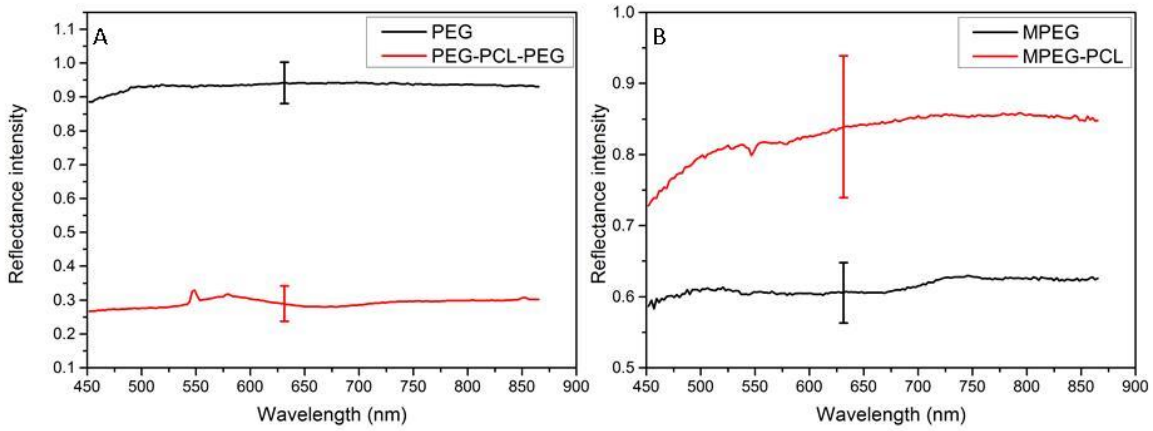


Figure 3-3 The mean reflectance spectrum (n=100) imaged in the visible region A) PEG-PCL-PEG and PEG polymers, B) MPEG-PEG co-polymer and MPEG. Error bars show the standard deviation of the results

The only chemical difference between the PEG and PEG-PCL-PEG samples is the addition of the PCL block polymer which was also added to the MPEG. If the strong difference in the spectrum was from a chemical change then this should affect both co-polymer samples. However the intensity of the copolymer in Figure 3-3B is greater than the intensity of the monomeric starting material. This difference in the intensities is not necessarily from any chemical difference between the samples but is most likely caused by a difference in the topography and particle size of the polymers. The scattering of light is affected substantially by particle size and these different compounds are likely to have different particle sizes. The error bars in the Figure 3-3 show the standard deviation of the 100 spectra averaged, a single error bar is used to prevent the spectra being obscured by the full set of error bars. The variation is even throughout the spectra and one error bar is a good representation of the error within across the spectrum. The largest errors are in the synthesised polymers which show a standard deviation of more than 12%. However, the variation is an intensity scaling factor as when plotted the standard deviations show a linear plot across all wavelength, thus the deviation is even across all wavelengths. If the difference was from different spectra then these will cause a non-linear plot of standard deviations across the data as the different peaks would result in greater variance in the data set.

To combat this variation in the ability of the polymer samples to reflect visible light normalisation was used. This process, which is detailed in the methods section 3.2.2, will reduce this variation between the data and gives a clearer image of the spectral similarities and difference between the polymers. All of the spectra were normalised as standard from this point on and the normalised data plotted again in Figure 3-4. As before a single error bar is included to show the standard deviation of the data. These error bars are much smaller than in Figure 3-3 with the PEG-PCL-PEG having the largest percent standard deviation of 4.14%, this shows that the process of normalisation has successfully reduced the variation between spectra from the same polymer.

The spectra in Figure 3-4A have pronounced differences, the spectrum of PEG shows a broad peak around 500 nm but little other spectral characteristics while the PEG-PCL-PEG contains a strong peak at 580 nm. The peak in the PEG-PCL-PEG data is visible in Figure 3-3A but it is much weaker than in Figure 3-4A which implies that the normalisation is enhancing this difference making polymer identification easier. The sharp peak seen at 546.1 nm is from mercury emission from the fluorescent lighting in the laboratory and shows the ability of HSI to detect specific emissions in the spectra. This peak is present in each of the spectra to some degree which highlights a potential issue with the correction factor not being able to remove this peak as it should.

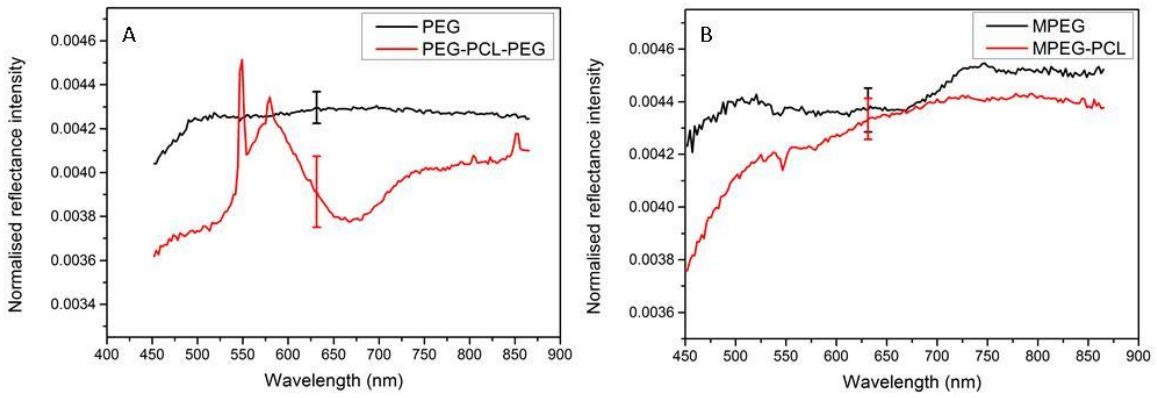


Figure 3-4 Mean reflectance spectrum (n=100) imaged in the visible region A) the PEG-PCL-PEG and PEG polymer B) MPEG-PEG co-polymer and MPEG. Error bars show the standard deviation of the results

In contrast the spectra shown Figure 3-4B have very little to distinguish them, due in part to a lack of spectral features for comparison. There is an increase in the signal intensity after 700 nm in the MPEG which is the major spectral change which could be used to distinguish between the data. The differentiation between samples will become less

pronounced as individual pixels are user rather than the mean of 100 pixels. This is due to the decreased signal to noise ratio when using individual pixels which could easily leas to the small variations between spectral signals being missed.

In both Figure 3-4A and B there is a small peak present in the synthesised copolymers around 550 nm which is not present in the starting materials. This peak is most likely caused by the inclusion of the PCL repeating unit into the copolymers and shows that the visible spectrum is able to differentiate some spectral differences between the samples.

3.3.1.2 Principal Component analysis

To improve separation of the data, principle component analysis (PCA) was used and the generated plots of scores are shown in Figure 3-5. Each point in the scores plot represents a unique pixel, and thus a unique spectrum, in the image and each polymer sample has 100 pixels which were used in the PCA analysis. The data was imputed into SIMCA-P and the PCA was then run using no other processing of the data. The scores plotted in Figure 3-5A are from the first two principal components computed from the PEG-PCL-PEG and PEG data. As can be seen, a good separation of the data is achieved in the first principle component (x axis) as the two separate data sets sit on either side of the origin. However, the data from the PEG-PCL-PEG sample is heavily dispersed in the direction of the second principal component (the y-axis). This implies that the first component is separating the samples but the second is highlighting a secondary (smaller) variation within the data.

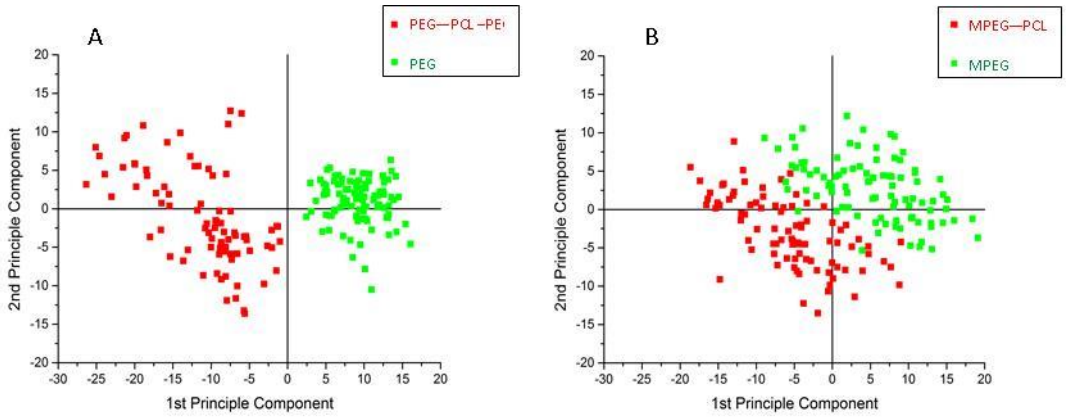


Figure 3-5 Plots showing the scores values of the first two principal components, the data in A) shows PEG-PCL-PEG co-polymer and PEG and B) the MPEG-PCL and MPEG polymers

A closer inspection of the scores plots reveals that the disperse pattern of the PEG-PCL-PEG polymer is not as random as it appears. Each sample point in the PCA scores plot is

from an individual pixel in the HSI image and the general position of these pixels from the sample area can be ascertained. The first ten samples are from the first line horizontally across the HSI image, the second ten samples are from the next line and so forth and included in the appendix is the scores plot with the entire sample locations annotated. The clusters are showing data from similar spatial locations in a vertical plane with the majority of the dispersed results being from the edges of the sample region. The symmetry of these results, with both sides of the sample area resulting in a dispersed pattern of scores indicates that the variation shown in this principle component is from topographical changes caused by the curved nature of the sample.

The data from MPEG-PCL and MPEG in Figure 3-5B shows poor separation and poor grouping with the two sample groups containing a small amount of overlap. No single component is showing the separation of this data rather the first two components are simultaneously separating out the results to create two distinct but close regions in the scores plot. The broad clustering of the data indicates that there is a great deal of inter-sample variation in both the synthesised polymer and the initial starting materials which are most likely due to topographical variation within the samples, chemical composition and particle size should be similar throughout the two polymers and so this shouldn't be a major cause of variation.

Principle Component	Percent of total sample variation (%)	
	PEG-PCL-PEG	MPEG-PCL
1	70.8	42.9
2	84.0	57.4
3	86.7	62.5
4	87.2	65.6

Table 3-1 The cumulative percentage of total sample variation contained within the principle components

The eigenvalue of a principal component is a measure of the variation within the data set that is contained within that component. By using these values a cumulative percentage of the total variation detailed within the principle components can be calculated, and this is done as standard in the SIMCA software and Table 3-1 shows these results. In the PCA of PEG-PCL-PEG and PEG the first three components show 86.7 % of the total variation in

the data with each subsequent component there after only containing a very small degree of the total sample variation. The second PCA was run on the MPEG-PCL and MPEG sample data and the cumulative variance are also shown Table 3-1. The cumulative totals for this PCA are much lower than the first samples with only 65.6 % of the variation within the sample described by the fourth principle component. However after this point the individual variation between principle components is small and no significant clustering of the data can be found. In each sample the first two components contain the largest share of the variation and so the scores values from these two components were analysed by plotting them in the graphs seen in Figure 3-5.

The loadings show which variables are causing the majority of the variation within the data the stronger the peak in the loadings the greater the contribution to the variation from the corresponding peak in the spectrum. The variables in this data are the wavelengths and the loadings are showing which wavelengths or set of wavelengths are different between samples which are analysed in the PCA. This may be a peak which is unique to one polymer sample, a change shift in peak position or a change in the gradient of a peak. The loadings were created simultaneously with the scores during the PCA, the loadings which are complementary to the scores data in Figure 3-5, are shown below in Figure 3-6.

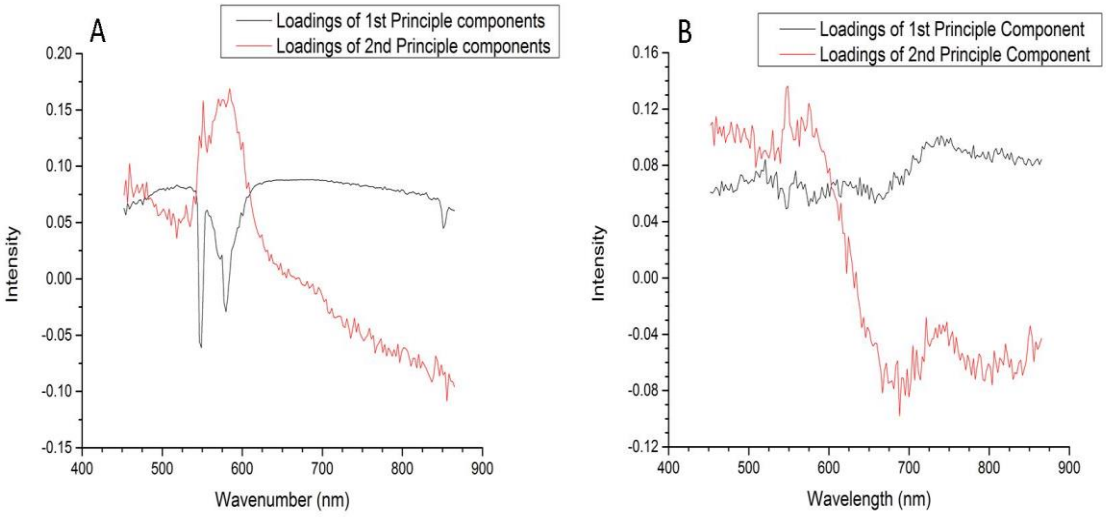


Figure 3-6 The loadings of the first two principle components from the PCA of A) PEG-PCL-PEG, PEG and B) MPEG-PEG, MPEG

Figure 3-5A shows a separation of the two samples, PEG-PCL-PEG and PEG, in the first principal component. The loadings for this component are shown in Figure 3-6A and the most prominent sections of this loadings are the peaks at 541 and 580 nm which correspond to the mercury peak and a secondary peak which can be seen within the main

broad peak in the spectrum of PEG-PCL-PEG in Figure 3-4. These results show that the spectra of the starting materials and synthesised products can best be separated using the peak at 580 nm. The loadings in Figure 3-6B do not show a single definitive peak for separation as with the other data, this is not surprising as the original spectral data contains very few peaks in the spectrum. There is however, a distinct difference between the two sets of loadings and it should be possible to use some combination of spectral locations to differentiate between the two original polymer samples with a high degree of success.

3.3.1.3 Investigating the changes in spectra at different locations on the polymer sample

The highly dispersed data from both the PEG-PCL-PEG and MPEG-PEG samples in Figure 3-5 may be due to imaging difficulties with the sample or problems relating to the region of polymer chosen for data collection. To investigate this two more areas of the polymer were selected as shown in Figure 3-7, these locations were chosen for their topographical and lighting properties.

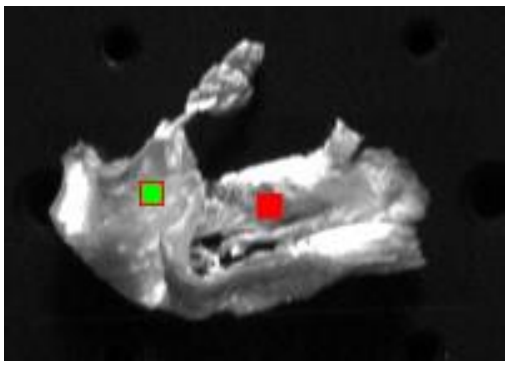


Figure 3-7 Grayscale image of PEG-PCL-PEG copolymer showing the regions selected for comparative analysis, these are known as the red and green regions

The green square is a 10x10 area of pixels which is referred to as the green region, it is relative smooth with little change in illumination across the sample. The red region is the red square in Figure 3-7 and contains a more curved surface with visible shadowing effects. The data in these regions was analysed in the same way as discussed above with scores plots generated for further comparisons. The regions were selected by visual inspection of the polymer to show which areas would best show the difference between different topographical locations. Histograms of the intensities from each pixel in the regions were also created and are shown in Figure 3-8. These histograms use the intensity at 725.9 nm to show the distribution of spectral intensities within the selected regions. This position was chosen as it represented the start of the plateauing of the data with minimal

variations caused by peak positions or shifts. Other positions were investigated and showed similar results but a larger spread in the data. The total pixel count in these histograms is 100 and therefore the count of pixels in the histogram is equal to percentage of pixels represented by that bar in the histogram.

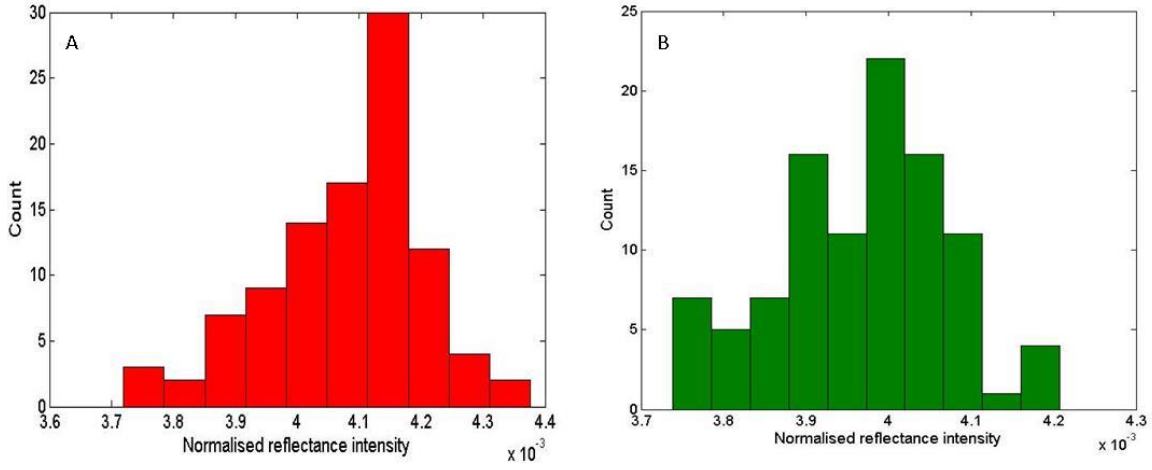


Figure 3-8 Histograms of normalised reflectance intensities for each pixel in the selected regions, A) the red region and B) the green region

The red regions histogram is shown in Figure 3-8A, there is a single intense bar which contains 30% of all the pixel intensities with a distribution around this intensity similar to a normal distribution. Figure 3-8B shows the distribution of spectral intensities in the green region and these are more evenly distributed across a smaller total range than the red region. While the total range of the green region is smaller the distribution of spectral intensities implies larger variation within the green region than the red region. No single bar in the histogram contains a majority of the spectral intensities, instead a more even distribution from 3.9 to 4.1 (x 10^{-3}) is observed. These regions were chosen by eye to be different but with the red region containing the greater variation especially from lighting effects. However, these results show that the red region is the most homogenous and that the reflectance correction and normalisation processes are effectively reducing the variability in the sample from lighting changes.

The Scores plots in Figure 3-9 show the results from the PCA of the data for the selected regions in Figure 3-7. Four separate PCA scores plots are shown in Figure 3-9 which are investigating the similarities and difference between the two regions in the PEG-PCL-PEG copolymer and a PEG sample.

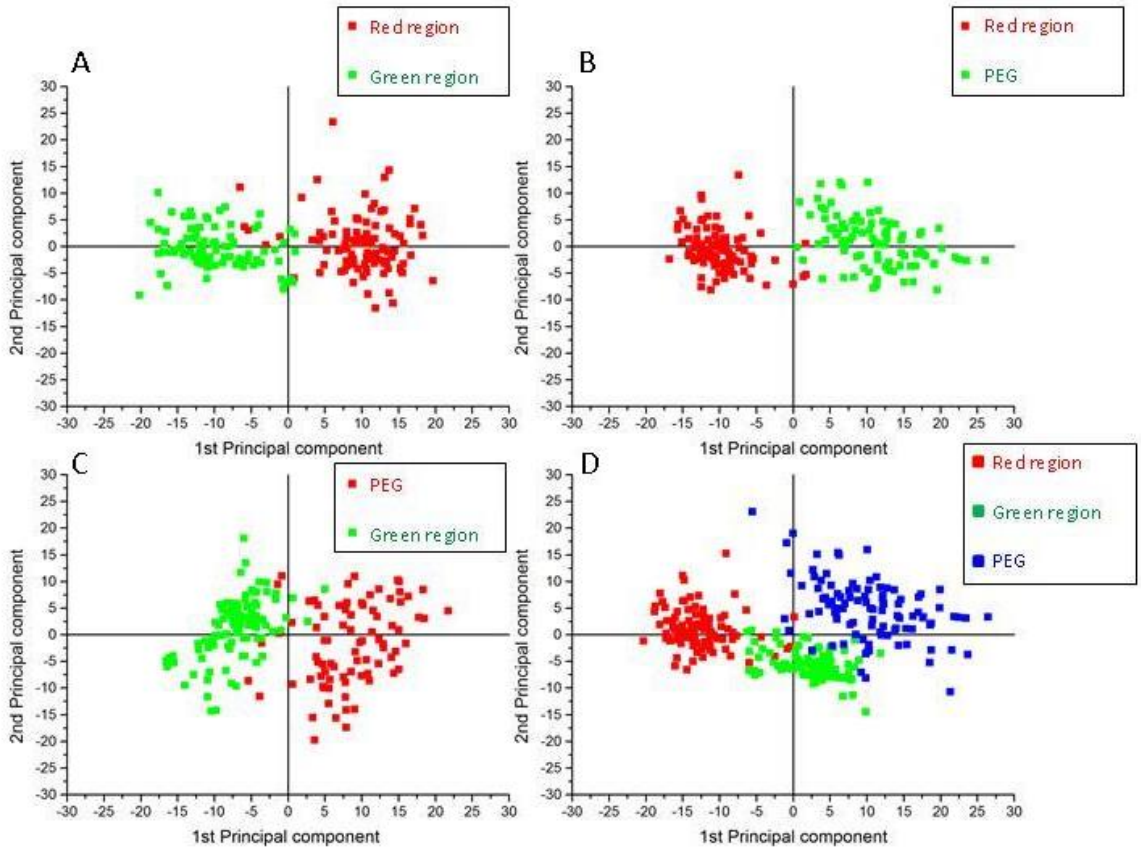


Figure 3-9 Scores plots showing the results from the PCA comparing different regions of the PEG-PCL-EPG copolymer both against each other and against a sample of PEG.

Figure 3-9A shows the scores from the PCA of the red and green regions, the chemical compositions of these two areas is the same and there should be no distinguishing features between the spectra. The PCA should not be able to clearly separate out the results into distinct groupings if only the chemical composition is affecting the spectra analysed in principal component analysis. However, the data is separated into two groupings which overlap slightly which implies that the topography or morphology is causing a variation in the spectral data. Figure 3-9B and Figure 3-9C both show the PCA scores plots of the regions in PEG-PCL-PEG against the PEG monomeric starting material. Again a good separation is seen with minimal overlap of the points in the scores plot which shows that while there may be intra sample variation it is not sufficiently large to affect inter sample separation in PCA. The spread of the green region in Figure 3-9C is greater than the red region in Figure 3-9B which implies that the pixels in the green region are less homogenous and contain greater intra sample variation. These plots show that while it is important to choose areas for sampling carefully it shouldn't affect the ability of the PCA to separate different polymeric species. Figure 3-9D is a PCA scores plot generated from the analysis of both the PEG-PCL-PEG regions and the PEG sample. This plot shows that

each of these different samples can be distinguished but with some overlap between the groups. The 1st principal component contains the majority of the variation and clearly separates the red region from the green region but the PEG overlaps the red slightly and the green heavily. This shows that a single component is insufficient to separate the copolymer and the monomeric polymer spectra but with two components there is sufficient separation in the data.

Figure 3-8 and Figure 3-9 illustrates the need for a strict selection process relating to the regions chosen for analysis to reduce the inter-sample variation. The nIR region should not suffer from these problems as severely due to the longer path length of the scattered light used. It is also possible that using a larger magnification (achieving smaller pixel sizes in the image) would reduce some of the topographical and lighting effects.

3.3.2 Results from the nIR HSI Camera

3.3.2.1 Analysis of sample spectra collected using the nIR HSI camera

The polymer samples were imaged a second time using the 835-1650 nm HSI camera to obtain data in the nIR region. The data was treated in the same way as visible HSI data with the same spatial region of 100 pixels being chosen and the average spectra of these regions are shown in Figure 3-10.

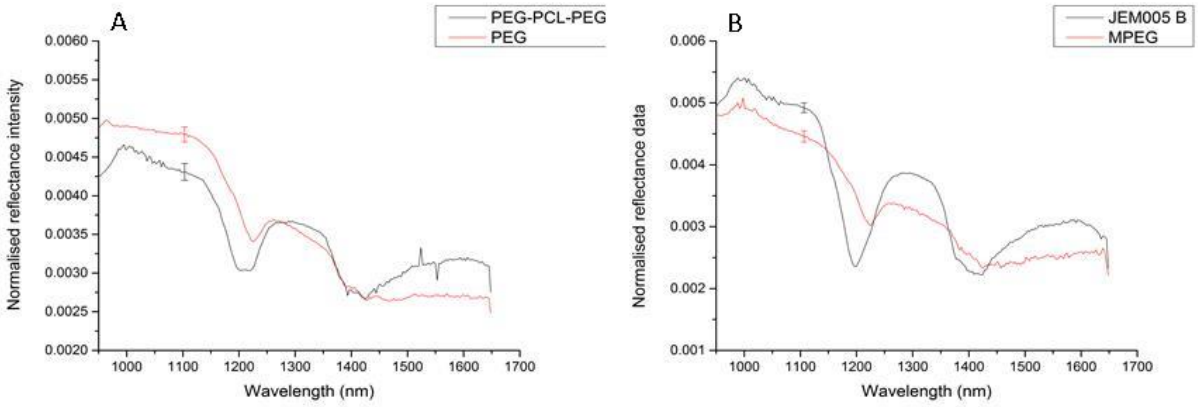


Figure 3-10 Plots containing the mean spectrum from polymer sample images which were collected using the nIR HSI camera. The data in A) is from the PEG-PCL-PEG and PEG polymers and B) contains the MPEG-PEG and MPEG polymers sample data

The spectra acquired using the nIR instrument shows much more spectral information than those obtained using the visible region camera. In the visible spectra the PEG sample gave very little spectral information across all wavelengths but the nIR spectrum contains

numerous peaks. Figure 3-10A shows the excellent separation between the spectra of the PEG starting material and the synthesised PEG-PCL-PEG polymer, this difference between the spectra is most notable from 1200 nm onwards where the size and shape of the peaks are very dissimilar between the two spectra. The peaks around 970, 1200 and 1450 nm are characteristic of OH overtone and combination vibrations. The broad peak between 1450 and 1650 nm is seen only in the polycaprolactone containing species and is most likely a combination vibration involving the carbonyl group PCL.

The improvement in spectral information with the nIR HSI camera compared to visible region HSI is particularly apparent in Figure 3-10B. In the visible spectrum there was very little to distinguish between the two different polymer spectra and even using PCA it was difficult to differentiate the results. In the nIR however there is a clear distinction between the averaged spectra of MPEG-PCL and its MPEG precursor which was not the case in data collected in the visible spectral range. The MPEG-PCL spectrum is overall of much greater intensity, and contains a peak at 1180 nm which dwarfs a neighbouring peak in the MPEG spectrum. The close location of these two peaks may imply that a peak that is present in the spectrum of MPEG at around 1260 nm is also present in the spectrum of the MPEG-PCL polymer, but its position is shifted due to the different environment in the MPEG-PCL. Once again, there is a broad peak present in the 1440-1680 nm region of the spectrum of the synthesised MPEG-PCL polymer but absent in the spectrum of the MPEG starting material.

3.3.2.2 Principal Component analysis of nIR data

PCA was performed on the data from the nIR HSI camera using the same pre-processing as earlier and the results are shown in Figure 3-11 as scores plots. The scores plots show good separation of the samples and the groupings are tighter than the equivalent data from the visible camera. The distribution of the PEG-PCL-PEG sample data shown in Figure 3-11A is much tighter than in the previous experiment, with excellent separation by the first principle component. The second component is again showing the variation from the different locations on the surface of the polymer but the spread is much tighter than previously. This shows that the nIR data is not affected as severely by the same topographical and lighting issues as the visible data.

The cause of the decrease in the variability of the data due to lighting issues is due to the different wavelength range used. During the imaging using the visible HSI system there are

a number of background sources of photons of visible wavelengths of light, these range from the sun to the lights in the room and contribute a large degree of variability as these lights do not emit a constant even flow of photons. However, the nIR has very few causes of background radiation which are strong enough to affect the results and so the photons emitted by the light source are the large majority of the total used in the experiment. The light source is made to emit photons at a constant rate and at high intensity thus reducing the variability in the lighting effects in the HSI data. The decrease in variability across the data due to the topographical change in the results is lessened when imaging in the near infrared compared to the visible. This change is from how the photons interact with the polymer surface; the visible photons will often be immediately reflected back from a surface at an angle related to the angle of incidence. So as gradient of the polymer surface changes so does the relative quantity of the total incident photons which are reflected back at an angle which can enter the instrument. The nIR photons however, will often penetrate through these samples and then be emitted at a location close to but not exactly the same as the incident point on the polymer surface. This means that the photons are not being reflected at the polymer surface which is the main cause of the topographical variations within the images.

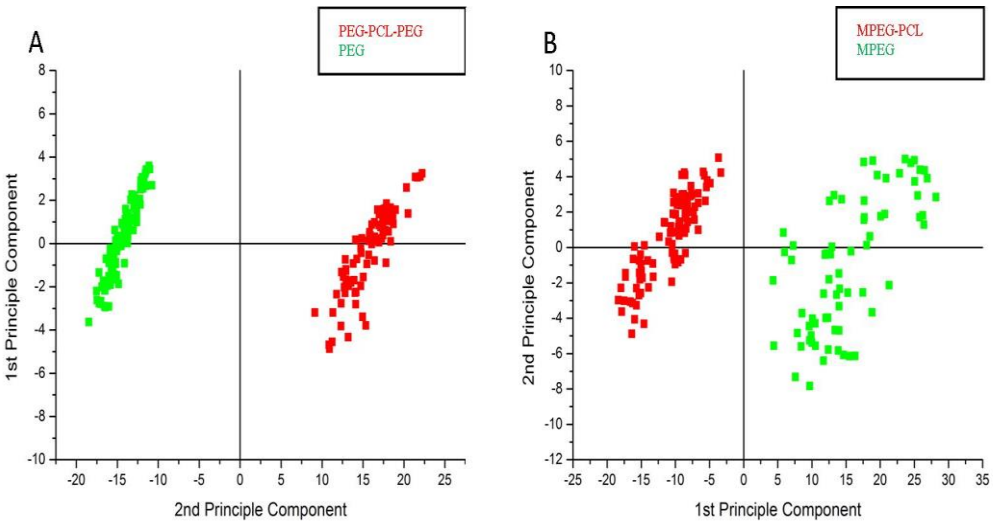


Figure 3-11 Scores plots showing the first two principal components of polymer sample data from nIR region as with previous plots A) contains the PEG-PCL-PEG, PEG data and B) contains the MPEG-PEG, MPEG data

The data from the MPEG sample in Figure 3-11B is unusually dispersed when compared to the other sample data captured using the nIR camera. The dispersion of the data may be due to imaging difficulties with the sample or problems relating to the region of polymer chosen for data collection. The MPEG was a purchased from Sigma-Aldrich and not

synthesised for the experiment so the powder is assumed pure. Therefore the separation seen in the data is from topographical issues which arrive from the powder having an uneven distribution rather than chemical variation within the sample. The scores biplot results was also dispersed when using the data captured with the visible camera, however on inspection the total dimensions of the spread it is clear that the nIR is a smaller cluster than the visible. This shows that while some problems still occur due to variation within the images these problems are minimised in the nIR.

The biggest improvement between the visible and nIR results is shown in sample data from MPEG-PCL and MPEG in Figure 3-11B, where the samples are grouped into two well separated clusters. With the visible data there was an overlap of results with a large spread in the PCA scores plot but the nIR is giving much better separation of the data. There is a clear grouping and also the separation between the samples is a function of the first and not the first and second principle components. This is important as it allows for easier separation of similar signals by investigation of a single region in the spectrum which would be shown by the loadings.

3.4 Conclusions

This study has shown that there are significant differences in the effectiveness of HSI in distinguishing between two block copolymers and their starting polymeric materials. The visible HSI system had an effective operating range of 450 - 860 nm and showed little spectral definition that could be used to separate the polymer signals. The nIR HSI system has an effective operating range of 900 - 1650 nm and the spectra contained more unique characteristics for the differentiation of the polymer signals. The nIR system shows functional groups present in the polymers which is the cause of the increased ability to differentiate the spectra in the nIR

Initially simple spectral analysis was tried but this was soon found to be insufficient for separating sample spectra due to the lack of distinguishing features in the visible wavelength range. PCA was used as a way to reduce the huge number of variables in the data to a much smaller sample size and to attempt to separate out the similar spectra signals of the starting materials and synthesised products. Using PCA to interpret the data, it was observed that, for the polymer systems in question, near infrared hyperspectral imaging was better able to distinguish between co-polymers and their starting materials than visible hyperspectral imaging. This result is not unexpected, as the polymers are white solids, and are therefore unlikely to have distinguishing features in the visible spectrum. In contrast, the nIR region corresponds to vibrational overtones in the polymers, which one would expect to vary significantly between samples.

This work was performed to show which wavelength range is most suited to the differentiation of similar biocompatible polymers, the cheaper more readily available visible hyperspectral imaging system or the more expensive and rarer nIR HSI systems. The results show that the nIR HSI system is much more effective than the visible HSI systems when imaging samples which are homogenous in colour and intensity such as tablets and powders. From results it was decided to use near infrared hyperspectral imaging systems for all future analysis on pharmaceutical products.

This work has also shown the importance of carefully choosing the best locations possible for the analysis of the HSI data as changes in light intensity and topography across the samples can have a substantial effect on a spectrum. The variability of the spectra due to these factors caused a great deal of the analysis issues but it is something which can be corrected for by simply choosing the idea locations on a sample surface.

The next step is to apply this knowledge to the field of pharmaceutical analysis and to investigate a process or processes using hyperspectral imaging. The majority of the existing literature which involves the use of hyperspectral imaging centres on analysing properties of simple unchanging systems such as the efficiency of the tablet blending by investigating the concentration of specific components across a tablet surface. These systems of work are novel ways to achieve something which can be performed using existing techniques such as Raman spectroscopy. The comparison of HSI and Raman spectroscopy is discussed in Chapter 6 where the two techniques are used to investigate the distribution of specific compounds within a tablet. To the knowledge of the author there is no known work on using hyperspectral imaging to investigate more complex dynamic systems which will be the aim of this work. By using the near IR HSI system and the multivariate analysis shown in this chapter the dissolution of a tablet will be followed. The dissolution will be performed in a flow cell which allows the tablet to be imaged throughout the experiment and the full dissolution will be followed. This will allow for investigation into the release of release from the tablets while also collecting information on the rate of tablet expansion and disintegration.

Chapter 4: Hyperspectral imaging of the dissolution of a range of acetaminophen containing tablets

4.1 Introduction

In Chapter 3 it was shown that a hyperspectral imaging (HSI) system which was attuned to detect photons in the near Infrared region (nIR) of the electromagnetic spectrum was best able to distinguish between and separate similar polymers. That the nIR spectrum is better than the visible spectrum is not unsurprising as standard nIR spectroscopy has been used by the pharmaceutical industry for some time as a powerful analytical tool. Near Infrared spectroscopy is a fast, none destructive technique which requires little to no sample preparation [127]. The investigating of tablet properties such as surface concentration, detection of excipient clusters and homogeneity of the tablet components is often performed using nIR or Raman spectroscopy.

The use of nIR spectroscopy for investigating the physical properties of a tablet is much less common with respect to following the release of drugs from the tablet during the dissolution of the tablet. The standard system for tracking the release of the active pharmaceutical ingredient (API) from the tablet is dissolution testing with a UV spectrometer. These procedures are common place and the *de facto* way to measure the release of an API and to calculate rates of release and the time from the tablet submersion to the drug being released.

Both dissolution testing and nIR spectroscopy can show a great deal of valuable information on the physical properties of a tablet but they have never been used together in the one system before. This work will use a hyperspectral camera which detects light from the near infrared spectrum to image the dissolution of a known tablet. This has the potential to combine these two techniques to investigate a number of physical processes. Using the imaging it should be possible to track the disintegration and expansion of the tablets, to determine how fast and for how long these processes occur and by using the spectral component then it should be possible to follow the release of the individual components of the tablets. By combining both the spectral and spatial information it should

be possible to track specific excipients within the tablet and follow them as they dissolve or break away from the tablet.

To create tablets in a laboratory setting can be time consuming as it requires a great deal of analysis of batches to ensure the tablets are uniform in their API content. There are also potential issues with components clustering within the tablet which could result in poor dissolution properties.

Commercially available tablets are created to a rigorous standard and must be within strict guidelines of purity, percentage error of APIs used and release properties. As such the use of these tablets as samples should remove any requirement to create tablets carefully in the lab. Using commercially available tablets a number of variants with slightly different excipients are available which may lead to different dissolution properties.

Paracetamol, or acetaminophen as it is also known, is a very common, readily available drug with a number of commercially available formulations and the chemical structure is shown in Figure 4-1. It is both an analgesic and antipyretic meaning that it helps to reduce pain and fevers in patients. Due to these properties it is often used for cold and flu medications to help lower any fever the patient may be suffering from while also helping with general pains. Commercially available paracetamol was chosen as the tablet for dissolution testing. Furthermore paracetamol is a fast acting tablet with release normally reaching 80% within 30 minutes. There are a number of paracetamol tablets which are also packaged with caffeine as a secondary API. Caffeine is a stimulant and is used to improve the performance of the drugs by helping to reduce any drowsy feeling which the imbiber may be suffering from.

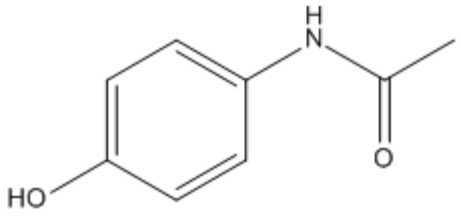


Figure 4-1 The chemical structure of paracetamol

4.2 Methods and materials

In this chapter the dissolution of known tablet systems will be followed using a hyperspectral camera attuned to detect nIR photons. There are three things required for this experiment to proceed which are the tablets, the camera and a setup which will allow imaging of the dissolution.

4.2.1 Materials

The chosen tablet brands cover a range of excipients and complexity with some basic tablets which contain little more than the API and starch as filler. Some of the tablets investigated are more complex containing a wide range of excipients which are used to improve the performance of the tablets. The total list of drugs imaged was Panadol Actifast, Panadol Advance, Panadol Extra, Tesco Paracetamol caplets, Tesco Paracetamol Extra, Boots Paracetamol and Boots Paracetamol Extra. These tablets all contain 500mg of paracetamol with the 'Extra' tablets containing an addition 60 mg of caffeine. These tablets use starch as a major excipient which works as a binder and filler to improve the stability of the tablets. Panadol Actifast is the only brand which includes an excipient in higher quantity than the starch which is sodium bicarbonate and this is used as a disintegrant. The sodium bicarbonate works by reacting with stomach acids to release carbon dioxide which forces the tablet apart and enhance the breakdown of the tablet [151].

The number of and types of excipients present in each tablet varies between but there are some common excipients such as starch, Polyvinylpyrrolidone (PVP) and magnesium stearate which are present in most tablets. The full list of excipients is detailed in Chapter 2.4.2. This variation between tablets is what causes the different dissolution properties and should allow for a variety of spectral changes to be visible. The common excipients and the two APIs were also imaged to provide a set of reference images which contained the spectra of these compounds which could be used in identification of spectral signals within the tablet. The different brands of paracetamol were given codes for ease of use and reference and these codes are displayed in Table 4-1.

Name of tablet	Panadol Actifast	Panadol Advance	Panadol Extra	Tesco paracetamol	Tesco paracetamol extra	Boots paracetamol extra
Code for tablet	Tablet A	Tablet B	Tablet C	Tablet D	Tablet E	Tablet F

Table 4-1 The naming codes used for each paracetamol brand

4.2.2 Experimental design

The dissolution apparatus for this experiment could not use a traditional USP dissolution system. All of these require the tablet to be placed into 900ml of aqueous solution (normally the buffer reservoir) or to be placed inside of a flow cell. These vessels are all curved to enhance the mixing properties of the system and prevent tablet material collecting in a corner. . All of these vessels would cause large complications with the quality of the imaging by causing distortion of or obstructing the imaging of the tablet. A 1cm quartz flow cell was bought which contained in and out flow ports along with a screw top to allow sample tablets to be placed inside once the flow cell was empty. The flat surface of the flow cell along with the low volume of liquid to image through, allows for effective imaging of the tablet dissolution using a flow cell of this design. Using this flow cell a setup was created which was modelled on the USP type IV with the buffer solution being pumped from the buffer reservoir through the flow cell and back to the reservoir. In these experiments the solution used was a simple pH2 hydrochloric acid solution; a more complex buffer system was not implemented as the camera was being used externally from the university and it was deemed unfeasible to transport the quantities of buffer solution needed. Instead the acid solution was made on site by dilution of a concentrated hydrochloric acid solution. Figure 4-2 contains two different diagrams of the experimental setup, Figure 4-2A shows a plan view of the setup with the buffer being pumped through the ports on the side of the flow cell. The two major changes from the USP type IV dissolution is that the flow cell is now a cubic 1cm path length flow cell and the UV spectrometer has been removed. While having UV measurements running alongside the HSI would be ideal there was no practical way to do this. Figure 4-2B shows a side on view of the camera being placed above the flow cell. This flow cell is placed onto a microscope slide which is itself placed onto a high precision mobile stage so that it can move under the camera and build up the images of the dissolution.

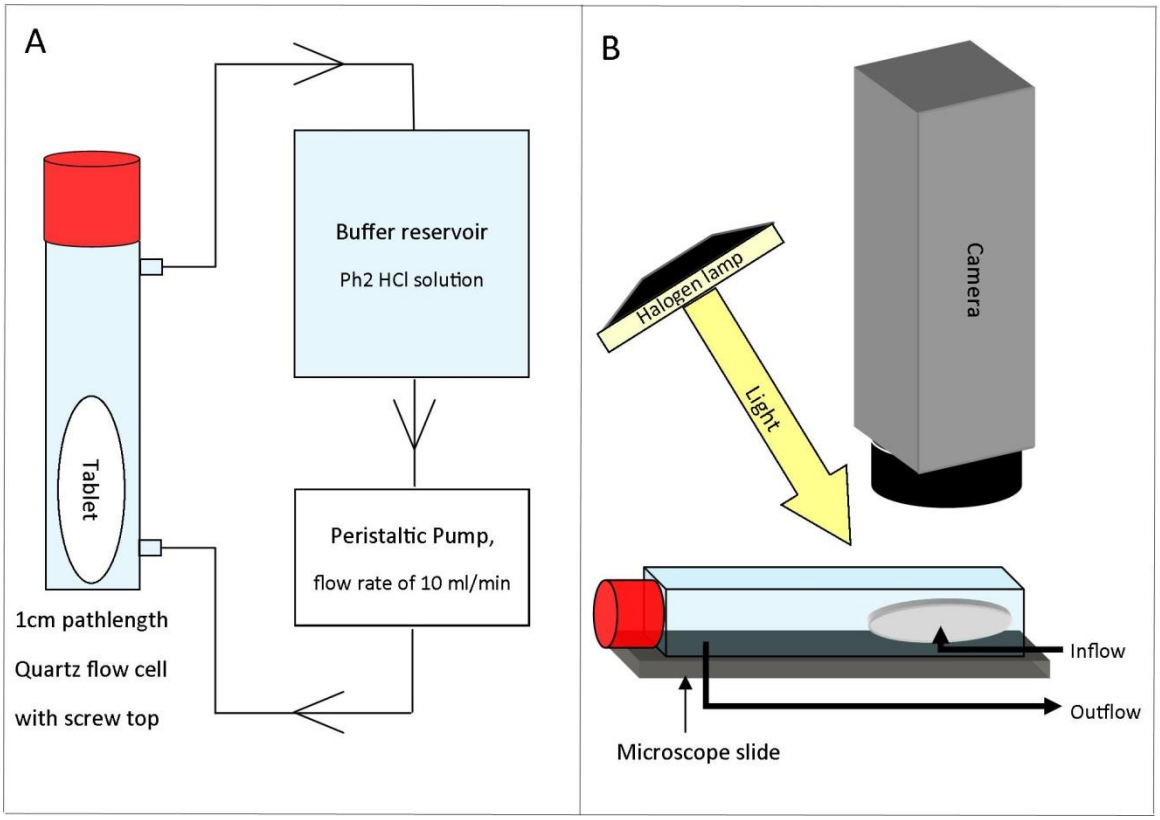


Figure 4-2 A schematic view of the experimental setup with A) a plan view of the equipment and B) a side view of the setup

By using this experimental setup it was possible to capture the dissolution of the tablet inside the flow cell. The walls of the flow cell are flat which removes the optical artefacts that a curved surface can create in the image and the total path length through the flow cell is 1cm which should remove any imaging obfuscation which the liquid may have caused.

4.2.3 Experimental procedure

Once the experiment had been devised the experimental procedure was created and followed for each experiment. The first step in the experiment was to calibrate the equipment using the procedures in Chapter 2. 3. Once the scanner speed and height of the instrument was calculated the flow cell was attached to a microscope slide using a piece of adhesive putty, the slide was then placed onto the motorised stage. The tablet was placed into the flow cell and the cell tightly closed up before the pump was activated and the cell filled with the hydrochloric acid solution. The imaging software was viewed prior to the experiment starting to ensure that the settings were correctly inputted.

The camera was positioned 100 mm above the stage and the images were scanned at a rate of 8 mm/s covering a total scan distance of 100 mm with a 4 ms exposure for each line of

the data. Using these values allowed for the greatest intensity in the pixels without saturation of the detector and also gave square pixels in the image with no distortion which can be a problem at an incorrect scan speed. The spatial size of the image pixels is 215 x 215 µm. As with previous experiments the images were captured using the Gilden in house HSI software SpectraSENS but unlike previous experiments the software was used to automatically capture a new image every 3 minutes. It was found that using a time interval of less than 3 minutes would cause the software to crash due to the large memory requirements of these images. From the beginning of the experiment a new image was created every 3 minutes for 30 minutes and then every 5 minutes until a total time of 45 minutes had passed. Using this methodology it was possible to capture the full breakdown and dissolution of the tablets. Previous UV dissolution experiments had shown that full release of the tablets should occur much more quickly than the 45 minutes allocated but this extra time was to allow for the confined steric condition in the flow cell. This was also the reason for the change in imaging rates for the last 15 minutes of the experiment as it was felt that using 5 minute intervals would help to reduce the total amount of data that needed analysis without being detrimental to the results. Once the camera was setup and the sample in position the dissolution was started. All experiments were run using a flow rate of 10 ml/min and 900 ml of pH 2 hydrochloric acid solution.

4.2.4 Equipment

The equipment used in this experiment is the same nIR HSI system as in Chapter 3. The camera is a Xenics Xeva InGaAs CCD detector with an effective array of 320 x 256 pixels. The spectrograph is a Specim V17E which comes with an inbuilt 16mm lens with f value of 1.4, the stage was a Zolix, KSA 11-200 S4N which was powered by a IMS Mdrive 17+ motor. The lighting was produced by a custom made halogen DC lamp created by the people at Gilden, this was used as the AC lamps used previously had been shown to cause flickering during testing which was discussed in Chapter 2.1.5. Tygon tubing and a Watson and Marlow 205U peristaltic pump was used for the dissolution apparatus.

4.3 Results

4.3.1 Tradition UV dissolution of tablets

Dissolution tests were run on each of the tablets using a USP type IV protocol which involves using a UV spectrometer to detect the release of the paracetamol from the tablet. These tests show the rate of release, how long a tablet takes to release and if the tablet reaches 100% release within the time allotted for these experiments. The standard non-caffeine containing paracetamol tablets were run first as the calculations for releases are simplest with these systems. Using the detected intensity at 270 nm and comparing this to a known standard a percentage of release was calculated for tablets A, B and D. These release profiles are shown in Figure 4-3 and each tablet formulation shows surprisingly different results. The error bars in the data show the average standard deviation for each tablet, the standard deviations for each tablet are low which shows little tablet to tablet variation.

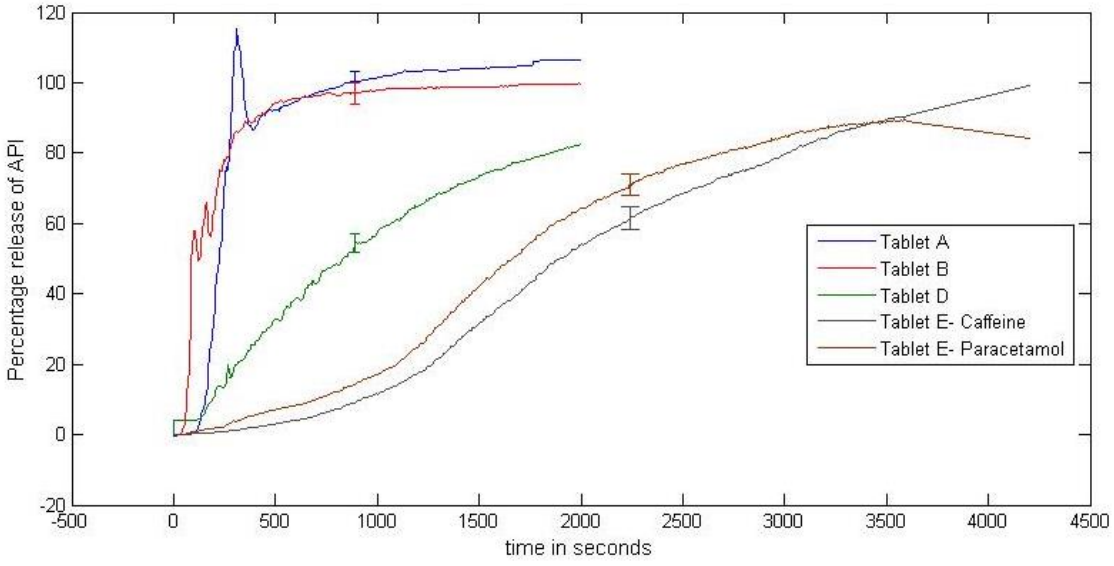


Figure 4-3 The percentage of release of paracetamol from four separate brands, n=3 and single error bars are used to show average standard deviation in the data.

Tablet A shows a peak sharp peak with a maximum at nearly 120% release, this is of course impossible. Two separate tablets were run through the system to verify the accuracy of these. The cause of the value being greater than 100% is from poor mixing and the rapid speed of the paracetamol release. The value of release is calculated against a sample made to contain the same ratio of drug in the acid solution as would be found when 500 mg of paracetamol were released into 900 ml of acid solution. The UV spectrometer detects the intensity of a specific wavelength (in this case 270 nm) from a sample of the acid buffer

solution pumped from the reservoir. This spike in the data is caused by a large collection of the tablet particles being released from the tablet and then quickly being drawn into the spectrometer before particles can effectively diffuse throughout the reservoir. As time progresses the reservoir begins to equilibrate with respect to the concentration and location of the paracetamol and thus the UV spectrometer accurately detects the percentage of release after 400 seconds. Tablet B begins to release faster than A but has a release rate slower than that of Tablet B and both tablet have greater than 90% release within 400 seconds. Tablet D is much more basic in design than tablets A and B and contains fewer excipients to expedite the release of the API. This is very clear when the dissolution profile is investigated with a release of approximately 80% being achieved by the end of the dissolution. These three different tablets have shown differing release profiles which should be reflected in the HSI data.

The caffeinated paracetamol tablets were also investigated to determine the release of the paracetamol and caffeine from the tablet matrix, the release from Tablet E is shown in Figure 4-3 alongside the non-caffeinated paracetamol tablets. The calculations required to determine the exact release of the paracetamol and caffeine are more complex than those used in a standard one drug system as both of these drugs absorb ultraviolet light and so the absorbance of the system is a function of two concentrations. This means that the spectral intensities recorded by the UV spectrometer is a function of the two component spectra adding together to create the detected signal. This process is additive and so can be solved by using standards, as before, and simultaneous equations with the exact derivations and equations shown in the appendix. Caffeine and paracetamol are both readily dissolved into acidic solutions and so their release profiles are similar with the only notable change coming at the end of the dissolution of Tablet E.

The release profiles of each of the tablets A – F is given in full in the appendix and show that the formulations are designed to release the APIs quickly which is expected from a pain relief medication. The release rates vary from one tablet to the next and this is a variation which should be notable in the HSI images. In general the tablets containing more excipients, and at a higher cost, begin releasing the drugs quicker and the rate of release is also much faster than the more basic formulations.

The hyperspectral dissolution experiments were run for 45 minutes or 2700 seconds which was chosen as a sufficient time period to image the release of the majority of the APIs

from the tablet. Using the UV dissolution data as a guide it should be possible to track changes in the images which can be attributed to the release of API. It is however, impossible to use these UV results to conclusively determine when the API is release in the HSI experiments due to the different reaction conditions most notably the different flow cell.

4.3.2 Hyperspectral dissolution testing of Non caffeinated tablets

The images were created using the methods described however, there were some complications with the software. Taking images every three minutes resulted in some images not capturing correctly as the software would freeze during the image capture. This results in some gaps in the imaging data where the images could not be captured but the majority of the imaging was successfully. The images were then investigated to see what information could be discovered from them by either simple visible inspection of the images or by more complex spectral analysis.

The images have been false coloured using a simple RGB colouring algorithm to better show the variation throughout the flow cell. In the algorithm three separate wavelengths have been chosen to create the colours of the image by assigning a colour to the relative intensities of the pixels at these wavelengths. The more intense the red wavelength, relative to the other two, the more red that pixel will appear. The wavelengths used were 1083, 1230 and 1601 nm and they were chosen because of their position on the average spectrum of the tablet which can be seen in Figure 4-4. These wavelengths are located on three separate peaks and should detail any changes in peak intensity throughout the dissolution process.

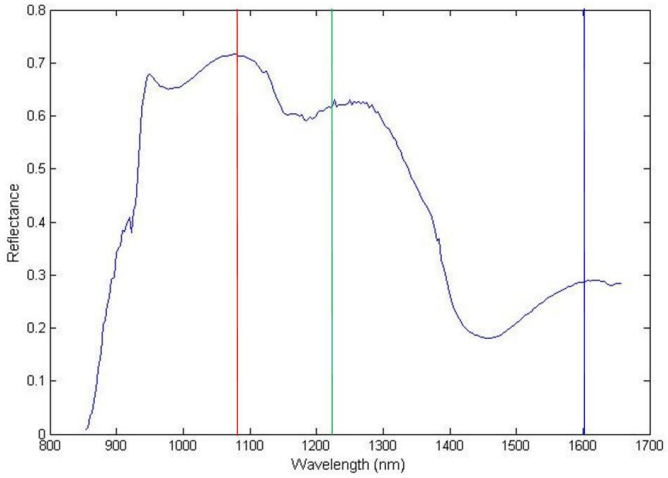


Figure 4-4 Spectra showing the wavelengths used to create false colour images. The coloured vertical lines denote the wavelength used for the corresponding false colour intensity

4.3.2.1 Tablet A

The first of the non-caffeinated tablet, Tablet A, was imaged in full and Figure 4-5 contains a selection of the images captured using the hyperspectral camera.

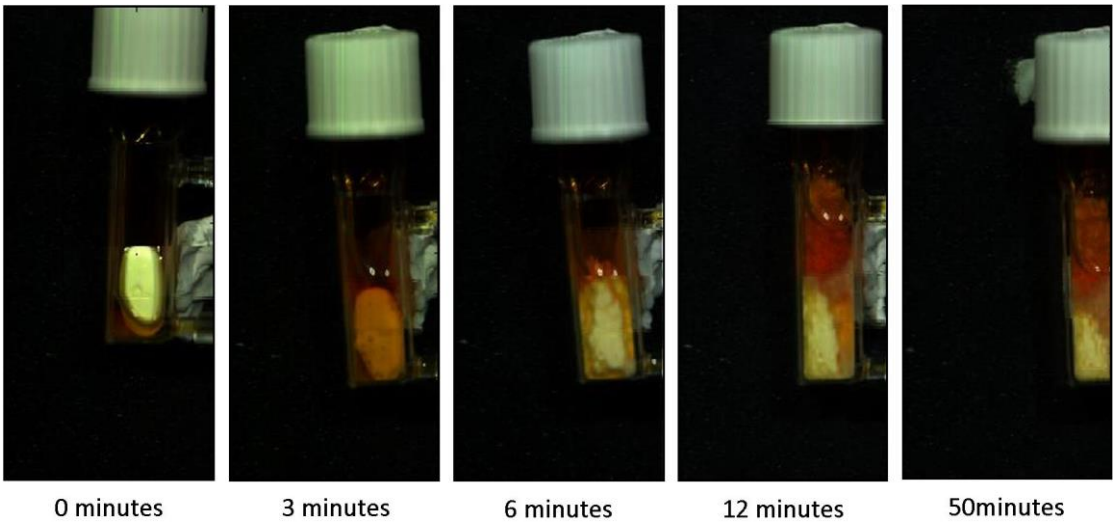


Figure 4-5 Key images from the HSI dissolution images generated during the dissolution of Tablet A

The images in Figure 4-5 show a range of interesting processes occurring during the experiment. The expansion of the tablet is visible as is the breakdown of the tablet with particles moving down the flow cell. This expansion could be from the hydration and swelling of the tablets or the disintegration and movement of tablet particles through the flow cell. It is not possible, using these images, to show which of these processes is occurring. The tablet at time 0 clearly shows the whole tablet before any dissolution has

occurred with the embossed P being distinguishable on the tablet surface. The tablet at three minutes has lost the sharp edges of the tablet which is indicating that the solution is starting to breakdown the outer layer of the tablet and that the tablets expansion is imminent. These results agree with the dissolution data in Figure 4-3 which shows a delay of around 200 seconds before the tablet began to expand and release the paracetamol within. By the 12 minute mark the tablet has completely disintegrated showing none of its original structure which again agrees with the UV data as it shows a fast and rapid release from the tablet. The amount of these particles appears to remain relatively constant even as the tablet mass is starting to get smaller as the experiment progresses.

There are some issues with the images which needed to be addressed; these are that the flow cell is clearly moving between images and that an air bubble is present at the top of the flow cell through much of the image. The air bubble reduces the useable volume of flow cell which can give viable spectral results as air bubbles can cause unusual scattering and dampening effects to the reflectance signal. The movement of the flow cell also means that it is very difficult, without additional complex image processing techniques, to investigate the same location on the tablets over time to see how they are changing. This is not so much of a problem at the bulk where the majority of the pixels are relatively homogenous but towards the edges of the tablet this can cause severe difficulties. Furthermore there also appears to be lines sporadically appearing in the data with the first image containing one line at the very top of the tablet, this is caused by the camera missing a line of data. The data was missed either by the software being unable to collect the line at that moment due to the memory allocation being insufficient or the speed of the stage is moving such that the data appears to miss a line. None of these issues are insurmountable however, and the data is still valid and of high quality.

The tablets are designed to absorb the surrounding liquids and to swell, this expansion of the tablets aids in the release of the API by speeding the breakdown of the tablets. The extent of expansion as a function of time is not a factor which can be tracked during tradition USP dissolution due to the lack of any imaging. However, using hyperspectral imaging it is a relatively simple process to follow the degree of expansion of tablet material through the flow cell. A program was written in MATLAB which would detect the edge of the flow cell and construct a set of boundaries such that only the flow cell would be investigated. Once this region had been selected any pixels with a reflectance value at a set wavelength (1263 nm) higher than 0.5 is counted. This value was high enough to remove any of the cell wall pixels but low enough to ensure that every pixel in

the image which contained tablet was counted. By counting the number of pixels containing tablet particles a quantifiable measure of the movement of the tablet particles thought the cell was determined, specifically the percentage of the total flow cell which contained tablet data. This process was repeated for every image created during the dissolution of the tablet and the percentage values plotted to generate a plot which can show the degree of expansion.

The plot of the movement of tablet particles from the dissolution of Tablet A can be seen in Figure 4-6, the y axis shows the percentage of the flow cell which contains tablet data and the x axis is time. The uneven spacing of the markers on the x axis is to represent the exact times the images were captured.

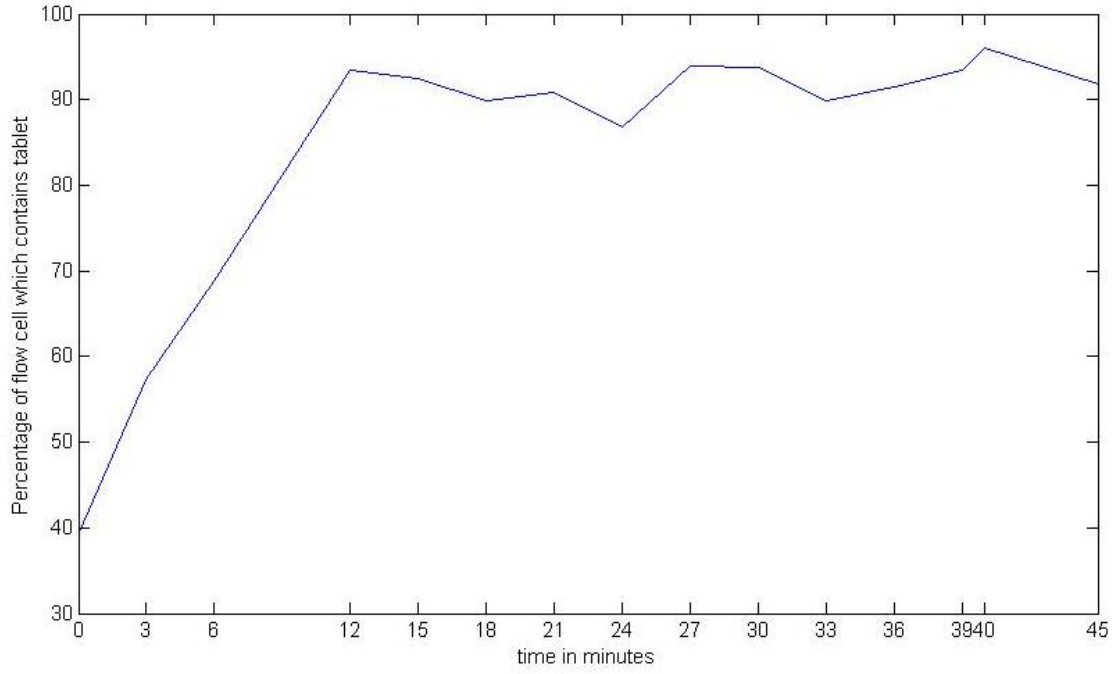


Figure 4-6 The percentage of the flow cell occupied by Tablet A throughout the dissolution process; the time points reflect the time the images were captured.

The tablet covers 40% of the flow cell at the start of the experiment and then rapidly appears to expand up to a total of approximately 93%. The expansion of the tablet and movement of particles through the flow cell increases in a linear fashion for 12 minutes when the maximum degree of expansion is reached and a plateauing of the data occur with some fluctuations around the maximum. These fluctuations are from natural variation in the data and some inaccuracy in the technique caused by the flow cell moving throughout the image which makes it difficult to define the exact same area. By inspection of the images in Figure 4-5 there is also an air bubble expanding and contracting in the flow cell,

the edges of which will have be causing a false positive in the counting of the pixels. The edge of the air bubble cause strong specular reflectance of the acid solution which give these pixels higher intensities than they would normally have. These two factors are causing the majority of the variation within the cell but there is also some natural loss of signal from the tablet material being expelled from the flow cell over time. Again even with these issues there is a clear pattern to the data and the problems are only visible in the areas where there is little change in the data, thus small variations become more impactful.

The preliminary dissolution experiments, which were performed on a USP IV dissolution apparatus (Figure 4-3), details a small lag period followed by a rapid release which began to plateau by the 8-10 minutes. The expansion data reveals a constant expansion of the tablet for the first 12 minutes, well after the tablet should of release the majority of the paracetamol. However, both the UV and expansion data show a plateauing of the data after 12 minutes which implies that either the release from the tablet is faster than the expansion or that the flow cell is impeding release in some due to the confined space relative to a USP type IV vessel. While the data may not correlate exactly between the UV and HSI results there is a clear correlation. These results also show that the release profiles created from the UV dissolution testing can only be used as a guideline to the release properties of the tablets in the HSI dissolution as the experimental conditions are not the same between the two experiments.

There is clearly breakdown and release of Tablet A within the flow cell and a spectral change which accompanies the breakdown and release of the paracetamol should be detectable. A region of 100 pixels was chosen in the bulk of the tablet and the mean spectrum calculated by averaging these 100 spectra, this it was hoped, would show how the tablet bulk is changing over time as the API is released from the tablet matrix. A mean spectrum was used as there is some level of natural variation within the data due to topographical difference across the tablet as well as each pixel containing a slightly different volume of the excipients. These slight variations mean that any one pixel could be a poor representation of the bulk as a whole and thus a mean was chosen. Figure 4-7 shows this data and apart from the data collected at 3 minutes there is very little change in the spectra over time. The error bars in the image show the mean standard deviation in the data which is approximately 0.015 for each sample.

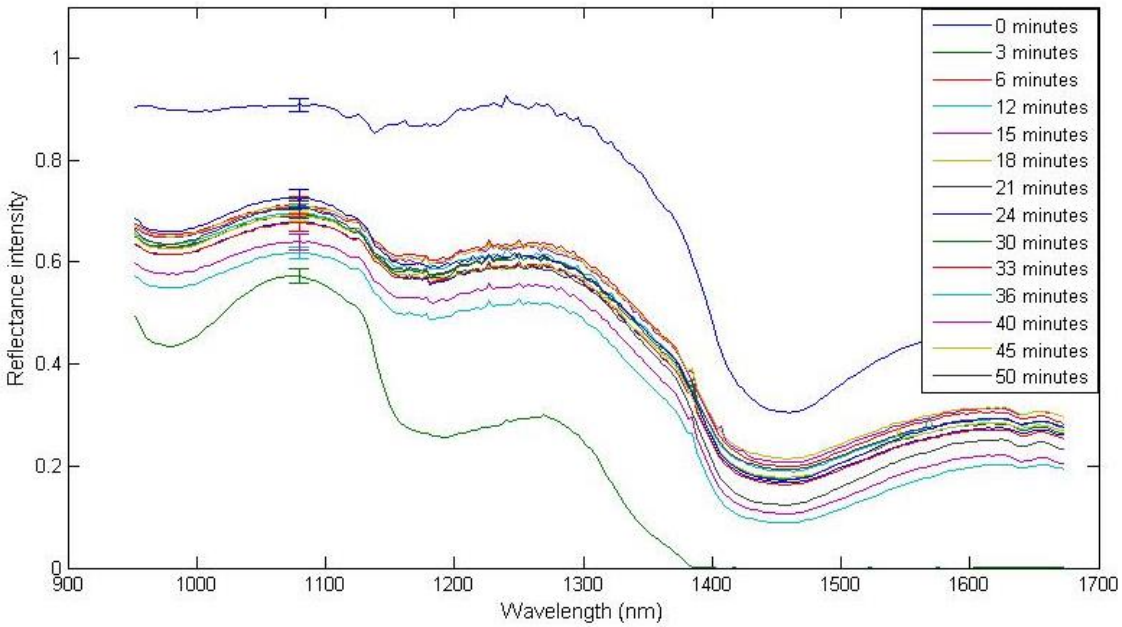


Figure 4-7 Average spectral data (n=100) from each of the different time images detailed in Figure 4-5

In an effort to clean up the plot a second set of reduced set of spectra were plotted and in this plot, shown in Figure 4-8, only the data from the images at 0, 3, 15, 30 and 45 minutes one used. The first two time points cover the hydration of the tablet and the disintegration and movement of the tablets through the flow cell and the remaining time points cover the remainder of the dissolution experiment. This generates a plot which is much easier to follow than Figure 4-7 but still contains the general flow of the spectra over time.

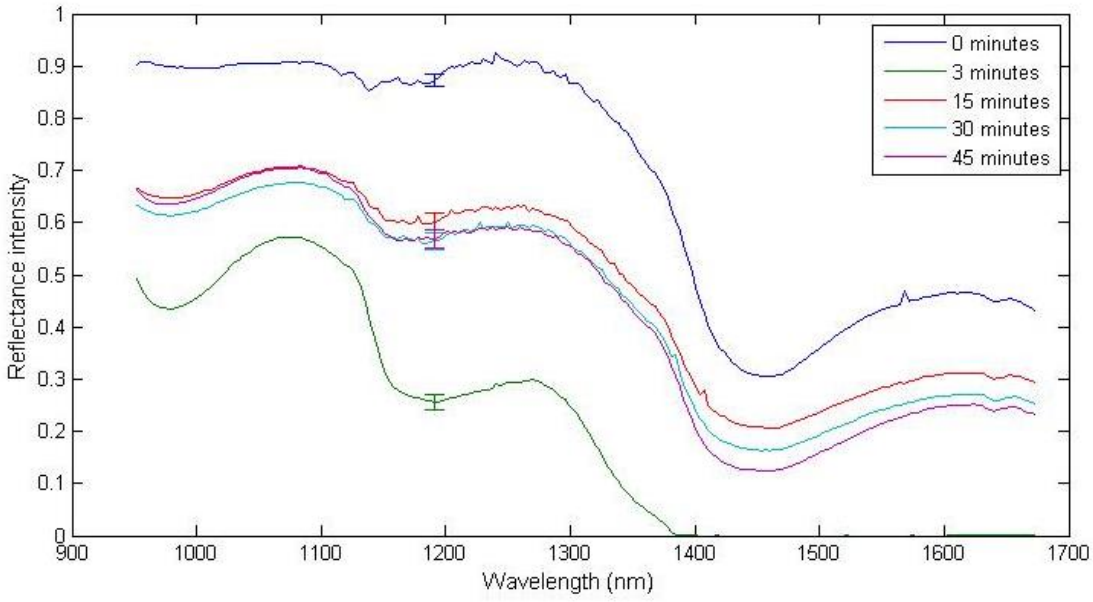


Figure 4-8 Average spectral data (n=100) from 5 different points in the dissolution of Tablet A

The spectra detail a change in the tablet as the initial dry tablet hydrates and expands; the spectra change substantially in both intensity and peak definition for the first three data points until reaching an equilibrium point. The spectrum of the 0 minute data is of the dry tablet before it is submerged and it shows a substantial difference to the other data. The tablet spectrum from 3 minutes into the experiment is much lower in intensity than the earlier dry spectrum. By this point the tablet has been submerged in the acid solution but there is only a minimal amount of expansion which implies that the change in the spectrum is not from the release of the API. The intensity increases from the 3rd to the 15th minute of the dissolution and from this point the data changes only slightly as the experiment progresses. An overall decrease in the signal intensity can be seen from the 15 to 45 minute spectra but this change is small with the difference between the 15 and 45 minute data being less than 0.05. It should be noted however, that this change is not as clear as it appears in Figure 4-8 which implies that the later the data in the central belt the lower the intensity. Inspecting the lower intensity spectra in Figure 4-7 shows that the minimum intensities are not from the ending data but from the 33rd, 36th and 45th time points with the 40th and 50th minute data having higher intensities.

The fluctuation seen in the data around this central band implies a resting spectrum with some variation between the images causing the small changes in the intensity of the average spectra plotted. In Figure 4-8 the 0 and 3 minute data shows substantial differences to the other spectra. By inspection of the images from 0 and 3 minutes in Figure 4-5 it is apparent that these spectral changes are caused by the physical state of the tablet. There are few signs of expansion or disintegration in the images of the tablets at these time points with the 0 minutes image showing the tablet before it is fully submerged. The 3 minute data is submerged but is still keeping the shape of the tablet which is not true in later images. This implies that the details in Figure 4-7 and Figure 4-8 may be showing more about the physical changes such as the state of hydration within the flow cell and less of the spectral changes which are a result of the target compounds leaving the tablet.

4.3.2.2 Tablet B

While the results from the dissolution of Tablet A are inconclusive with respect to loss of API there were two other non-caffeinated drugs imaged. The dissolution images at four points through the dissolution of Tablet B can be seen in Figure 4-9. Figure 4-9 shows that this tablet is so fast to dissolve that even the initial images show a tablet which has begun to disintegrate. The tablet has been able to disintegrate so quickly like this due to the way

in which the imaging is performed. There is a slight operational lag of approximately 1 minute between the flow cell being filled with the acid solution and the imaging starting. Once the image starts it takes a few seconds for the tablet to pass under the detector so that by the time the tablet passes under the detector the disintegration of the tablet has already begun.

Figure 4-9 The dissolution of Tablet B at 4 time points which are 0, 15, 30 and 45 minutes

The tablet at the start of Figure 4-9 has clearly begun to breakdown but the shape of the tablet remains and is clearly visible on the top edge of the tablet. However, it is clear that within a few minutes the tablet has begun to lose all shape and is simply a mass of particles which slowly spread through the flow cell. The direction of flow is from the tablet to the cap which explains the movement of the particular matter.

While the images displayed in Figure 4-9 show only the data from four of the dissolution images an average spectrum for each images was created and plotted together to create Figure 4-10. The data for the first 100 nm is not reliable, and is omitted, due to the detector having poor efficiency in this region which is not ideal as this is the only region where the spectra of the different time images differ from each other. It is possible that the only cause of this differentiation is the error within the detector and so all of this should be discounted

Much like in Figure 4-7 there is a central band which contains the majority of the data, unlike Figure 4-7 this band is very tight and there is no deviation in any spectrum from this band.

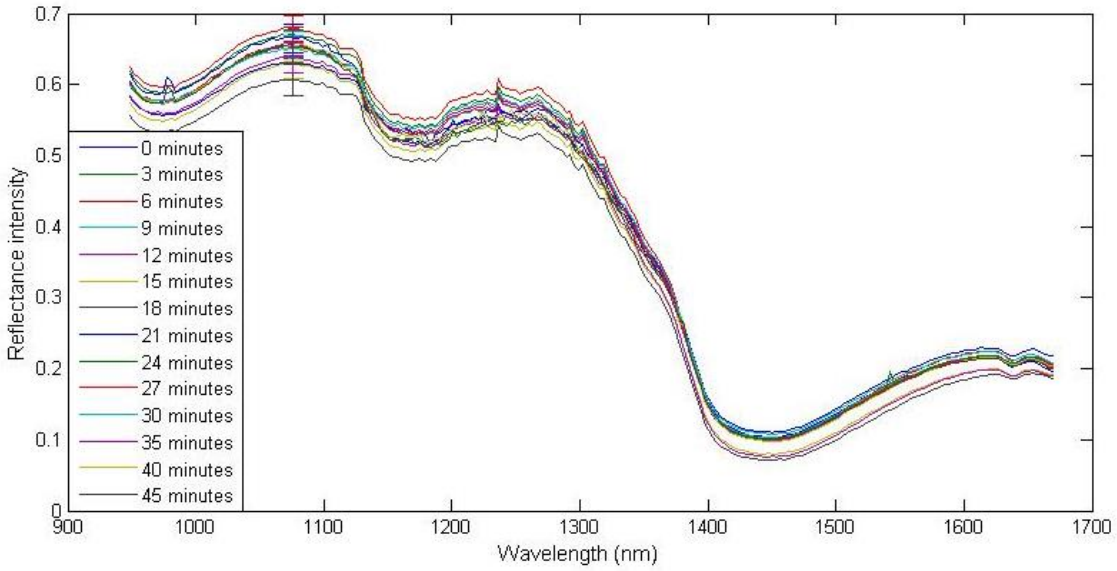


Figure 4-10 A plot showing an average spectrum (n=100) from each time point in the dissolution of Tablet B

It is possible that the speed of the disintegration and expansion of tablet material shown by this tablet causes the release of any APIs to be faster than the camera can detect as no visible changes over time are being seen in the tablet spectrally. The error bars in Figure 4-10 represent the mean standard deviation of each spectrum. These are between 0.015 and 0.025 which is similar to the standard deviation of the other tablet spectra which indicates variation within a single image is caused by noise. The expansion of the tablet particles in the flow cell was measured, as with Tablet A, and the percentage of the flow cell containing tablet is shown in Figure 4-11. The spectral analysis showed very little changes through time, something which is also shown by the expansion analysis. There is a fast expansion state for the initial 3 minutes of the dissolution which is preceded by a plateau and then a decrease in intensity after 21 minutes.

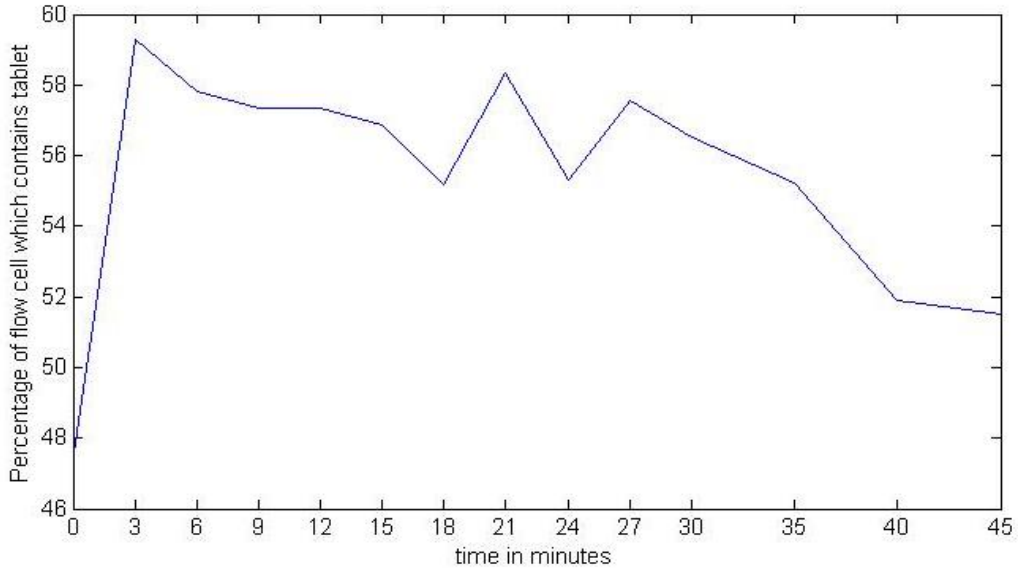


Figure 4-11 The expansion of tablet particles from the dissolution of Tablet B as a function of total flow cell coverage. The time points reflect the time the data was captured

4.3.2.3 Tablet D

Tablet D is the third non-caffeinated tablet formulation that was imaged. Figure 4-12 contains a selection of images from the tablet dissolution, the broad time points were chosen as they best show the slow progress of the tablet which is in stark contrast to Figure 4-9. Even by 45 minutes there is still a clear shape to the tablet as it has expanded but is showing no visible signs of braking down and disintegration.

Tablet D is a cheaper brand of paracetamol tablet and contains fewer excipients than the two previous tablets, the UV showed that the release was much slower that Tablets A and B with it taking almost 30 minutes to release 80% of its paracetamol. The images show that the expansion of the tablet after being submerged in the acid solution is also substantially slower than Tablets A and B which correlates with previous results showing that disintegration, expansion and release are strongly correlated. Also visible in this image is the difficulty of imaging with this setup as the flow cell has clearly moved between the images. The reason for this movement is motion of the stage caused the tubing to become stretched and generating a pulling force onto the flow cell which could not be completely controlled.

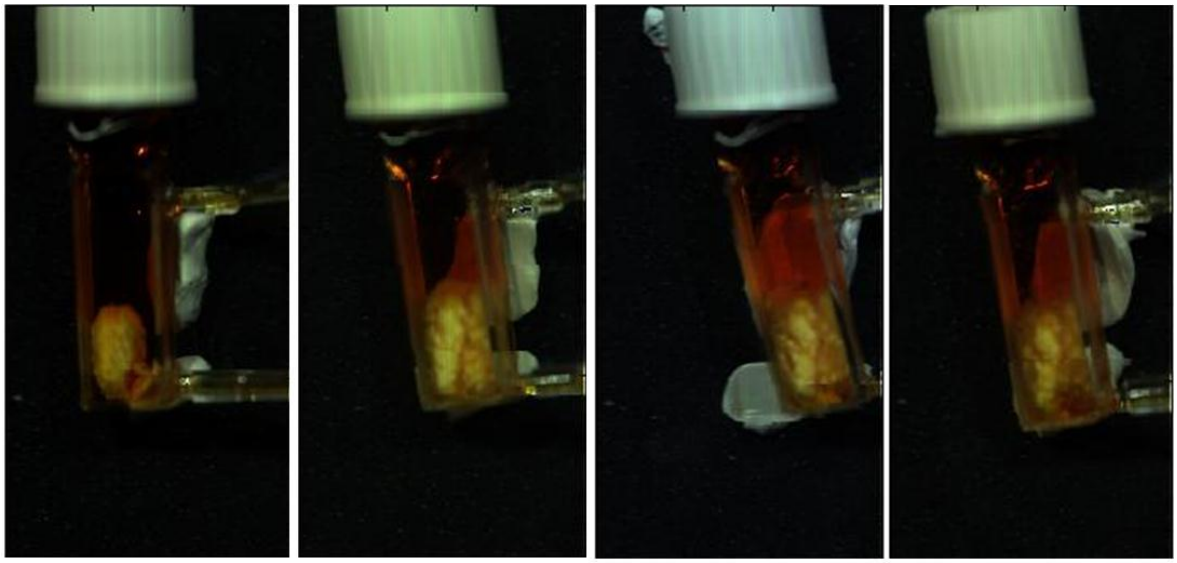


Figure 4-12 The dissolution images from the dissolution of Tablet D. The images are from the 0, 15, 30 and 45 minutes points of the experiment

One of the few notable changes in the tablets in Figure 4-12 is the presence of a darker region in the bottom right which is caused by the tablet braking down slightly from the force of the water being pumped into the flow cell from this location.

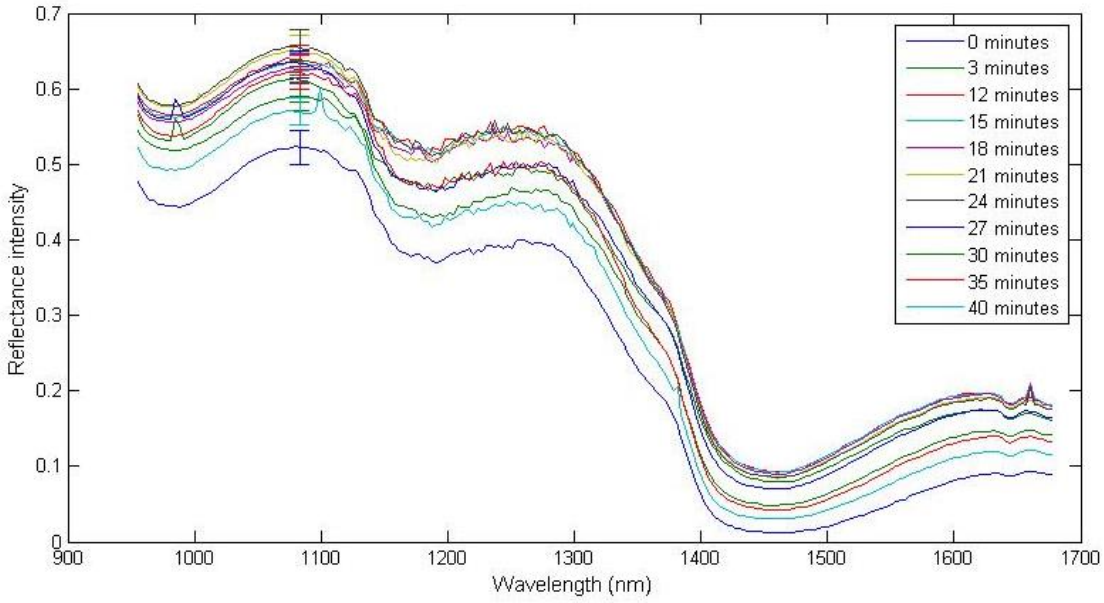


Figure 4-13 The mean spectrum (n=100) from each different image created during the dissolution of Tablet D

As before a mean spectrum was created for each tablet image and the spectra plotted to make Figure 4-13. As with all of the previous data there is a similar shape and distribution to the spectra with the majority of the data being within a very small band. Also similar to

other data the mean standard deviation are shown using error bars and have values between 0.012 and 0.024. However, the data from the first 15 minutes can be tracked in Figure 4-13 as the 0 minutes is the lowest intensity spectrum and the proceeding intensities are that of the 3, 12 and 15 minute data. This implies that from an initial time point the intensities are slowly increasing until equilibrium is reached with the intensity deviating only slightly around a fixed point. Figure 4-7 and Figure 4-8 also showed a similar pattern but with a much more rapid approach to this equilibrium state, this implies that the same process is giving the same results from these tablets.

A loss of paracetamol should result in a decrease in signal intensity at key locations not a uniform increase across the spectrum. The increasing intensities imply that this initial change is a structural one. These tablets are designed to swell to facilitate the release of the APIs contained within so this physical change could be the swelling of the tablet however, without more evidence it is not possible to conclusively say it is swelling and not expanding post disintegration. The expansion of tablet material as a function of total flow cell coverage over time is shown in Figure 4-14. The first 3 minutes show a rapid increase in tablet size as it expands and fills the flow cell. After 3 minutes the tablet continues to expand but at a slower rate for the remainder of the dissolution. Unlike Tablets A and B there is no plateauing to the expansion of the tablet material with a constant change throughout the dissolution. This result agrees with the UV dissolution which shows a fast release for Tablets A and B and a slower but more constant release for Tablet D. These results agree with the theory that the early changes to the spectrum are from the changes in the physical state of the tablet, more concisely the degree of expansion in the tablet.

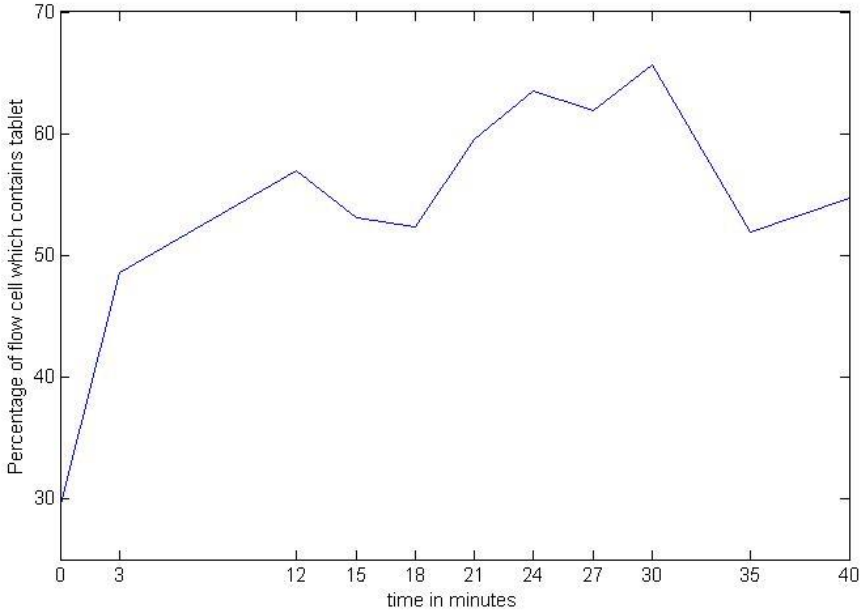


Figure 4-14 The expansion of Tablet D during the dissolution experiments as a function of the percentage of flow cell containing tablet data.

The tablet expansion and particle migration data from Tablets A, B both show a rapid increase from the initial data before any disintegration or expansion has occurred to a more constant plateau from the point at which the expansion has reached a maximum level. Tablet D is constantly expanding with tablet material moving along the cell but the rate is lower than Tablets A and B. There is also evidence of the particles leaving the flow cell of Tablets A and B towards the end of the experiments where a drop in the volume of the flow cell containing tablet data can be seen.

The spectral data from the tablets shows that most of the data lies within a narrow band except the earliest time points which often have an intensity below that of this band.

By comparing the spectral data from the earliest time points before the narrow band is reached alongside the plots of the tablet expansion it is clear that the spectra are also indicative of the disintegration and expansion of the tablet. When the tablet is expanding rapidly the spectra increase in intensity and once the rate of expansion has plateaued so too the spectra collect into a small band of intensities.

4.3.3 Hyperspectral dissolution imaging of caffeine containing paracetamol tablets

4.3.3.1 Tablet E

The caffeinated tablets were imaged using the same protocol as the non-caffeinated tablets. Tablet E is a standard non-caffeinated paracetamol and a full set of the images from the hyperspectral dissolution imaging is shown below in Figure 4-15. The different states of the tablet can be clearly seen in these images; the tablet starts the dissolution as a solid tablet, unaffected by expansion or disintegration and then slowly expands as from the uptake of the acid solution into the tablet matrix. The tablet eventually loses its form and begins to move along the flow cell as a mass of tablet debris after the 18 minute mark. The UV dissolution in Figure 4-3 details a slow release of paracetamol by this tablet brand. The first 15-20 minutes being the slowest and after this point the release begins to increase in rate. This would coincide with the tablet losing its form and expanding and breaking down into multiple smaller pieces of tablet. These images show that the tablet expands at a rate between that of Tablets A/B and D and is an example of the ability of this imaging to follow the various states of tablet dissolution and to allow comparison between different tablets while also gathering a huge volume of separate useful data.

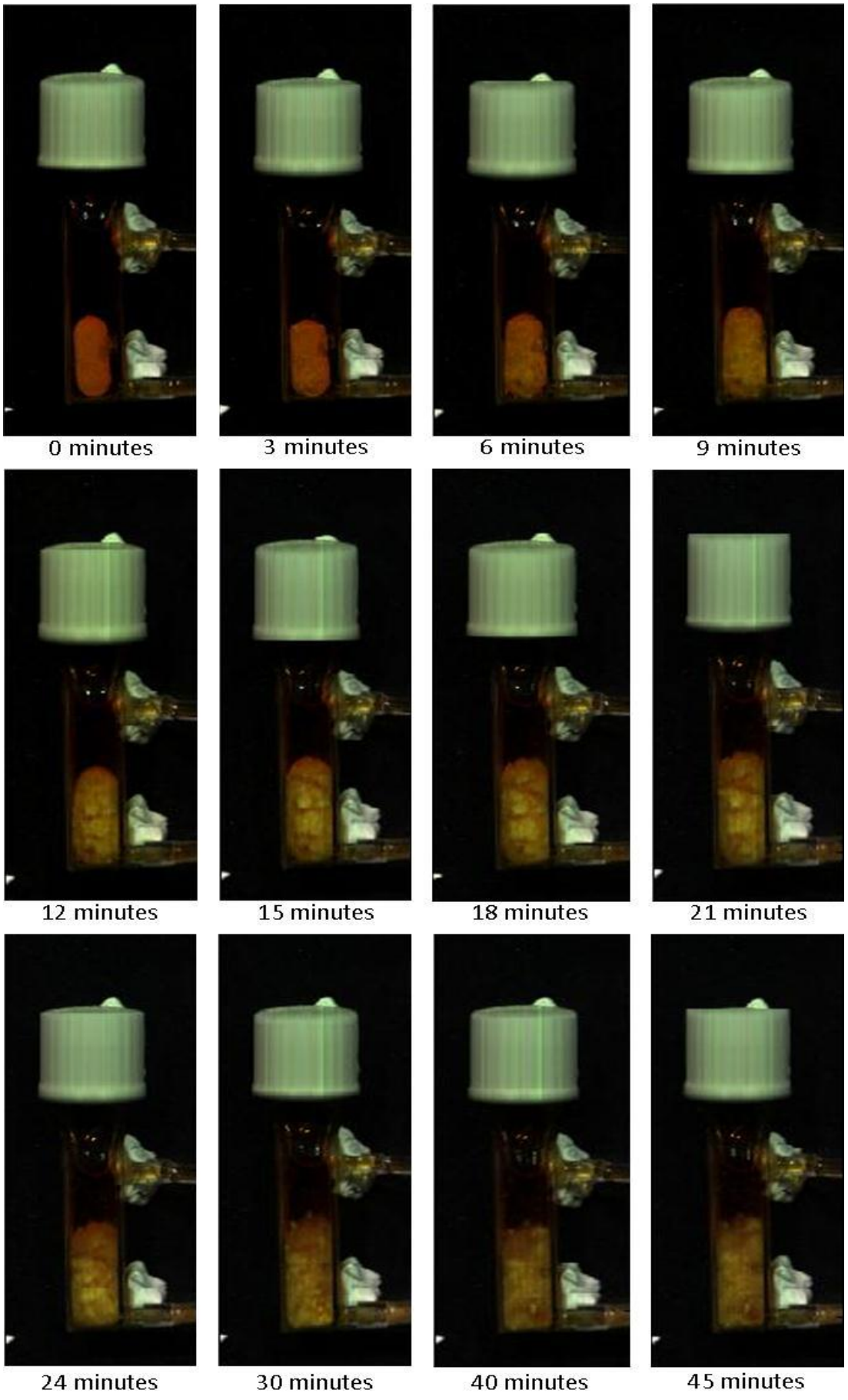


Figure 4-15 The full set of dissolution images from the dissolution of tablet E

The images of Tablet E's dissolution look similar to those of Tablets A, B and D. This is not an unexpected result as the caffeine is incorporated into the tablet for improved physiological effects not improved dissolution properties. The breakdown and expansion of Tablet E can be clearly followed as it begins to lose its shape inside the flow cell. The degree of expansion is shown in Figure 4-16 where once again the flow cell has been investigated and the number of tablet containing pixels identified. These were calculated as a percentage of the total flow cell and then plotted against time to show the movement of tablet materials through the flow cell during the experiment.

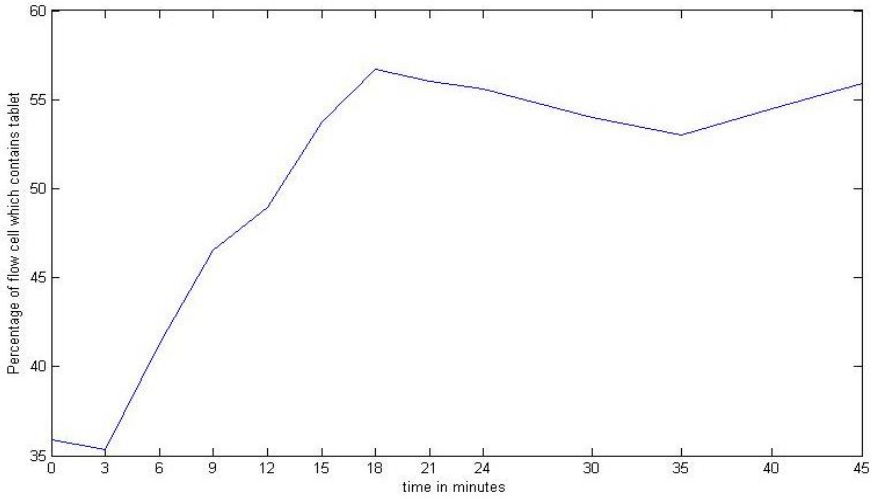


Figure 4-16 The percentage coverage of the tablet data within the flow cell during the dissolution of Tablet E

The expansion data shows a short lag period before the tablet begins to expand and tablet material move through the cell, once this period has passed the tablet expands at a linear rate until the 18 minute mark. The data after the 18 minute mark decrease slightly and then increase which is most likely showing natural fluctuations in the tablet particles locations.

The mean spectrum was calculated for each of the images in Figure 4-15 and then plotted to create the data shown in Figure 4-17. Again the standard deviations are shown in the image a mean value for each spectrum and the spread of these standard deviations is between 0.015 and 0.028 which is similar to the non-caffeinated formulations. The data here appears to be following the same trend as the non-caffeinated tablets of increasing signal intensity as the experiments progresses. However, it is unclear exactly which time points are causing the highest and lowest intensity bands due to the number of different spectra.

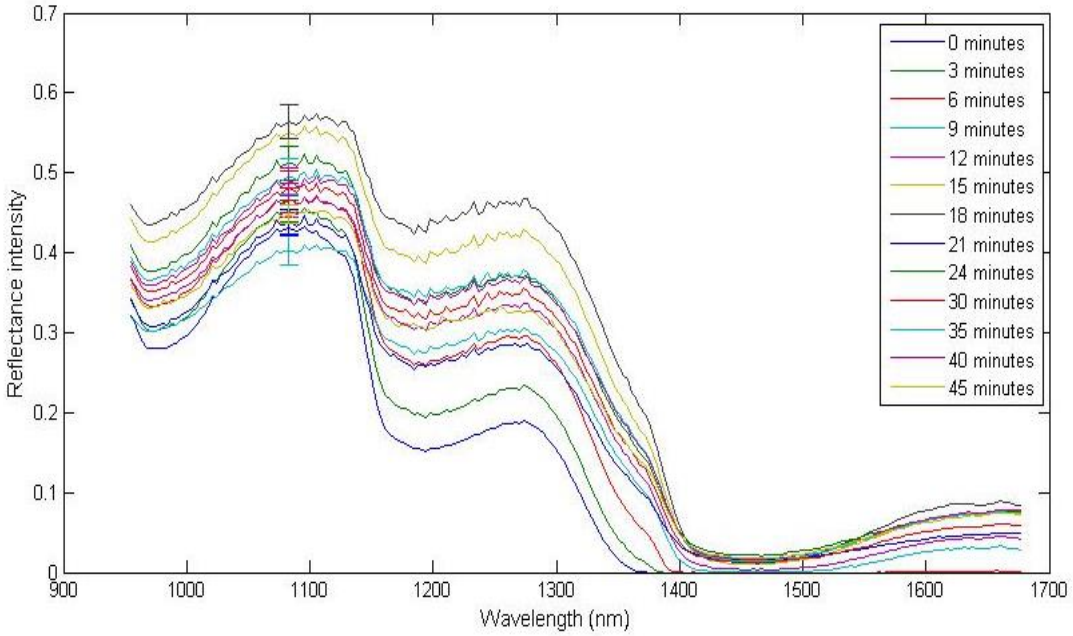


Figure 4-17 The average spectra (n=100) collected from the dissolution of Tablet E

The lowest intensity spectrum, at all points expect near 1153 nm, is from the 0 minutes image. The highest intensity spectrum is from 18 minutes which shows a large spread of data within the first 18 minutes and the spectra collected after 18 minutes sit in the central band in Figure 4-17 with very little variation. To better show this effect the first 7 spectra were re-plotted and the result is Figure 4-18 where the progression of the spectra can be easily followed.

The tablets expand and begin to breakdown during the dissolution process, this expansion and subsequent distribution of tablet particles throughout the cell is shown in Figure 4-16. The tablet material expands during the earliest stages of the dissolution, the data in Figure 4-16 shows this to occur for 18 minutes. The intensity of the tablet spectra is increasing during this same time period, but the increase is not occurring in a linear fashion with the distance between spectra changing as the experiment progresses. One explanation for this is the physics of how these tablets are expanding and breaking down. The intensity increases for every spectrum in the first 18 minutes however, inspection of the data from 6, 9 and 12 minutes shows that it these spectra are located within the central band which contains the majority of the spectra in Figure 4-17. This implies that the cause of the highest intensities is something which is reached briefly but a change causes the spectral intensity to decrease back down to the level of the central band once more. Figure 4-18 contains data from the first 18 minutes of the experiment which Figure 4-16 shows is the same time period as when the tablet expansion is occurring. The majority of the data in the

narrow central band is from after the tablet has stopped expanding. The exception to this is the data from 6-12 minutes. The tablet is expanding during these 18 minutes and so the change in the spectra is most likely a function of the height of the tablet surface as it expands. The taller the tablet, the shorter the path length of the photons though the hydrochloric acid solution which will improve signal strength. If the spectral intensities are showing a measure of the degree of expansion upwards in the flow cell then the height of the tablet surface in each of the results contained within the central band must be similar. This implies that the tablet is expanding upwards to a maximum height which is reached at 18 minutes. Once this is reached the tablet disintegration is a greater force than the expansion of the tablet and the particles settle down to a height equal to the 6-12 minute region thus the corresponding spectral intensities. While the height of the tablet during the experiment cannot be determined in these images due to the orientation of the camera it is possible that the breakdown of the tablets occurs after 18 minutes and the tablet materials settles at a height equal to that of the tablet material at the 6^{th} to 12^{th} minutes. Change in the height of the samples can explain why it is that the fluctuations in the spectrum can be seen.

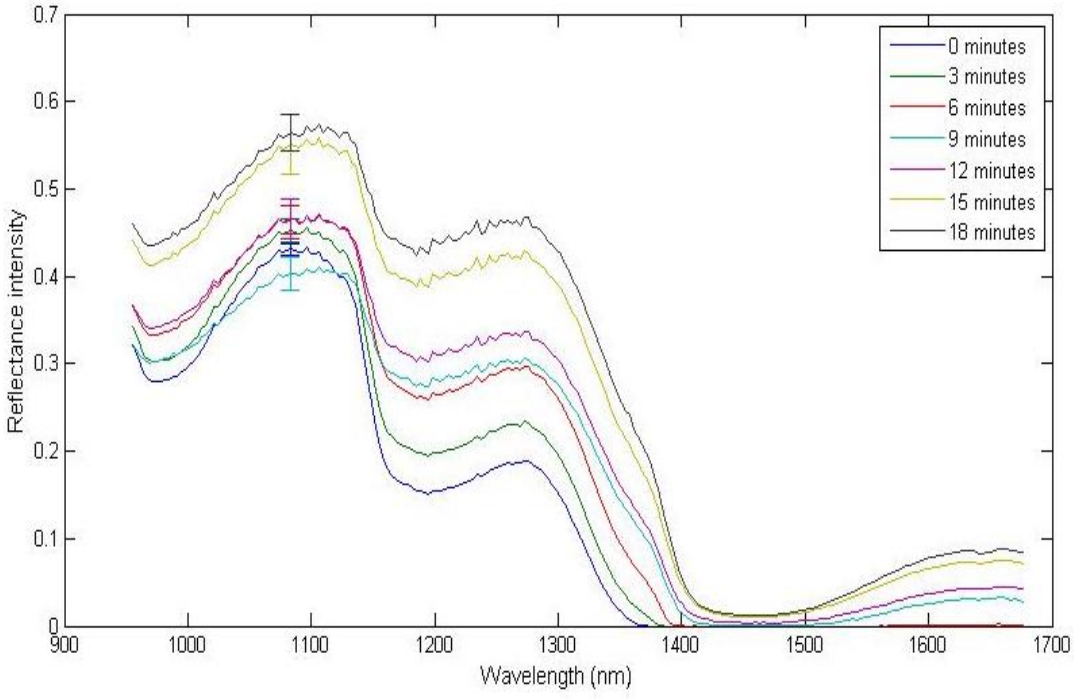


Figure 4-18 The mean spectrum (n=100) from the first 7 HSI dissolution images from Tablet E

This result shows why the expansion data and the spectral plots used in conjunction are very powerful. The expansion plot shows that the tablet expands for 18 minute and then

stops expanding and stays constant. The spectral data shows a change for the first 18 minutes and then a reduction in spectral intensities to reach the central band which is common in these plots. The cause of these results cannot be determined fully however, it appears that the tablet is expanding both along the flow cell and upwards in the flow cell until it reaches a point near the top of the cell. After this point the tablet collapses down to a height equal to that experienced at approximately 9 minutes into the dissolution. While this process is probably the swelling of the tablet before the disintegration begins this cannot be proven, it is also possible that the tablet has disintegrated already and is simply expanding in all direction as a mass of loose tablet particles. The change in the environment of the particle and their size will affect their light scattering properties and subsequently the intensity of the spectral signal. The expansion data does not show this collapse because it is determining how much the tablet material has moved through the flow cell, a 3 dimensional process, using only a 2 dimensional image. This is a weakness with this approach which the spectral data helps to compensate for.

While the caffeinated paracetamol spectrum differs from the earlier non-caffeinated it is a small change and the general form and progression of the data through the dissolution appears to be the same. This result is repeated again in the other caffeinated drugs (the plots can be seen in the appendix) and so it seems that there is no way to use the spectrum to simply watch changes in the composition of the tablets as a function of the changing spectrum over time.

There is also a clear change in the colour of the images in Figure 4-15, these colours are a function of the intensity at three separate wavelengths (1083, 1230 and 1601 nm). A change in the colour is caused by a change in the relative intensities at these three wavelengths. In the first 6 minutes the tablet appears red in the image however, from 9 minutes the tablet appears orange. Inspection of the Figure 4-18 shows that the ratio of the peaks at 1083 and 1230 nm is not constant and that the intensity of the 1230 nm peak increase more rapidly than the 1083 nm peak. This is best seen by inspection of the 0 and 12 minute spectra, the peak at 1083 nm has only increased by approximately 0.05 while the 1230 nm peak has increased by approximately 0.10. It is not clear what is causing the change in these results, but there is clearly some affect which is causing the spectra to change. The change may be caused by the chemical composition of the tablet being altered during the dissolution (loss of API or excipient) or some other physical factor. As the intensity of the entire system is increasing during the expansion phase it is likely that the

spectral change seen is a function of the wavelength dependency of this process which is causing some wavelengths to change more rapidly than others.

4.3.4 Principle Component Analysis

The basic spectral analysis can be used to track the physical changes occurring during the dissolution however, it has not been possible to distinguish any specific changes in the spectra of the tablet which can be attributed to the loss of API or excipient. By using principle component analysis (PCA) small variations between the data which may be difficult or impossible to spot by eye can be discovered. These small variations could be the release of API from the system and this would show the effectiveness of HSI as a novel approach to dissolution testing.

The PCA of Tablet A is shown below in Figure 4-19 and it uses the mean spectral data from Figure 4-7, each mean spectrum is a single point in the scores plot. In Figure 4-19 the 3 minute data is far removed from any other data points and this indicates a severe variation between their spectra and the bulk of the data. By investigating Figure 4-7 it is clear that the 3 minute spectrum is dissimilar to the majority of the data and thus the separation in the PCA.

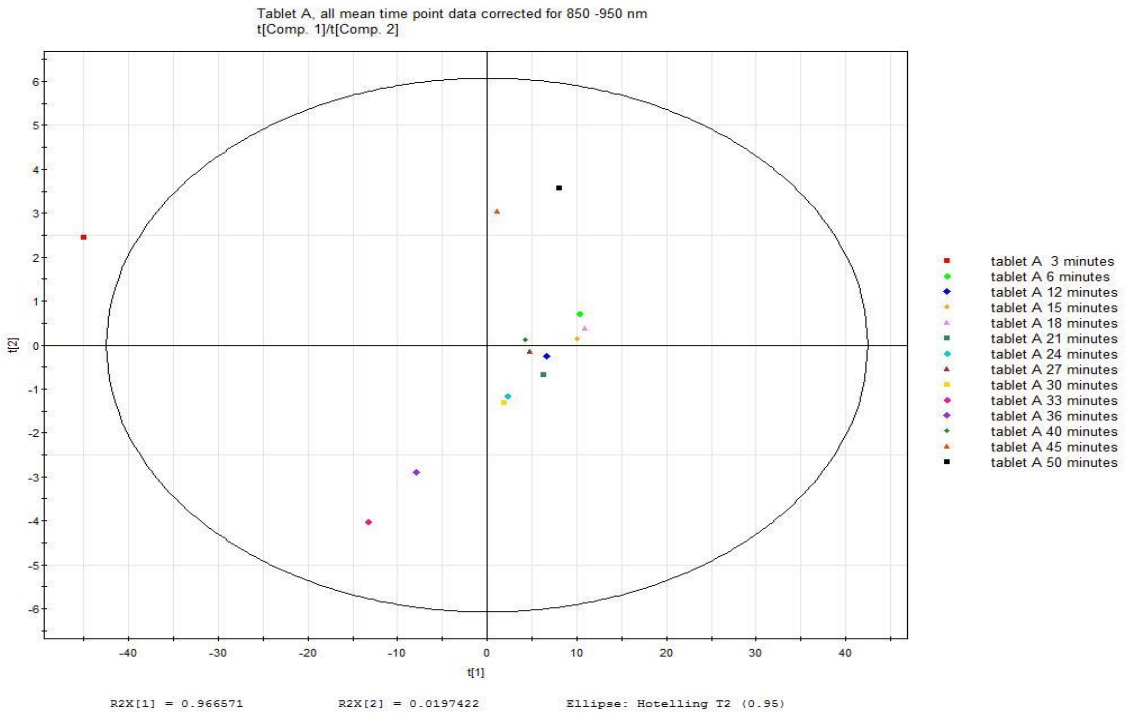


Figure 4-19 A scores biplot generatered from the first two components of the PCA of the dissolution of Tablet A

The majority of the data is centred on the origin of the plot but there are some outlying groupings which can be seen. At the top of plot the data two points can be identified as being that from the 45th and 50th minute spectra. Two points can also be seen in the bottom left which are the 33th and 36th minute data. If you imagine a line of best fit through these 4 data points then two things are of note, firstly the data points line up chronologically so that the earliest is at the bottom left and the latest is at the top right. The second point to note is that the green square denoting the 40 minute data is also near that line but appears to be clustered within the remaining earlier data. The loadings for these principle components should help to show what is causing this linear escalation of sample intensities from the bottom left quadrant to the top right. The loadings detail which variables, in this case wavelength, have the greatest effect on scores data. A large peak in the loadings at a specific wavelength infers that this wavelength causes the variation depicted in the scores plot. The loadings for the first principle components is shown in Figure 4-20A and the second component is plotted in Figure 4-20B. While the intensities of these loadings are very dissimilar the shape of the second loading plot is close to the inverse of the first. A positive value in the first component is caused by a variation in the same location as negative value in the second component.

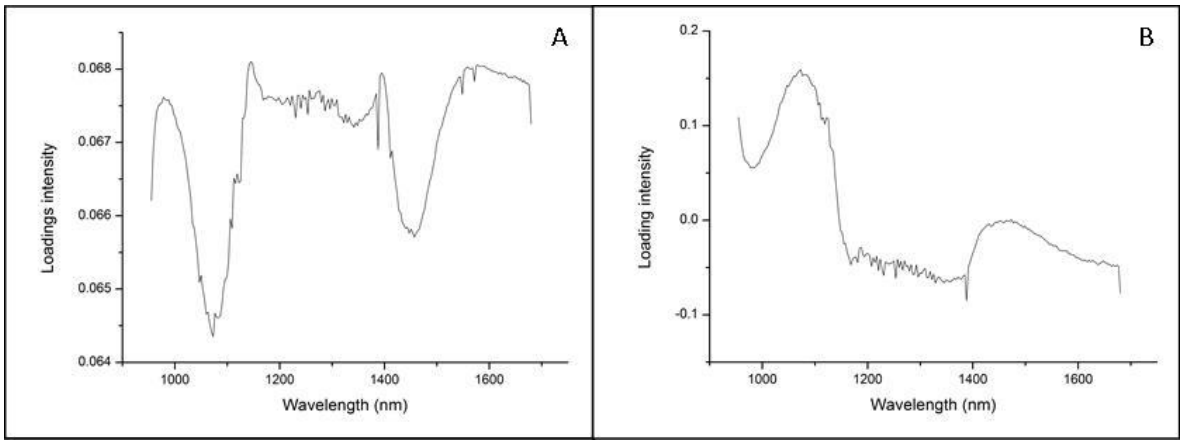


Figure 4-20 Loadings from the scores plot in Figure 4-19 with the 1st component shown in A) and the second in B)

The major regions of interest in these loadings are the 1000 - 1150 nm and 1400 - 1600 nm regions as these give the strongest change in the data. The lower the intensity of the spectrum in these regions the lower (or more negative) the value in the scores for the first principle component and he higher the intensity the larger and more positive the scores value in the second principle component. So as the data from the 33rd minute was the most negative it follows that this spectrum should be the least intense with the 50 minute data

being the most intense. The spectral data from these regions is shown in Figure 4-21 and as the loadings have shown the spectra follow a pattern of increase intensity as time progresses which led to the spread of the data in the PCA scores plot.

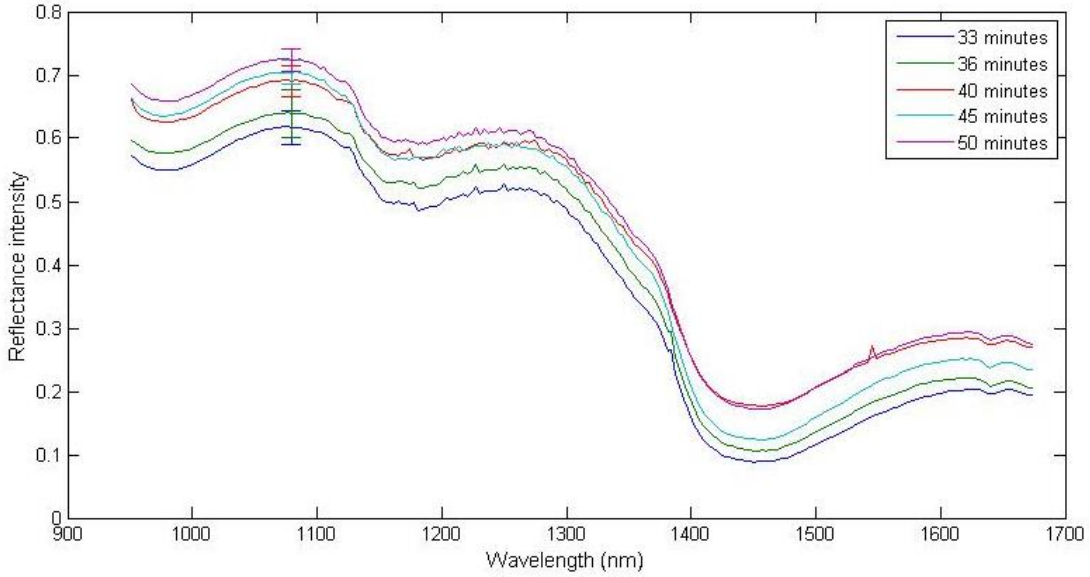


Figure 4-21 The mean spectrum (n=100) from the last 5 time points collected during the dissolution of Tablet A

This change in intensity is the same trend as was seen in early data with the breakdown and expansion of the tablet causing these changes in intensity. If this was a spectral change from the loss or release of API or excipient then it would be seen as a decrease in peak intensities at specific locations and not as an overall gain in intensity. This shows how an even greater degree of structural information can be determined from more advanced analysis of the results. This change is from either a secondary expansion phase or some other physical process which is mimicking the effects caused by the tablet particles expanding through the flow cell in all directions. These results agree with the expansion data in Figure 4-6 that more of the flow cell contained tablet data in the latter stages of the dissolution than earlier. It is unlikely that a second weaker expansion phase has caused this effect as starch would already be fully hydrated and swollen before 30 minutes into the experiment. It is possible that the breakdown of the tablet is becoming more severe and this is somehow causing the bulk, where the data was gathered, to show an increase in intensities. The location of the bulk of the tablet data in the scores plot (Figure 4-19) indicates that the majority of the data has a greater intensity than the data investigated here which may also help to explain the change. A substantial loss of tablet material could lead to a decrease in the spectral intensity at the 33rd minute, this was then bolstered by tablet

being deposited at the bulk location and the intensity again rose which resulted in the strange linear clustering seen in the score plot.

The analysis of this data shows that the PCA can be used to classify patterns of spectral change which could not be seen in the normal spectral analysis. This is due to the nature of the data in the tight band of spectra in which small intensity changes are almost impossible to track.

4.4 Conclusion

The aim of this experiment was to use hyperspectral imaging to follow the dissolution of tablets and show the physical and spectral changes occurring over time to these systems. To this end a full set of tablet dissolution was run to give a set of reference data for the speed of API release. It was hoped that the loss of, or location of, APIs within the tablet could be identified using this technique however, this was not possible. The UV dissolution data showed that the release of API from these systems often occurs quickly and the time frame for the HSI images was too long. Figure 4-22 shows the hyperspectral image from the dissolution of the four tablets discussed in this chapter at four different time points. The 0 minute shows how much can occur in the initial few seconds where the flow cell fills with the acid solution and the image capture is started. Most tablets remain whole, intact and show few signs of any change except Tablet B which is clearly showing signs of the tablet disintegrating. This details the difficulty of tracking the physical changes in these systems without a change to the time frame, some tablets react faster than the current image capture procedure. However, while the time frame needs to be reduced the length of the experiment does not. The 15 minute data shown in Figure 4-22 shows that some tablet formulations still have some a definitive tablet shape and show no signs of tablet material being dispersed through the flow cell. A new experimental; protocol is needed which modifies the experimental technique to lower the time frame between data capture without changing the dissolution time which is a central idea for the next chapter of work.

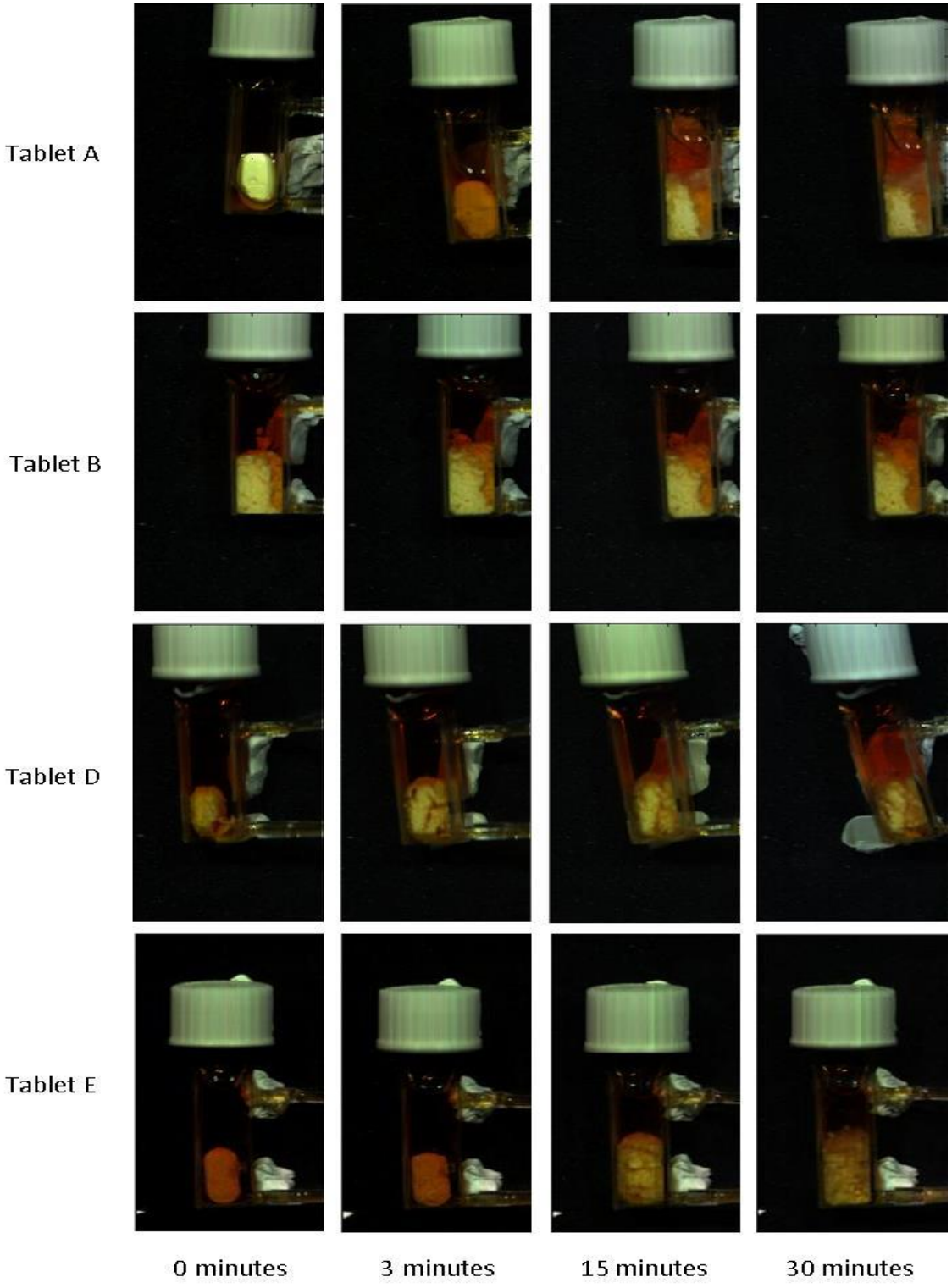


Figure 4-22 The hyperspectral images of the four tablets investigated in this chapter at 0, 3, 15 and 30 minutes

It was however, possible to closely follow the changes in the tablet on a structural level with both the image data and spectral data being of use in following the expansion and breakdown of the tablets over time. The rate of expansion of tablet material was also shown to be closely related to the UV dissolution result which implies that the experimental setup is not hindering the dissolution properties of the tablets. By using the

spatial and spectral sides of the HSI data it was possible to generate a more complete understanding of the degree and rate of the tablets expansion than would be possible with only one of them. The spectral intensities changes in the data are difficult to show as being a function of expansion of the tablet without the images and by using the spectral data the relative height of the tablet could be identified which could not be done using the 2 dimensional images of the tablet in the flow cell. The movement of the flow cell during the imaging was also a cause for concern as it lowered the effectiveness of the algorithms for tracking the expansion of tablet material. The fixing of the flow cell to the stage is important to prevent these outcomes recurring so a holder is needed for the flow cell. This is also a critical component of future work which is addressed in the following chapters.

The results in this chapter show that the breakdown and expansion of Tablet B is the quickest with Tablet A also showing a fast and constant expansion of the tablet. The other tablets imaged showed good release which was significantly slower than that of Tablets A and B which agrees with UV dissolution results. These results are as expected as Tablet A and B both contain a variety of excipients to aid in their release while the other tablets are more basic and contain fewer excipients.

In a traditional UV dissolution experiment there are none of these expansion and disintegration parameters that can be ascertained, only the rate of release of specific APIs. This shows some of the ability of the HSI to give a bigger and more in depth overview of the processes occurring in the tablets during the stages of dissolution.

The hyperspectral imaging of the tablet dissolution showed a great deal of information about the expansion of the tablet material however, the time period between image capture is too long. The majority of the physical changes and the release from the tablets can be seen to occurring rapidly in both the HSI images and the UV dissolution data. Often there is little more than a single image during the expansion stages and in the case of Tablet B there are no images that cover the start of the expansion process. There is more information which can be gathered in this region as this is where the majority of the physical changes are occurring and so a better time resolution is imperative for more accurate analysis of these images.

Chapter 5: Imaging tablet dissolution using high temporal resolution.

5.1 Introduction

There is very little literature on the hyperspectral imaging of pharmaceutics in any form[14,130,152] and what little there is concentrates on investigating surface concentrations, purity and mixing efficiency in solid tablet forms. None of the literature contains any time dependency or anything more complex, from an imaging viewpoint, than simply taking one picture per sample and analysing it. This is not the case for this work as following the dissolution of a tablet requires some way to track the changes over time in a dynamic system. As far as we are aware this work represents the first use of hyperspectral imaging in pharmaceutical analysis to track a dynamic process occurring in real time within a liquid flow system.

In Chapter 4 a series of traditional hyperspectral images were taken, these made full use of the spatial information to show both the rate of expansion of tablet particles and the process of tablet breakdown. However, due to limitations with the software the images were captured once every 3 minutes and this is insufficient to accurately follow the release of the active pharmaceutical ingredient (API) from the tablet. The majority of the expansion and release is primarily completed within 5-10 minutes and so there was very little information captured between the initial whole tablet and when the expansion of tablet material has ceased. To improve the temporal resolution a novel approach was needed, one that would give faster image acquisition while still retaining the spectral fidelity of the hyperspectral data. The solution is to sacrifice some of the spatial information in the images to acquire a substantial improvement in the time resolution with data acquisition changing from 1 per 180 seconds to 1 per 100 milliseconds.

5.2 Materials and Methods

5.2.1 Materials

The six different paracetamol brands used in the previous chapter are Panadol Advance, Panadol Actifast, Tesco paracetamol, Tesco paracetamol extra, Boots paracetamol extra and Panadol Extra. The tablets are evenly distributed between caffeinated and none

caffeinated varieties which allows for the potential of an interesting set of comparative results about the expansion of tablet material and release rates of the two different kinds of commercially available paracetamol drugs. They are also a range of prices and complexity between the tablets, with some formulations containing a much greater number of excipients to help with the tablet function. The exact tablet formulation can be seen in the Chapter 2.4.2, the major constituents for each tablet are the active pharmaceutical ingredients and starch which is used as a filler and binder. A small number of common excipients are present in most tablets such as stearates and polyvinylpyrrolidone, these are used to improve the breakdown of the tablets. The one tablet formulation that contains a major excipient that is unique is Panadol Actifast which contains a large quantity of sodium bicarbonate to aid in the breakdown of the tablet. For ease each of the tablets is referred to by its code which is shown in Table 5-1.

Name of tablet	Panadol Actifast	Panadol Advance	Panadol Extra	Tesco paracetamol	Tesco paracetamol extra	Boots paracetamol extra
Code for tablet	Tablet A	Tablet B	Tablet C	Tablet D	Tablet E	Tablet F

Table 5-1 The naming codes used for each paracetamol brand

5.2.2 Experimental setup: Use of a single line

To circumvent the problems in Chapter 4 the 30 minutes of the dissolution were saved to a single file. This would mean there is a single 30 minute image with all the data contained within. While it was a relatively simple procedure to keep the image running for this length there was no way to move the stage to build up the images. This means that the images produced are from a single spatial line across the flow cell which is imaged constantly for the duration of the experiment.

The near infrared (nIR) camera used in Chapter 4 was used again as this wavelength range was shown to have the better potential for useful results than the HSI systems which are sensitive to visible light. The use of nIR is common place in pharmaceutical analysis due to it being non-destructive, fast and easy to use [148,149]. The camera used for the dissolution imaging was a Xenics Xeva InGaAs CCD attached to a Specim V17E spectrometer which

has a wavelength range of 850 – 1680 nm. The stage was a Zolix KSA 11- 200 S4N with an IMS Mdrive 17+, while the stage was not used for movement it was used to accurately position the flow cell in the same location for each experiment. At the start of each day and after any movement of the system it was refocused using a simple checkerboard or diamond pattern as described in Chapter 2.1.3. The lights were turned on at least 10 minutes prior to any imaging being performed so that they had time to equilibrate the output of infrared photons. Any imaging before the lights have been able to warm up could easily give results with weaker or varying intensities.

The dissolution of the paracetamol tablets was performed as in the previous chapter by pumping a pH 2 HCl acid solution through the flow cell with a constant flow of 10 ml min^{-1}. The flow cell from the previous experiments was again used for these experiments. The flow cell was placed into the holder, described below, and then locked in place onto the camera's stage as pharmaceutical analysis requires a high degree of repeatability between measurements. This is not normally an issue but when imaging a single line of height ~300 µm it is not possible to guarantee repeatability of the location being imaged by the camera without something to position the target. A holder was designed which could be bolted onto the stage to prevent any movement and which would keep the flow cell positioned at exactly the same location. Once the holder was attached to the stage it was moved to the stage's starting location, 0 mm movement from home location, and then moved forward by 90 mm and this process is illustrated in Figure 5-1.

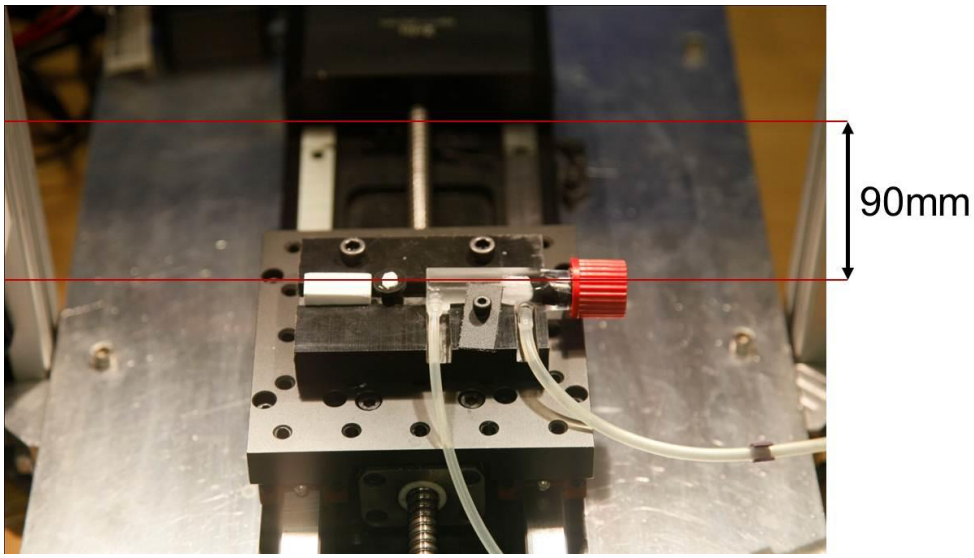


Figure 5-1 Picture showing the location of the holder and how the motorised stage is used to position flow cell for each experiment.

5.2.3 Experimental setup: Experimental procedure

The hyperspectral dissolution imaging system was configured to capture a new line of data every 100 ms for 30 minutes for a total of 18000 lines of data. Each line had an exposure of 90 ms which gave a maximum intensity to the data without the detector becoming saturated. The sample to be imaged was placed into the flow cell which was itself placed into the flow cell holder. This was then bolted onto the stage and moved into position so that the imaging could begin. Once the flow cell was in place and all connecters had been checked the imaging would being and the peristaltic pump would be switched on so that the flow cell began to fill with the acid solution. After the experiment was completed the flow cell was emptied and cleaned with distilled water along with the acid reservoir. Distilled water was then pumped through the tubing to remove and tablet material which remained.

Before the experiments were run a set of reference spectra were collected to show the reflectance signals from the various components of the system before the dissolution experiment. The measurements are of the empty flow cell, the flow cell filled with acid solution, the white reflectance calibration tile and the dark reference (which was taken by placing the lens cap on the camera and switching the lights off).

Before the tablet dissolution imaging began a set of spectra were also collected for the major excipients and active pharmaceutical ingredients (APIs) present in the tablets. There are a wide range of excipients used across the tablets so the most common ones were used which are starch, polyvinylpyrrolidone (PVP) and magnesium stearate. These images were created by simply inserting a large quantity of the powdered compound into the flow cell and then letting the flow cell fill up. This generated analogous data for the excipients both before the dissolution and during which could be compared to the tablet data as a way to analyse and assign any spectral changes.

Each tablet type was imaged three times to ensure that there was no large variation between the images of the same tablet types. Any substantial intra tablet variation with respect to the tablet spectrum or dissolution behaviour would indicate an area of concern. If these images show no major variation in these areas then the imaging setup can be seem as repeatable with respect to the conditions within the flow cell.

5.2.4 Experimental setup: The tablet holder

It is imperative in pharmaceutical analysis that the procedures by highly repeatable. These experiments image a single line across the flow cell and so it is necessary to create a holder or cradle which can be used to position the tablet in the same spatial location each time the experiment it run. Figure 5-2 is a picture of the flow cell which has been placed into the holder which was made using Acetal Polyoxymethylene copolymer and was designed to perform 3 different functions. The flow cell is placed into the deep groove which is machined to be an exact fit with no room for lateral movement. Above the flow cell two bolts can be seen, these are used for attaching the holder to the stage. In this manner the holder can be guaranteed to be in the exact same location on the stage for each experiment; the bolt holes in the stage were also marked to ensure the same ones were used each experiment.

To the left of the flow cell the holder contains a space for a white block, and a small elliptical white object. These are used as calibration aids in the images, it is important to show that there is no change in the imaging environment throughout the experiment. To show this the two reference material are used should never change throughout the experiment. The white block is a piece of the reflectance PTFE (polytetraflouroethylene) calibration tile which is used to convert the raw data collected into the corrected reflectance data. In the final images this should always have a value of 1.0 (± 0.05 for machine error). The smaller white object encased in a piece of black tubing is a tablet, of the same make as that the one in the flow cell, which has been split in half. There are two uses for this; firstly it should be a constant signal throughout the experiment as there is no interaction between this tablet and any acid solution. Secondly it gives a reference spectrum for the beginning of the experiment before any dissolution has occurred and also shows any change in signal strength which is occurring as the photons pass through the glass.

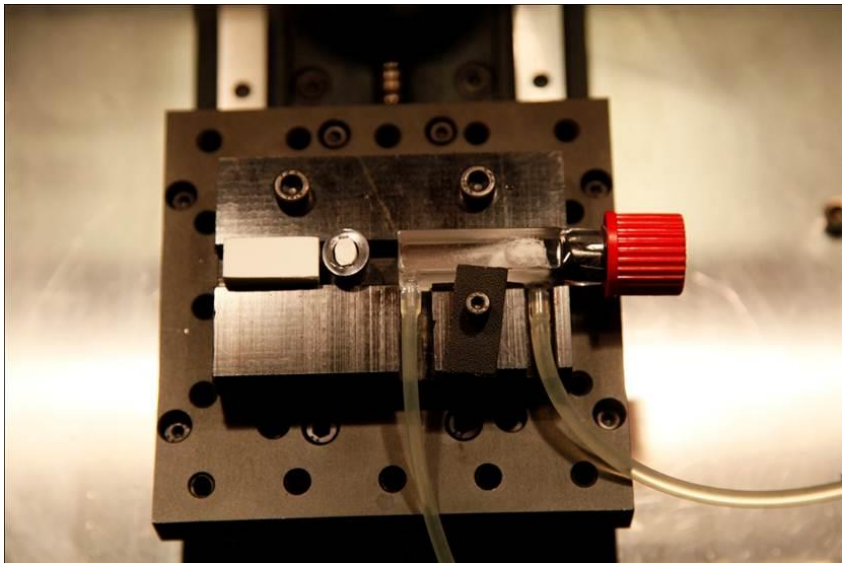


Figure 5-2 The flow cell inside its holder which has been bolted onto the mobile stage.

5.2.5 Corrections

Once the image capture was completed a set of calibration images were captured using a white PTFE tile for the 100% reflectance and a 0% reflectance image was recorded by switching off the lights and placing the lens cap onto the camera. These two values are used to convert the raw data which is expressed as a number of counts into the reflectance of the nIR radiation from the sample.

The calculation for determining the reflectance of any point is shown below

$$R_{s\lambda} = \frac{I_{s\lambda} - D_{\lambda}}{I_{o\lambda} - D_{\lambda}}$$

Where $I_{0\lambda}$ and $I_{s\,\lambda}$ are the intensity (in counts) of the 100% reflectance standard and the target pixel respectively at a set wavelength λ. D is the intensity of the dark or 0% reflectance standard and $R_{s\,\lambda}$ is the reflectance value for any specific pixel in the image at a set wavelength. This data correction was performed using an inbuilt command in the imaging software.

5.3 Results

5.3.1 Creation of line scan results

The hyperspectral images created in Chapters 3 and 4 contain a more traditional image in which a single snapshot was taken of a sample with x and y spatial dimensions and the λ spectral dimension. The results created using this novel HSI imaging process are quite different; a single line across the cell has been imaged repeatedly for 30 minutes thus the x and y dimensions are not both spatial. The x dimension contains the spatial information across the cell, but the y axis shows the detail in the temporal domain. The changes in the flow cell are characterised by the evolution of the data at a set x location as the data proceeds down the y axis. The image in Figure 5-3 has two parts, the top shows the tablet in the flow cell and the red line depicts the location that the camera is imaging.

The bottom part to Figure 5-3 is a brief section from the hyperspectral dissolution imaging from Tablet F, the numbers on the y axis show the time in seconds for the first 40 seconds of Tablet F's dissolution. The two white bands on the left of the image are the reference materials placed inside the holder and to the left of the flow cell. The leftmost band is the calibration tile and the second darker band is the broken tablet. The yellow band is the tablet, the very first line of data is at the top of the image and every successive line of data is 100 ms after its predecessor. At 20 seconds into the dissolution there is a clear artefact created within the data which is an air bubble moving both in line with and perpendicular to the imaging. This generates the slanted left to right angle to the bubble and also explains how it can appear and then disappear in the centre of the tablet mass. This example details the power of using such a high temporal resolution approach to follow the dissolution of tablets as small fast moving objects like bubbles can be tracked accurately.

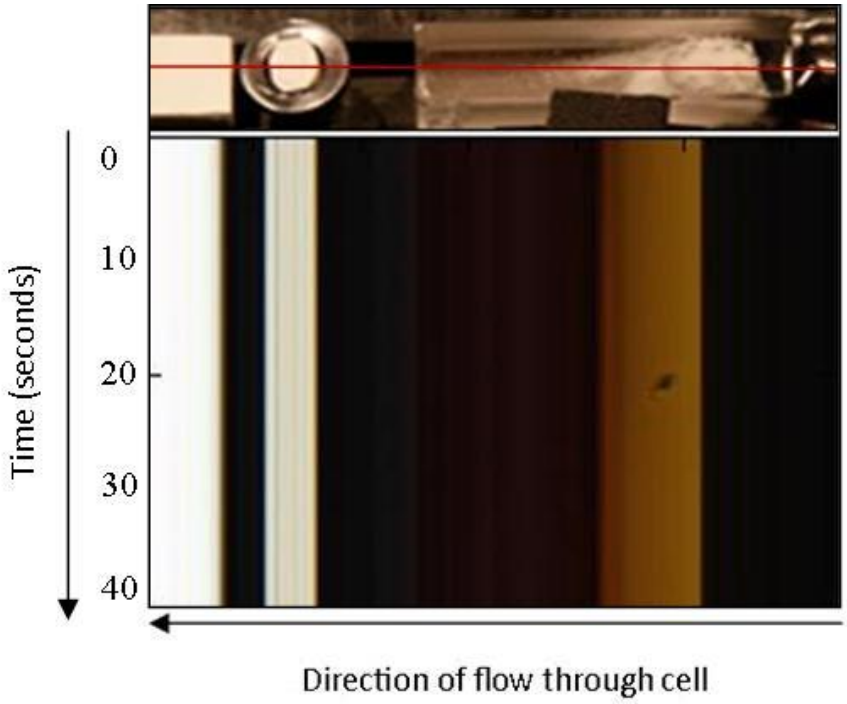


Figure 5-3 An illustration of how the hyperspectral images are created during this imaging approach. The top image is a standard picture, captured with a digital camera, of the experimental setup and the bottom image is a short piece of the corresponding hyperspectral dissolution image. The red line in the upper image denotes the location at which the camera collects the hyperspectral data.

The data in hyperspectral images is often output in grayscale where the intensity of one wavelength is used to denote the intensity of each pixel. However, these images have been shown in a false colour to better detail any changes in the spectra over time. To generate this false colour image three separate wavelengths are used and the intensities in each pixel at these wavelengths denotes the intensity of the false red, green and blue colour at that pixel. The wavelengths need to be chosen carefully or any valuable changes in the spectrum could be missed. Figure 5-4 is a spectra gathered by averaging the spectra of a 10 by 10 region of pixels within the tablet. This spectrum shows the shape and form of the average tablet spectrum without the noise often found in individual spectra. The false colouring used the wavelengths of 1073 nm, 1279 nm and 1407 nm to provide the intensities for the false red, green and blue colours. These regions were chosen as being the maxima of the one peaks, a saddle position in the data and also a minimum. Any change in peak size, locations or the minima should be shown as a perceivable change in the colour of the data.

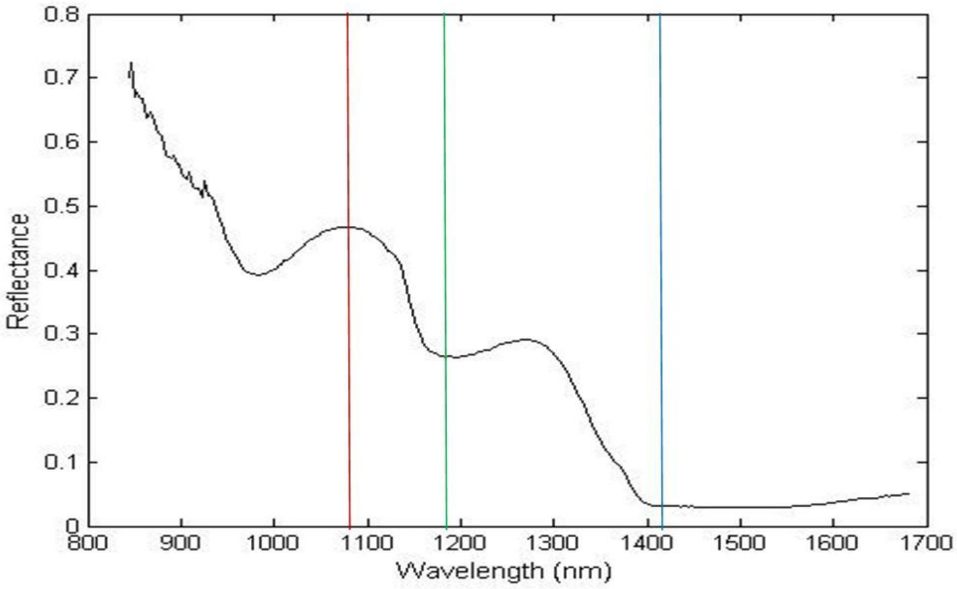


Figure 5-4 An average spectra generated from viewing a 10 x 10 region in the centre of the tablet mass. The three coloured lines indicate the wavelengths used to create the false colour HSI images

5.3.2 Analysis of spectral data from tablet dissolution

5.3.2.1 Tablet F's hyperspectral dissolution images

The use of false colours allows for quick inspection of the data and has the potential to yield interesting results dependant on the spectral changes and movement of the tablet through the flow cell.

Tablet F contains a number of excipients and two major components, the release of which can hopefully be followed throughout the experiment. These APIs are paracetamol and caffeine with the paracetamol being in the large majority (a 500:65 ratio). The spectrum of two separate locations in the image from the same time period was extracted and investigated to determine what spectral changes are occurring throughout the flow cell. The spectra can be seen in Figure 5-5 with an inlaid image of the tablet dissolution furthermore the location used to extract the spectral information is marked by the arrows to the image. There are two separate and distinct spectra present in this HSI data within bounds of the flow cell and the differences between these spectra cause the colours in this false coloured image. The first and most striking difference is an intensity change with the region coloured yellow in the image having a more intense spectrum at all wavelengths however, on closer inspection the intensity change is not even at all wavelengths. The

wavelengths used to assign the colour to each pixel in the image are 1073, 1279 and 1407 nm. The inspection of the data shows that the ratios of intensities of the two spectra at these wavelengths are not constant. The ratio of spectral intensities at 1073 and 1407 nm is around 1:1.3 but the ratio at the 1279 nm point is closer to 1:3 therefore the second spectrum has a lower intensity at the middle section of the spectrum which can be seen from 1160 to 1400 nm. The reason this gives a red colour to the second spectrum is that intensity of the 1073 nm 'red' wavelength is only slightly changed while the 1279 nm 'green' wavelength is severely reduced resulting in little contribution to the false colour image. This shows how the use of false colour can easily show a major spectral change in the image quickly and to see how this change is spreading throughout the image over time.

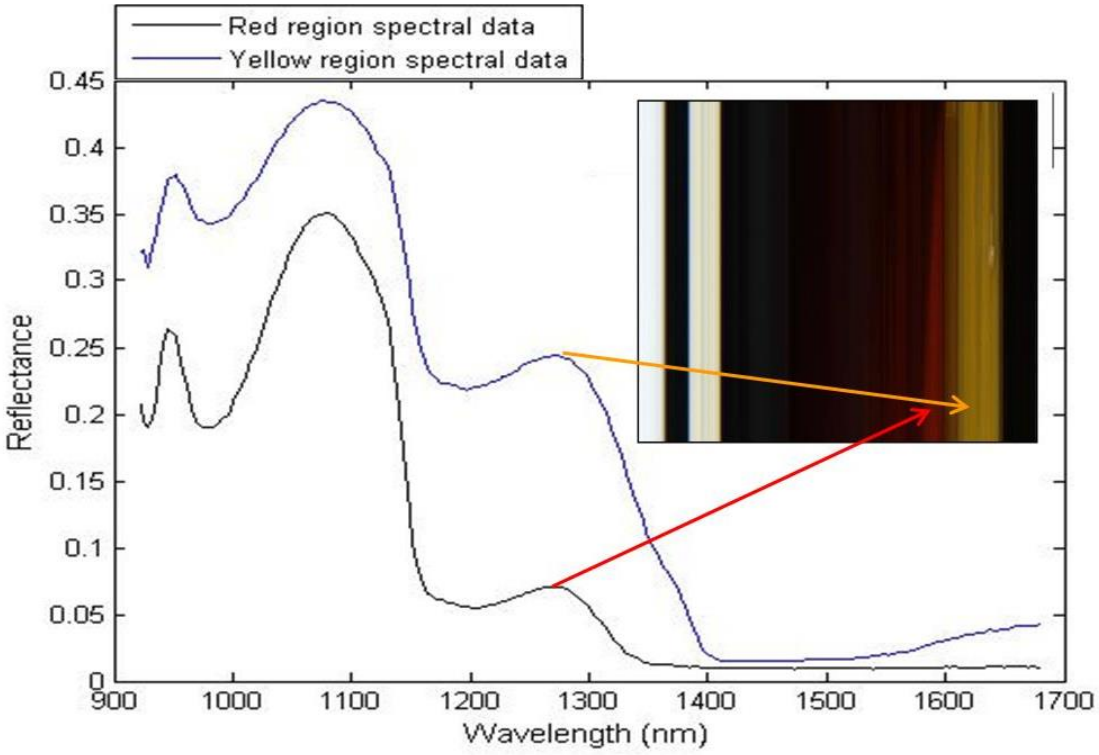


Figure 5-5 A plot containing the spectra of two distinct locations in the HSI data which is inset into the figure. The difference between these two spectra causes the different colours seen in the inset false coloured HSI image.

The dissolution of Tablet F is shown in full in Figure 5-6A with a shorter 400 second image shown in Figure 5-6B. UV dissolution was performed on these tablets (and is shown in the appendix) and they showed that the first 400-500 seconds contained a large portion of the release of the drugs from within the tablets. The time period chosen for Figure 5-6B was one which should show the majority of the release from the tablets with good clarity.

The, false colour, yellow band in the image is the tablet and as the image progresses it is showing how the same specific location in the tablet is changing over time. This yellow colour is the output from the false colouring of these pixels and not the true colour of the tablet which is white. As the spectrum of the pixels change there is a change in the false colouring of the data which is seen with the evolution of a false coloured red region in the images.

While the tablet initially appears homogenous within 60 seconds there are changes appearing in the tablet surface.

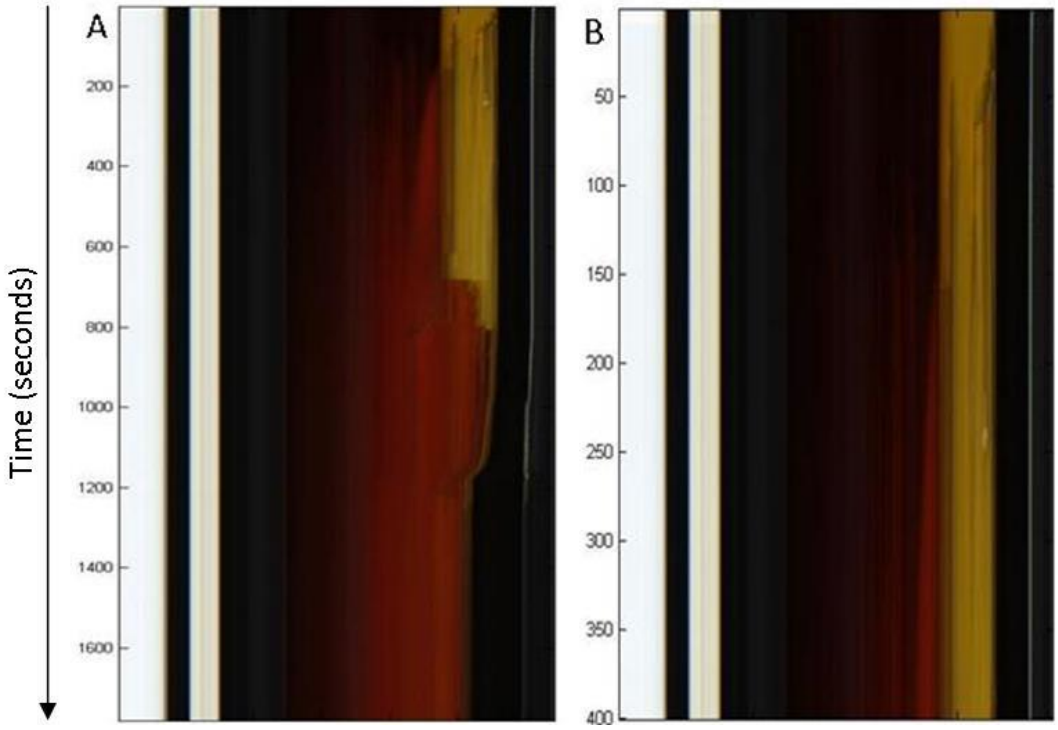


Figure 5-6 The hyperspectral image of the dissolution of Tablet F using two time scales A) the full 30 minutes the experiment was run and B) the initial 400 seconds of the dissolution

A change in the image as the false colouring of the tablet changes from the yellow to red is also visible within the first 200 seconds of Figure 5-6B. This colour change is a visible indication of the spectral change within the data described in Figure 5-5. The ability to easily follow the change in the spectrum of each pixel in the flow cell allows for this change in spectra to be tracked through time when using these false colour images.

During the initial 400 seconds of the dissolution of Tablet F there are a number of processes occurring during the early stages of the dissolution. At the start of the experiment the tablets appears homogenous with no changes in the colour or intensity.

However, as the experiment progresses darker regions can be seen at the edges of the tablet which indicate a lower spectral intensity. These are most likely from the tablet surface starting to hydrate and crack as the expansion begins but this cannot be proven without additional imaging. The evolution of a secondary spectrum can also be seen as a band of red in the image, which is caused by the tablet having a changed spectral signal. By investigating the whole data set it was hoped that more useful data could be attained either as a second spectral change, denoted by a second colour change, or increased spatial information could show more about the tablet expansion and the movement of tablet particles. The image in Figure 5-6A contains the full hyperspectral image of the experiment and the time is shown in seconds on the side of the image. There appears to be a slow change in the tablet for the first 700 seconds and then a rapid change occurs in the tablet bulk. Within 100 seconds the false colours derived from the tablet spectrum has progressed from being predominantly yellow to almost completely red, with only a very small yellow streak present at the edge of the tablet. This end of the tablet is held within a small piece of plastic tubing to prevent movement of the sample and it is highly possible that this yellow streak is the edge of the tablet being held in this tubing. This represents a rapid change within the physical conditions of the tablet which is causing a spectral change in tablet. By the end of the experiment the entire flow cell contains the secondary spectrum with the lower relative intensity at 1279 nm and the tablet appears to be completely broken down with a distribution throughout the entire cell. The distribution does not appear homogenous as the right hand side of the flow cell is brighter which would indicate a higher spectral intensity caused by larger quantity of the material. These results shows that this technique is able to view the tablets and visualise there dissolution properties within the flow cell. However there has only been one change in the flow cell, no other spectral variation is easily notable throughout the full 30 minutes.

5.3.2.2 Tablet A's hyperspectral dissolution images

The next tablet investigated was Tablet A which is non-caffeinated and is unique amongst the tablet as it contains a large quantity of bicarbonates. These are used to cause the tablet to effervesce and quickly break down within the acidic condition of the stomach, aiding release of the paracetamol. The image in Figure 5-7 is false coloured using the same wavelengths as previously and begins with a white section of dry tablet which is quickly submerged and changes to the yellow colouration as the spectrum of the tablet is changed by the tablets submersion. In previous experiments there was a slow expansion and development of a secondary spectrum which is characterised in the image as a red

colouration in the image. In Figure 5-7 however, within 100 seconds of the tablet being submerged it has begun to expand through the entire cell. The expansion is rapid and the image reflects this with the tablet appearing to rupture and split into a number of bright streaks which run though the image. These streaks are thin, and are much brighter in the image with less of the yellow than previous images which indicates a spectral change has occurred at these locations, and their presence can be seen through the entire flow cell. There is still evidence of the secondary spectrum occurring by the existence of the red colour appearing in the image as the tablet expands but this is quickly smothered by these streaks in the image.

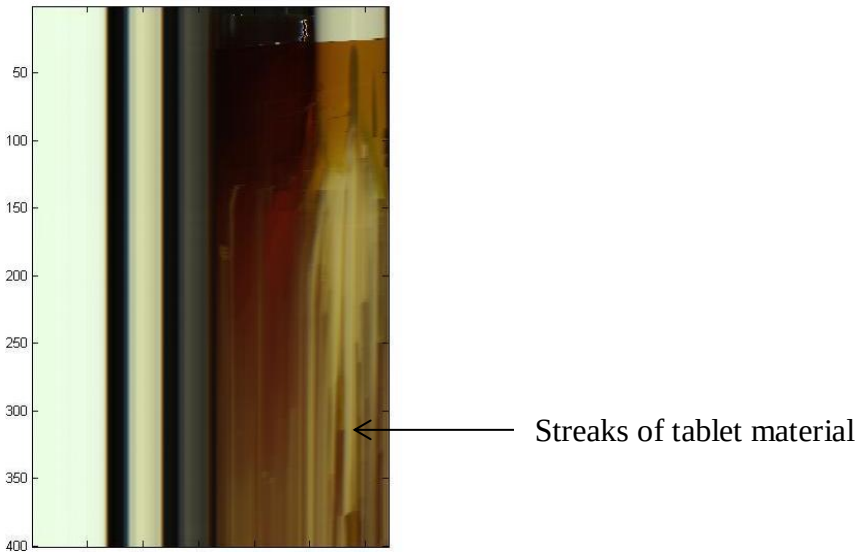


Figure 5-7 The hyperspectral dissolution image of tablet A. The first 400 seconds of the experiment are shown

Investigation of the spectra of these streaks was used to better understand what they are showing and the plot is shown in Figure 5-8. As with previous results an average spectrum was creates from a 10 x 10 region of pixel and three different areas were investigated. The areas of the dissolution which were investigated were the tablet before submersion, the bulk of the tablet and a section of the streaks which are prominent throughout the image.

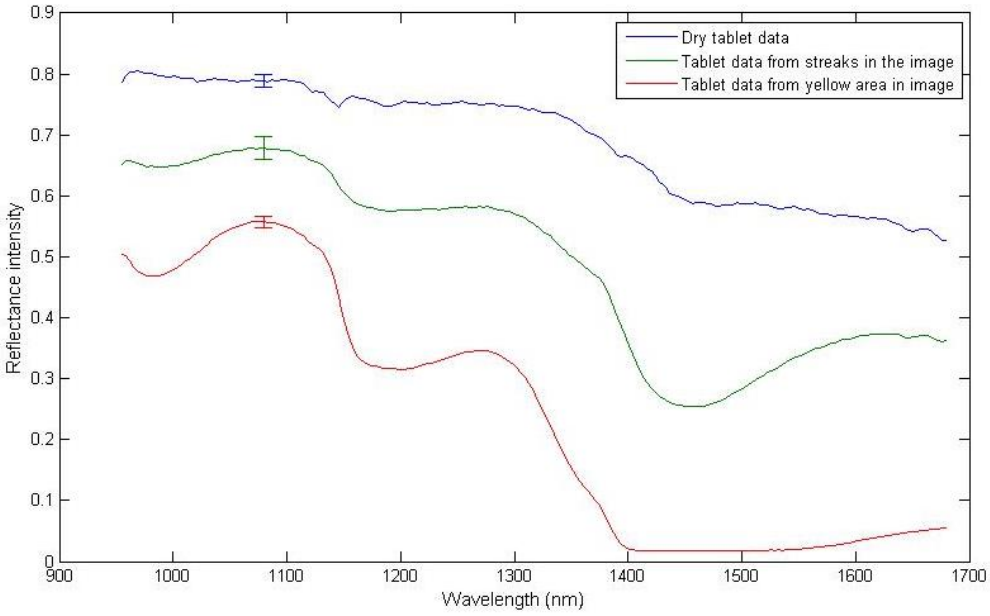


Figure 5-8 The average spectrum of the three regions of interest in Figure 5-7

The spectrum of the streak data has characteristics of both the original pre dissolution tablet and the submerged tablet. The intensity lies in the middle of the two and for the 900 – 1400 nm region there is no individual peak or shoulder which cannot be seen in one of the other spectra. There is however a peak in the data around 1600 nm which is either a unique spectral peak or some characteristic of the process which is generating the spectra in the yellow areas of the image. In the tablet spectrum post submersion there is little to no spectral signal after the 1400 nm region, which is a very strong loss of signal relative to how the spectrum looked before the dissolution began. From visual inspection during the dissolution these tablets bubble and expand rapidly leaving a tablet residue on the flow cell which is, in all likelihood, what is causing the streaks in these results. From observation and analysis of the spectrum these streaks in the images are either from the centre of the tablet being forcibly ejected by the evolution of large quantities of carbon dioxide or they could be showing the effect of the depth of water on the signal as they are most likely attached to the surface of the flow cell with a minimal path length for the reflected nIR photons. The depth penetration of a photon is a function of its scattering properties which are proportional to λ^{-4} where λ is the wavelength. The photons of lower energy have longer wavelength and are more readily scattered which reduces the effective path length of these photons. This would explain why the vein data, which if it is near the surface has a shorter path length, has more spectral definition in the longer wavelength region of the spectra. The longer path length to the bulk of the tablet leads to greater scattering for all

wavelengths but the effect is most severe at the longer wavelengths which causes to loss of spectral information after 1400 nm.

Investigation of the full image will indicate how this initial expansion progresses and if any more changes occur within the data. The full HSI data is shown in the appendix and it shows that the streaks in the data are clearly present for a long time with there being no real sign of them changing throughout the experiment.

There is a small number of pixels coloured red in the 30 minute image which is unusual as the remainder of the tablet hyperspectral images show the evolution of the tablet data from the starting 'yellow' spectrum to the 'red' spectrum. The tablet is designed to release the API as fast as possible and the shift in the spectrum which leads to the red colour in the images is believed to be a function of this process. This implies that Tablet A, which is marketed as a rapid release formulation, is not breaking down and releasing the API in a way which is consistent with the other tablets. The two possibilities for this discrepancy are that the secondary spectrum, giving the red colour, in which the intensity ratio of the 1073 to 1279 nm wavelengths is much lower is not a function of release from the tablet or that the large steric bulk of Tablet A is causing significant hindrance with the release in the relatively tight flow cell. On closer inspection of the streaks in the image there is a shift in many of them down the cell in the direction of flow. These streaks are moving through the cell in the direction of the flow and appear to show that the steric bulk is not fully hindering movement through the cell. There is also a small region of coloured red in the early portion of the image (between 100 and 300 seconds) which indicates that while this was formed it was in some way stopped, removed or masked by the formation of these streaks within the dissolution images.

The tablets not discussed here (tablets B, C, D and E) also showed similar results to tablets A and F. The initial spectrum of the tablet bulk changed over time with the intensity at 1279 nm decreasing relative to the intensity at 1073 nm. This change to the spectrum of the tablet is seen throughout the hyperspectral data as a change in the false colouring of the image. Furthermore this spectral change in the images appears to show the point at which the tablets are expanding and dissolving as it is the edges of the tablet which change first. The images from these experiments are all shown in the appendix.

5.3.2.3 Comparison of bulk tablet spectra

An average spectrum for each of the tablets was created and plotted together to show any similarities or difference between the spectra. These spectra are shown in Figure 5-9 and at a cursory glance the spectra appear different. However, on closer inspection the data is very similar but with an intensity scaling difference which is easily explained by the variations in signal intensities across the tablets. By visual inspection of Figure 5-6 it is clear that there are regions of higher (brighter) intensity and regions of lower (darker) intensity across the tablet. These variations are prevalent throughout the images and across the individual tablets which could easily cause the intensity of a random spectrum to change in both an intra or inter tablet situation.

Other than the intensity difference between the spectra there is very little to distinguish between the tablets. The two major regions of difference in the spectra are the gradient between 1200 and 1300 nm and the broad peak at 1600 nm. There is a large difference in the compounds used to make these tablets with tablet A being especially complex. This is not reflected in the post submersion spectrum in any way and details a potential issue with the data.

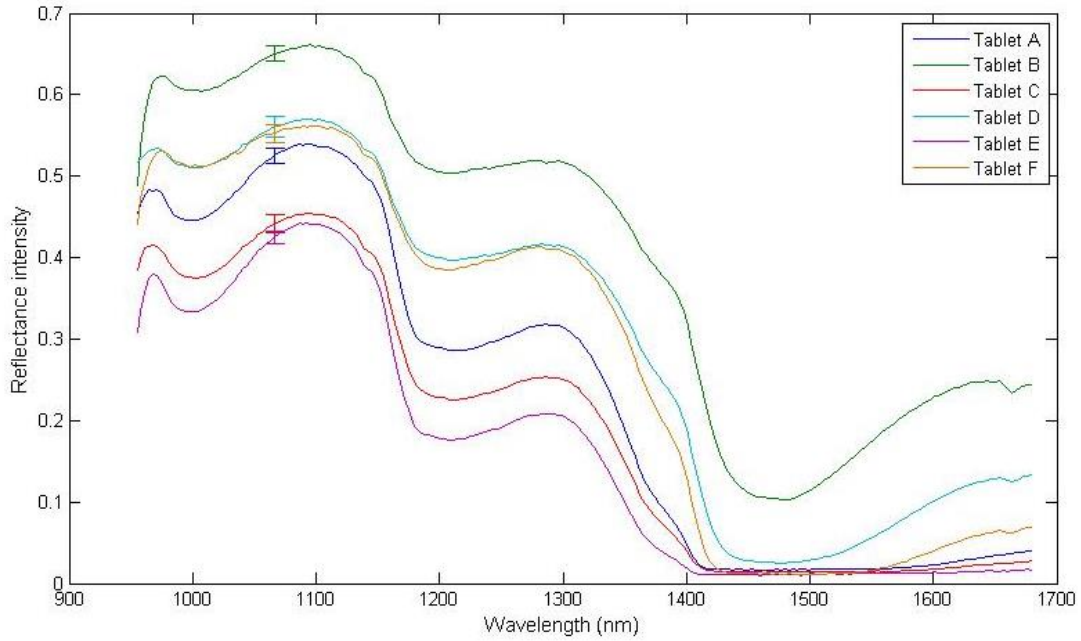


Figure 5-9 The average spectrum from each of the tablets after submersion in acid solution

5.3.3 Using mesh plots to show the changes occurring to the tablet as the experiment progresses

5.3.3.1 Introduction to mesh plots and how they were generated

The hyperspectral images of the tablet dissolution details a similar dissolution progression over time for each different tablet formulation. The tablet spectrum with the more intense signal at 1279 nm appears in the image as a yellow band and is seen in the initial portion of the image. Over time the spectrum is changed as the intensity of the 1279 nm is lowered which causes a colour change creating a red band in the false colour image as the dissolution progresses. The sole exception to this is Tablet A, the red colouring does appear briefly in the image before it is masked by intense streaks of tablet data. The shift in the false colours in the image from yellow to red clearly shows a spectral change within the tablet data and this can be seen from 1200 nm onwards. By closer inspection of the data both across the flow cell at set points in time and a set point in the flow cell as the experiment progresses it may be possible to better characterise this change in the spectral data. This was performed using mesh plots in MATLAB which plot consecutive spectra to create a 3D plot. The first tablet formulation investigated was tablet F as this was seen as the best 'average' tablet for the rates of change and tablet disintegration and expansion.

The hydration and expansion of the tablet which appears to create the secondary spectrum and moreover the red colouration is clearly seen within the first 400 seconds of Figure 5-6 and this same image was used to locate the best positions for the mesh plots. The x axis mesh plots dissect through the flow cell at a specific time point and will show how the tablet is changing, how it is expanding and how uniform the changes are across the tablets. Four of these locations were analysed and the positions were chosen to allow for comparison of data from multiple points in the tablet expansion process from the early tablet through to the later time points where the expansion and breakdown of the tablet is the prominent feature through the entire flow cell. The y axis mesh plots dissect the y 'time' axis at a set spatial pixel location; three of these locations were chosen to detail how the spectrum changes over time. The locations used are from the bulk tablet, the interface of the tablet between the edge of the tablet and the surrounding acid solution and also a point slightly away from the early tablet data but within the later expanded tablet data. The exact locations for these mesh plots are shown in Figure 5-10, the x axis mesh plots are shown in green and the y axis mesh plots are shown in blue. The y axis mesh plots cover the whole of the data set not just the 400 seconds shown in this image.

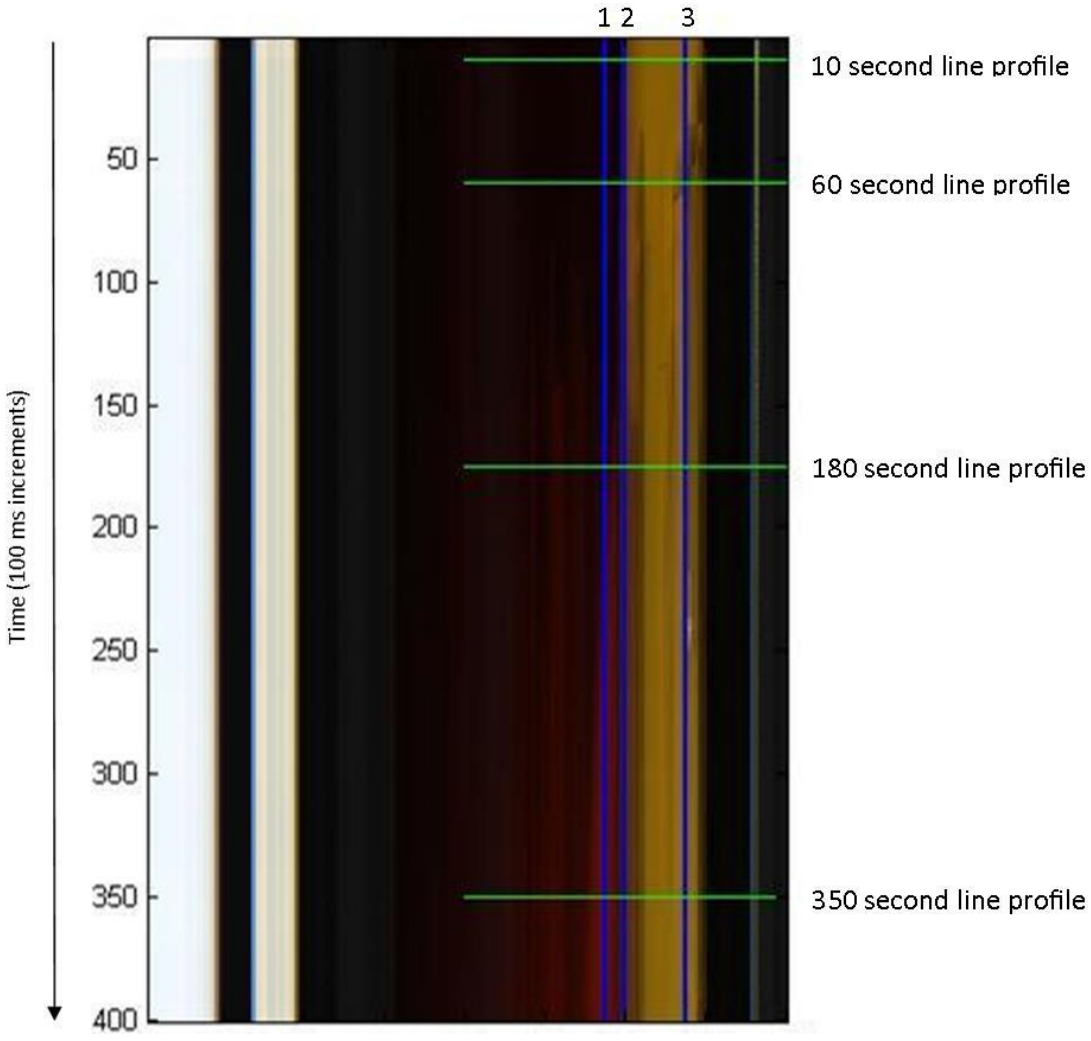


Figure 5-10 HSI Image of tablet F's dissolution which has been marked to show the locations chosen for mesh plot analysis. The green lines show the x axis plot locations at 10, 60, 180 and 350 seconds and the blue lines show the y axis plots at Positions 1, 2 and 3

5.3.3.2 X axis mesh plots of Tablet F

The x axis mesh plots are all shown in Figure 5-11 with the earliest time point shown in Figure 5-11A etc. The earliest time data is from 10 seconds into the experiment and is used to show the tablet data just after submersion when little to no changes have occurred within the tablet. The tablet itself is between the 9 and 19 mm markers in the cell and is uniform across the entire tablet which gives a good starting plot for comparative analysis. There is also an individual spectrum around 22 mm which appears to be an artefact or a small piece of something which has moved into the flow cell.

The second x axis mesh plot was created using data from 60 seconds after the experiment began, and can be seen in Figure 5-11B. This plot shows how the expansion and disintegration of the tablet is progressing after the tablet has been submerged for only 60

seconds. The once uniform level of tablet spectra is gone by this time point and has been replaced by a range of intensities. The edges of the tablet have lost intensity rapidly while the core is staying mostly unchanged from the data shown in Figure 5-11A. The dimension of the tablet has not changed so no expansion has yet occurred in the direction of the flow which implies that these changes are either pre-expansion or showing upwards expansion to fill the flow cell before it spread outwards down the flow cell. One limitation of the imaging protocol is the lack of the spatial information which could be used to show exactly what this change is showing. It is not possible to definitively determine if the expansion of the tablet is the expansion of a complete tablet or the movement of tablet particles which are the product of the tablet breakdown without the full spatial information. Whatever change is being seen in this plot it is clear that the change occurs at the end of the tablet where the surface area for interaction with the acid solution is greatest.

Figure 5-11C is the mesh plot created from taking a cross section of the flow cell at 180 seconds into the dissolution. Inspection of Figure 5-10 shows that this time point is where the false red colour starts to become prominent on the edge of the tablet but that the tablet itself is still mostly the same spectrum which gives the yellow colour in the image. Previously there was no evidence of tablet material in the first 9 mm of the flow cell but Figure 5-11C does include some spectral data in this region. The intensity is very low but it is clear that the tablet is beginning to break apart and some of the tablet particles are causing these spectral signals. The end of the tablet is also beginning to break apart, something which can be seen in the intensity of the bulk tablet region. In Figure 5-11B the edge of the spectral data was still sharp with a very clear end to the tablet which is not the case in Figure 5-11C. In this data the end of the tablet now shows a much gentler gradient from the tablet bulk to the region of the flow cell which contained no tablet at the start of the experiment. This change is evidence of the tablet breaking apart from the edge and also potentially the start of the swelling process as the tablet begins to move further down the flow cell.

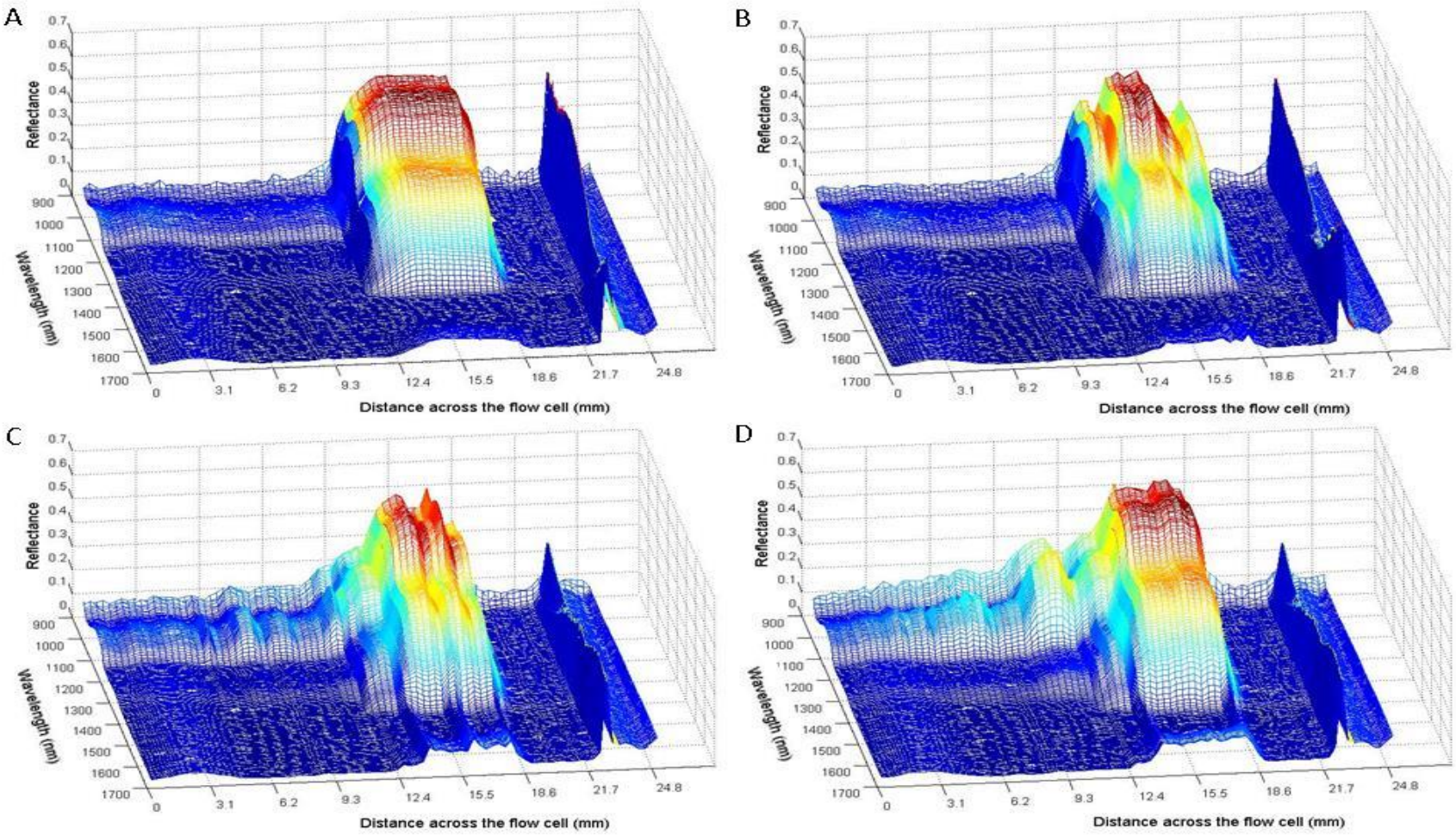


Figure 5-11 Four mesh plots which show the spectra from a cross section of the HSI dissolution data of Tablet F. The exact locations used are shown in Figure 5-10 where A) 10 seconds into the experiment, B) is the 60 seconds into the experiment, C) is 180 seconds into the dissolution and D) uses data collected 350 seconds after the experiment began.

The final x axis mesh plot is Figure 5-11D and the data used for this mesh plot is taken from 350 seconds into the experiment. This time point was chosen to show the evolution of the secondary spectrum in which the 1279 nm intensity is lower, relative to previous data, and is becoming more prominent and should also show the later stages of the tablet swelling process. The spectral data looks most like Figure 5-11A as the bulk of the tablet is mostly uniform with no sign of the strong intensity changes shown previously. This change in uniformity is due to the expansion of the tablet coming to an end with the tablet material being either pressed against the wall of the flow cell or settling to a constant level after the tablet has broken down. There is no way to ascertain the height of the data so either reality is possible but the end result is the same, a return to the uniform intensities in the bulk of the tablet. To the side of the tablet is more evidence of tablet material being moved along the flow cell either by the swelling and breakdown of the tablet or by tablet particles being moved by the flow of the acid solution. The secondary spectral signal in the false colour image of Figure 5-10 is generated by these small amounts of tablet material which are deposited throughout the flow cell and this indicates that the red spectrum is in some way reflecting the tablet after breakdown.

All of these mesh plots show interesting information about the environment throughout the flow cell at a specific time but by considering them together it is possible to see the changes in the tablet. The mesh plots show that the tablet starts as a single uniform spectrum which quickly changes once the acid solution begins to flow. The tablet appears to expand and break down with small objects being distributed downstream of the tablet, which causes a slight change in the signal at specific regions, as the experiment progresses larger pieces of tablet can be seen.

5.3.3.3 Y axis mesh plots for Tablet F

The Y axis mesh plots can show the changes in a specific point in the flow cell as time progresses. Unlike the X axis mesh plots these plots don’t show the change for the entire cell but rather the changes at a specific location such as the bulk of the tablet. Figure 5-12 was created by taking a cross section of the data along the time axis at Position 1 which can be seen in Figure 5-10. This region was chosen to give more information on the expansion, breakdown and clearance of the tablet within the flow cell during the dissolution.

The earliest data shows very little signal which is to be expected as at the start of the experiment there should be nothing in the flow cell at this location. As the experiment progresses there is a clear increase in the signal response from no signal to a maximum reached after 900 seconds.

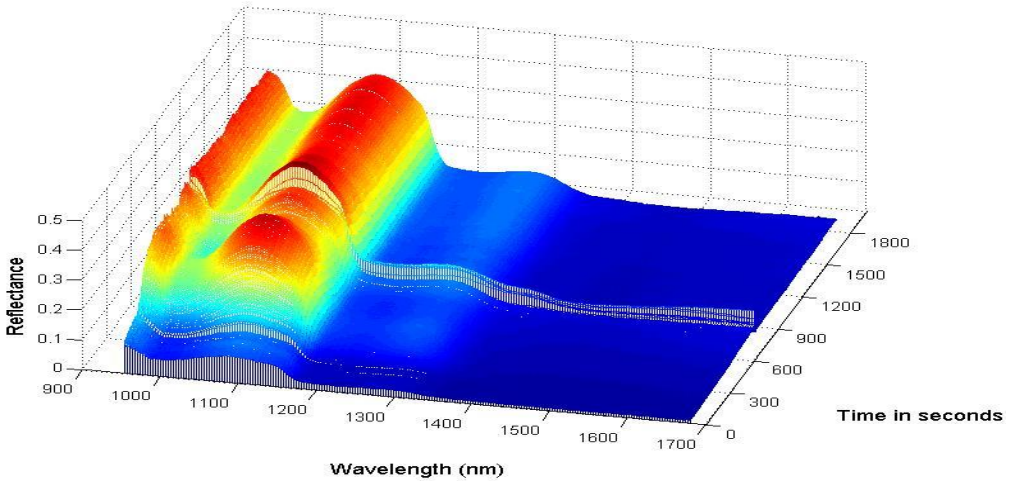


Figure 5-12 A mesh plot generated using a cross section of the time axis in the image from the dissolution of Tablet F. The exact location used for this plot is shown in Figure 5-10 as the Position 1

The increase in signal strength is gradual in the earliest stages, which indicates that the tablet is slowly expanding and that tablet material is being deposited at this location at a steady rate. It appears that the some of the tablet material is lost in the period between 600 and 900 seconds but that it is quickly replaced and after this point the intensity doesn't appear to change for the duration of the dissolution. The tablet has completely broken down by 900 seconds and no more changes are occurring with respect to changing signal intensity after this point. To investigate this change the intensity of the peak at 1076 nm over time was plotted and is shown in Figure 5-13. As with Figure 5-12 there is a clear lag period before the tablet moves across into the selected region of the flow cell. This lag period is followed by the first of the three different stages which can be clearly seen in the plot. This region has the steepest gradient and thus the rate of change is quickest during this region, this is where the tablet material is initially expanding across the flow cell. The second region lasts for approximately 600 seconds and shows a slow constant increase in the single intensity over this time. This increase is seen in Figure 5-12 and is caused by the tablets disintegration causing more material to lie at this point in the flow cell. The final region is mostly flat with no change and shows that the tablets disintegration and the distribution of tablet material have mostly stopped by this point in the experiment.

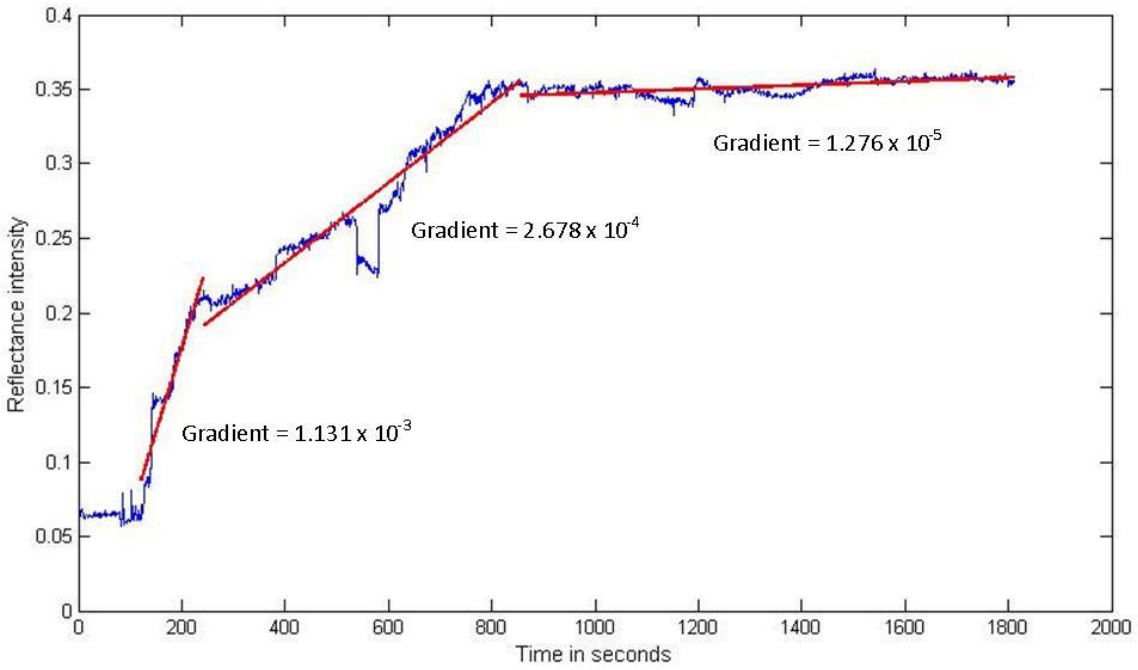


Figure 5-13 A plot following the change of intensity at 1076 nm over time using the data shown in Figure 5-12

The other y axis mesh plots (which are shown in the appendix) show similar results with the majority of the variation occurring within the initial 900 seconds of the experiment and very little being visible after this point. The results of mesh plot analysis for each of the tablets is similar with a change across the tablet as it breaks down and the secondary red spectrum begins to become prominent. The rate of this change and spread throughout the cell is different for each tablet brand but the general information shown is consistent.

The loss of the API or excipients from the tablet should cause a slow change in specific peak locations in the data a process which is not seen in any of these results. Also there is a significant change in the spectrum of each tablet as soon as it is submerged in the acid solution; this change is so rapid it could be used in some images to track the rate of the acid flow in the flow cell. These results imply that there is interference in the flow cell which is caused by the acid solution. The exact nature of how this interference is interacting with the signal must be calculated so that a correction factor can be created to return the attenuated tablet spectra back to the original form. Without this correction then this data cannot be used to follow the release of paracetamol and caffeine from the tablet. The data can still be used to investigate the disintegration and expansion properties of the tablets in great detail due to the high time resolution used in the experiments.

5.4 Calculating and correcting for the signal attenuation present in the data

On closer inspection of the results from 5.3.2 it is clear that some physical effect is altering the spectral data which is creating difficulties with the analysis of the results. By following the motion of the acid solution through the cell and over the tablets in the initial stages of the experiment it is clear that the acid solution is causing some kind of filtering effect. The spectrum is seen to change as soon as the acid is between the tablet and the camera, such that the photons travelling through the acid are undergoing some form of filtering effect.

A series of different corrections were attempted to remove this effect which are detailed in the appendices. The outcome of these corrections is that the acid solution is attenuating the signal from the tablets and this effect cannot be easily corrected. This effect is not a simple combination of two signals interacting to make the output spectral response as the attenuation changes during the experiment.

5.4.1 Determination of the cause for the signal attenuation

The correction results in the appendix show that the signal is being affected by more than a simple masking effect or filtering from a second signal. The spectral response from the nIR reflectance is being attenuated with respect to some factor. The most likely cause of this attenuation is the path length of the nIR photons through the acid solution; as the path length increases the attenuation is also increased.

The reflectance value detected by the hyperspectral camera is a mixture of different reflectance terms which are shown in Figure 5-14. The light which enters into the flow cell from the lamp is denoted as I_0 and is a 100% transmission of the nIR photons with no loss until the flow cell is reached. Once the nIR photons reach the wall of the flow cell there is a small amount of reflection and absorption of the nIR photon which causes a slight reduction to the total number of photons. This happens each time the photons reach the edge of the glass the flow cell and they are shown in the diagram as R_G. The loss of the spectral signal due to the absorption of the photons through the acid solution is shown as R_A and the final reflectance value R is the detected reflectance from the sample. The loss in the total signal intensity of R is a product of each of the different absorption or reflectance process; the loss from passing through the glass and the acid solution is a percentage of the total intensity and so R can be expressed as:

$R = I_0 \times R_G \times R_G \times R_A \times R_{Sample} \times R_A \times R_G \times R_G$

Where each R term has a value between 0 and 1 such that R has a non-negative value which is lower than I_0.

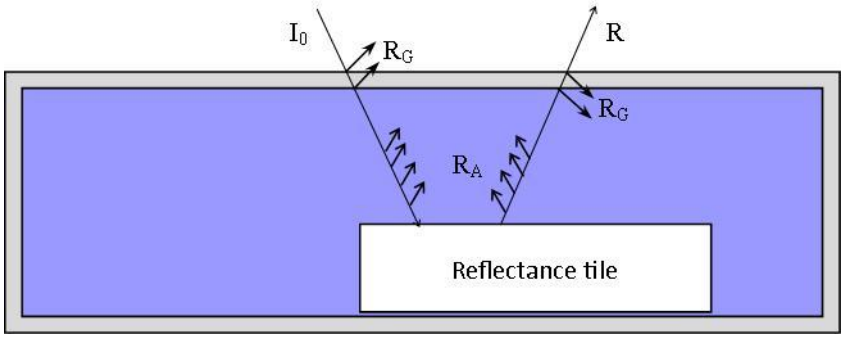


Figure 5-14 A diagram showing the different ways that the signal attenuated

The reflectance value from the wall of the flow cell (R_G) is measured during the pre-dissolution calibration images. This is constant both between and during experiments and can be considered as a singular unchanging factor. While the signal is attenuated from this process it is by a fixed amount in each experiment and cannot be the cause of the change over time to the spectral signal.

The intensity of I_o is also a constant throughout the images which leaves only two factors which could explain the attenuation of the signal; the reflectance of the sample changing or the reflectance of the acid changing. The sudden change in the spectrum implies that the change is not from the reflectance of the sample as it is unlikely that the sample would change so severely so fast. If the path length of the acid is causing this attenuation of the spectral signal that is seen in the images above then the value of R_A is changing.

5.4.1.1 Path length data

To test if the path length is the factor controlling the attenuation an experiment was carried out. A small piece of the reflectance calibration tile was placed into the flow cell and the cell filled with acid solution. The reflectance of the tile was then measured and the flow cell emptied, the tile was then raised by placing it onto a thin piece of polyvinylchloride (PVC) and the process repeated. This was done until the calibration tile could be raised no higher and the reflectance values were then plotted. The calibration tile and the PVC raisers were used as they are both inert in a hydrochloric acid solution. As the plastics are

inert the spectral response of the tile will be constant throughout the experiment and the height will not be changed by any dissolution of the raising materials.

Once the imaging was completed a region of 10x10 pixels was chosen to create an average spectrum, this same region was used in each image to minimise the variability, if any, across the sample. The average spectra were then plotted as seen in Figure 5-15 where it is clear that an increase in path length, or a decrease in height which is how the data is represented in the plot, has a definitive and seemingly ordered effect on the intensity. The plot also shows that the data from the second height increment appears to be erroneous as it changes in intensity to a greater degree than any other signal, this data was removed from further calculations to prevent any skew or bias to these incorrect values. It is believed that this result is from the calibration piece moving so that a different edge is facing upwards, which would change the spectral signal substantially.

The degree of attenuation at each wavelength is separate and unique as is shown by the difference between the intensity of the signals at the wavelengths of 1255 and 1650 nm. The change in intensity at the 1255 nm region of the spectra shows a near linear height change with a loss of reflectance intensity of approximately 0.1 at each height step. The data from the 1650 nm region however, displays a substantial difference in the attention effects with the signal intensity quickly minimising to a value of below 0.1 within 3 increments of height.

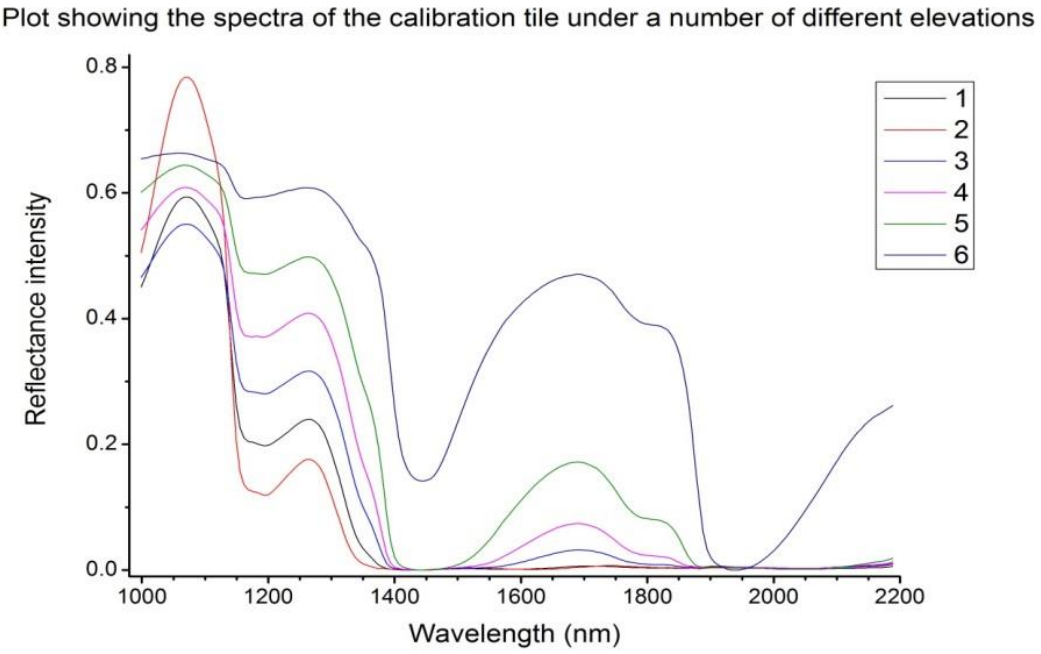


Figure 5-15 Plot showing the incremental height data used to determine the effect of path length in the data

These results demonstrate that the path length is having an effect on the attenuation of the signal. The next step was to attempt to calculate how the path length is affecting signal as the degree of attenuation is not the same in each wavelength. Mathematical models were created to investigate the signal attenuation and to discover the relationship between the path length and the degree of the attenuation. These models use the same data as in Figure 5-15 and are included in the appendix.

The models show that the path length and the signal attenuation are related by the equation below where y is the spectral response, c, A and d are constants and x_i is the path length.

$$y_i = c^{x_i + A} + d$$

This equation shows that the magnitude of the signal attenuation is caused by a power relationship which highlights the importance for minimising the path length of the photons through an absorbing medium.

This result is not unexpected as it is showing that the reflectance of the sample can be calculated by considering the path length of the light through an absorbing medium and a factor which is related to the absorbance of a signal at a set wavelength with no

interference. This equation is very closely related to the Beer-Lambert law which states that:

$$R = \frac{I}{I_0} = 10^{-\alpha l}$$

Where R is the reflectance of the system which is a function of the measured light intensity divided by the total light intensity, l is the path length of the photons through the absorbing medium and α is the absorption coefficient. The Beer-Lambert law shows the same relationship between signal strength and path length therefore the α term is closely related to the c term.

5.5 Conclusions

During previous work the dissolution of paracetamol tablets had showed a change over time but the time between the different hyperspectral images was too long and little useful spectral information could be gathered. In this chapter a novel approach to the imaging of these tablets was used, in which the spatial information was degraded but the temporal information was improved substantially. The new images create a new line of data every 100 ms by only imaging a single spatial location. To ensure the accuracy and repeatability of these experiments a number of procedures were used and the creation and use of a specialised holder along with the taking of pre imaging reference spectra being the two major procedures.

The very high time resolution does, however, come at a price and due to the change in the form of the hyperspectral imaging the data produced can be difficult to understand. The loss of spatial information results in no way to clearly see the full breakdown of the tablets in the flow cell, instead there is a change in a band across the tablet which does not reflect the full complexity of the tablet breaking down. There is no knowledge of how the rest of the tablet is moving and if the band seen in the images is representative of the whole. However, the temporal resolution in this data is extremely good and can lead to the tracking of small fast moving objects which pass through the flow cell during the dissolution. This resolution is better even than some UV dissolution experiments and thus has the potential to better show the rate and distribution of API loss from the tablet matrix.

While there is a loss of spatial information and a difficulty in understanding and reading the images they have clearly shown more about the system than the previous, more standardised images. The results from Tablet A in particular emphasise the level of understanding which can be gleamed in this technique. Within the dissolution of this tablet a complex system of the acid egress and rapid tablet expansion can be tracked along with the movement of particles through the flow cell and even slight changes over time in the spectra of set locations.

The use of the high temporal resolution allowed for mesh plots detailing the changes within the flow cells at set points, these mesh plots can show how the tablet environment is changing such as the breakdown and expansion of the tablets over time. The hydration and tablet expansion can be tracked with high precision using this technique and with further analysis it should be possible to generate a measure of the expansion as was done in

previous chapters. The loss of API from the tablet is not visible in this data and appears impossible to track with the current imaging setup. However, the increased time resolution means that once the attenuation has been removed or reduced then it should be possible to use HSI to accurately follow the release of the API from within the tablet matrix. The mesh plots also show that is should be possible to follow the loss from specific pixels in the tablet data rather than as an average of the whole like in the current dissolution systems.

This shows that there is a potential to use hyperspectral imaging in tracking the progress of moderately high speed reaction, and not just the breakdown and release from tablets, if the problems in the imaging of the tablets can be resolved.

It became clear during the analysis that the spectrum of the tablets was being attenuated and that this was most likely due to the path length of the nIR photons through the acid solution. Further analysis of the results showed that the path length was indeed having an effect on the spectral signal by attenuation. This attenuation is not the same at every wavelength and cannot be removed by a simple subtraction or filtering of the data through a transfer function.

By using a model it was shown that the severity of the attenuation is proportional to a constant raised to the power of the path length. This explains why the signal becomes weaker as the experiment progresses because as the tablet breaks down the path length through the acid to the tablet increase. The increase in the path length causes a large increase in the degree of attenuation to the signal.

These results were predicted by the Beer-Lambert law but the severity of the attenuation unexpected, the absorption coefficient of the water in the system is large enough that it is causing a masking of all over data signals.

The cause of the attenuation is the water molecules in the acid solution strongly absorbing the nIR photons which mask any smaller absorptions from the surface of the tablet. This attenuation is a crippling problem for using nIR to follow the dissolution of the tablets and there is no way to stop the attenuation from occurring when using aqueous systems and near infrared detection.

There are only two possible paths to take from these results, either the attenuation must be reduced by minimising the path length or a new wavelength range should be used. The new

wavelength range would require a UV hyperspectral camera as any nIR or IR based system would be affected and in chapter 3 it was shown that the HSI cameras which operate in the visible are not sufficient for the task. UV hyperspectral systems are rare and very expensive so gaining access to a system is unlikely, also the need for a UV light source could cause further complications.

Minimising the path length could be attained by either using a narrower dissolution vessel or by using gravity or imaging from below the flow cell. Both of these approaches have complications and would not completely remove the effect of the attenuation however, it may well be that this is the most realistic approach to improving the results. The approach chosen in Chapter 6 is to use gravity to pull the tablet material to the bottom of the flow cell and to image in a bottom up style. This system is able to minimise the path length of the nIR photons without changing the fluid dynamics of the system by creating a more restricted flow.

If the attenuation of the signal can be averted then the high temporal resolution and good spectral resolution could be used to great effect to determine the release properties of tablets in new and novel ways.

Chapter 6: Hyperspectral determination of caffeine within a tablet during dissolution.

6.1 Introduction

The use of nIR as a tool for hyperspectral dissolution imaging has been shown in previous chapters to have great potential but it is hampered by signal attenuation. Water absorbs near IR radiation strongly at two specific wavelength which are 1440 nm and 1930 nm[153]. The absorption of these photons causes specific regions of the nIR spectrum to be masked by broad peaks which prevent any useful information being identified. This attenuation of the spectral data causes a loss of meaningful results within the dissolution imaging.

The best way to improve the hyperspectral dissolution imaging is to lower the effect of the signal attenuation by the water in the flow cell. The simplest and potentially most effective way to achieve this aim is by changing the experimental setup. The absorption of any wavelength of light through a liquid is proportional to the path length through the absorbing medium which the photons must pass. By reducing the path length of the photon through the liquid in the flow cell then the amount of absorption by the water, and thus the severity of the attenuation, can be diminished.

This body of work attempts to use this simple premise to effect a reduction in the signal attenuation present in the data and thus to produce results with a greater degree of useful spectral information. By lowering the degree to which the signal is attenuated it should be possible to see peaks of interest in key areas, these peaks could allow for more in depth analysis of the tablets. This analysis could show anything from seeing an excipient breaking apart from the cell to being able to track the release of active pharmaceutical ingredients or even acquiring greater information about the change in the tablet during the expansion and disintegration stages. Any one of these potential analytical areas could provide very useful information about how the tablet releases its payload, which could in turn be used to better design the tablets.

The hyperspectral imaging systems used in these experiments is configured to detect reflected nIR photons. This was utilised to minimise the path length through the acid solution by imaging from beneath the flow cell. Gravity will pull the material to the base of

the flow cell and thus the nIR photons should only have the acid solution which has permeated the tablet matrix causing attenuation. There will be very little acid solution for the nIR photons to pass through before reaching the detector which should minimise the effect of the signal attenuation. The limitation of this system is that the flow across the surface of the tablet is affected relative to the previous experiments. In Chapters 3 and 4 the image is of the top of the tablet which has no restriction on the flow of liquid over the surface. However, this is not true with the experimental setup used in this chapter as the experiment is imaged from underneath the flow cell. The face of the tablet being imaged is pressed against the wall of the flow cell with minimal space for liquid to flow through. The tablet will not have the same rate of flow from the liquid over the surface of the tablet than in previous chapters. This could affect the dissolution properties of the tablets however, the tablets chosen are fast acting and swell rapidly which will minimise these problems as swollen tablets can also allow for the solution to pass through them.

6.2 Materials and Methods

6.2.1 Materials

Due to restriction with the availability of the HSI system a reduced number of tablets were imaged in these experiments. Previous experiments contained a mixture of caffeinated and non-caffeinated tablets as well as a range of complexity to the tablet formulations. It was decided that the three best tablets to use would be Panadol Actifast, Panadol Advance and the Panadol Extra; these were known as Tablet A, Tablet B and Tablet C respectively. These three tablets are fast acting and have shown good release profiles, their formulations are designed to maximise the breakdown and release of the API from the tablet and a full list of excipients is available in Chapter 2.4.2. Panadol Actifast is unique because it contains sodium bicarbonate which is used to improve the breakdown of the tablet by reacting with stomach acids, creating carbon dioxide gas. There is also a mixture of caffeine containing and none caffeine containing paracetamol tablets, the caffeine is added as a stimulant into certain brands of paracetamol to improve the performance of the drugs and to help lessen any drowsy and groggy feelings which the taker may be suffering from. The inclusion of caffeine into the tablet could show regions of great interest or the increased complexity may mask other results, which is why both types of tablet were used.

6.2.2 Equipment

The nIR HSI system used in these experiments was a Specim SWIR Spectral Camera which consists of MCT camera with active pixel dimensions of 320 (spatial) x 256 (Spectral) pixels. The spectrograph was an ImSpector N25E with OLES Macro lens attached.

The software used for the data capture was SpectralDaq® with MATLAB® being used to control the scanner speeds for later data analysis. Initial image verification and analysis was performed using the Exelisvis ENVI® software package and SIMCA-P 11 was used to generate scores plots.

6.2.3 Experimental setup

The flow cell apparatus remained unchanged from Chapter 5 with the same pump, flow rate, tubing and holder. However, the flow cell holder was suspended in the air above the HSI camera so that the imaging of the base of the flow cell could take place. The flow cell

was placed into the grips of a modified clamp stand, which contained a bolt to attach it to the motorised stage. Subsequently the x-axis and y-axis movement of the flow cell can be accurately controlled by the movement of the motorised stage. A picture of the experimental setup is shown in Figure 6-1, the flow cell and holder can be seen positioned close to and above the HSI camera.

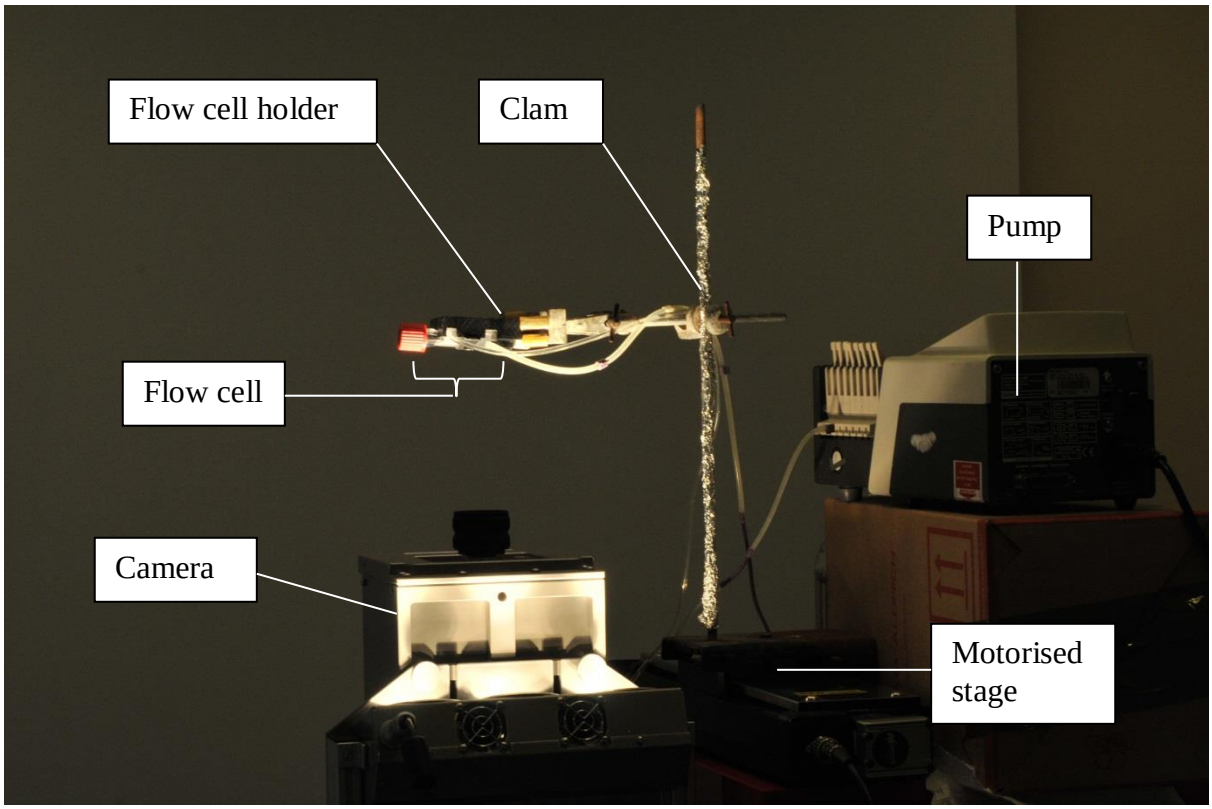


Figure 6-1 Picture showing the experimental setup used in this chapter

6.2.4 Experimental procedure

The dissolution of Tablet A, Tablet B and Tablet C was run for 30 minutes following the same experimental procedure as shown in Chapter 5 with a flow rate of 10ml/min. Previous experiments had shown that the majority of the API is release from these tablets within a 15 – 20 minute window. The extra time was to allow for any reduced release rate due to the more confined position of the tablet with relation to hydration and expansion. The camera captured a new line of data every 100 ms with an exposure of 2.5 ms, the motorised stage was set to move 1.97 mm/s causing a full pass of the flow cell every 6 seconds. The movement of the stage was controlled using special functions coded in MATLAB, this allowed for very tightly defined control of the speed and location of the stage.

The nIR photons were generated by a set of high powered DC lamps attached to the camera to ensure a maximum illumination of the sample. It is important that the sample is well illuminated to ensure a strong and constant signal response from the reflectance off the sample.

The system was also used to collect a series of reference spectra for the active pharmaceutical ingredients (API) and major excipients present in the tablets. The image capture began as the solution entered the flow cell and continued until 1 minute after the flow cell had completely filled. This gave an accurate spectrum for both the dry and wet compounds.

Once the data had been captured it was viewed using ENVI to ensure that there were no imaging artefacts which would ruin the data processing such as the flow cell 'slipping' out of view or some particles on the lens.

6.2.5 Data processing

6.2.5.1 Data correction

Once the image capture was completed a set of calibration images were captured using a white PTFE tile for the 100% reflectance and a 0% reflectance image was recorded by switching off the lights and placing the lens cap onto the camera. These two values are used to convert the raw data which is expressed as a number of counts into the reflectance of the nIR radiation from the sample.

The calculation for determining the reflectance of any point is shown below

$$R_{s\lambda} = \frac{I_{s\lambda} - D_{\lambda}}{I_{o\lambda} - D_{\lambda}}$$

Where $I_{0\lambda}$ and $I_{s\,\lambda}$ are the intensity (in counts) of the 100% reflectance standard and the target pixel respectively at a set wavelength λ. D is the intensity of the dark or 0% reflectance standard and $R_{s\,\lambda}$ is the reflectance value for any specific pixel in the image at a set wavelength. This data correction was performed using a bespoke algorithm written in MATLAB.

6.2.5.2 Pre-treatment of spectra

It is common with nIR analysis to investigate the first and second derivative of the spectra with respect to wavelength ($\delta R/\delta\lambda$) [127,148]. The peaks in the nIR spectra are often broad, this is due to them being comprised of a number of smaller slightly shifted peaks. The change in any one of these smaller peaks will affect the overall peak shape but not necessarily in a large easily visible manner. By investigating the derivatives of the peaks any changes in the gradient of the peak (first derivative data) or even the change in the gradient (second derivative data) can be seen analysed.

6.3 Results

6.3.1 Creation of tablet dissolution images

Tablets A, B and C were imaged and the data stored and corrected to generate a value of reflectance relative to the 100% PTFE standard.

The full datasets contain multiple repetitions of the flow cell over time which builds up to create the hyperspectral dissolution image. A single repetition is shown in Figure 6-2 which has been annotated to show the different areas within the image. The tablet inside the flow cell can be seen to the right of the image along with the PVC holder used to position the tablets and to the left are two other distinct objects. The very left edge contains a piece of white calibration tile, which is used to ensure the accuracy of corrections. The reflectance value of any pixels randomly chosen within these regions should equal 1 ± 0.05 for detector error. The second, more rounded object in the image is a tablet, specifically the same brand as that being imaged, which has been split in half showing a cross section of the tablet. This piece of tablet can be used as a secondary reference sample which is unchanged throughout the experiment.

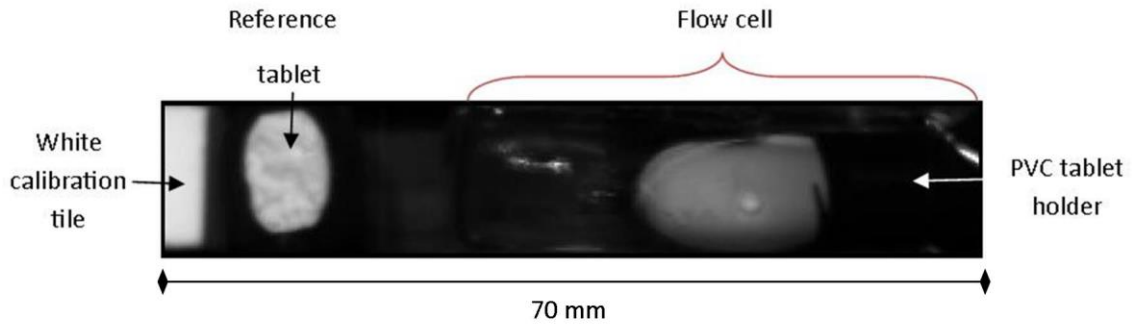


Figure 6-2 Annotated diagram detailing the part of the image

The start of the experiment is always the top of the image, with each new pass of the flow cell over the camera generating the next tablet down in the image. This building up of the image is illustrated in Figure 6-3, where the first 4 repetitions of the dissolution of Tablet A can be seen. The motorised stage moves back and forth to capture the data which creates a degree of symmetry within pairs or couplets of tablets. Every second tablet in the image is inverted as the direction of the imaging has also been inverted an effect most notable in these tablets. A bright spot can be seen on the tablet from the centre of the letter P which is

embossed on the tablet surface. It is possible to follow the location of this P and to see the mirroring of the images as time progress down the image.

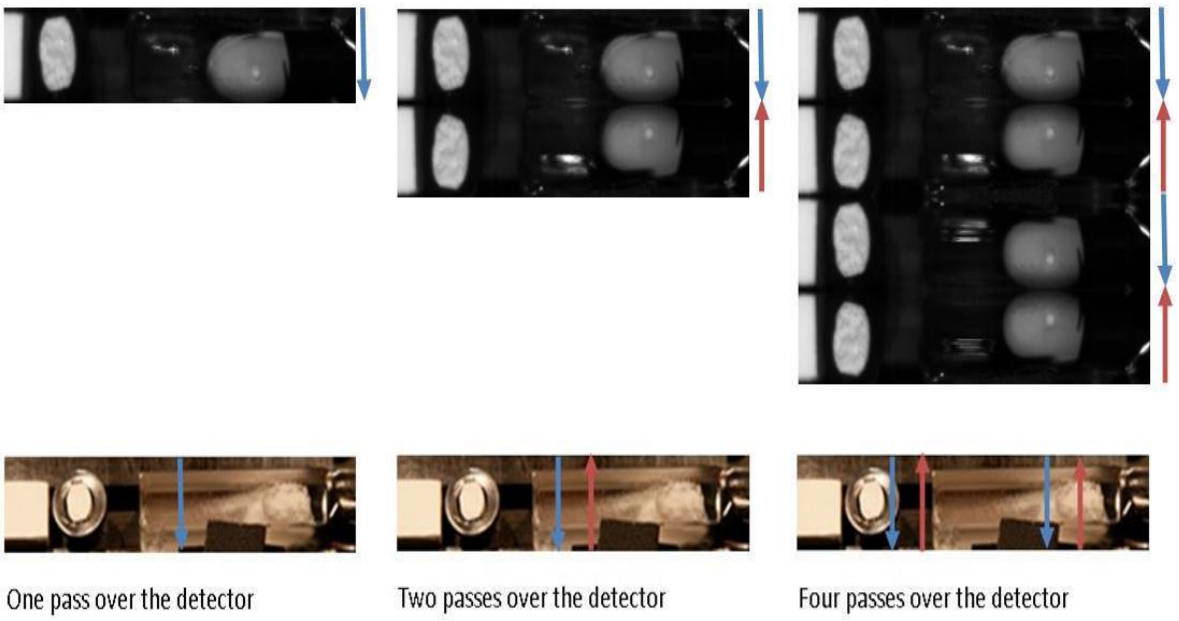


Figure 6-3 A diagram showing how the hyperspectral images are created. The direction of the stage's movement is shown by the arrows and each new pass across the detector generates a newflow cell image.

The hyperspectral dissolution images of the three tablets were all built up in this manner and the resulting images are shown in Figure 6-4 The images, created in MATLAB, are greyscale with the intensity of each pixel being determined by the intensities of the 1281 nm band. These images contain only the first 2000 lines of data which shows most of the initial expansion and breakdown of the tablets. The frame rate of the camera was 10 frames per second so the first 2000 lines contain the data from the first 200 seconds of the experiment.

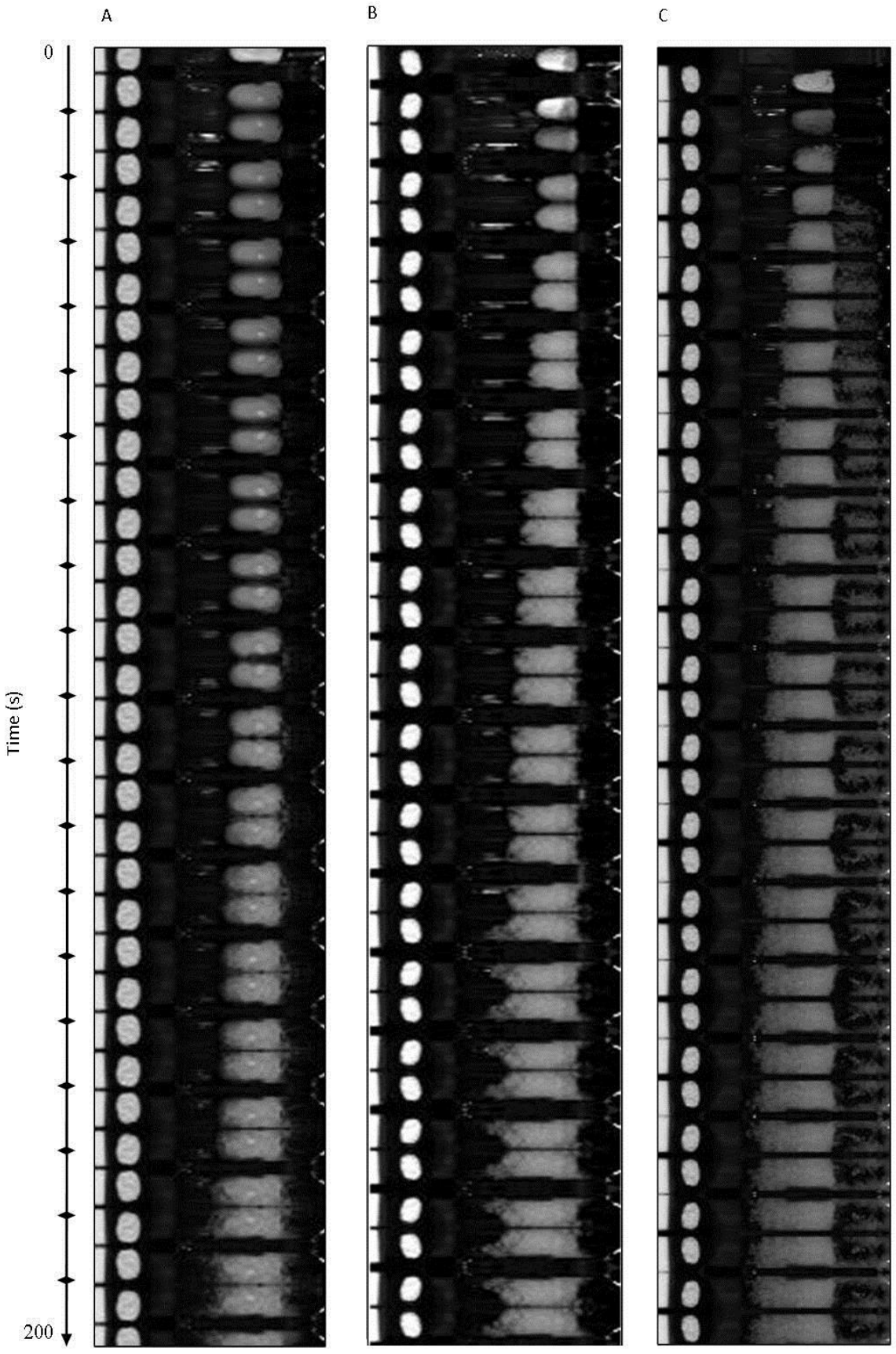


Figure 6-4 The hyperspectral images created during the dissolution experiments. The image is a greyscale relative to the intensity at 1281nm. The annotation on the image denotes the tablet being imaged. Tablet A contains sodium bicarbonate and paracetamol, Tablet B contains just paracetamol and Tablet C contains paracetamol and caffeine

Figure 6-4A shows the dissolution of Tablet A during the first 200 seconds of the experiment. The edges of the tablet are becoming less defined as they begin to break away and dissolve during the experiment and the tablet loses its shape completely within 200 seconds. The changes in this tablet are slower than is seen in Tablets B and C however, the changes in the tablet are still rapid as there is clearly expansion of the tablet material within 200 seconds of tablet being submerged in the acid solution. The expansion of the tablet is tracked in greater detail below (Figure 6-5) where it is quantified and a rate of expansion determined.

The dissolution of the tablets was measured previously using a USP type IV apparatus and the data is shown in the appendix. These results show that the fastest releasing formulation is Tablet B which begins to release almost immediately and reaches 80% release within 300 seconds. Tablet A has a longer lag period before any release of the drug is seen but then very rapidly releases the drug with less than 200 seconds needed to move from no release to full release of the paracetamol. The release data for Tablet C is much slower with approximately 1200 seconds needed for 80% release of the paracetamol.

Tablet A contains carbonates which in the presence of acid will quickly dissolve releasing CO_2 gas. The release of the carbon dioxide gas is designed to aid the dissolution of the API within the tablet by causing rapid tablet breakdown. The rapid rate of release is aided by the carbon dioxide gas being forced out of the tablet. During the UV dissolution the tablet completely dissolves but the HSI shows that there is still tablet material present even after 30 minutes. The change in dissolution properties between the UV and the HSI is most likely due to the steric confinement of the flow cell used in the HSI experiments. Tablet A is a large bulky tablet which uses the effervescence of gas to force apart the tablet and increase the rate of tablet breakdown. When the tablet is placed into the flow cell it fills a large portion of the cross sectional area of the flow cell before any expansion has occurred and so the expansion of tablet material will quickly fill the flow cell, impeding the full dissolution properties of the tablet. However the tablet, while potentially slower releasing than in the UV, still fully breaks down within 4 minutes and the tablets expands more in the direction of flow and less uniformly in all directions

The dissolution of Tablet B is shown in Figure 6-4B, it can be seen that this tablet is much smaller than Tablet A and that the expansion of the tablet appears to occur much quicker. The initial six tablets after the flow cell is filled show tablet expansion but with the tablet remaining in its initial shape. After this period of approximately 45 seconds from the start of the experiment the shape of the tablet is lost and a flat leading edge is formed as the tablet expands in the direction of the flow. The expansion appears to stop around 75% of the way down the flow cell, this is where the outlet for the cell is found and most likely this is why the expansion of tablet material appears to have stopped. Any further expansion in the direction of the flow would most likely expand into the outlet of the cell and be flushed out of the system.

The last dataset in Figure 6-4 is from Tablet C, this is caffeine containing paracetamol tablet. The rate of expansion and breakdown appears to be the highest for all three tablets, which later analysis proved by investigating the rate of tablet expansion. Inspection of even the earliest wet tablet images shows that the tablet is not just expanding. There is evidence of particles in and around the tablet from its breakdown while the expansion of the tablet material is occurring; this tablet breaks down into small particles as it expands. This is contrary to the UV dissolution data which shows that release of API from Tablet C is much slower than Tablets A and B. The UV dissolution data is not run alongside the HSI but was collected at a much earlier date using a USP type IV dissolution apparatus and a different batch. The disparity between the dissolution data and the HSI data is caused by this variation. Either the change in the dissolution vessel has affected the dissolution rate of either Tablet C or Tablets A or B or the formulation of the tablet has changed between the two different batched used in the two different experiments.

The other unusual observations in this image are that there are particles moving to the right of the tablet which is against flow and that the tablet is able to distribute the particles throughout the flow cell. Both of the previous tablets generated images which show the expansion stopping around the out flow point of the flow cell, while Tablet C has particles which go past this point.

The HSI data clearly shows the tablets beginning to expand and then breakdown as the experiment progresses. In Chapter 4 the degree of expansion and breakdown within the flow cell was calculated as a function of how many pixels, within the boundaries of the flow cell, contained tablet data. A function was written in MATLAB for this purpose which was able to separate the individual flow cell repetitions and then count the number

of tablet containing pixels. The number of pixels was then converted into a percentage of the total flow cell to give a uniform metric to compare the different tablet release rates. Figure 6-5A is a table containing images of tablets A, B and C every 100 seconds for the first 400 seconds. This time period shows the initial expansion and breakdown of the tablets and how the tablet material spreads through the flow cell. Figure 6-5B is a plot which shows the rate of tablet expansion as a function of the percentage of the flow cell which contains a significant amount of tablet. This data is for the full 30 minutes of the experiment and each tablet shows a sharp increase in how much of the flow cell contains tablet data over the first 400 seconds, after this time there is also a loss of tablet particles in the flow cell.

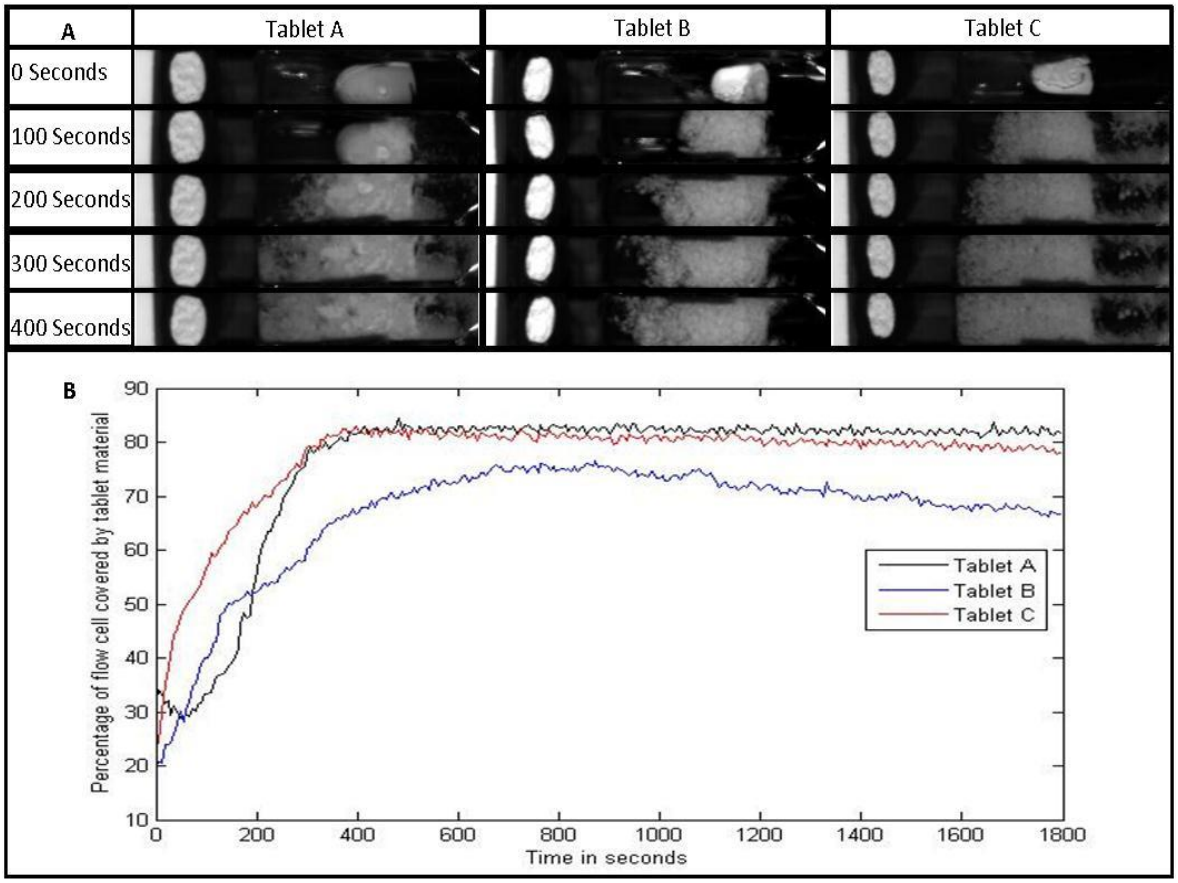


Figure 6-5 The expansion of tablets as A) images from the first 400 seconds which show a snapshots of the expansion and breakdown and B) a plot of the percentage coverage of the flow cell

The earliest stages of Tablet A's dissolution show much a decrease in flow cell coverage which may be caused by a passing air bubble as these cause intense signals which would fool the algorithm. After this initial blip the data looks good and the rates of the expansion can be calculated assuming a first order rate and these results are shown in Table 6-1. The rates were calculated by applying a line of best fit to the data and these images are included

in the appendix. The rate of expansion for Tablet A is the highest which can be seen by the sharp increase in Figure 6-5B, this formulation is specifically designed to use carbonates to increase the rate of breakdown and release which this data clearly shows. Tablets B and C are both complex formulation which use a number of excipients to increase their rate of release over the more basic tablets on the market. Their release rates may be similar and slower than Tablet A but they are still showing complete breakdown within 400 seconds which is a rapid expansion of the tablet. A rate of loss was also calculated which shows how fast the tablet material is clearing the flow cell. Tablet B has material leaving the flow cell at the fastest rate while Tablet A is much slower. The material that is being cleared from the cell should mainly consist of the none-soluble excipients within the tablets such as magnesium stearate and their clearance is not necessarily indicative of the release of the API from the tablets. The rates of release are much lower for each of the tablets than the expansion however this does not mean that the tablet is moving out of the flow cell at this rate only that the pixels in the flow cell are clearing at this rate. Each pixel is over 200 μm in diameter and can contain a great many particles of the tablet stacked on top of each other. The volume of tablet in a pixel is not calculable due to the signal attenuation affecting the signal strength and so the algorithm is unable to account for the movement of tablet particles from a pixel which does not constitute the complete vacating of that pixel. What this means is that the rate of loss from the flow cell is most likely much higher than the results calculated here and that these values are guideline approximations.

	Tablet A	Tablet B	Tablet C
Rate of tablet expansion and breakdown (percentage change/s)	0.1711	0.1102	0.1226
Rate of loss from the flow cell (percentage change/s)	- 0.0005	- 0. 0076	- 0.0022

Table 6-1 Rates of tablet expansion (percentage change per second) through the flow cell and tablet loss from the flow cell

6.3.2 Analysis of the spectra for the tablet data

6.3.2.1 Average spectral profile comparisons

The first spectral analysis performed on these results was to simply inspect and investigate the spectra throughout the data at different times. This experiment was designed such that the tablet spectrum should be less affected by signal attenuation than in previous experiments.

The spectra of 4 different regions in the images were analysed, these were chosen to best represent the data in the image as a whole. The 4 points were: the initial dry tablet, the tablet immediately after wetting, the tablet during breakdown and the tablet after breakdown. The spectra of the data in these regions should detail the major changes present in the tablet as the dissolution progresses while also showing that the experimental design has been successful. The specific locations used are highlighted as black squares in Figure 6-6, these locations all lie on the same vertical line down the image. Choosing the pixels in this way prevents any pixel to pixel variation which may be present in the detector or any differences in the light intensity across the image. The regions chosen are a 10 x 10 square of pixels which covers an area of approximately 2360 μm^2. A 10 x 10 region was chosen as it is a region large enough to give an accurate mean spectrum of the bulk environment while also being small enough to give multiple regions in the tablets for comparison is required.

Figure 6-7 contains the spectral data isolated from the dissolution of Tablet A at the locations shown in Figure 6-6. The first major result from this image is that there is a distinct difference between the dry tablet data and any data captured from after the flow cell has filled with the HCl solution. This shows that there is still a significant attenuation of the signal caused by the solution in the flow cell; however, the degree of attenuation appears less than in previous results (Chapter 5).

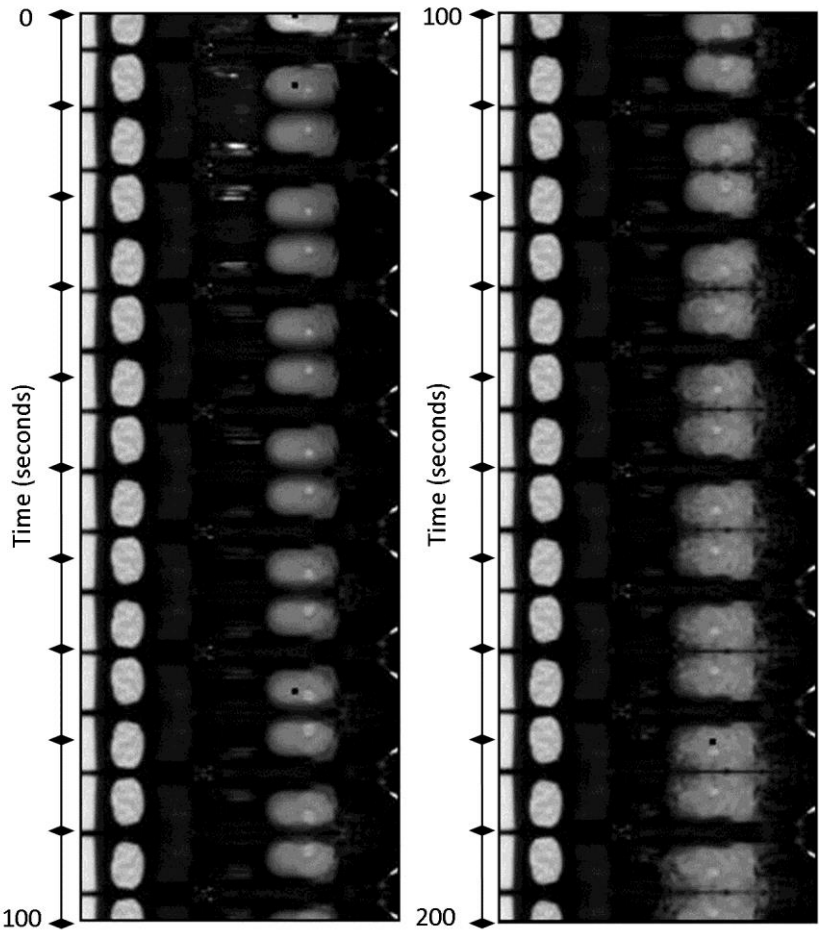


Figure 6-6 The first 3 minutes of the dissolution of Tablet A, the black regions highlight the areas used for spectral analysis

The hyperspectral dissolution image in Figure 6-6 details the dissolution pathway of a single tablet which is imaged repeatedly to create the data. The pre expansion tablet data has a spectrum very similar in intensity to that of the post tablet breakdown data with the peak around 1930 nm being lower in intensity than any of data. The pre expansion data was taken from first wet tablet, the exact location can be seen in Figure 6-6 as the black square within the second tablet repetition. Visual inspection of Figure 6-6 shows that the tablet has not yet begun to expand by this point in the image. It is believed that the low intensity values of the spectra post tablet submersion but pre expansion are due to the longer path length of the nIR photons through the acid solution to reach the tablet. Once the tablet begins to expand the path length of the photons decrease and the attenuation of the signal becomes less pronounced.

This experiment was designed to minimise the path length of the nIR photons through the acid solution, thus minimising the attenuation that the spectral signal receives from the nIR absorbance by the water molecules. Unfortunately while the signal attenuation has been

lowered there is still significant attenuation and loss of characterising peaks in relation to the dry spectra in Figure 6-7. This is not surprising as tablets are designed to allow water to be absorbed into the tablet which aids in disintegration. This means that even with a minimal path length through the acid solution there are still attenuating water molecules present inside the tablets once hydration and swelling has begun. There is one set of peaks which can be seen in the 1600 – 1700 nm region, these small peaks may be useful in identifying the changes in the tablets throughout the dissolution. It is possible that these peaks could be used to follow any changes over time such as a loss of peak intensity or a change in the rate of loss for any one of the peaks. The peaks may all be caused by the same compound with a number of similar but distinct vibrational regions or by different compound with similar vibrational regions. The former can help to track release over time, the latter could show the loss of specific excipients or APIs.

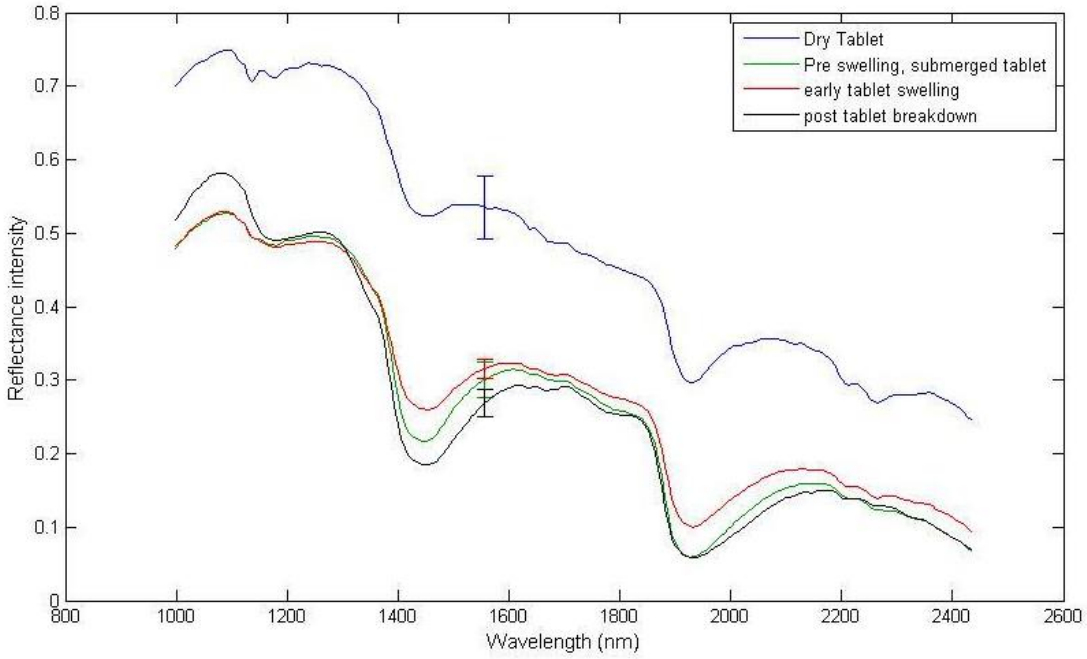


Figure 6-7 The spectra of the four predetermined location in the dissolution of Tablet A

The data used in these plots is generated by averaging 100 points so that a fair representation of the bulk tablet is given in the data. However it is possible that variation across the tablet is significant and that this process is insufficient so the standard deviation of the data was calculated. The standard deviation was calculated for each spectra and found to be 2.5 % (±0.4) of the data, this is a very small standard deviation and confirms that the data is mostly uniform and that the averaging is a valid technique. The standard

deviation is not included in figures as it is just adds unnecessary confusion into the plots, especially in the regions where the different spectra are very close in intensity.

The same analysis was performed on Tablet B with points at specific time being plotted as shown in Figure 6-8, the same distinct pattern of intensity changes can be seen. There appears to be a greater number of peaks in the dry tablet which could be used for identification than in Figure 6-7 but still after attenuation only the 1600 -1700 nm region shares any peaks in common with the original dry data.

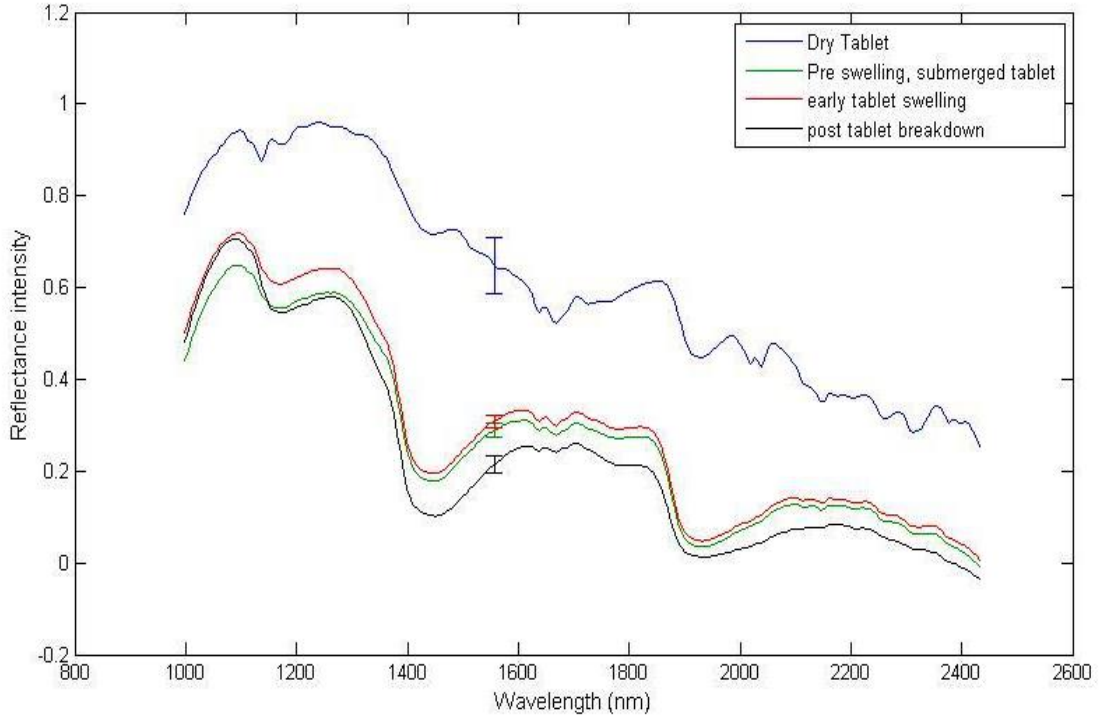


Figure 6-8 The spectra of the four predetermined location in the dissolution of Tablet B

The last tablet, Tablet C was analysed in the same manner and the plots are shown in Figure 6-9. Again there is a significant degree of attenuation to the signal after the flow cell has filled and the tablet hydration and expansion has begun. The post breakdown data appears to show the greatest intensity of all the 'wet' data points which is unusual and it is possible that this effect comes from the manner of Tablet Cs breakdown. Tablets A and B expand in the direction of flow, this mass is then slowly eroded over time resulting in the eventual breakdown of the tablet. Close inspection of Figure 6-4C shows that the tablet does not expand and breakdown in the same manner as Tablets A and B but instead quickly breaks down into large granular components which then move in the direction of the flow. These granules are easiest seen at the edges of the tablet and at the back of the

flow cell. The granules may take some time to settle causing the increase in intensity to a maximum post hydration value to take longer than the other tablets.

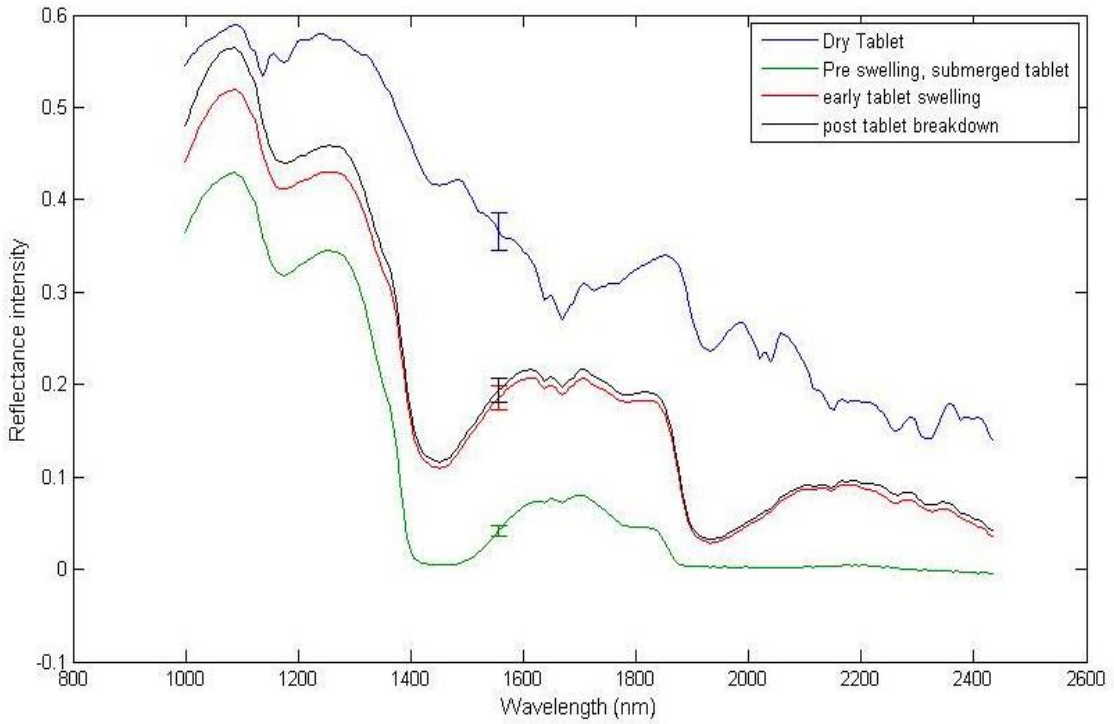


Figure 6-9 The spectra of the four predetermined location in the dissolution of Tablet C

By investigating the spectra in its simplest form there is limited information about the overall release of drug which can be attained. The UV dissolution data shows that as the tablet beings to break down the API are release at a rapid rate. The initial expansion and breakdown shown in the HSI is releasing a large portion of the API from the tablet and these results show that we cannot simply use signal intensities until after this release has occurred. Thus simple spectral analysis like this cannot be used to show the release of the drug as the spectra of the tablet will not settle until after this time and the very early data is still too attenuated to be of use.

6.3.2.2 Comparison of the different tablet spectra at different time points

The difference between the tablets at set time points can show how the spectral profile of the tablets changes over time with the breakdown and release. It is also possible to compare the spectra of each tablet at the same time point. This analysis can show the differences and similarities between the tablets through time. Figure 6-10 shows a comparison of the three dry tablet spectra.

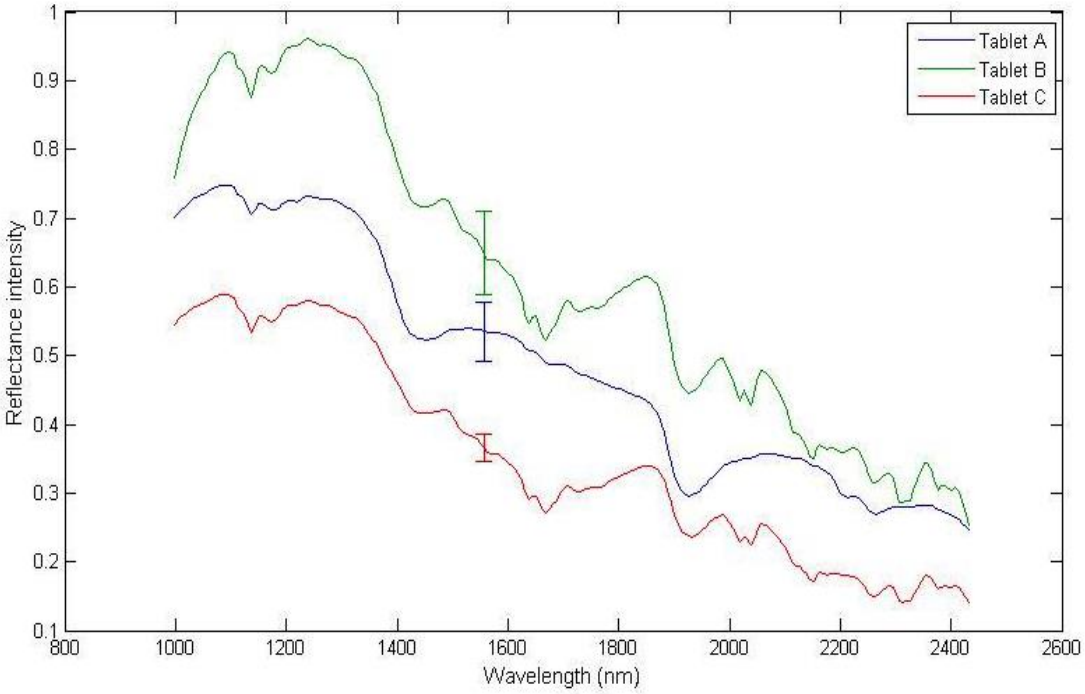


Figure 6-10 Comparison of the dry spectral data from the three tablets

Peak Position	Functional Group
1137 nm	Methyl group 2nd overtone stretch
1425 nm	Hydroxyl group 1st overtone stretch.
1660-1700 nm	Aryl and alkyl methyl 1st overtone stretch
1900 nm onwards	O-H, C-H, N-H and C-C combinations

Table 6-2 Tablet of functional groups stretches present in Figure 6-10

The position of the major peaks in Figure 6-10 and their corresponding functional group are shown in Table 6-2. The majority of the peaks are from various C-H and O-H stretches however, the region after 1900 nm is difficult to identify which stretches, or vibrational modes, are causing the peaks. This region consists of numerous combinations of different vibrations which are energetically very similar and so it is difficult to say with certainty which peak is caused from which combination. Each tablet also shows a trend to decreasing signal intensity; this is due to the wavelength dependence on the scattering of

photons. The intensity of scattered light is proportional to λ^{-4} so the longer wavelengths of light will have a lower intensity which explains the trend in the data.

Tablet B and Tablet C have a very similar spectral profile with the majority of the constituent of these tablets being acetaminophen with all other compounds being in a minority of the total tablet constituents. Tablet C does include one major difference which is that it contains 60 mg of caffeine. Tablet A shows much fewer spectral characteristics than the other two tablets which may be due to the structure of the tablet. A major component of the tablet is sodium bicarbonate and this comprises of only a single set of active vibrational modes, visible in the near IR, coming from the O-H bonds. This reduced number of stretching frequencies could be causing the smoothing effect on the tablet data.

One of the major variations between the tablets is the intensity of the spectra, this could be a physical difference between the tablets but is most likely simply a photonics issues. One approach to determining if the difference between the spectra is a scaling factor is to mean centre the data. This process shifts each spectrum individually so that their mean spectral intensity lies at 0 and the spread of the data is reduced by an amount relative to the distance from the mean to the 0 point. When this is done the spectra of Tablet B and C will overlay and any spectral difference can be clearly seen, if there are no differences then the results seen in Figure 6-10 is simply a scaling effect. The mean centred data was plotted in Figure 6-11 and Tablet B and C now lie closer to each other but not on top of each other. This result implies that the spectra are different and that the intensity differences are a real and genuine effect. Such a result is not unexpected as while the major component of two the tablet is the same they both contain a number of different excipients which should effect their nIR spectrum.

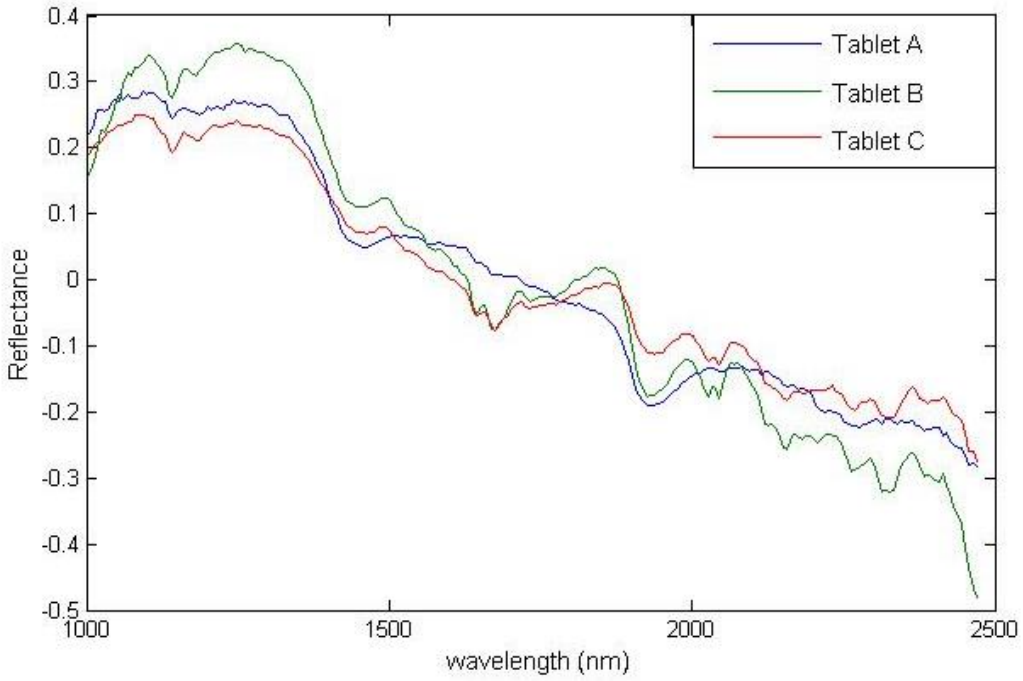


Figure 6-11 A plot showing the mean centred dry tablet data

The dry tablet data in Figure 6-10 details any difference that the HSI can detect in the normal tablets before any expansion or release of drugs. Once the acid has entered the flow cell, expansion begins and the spectra show how the tablet is reacting. The comparison of these spectra is shown in Figure 6-12 and these spectra are much weaker than those seen in the previous plot which contained data from only a few seconds earlier in the experiment. The cause of this variation is the attenuation caused by the water molecules in the flow cell which absorb the nIR photons. The longer the path length of the photons through the acid solution in the flow cell the greater the degree of absorption of photons by the molecules of water in accordance with the Beer-Lambert law. Figure 6-12 implies that Tablets B and C have yet to expand significantly, although a small degree of tablet expansion may be present, and the path length through the flow cell to the tablet is at a maximum. This maximum causes greater attenuation resulting in the minimum tablet intensities in the image. Tablet A however, has a much stronger spectral signal which implies that the tablet has begun to expand very rapidly almost immediately after the liquid has entered the flow cell.

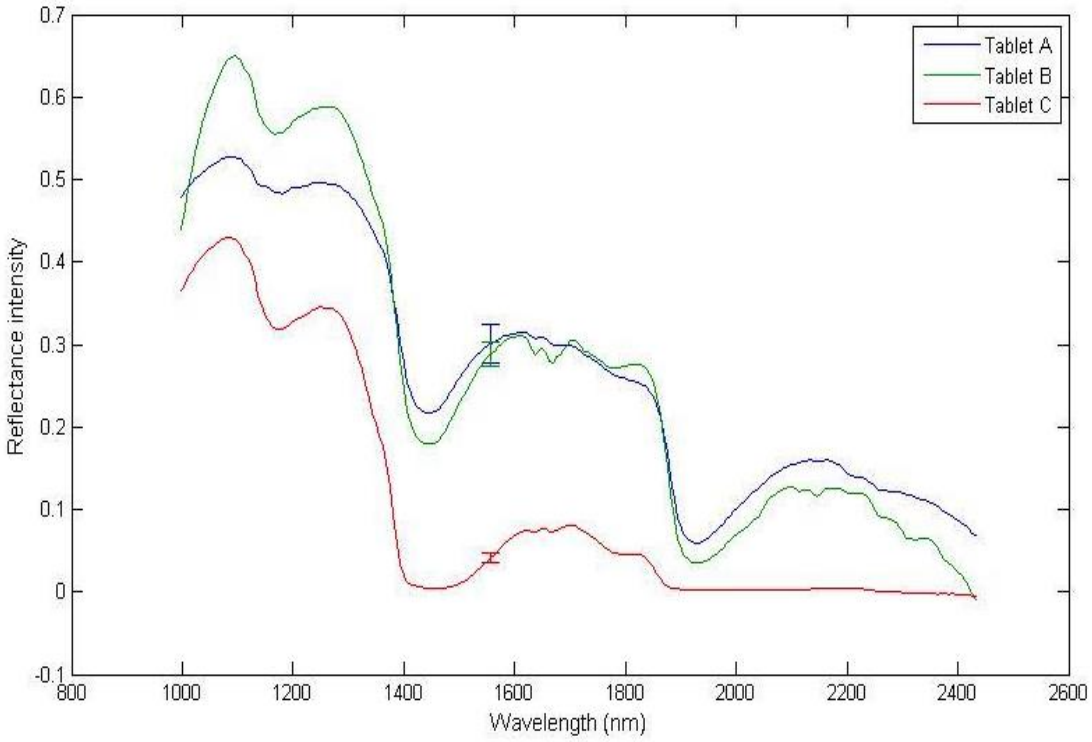


Figure 6-12 The spectra of each tablet as it has been submerged in hydrochloric acid.

The next data point was a few seconds after each tablet had begun to show physical signs of expansion, these points are shown in Figure 6-13. Again Tablet A has a very different spectrum compared with the others, there is substantially less peak definition with most of the spectrum being flat. The peaks around 1450 and 1925 nm are from the water attenuation and other than those there are only a couple of small very broad peaks. Tablets B and C however, show much more spectral definition and again are very similar but with Tablet B being the more intense of the tow as with Figure 6-10. One of the largest differences which is appearing in Tablet B is the region after 2200 nm, in which there is a strong negative gradient with the signal intensity quickly trending to 0. Both Tablet A and Tablet C have a much shallower negative gradient in this region.

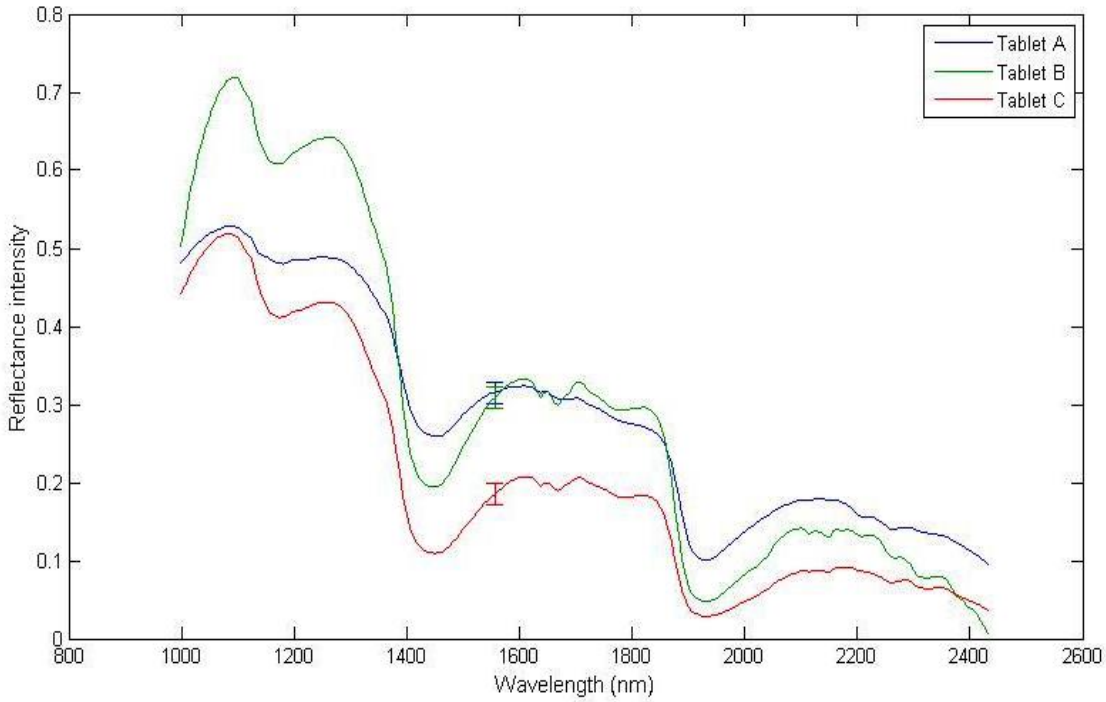


Figure 6-13 The spectra of each tablet at 120 seconds into the dissolution

The last time point was used to show how the spectra can compare once the tablets have completely broken apart and started to disperse across the flow cell. The plot of these results is shown below in Figure 6-14 and it is clear that the spectra are beginning to become more homogenous. The intensities of each tablet are closer than in previous results and this allows for a more accurate comparison between the different spectra. The overall shape and form of the spectra is very similar between each tablet which is a change for Tablet A which has previously looked different to the other tablets. The intensity of Tablet B is different from Tablets A and C as it has the highest intensity at the lower wavelengths, the lowest intensity at the highest wavelengths and has minima in the water peaks equal to Tablet C.

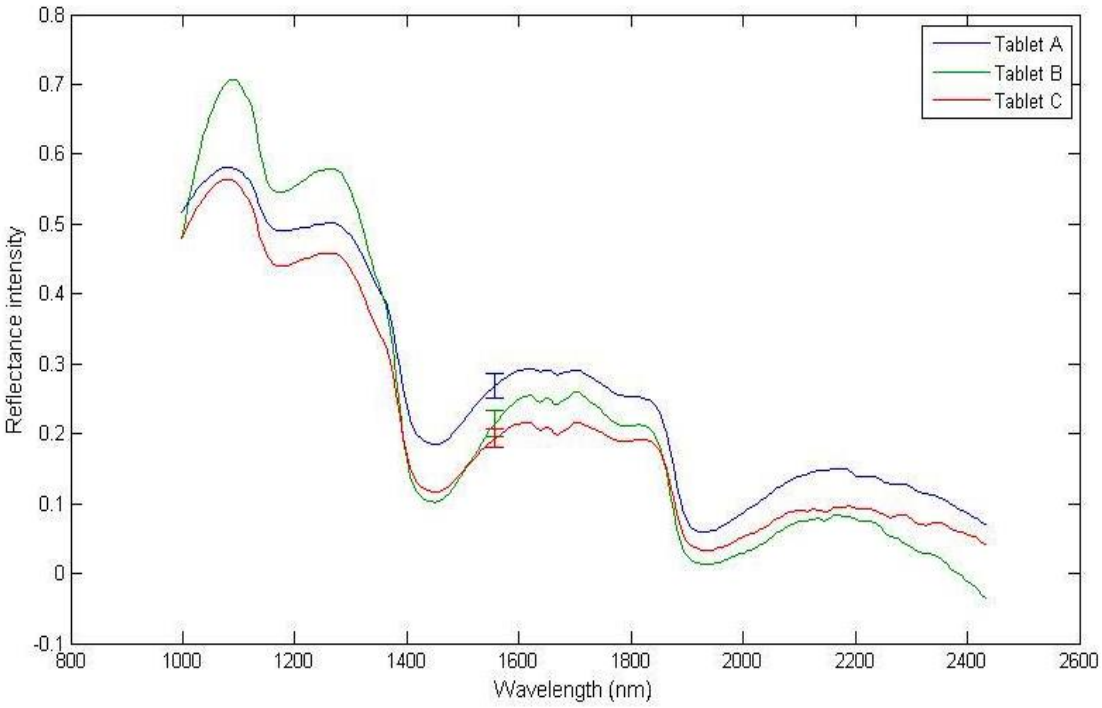


Figure 6-14 A comparison of the spectra from each tablets once the breakdown is nearly completed

6.3.2.3 Line profile across the flow cells

Simply investigating one point in time or one point in the cell is not sufficient to show the real changes occurring in the flow cell. By taking a line through the cell and plotting the spectrum at each point onto a single plot it is possible to see changes in the tablet. If any major excipient or the API is lost from any point along the line it should be possible to see this as an anomaly in the plot.

The first line profiles to be attempted were to see any changes across the flow cell at a set time by investigating the spectra along the flow cell from left to right. The location for this line of pixels is one that dissects the central point in the tablet as this increases the volume of useful spectra and should minimise any edge effects. Figure 6-15 shows an example of where the data for the line profiles is located. The spectra at each pixel is used to create the line profile, a total of 120 spectra were used in each line profile. Each pixel is 216 μm long so the total area investigated with this technique is 25.9 mm. These spectra are then placed alongside each other in a 3D plot to generate the line profiles.

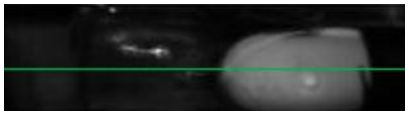


Figure 6-15 A diagram showing an example location of the line profiles

Four line profiles were generated each at a distinct and different time in the image with the exact locations being very close to those used in 6.3.2.1. These plots were created using the mesh plot function in MATLAB and some are shown in Figure 6-16, the front facing axis is the spatial location down the tablet in the flow cell and the retreating axis is the wavelength.

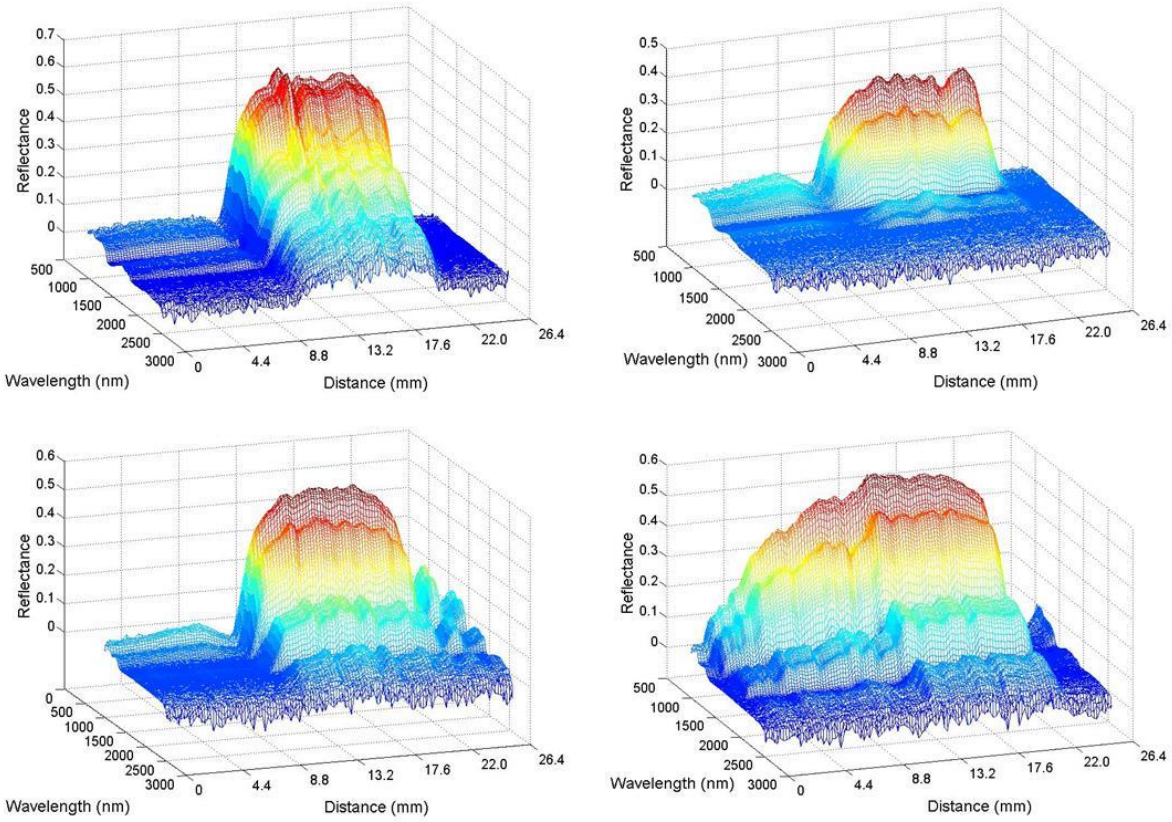


Figure 6-16 Mesh plots showing the spectra across Tablet C at set points in the image which are: A) dry data (5 seconds), B) wet data (10 seconds), C) early expansion (22 seconds) and D) late expansion (100 seconds).

The plots in Figure 6-16 are from Tablet C but the results from Tablet A and Tablet B are very similar. Plot A shows the initial dry tablet when no expansion of the tablet or change should has occurred as the tablet is not yet submerged in the acid solution. Along the tablet (left to right in the image is left to right on the tablet) there are some small changes in the intensity of the signal which could easily be background noise as the deviations are slight. The edges of the central block fall away quickly which is as expected when the tablet is

whole. Figure 6-16 B is only 5 seconds later but already a small amount of material can be seen down the flow cell, with a small set of peaks visible to the left of the plot. The intensities across the bulk of the tablet are less uniform than before, this may well be showing the tablet starting to break apart.

By plot C we are again seeing a more uniform distribution, this is because the tablet has now expanded to fill the flow cell and the spectral intensities have all risen just like Figure 6-9 describes. Thus the none uniform distribution in plot B is most likely showing the expansion and breakdown of the tablet to the surface of the flow cell at differing rates. The last plot in Figure 6-16 contains data from much later in the dissolution. The right side of the tablet still appear to be largely unaffected however, the left side of the tablet has broken down resulting in the shift of tablet particles through the cell, thus the sloping change in intensities on the left of the image.

Each of the tablets shows this same progression, starting with a homogenous intensity across the tablet then progressing so the left side of the plot beginning to show signs of the tablet breakdown. The other images can be found in the appendix for reference.

6.3.3 Principle component analysis of the HSI data

6.3.3.1 Introduction to PCA methods used in this work

Analysis of the spectral data has produced data on the rate of the tablet breakdown and expansion, it has also shown how the release of each tablet is different with Tablet A being of note for its dissimilar spectrum and expansion relative to Tablets B and C. There was not however, any information on the release of the API from the tablets which is the aim of this work. By using more complex analytical techniques it should be possible to investigate the changes within the tablets which cannot be easily seen by the more basic analysis of the spectrum used previously. Using principle component analysis (PCA) allows for the slight variations between the spectra to become more apparent and also shows the reasons for these variations, the exact wavelengths in this case. Principal component analysis is explained in greater detail in Chapter 2.5.1 and the reader is encouraged to read this section for clarification on terms and how the analysis works if they are unsure.

These images contain a wealth of data and two different styles of PCA are required to best analyse the data. The first approach is to take an individual or mean spectrum from each tablet in the image and to compare these in the PCA. Using the individual pixels allows for

the separation of the data through time so that any variations which are occurring during the dissolution can be tracked. This approach uses a biplot of the scores values for each tablet sample to show the clustering and separation of the data so that the patterns and variations between the tablets as the dissolution progresses can be followed. The advantage to this style is that it is quick and able to show the small variations between tablets in a clear fashion. The only real disadvantage is that it requires more work to distinguish the cause of the variation between the tablets than other approaches as loadings must be inspected to better determine the reasons for the splitting and clustering shown in the biplots.

The second approach to the PCA of this data is to run the PCA on the entire image and this can highlight any difference within the tablets over time such as the release of any one component from the tablet. The visualisation of these results is not, as in the other approach, in a biplot but instead the entire image is recreated with the scores values used as pixel intensities. This gives clear way to see exactly what variations the PCA is using at each component which is the major advantage to this style of PCA it also has the potential to distinguish variations within the tablet which the single spectra approach would not do. The disadvantage to using PCA in this way is the increased processing time required to run the analysis and to create the images.

6.3.3.2 Principle component analysis using individual spectra

The data from a single pixel location over time should show any changes or groupings in the spectral data throughout the dissolution. To better improve the data and to make it cleaner and easier to interpret only one sample per tablet was taken and plotted. The sample is a mean spectrum taken from a 10 x 10 area of pixels within the centre of the tablet. An average spectrum is used as it best represents the tablet a whole without any individual tablet variations causing an unintentional skewing of the results. A mean sample spectrum was collected from the centre of each tablet repetition in the HSI image for the first 2 minutes of the dissolution. It took 6 seconds for the flow cell to complete one full pass over the HSI camera and so 20 sample spectra were created to be put through principle component analysis.

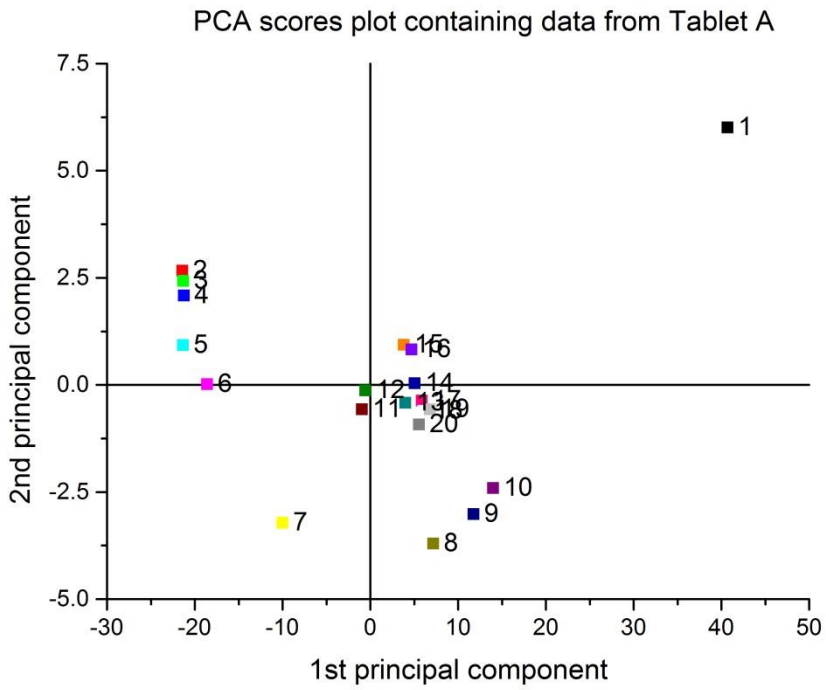


Figure 6-17 PCA scores biplot showing the scores of the first two components generated from the PCA of the data from the HSI dissolution of Tablet A.

Figure 6-17 is the scores biplot of the first two principle components generated by the PCA of Tablet A. Each sample in the plot is numbered to show where the data was extracted from with number 4 being the 4th tablet (from the top) visible in the HSI dissolution image of Tablet A (Figure 6-4A).

The most notable point in Figure 6-17 is the first data point, this is located quite far from the rest of the data. This change reflects the difference between the dry and wet tablet spectra and is a very strong outlier. PCA analysis involves the comparison of variables from different samples to generate a matrix of variance; therefore one value which is strong outlier has the potential to skew the data significantly. To correct for this the data point was removed and the PCA analysis was recalculated.

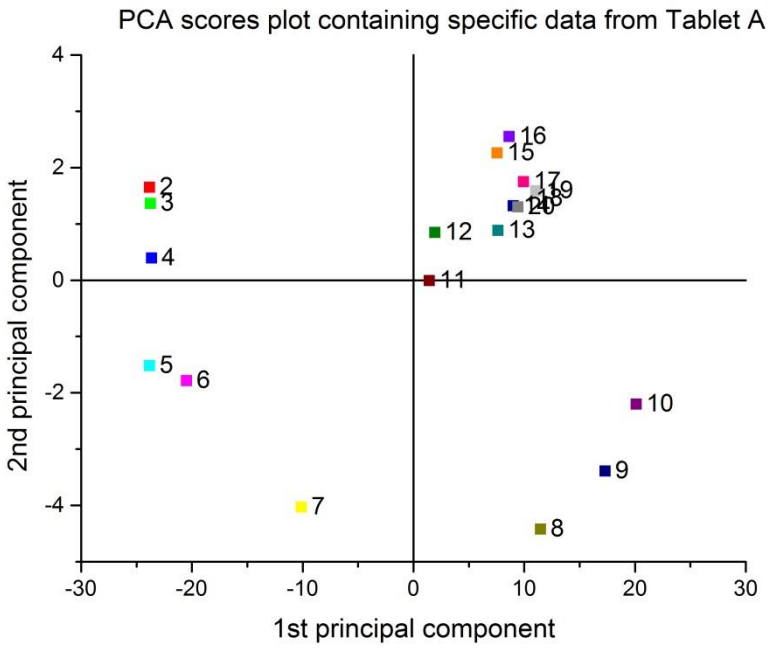


Figure 6-18 The data from Figure 6-17(for Tablet A) recalculated to remove outlying effects

The recalculated scores plot is shown in Figure 6-18; the numbers assigned to each point are the same as in the previous figure. To the left of the plot is a series of data points with three very tightly clustered and 3 more moving away in linear fashion down the y axis. These are the data points from tablets 2 – 6 and they appear to show the change in the tablet over time as it begins to expand and change. There is also a linear progression from points 8 – 10 which are the last data points outside of the central cluster. The central cluster contains every sample from number 11 – 20 which implies that at from point 11 there is very little change in the tablets.

When principal component analysis is performed it is standard to also generate an eigenvalue for each principal component which shows the relative size of that component. This is a measure of the degree of the total variation within the data which can be described using that principal component. It is a simple process to use these eigenvalues to generate a percent of the total variation described within a component and these values for the previous two PCAs are shown in Table 6-3. The first component is showing the majority of the data with over 96.5% of the total variation explained in just this component in both analyses. The tables also shows that the removal of the first erroneous data point makes very little difference to the scale of the components with only a 0.3% change in the total variance explained in the first principal component.

	Percentage of total variation explained by the principle components.	
Principle component	Figure 6-17	Figure 6-18
1	96.6	96.9
2	98.7	98.8
3	99.6	99.4

Table 6-3 The percentage of total variation shown by the first three principal components in Figure and Figure 6-18

The loadings are also created during PCA and can be used to determine which variables are having the greatest effect on the principal components. The more intense the peak in the loadings the greater the variation between the samples at these wavelengths and using this information the cause of groupings in the scores plots can be identified. The loadings from the first two principle components are shown in Figure 6-19.

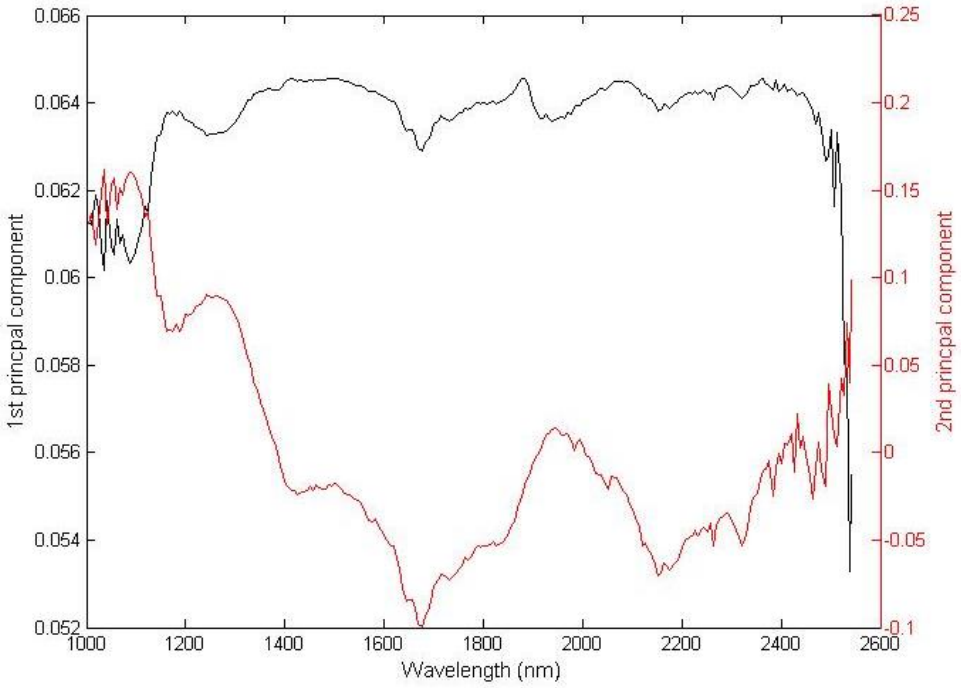


Figure 6-19 The loadings for the first two principle components from the PCA analysis of Tablet A

The loadings from the first component show very little to distinguish between the results; the peaks at 1250 nm and 1700 nm are the most prominent and should be the cause of the variation between the samples. The loadings from the second principle component detail a much greater range of useful peaks with two peaks being of particular interest. The peaks at 1250 nm and 1950 nm are intense and lie in a position similar to peaks in the tablet spectrum in Figure 6-7. These results imply that the variations seen in the scores plot are caused by changes in the peaks at 1250, 1700 and 1950 nm during the dissolution. Inspection of these wavelengths in the spectrum of the tablet highlights three major areas of intensity change throughout the experiment.

In previous chapters PCA has shown how the tablet is expanding as a function of the signal attenuation. These PCA results are complimentary and are showing how the tablet breaks down with clear separation of the initial dry data, the expansion of the tablet and then a steady state where the tablet data clusters together. This clustering is coming from the end of the tablet expansion at a point where the tablet should be breaking apart. The steric problems of a large bulky tablet within a relatively small flow cell where seen in the image analysis of Figure 6-4 and these problems are what keep the samples clustered in the PCA scores biplot. The tablet is unable to break apart as rapidly due to the confinement of the tablet and so there is little change in the physical composition in the bulk of the tablets.

The scores plot from Tablet B and tablet C show also show very little other than a separation of the dry data, early expansion and late expansion data. Inspection of data from after this point or of every data point for the first 120 seconds still shows the same shape and form to the results with no discernible pattern other than the change in the physical state of the tablet. This technique is much more advanced than the simple spectral analyse implemented in earlier approaches but is still very limited in the amount of information it can produce.

6.3.4 Processing and Principle component analysis of the image data

The second style of principle component analysis used is to run the analysis on the entire image to investigate any spatial variation in the data and also potentially any changes over time. In previous experiments a select number of mean spectra were used in the principal component analysis but for this approach every pixel in the image was inputted into the PCA. Once the PCA was run and scores values for each pixel created, images could be

made by using these scores values rather than the wavelength intensities used in the previous images. The initial approach to this used the same corrected data as in 6.3.3.2 and showed very little variation of interest. Most of the results simply detailed variation between the different components of the system with no new useful data. The first (and second) derivative with respect to the change in intensity across the spectrum is a common form of pre-processing applied to nIR datasets. By using the derivative of the data it is possible to better determine changes in the shoulder peaks which are common throughout nIR spectra. The derivative was calculated using the methods discussed in 6.2.5.2 and then the PCA was rerun on this new dataset. Analysis of the PCA data was performed by investigating the scores values for each pixel at each principal component. The scores values were used to generate greyscale images which could be examined for any anomalies or regions of difference.

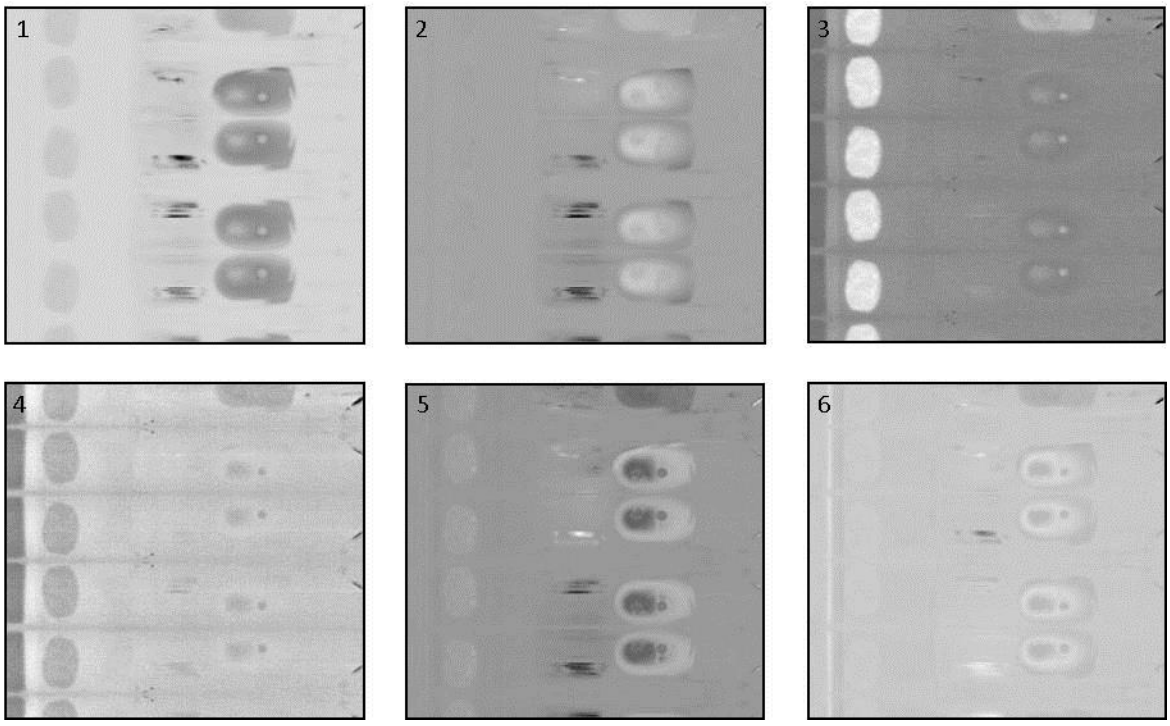


Figure 6-20 The scores results from the first six principle components generated by the PCA of Tablet A. Shown as the greyscale images with the scores values being the scaling factor

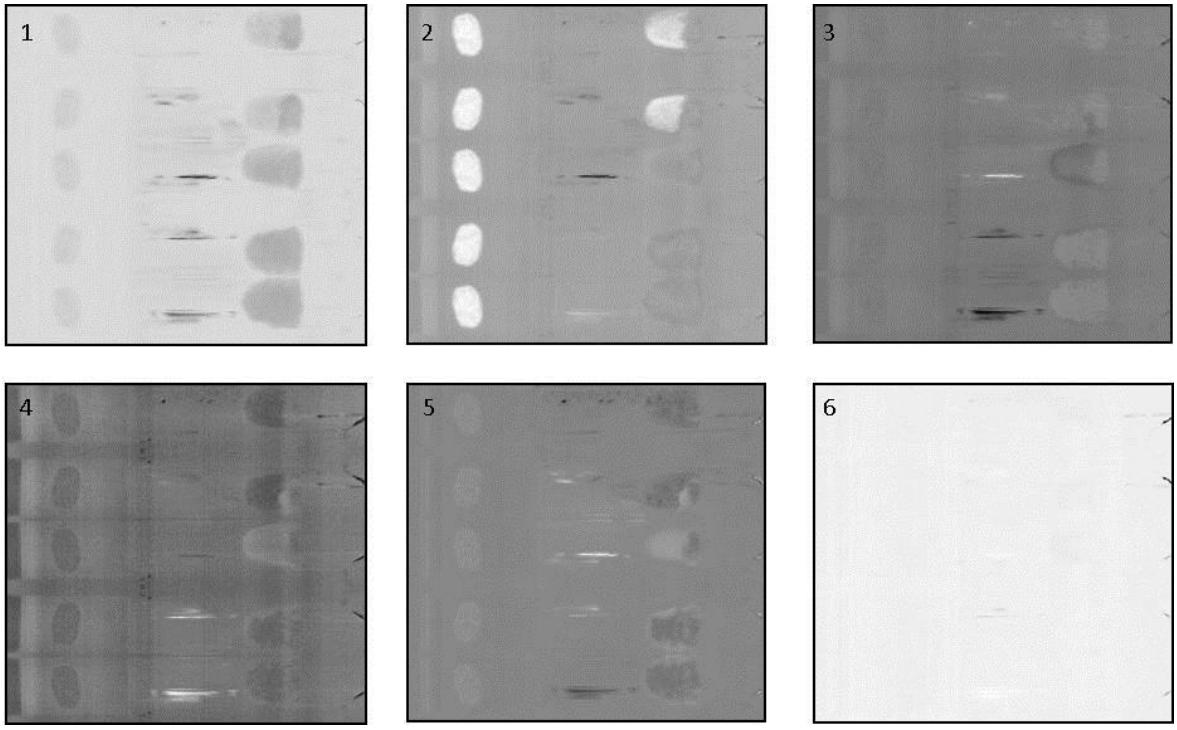


Figure 6-21 The scores results from the first six principle components generated by the PCA of Tablet B. Shown as the greyscale images with the scores values being the scaling factor

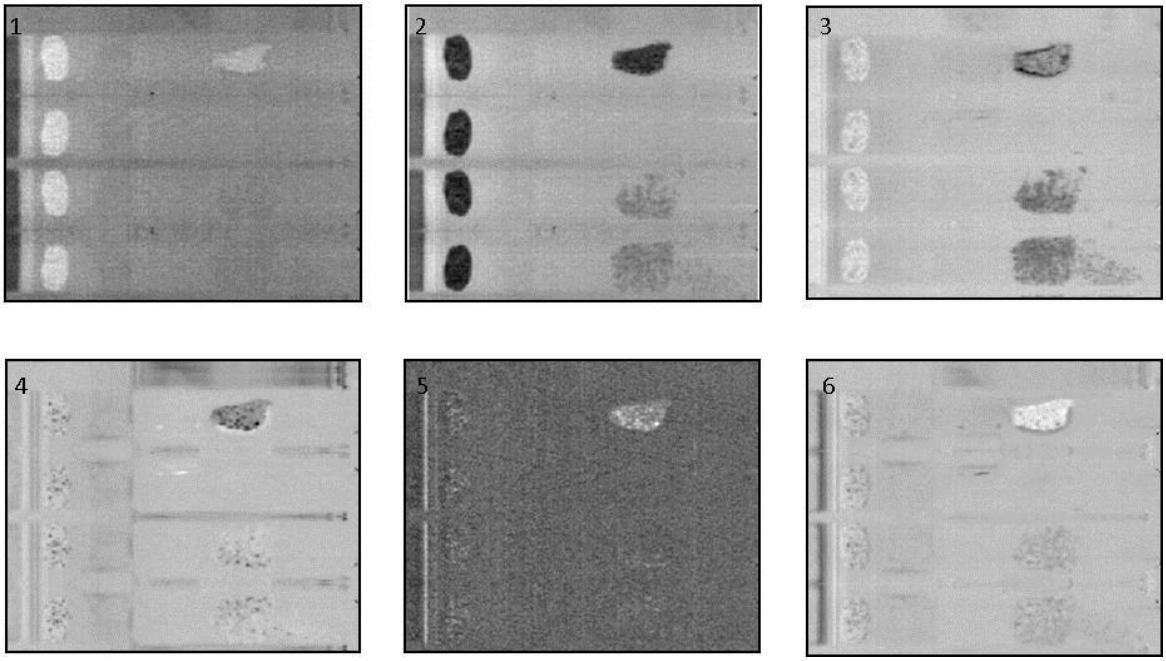


Figure 6-22 The scores results from the first six principle components generated by the PCA of Tablet C. Shown as the greyscale images with the scores values being the scaling factor

	Percentage of total variation explained by the principle components.		
Principle components	Tablet A	Tablet B	Tablet C
1	57.25	50.12	56.48
2	71.32	69.01	73.35
3	83.58	80.38	84.05
4	92.11	89.82	92.91
5	95.11	92.39	94.91
6	95.99	93.66	95.77

Table 6-4 A table showing the cumulative variance detailed within the principal components for the three tablets

Figure 6-20 shows 6 separate images which have been generated from the PCA analysis of the Tablet A data; the results in Table 6-4 shows that the first four principal components contain over 90% of the total variation within the data. The images contain data from the first 30 seconds of the experiments and the variations seen in these images are between different objects in the flow cell and not a change in the data over time. This is not a surprising result as PCA separates out the data with the greatest variations being shown first, the second greatest variation being second etc. The biggest variation in the image of the flow cell is between the different physical components within the cell and not the small changes in the tablet data over time. The exception to this is that the top tablet is of the dry tablet before submersion and this tablet is often shown as being considerably different to the remainder of the data which is not unexpected.

The first two scores images separate the tablets from the surrounding environment and also show a degree of homogeneity between the tablets, an air bubble is also clearly marked as a region of very dark intensity near the centre of the images. This air bubble can be viewed by inspection of Figure 6-4A, with the air bubble being slowly drawn out of the flow cell

by the flow of the acid solution. The third and fourth principle components appear to separate the two reference objects to the left of the image. These scores plot also show that the top tablet is not submerged as it is closely linked to the strong signals coming from the dry tablet ad the calibration tile. Once the tablet is submerged the attenuation of the nIR signal by the water molecule causes the tablet spectrum to be altered, as shown in chapter 5, and so these tablets do not show up in these scores plots. The final two components show the variation within Tablet A as a function of the topographical change throughout the tablet, the most striking variation is between the edge and the central bulk of the tablet. The scores plot clearly shows the edges of the tablet to be very pale with the centre being much darker which implies a very large gap in intensity values. This difference is explained by the change in the path length of the nIR photons through the acid solution to reach the tablet. This slight difference in the path length and reduced thickness in the tablet for nIR scattering and absorption causes a significant difference in the attenuation of the signal and the spectral response detected for these regions of the tablet.

Figure 6-21 contains the scores images from the first derivative data of Tablet B, both this tablet and Tablet A contain acetaminophen as the major ingredient. However, the scores plots show some stark differences. The first component image highlights the tablets but it also shows some degree of variation within the tablets with the first two tablets being much fainter on the left edge. The second scores images separates the dry reference tablet and the first two tablet repetitions, this implies that the reason for the variation in the first scores image is the hydration and wetting of the tablets. The remaining 4 scores images in Figure 6-21 show very little useful separation despite being responsible for almost 14% of the total variation.

The scores images for Tablet C are shown in Figure 6-22 and as with Figure 6-20 90% of the total variation within the data is explained by the first four principle components. In each component the seconds tablet is not visible and the images imply that the flow cell is empty. It is not clear what is causing this disparity but it is most likely that the tablet is raised up by the flow of the acid solution filling the cell. As the tablet is raised up by the acid solution the path length through the acid is maximised and the signal attenuation severe. Once the flow cell is filled the tablet settles back on the base of the flow cell and is visible again. The first component details the variations within the image from the dry tablet data, thus the reference tablet as well as the initial tablet data is pronounced in the image. The second scores image describes two separate sets of variation; the first is the same as the first scores image, with the dry data being coloured black. The second

variation with the scores plot is the bottom two tablets which have already started to expand and break apart. The fourth principle component shows black pixels within the tablet images which are distributed in a none-uniform manner throughout the tablet. The location of these black pixels is constant throughout the image (the fourth tablet is inverted due to the back and forth motion of the scanner). This implies that these black pixels are a physical effect inside or on the surface of the tablets. Figure 6-23 is the same image as the one in Figure 6-22 but showing data for a much longer period of time. The black pixels are visible throughout but the number of them is decreasing over time, which implies so kind of loss or change over time. It is important to note that the principle component analysis itself is not showing a change over time but rather it is showing a spectral difference within the tablet. The change within the tablet which is generating these regions is much more significant than the change in the tablet over time as these regions disappear. Thus the PCA is showing the greater variation of the regions and not the lesser variation of the changes within the tablet over time. Once created however, the scores images can be analysed as to determine a change in the tablet as the dissolution occurs.

The pixels which have been coloured black in this PCA scores images have a lower intensity than the bulk of the tablet, thus the colouration in the greyscale image. It is possible to analyse the number and intensity of these pixels in each of the repetitions of the flow cell to generate an understanding of how the tablet is changing throughout the dissolution.

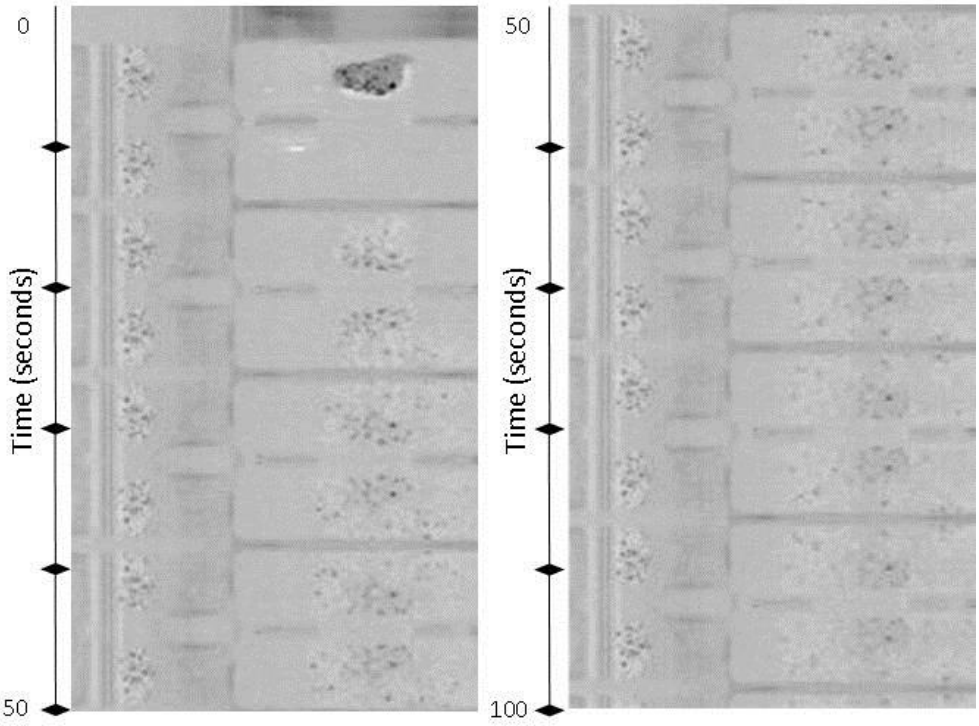


Figure 6-23 The scores plot from the fourth PC of Tablet C, the right hand image continues from the bottom of the left hand image.

The reference tablets on the left of the image also show these black regions throughout and the reference tablet is a tablet which has been split down the centre so that the image contains a cross section of the inside of the tablet. That the reference tablets also shows these black regions proves that this effect is not just an artefact on the surface of the tablet or within the flow cell itself. These black pixels never move or decrease in intensity which shows that the change within the flow cell is due to a physical change in the system.

6.3.5 Spectral analysis of the black pixels

Figure 6-22 shows that Tablet C contains some very interesting regions which are highlighted as black pixels. By using the data in Figure 6-22 the locations of the black pixels were calculated and the spectra of these regions investigated. An average spectrum was created from the data of 8 separate pixels and was then plotted against a mean bulk tablet spectrum. This graph is shown in Figure 6-24; the shape of these two spectra is very similar with little to no difference at most of the wavelengths. However, around 2250 nm there is a clear dip in the nIR spectrum of the black pixel which is not found in the bulk spectrum. This is shown in Figure 6-24 as a blue region with an arrow and it is this spectral difference which the PCA is using to separate the data, giving the black pixels. The difference in the spectra is slight but distinct which implies that this is showing a change in

the relative chemical composition of the tablet. A new or increased concentration of a compound is shown in these regions which is not present throughout the rest of the tablet. The most probably cause of this is an increase in caffeine concentration, this is the second most common compound in the tablet and if the mixing is not homogenous it is possible that areas of higher caffeine concentrations can be found.

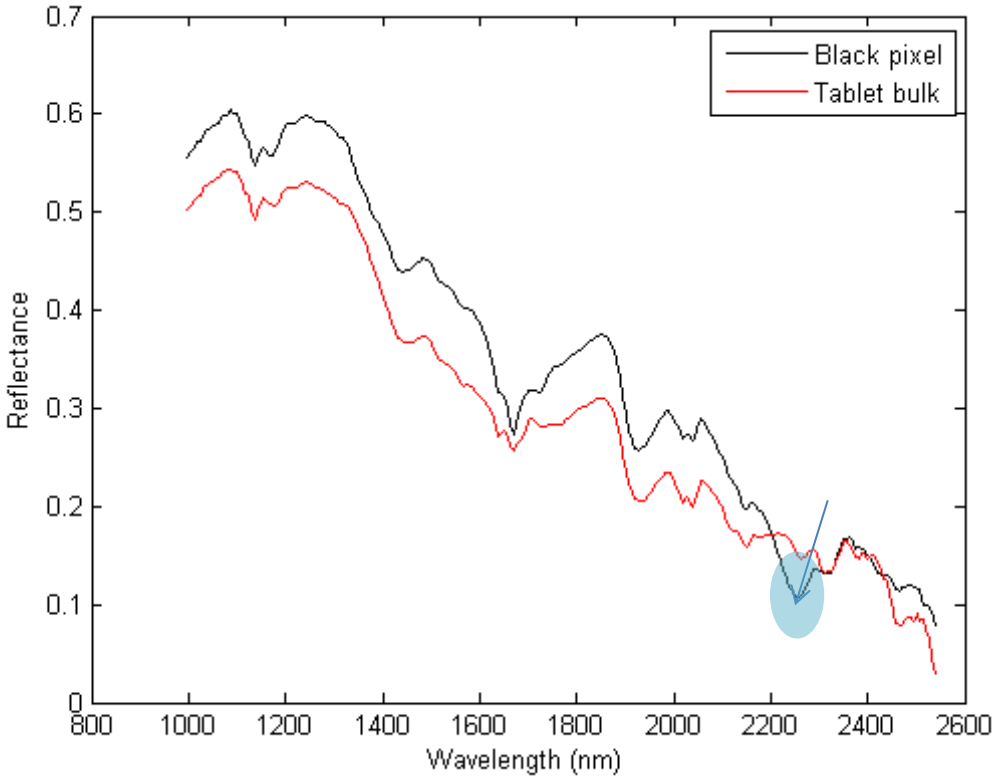


Figure 6-24 The spectra of the bulk and black pixels, the ellipse marks the region of interest that the PCA is using to separate out like spectra in Figure 6-23.

The loadings for the fourth principle components are shown below in Figure 6-25, the more intense the peak the greater the affect that variable has on the variation depicted in the scores images. The loadings clearly show corroborating evidence that the region of the tablet spectrum which is causing the separation in the PCA, and thus creating the black pixels, is the peak at 2257 nm.

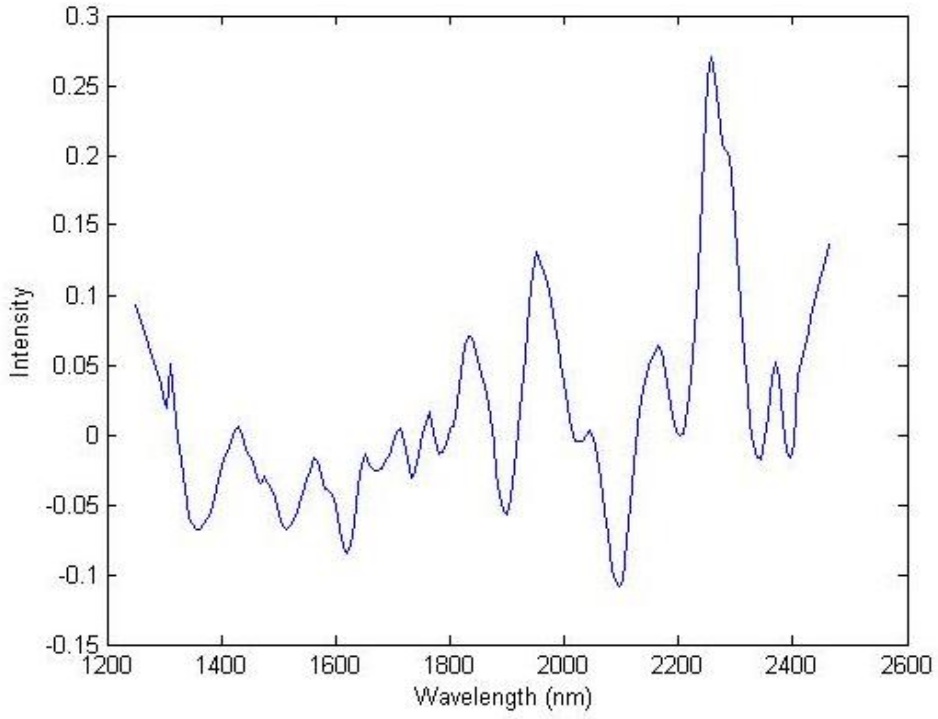


Figure 6-25 The loadings from the 4th principle component analysis of Tablet C

This change in the spectra of certain pixels implies that there is something different within these areas of the tablet than the bulk; the most likely cause is that an active pharmaceutical product (API) or excipient is more concentrated in these areas than the bulk. This would imply a potential mixing issue within the tablet and a non-homogenous distribution.

The two APIs in Tablet C are acetaminophen and caffeine and the most common excipient is starch. The excipients were also imaged using the same experimental procedure to gather a set of reference spectra. The spectra for these three compounds were plotted to see if the peak in Figure 6-24 could be identified; the mean black pixel spectrum was also plotted for comparison and the results are shown in Figure 6-26.

There is only one spectrum which contains the peak at 2257 nm, which is the caffeine. The peak and adjoining shoulder in the black pixel spectrum are very similar to caffeine however the rest of the spectrum is closer to the acetaminophen signal. This implies that these black pixels contain a high concentration of caffeine but are still mainly acetaminophen.

The pixel size in these images is 219 x 219 μm^2 which is large enough to contain clusters of caffeine particles which have not been mixed thoroughly. However before any

conclusion regarding the nature of the black pixels can be postulated more results are required.

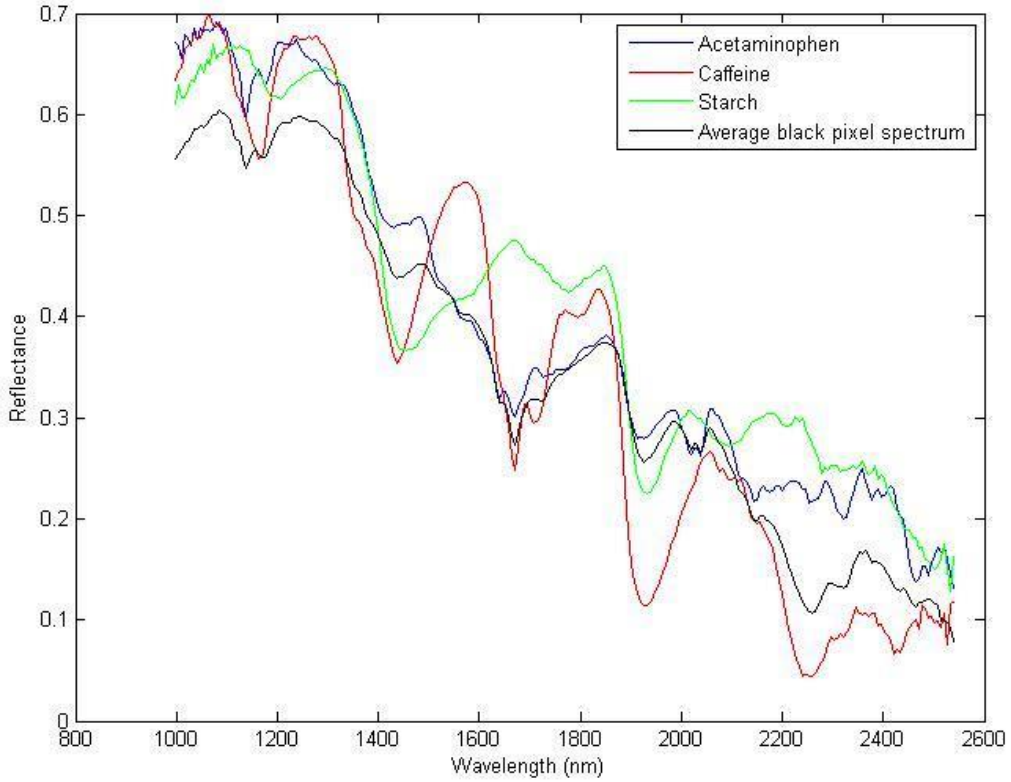


Figure 6-26 Spectra of Tablet C's major components alongside a reference spectrum from the black pixels

6.3.6 Tracking the change in intensity of the black pixels over time

6.3.6.1 Investigating the change in the number of black pixels within the tablet over time in an attempt to infer changes within the tablet

The basic spectral analysis of the tablet data performed previously in this chapter and other has shown that while the expansion can be tracked with a high degree of time resolution it is not possible to follow the spectral data to generate an accurate dissolution profile. Two separate factors cause issue with the spectral data firstly the change in the tablet spectra due to the attenuation of the signal and secondly the requirement for the tablet to expand before a stable 'maximum' spectrum can be achieved. This effect is seen clearly in Figure 6-7,8 and 9 where the initial dry spectrum is dissimilar to the other data and there is a period of fluctuations in the spectral results before the spectra settle. To determine a release from the data it would be necessary to have a 100% drug reading from a sample

submerged in the acid solution. However the expansion of the tablet changes the degree of attenuation and therefore the spectral response. A tablet containing 100% of the drug while also swollen fully so that the attenuation is not changing the spectrum would be required to give a true starting value of 100% drug in the tablet. The UV results in the appendix show that the release of the acetaminophen and caffeine from these tablets is rapid and that as soon as the tablet has begun to expand it has started to release. These factors result in a portion of the data which has misleading intensity values and no true starting intensities for the attenuated data.

However, the black pixels in the image can be seen from the very first image and followed throughout the experiment which potentially gives a novel approach to following the release from the tablet. These black pixels eventually begin to diminish in number however, and the rate at which these pixels leave the cell should be indicative of the release rate of caffeine from the tablet. Therefore by investigating the release of these black pixels if could provide a possible approach to calculate some form of rate of release for the tablets.

The intensity of the black cells is lower than any other values within the flow cell and it is this property can be used to count how many black pixels are present in any flow cell repetition, The number of black pixels in any individual flow cell was calculated using code written in MATLAB, which is available in the appendix, this code used a thresh holding approach to find every pixel with a value below 0.18. By finding every pixel below a specific value it was possible to find only the black pixels. There is a small issue with this approach however; it requires manual interpretation of the data to find a threshold value. This can be changed however, as an automatic thresh holding system could be implements for further analysis which would remove the need to manually choose a value. It was decided that for a single tablet experiment, this was not necessary and so a manual approach was used.

The count of pixels was calculated and then plotted against time to show the change in numbers as the dissolution progressed; the plot is shown in Figure 6-27. The initial two flow cell repetition in the HSI dissolution image are very dissimilar to the bulk of the results, as can be seen in Figure 6-23, which is the cause of the sharp peak at the very beginning of the plot. The general trend of the data is a loss of black pixels over time with the loss of black pixels being fastest in the first 1 – 2 minutes. The gradient of loss becomes shallower after the 2 minute mark until after 6 minutes no more black pixels can

be seen. This pattern can be thought of as the inverse of a standard dissolution UV spectrum in which there is a rapid release followed by a more gradual gradient until it plateaus at 100 % release.

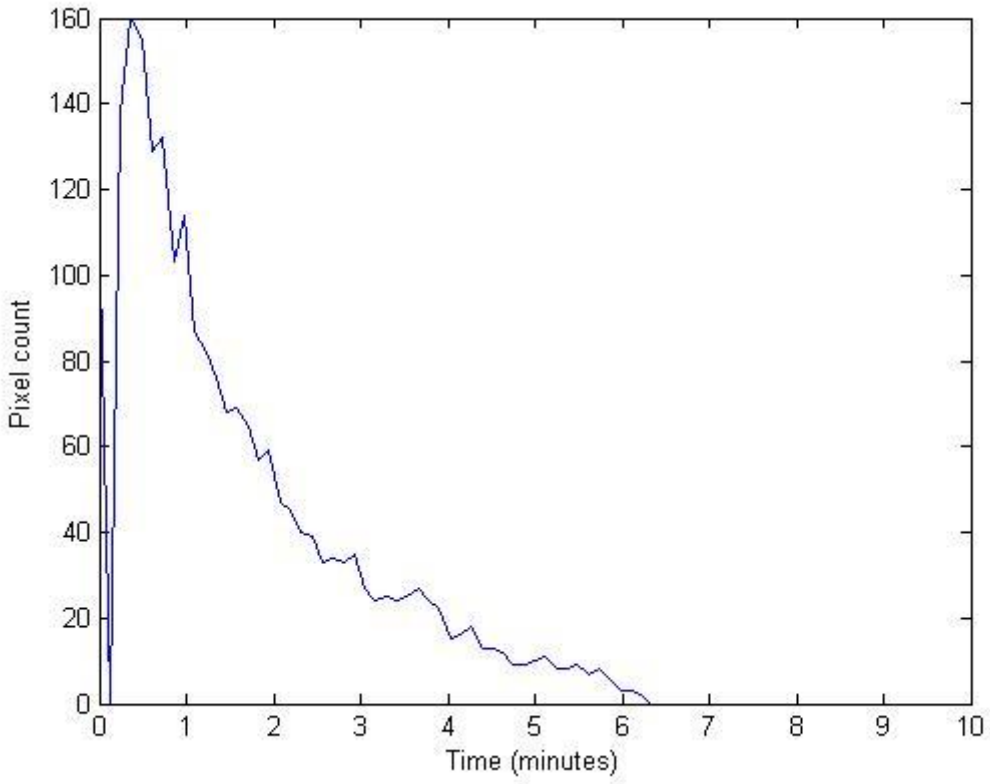


Figure 6-27 A plot showing the count of black pixels present in the scores plot of Tablet C over 10 minutes

6.3.6.2 The rate of caffeine release from the tablets

These results have shown that a change within the image can be followed as a function of time. It is important to investigate if this change is the release of the caffeine from the tablet, a combination of release from both of the APIs or some other physical effect.

To compare the change in black pixel counts with the release from the tablet systems it is necessary to perform a standard UV dissolution with the same setup as the one used in the experiments. While the dissolution testing of Tablet C was performed previously, and is shown in the appendix, this was on a different system and the values cannot be used for comparison. Tablet C contains two UV active compounds, acetaminophen and caffeine, which makes calculating the release of one of these more difficult. The concentrations were calculated by using the intensities at two separate wavelengths and then using

simultaneous equations to solve for the caffeine concentration. The simultaneous equations are based on the Beer Lambert law and are shown in the appendix.

For the UV dissolution the imaging experiment was repeated using the same experimental setup as the hyperspectral imaging but without the camera, a second line of tubing was also run out of the dissolution reservoir and into a UV spectrometer. The UV spectrometer measured the intensity at 4 different wavelengths every 10 seconds for 30 minutes. These spectral intensities were then used to calculate the release of caffeine from the tablets by using a simultaneous equation to separate out the two UV active APIs.

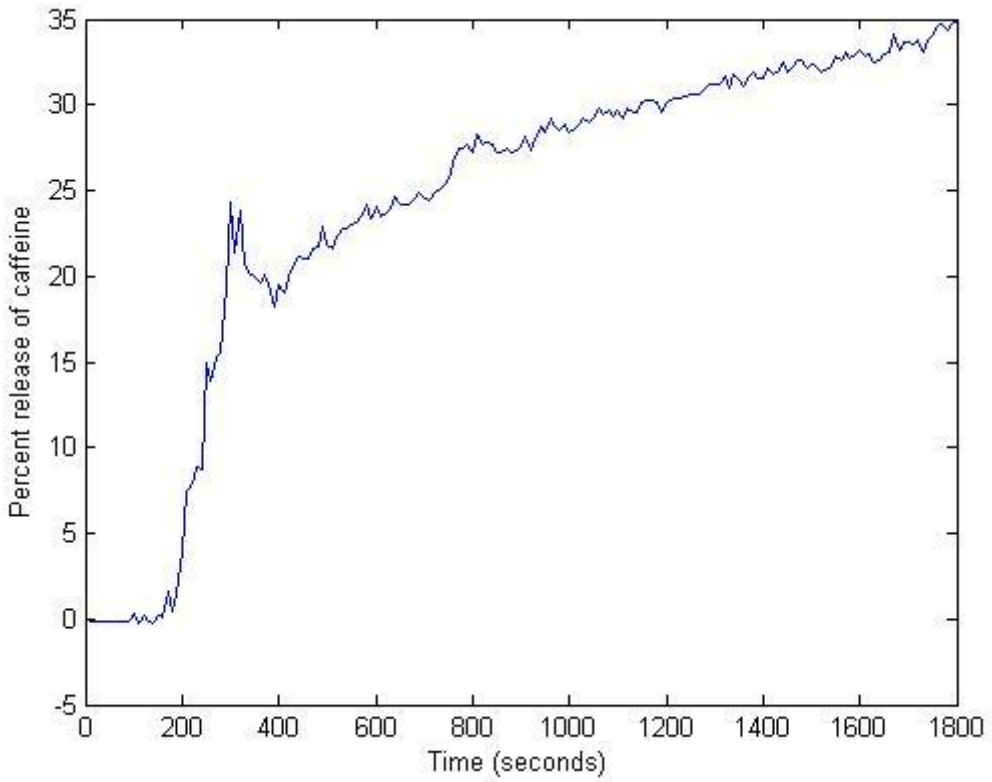


Figure 6-28 The percentage release of caffeine from Tablet C, measured by UV dissolution

The release of the caffeine from Tablet C as a function of the percentage of total release against time is shown in Figure 6-28. If the decrease in the number of black pixels is analogous to the release of caffeine from the tablet then the two plots should show similar results. The UV release data shows that after a short lag period the majority of the release occurs in a 100 – 200 second window after which the release slows. The black pixels decrease over time with the majority occurring within the first 2 – 3 minutes, this loss is analogous to the release of caffeine. However the results in Figure 6-27 show that there is complete loss of the black pixels within 7 minutes while the dissolution in Figure 6-28

only reaches a maximum of 37% after 30 minutes. The reason for this discrepancy in the total percent of release is that the UV dissolution is a very sensitive technique which can detects very small volumes of dissolves drug while the HSI is showing pixels which contain a higher than average caffeine concentration. The spatial resolution of the HSI image gives pixels which are ~219 x 219 μm^2 in length which means that each black pixel covers a large area within the tablet and contains a huge number of caffeine particles. When the tablet expands the caffeine in these pixels is released and the high concentration of caffeine begins to dissolve and dissipate through the flow cell. Once the total amount of caffeine in a pixel is below a threshold it will no longer be detectable by the camera thus giving a 100% release reading when the UV data shows only 37%.

This UV dissolution should be smooth with no decrease in intensity at any point as it shows the cumulative release. There was no mixing in the reservoir which stored the hydrochloric acid and this resulted in poor distribution of the APIs through the solution. With this poor mixing it is possible to get inaccuracy in the UV spectroscopy results which are what causes the noise in the data. The general shape of the data however, follows the traditional dissolution format with a fast release followed by a much slower release. There is a major issue that these results are showing which is that within 30 minutes there is only 35% release of the caffeine from the tablet. Multiple dissolution experiments were tried and multiple standards were used to double check the accuracy of these results.

The flow rate used in the UV dissolution tests was the same as the flow rate used in the HSI experiments to ensure accuracy and comparability between the experiments. The flow through the cell is limited during the experiment and the flow rate, during the experiment, of 10ml/min may be too low to properly allow for the tablet to dissolve and for particles to be swept away from the flow cell. The UV dissolution was repeated using a flow rate of 20ml/min in an attempt to improve the release characteristics of the experiment and to prove that the poor release was due to the flow rate. The release was calculated for the increased flow rate and the results can be seen in Figure 6-29. The original dissolution data is shown in Figure 6-29A for comparison and the 20mn/min UV dissolution data shown in Figure 6-29B. With the higher flow rate a maximum release much closer to 100% can be viewed, which shows that the quicker flow rate is allowing for the tablet to better dissolve and flow through the system. There is very little difference in these two release profiles other than the total degree of release and the speed of the initial release. In Figure 6-29A and Figure 6-29B the end of the steep fast release period is at around 200 seconds which

also shows that the flow rate has not affected the physical breakdown of the tablet only the quantity of caffeine which has been able to dissolve into the bulk of the solution.

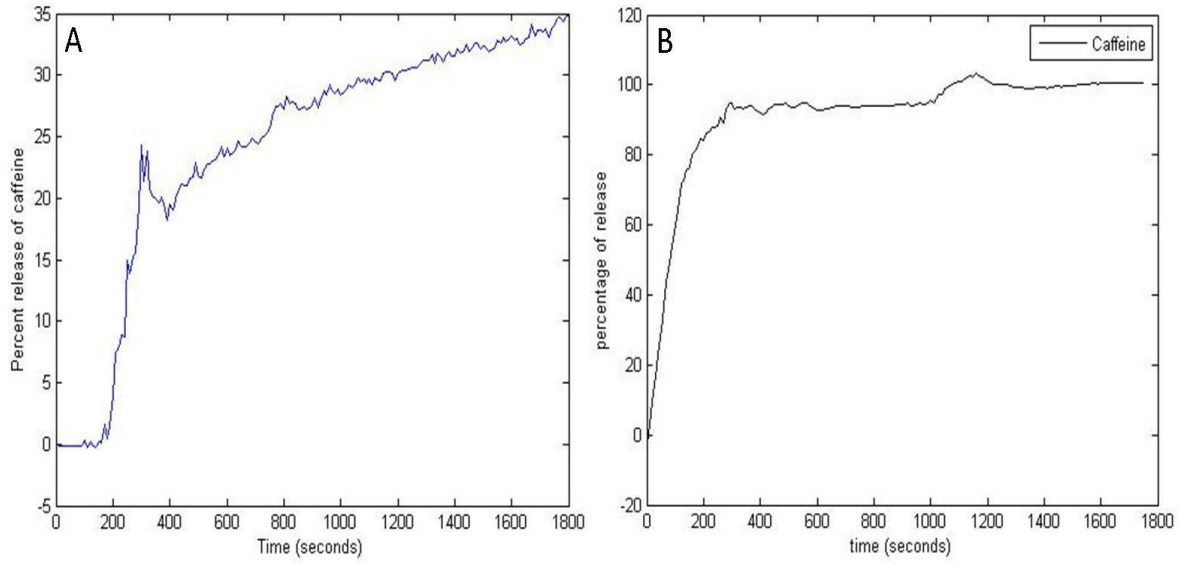


Figure 6-29 The release of caffeine from Tablet C as a percentage of total release with a flow rate through the system of A) 10ml/min and B) 20ml/min

The black pixels are detailing regions which contain a higher than average mass of caffeine but they are not pure caffeine. When the tablet expands and begins to dissolve into the acid solution the quantity of caffeine in these pixels will begin to decrease. As the concentration decreases there will be a point at which there is no net change to the spectral signal and the PCA will no longer be able to distinguish these regions as being different to the bulk. At this point the pixel will change from black to the same grey as the bulk thus lowering the total number of black pixels. These results show that the reduction in the number of the black pixels can be correlated to the release of caffeine from the tablets either as a function of dissolution or expansion so that the particles spread through the flow cell. While the rate of change cannot be exactly matched to any dissolution data it does show the ability of the system to follow a physical change within the tablet at specific spatial locations, which can be identified spectrally, over time.

6.3.7 Raman mapping of the tablet surface

Raman mapping is a process which generates a hyperspectral image containing Raman spectra. It is used for frequently to determine what chemicals can be identified on the surface of the tablet. Using Raman spectroscopy it is possible to scan a region of the tablet

surface and to use the map generated to look for regions containing high caffeine concentrations.

6.3.7.1 Raman spectroscopy of tablet APIs

Raman spectroscopy produces maps with a much higher spatial and spectral resolution than HSI.

The Raman system was a Thermo DXR Raman spectrometer which used a 532 nm laser for excitation with a power of 10 mW through a 25μm pinhole aperture. The sample was exposed for 20 seconds and an average of 4 measurements was used to create each recorded Raman spectrum. The spectra were corrected for fluorescence using a second order polynomial correction algorithm which is built into the Raman OMNIC software. The mapping was performed on a 1500 μm by 2000 μm region of the tablet with spectra captured every 100 μm. Due to the curvature of the tablet it was necessary to capture each line separately and to re-focus the microscope for each new line. To use the Raman maps effectively specific wavelengths which contain peaks unique to caffeine or acetaminophen are needed. The Raman spectrum for each API was generated just prior to the mapping to reduce any environmental or experimental deviation; the plot of the spectra is shown in Figure 6-30.

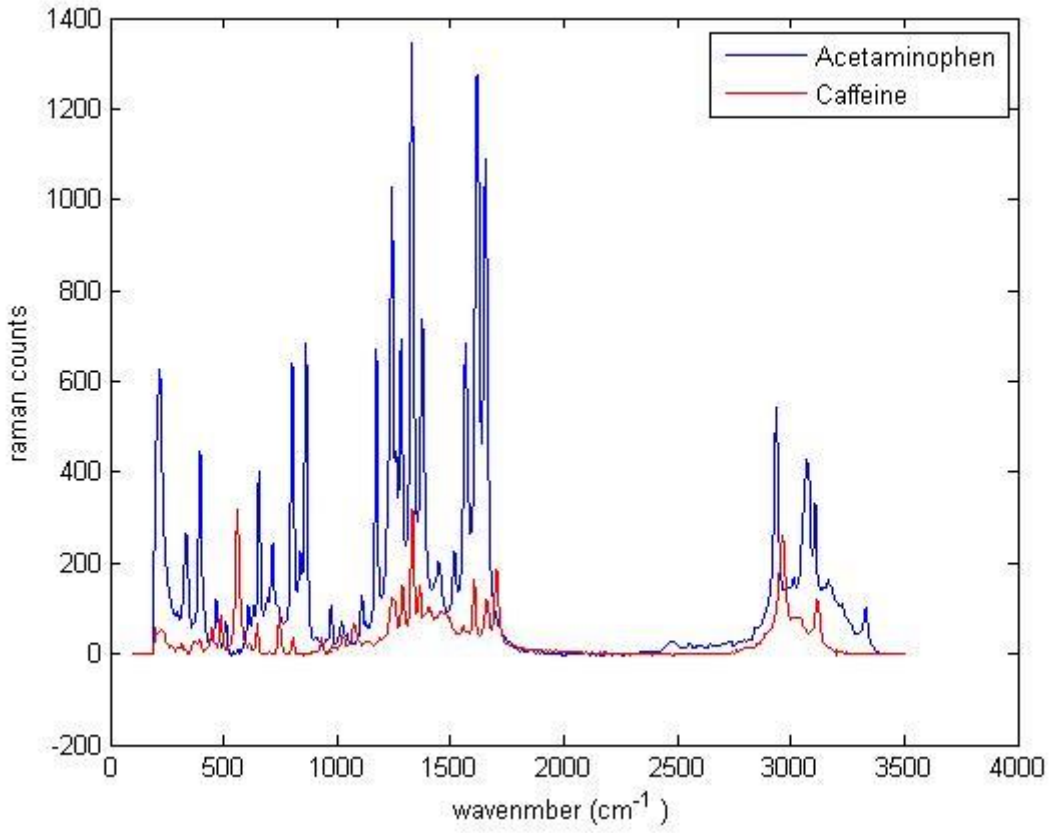


Figure 6-30 Raman spectrum of acetaminophen and caffeine

6.3.7.2 Raman map of Tablet C: determining the distribution of caffeine throughout the sample.

At 558 cm^{-1} there is a peak in the caffeine spectrum which is strong and unique and there is no such peak in the acetaminophen. By looking at the intensity at this wavelength for the captured data it is possible to create a map which shows where the concentration of caffeine is highest.

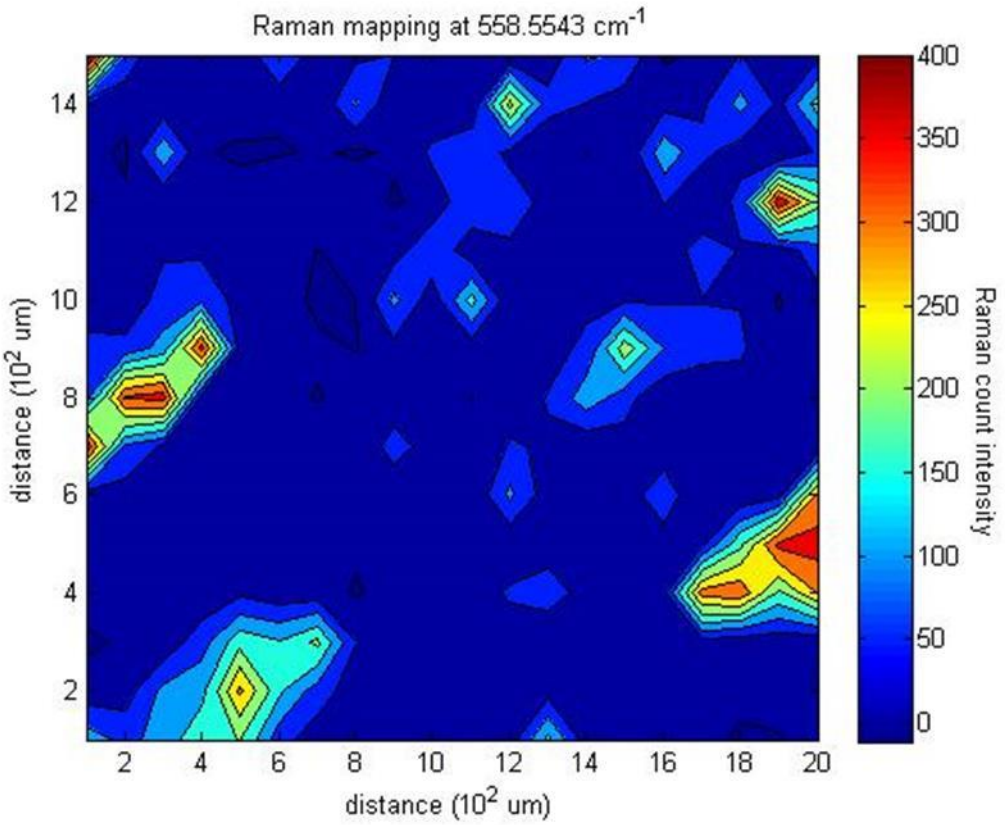


Figure 6-31 Raman map of Tablet C, showing the intensity at 558 cm^{-1}

The Raman data in Figure 6-31 is shown as a contour map with the red areas showing the most intense peaks. There are a number of areas in map which correlate to high caffeine concentrations, these regions are spaced throughout the image in a none-homogenous manner. The sporadic placement of the caffeine clusters throughout Figure 6-31 matches the distribution of the black pixels in Figure 6-22; the caffeine should be evenly distributed throughout the tablet.

The size of the caffeine regions vary from a single pixel like in the top right with a dimension of approximately 100 x 100 μm^2 to much larger particles like in the bottom right with dimensions of approximately 300 x 200 μm^2. This range of caffeine region sizes also agrees with the image data in Figure 6-22, were the black pixel regions range from a single pixel to groupings of 8 – 10 pixels. The larger pixels sizes from Figure 6-22 can be attributed to the long path length and scattering of the nIR photons which can cause the caffeine signal to from one pixel to be detected within the surrounding pixels.

6.3.7.3 Raman map of Tablet C to show the distribution of acetaminophen throughout the tablet

It is possible that the regions in Figure 6-31 are not specific to caffeine but are instead just areas of with higher signal intensity. To correct for this possibility a second Raman map was created using the same data and the results are shown in Figure 6-32. The first map was tailored to find areas of caffeine by using a unique peak; the second map uses a peak unique to acetaminophen, specifically the one at 886 cm^{-1}. This peak results in a map which will show either an even distribution of acetaminophen throughout or the intense peaks in Figure 6-31 will be either fully or partially replicated.

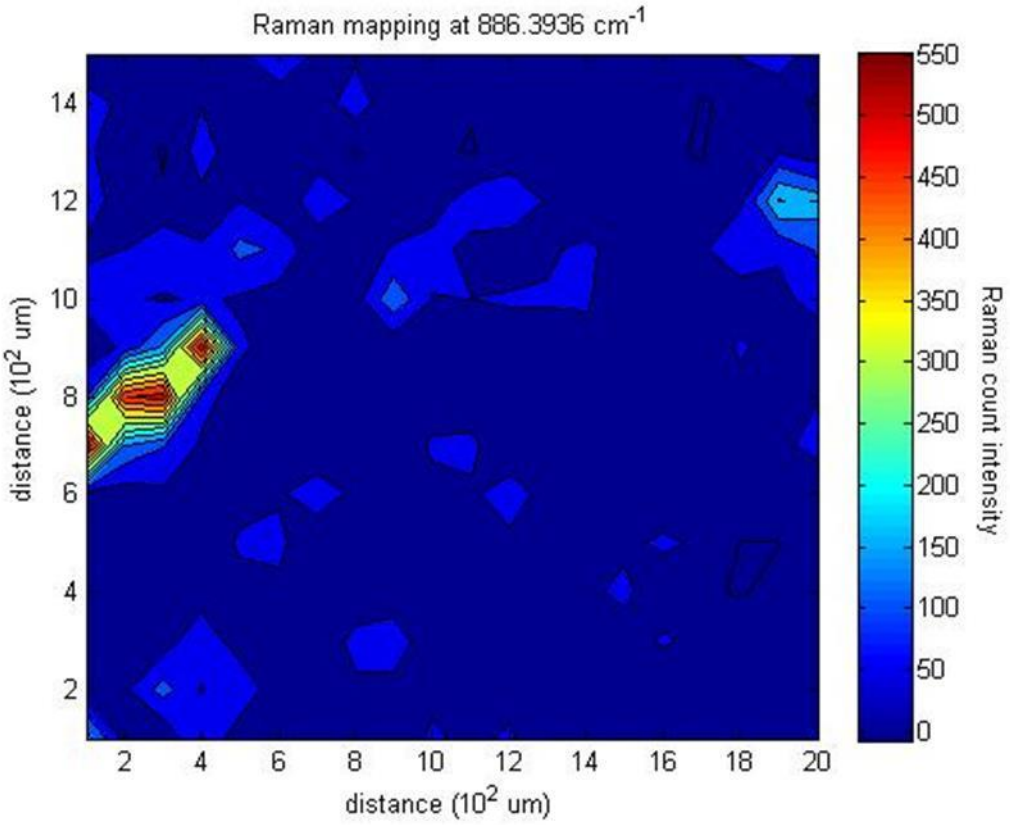


Figure 6-32 Raman map of Tablet C, showing the intensity at 886 cm^{-1}

The Raman map of the acetaminophen data shows an intense region on the left of the image which corresponds to Figure 6-31. However, all of the other intense regions in Figure 6-31 have no corresponding intense region in Figure 6-32. These results show that the caffeine regions are not simply areas of higher intensity but are actual areas containing high caffeine concentration. Thus by two independent techniques it has been shown that these regions containing black pixels in the scores image are consistent with high caffeine concentrations.

6.3.7.4 Degree of caffeine coverage detected using HSI and Raman mapping.

The British Pharmacopeia states that any caffeine containing paracetamol tablet must have a mass of caffeine within the tablets to be 60mg ± 5%. The average weight of Tablet C is 687 mg and using this information any single Tablet C should contain 8-9.5% caffeine in any one region if the mixing is even.

To calculate the ration of the black pixels to the bulk a threshold algorithm was used on the scores data which first showed the black regions. The algorithm simply finds every pixel with a value within certain bounds; this value was determined by simple observation of the data. The algorithm was used on the first 'dry' tablet in the image as this has not undergone any expansion or potential loss of any API or excipient components. It was calculated that 14.4 – 22.5% of the tablet surface contains black pixels a value higher than the British Pharmacopeia standard values. This is not unexpected, the nIR photons have a large path length as described early which could easily cause a number of false positive results. The total size of the tablet is also difficult to determine as some is hidden in the holder on the far right of the image. These two factors could be causing a value much higher than expected.

The same treatment was given to the Raman mapping data shown in Figure 6-31, this time the threshold algorithm was used to find how many pixels contains a value above 150 counts. This value is high enough that any background will be removed while low enough that any region which shows a higher than average caffeine concentration will be counted. This threshold finds exactly 27 pixels out of the total 300, giving a percentage coverage of 9% which also corresponds to the previous results.

6.4 Conclusions

In this work hyperspectral imaging has been used to create an image which contains spectral information bound within 2 spatial dimensions along with a time axis with good resolution (100 ms). The way this image was generated with the same region imaged over time is a novel use of the HSI systems and has resulted in some very powerful images. The dissolution of a tablet can be seen as a real time image which can be made into a pseudo video with spectral information at every pixel.

Using HSI and PCA caffeine has been shown to exist in Tablet C in a non-uniform manner and this determination was also possible with the strong signal attenuation caused by the acid solution in the flow cell. The caffeine was first identified as a none uniform distribution of black pixels in the scores images from the fourth principle component generated by principle component analysis of the hyperspectral imagine of Tablet C's dissolution. Spectral analysis of these black regions showed that there was a key difference in the spectrum and comparison with the loading data confirmed that this region was the cause of the separation in the image. Further investigation was performed using UV and Raman techniques to verify that these black pixels were caffeine containing areas of the tablet. For the first time hyperspectral imaging has been used to successfully isolate, identify and track a known compound during the dissolution of a pharmaceutical grade tablet.

To image the dissolution of a tablet within an aqueous fluid system was a problematic proposition due to the effect the water in the flow cell has on any nIR signals. By positioning the camera underneath the flow cell the issues caused by the attenuation of the signal by the water present in the flow cell, in the form of the acid solution, have been reduced. A significant increase in the signal intensity was visible as well as some small regions of identifiable peaks which were seen in the original data. This is a massive improvement on the earlier data and resulted in the identification of caffeine in the tablet being possible.

Simple analysis of the unprocessed spectral data was insufficient to determine any changes within the image caused by release of API but could show changes that can be attributed to the physical processing occurring within the tablet during dissolution. The expansion and hydration of the tablet can be easily tracked and followed with this techniques but the analysis shows little novel or unexpected.

By processing the data and generating the first derivative of the image combined with principle component analysis it was possible to detect regions within a tablet which could be attributed to a specific tablet component. This was achieved without the need for a zoom lens or other magnification which would be necessary in Raman spectroscopy, the other technique which could be potentially used for this purpose.

Using a standard dissolution experiment it is only possible to gather the spectral and temporal information from the release of the drugs. Using Raman or traditional nIR

spectroscopy allows for the spatial and spectral information to be mapped but loses the temporal information due to the long times required for data acquisition. However, with hyperspectral imaging it has been shown that you can collect a huge amount of spatial and spectral information while also recording this data in the temporal domain. This is a massive advantage of this technique over the major systems for tablet analysis which are currently used.

There is also a compromise with HSI as the spectral and spatial resolution is often lower than the normal spectrographic techniques. Without the use of an expensive zoom lens the pixel size of 219 x 219 μm^2 is one of the best that can be achieved, Raman can easily be used on much smaller scales than this. The spectral resolution with this HSI equipment was ~ 6.3 nm which is much lower than a standard nIR, which can cause issue with smaller finer peaks which may be lost within the data when the resolution is lowered. Again this is a compromise of sorts; while the resolution may be poorer in some areas it is substantially greater in others. A normal nIR spectrometer would struggle to collect images at the speed utilised in this work and most Raman would not be able to collect high quality spectra within 100ms.

However, despite the complications brought on by the use of nIR spectroscopy in an aqueous environment it has been possible to show the physical location of an API inside a tablet during dissolution. This API was characterised using PCA and the existence of these regions can be tracked throughout the entire dissolution until the concentration of them becomes too low and they disappear from the image. This is, to the authors knowledge, the first time that the dissolution of a sample has been followed with the whole tablet being imaged and with APIs being tracked within the tablet in real time at the sub millimetre level.

Summary and future work

7.1 Summary

In this work hyperspectral imaging has been used as a novel approach in tracking the processes occurring within a tablet during dissolution and to follow the loss of a known compound from the tablet. The earliest stages of the work centred on analysis of different HSI systems to determine which was most effective for pharmaceutical analysis. The results showed that the nIR was a substantially better system for investigating white solids or powders than a system working at visible wavelengths. It also showed that principal component analysis is necessary when analysing these images due to the large number of variables and small differences between the spectra of the samples. Once the best system and methods of analysis had been determined then the investigation of the dissolution process was performed. The established USP dissolution apparatus was not suitable for imaging due to the curved dissolution vessels and large volumes of liquid which would affect the quality of the imaging. A flow system was designed which would allow imaging to be performed by placing the tablet inside a modified 1cm quartz flow cell. Using this flow system the dissolution was imaged and while APIs could be isolated the physical changes in the tablet were tracked.

The time resolution of the first experimental setup was poor with respect to the fast release from these tablets which had been shown in separate UV dissolution tests. A system was devised that would sacrifice spatial information for enhanced time resolution to improve the information gathered on the release from the tablet. This system showed a definitive affect from the acid solution, which caused attenuation of the signal and affected the quality of the results. The degree of signal attenuation was shown to be proportional to the path length travelled through the absorbing medium using mathematical models.

A protocol was devised which combined the spatial and time resolutions from earlier experiments to create a hyperspectral image which accurately portrayed the dissolution of a paracetamol tablet. Principal component analysis was performed on the images and investigation of these results showed the locations of caffeine rich regions in a tablet. These regions were verified as being caffeine containing by analysis of Raman mapping of the tablet surface. The rate of loss of these regions was also investigated and compared to that of the UV dissolution data which gave comparable results.

These results showed the potential of using a nIR HSI system to follow a spectral signal which would normally be too attenuated by water in the flow cell to give any meaningful data. Overcoming this problem, by inverting the camera and placing it under the experiment, allows for analysis of a number of spectral regions which could potentially be used to follow release. In the data presented above there was a set of peaks in the 1600 – 1700 nm region which were never used for analysis but which have the potential to be a second point of interest for further analysis. These peaks and the ones at 2257 nm also show that this approach to nIR spectroscopy can result in useable signal peaks through such a strong attenuating medium as water.

All of these areas show that hyperspectral imaging has the potential to be a very powerful tool for investigating the breakdown and release of tablets. There are however, some problems with HSI that must be considered. The technique while fast generates a huge quantity of data which can take a large amount of time and effort to successfully mine for useful results. It is also very possible that potentially useful results could be missed if the correct analysis is not performed a problem which is much rarer in the more standard spectroscopy techniques.

However, despite the complications and difficulties inherent to this work it has been possible to show the physical location of an API inside a tablet during dissolution. This API was characterised by a specific wavelength region which allows the whole dissolution to be tracked. This is the first time that the dissolution of a sample has been followed with the whole tablet being imaged and with APIs being tracked at a micron level. While it has long been possible to follow the release of a drug during dissolution, following the whole tablet changing spectrally is a novel result. It has taken a number of years to progress the work to this point due to the inherent difficulties present when using nIR in an aqueous system. The results do show the potential of this technique to find a place in improving both the manufacture and pharmokinetics of the tablets. These data show how the API is distributed throughout the tablet over time and how these clusters of API move, dissolve or stay resistant to the dissolution process over time.

7.2 Future work

The results in this thesis give a number of potential avenues for future works to further investigate the ability of this HSI system as a novel technique for the analysis of dynamic systems.

Greater analysis of the data presented here could be performed to enhance the knowledge of how the caffeine is being lost from the tablet. Analysis of the intensities over time of the different caffeine regions could give a measure of the homogeneity of the release from the tablet. This could lead to a better understanding of the processes occurring in the dissolution and if the uneven distribution of the caffeine in the tablets also affects the release rate of the caffeine. In this work a single caffeine containing tablet formulation was imaged during the last experiment. By investigating the other caffeine containing paracetamol formulations the results could be verified and the protocol shown to be a good identifier of known compounds in a tablet system. The expansion of the imaging methodology to include new tablet variants which included known excipients or APIs in sufficiently large concentrations is a logical progression to this work. This would reveal whether the system is an effective method for tracking these compounds through the dissolution in a spatial resolved manner.

The use of magnification to improve the spatial resolution of the hyperspectral imaging is also a possible area of interest for improving the quality of the results. By improving the spatial resolution, it should be possible to get more separation of the components within the tablets as current pixel sizes are in the order of hundreds of micrometres. An increased spatial resolution without using magnification can also be achieved using a higher pixel density in the detector. This is especially useful for investigating the location of excipients which are a small quantity of the tablet, in a large pixel the majority of the signal is API but a small pixel may contain a higher proportion of excipient.

The biggest problem in the analysis of these systems is the presence of water, especially in as an acid solution, which caused severe signal attenuation. Investigating the use of other non-aqueous solvents could also be a possible area of interest. This would require the identification of a solvent that the tablets can dissolve into and that don’t cause corrosion of the equipment. By using this system it should be possible to follow the breakdown and release of API and excipients with much greater spectral definition than is possible when using aqueous solutions. The major component in the tablets used in these experiments is

paracetamol which is soluble in a number of polar organic solvents such as acetone, ethanol, ethylene glycol, butanone, tetrahydrofuran and formamide. These solvents contain either a hydroxyl or carbonyl group which helps to facilitate the uptake of the paracetamol into solution making them suitable replacement dissolution solvents. It would be necessary to test these solvents individually for their signal attenuation and efficiency at dissolving the paracetamol tablets as the solubility of excipients will be variable for the different solvent systems. However, it is seems probable that a solvent could be chosen to replace the aqueous solution and thus remove or lower the signal attenuation caused by the absorption of near IR photons by water.

All of these possible areas of interest centre on the use of the HSI system to image more tablets systems but it is also possible to use the time resolved nature of the system to follow other dynamic process. The investigation of chemical reactions or the tracking of particle size in a solution that is aggregating should be possible and is a novel, none pharmaceutical use of the imaging setup.

References

1. Buttingsrud B, Alsberg BK 2006. Supperresolution of hyperspectral images. Chemometrics and Intelligent Laboratory Systems 84(1-2):62-68.
2. Erives H, Targhetta NB 2009. Implementation of a 3-D Hyperspectral Instrument for Skin Imaging Applications. Ieee Transactions on Instrumentation and Measurement 58(3):631-638.
3. Wong G snapshot hyperspectral imaging and practical applications. Journal of physics: conference series 178(012048):1-5.
4. Hirohara Y, Okawa Y, Mihashi T, Yamaguchi T, Nakazawa N, Tsuruga Y, Aoki H, Maeda N, Uchida I, Fujikado T 2007. Validity of retinal oxygen saturation analysis: Hyperspectral imaging in visible wavelength with fundus camera and liquid crystal wavelength tunable filter. Optical Review 14(3):151-158.
5. Pan ZH, Healey G, Prasad M, Tromberg B 2003. Face recognition in hyperspectral images. Ieee Transactions on Pattern Analysis and Machine Intelligence 25(12):1552-1560.
6. Martin ME, Wabuyele M, Panjehpour M, Overholt B, DeNovo R, Kennel S, Cunningham G, Vo-Dinh T 2006. An AOTF-based dual-modality hyperspectral imaging system (DMHSI) capable of simultaneous fluorescence and reflectance imaging. Medical Engineering & Physics 28(2):149-155.
7. Gupta N, Voloshinov V 2005. Hyperspectral imaging performance of a TeO2 acousto-optic tunable filter in the ultraviolet region. Optics Letters 30(9):985-987.
8. Qin JW, Lu RF 2008. Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. Postharvest Biology and Technology 49(3):355-365.
9. Sutherland VL, Timlin JA, Nieman LT, Guzowski JF, Chawla MK, Worley PF, Roysam B, McNaughton BL, Sinclair MB, Barnes CA 2007. Advanced imaging of multiple mRNAs in brain tissue using a custom hyperspectral imager and multivariate curve resolution. Journal of Neuroscience Methods 160(1):144-148.
10. Gordon KC, McGoverin CM 2011. Raman mapping of pharmaceuticals. International Journal of Pharmaceutics 417(1-2):151-162.
11. Sasic S 2007. Raman mapping of low-content API pharmaceutical formulations. I. Mapping of alprazolam in Alprazolam/Xanax tablets. Pharmaceutical Research 24(1):58-65.
12. Sasic S 2007. An in-depth analysis of Raman and near-infrared chemical images of common pharmaceutical tablets. Applied Spectroscopy 61(3):239-250.
13. Sasic S 2011. Parallel imaging of active pharmaceutical ingredients in some tablets and blends on Raman and near-infrared mapping and imaging platforms. Analytical Methods 3(4):806-813.
14. Cruz J, Blanco M 2011. Content uniformity studies in tablets by NIR-CI. Journal of Pharmaceutical and Biomedical Analysis 56(2):408-412.
15. Stuffler T, Forster K, Hofer S, Leipold M, Sang B, Kaufmann H, Penne B, Mueller A, Chlebek C 2009. Hyperspectral imaging-An advanced instrument concept for the EnMAP mission (Environmental Mapping and Analysis Programme). Acta Astronautica 65(7-8):1107-1112.
16. Kokaly RF, Asner GP, Ollinger SV, Martin ME, Wessman CA 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. Remote Sensing of Environment 113:S78-S91.
17. Schmidt KS, Skidmore AK 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment 85(1):92-108.

18. van der Meer FD, van der Werff HMA, van Ruitenbeek FJA, Hecker CA, Bakker WH, Noomen MF, van der Meijde M, Carranza EJM, de Smeth JB, Woldai T 2012. Multi- and hyperspectral geologic remote sensing: A review. International Journal of Applied Earth Observation and Geoinformation 14(1):112-128.
19. Sabins FF 1999. Remote sensing for mineral exploration. Ore Geology Reviews 14(3-4):157-183.
20. Hakvoort H, de Haan J, Jordans R, Vos R, Peters S, Rijkeboer M 2002. Towards airborne remote sensing of water quality in The Netherlands - validation and error analysis. Isprs Journal of Photogrammetry and Remote Sensing 57(3):171-183.
21. Herold M, Gardner ME, Roberts DA 2003. Spectral resolution requirements for mapping urban areas. Ieee Transactions on Geoscience and Remote Sensing 41(9):1907-1919.
22. Benediktsson JA, Palmason JA, Sveinsson JR 2005. Classification of hyperspectral data from urban areas based on extended morphological profiles. Ieee Transactions on Geoscience and Remote Sensing 43(3):480-491.
23. Maathuis BHP, van Genderen JL 2004. A review of satellite and airborne sensors for remote sensing based detection of minefields and landmines. International Journal of Remote Sensing 25(23):5201-5245.
24. Tiwari KC, Arora MK, Singh D 2011. An assessment of independent component analysis for detection of military targets from hyperspectral images. International Journal of Applied Earth Observation and Geoinformation 13(5):730-740.
25. Mehl PM, Chen YR, Kim MS, Chan DE 2004. Development of hyperspectral imaging technique for the detection of apple surface defects and contaminations. Journal of Food Engineering 61(1):67-81.
26. Gosselin R, Rodrigue D, Gonzalez-Nunez R, Duchesne C 2009. Potential of Hyperspectral Imaging for Quality Control of Polymer Blend Films. Industrial & Engineering Chemistry Research 48(6):3033-3042.
27. Qin JW, Burks TF, Ritenour MA, Bonn WG 2009. Detection of citrus canker using hyperspectral reflectance imaging with spectral information divergence. Journal of Food Engineering 93(2):183-191.
28. Wang WL, Li CY, Tollner EW, Gitaitis RD, Rains GC 2012. Shortwave infrared hyperspectral imaging for detecting sour skin (Burkholderia cepacia)-infected onions. Journal of Food Engineering 109(1):38-48.
29. Liu YL, Chen YR, Wang CY, Chan DE, Kim MS 2005. Development of a simple algorithm for the detection of chilling injury in cucumbers from visible/near-infrared hyperspectral imaging. Applied Spectroscopy 59(1):78-85.
30. Gowen AA, Taghizadeh M, O'Donnell CP 2009. Identification of mushrooms subjected to freeze damage using hyperspectral imaging. Journal of Food Engineering 93(1):7-12.
31. Liu YL, Chen YR, Kim MS, Chan DE, Lefcourt AM 2007. Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. Journal of Food Engineering 81(2):412-418.
32. Liu YL, Windham WR, Lawrence KC, Park B 2003. Simple algorithms for the classification of visible/near-infrared and hyperspectral imaging spectra of chicken skins, feces, and fecal contaminated skins. Applied Spectroscopy 57(12):1609-1612.
33. Naganathan GK, Grimes LM, Subbiah J, Calkins CR, Samal A, Meyer GE 2008. Visible/near-infrared hyperspectral imaging for beef tenderness prediction. Computers and Electronics in Agriculture 64(2):225-233.
34. Polerecky L, Bissett A, Al-Najjar M, Faerber P, Osmers H, Suci PA, Stoodley P, de Beer D 2009. Modular Spectral Imaging System for Discrimination of Pigments in Cells and Microbial Communities. Applied and Environmental Microbiology 75(3):758-771.

35. Volent Z, Johnsen G, Sigernes F 2009. Microscopic hyperspectral imaging used as a bio-optical taxonomic tool for micro- and macroalgae. Applied Optics 48(21):4170-4176.
36. Schultz RA, Nielsen T, Zavaleta JR, Ruch R, Wyatt R, Garner HR 2001. Hyperspectral imaging: A novel approach for microscopic analysis. Cytometry 43(4):239-247.
37. Chaudhari AJ, Darvas F, Bading JR, Moats RA, Conti PS, Smith DJ, Cherry SR, Leahy RM 2005. Hyperspectral and multispectral bioluminescence optical tomography for small animal imaging. Physics in Medicine and Biology 50(23):5421-5441.
38. Zuzak KJ, Naik SC, Alexandrakis G, Hawkins D, Behbehani K, Livingston EH 2007. Characterization of a near-infrared laparoscopic hyperspectral imaging system for minimally invasive surgery. Analytical Chemistry 79(12):4709-4715.
39. Kong SG, Martin ME, Vo-Dinh T 2006. Hyperspectral fluorescence imaging for mouse skin tumor detection. Etri Journal 28(6):770-776.
40. Zavattini G, Vecchi S, Mitchell G, Weisser U, Leahy RM, Pichler BJ, Smith DJ, Cherry SR 2006. A hyperspectral fluorescence system for 3D in vivo optical imaging. Physics in Medicine and Biology 51(8):2029-2043.
41. Zuzak KJ, Gladwin MT, Cannon RO, Levin IW 2003. Imaging hemoglobin oxygen saturation in sickle cell disease patients using noninvasive visible reflectance hyperspectral techniques: effects of nitric oxide. American Journal of Physiology-Heart and Circulatory Physiology 285(3):H1183-H1189.
42. Zuzak KJ, Schaeberle MD, Lewis EN, Levin IW 2002. Visible reflectance hyperspectral imaging: Characterization of a noninvasive, in vivo system for determining tissue perfusion. Analytical Chemistry 74(9):2021-2028.
43. Lopes MB, Wolff JC, Bioucas-Dias JM, Figueiredo MAT 2011. Study on the Effect of Pixel Resolution and Blending Grade on Near-Infrared Hyperspectral Unmixing of Tablets. Applied Spectroscopy 65(2):193-200.
44. Ravn C, Skibsted E, Bro R 2008. Near-infrared chemical imaging (NIR-CI) on pharmaceutical solid dosage forms-Comparing common calibration approaches. Journal of Pharmaceutical and Biomedical Analysis 48(3):554-561.
45. Cruz J, Bautista M, Amigo JM, Blanco M 2009. Nir-chemical imaging study of acetylsalicylic acid in commercial tablets. Talanta 80(2):473-478.
46. Lopes MB, Wolff JC 2009. Investigation into classification/sourcing of suspect counterfeit Heptodin (TM) tablets by near infrared chemical imaging. Analytica Chimica Acta 633(1):149-155.
47. Lopes MB, Wolff JC, Bioucas-Dias JM, Figueiredo MAT 2010. Near-Infrared Hyperspectral Unmixing Based on a Minimum Volume Criterion for Fast and Accurate Chemometric Characterization of Counterfeit Tablets. Analytical Chemistry 82(4):1462-1469.
48. United States Pharmacopeia and National Formulary USP 29-NF 24. The United States Pharmacopeial Convention I, Rockville, MD, 2007., ed.
49. Carstensen JT, Lai TYF, Prasad VK 1978. USP DISSOLUTION .4. COMPARISON OF METHODS. Journal of Pharmaceutical Sciences 67(9):1303-1307.
50. de la Pena MHY, Alvarado YV, Dominguez-Ramirez AM, Arroyo ARC 2003. Comparison of dissolution profiles for albendazole tablets using USP Apparatus 2 and 4. Drug Development and Industrial Pharmacy 29(7):777-784.
51. Goldberg AH, Gibaldi M, Kanig JL 1966. INCREASING DISSOLUTION RATES AND GASTROINTESTINAL ABSORPTION OF DRUGS VIA SOLID SOLUTIONS AND EUTECTIC MIXTURES .2. EXPERIMENTAL EVALUATION OF A EUTECTIC MIXTURE - UREA-ACETAMINOPHEN SYSTEM. Journal of Pharmaceutical Sciences 55(5):482-&.
52. Mattok GL, McGilver.Ij, Mainvill.Ca 1971. ACETAMINOPHEN .3. DISSOLUTION STUDIES OF COMMERCIAL TABLETS OF ACETAMINOPHEN

AND COMPARISON WITH IN-VIVO ABSORPTION PARAMETERS. Journal of Pharmaceutical Sciences 60(4):561-&.
53. Gray V, Kelly G, Xia M, Butler C, Thomas S, Mayock S 2009. The Science of USP 1 and 2 Dissolution: Present Challenges and Future Relevance. Pharmaceutical Research 26(6):1289-1302.
54. van der Weerd J, Kazarian SG 2004. Combined approach of FTIR imaging and conventional dissolution tests applied to drug release. Journal of Controlled Release 98(2):295-305.
55. Ostergaard J, Meng-Lund E, Larsen SW, Larsen C, Petersson K, Lenke J, Jensen H 2010. Real-Time UV Imaging of Nicotine Release from Transdermal Patch. Pharmaceutical Research 27(12):2614-2623.
56. Boetker JP, Savolainen M, Koradia V, Tian F, Rades T, Mullertz A, Cornett C, Rantanen J, Ostergaard J 2011. Insights into the Early Dissolution Events of Amlodipine Using UV Imaging and Raman Spectroscopy. Molecular Pharmaceutics 8(4):1372-1380.
57. Herschel W 1800. Experiments on the Refrangibility of the Invisible Rays of the Sun. Phil Trans R Soc Lond 90:284-292.
58. Williams PC, Norris KH, Zarowski WS 1982. INFLUENCE OF TEMPERATURE ON ESTIMATION OF PROTEIN AND MOISTURE IN WHEAT BY NEAR-INFRARED REFLECTANCE. Cereal Chemistry 59(6):473-477.
59. Brashear RL, Flanagan DR, Luner PE, Seyer JJ, Kemper MS 1999. Diffuse reflectance near-infrared spectroscopy as a nondestructive analytical technique for polymer implants. Journal of Pharmaceutical Sciences 88(12):1348-1353.
60. Kramer K, Ebel S 2000. Application of NIR reflectance spectroscopy for the identification of pharmaceutical excipients. Analytica Chimica Acta 420(2):155-161.
61. Ulmschneider M, Barth G, Trenka E 2000. Building transferable cluster calibrations for the identification of different solid excipients with near-infrared spectroscopy. Pharmazeutische Industrie 62(5):374-376.
62. Thosar SS, Forbess RA, Kemper M, Shukla AJ 1999. Determination of copolymer ratios of poly(lactide-co-glycolide) using near-infrared spectroscopy. Journal of Pharmaceutical and Biomedical Analysis 20(1-2):107-114.
63. Gustafsson C, Nystrom C, Lennholm H, Bonferoni MC, Caramella CM 2003. Characteristics of hydroxypropyl methylcellulose influencing compactibility and prediction of particle and tablet properties by infrared spectroscopy. Journal of Pharmaceutical Sciences 92(3):494-504.
64. Ilari JL, Martens H, Isaksson T 1988. DETERMINATION OF PARTICLE-SIZE IN POWDERS BY SCATTER CORRECTION IN DIFFUSE NEAR-INFRARED REFLECTANCE. Applied Spectroscopy 42(5):722-728.
65. O'Neil AJ, Jee RD, Moffat AC 1998. The application of multiple linear regression to the measurement of the median particle size of drugs and pharmaceutical excipients by near-infrared spectroscopy. Analyst 123(11):2297-2302.
66. Otsuka M 2004. Comparative particle size determination of phenacetin bulk powder by using Kubelka-Munk theory and principal component regression analysis based on near-infrared spectroscopy. Powder Technology 141(3):244-250.
67. Szalay A, Antal I, Zsigmond Z, Marton S, Eros I, Regdon G, Pintye-Hodi K 2005. Technical note: Study on the relationship between particle size and near infrared diffuse reflectance spectroscopic data. Particle & Particle Systems Characterization 22(3):219-222.
68. Otsuka M, Fukui Y 2010. Determination of carbamazepine polymorphic contents in double-layered tablets using transmittance- and reflectance-near-infrared spectroscopy involving chemometrics. Drug Development and Industrial Pharmacy 36(12):1404-1412.
69. Patel AD, Luner PE, Kemper MS 2001. Low-level determination of polymorph composition in physical mixtures by near-infrared reflectance spectroscopy. Journal of Pharmaceutical Sciences 90(3):360-370.

70. Seyer JJ, Luner PE, Kemper MS 2000. Application of diffuse reflectance near-infrared spectroscopy for determination of crystallinity. Journal of Pharmaceutical Sciences 89(10):1305-1316.
71. Bai SJ, Rani M, Suryanarayanan R, Carpenter JF, Nayar R, Manning MC 2004. Quantification of glycine crystallinity by near-infrared (NIR) spectroscopy. Journal of Pharmaceutical Sciences 93(10):2439-2447.
72. Buckton G, Yonemochi E, Hammond J, Moffat A 1998. The use of near infra-red spectroscopy to detect changes in the form of amorphous and crystalline lactose. International Journal of Pharmaceutics 168(2):231-241.
73. Hogan SE, Buckton G 2001. Water sorption/desorption - near IR and calorimetric study of crystalline and amorphous raffinose. International Journal of Pharmaceutics 227(1-2):57-69.
74. Singh P, Jangir DK, Mehrotra R, Bakhshi AK 2009. Development and validation of an infrared spectroscopy-based method for the analysis of moisture content in 5-fluorouracil. Drug Testing and Analysis 1(5-6):275-278.
75. Mantanus J, Ziemons E, Lebrun P, Rozet E, Klinkenberg R, Streel B, Evrard B, Hubert P 2009. Moisture content determination of pharmaceutical pellets by near infrared spectroscopy: Method development and validation. Analytica Chimica Acta 642(1-2):186-192.
76. Berntsson O, Zackrisson G, Ostling G 1997. Determination of moisture in hard gelatin capsules using near-infrared spectroscopy: Applications to at-line process control of pharmaceutics. Journal of Pharmaceutical and Biomedical Analysis 15(7):895-900.
77. Gupta A, Peck GE, Miller RW, Morris KR 2005. Real-time near-infrared monitoring of content uniformity, moisture content, compact density, tensile strength, and Young's modulus of roller compacted powder blends. Journal of Pharmaceutical Sciences 94(7):1589-1597.
78. Zhou GX, Ge ZH, Dorwart J, Izzo B, Kukura J, Bicker G, Wyvratt J 2003. Determination and differentiation of surface and bound water in drug substances by near infrared spectroscopy. Journal of Pharmaceutical Sciences 92(5):1058-1065.
79. Sinsheim.Je, Poswalk NM 1968. PHARMACEUTICAL APPLICATIONS OF NEAR INFRARED DETERMINATION OF WATER. Journal of Pharmaceutical Sciences 57(11):2007-&.
80. Jones JA, Last IR, Macdonald BF, Prebble KA 1993. DEVELOPMENT AND TRANSFERABILITY OF NEAR-INFRARED METHODS FOR DETERMINATION OF MOISTURE IN A FREEZE-DRIED INJECTION PRODUCT. Journal of Pharmaceutical and Biomedical Analysis 11(11-12):1227-1231.
81. Chen YX, Thosar SS, Forbess RA, Kemper MS, Rubinovitz RL, Shukla AJ 2001. Prediction of drug content and hardness of intact tablets using artificial neural network and near-infrared spectroscopy. Drug Development and Industrial Pharmacy 27(7):623-631.
82. Otsuka M, Yamane I 2006. Prediction of tablet hardness based on near infrared spectra of raw mixed powders by chemometrics. Journal of Pharmaceutical Sciences 95(7):1425-1433.
83. Donoso M, Kildsig DO, Ghaly ES 2003. Prediction of tablet hardness and porosity using near-infrared diffuse reflectance spectroscopy as a nondestructive method. Pharmaceutical Development and Technology 8(4):357-366.
84. Kirsch JD, Drennen JK 1999. Nondestructive tablet hardness testing by near-infrared spectroscopy: a new and robust spectral best-fit algorithm. Journal of Pharmaceutical and Biomedical Analysis 19(3-4):351-362.
85. Blanco M, Alcala M 2006. Content uniformity and tablet hardness testing of intact pharmaceutical tablets by near infrared spectroscopy - A contribution to process analytical technologies. Analytica Chimica Acta 557(1-2):353-359.

86. Cahyadi C, Karande AD, Chan LW, Heng PWS 2010. Comparative study of non-destructive methods to quantify thickness of tablet coatings. International Journal of Pharmaceutics 398(1-2):39-49.
87. Romer M, Heinamaki J, Strachan C, Sandler N, Yliruusi J 2008. Prediction of Tablet Film-coating Thickness Using a Rotating Plate Coating System and NIR Spectroscopy. Aaps Pharmscitech 9(4):1047-1053.
88. Wray PS, Clarke GS, Kazarian SG 2013. Dissolution of tablet-in-tablet formulations studied with ATR-FTIR spectroscopic imaging. European Journal of Pharmaceutical Sciences 48(4-5):748-757.
89. Tabasi SH, Moolchandani V, Fahmy R, Hoag SW 2009. Sustained release dosage forms dissolution behavior prediction: A study of matrix tablets using NIR spectroscopy. International Journal of Pharmaceutics 382(1-2):1-6.
90. Moltgen CV, Puchert T, Menezes JC, Lochmann D, Reich G 2012. A novel in-line NIR spectroscopy application for the monitoring of tablet film coating in an industrial scale process. Talanta 92:26-37.
91. Moes JJ, Ruijken MM, Gout E, Frijlink HW, Ugwoke MI 2008. Application of process analytical technology in tablet process development using NIR spectroscopy: Blend uniformity, content uniformity and coating thickness measurements. International Journal of Pharmaceutics 357(1-2):108-118.
92. Souza JAL, Albuquerque MM, Grangeiro S, Pimentel MF, de Santana DP, Simoes SS 2012. Quantification of captopril disulphide as a degradation product in captopril tablets using near infrared spectroscopy and chemometrics. Vibrational Spectroscopy 62:35-41.
93. Drennen JK, Lodder RA 1990. NONDESTRUCTIVE NEAR-INFRARED ANALYSIS OF INTACT TABLETS FOR DETERMINATION OF DEGRADATION PRODUCTS. Journal of Pharmaceutical Sciences 79(7):622-627.
94. Moffat AC, Assi S, Watt RA 2010. Identifying counterfeit medicines using near infrared spectroscopy. Journal of near Infrared Spectroscopy 18(1):1-15.
95. Petri J, Kaunzinger A, Niemoller A, Karas M 2005. Quality control of tablets by Near Infrared (NIR)-Spectroscopy. Pharmazie 60(10):743-746.
96. Alvarenga L, Ferreira D, Altekruse D, Menezes JC, Lochmann D 2008. Tablet identification using near-infrared spectroscopy (NIRS) for pharmaceutical quality control. Journal of Pharmaceutical and Biomedical Analysis 48(1):62-69.
97. Hennigan MC, Ryder AG 2013. Quantitative polymorph contaminant analysis in tablets using Raman and near infra-red spectroscopies. Journal of Pharmaceutical and Biomedical Analysis 72:163-171.
98. Karande AD, Heng PWS, Liew CV 2010. In-line quantification of micronized drug and excipients in tablets by near infrared (NIR) spectroscopy: Real time monitoring of tabletting process. International Journal of Pharmaceutics 396(1-2):63-74.
99. Aldridge PK, Mushinsky RF, Andino MM, Evans CL 1994. IDENTIFICATION OF TABLET FORMULATIONS INSIDE BLISTER PACKAGES BY NEAR-INFRARED SPECTROSCOPY. Applied Spectroscopy 48(10):1272-1276.
100. Dempster MA, Macdonald BF, Gemperline PJ, Boyer NR 1995. A NEAR-INFRARED REFLECTANCE ANALYSIS METHOD FOR THE NONINVASIVE IDENTIFICATION OF FILM-COATED AND NON-FILM-COATED, BLISTER-PACKED TABLETS. Analytica Chimica Acta 310(1):43-51.
101. Raman CV 1928. A change of wave-length in light scattering. Nature 121:619-619.
102. Widjaja E, Seah RKH 2008. Application of Raman microscopy and band-target entropy minimization to identify minor components in model pharmaceutical tablets. Journal of Pharmaceutical and Biomedical Analysis 46(2):274-281.

103. Vajna B, Farkas I, Szabo A, Zsigmond Z, Marosi G 2010. Raman microscopic evaluation of technology dependent structural differences in tablets containing imipramine model drug. Journal of Pharmaceutical and Biomedical Analysis 51(1):30-38.
104. Henson MJ, Zhang L 2006. Drug characterization in low dosage pharmaceutical tablets using Raman microscopic mapping. Applied Spectroscopy 60(11):1247-1255.
105. Sasic S, Whitlock M 2008. Raman mapping of low-content active-ingredient pharmaceutical formulations. Part II: Statistically optimized sampling for detection of less than 1% of an active pharmaceutical ingredient. Applied Spectroscopy 62(8):916-921.
106. Windbergs M, Haaser M, McGoverin CM, Gordon KC, Kleinebudde P, Strachan CJ 2010. Investigating the Relationship between Drug Distribution in Solid Lipid Matrices and Dissolution Behaviour Using Raman Spectroscopy and Mapping. Journal of Pharmaceutical Sciences 99(3):1464-1475.
107. Zhang L, Henson MJ, Sekulic SS 2005. Multivariate data analysis for Raman imaging of a model pharmaceutical tablet. Analytica Chimica Acta 545(2):262-278.
108. Gendrin C, Roggo Y, Collet C 2008. Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review. Journal of Pharmaceutical and Biomedical Analysis 48(3):533-553.
109. Vajna B, Farkas I, Farkas A, Pataki H, Nagy Z, Madarasz J, Marosi G 2011. Characterization of drug-cyclodextrin formulations using Raman mapping and multivariate curve resolution. Journal of Pharmaceutical and Biomedical Analysis 56(1):38-44.
110. Sasic S, Clark DA, Mitchell JC, Snowden MJ 2004. A comparison of Raman chemical images produced by univariate and multivariate data processing - a simulation with an example from pharmaceutical practice. Analyst 129(11):1001-1007.
111. Beebe KR, Kowalski BR 1987. AN INTRODUCTION TO MULTIVARIATE CALIBRATION AND ANALYSIS. Analytical Chemistry 59(17):A1007-&.
112. Hotelling H 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 24:417-441.
113. Hotelling H 1933. Analysis of a complex of statistical variables into principal components. Journal of Educational Psychology 24:498-520.
114. Pearson K 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine 2(7-12):559-572.
115. Bajwa SG, Bajcsy P, Groves P, Tian LE 2004. Hyperspectral image data mining for band selection in agricultural applications. Transactions of the Asae 47(3):895-907.
116. Ye X, Sakai K, Sasao A, Asada S 2008. Potential of airborne hyperspectral imagery to estimate fruit yield in citrus. Chemometrics and Intelligent Laboratory Systems 90(2):132-144.
117. Winson MK, Goodacre R, Timmins EM, Jones A, Alsberg BK, Woodward AM, Rowland JJ, Kell DB 1997. Diffuse reflectance absorbance spectroscopy taking in chemometrics (DRASTIC). A hyperspectral FT-IR-based approach to rapid screening for metabolite overproduction. Analytica Chimica Acta 348(1-3):273-282.
118. Williams PJ, Geladi P, Britz TJ, Manley M 2012. Near-infrared (NIR) hyperspectral imaging and multivariate image analysis to study growth characteristics and differences between species and strains of members of the genus Fusarium. Analytical and Bioanalytical Chemistry 404(6-7):1759-1769.
119. Williams P, Geladi P, Fox G, Manley M 2009. Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. Analytica Chimica Acta 653(2):121-130.
120. Usenik P, Burmen M, Fidler A, Pernus F, Likar B 2012. Automated Classification and Visualization of Healthy and Diseased Hard Dental Tissues by Near-Infrared Hyperspectral Imaging. Applied Spectroscopy 66(9):1067-1074.
121. Nuffer LL, Medvick PA, Foote HP, Solinsky JC 2006. Multispectral/hyperspectral image enhancement for biological cell analysis. Cytometry Part A 69A(8):897-903.

122. Liu ZY, Shi JJ, Zhang LW, Huang JF 2010. Discrimination of rice panicles by hyperspectral reflectance data based on principal component analysis and support vector classification. Journal of Zhejiang University-Science B 11(1):71-78.
123. Feng YZ, Sun DW 2012. Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review. Critical Reviews in Food Science and Nutrition 52(11):1039-1058.
124. Baer I, Gurny R, Margot P 2007. NIR analysis of cellulose and lactose- Application to ecstasy tablet analysis. Forensic Science International 167(2-3):234-241.
125. Watanabe A, Morita S, Ozaki Y 2006. A study on water adsorption onto microcrystalline cellulose by near-infrared spectroscopy with two-dimensional correlation spectroscopy and principal component analysis. Applied Spectroscopy 60(9):1054-1061.
126. Luypaert J, Massart DL, Heyden YV 2007. Near-infrared spectroscopy applications in pharmaceutical analysis. Talanta 72(3):865-883.
127. Reich G 2005. Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. Advanced Drug Delivery Reviews 57(8):1109-1143.
128. Blanco M, Coello J, Iturriaga H, Maspoch S, de la Pezuela C 1998. Near-infrared spectroscopy in the pharmaceutical industry. Analyst 123(8):135R-150R.
129. Reid LM, O'Donnell CP, Downey G 2006. Recent technological advances for the determination of food authenticity. Trends in Food Science & Technology 17(7):344-353.
130. Amigo JM 2010. Practical issues of hyperspectral imaging analysis of solid dosage forms. Analytical and Bioanalytical Chemistry 398(1):93-109.
131. Cozzolino D, Smyth HE, Gishen M 2003. Feasibility study on the use of visible and near-infrared Spectroscopy together with chemometrics to discriminate between commercial white wines of different varietal origins. Journal of Agricultural and Food Chemistry 51(26):7703-7708.
132. Liu MY, Meng Y, Ren YL, Zhang HQ 2007. Nondestructive quantitative analysis of cofrel medicines by improved partial least squares-NIR spectroscopy. Spectroscopy and Spectral Analysis 27(6):1098-1101.
133. Baptistao M, Rocha WFD, Poppi RJ 2011. Quality control of the paracetamol drug by chemometrics and imaging spectroscopy in the near infrared region. Journal of Molecular Structure 1002(1-3):167-171.
134. Gendrin C, Roggo Y, Collet C 2007. Content uniformity of pharmaceutical solid dosage forms by near infrared hyperspectral imaging: A feasibility study. Talanta 73:733-741.
135. Chapelle O, Haffner P, Vapnik VN 1999. Support vector machines for histogram-based image classification. Ieee Transactions on Neural Networks 10(5):1055-1064.
136. Fauvel M, Benediktsson JA, Chanussot J, Sveinsson JR 2008. Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles. Ieee Transactions on Geoscience and Remote Sensing 46(11):3804-3814.
137. Burbidge R, Trotter M, Buxton B, Holden S 2001. Drug design by machine learning: support vector machines for pharmaceutical data analysis. Computers & Chemistry 26(1):5-14.
138. Bao L, Sun ZR 2002. Identifying genes related to drug anticancer mechanisms using support vector machine. Febs Letters 521(1-3):109-114.
139. Byvatov E, Fechner U, Sadowski J, Schneider G 2003. Comparison of support vector machine and artificial neural network systems for drug/nondrug classification. Journal of Chemical Information and Computer Sciences 43(6):1882-1889.
140. Franch-Lage F, Amigo JM, Skibsted E, Maspoch S, Coello J 2011. Fast assessment of the surface distribution of API and excipients in tablets using NIR-hyperspectral imaging. International Journal of Pharmaceutics 411(1-2):27-35.
141. F. GH, Paul G. 2007. Hyperspectral image analysis. ed.: John Wiley and Sons Ltd.
142. Eugene H. 1998. Optics. ed., Reading, Mass.: Reading, Mass. : Addison-Wesley.

143. Jack LK. 1999. Spectroscopy of polymers. ed., Amsterdam ; New York: Amsterdam ; New York : Elsevier.
144. Koenig JL. 1999. Spectroscopy of Polymers. 2 ed.: Elsevier Science.
145. Kaye GWC. 1995. Tables of physical and chemical constants. ed., Essex, England ; New York: Essex, England ; New York : Longman.
146. Abdi H, Williams LJ 2010. - Principal component analysis. - 2(- 4):- 459.
147. Hansen PM, Schjoerring JK 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sensing of Environment 86(4):542-553.
148. Blanco M, Villarroya I 2002. NIR spectroscopy: a rapid-response analytical tool. Trac-Trends in Analytical Chemistry 21(4):240-250.
149. Rasanen E, Sandler N 2007. Near infrared spectroscopy in the development of solid dosage forms. Journal of Pharmacy and Pharmacology 59(2):147-159.
150. Gou ML, Gong CY, Zhang J, Wang XH, Gu YC, Guo G, Chen LJ, Luo F, Zhao X, Wei YQ, Qian ZY 2010. Polymeric matrix for drug delivery: Honokiol-loaded PCL-PEG-PCL nanoparticles in PEG-PCL-PEG thermosensitive hydrogel. Journal of Biomedical Materials Research Part A 93A(1):219-226.
151. Rostami-Hodjegan A, Shiran MR, Tucker GT, Conway BR, Irwin WJ, Shaw LR, Grattan TJ 2002. A new rapidly absorbed paracetamol tablet containing sodium bicarbonate. II. Dissolution studies and in vitro/in vivo correlation. Drug Development and Industrial Pharmacy 28(5):533-543.
152. Shi ZQ, Anderson CA 2011. 2-D Image Localization in Hyperspectral Image Analysis of Pharmaceutical Materials. Journal of Pharmaceutical Innovation 6(1):2-9.
153. Buning-Pfaue H 2003. Analysis of water in food by near infrared spectroscopy. Food Chemistry 82(1):107-115.

Appendices

Appendix 1: Hyperspectral images

A1.1 Chapter 3

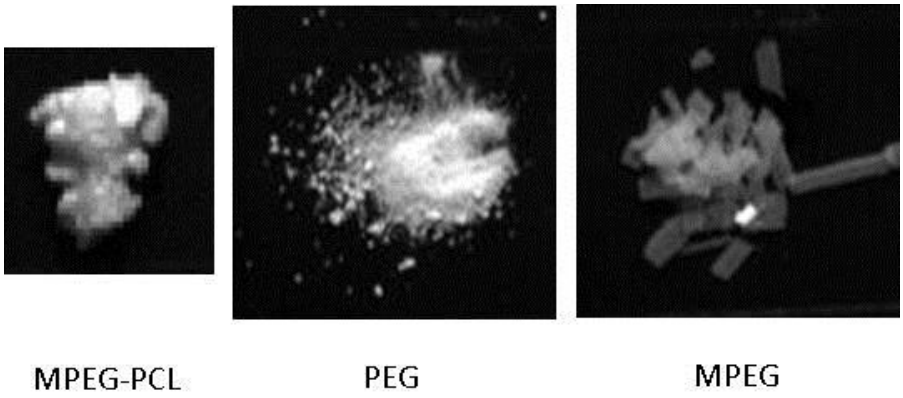


Figure A1-1 The greyscale images of the starting materials and copolymers not shown in Chapter 3

A1.2 Chapter 4

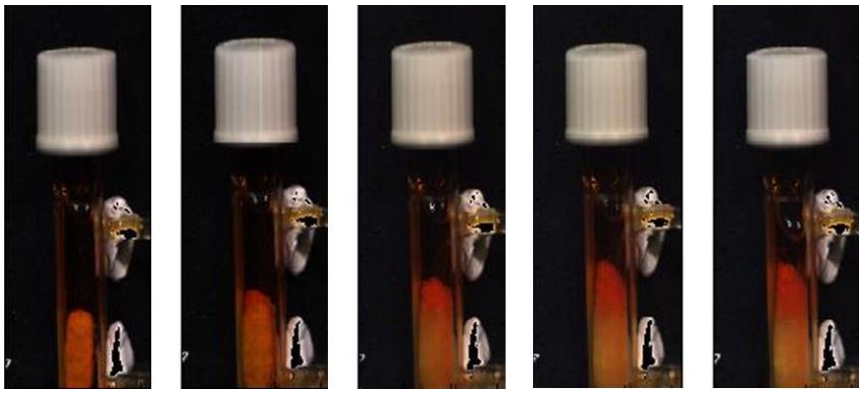


Figure A1-2 False colour images of Tablet C at 0, 3, 15, 30 and 45 minutes

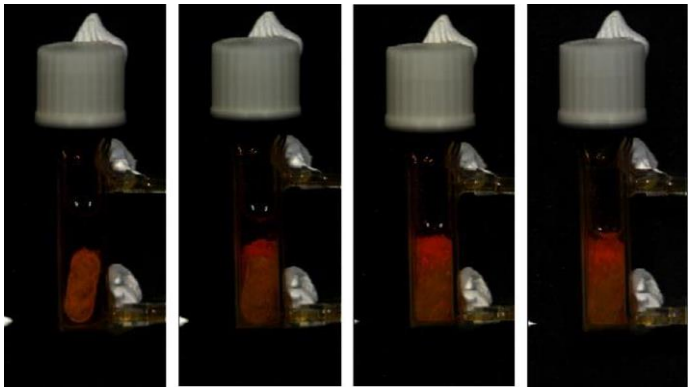


Figure A1-3 False colour images of Tablet F at 0, 15, 30 and 45 minutes

A1.3 Chapter 5

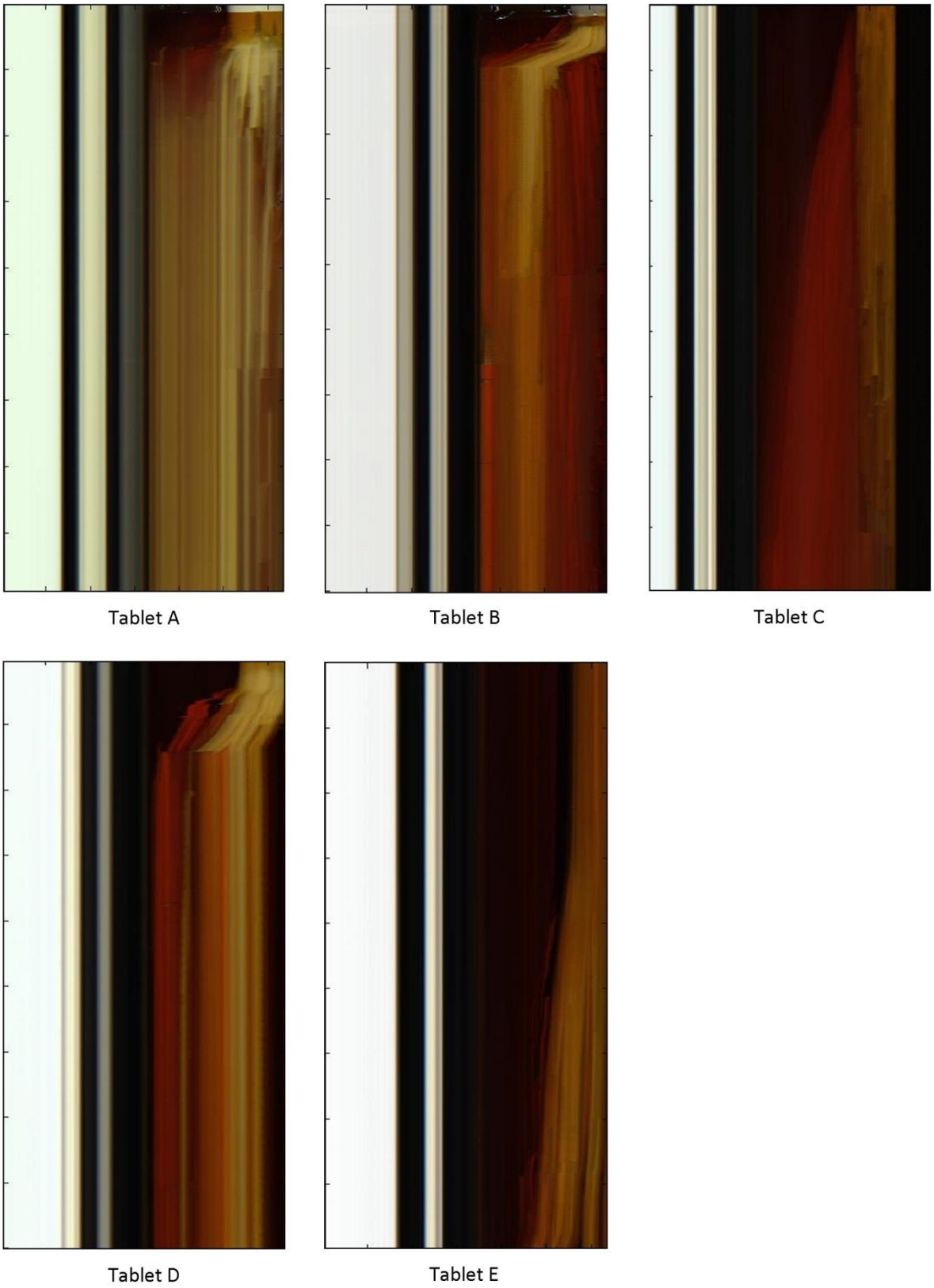


Figure A1-4 False colour images of the full data captured during the dissolution of Tablets A-E

Each tablet also had a full set of reference spectra created before the experiments began and the reference spectra for Tablet B are shown here:

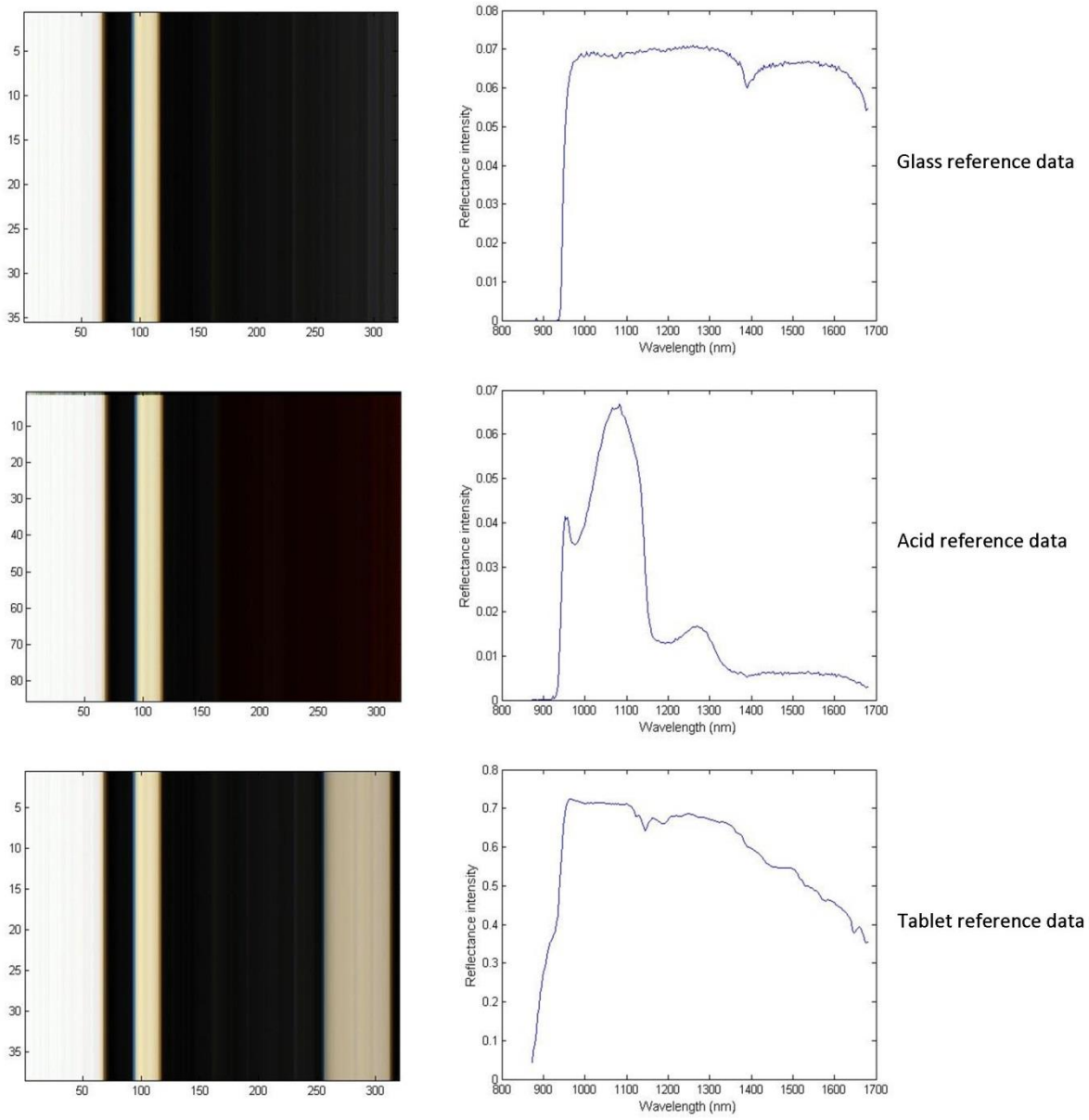


Figure A1-5 A set of reference images and their corresponding spectra

A1.4 Chapter 6

Figure A1-6 The full hyperspectral image from the dissolution of Tablet A

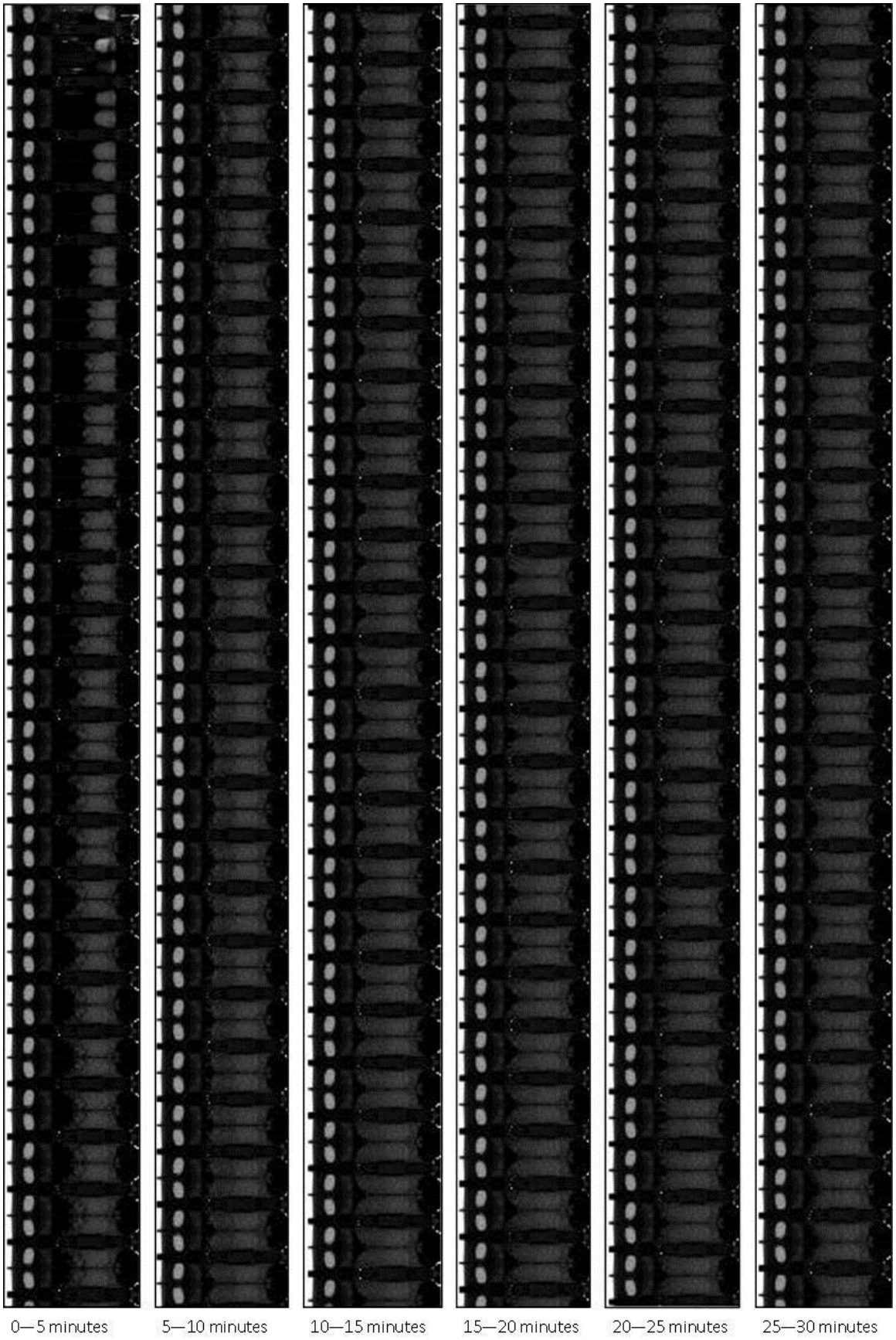


Figure A1-7 The full hyperspectral image from the dissolution of Tablet B

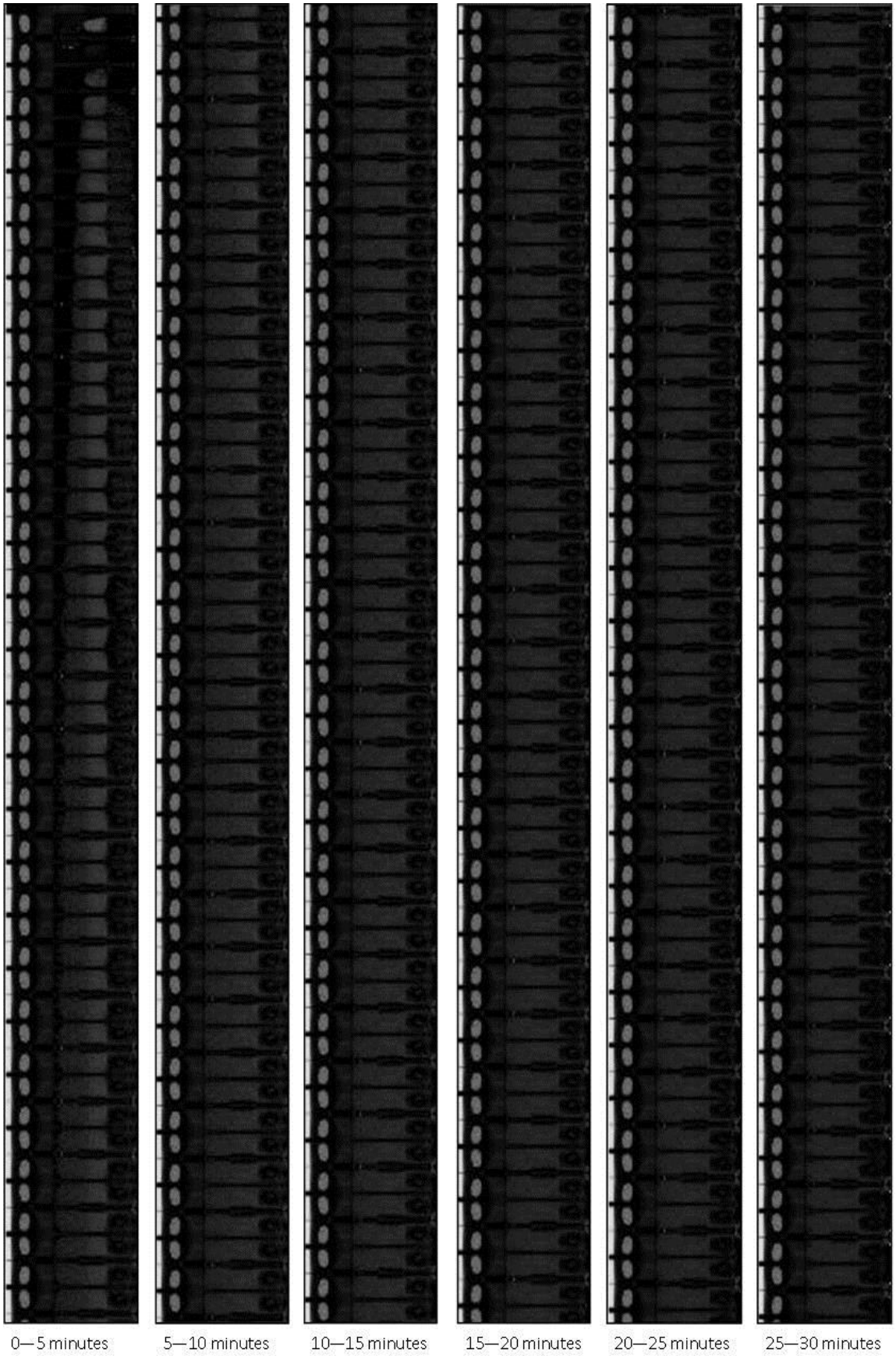


Figure A1-8 The full hyperspectral image from the dissolution of Tablet C

Appendix 2: Spectral data

A2.1 Chapter 3

A2.1.1 Scores plots

The labelled scores plot from the principal component analysis of the data collected using the visible HSI system. The labels used denote the locations of the pixel from the spectra were collected.

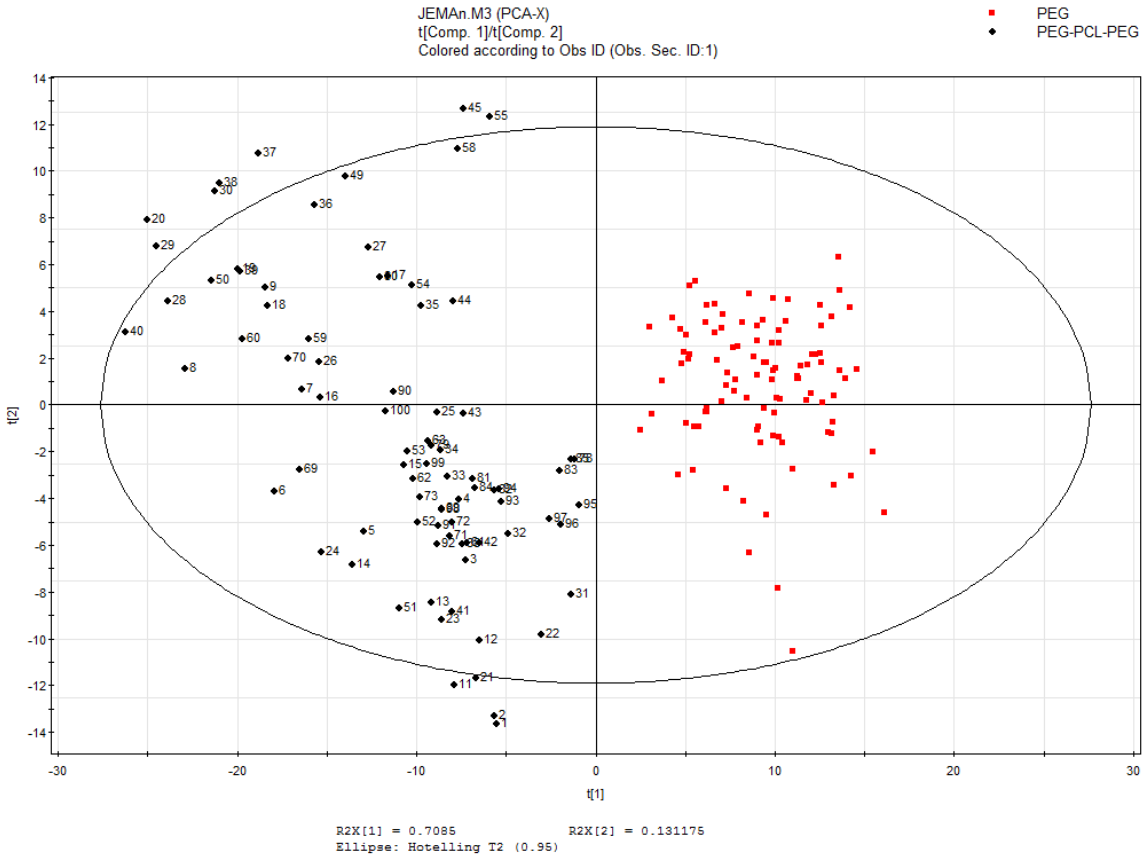


Figure A2-1The labelled scores plot of PEG and PEG-PCL-PEG

A2.2 Chapter 4

A2.2.1 Tablet Spectra

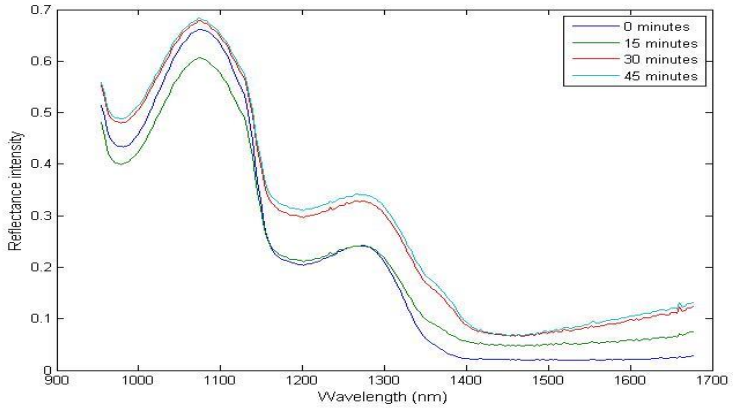


Figure A2-2 The spectra of the bulk of Tablet C at 0, 15, 30 and 45 minutes

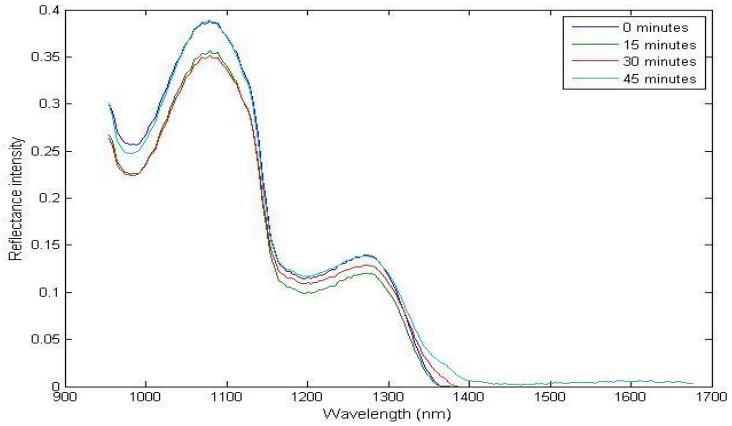


Figure A2-3 The spectra of the bulk of Tablet F at 0, 15, 30 and 45 minutes

A2.2.2 **Swelling data**

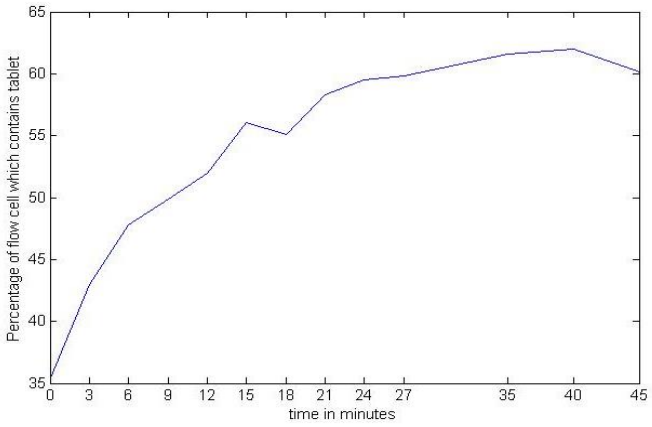


Figure A2-4 The percentage of the flow cell covered with tablet material during the dissolution of Tablet C

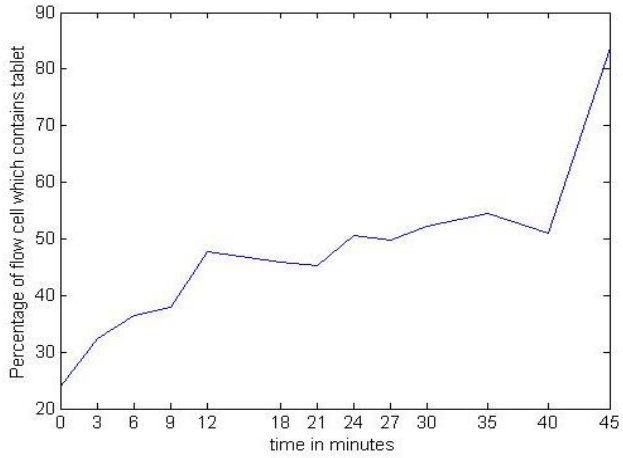


Figure A2-5 The percentage of the flow cell covered with tablet material during the dissolution of Tablet F

A2.2.3 **PCA scores plots**

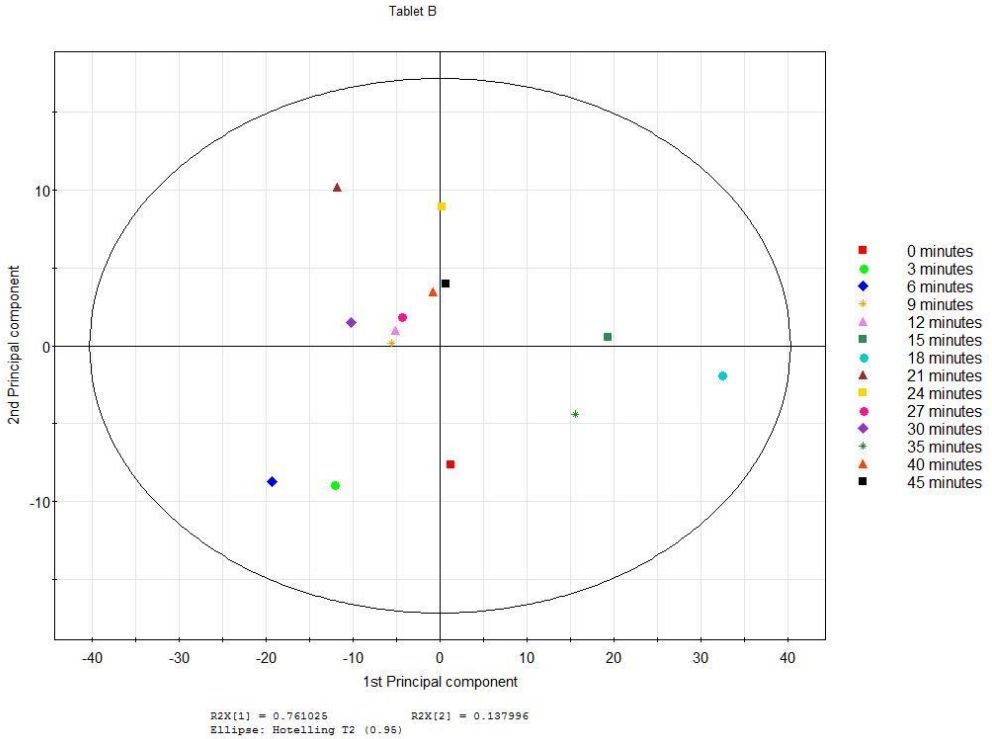


Figure A2-6 The scores plot from the PCA of Tablet B

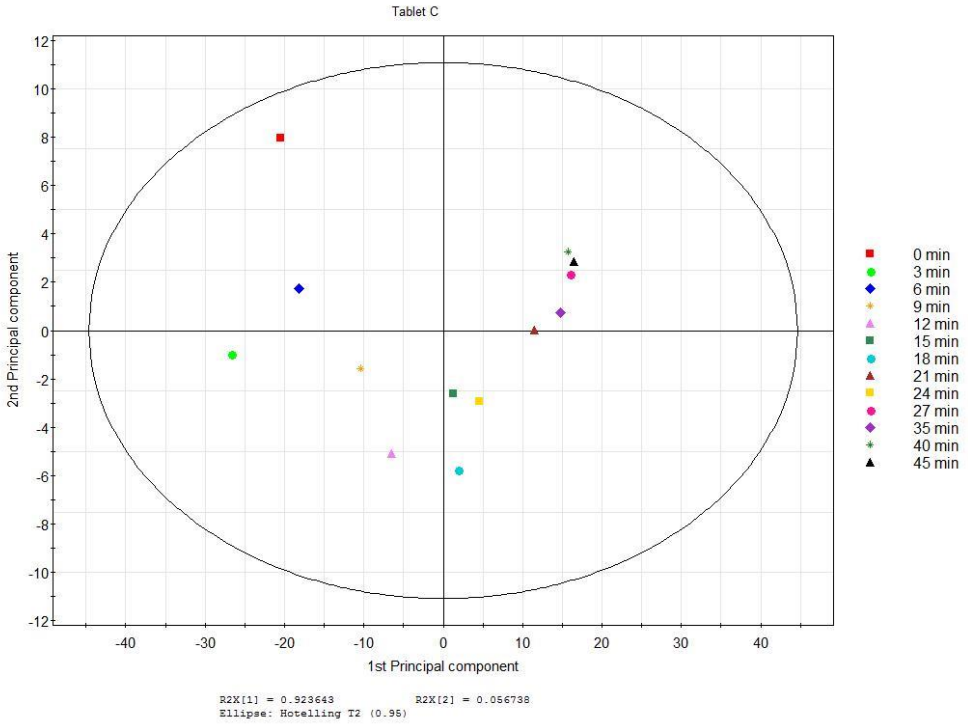


Figure A2-7 The scores plot from the PCA of Tablet C

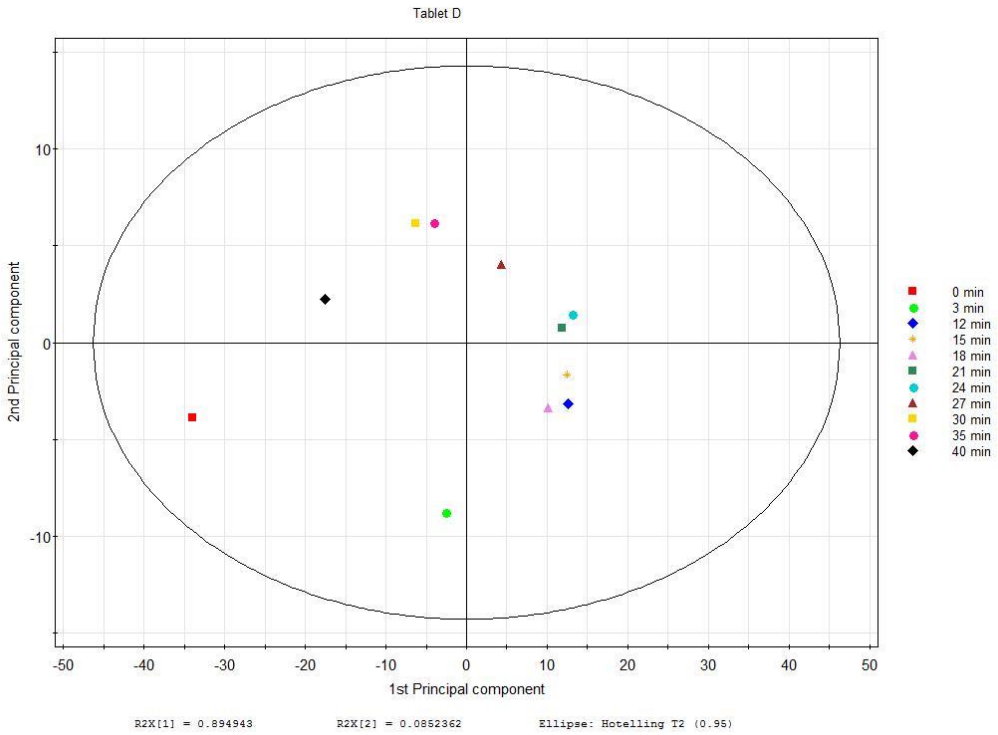


Figure A2-8 The scores plot from the PCA of Tablet D

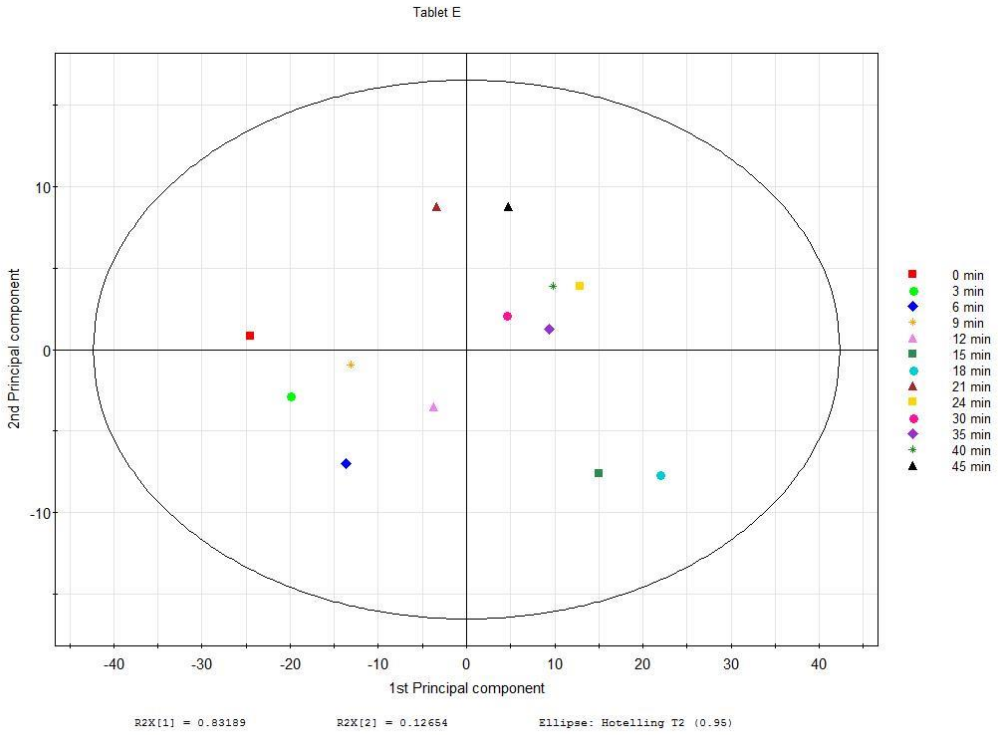


Figure A2-9 The scores plot from the PCA of Tablet E

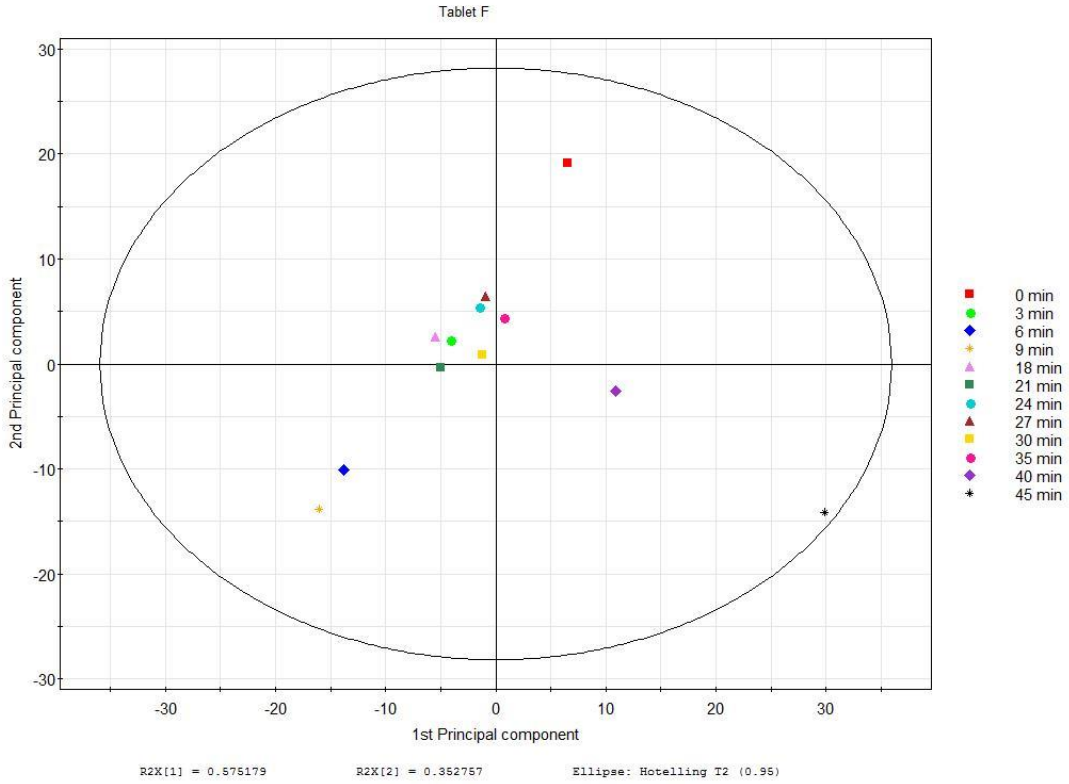


Figure A2-10 The scores plot from the PCA of Tablet F

A2.2.4 **PCA loadings plots**

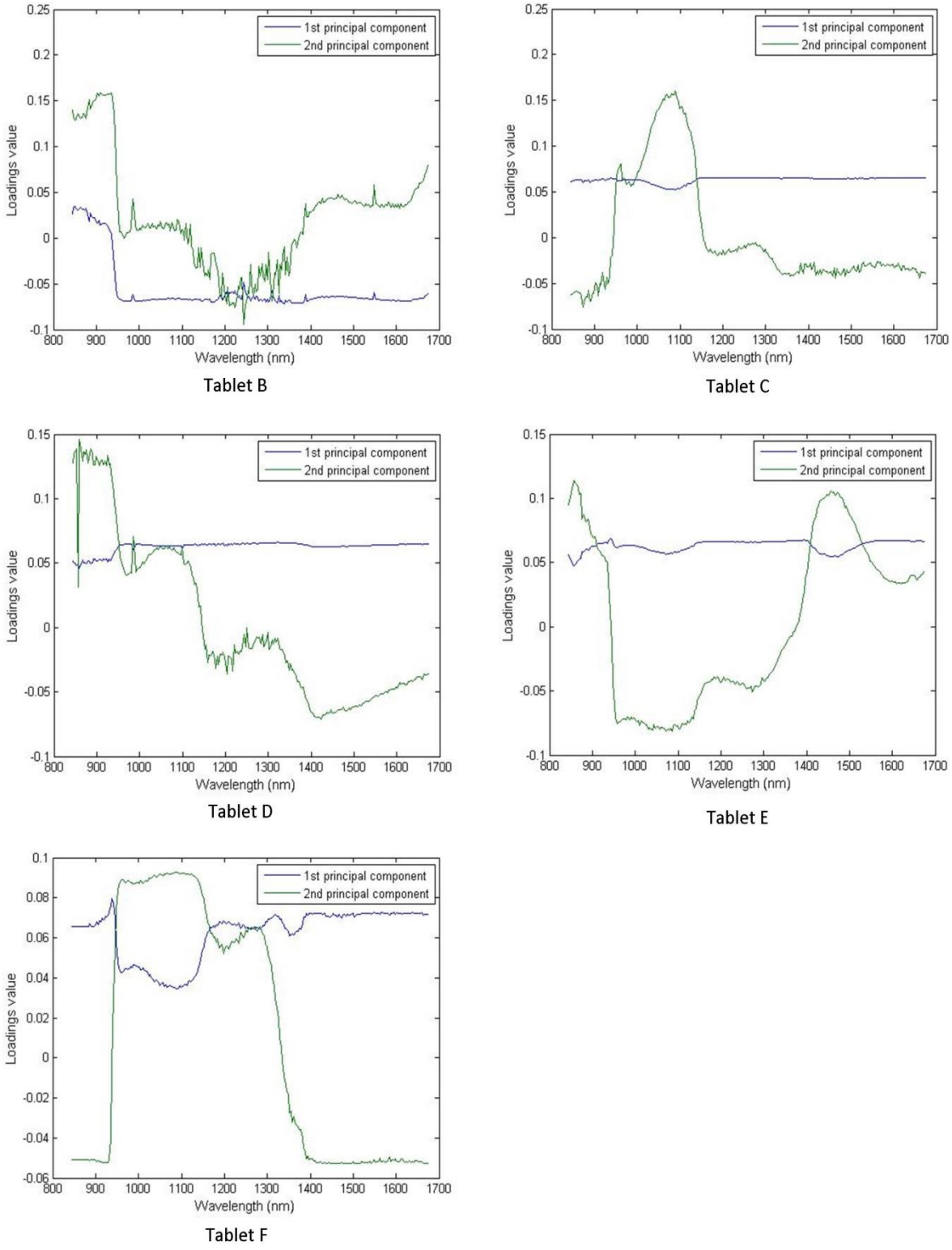


Figure A2-11 The loadings from the PCA of Tablets B-F

A2.3 Chapter 5

A2.3.1 Spectra of tablets

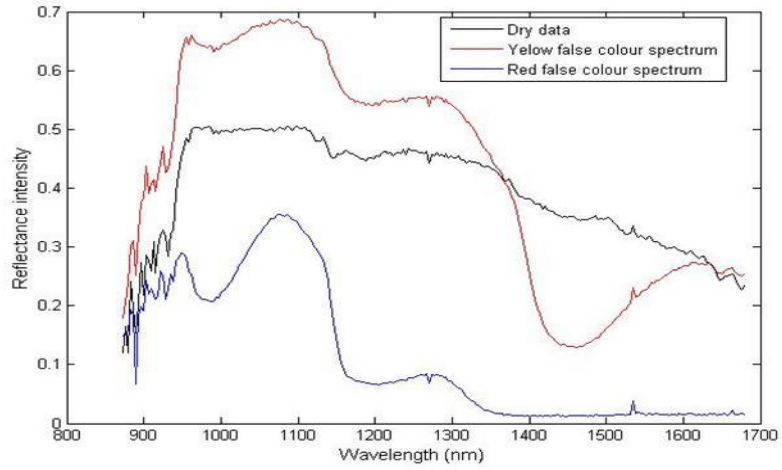


Figure A2-12 The spectra of 3 different regions in the hyperspectral image of the dissolution of Tablet B

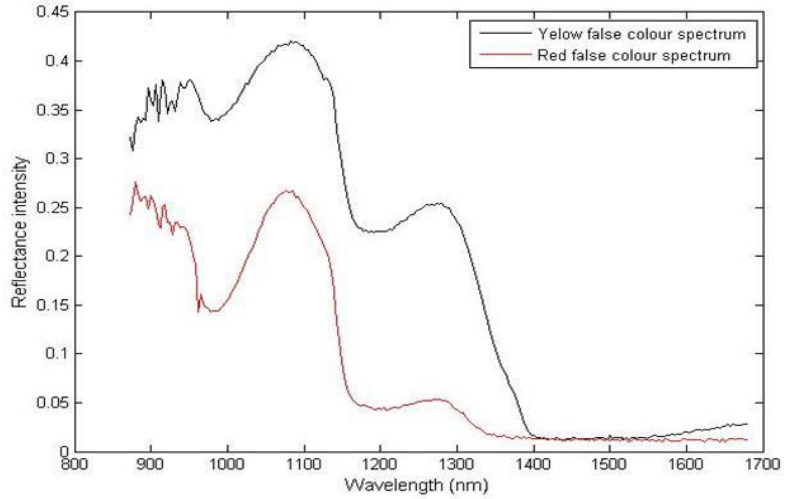


Figure A2-13 The spectra of 2 different regions in the hyperspectral image of the dissolution of Tablet C

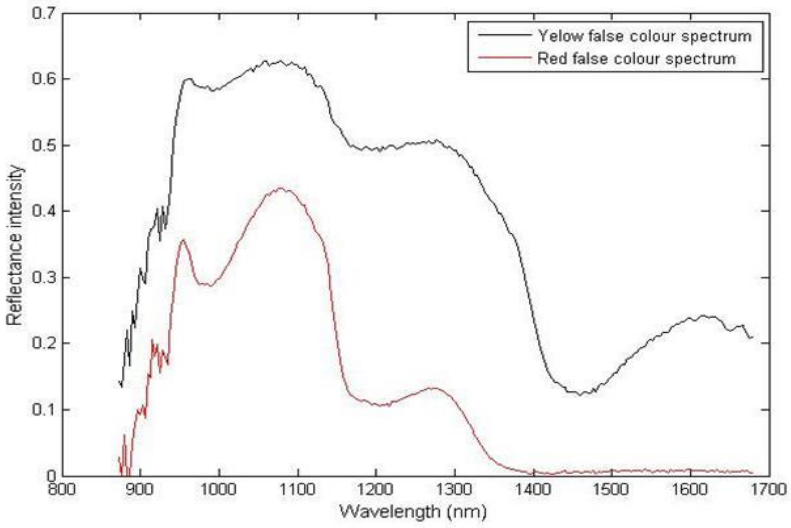


Figure A2-14 The spectra of 2 different regions in the hyperspectral image of the dissolution of Tablet D

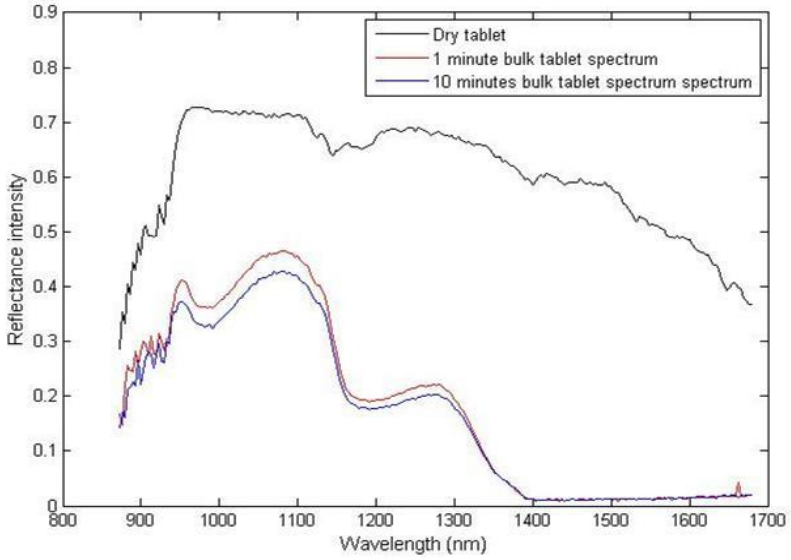


Figure A2-15 The spectra of 3 different regions in the hyperspectral image of the dissolution of Tablet E

A2.3.2 **Mesh plots detailing the spectra in the y axis of Tablet F**

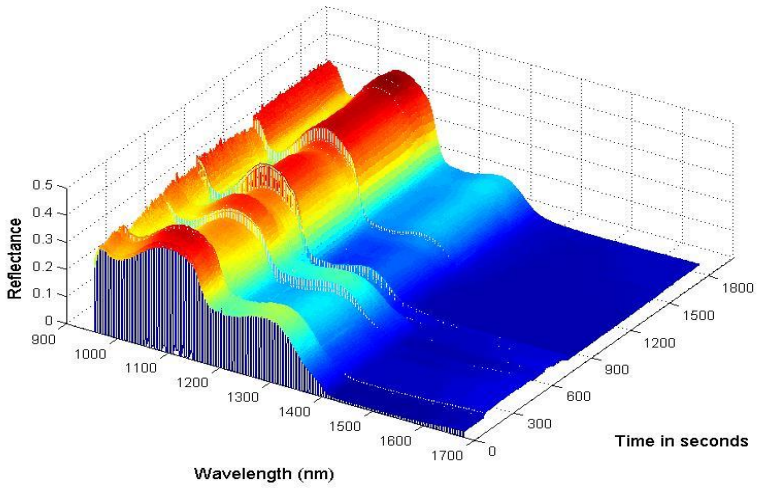


Figure A2-16 The y axis mesh plot created using the hyperspectral image of Tablet F at position 2. The relative location of this position is shown in Figure 5-10

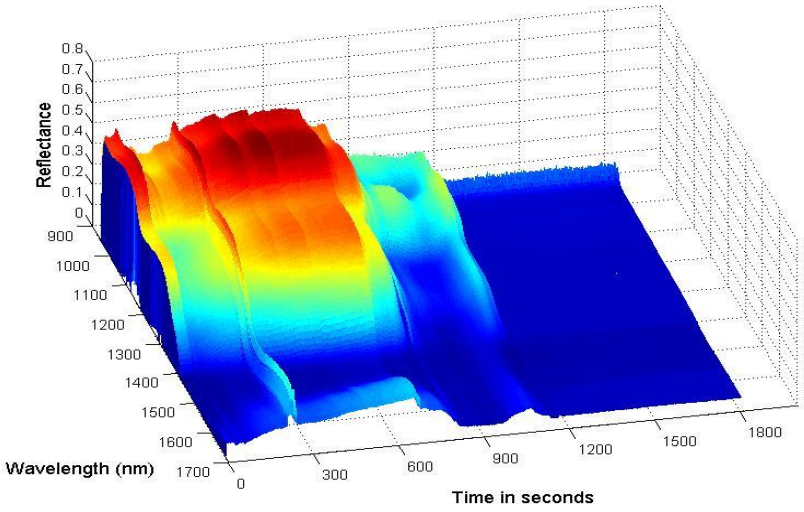
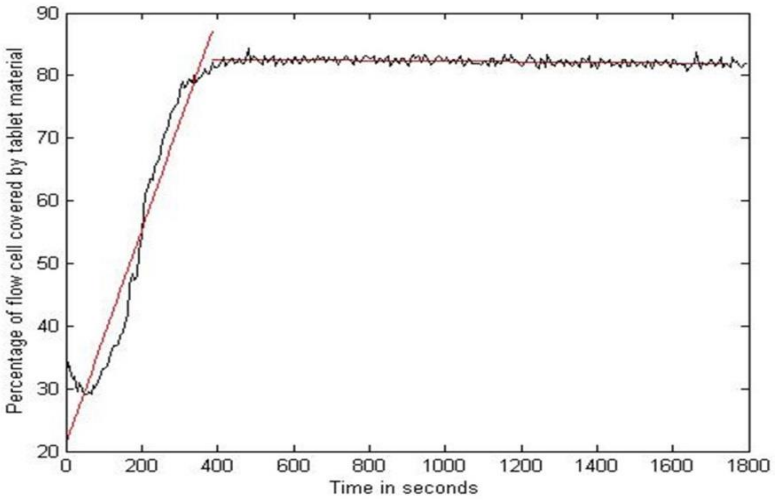


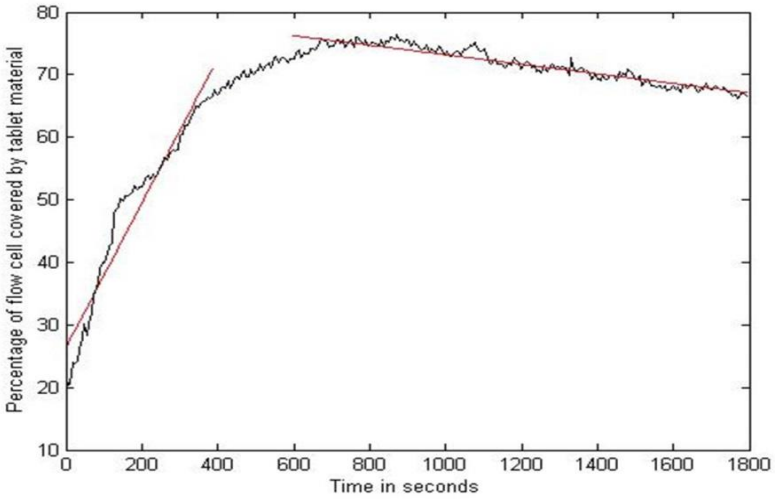
Figure A2-17 The y axis mesh plot created using the hyperspectral image of Tablet F at position 3. The relative location of this position is shown in Figure 5 10

A2.4 Chapter 6

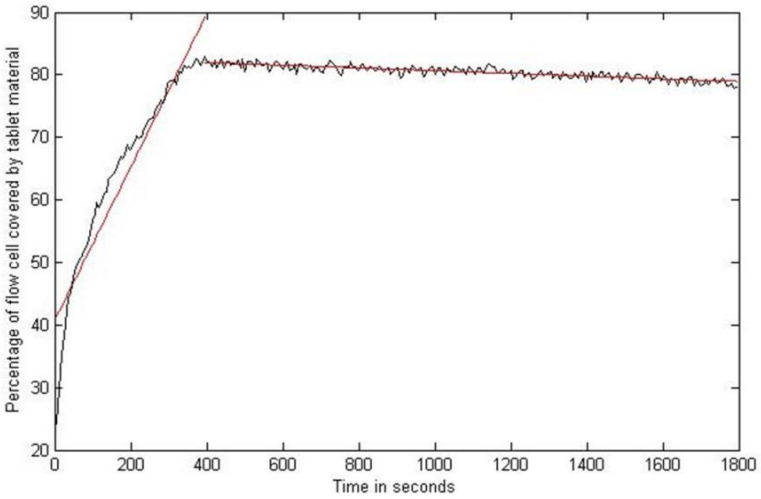
A2.4.1 Rate of expansion within tablet



Tablet A



Tablet B



Tablet C

Figure A2-18 The plots of tablet expansion during dissolution with the linear fits used to determine the rates of expansion shown in Table 6-1

A2.4.2 **Mesh plots**

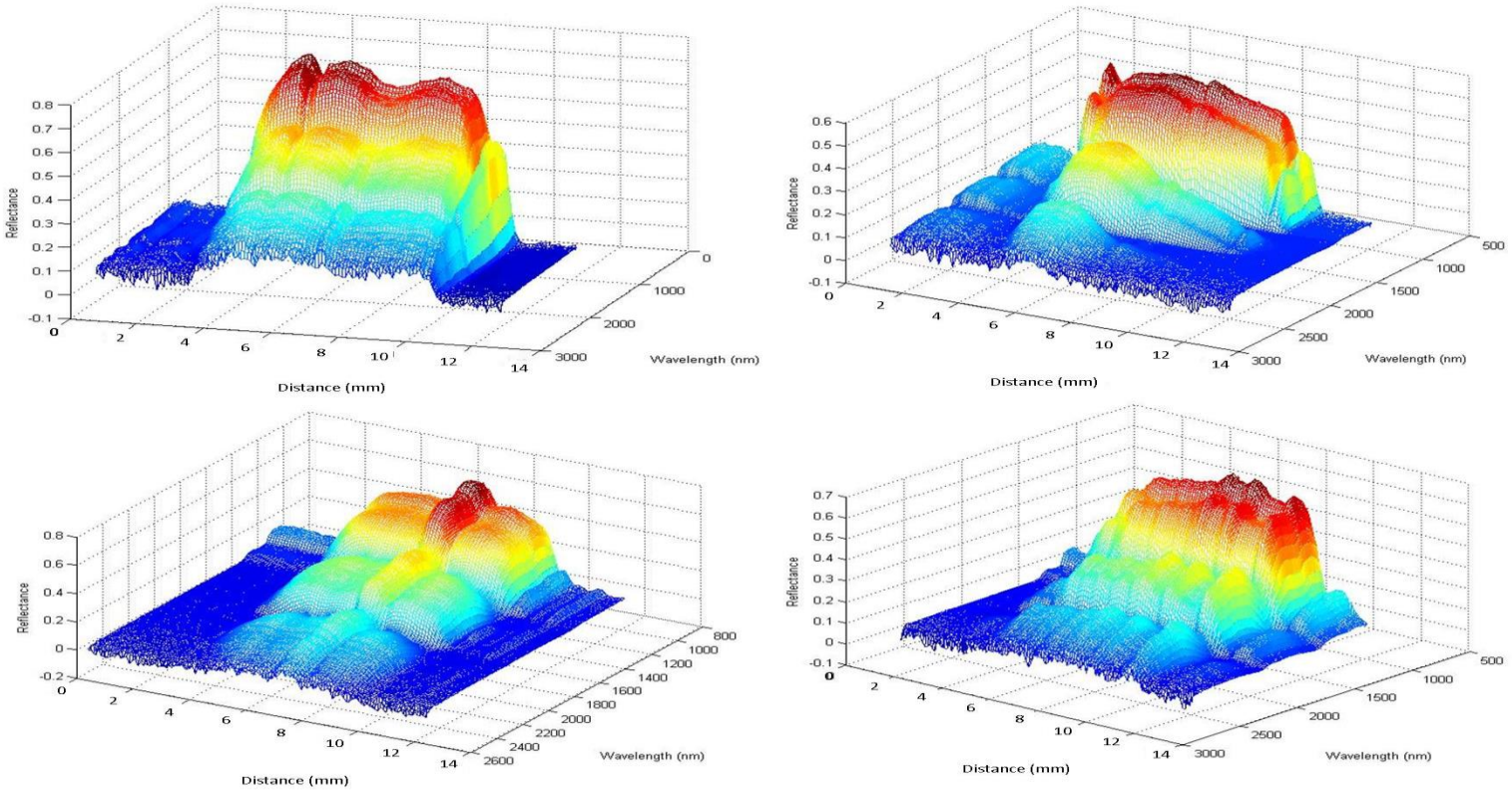


Figure A2-19 The mesh plots across the flow cell during the dissolution of Tablet A. The four mesh plots are created using the data across the flow cell, from top left to bottom right, at 1 second, 7 seconds, 70 seconds and 160 seconds.

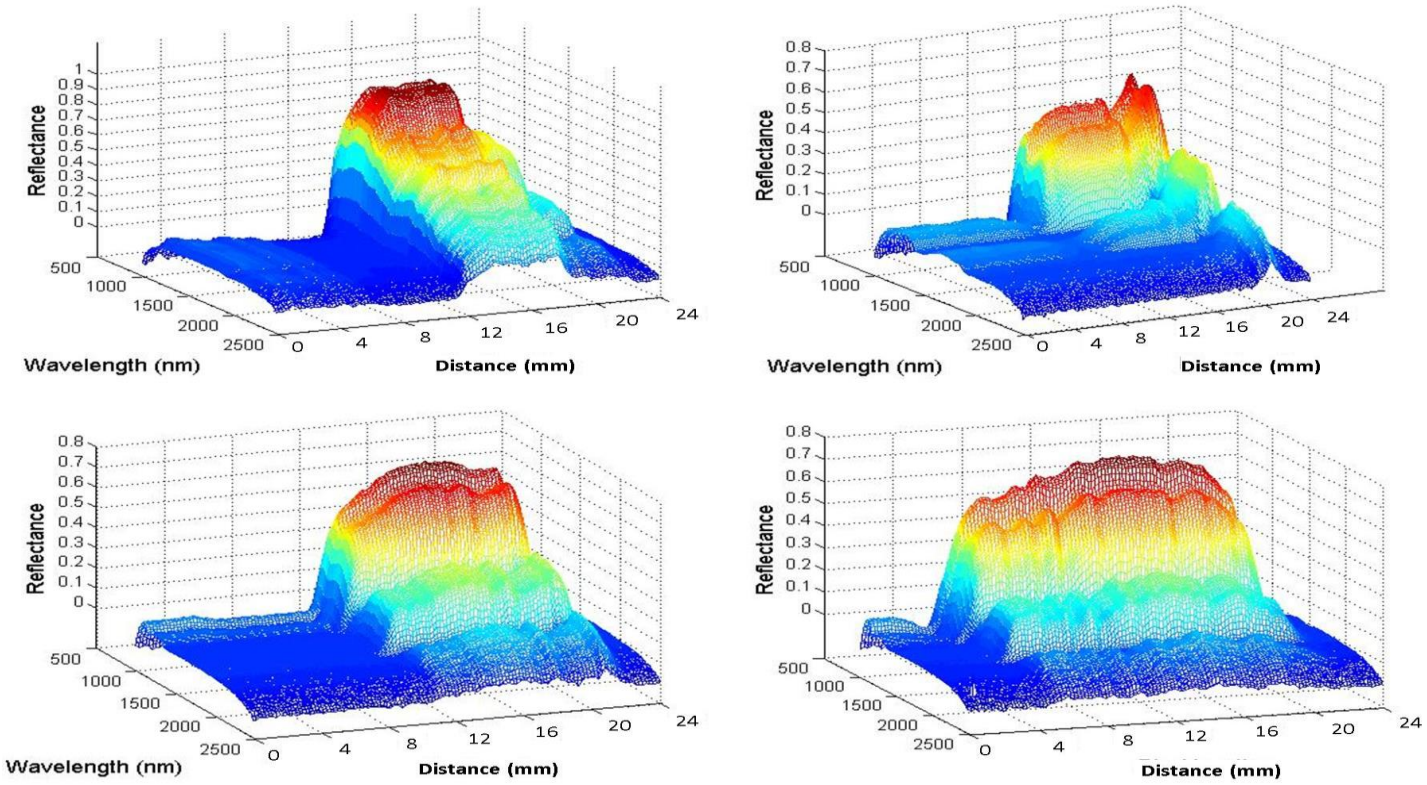


Figure A2-20 The mesh plots across the flow cell during the dissolution of Tablet A. The four mesh plots are created using the data across the flow cell, from top left to bottom right, at 2 second, 14 seconds, 35 seconds and 120 seconds

A2.4.3 **PCA data**

Tablet B

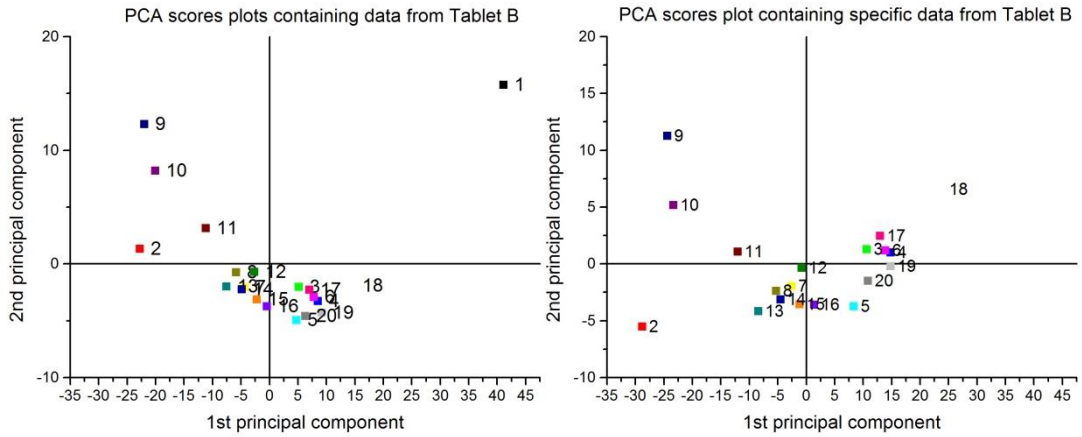


Figure A2-21 The scores plots from the PCA on different time points during the dissolution of Tablet B

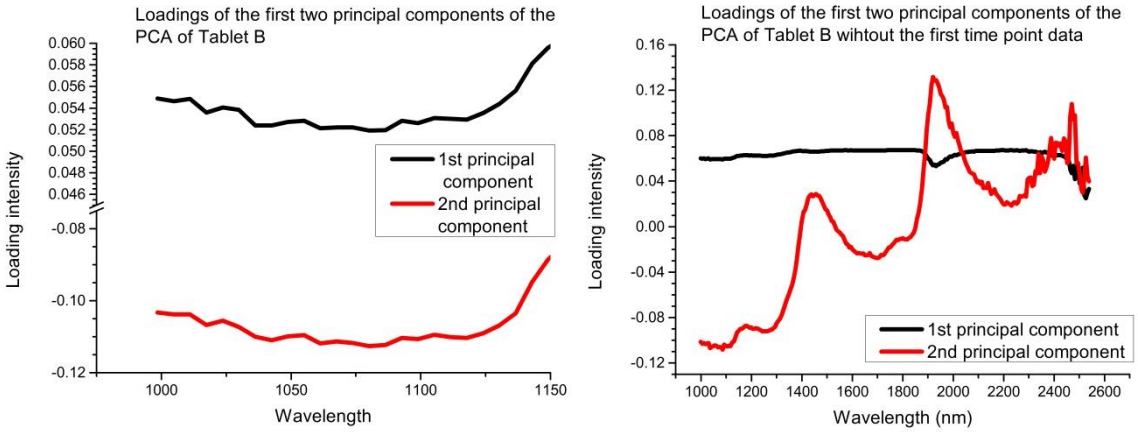


Figure A2-22 The loadings plots from the PCA on different time points during the dissolution of Tablet B

Tablet C

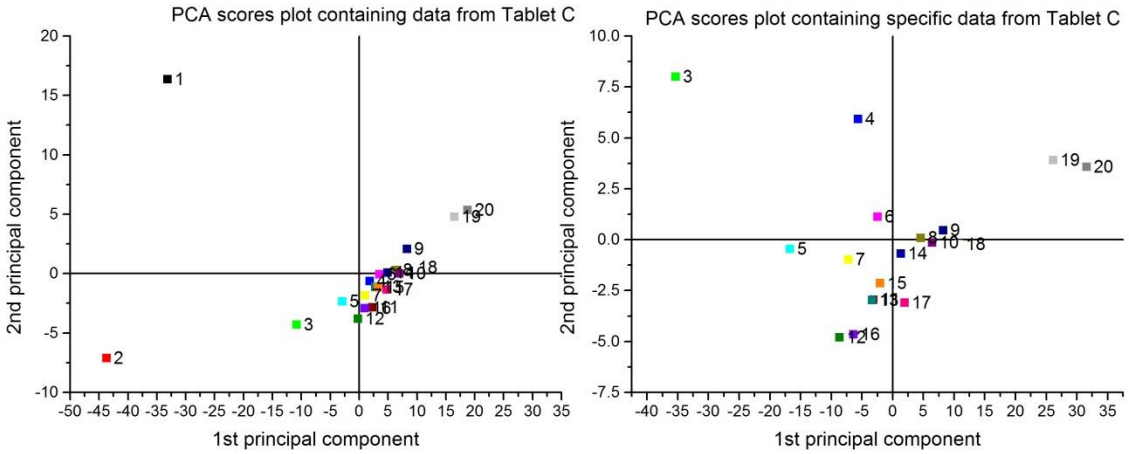


Figure A2-23 The scores plots from the PCA on different time points during the dissolution of Tablet C

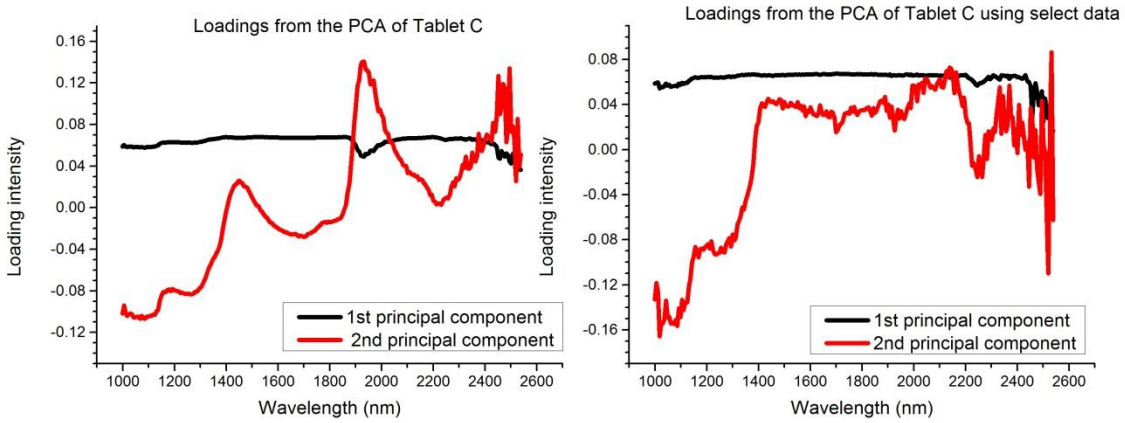


Figure A2-24 The loadings plots from the PCA on different time points during the dissolution of Tablet C

A2.5 UV data

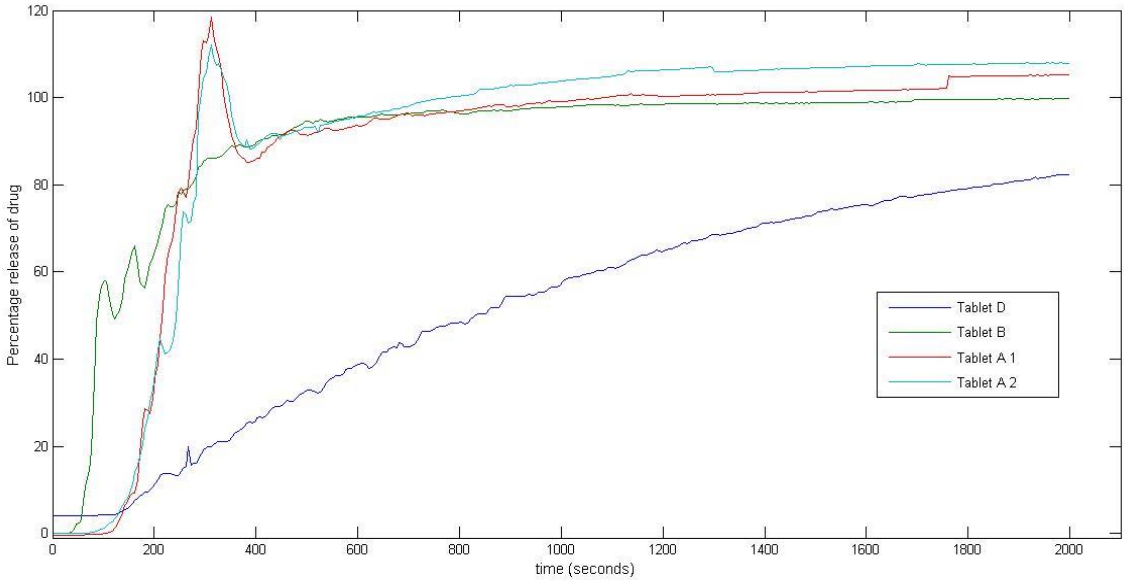


Figure A 2-25 Release profiles of the non-caffeinated paracetamol formulations determined using a USP type IV dissolution apparatus.

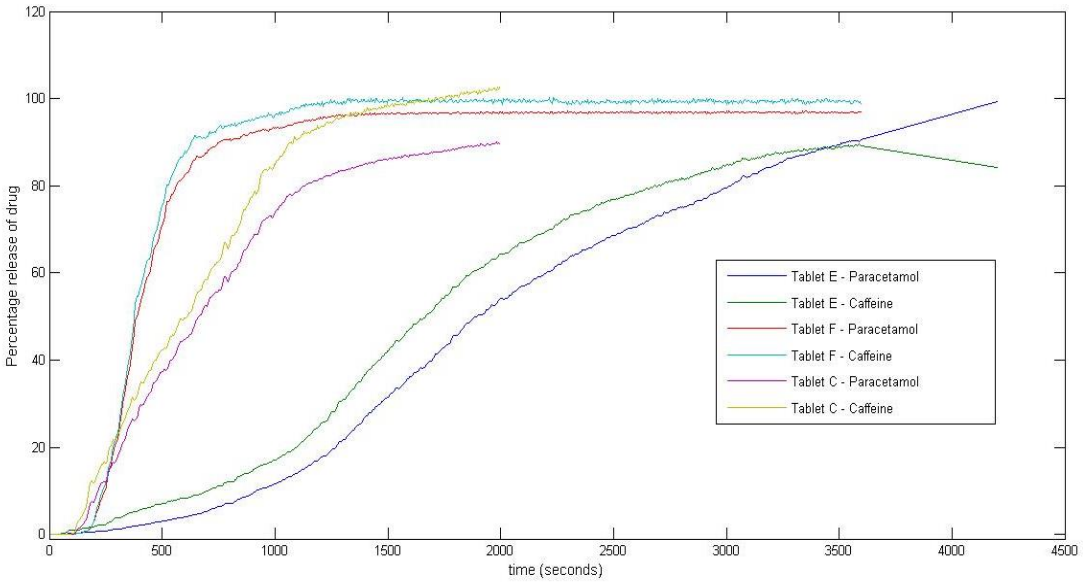


Figure A2-26 Release profiles of the caffeinated paracetamol formulations determined using a USP type IV dissolution apparatus.

Appendix 3: Mathematical functions: Dissolution corrections functions and the Data correction methods from Chapter 5

A3.1 Simultaneous equations to calculate the concentration of two compounds which both contribute to the absorbance of a solution in UV spectrometry

A = Absorbance

C = concentration

L = path length

ε = molar absorbance coefficient

$$A = \varepsilon cl$$

As the path length is unchanging throughout the experiment a new term *A'* is created by the form of:

$$\frac{A}{l} = A'$$

The molar absorbance coefficient can be calculated for the different absorbing compounds by using standards of known concentration. By using this value of the molar absorbance coefficient it is possible to calculate the absorbance value in the solution. In a two sample system where two separate absorbing species exist then the absorbance can be seen as being equal to:

$$A_t = A_1 + A_2$$

$$A'_t = \varepsilon_1 \times c_1 + \varepsilon_2 \times c_2$$

If we rearrange this further then the concentration of one species can be determined as

$$C_1 = \frac{A_t'}{\varepsilon_1} - \frac{\varepsilon_2 C_2}{\varepsilon_1}$$

By using two sets of wavelengths then it is possible to create a value for C_2 which can be solved to generate a value for C_1.
If the two wavelengths are called x and y then:

$$C_1 = \frac{A'_{tx}}{\varepsilon_{1x}} - \frac{\varepsilon_{2x} C_2}{\varepsilon_{1x}}$$

$$C_2 = \frac{A'_{ty}}{\varepsilon_{2y}} - \frac{\varepsilon_{1y} C_1}{\varepsilon_{2y}}$$

$$C_{1x} = \frac{A'_{tx}}{\varepsilon_{1x}} - \frac{\varepsilon_{2x} C_{2x}}{\varepsilon_{1x}} \times \left(\frac{A'_{ty}}{\varepsilon_{2y}} - \frac{\varepsilon_{1y} C_1}{\varepsilon_{2y}} \right)$$

By expansion of the terms and rearrangement the equation can be expressed as:

$$C_1 \left(1 - \frac{\varepsilon_{2x}}{\varepsilon_{1x}} \times \frac{\varepsilon_{1y}}{\varepsilon_{2y}} \right) = \frac{A'_{tx}}{\varepsilon_{1x}} - \frac{\varepsilon_{2x}}{\varepsilon_{1x}} \times \frac{A'_{ty}}{\varepsilon_{2y}}$$

All of these terms can be discerned during the UV dissolution testing and once the concentration of C_1 is calculated it can be inputted back into the earlier equations to calculate C_2.

A3.2 Calculating and correcting for the signal attenuation present in the data

On closer section of the results from 5.3.2 it is clear that some physical effect is altering the spectral data which is creating difficulties with the analysis of the results. By following the motion of the acid solution through the cell and over the tablets in the initial stages of the experiment it is clear that the acid solution is causing some kind of filtering effect. The spectrum is seen to change as soon as the acid is between the tablet and the camera, such that the photons travelling through the acid are undergoing some form of filtering effect.

The spectra of the tablet at two distinct, different times are plotted in Figure A3-1. The initial spectrum before the tablet has been submerged is labelled as the dry data and the

spectrum after the tablet has been submerged is called the wet data. Figure A3-1 shows that the two signals are quite dissimilar with a great deal of the spectral information being lost after submersion in the dissolution liquid. The major goal of this work is to follow the breakdown and release of active pharmaceutical ingredients (APIs) and excipients from the tablets over time. The results of Figure A3-1 make it clear that some form of correction needs to be applied to the data to account for the effect of the acid solution before any further analysis of the release of excipients can be performed.

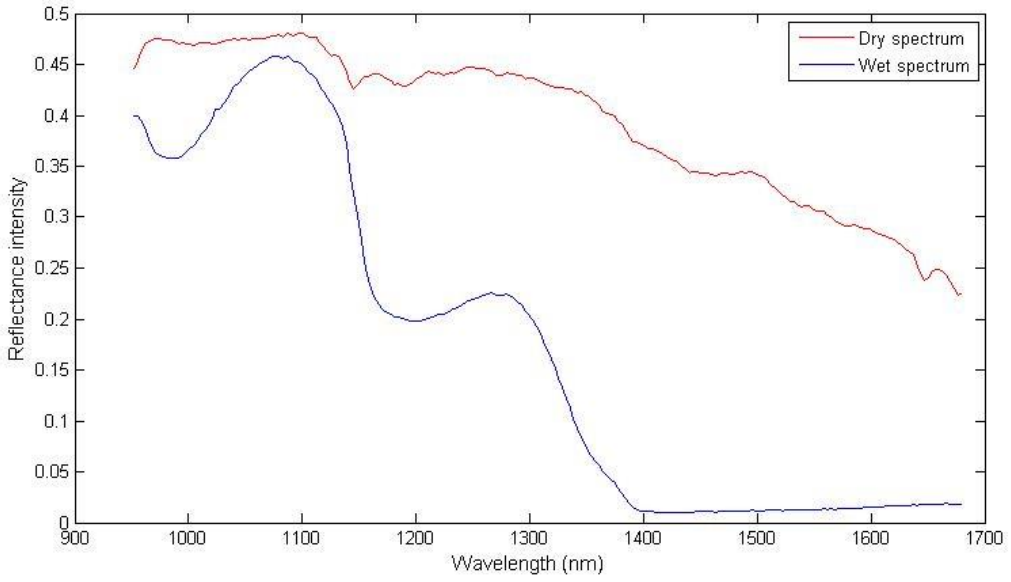


Figure A3-1 A plot of the data contained both before and after tablet submersion. The tablet used for these spectra was Tablet B

Subtraction of background signal

The results shown in Figure A3-1 imply that the spectral profile of the target is being modified in some way by the presence of the acidic media which the photons are passing through. This effect needs to be corrected using some form of manipulation of the data; the first such correction used was the simplest, subtracting the spectrum of the acidic medium from the tablet data. If the effect of the acid is additive then it can be considered as two separate overlapping spectral signals which generate a single spectrum with intensity equal to the sum of the component spectra. If this holds true then the subtracting the acid spectrum should reveal the original tablet spectra.

Every dissolution experiment contains a series of pre-experiment images which were created to allow for the correction of the data. One of these images is the acid flowing through the empty flow cell and it is this data that is used for the correction and an example

of the acid spectrum can be seen in the appendix. For each correction a unique acid spectrum was calculated to account for any changes which may have occurred between images. There should however, be no change over time within the image as there is nothing in the system other than the acid, unlike in the dissolution when the acid contains dissolved chemicals.

During the dissolution process the tablet undergoes complex spectral changes which are a result of the loss of both API and excipients from within the tablet which would add difficulty when trying to determine the validity of the corrections. This is not a problem with the pure tablet APIs or excipients as these components should not undergo any change when submerged in the acid solution over than to dissolve. Due to this the correction was run on the tablet component spectra initially; if the correction is not successful with these spectra then there is little advantage in using the correction on the tablet data. There are a large number of excipients in each tablet and none of the tablets contain exactly the same excipients. Therefore the excipients used for these experiments were chosen as the ones which were used by the largest number of the tablet brands and were readily available. A total of three excipients and the two APIs have been imaged for classification of the tablet data: Acetaminophen, Caffeine, Starch, Polyvinylpyrrolidone (PVP) and Magnesium stearate. Two distinct spectral regions of PVP were seen in the data, one with a much higher intensity to the spectral response than the other. These regions were classified as being either PVPL (low intensity) or PVPH (PVP high intensity). The data was then corrected by subtracting the acid spectrum and the results were plotted, alongside the original data, which is shown in Figure A3-2. The spectra of the excipients before any corrections were applied are shown in Figure A3-2(A) with the corrected excipient spectra in Figure A3-2(B). The majority of the excipient data shows very little spectral change other than a decrease in intensity, there are no new peaks formed and no correction of the data. The only exception is the PVPL data which became virtually flat; this implies that either the acid data in the flow cell was incorrectly assigned as being PVP, or that the signal of the PVP is too weak and only the acid spectra can be seen.

These results clearly detail that a simple subtraction of the acid spectrum to remove any additive issues is not sufficient. Without the ability to correct for the change in the spectral signal caused by the acid solution it may not possible to detect the change in tablets as the APIs is released. A more complex system needs to be derived and tested to discover if this can allow the original data can be re-calculated.

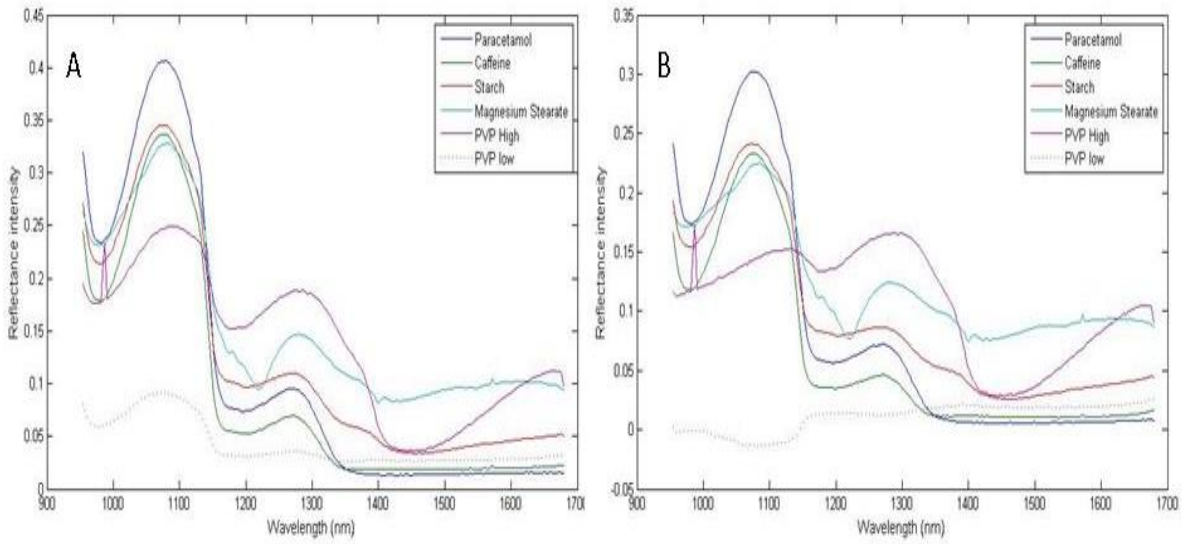


Figure A3-2 The average spectra of excipients before (A) and after (B) acid subtraction

The Transfer Function

The transfer function is a mathematical approach to correcting the spectrum for the effects of the acid signal. The idea is that there is a relatively simple mathematical change occurring to the signal which is changing it from the initial spectrum to the altered one. This idea is expressed as a schematic in Figure A3-3 and if this holds true then it should be a simple process of reversing the mathematical process to get to the correct data.

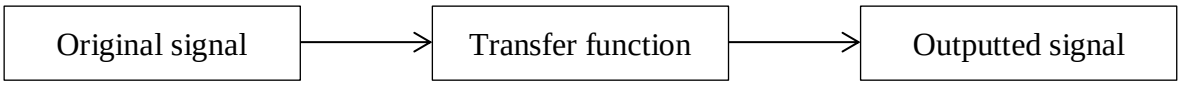


Figure A3-3 A schematic view of how the transfer function works with the data

The change has been shown to not be a simple addition of a secondary spectrum so the next approach was to test is the change is multiplicative, that is that the change is from the original spectrum multiplied by a filtering factor which would be unique for each wavelength. As the original signal and the outputted signal are readily available in the images then it should be a simple process to calculate a transfer function. If the original signal is being multiplied by the transfer function to generate the outputted signal then it is a simple matter to rearrange this equation to calculate a transfer function.

The transfer function was generated using data from over 500 pixels at both dry (original) and wet (outputted) regions of the tablet spectra. The wet spectrum refers to any spectral data after the target is submerged in the acid solution and the dry is tablet data from before the tablet is submerged. The dry and wet spectral data used to generate the transfer

function must come from the same image because intra image variation could cause the transfer function to be incorrectly calculated. When generating the transfer function the points of wet data were from immediately after the wetting of the tablet to minimise any physical changes which may have occurred to the tablet.

$$\text{Transfer function} = \frac{\text{Wet Spectra}}{\text{Dry Spectra}}$$

While the data used to calculate the transfer function was from 500 pixels a single mean spectrum was created for use in these equations. This mean spectrum is generated from a large data set so should be a good indication of the average for entire tablet. The pixel to pixel variations in the data are minimised in this way.

A transfer function was calculated for Tablet B and then tested on the rest of the image to try and correct the 'wet' spectra back to their original, unaffected spectra. The first data recreated was from a time immediately after the wet data used in the calculations and the corrected dry data was a good match for the original pre submersion spectrum. The next test was to apply the correction to data much later on in the dissolution process to see if the change holds. The times chosen were 1 minute after the initial submersion and the last minute of the dissolution as these represent both a small and large change to the spectra. The data after 1 minute has not changed significantly to produce the characteristic red colouring in the HSI image but by the end of the experiment the spectrum is completely changed. The results of this test are shown in Figure A3-4 and it is clear that the transfer function is not an accurate tool for correcting the full data set. The data from shortly after the dissolution of the tablet began is still closely related to the original dry spectrum, while the wet data collected from the ending stages of the experiment displays a significant difference in spectral responses. The data from the ending stages represents the tablet once it has broken down fully and contains only the spectrum of the 'red' colour in the false RGB images. This implies that the transfer function cannot account for these changes in the spectra, especially in the region after 1300 nm.

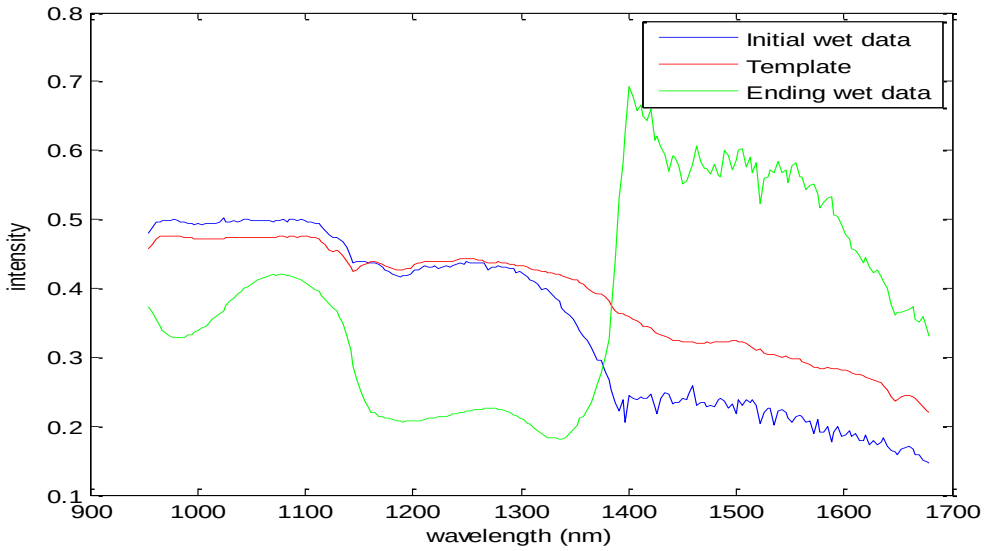


Figure A3-4 A graph showing three separate dry spectra from Tablet B: the actual dry spectra, one calculated from the initial portion of the dissolution experiment and dry data calculated from the later stages of the dissolution.

There are two main possibilities for this extreme variation between the data; the first is that the release of the APIs and excipients from the tablet matrix is changing the spectra of the tablet. While this spectral change should be best seen as a shift in peak position or a loss in peak intensity this is not what is shown by the transfer function. The second possibility is that as the tablet expands and then begins to break down there is a notable change in the path length of the photons through the acid solution. This increased path length is having a pronounced effect on the spectrum of the tablet.

The second explanation for these results appears the most likely as the loss of tablet should not cause such a severe and drastic change in the effectiveness of the transfer function. The loss of any one excipient or API would cause a change at key locations throughout the tablet spectrum but not such a severe change through the whole spectrum. If the path length is the problem however, it is quite possible that the whole spectrum could be altered by this change which agrees with the results of the transfer function.

While the second explanation appears the most likely it is worth noting that that the change to the data is not a single finite mathematical operator which is the assumption made in the previous analysis. As the dissolution progresses the correction is becoming less accurate, this indicates that there is at least one more term effecting the correction and that without knowledge of the terms and a system to empirically calculate these new factors then it would be impossible to correct the data.

Calculating the effect of the path length on signal attenuation

These results demonstrate that the path length is having an effect on the attenuation of the signal. The next step was to attempt to calculate how the path length is affecting signal in the form of a simple algebraic equation. Four different wavelengths were chosen and the intensity of these wavelengths at the different height increments was plotted as a function of the path length. If a best fit of these different intensities show a linear gradient then the path length is affecting the attenuation in the form of:

$$y = cx_i + d$$

In this equation the signal intensity (y) is the intensity of a given pixel at a given wavelength and the *x* value relates to the *i*th path length of the photons through the acidic dissolution media. The *d* value is a constant which is used to account for the changes in the signal intensity due to the photons having to pass through the cell wall of the flow cell; the *c* value is also a constant and is the correction factor which can give the value of *y*. The value of *c* must be between 0 and 1 which is also true of *y*.

These linear plots can be found in the appendix and they show that the data points do not trend towards a linear relationship in this data but rather an exponential one. Such results clearly indicate that the effect of the path length is not a simple linear multiplicative as explained above but instead follows some kind of power law or possibly something more complex. The log of the data was then plotted against the path length; if this data is linear then the effect of the path length is most likely given by:

$$y_i = c^{x_i} + d$$

These results are displayed in Figure A3-5 and show a substantially better fit to a linear fit than the previous data. This illustrates that the path length is causing the attenuation effect on the spectral signal from the tablet data and that the degree of that attenuation is raised to the power of the path length. The greater the path length the more severe the attenuation which agrees with all of the results from earlier, but the results must first be tested to ascertain their validity. This was done by creating a model of the data which can accurately recreate the real values of *y* using the modelled values of c and d.

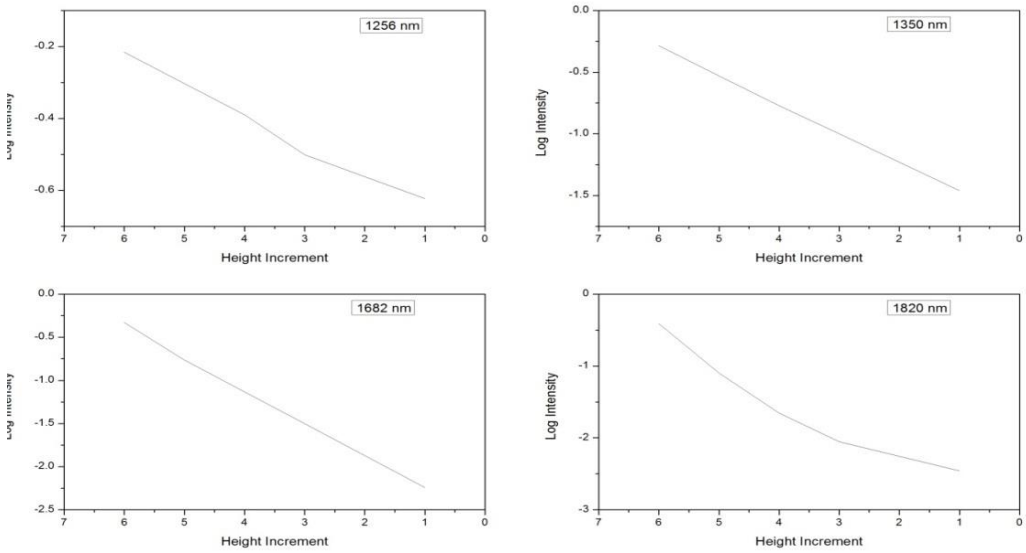


Figure A3-5 Plots of logy versus path length

A3.2.1 **Modelling the attenuation of the spectral signal**

Models were created to show a correlation between the spectral response recorded *y*, the path length x_i and a constant *c* which corresponds to the original spectral signal in a manner according to the proposed path length equation shown below. Two different models were used to look into the data and to try and determine a set of values which could be used to show that the equation below holds true. The values of x_i where chosen to be 1, 2, 3, 4, 5, 6 as this represented the identical change in the height of the reflectance tile. The values of *c* and *d* must be within the bounds of $0 < x > 1$ as the value of y must also be within these bounds as it is in units of reflectance.

$$y_i = c^{x_i} + d$$

The Rearrangement Model

The Rearrangement Model uses simple algebraic re-arrangement of the equation above to calculate the values of *c* and *d* which is shown below. By using three different adjoining height data sets then it is possible to generate the equality at the bottom of the equations. This shows that by simply using the known values of *y* a value of *c* can be calculated and this can subsequently be used to calculate the value of d by a basic rearrangement of the initial equation.

$$y_1 = c^{x_1} + d$$
$$y_2 = c^{x_2} + d$$
$$y_3 = c^{x_3} + d$$

Where $x_1 = 1$, $x_2 = 2$ …

$$y_1\text{-}y_2 = c^1 + d\text{-}c^2\text{-}d$$
$$y_1\text{-}y_2 = c(1\text{-}c)$$

$$\frac{y_1\text{-}y_2}{y_2\text{-}y_3} = \frac{c(1\text{-}c)}{c^2(1\text{-}c)} = \frac{1}{c}$$

This model was able to quickly compute the value for *c* however, the problem with the second height level data severely limited the number of calculation which could be used to generate *c*. The model was run using data from three separate wavelengths to increase the number of results and to better show the accuracy of this model.

The results from the three wavelengths are shown in Table A3-1 and detail both possible values of *c* generated by this model; the mean value is also shown with the standard deviation in brackets. The values for *c* are mostly within the $0 < c > 1$ bounds however, they have substantial percentage variations between samples. The values for 1255.77 nm contain a standard deviation of 14.9% which is too high a degree of variation for the model to be seen as accurate in its current form.

Wavelength(nm)	**1255.77**	**1349.83**	**1681.68**
Y1-Y2/Y2-Y3	1.230	1.803	3.057
Y2-Y3/Y3-Y4	0.9776	1.817	2.320
C	0.8127	0.5547	0.3271
	1.022	0.5503	0.4311

Mean C value	0.9178 (0.1486)	0.5525 (0.0031)	0.3791 (0.0735)

Table A3-1 Results from the Rearrangement Model

The variation this model has shown is due to the lack of an error term, this means that any slight variation within the imaging becomes transferred into the model. A good example of this problem is that one of the values of C is slightly above 1 which is not within the bounds of true reflectance values but is within experimental detector error of 5%.

The Iterative Model

The second model was named the Iterative Model[1] and uses a very different approach to calculating the values of *c* and *d*. The first step to this model is again to rearrange the proposed path length equation so that the log of the equation can be used to calculate the value of *c* for a given value of x:

$$\log(y\text{-}d) = \log c^{x}$$
$$\log(y\text{-}d) = x\log c$$

When plotted the values from this equation should generate a simple linear progression with gradient *logc*, given by the standard linear equation of $y = mx+c$. By calculating a value for *d* which minimises the variation between the log(y-d) and the linear regression fit of the data then the value of *c* can be calculated.

The model was created using MATLAB and follows a Newton-Raphson process of iteratively changing a value to get the best fit for the data possible. The model first selects a semi random value of d (d is within the limits of $0 < d > 1$) and *log(y-d)* is calculated, once this is complete a linear fit is applied to the data and residuals calculated. The residual is defined for this work as being the modulus of the difference between the value of *log(y-d)* and the fit of the data. A smaller residual will mean that the linear fit is closer to the real data than it would be if the residual was larger and this is the ideal result. Once the residual has been calculated the value of *d* is changed slightly, the process repeated and the residuals compared. The model continues this iterative process of constantly changing the value of *d* until a threshold number of iterations have been run with no change to the minimum residual achieved. Once this occurs the value of *d* used to generate the minimum value of the residual is output along with the corresponding value of *c*.

The model was tested for accuracy and the optimum number of iterative cycles before completion to obtain the ideal balance between processing time and accuracy of the data. It was found that the model worked very well using a relatively small number of cycles and it was decided that 1000 was a valid cycle number to use. If the required number of iterative cycles is much higher than 1000 then this can cause the processing time to increasing dramatically so that to calculate the value of *c* and *d* for each wavelength can take hours. It was also discovered that much below this threshold number of cycles and the outputted values of interest could be inaccurate and showed a large degree of variation between runs.

The model was run using data from the 1249 nm region of the hyperspectral image and a plot of these data points as a function of path length is shown in the appendix. This data was input into the Iterative Model and run with until the fit of the data had not been improved in 1000 iterations. Once the necessary data had been output the plot of log(y-d) versus the path length which represented the best possible fit of the data was plotted; the linear regression fit was also plotted onto the graph to show how well the data is fit. This plot can be seen in Figure A3-6A and it is clear that there is very little difference between the residual values and the original log(y-d) data. This very close fit means that the value for *c* is very accurate as the fit is good and that the values can be trusted.

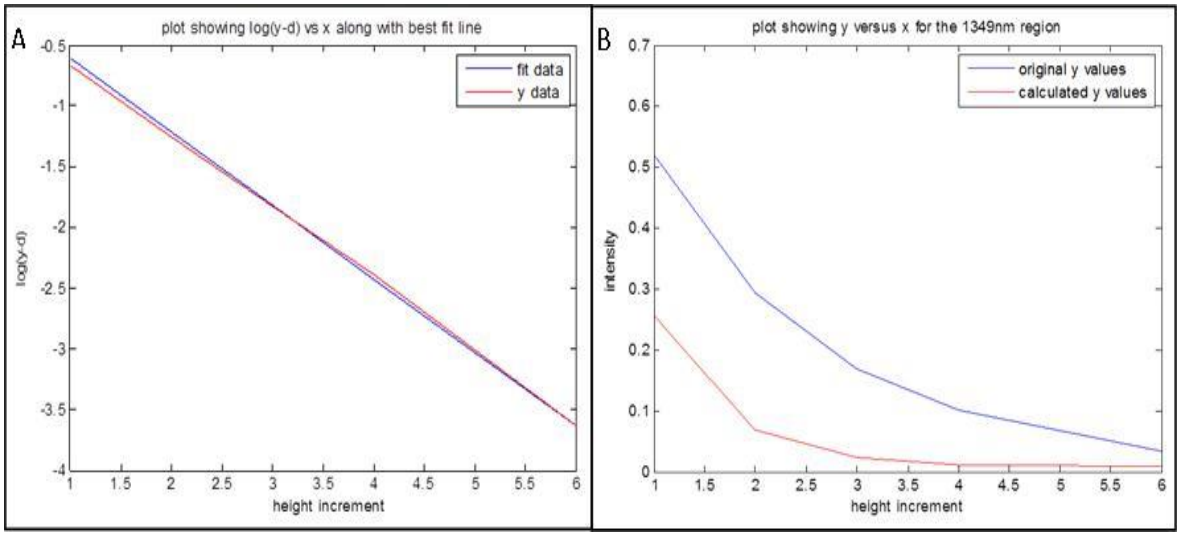


Figure A3-6 A) a plot which details the accuracy and closeness of fit for the values calculated in with the iterative Model and B) Calculated and real values for y at 1349 nm

A final test was to use these values of *c* and *d* to generate a new set of *y* values and to plot these against the data real values of y. The result from this test can be seen in Figure A3-6B with the red line showing the calculated values of *y* and the blue line are the real

values of y. The intensity of these calculated values is constantly below that of the real values but they do follow the same trend perfectly. These results imply that there may be a co-efficient to the c^x term which would scale the results; this is not unexpected as the modelled path length value does not in its current form account for the path length through the cell wall.

This change to the path length to account for the absorbance and reflectance from the glass wall of the flow cell should be a constant. This is because the glass remains unchanged during the dissolution and its reflectance properties would also be unaffected. By using this knowledge a new form of the path length equation can be proposed in which *A* is used to account for the path length changes due to the cell wall and the equation is:

$$y_i = c^{x_i + A} + d$$

The Iterative Model is able to generate the values of *c* and *d* with a high degree of repeatability and accuracy; the testing of the model has shown that a new value of *y* can be calculated which follows closely the trend of the data but with an intensity lower than the original.

This result is not unexpected as it is showing that the reflectance of the sample can be calculated by considering the path length of the light through an absorbing medium and a factor which is related to the absorbance of a signal at a set wavelength with no interference. This equation is very closely related to the Beer-Lambert law which states that:

$$R = \frac{I}{I_0} = 10^{-\alpha l}$$

Where *R* is the reflectance of the system which is a function of the measured light intensity divided by the total light intensity, *l* is the path length of the photons through the absorbing medium and α is the absorption coefficient. The beer lambert law is the same equation with the *c* term being related to α.

1. In consultation with Dr Greenhalgh & Dr Gray, Department of Maths and Statistics, University of Glasgow

Appendix 4: MATLAB code

The code used to manipulate the hyperspectral data is different between chapters and is tailored to the specifics of the images and the wavelength ranges used. Shown here are examples of the most common coding templates that were changed for the specific needs in the different chapters.

A4.1 Code for opening HSI images and creating a false colour image with the data

```
clc, clear all, close all

%Code to open image files
%start by moving the active MATLAB window to the same location as the
file
%to open
location = 'E:\PhD\Chapter 6\Sorted Good Data\Linescan Videos of
Panadol\PaEx Video and Corrections';
cd(location)

%set up the filename and header for reading
Filename = 'PaEx_refl';

InputFile = [Filename '.raw'];
Inputhdr = [Filename '.hdr'];

%Read the file header for the bands, lines and samples

fid = fopen(Inputhdr,'rt');
out = textscan(fid,'%s = %f', 'headerlines', 8);
fclose(fid);
ind = strmatch('samples',out{1});
S = out{2}(ind) ;
ind = strmatch('lines',out{1});
L = out{2}(ind) ;
ind = strmatch('bands',out{1});
B = out{2}(ind) ;

%open the Image, subset the image to make the file a manageable size
Data = multibandread(InputFile, [L S B], 'float', 0, 'bil', 'ieee-le',
{'row', 'range' [1 1 2000]}, {'column', 'range', [1 1 S]}, {'band',
'direct', [10 1 B]} );
%Clear unwanted data for memory
clear out fid ind L S B F
%%
% Create a false coloured RGB image

%Make the matrix Y which contains the three bands used for the image and
%create matrix of size of data
Y = Data(1:1000,:,[71 111 173]);
[rows cols bands] = size(Y);

%Take any values which are above a threshold and turn pixel data to 0.
This
%is to remove erroneous values from the edge of the air bubbles
for a= 1:rows;
```

```
    for b = 1:cols;

        for c = 1:bands;

            if(Y(a,b,c) >= 1.1);
                Y(a,b,:) = 0;
            end

        end

    end
end

%Image must have values between 0 and 1. Use the min and max functions to
%normalise the values to between 0 and 1
Max = max(max(Y));
Min = min(min(Y));

ImData = zeros(rows, cols, 3);

for i = 1:3

    Y(:,:,i) = Y(:,:,i)- Min(i);
    ImData(:,:,i) = Y(:,:,i)./(Max(i)-Min(i));

end

%output image
image(ImData)
```

A4.2 Code for determining the degree of tablet expansion within the flow cell

```
% The areas of the individual flow cell repetitions must be identified,
this
% is done by finding the edges in the white tile to the left of the image

% The WhiteData is a strip of data which disects the location of the
% refrence claibration tile in the image. Using this data the location of
% the flow cells can be determined
%
X = find(WhiteData(:,4) <= 0.6);

[rows cols] = size(X);

%When there is a gap of 2 or more pixels which arnt white then this is an
%edge, these edges position are stored as pixel locations in the matrix Q

for i = 1:rows-1
    if (X(i+1)-X(i)) >= 2;
        Q(i) = X(i);
        Q(i+1) = X(i+1);
    end

end
```

```
%find the locations of the Q values and make a matrix of these values
YY = find(Q >= 5);

Loc = X(YY);

%clear X Y YY WhiteData

%%
% Next run the analysis on the flow cells split into sections
% Use the locations generated to calculate the location of each flow cell
% and then simply find the number of pixels within the flow cell which
have
% a value = black region

[rows cols] = size(Loc);

for i = 1:rows-1

    if Data(Loc(i)+1, 8) <= 0.6;
        i = i+1;
    end

    %create a redcuded size cell
    Cell = Data(Loc(i):Loc(i+1),:,:);
    %imshow(Cell)
    Cell = Cell(:,115:245);

    %Now use the size of the cell to caluclate a total pixel count
    [rowsC colsC] = size(Cell);
    TotalPixels = rowsC*colsC;

  Totalswell = find(Cell(:,:) >= 0.18);
   [swell misc] = size(Totalswell);

 %calculate a % swelling
 PercentCoverage(i) = (swell/TotalPixels)*100;


  End
```

A4.3 Creation of mesh plots from the images

```
%Input a range of locations which are then used to create mesh plots

%Open specific locations containing the wavelength data, this is used
%during the mesh plot function
location = 'E:\Sorted Good Data\';
cd(location)

load SWIR_Wavelengths
W = SWIR_wavelengths(10:256);

%Choose the lines for creating a mesh plot
X = [45, 115, 235, 1080];

%Use a for loop to select each location in turn and then create a
%mesh plot of the data.

for i = 1:4;
```

```
    Y = X(i);
    figure, mesh(W, 0:120, squeeze(Data(Y, 140:260,:)))
    xlabel('Wavelength (nm)','fontsize', 13)
    set(gca, 'fontsize', 11)
    zlabel('Reflectance', 'fontsize', 13)
    set(gca, 'Ztick', 0:0.1:1, 'fontsize', 11)
    ylabel('Pixel location', 'fontsize', 13)
    set(gca, 'YTick', 0:20:120, 'fontsize', 11)
end
```