



## Quantitative polymorph contaminant analysis in tablets using Raman and near infra-red spectroscopies

Title	Quantitative polymorph contaminant analysis in tablets using Raman and near infra-red spectroscopies
Author(s)	Ryder, Alan G.;Hennigan, MC
Publication Date	2013
Repository DOI	<a href="https://doi.org/10.1016/j.jpba.2012.10.002">10.1016/j.jpba.2012.10.002</a>

1 This is the final author version (revised & corrected), the definitive version is the published  
2 document on the JPBA website.

3  
4 **QUANTITATIVE POLYMORPH CONTAMINANT ANALYSIS IN TABLETS USING RAMAN**  
5 **AND NEAR INFRA-RED SPECTROSCOPIES.**

6  
7 Michelle C. Hennigan and Alan G. Ryder.\*

8  
9  
10 Nanoscale Biophotonics Laboratory, School of Chemistry, National University of Ireland Galway,  
11 University Road, Galway, Ireland.

12  
13  
14 \* To whom all correspondence should be addressed.

15 **Tel:** +353 91 49 2943 **Fax:** +353 91 49 4596 **Email:** alan.ryder@nuigalway.ie

16 **Postal Address:** Nanoscale Biophotonics Laboratory, School of Chemistry, National University  
17 of Ireland Galway, University Road, Galway, Ireland.

18  
19  
20 **ABSTRACT**

21 The detection and quantification of alternate polymorphs of Active Pharmaceutical Ingredients  
22 (API), particularly at low concentrations is a key issue for the manufacture and analysis of solid  
23 state formulations. Each polymorph can possess unique physical and chemical properties which in  
24 turn can directly affect factors such as solubility and bioavailability. Near Infra-red (NIR) and  
25 Raman spectroscopies can be used for the rapid characterisation and quantification of polymorphs  
26 in solid samples. In this study we have generated a model tablet system with two excipients and a  
27 10% API concentration, where the API is a mixture of the FII and FIII polymorphs of Piracetam.  
28 Using Transmission Raman Spectroscopy (TRS) and NIR spectroscopy it was possible to detect  
29 FII polymorph contamination in these model tablets with Limits of Detection (LODs) of 0.6 and  
30 0.7% respectively with respect to the total tablet weight (or ~6-7% of the API content). The TRS  
31 method is the superior method because of the speed of analysis (~6 seconds per sample), better  
32 sampling statistics, and because the sharper, more resolved bands in the Raman spectra allowed for  
33 better interpretation of the spectral data. In addition the TRS used here provides facile access to  
34 the low frequency wavenumber region for analysis of solid-state lattice modes.

35  
36 **Keywords:** Polymorph, Contamination, Spectroscopy, Raman, Near Infra-red, Chemometrics.

## 1. Introduction

Tablets are one of the most popular pharmaceutical solid dosage forms in production today. The knowledge of which solid state form of the Active Pharmaceutical Ingredient (API) present is critical from production line of the product to its therapeutic response in the patient's body. Solid state forms include polymorphs, amorphous solids, salts, solvates, co-crystals and all can have different physicochemical properties due to the variations in their free energies and inter- and intra molecular bonding.[1] Properties such as solubility, dissolution rates, and bioavailability of the API can be affected by a change in solid state form and thus accurate characterisation and quantification of the solid state present and/or of any changes in solid state form is of great importance. Incidences of unwanted polymorphism (which can arise from the ability of a molecule to crystallise in different crystal structures) in pharmaceutical products have been widely reported and can have major ramifications.[2-4] From a regulatory standpoint, drug products must be manufactured in a reproducible and validated process compliant with legal requirements.

The implementation of real-time measurements of pharmaceutical processes and there is a need for rapid, non-invasive and non-destructive technologies to achieve this. High performance liquid chromatography (HPLC), the traditional means of quantitative analysis of pharmaceutical solids is destructive, can be time consuming as regards sample preparation and dissolution, and crucially provides no information pertaining to polymorphic content and the solid state form of the API. Powder X-ray Diffraction (PXRD) is considered the definitive method of differentiation between polymorphs and solid state forms as a difference in crystal structure signifies the presence of an additional solid form. It has been used successfully for the quantification of polymorphic mixtures, [5, 6] however it can be time consuming (compared to spectroscopic methods) as samples must be generally analysed off-line. Another issue with PXRD is the sensitivity to effects such as preferred orientation where the particles in the sample may become arranged such that certain faces of the crystals/fine powder are better presented for analysis resulting in diffraction patterns that are not truly representative of the entire crystalline material. This effect can be reduced by sample grinding to reduce particle size (but only if grinding causes no solid state changes) or sometimes by rotation of the sample during measurement. In any event this introduces another sample handling step in the process which is often impractical for in-situ tablet analysis.

Vibrational spectroscopies such as Near Infra-Red (NIR) and Raman are potentially rapid methods that can be used for the non-contact, non-destructive, *in-situ*, characterisation and quantification of drug substances. NIR spectroscopy measures overtones and combination modes of fundamental molecular vibrations that occur in the 700-2500 nm region.[7] The main positive attributes of NIR are that often little to no sample preparation is necessary, the use of NIR radiation enables the use of relatively inexpensive optics, standard fibre optics enables remote sampling, and overall the instrumentation is relatively robust and often inexpensive. Spectral interpretation is however not straight forward due to the broad, overlapping nature of the bands and thus

1 chemometric analysis is usually required. Nevertheless this method is widely used with success in  
2 a wide variety of applications such as reaction monitoring, quality control and quantification.[8, 9]

3 The application of Raman spectroscopy to pharmaceutical manufacturing has grown  
4 substantially in the past two decades with a wide variety of uses including quantification of  
5 polymorphic mixtures, tablets, and capsules.[10-13] Many APIs are ideal for study by Raman as  
6 they are typically strong Raman scatters due to the presence of aromatic functional groups with  
7 symmetric vibrational modes.[10] Raman, like NIR, allows for minimal if any sample preparation  
8 and *in-situ* analysis, however, the sharper and more resolved spectral bands allow for more facile  
9 spectral interpretation which is very useful in the context of contaminant analysis. One of the main  
10 drawbacks with Raman spectroscopy is the issue of fluorescence in some samples, which can  
11 obscure useful spectral information by swamping the spectrum with the stronger, broader  
12 fluorescence signal. This effect can be mitigated or eliminated (depending on the intensity of the  
13 fluorescence signal) via a number of instrumental routes, but the use of NIR excitation [14] which  
14 is by far the most common and least expensive instrumental method at e present.

15 The two most common sampling geometries employed in the Raman spectroscopy of solid  
16 materials are Backscattering and Transmission. In Backscattering Raman (BR) the excitation light  
17 is delivered to the sample surface and the Raman signal is collected along the same optical  
18 pathway, *i.e.* at an angle of  $180^\circ$  to the incident beam. Therefore for BR the sampling volume is  
19 located primarily at the surface of the sample and is determined by a combination of opacity and  
20 scattering potential of the material. In Transmission Raman Spectroscopy (TRS) a collimated laser  
21 beam is incident on one side of the sample and the Raman signal is collected from the opposite  
22 side after passing through the sample *i.e.* at an angle of  $0^\circ$  to the incident beam.[15] The critical  
23 differences between these two geometries are the sampling volume and location. In BR, the spot  
24 sizes tend to be small (depending on the optics, but typically  $< 100 \mu\text{m}$  for microscope based  
25 systems) and the Raman signal generated is typically representative of the surface layer and the  
26 very immediate layers just below the surface where the excitation light has penetrated.[16] This  
27 sub-sampling is a limitation of many BR based systems; however, this can be mitigated somewhat  
28 by the use of mapping. Unfortunately mapping generally does not solve the penetration depth  
29 issue and is also accompanied by a severe time penalty. A further issue is the fact that BR efficacy  
30 can be degraded if there is a surface coating or film on the tablet surface leading to a composite  
31 signal of film and tablet. It depends on the thickness and composition of the film as to how much  
32 of the tablet signal is obtained. The best option in BR will tend to be the use of large sample spot  
33 sizes using dedicated accessories such as the PhAT probe (Kaiser Optical Systems Inc).[17] In  
34 TRS the Raman signal is more representative of the entire tablet sample as Raman photons can be  
35 generated at any point where the collimated laser beam passes through the complete sample.[15]  
36 In TRS one typically can use larger spot sizes ( $\sim 8 \text{ mm}$  in this case) and thus one samples a much  
37 greater proportion of the tablet, which reduces the issue of sub-sampling leading to more  
38 representative spectra. The popularity of TRS has increased significantly in the past decade,  
39 particularly for pharmaceutical analysis such as the quantification of APIs and excipients in

1 pharmaceutical tablets and capsules and analysis of polymorphs in pharmaceutical  
2 formulations.[12, 18-21] Johansson *et al.* compared the efficacy of the BRS and TRS methods  
3 using tablets and they observed a significant improvement in relative RMSE from 2.9%(BRS) to  
4 2.2% (TRS) for API content prediction which is largely attributable to the better sampling statistics  
5 [12]. Hargreaves *et al.* have utilised TRS for quantitative analysis of pharmaceutical capsules  
6 achieving a relative RMSE for prediction of API concentration of 1.5% [22]. They found the  
7 technique to be insensitive to capsule fill weights which made the method more robust. TRS has  
8 also been used for the quantification of analysis of ranitidine hydrochloride polymorphs in tablets  
9 and capsules, with reported RMSECVs of 2.40 and 4.80% respectively which in the case of the  
10 tablets was an improvement on a previous NIR study (RMSECV of 2.89 % w/w reported). [23, 24]

11 Piracetam (*2-oxo-1-pyrrolidine acetamide*) is a nootropic drug which has five reported  
12 polymorphs, and at least two hydrated forms (a mono- and a dihydrate).[25-27] Two forms, FIV  
13 and FV are generated under high pressure while FI is generated when either FII or FIII are heated  
14 to 127 °C (below the melting points of both polymorphs) and then quenched to room  
15 temperature.[27, 28] The FII and FIII polymorphs used in this study are both kinetically stable at  
16 room temperature and Kuhnert-Brandstatter *et al.* determined by melting point analysis that FIII  
17 was the thermodynamically more stable polymorph with the FII being metastable.[26] The  
18 packing arrangements of FII and FIII are similar, featuring networks of centrosymmetric hydrogen-  
19 bonded dimers of piracetam molecules but with different hydrogen bonding motifs for each  
20 polymorphs.[29] Previous work has shown that one can easily quantify polymorph ratios in pure  
21 piracetam using BR, NIR, and PXRD, with BR and NIR giving LODs of 1.48 and 0.84 % of the  
22 FII polymorph respectively.[30]

23 Quantification of APIs in the presence of excipients in a dosage form where the API is  
24 present as a low percentage by weight is an important issue to address during manufacture from a  
25 regulatory affairs standpoint. In addition, a common problem can be the control (and/or  
26 measurement) of the proportion of polymorphic forms present in a drug formulation at low  
27 concentrations. This problem is exacerbated by the fact that one needs to be able to quantify any  
28 changes in polymorphism over the shelf life of the product, thus there is a need for a reliable, and  
29 rapid testing method capable of measuring very small levels of polymorph contamination. It is  
30 therefore essential to critically evaluate competing, non-destructive, spectroscopic techniques for  
31 the quantification of low level polymorph contamination in low dosage solid formulations such as  
32 tablets. This manuscript assesses the efficacy of NIR and Raman (Backscattering and  
33 Transmission modes) spectroscopies coupled with chemometric methods for the quantitative  
34 measurement of low level polymorph contamination in a model tablet system.

35  
36

## 2. Materials & Methods

### 2.1 Polymorph and Tablet preparation:

Piracetam (FIII polymorph form) was supplied by Sigma Aldrich and used as received. The FII polymorph was prepared by first generating the FI polymorph by heating the FIII (previously ground roughly using an agate pestle and mortar for 1 minute to decrease particle size) to 140 °C in an oven for 72 hours. The FII form was then formed from the unstable FI form at room temperature over four days with the conversion process monitored by Raman spectroscopy. [28, 31] The hold time at 140°C of 72 hours is probably excessive, as the conversion was probably complete in less than this, however, we wanted to ensure that we got 100% conversion, and we thus we used the procedure published by Croker *et al.* [30]. Piracetam (mixtures of FII and FIII), Calcium Carbonate (CaCO<sub>3</sub>), and Microcrystalline Cellulose (MCC) were mixed in a 10:10:80 ratio by weight (Table 1). The Piracetam component (10% by weight) comprised of mixtures of FII and FIII varying from 1% FII with 99% FIII to 99 FII with 1% FIII. These were prepared by weighing out the appropriate amounts of each polymorph and then mixing, this was then combined with the excipients. For each tablet composition the total mass of the prepared solid mixture was 2.0 g so as to minimise weighing errors. To ensure homogeneity of the powder mixtures before pressing into tablets, each powder sample was then mixed thoroughly using a Vortex mixer. The homogeneity of each mixture was measured using NIR every thirty seconds, and mixing was continued until reproducible spectra were obtained as determined by Principal Component Analysis (*see supplemental information, Figures S1 and S2*). The final FI and FIII concentrations in the powders varied from 0.1% to 10% by weight. These mixture ratios were chosen because we wanted to investigate low levels of polymorph contamination in API formulations and see if it was feasible to accurately measure low levels (1-5%) of API polymorph contamination using conventional Raman and NIR sampling systems. In the manufacturing context, the precise concentration level that needs to be measured will of course vary according to the API, polymorph contaminants and specific drug formulation.

Tablets of different thicknesses (1, 2, and 3 mm) were prepared with 15 different FII:FIII concentration ratios (Table 1) using a standard IR hydraulic press (Perkin-Elmer) with a 13 mm die set. Each tablet was formed by compressing the powder at 253 kPa for 10 minutes. The thickness of the tablets was varied by weighing out different amounts of the powder mixture. The exact mass of material for each tablet thickness was first determined by trial and error measurements. For each thickness-concentration sample, three replicate tablets were prepared, generating a complete sample set of 135 tablets which were used for calibration sample sets. Four tablets with 10% Piracetam with FII:FIII ratios of 0.4:9.6, 1.0:9.0, 9.0:1.0 and 9.2:0.8 at each of the three thicknesses were prepared in an identical manner and used as the testing/validation set. The variation in tablet thickness was low, with a maximum % error of 0.8% for the 1 mm, 0.3% for the

1 2 mm and 0.2% for the 3 mm tablets (see *supplemental information, Table S1, for more details of*  
2 *thickness measurements*).

## 4 **2.2 Instrumental Methods:**

5 NIR reflectance spectra were collected using a Perkin Elmer NTS ATR-FT-NIR at a  
6 resolution of 8 cm<sup>-1</sup> from 4000 to 10000 cm<sup>-1</sup> with 32 scans per spectrum. Interleaved scanning  
7 mode was used for collection of all spectra which allows for a ratio of the sample spectrum against  
8 a background that is recorded almost simultaneously, eliminating residual atmospheric absorptions  
9 for the final spectrum. Three spectra were collected from each tablet where the tablet was rotated  
10 120° between measurements. These spectra were then averaged prior to use in the chemometric  
11 modelling.

12 Backscattered Raman Spectroscopy (BR) data were collected using a RamanStation spectrometer  
13 (AVALON Instruments Ltd., Belfast, Northern Ireland; now PerkinElmer) with 785 nm excitation  
14 and a spot size of ~200 µm in diameter. A laser power of 80 mW (at the sample) with an exposure  
15 time of 10 acquisitions x 2 seconds was used and spectra were collected at a 4 cm<sup>-1</sup> resolution from  
16 250 to 3311 cm<sup>-1</sup>. A 4 x 4 sampling grid was used to minimize spectral variance and reduce sub-  
17 sampling effects, resulting in 16 spectra in total which were then averaged for subsequent data  
18 analysis.

19 Transmission Raman Spectroscopy (TRS) data were collected using a Cobalt Light Systems Ltd.  
20 TRS100 Transmission Raman Optical Engine at an excitation wavelength of 830 nm. Tablets were  
21 arranged on a multi-tablet holder plate for analysis. A spot size of 8 mm was used and spectra  
22 were collected at a resolution of 7 cm<sup>-1</sup> from 50 to 2400 cm<sup>-1</sup>. An exposure time of 32 acquisitions  
23 x 0.2 seconds was used, resulting in a 6.4 second analysis time per tablet.

## 25 **2.3 Data Analysis:**

26 Multivariate data analysis was carried out using The Unscrambler software version 8  
27 (Camo, Norway) and the PLS toolbox (Eigenvector) running on Matlab version 8 (Mathworks). A  
28 variety of several pre-processing methods were assessed for each analytical method and these  
29 included 1<sup>st</sup> order and 2<sup>nd</sup> order Savitsky-Golay derivatives, standard normal variate (SNV) and  
30 multiplicative scattering correction (MSC).[32] Principal Component Analysis (PCA) was used  
31 for qualitative analysis and standard Partial Least Squares (PLS) regression was used for  
32 quantitative modelling.[32] For each calibration model, all 45 samples (15 concentrations  
33 replicated 3 times each) were used. These models were then tested using an external validation set  
34 of 4 different tablets with varying API concentrations. The performance of the various PLS  
35 calibration models generated from data employing the different pre-processing methods was  
36 evaluated using the correlation coefficients and the root mean square errors for calibration  
37 (RMSEC) and validation (RMSEV). The accuracy of the test set predictions was quantified using  
38 the root mean square error of prediction (RMSEP).

### 3. Results & Discussion

#### 3.1 Polymorph Conversion:

FII was prepared by first heating FIII to 140 °C, generating FI as determined by PXRD (data not shown), which once cooled to room temperature begins to rapidly transform into FII.[28, 31] BR was used to monitor this transformation at 24 hour intervals from before the FIII sample was heated (Figure 1). It appeared that the FIII to FI transformation was complete after 24 hours of heating while the FI to FII transformation was complete 24 hours after removal of the FI form from the oven. PCA (not shown) of the Raman data showed very clear differences between all three polymorphs with no overlap of the sample cluster 95% ellipses. PXRD diffraction patterns and DSC thermograms (*supplementary information, Figures S3 and S4 respectively*) of the two polymorphs prepared here were compared to previously published data and confirmed that the pure forms of FII and FIII polymorphs had been generated with no impurities detectable.[26, 31]

#### 3.2 Spectral Analysis:

The NIR spectra (Figure 2A) of the polymorphs show some differences in the 5870–5600 and 4314–4080  $\text{cm}^{-1}$  regions. In the FII spectrum we see single peaks at 5724 and 4364  $\text{cm}^{-1}$ , while in the FIII spectra these are both split into two peaks at 5748 / 5708  $\text{cm}^{-1}$  and 4380 / 4358  $\text{cm}^{-1}$  respectively. The Backscattered Raman (Figure 2B) and Transmission Raman (Figure 2C) spectra show much more detail and we observe a lot of sharp bands over the 250  $\text{cm}^{-1}$  to 3310  $\text{cm}^{-1}$  spectral range. Khamchukov *et al.* assigned Raman bands to Piracetam ascribing the bands around 3140, 1680 and 1650  $\text{cm}^{-1}$  to symmetric stretching vibrations of  $\text{NH}_2$ , and the amide I  $\text{C}=\text{O}$  stretching vibration respectively in the ring and acetamide.[33] The bands around 2750–2990  $\text{cm}^{-1}$  can be assigned to the symmetric and anti symmetric stretching vibrations of the  $\text{CH}_2$  groups. It is noted that there are some differences in the Raman spectra of these two polymorphs which are easy to distinguish due to the sharp, well resolved nature of the Raman bands. FII has a single Amide I band at 1654  $\text{cm}^{-1}$ , while in the FIII form this is split into a pair of peaks at 1658 and 1648  $\text{cm}^{-1}$ . A peak specific to FIII is present at 1410  $\text{cm}^{-1}$  and band shifts are observable in the 1530–1370  $\text{cm}^{-1}$  spectral region. Also evident are band shifts in the 890–750  $\text{cm}^{-1}$  spectral region which are due to changes in the bond lengths and angles made by atoms adjacent to the carbonyl groups. The remaining Raman and Infra-red band assignments can be found in the literature. [33, 34]

Figure 2D shows the TRS spectra from the low wavenumber range (44 - 450  $\text{cm}^{-1}$ ) which typically relates to lattice and phonon modes. Access to this spectral region reveals very significant differences between the two polymorphs. There are 3-5  $\text{cm}^{-1}$  differences observed for several of the bands at ~380, 300, 160, and 110  $\text{cm}^{-1}$ . The other major difference is a splitting of the 91  $\text{cm}^{-1}$  band in the FII polymorph where we see a reduction in the intensity at 91  $\text{cm}^{-1}$  coupled with the appearance of a strong shoulder at 79  $\text{cm}^{-1}$ . Khamchukov *et al.*[33] observed the same bands at low wavenumbers, and ascribed the vibrational modes at 42, 84, 125  $\text{cm}^{-1}$  to crystal lattice vibrations. We note that here we observe a band at 91  $\text{cm}^{-1}$  and not at 84  $\text{cm}^{-1}$ . The peak shift



1 around  $106\text{ cm}^{-1}$  is attributed as an out-of-plane vibration in the amide group, which is due to the  
2 change in the angle between  $\text{C}_6\text{C}_5\text{N}_3$  (the Carbon adjacent to the ring, C=O and amide N) and  
3  $\text{O}_2\text{C}_5\text{N}_3$  (O of carboxyl group, C of carboxyl group and amide N) planes. This ability to observe  
4 the low frequency, fundamental vibrational modes is a very significant advantage for TRS  
5 compared to NIR and FT-IR spectroscopy where it is not possible to measure these low energy  
6 transitions. The BRS instrumentation we used could not address this spectral region because its  
7 Rayleigh rejection filters are of a lower standard (the system was manufactured in 2005). However,  
8 many newer Raman systems have much improved Rayleigh rejection filtering which provide  
9 access to this region and it will become much more common for the routine analysis of  
10 pharmaceutical solids using Benchtop instrumentation.

11 It is easy to distinguish the differences between polymorphs from the spectra of the pure  
12 materials, and for simple binary mixtures it is also comparatively easy to observe the presence of  
13 both polymorphs by visual inspection and simple chemometrics.[30] However, once we introduce  
14 these polymorphs into a model system with a large excess of excipient the situation becomes  
15 altogether much more complex (Figure 3). As expected, it is the MCC which has the greatest  
16 impact on all the spectra because of its high concentration. In the NIR spectra (Figure 3A), all we  
17 see essentially are the broad bands of the MCC and it tends to visually obscure the regions of  
18 interest for polymorph analysis. This highlights one of the problems with NIR methods, in that we  
19 have to use chemometric methods (both to pre-process the spectra and to statistically analyse the  
20 variance) to extract any useful data.

21 In the raw Raman spectra the MCC again dominates, but this time its major impact lies in  
22 the large sloping baseline that it imparts on both the BR (Figure 3B) and TRS spectra (Figure 3C  
23 and D). This background can largely be ascribed to diffuse or Mie scattering (and other optical  
24 effects) from the MCC which has a particle size in the micron range ( $\sim 20\text{ }\mu\text{m}$ ). This diffusely  
25 scattered radiation is not collimated and this thus leads to the excitation light being dispersed  
26 across the CCD detector, as described by Bonnier and co-workers.[35] We can discount the  
27 likelihood of fluorescence being the source of the background because there is comparatively little  
28 difference in the background when 830 or 785 nm excitation is used. The extra shot noise  
29 associated with this background scattered light signal will limit sensitivity (for discriminating very  
30 low contaminant concentrations) because it is not possible (with conventional Raman methods) to  
31 separate the shot noise from Raman signal. However despite this large background, much of the  
32 Raman band detail is clearly visible and it is comparatively easy to identify bands due to each of  
33 the polymorphs.

34 In comparing the Raman signal quality from the two methods, it is clear that there is not a  
35 lot to choose between the two, particularly since we are dealing with non-fluorescent samples.  
36 Intrinsically on the excitation side, the 830 nm TRS will be  $\sim 80\%$  as efficient (due to the 4<sup>th</sup> power  
37 wavelength dependence) compared to 785 nm and thus requires more excitation power, likewise  
38 on the detection side the CCD detectors have lower quantum efficiencies in the equivalent Raman  
39 spectral regions. As regards the number of acquisitions used for each instrument, the average TRS

1 spectra were a result of using 32 acquisitions per sample while the BR data was an average of 160  
2 spectra (10 acquisitions per point using 4x4 mapping grid) led to the TRS data being ~6 times  
3 noisier than the BR data. However, for this type of application, these issues are not as important as  
4 the sampling volumes addressed by the two systems. Therefore the intrinsic advantage of 785 nm  
5 excitation was overcome by the better sampling and much lower acquisition times in TRS mode  
6 (*vide infra*). Analysis times per tablet for BR required a total analysis time of 3 minutes using a 4  
7 x 4 grid of spectra, NIR required 2.5 minutes per tablet, while TRS just required 6.5 seconds per  
8 tablet. This dramatic reduction in acquisition time enables the high-throughput screening of large  
9 numbers of samples. Eliminating the noise issue for the TRS data can easily be achieved by the  
10 averaging of multiple spectra, although this would reduce the high throughput somewhat.

11 Apart from the purely signal based considerations, TRS is also better from a sampling  
12 perspective in that with an 8 mm diameter spot size compared to the 200  $\mu\text{m}$  the surface area  
13 sampled is ~100 times greater. Overall the TRS samples ~38% of the tablet surface whereas the  
14 BR only manages a paltry 0.4%, (Calculated by calculating surface area of 200  $\mu\text{m}$  spot size and  
15 multiplying by 16 as a 4x4 grid was used here), and the consequences of this are also evident in the  
16 quantitative analysis discussed below. The effective sampled volume is even greater because the  
17 BR essentially samples the top surface (the estimated depth of field for the BR systems optics is  
18 ~4-600  $\mu\text{m}$ ) whereas in TRS the complete tablet thickness is sampled. This means that the TRS  
19 data should be essentially free from sub-sampling effects compared to BR. The variation in the  
20 raw TRS spectra compared to the BR spectra is also much less when we analysed replicate data  
21 using Principal Component Analysis (see *Figure S5, supplemental information*).

### 22 23 **3.3 Quantitative Analysis by NIR:**

24 The NIR spectra collected from all tablet sample sets show some differences in baseline  
25 due to differences in the scattering of light. The NIR spectra were thus first pre-processed using  
26 Multiplicative Scatter Correction (MSC) to eliminate the baseline offset and scattering effects. In  
27 general we found that the use of second derivative pre-processing was the best method for  
28 amplifying the spectral differences arising from the polymorphs in the 6000 – 5800  $\text{cm}^{-1}$  region.  
29 For the quantification of FII in FIII a variety of discrete spectral regions of interest were used to  
30 build quantitative models (Table 1). The best results (based on an appropriate combination of low  
31 error values and numbers of latent variables) were generally obtained when the 5870 – 5600 and  
32 4314 – 4080  $\text{cm}^{-1}$  spectral regions of the NIR spectra were used. For the NIR data, the best  
33 quantitative model used the 6000 – 5600  $\text{cm}^{-1}$  spectral region (Figure 1) and required only one  
34 latent variable (Table 2 and supplemental information). Analysis of the PC1 loadings plot (Figure  
35 4B) does not generate much in the way of specific information apart from the fact that the  
36 significant variables are evenly spaced over this region. On testing this model with the external  
37 prediction set a RMSEP of 0.54 % FII was obtained which indicates that the FII contamination can  
38 be reasonably well quantified. This 0.54% RMSEP value represents a ~5.4 % level of FII  
39 contamination in the FIII component. These results are comparable to that achieved when NIR

1 was used previously for the determination of bromazepam and clonazepam concentrations in low-  
2 dosage tablets. Here they were able to achieve Standard Errors of Prediction of 0.59 and 0.57 %  
3 w/w respectively.[8]  
4

### 5 **3.4 Quantitative Analysis by Raman:**

6 The BR Raman spectra required a combination of pre-processing methods which were  
7 MSC and a Savitzky Golay second derivative which maximised differences in the spectra between  
8 the differing ratios of polymorphs present particularly in the 1730 – 1370  $\text{cm}^{-1}$  region. The best  
9 BR quantitative model is shown in Figure 5A/B, and it shows a reasonable correlation, however,  
10 the large degree of scatter and high RMS errors are probably due to the sampling issues. It is  
11 therefore clear that a 16 point grid is insufficient to accurately sample the tablet surface. A revised  
12 experimental design using a 12 x 12 grid with a single exposure of 2 seconds at each point would  
13 improve matters, and might generate a more accurate result. However, this increase in sampled  
14 area by a factor of 9 (*i.e.* ~3.4% of total tablet surface sampled) is still a long way from the 38% of  
15 the surface sampled by TRS. The simplest option for the BR method would be to change the  
16 focusing optics to enable a large 1 or 2 mm diameter spot size to be used, or use a PhAT type  
17 probe with a much larger sample spot size,[17] unfortunately neither facility was available for this  
18 study. The PC1 loadings plot from this BR quantitative model (Figure 5B) also demonstrates one  
19 of the problems with the BR measurement, in that it is very difficult to identify the bands  
20 specifically due to the individual polymorphs.

21 TRS spectra were first normalised then a variety of pre-processing methods were applied  
22 including MSC and a 9 point second derivative Savitzky Golay. These methods were used to try  
23 and maximise differences between the spectra as a result of changing FII polymorph concentration.  
24 The 1 mm tablets producing the best quantitative models (Table 4) used the 801 – 1730  $\text{cm}^{-1}$   
25 spectral region, and there is no real statistical difference between the use of the normalised raw  
26 data and the MSC/2<sup>nd</sup> derivative processed spectra, except that the normalised data required 3 LV's.  
27 This spectral region included the majority of the large spectral differences between the polymorphs  
28 such as the band shifts in the 1530–1370  $\text{cm}^{-1}$  spectral region, the differences between the  
29 polymorphs in their respective carbonyl stretching vibrations around 1650  $\text{cm}^{-1}$  and the peak  
30 specific to FIII at 1410  $\text{cm}^{-1}$ . The relatively high degree of consistency in the PLS models built  
31 using the TRS data is testament to the intrinsically high quality of the data collected. Improving on  
32 these results might be achieved by a revised data collection method to reduce the intrinsic noise in  
33 the raw spectra (average ~128 or 256 acquisitions per sample). This could be achieved without an  
34 excessive time penalty and thus maintain the high throughput.

35 A model built using the low wavenumber spectral region (44 – 250  $\text{cm}^{-1}$ ) also generated  
36 good results (Figure 5C/D). The predicted versus measured concentration plot does not show too  
37 much scatter and the PC1 loadings plot indicates that the Piracetam bands at 79 and 106  $\text{cm}^{-1}$  have  
38 the largest influence on the model. One significant advantage of using this spectral region is that  
39 there are very few interfering bands from the MCC and stearate excipients and it is very clear from

1 the loadings plot which are the key spectral variables for modelling. So apart from the slightly  
2 increased noise/signal ratio due to the scattering background the bands changes due to each  
3 polymorphic form are very clear. It is also worth noting that there is very little difference in the  
4 quantitative modelling if one is looking for FII in FIII or visa-versa.

5 Overall, the results from our comparative TRS and BR study are similar to that reported by  
6 Johansson and co-workers for a simple binary tablet system.[12] They used a dedicated tablet  
7 press and mixtures of Propranolol (16 to 24 % w/w) and Mannitol to produce ~3 mm thick tablets.  
8 They found that the TRS outperformed the BR method by ~20% with RESECV values of between  
9 0.4 and 0.6 % w/w being obtained. As in our case, the improved results from TRS are ascribed to  
10 better sampling statistics. It is also worth noting that the quantification accuracy was nearly as  
11 good here, despite the facts that our tablet system has a very significant baseline artefact induced  
12 by the MCC, that the BR spot size used here was a miniscule 200  $\mu\text{m}$  (compared to the 6 mm  
13 diameter in the Johansson study) and that we only used a small 4x4 mapping grid.

### 14 15 **3.5 Effect of Tablet thickness:**

16 We also investigated the effect of tablet thickness on the accuracy of the quantitative  
17 measurements (Figure 6). Overall, we found that the TRS outperformed NIR and BR in terms of  
18 consistency once the data had been pre-processed using normalisation, MSC, and second  
19 derivative. The TRS models are all very similar with only a small increase in RMSEC for the  
20 thicker tablets as might be expected.[12] For the NIR measurements, the 1 mm thick tablets gave  
21 generally the best models as this is due to a more complete sampling of the whole tablet.  
22 Transmission NIR (*see supplemental information*, Figure S6) of the different thickness samples  
23 show that a small fraction (0.1 to ~0.003%) of the NIR incident light passes through the 1 mm  
24 tablets but that the thicker tablets this is further reduced to 0.01 to 0.003%. This indicates that for  
25 the 2/3 mm tablets the effective sampled volume only comprises the top ~1 mm layer and that the  
26 rest of the sample is not analysed in the trans-reflectance mode used. The NIR spectra when  
27 analysed by PCA show systematic variation according to tablet thickness with the 1 mm data being  
28 very distinct from the 2 and 3 mm tablets. However, once the data is normalised, MSC corrected  
29 and a second derivative taken, this variation appears to disappear. For the BR models, the sub-  
30 sampling issue tends to obscure any systematic variation and we cannot see any thickness effect.

## 31 32 33 **4. CONCLUSIONS**

34 Raman and NIR methods were successfully employed for the analysis of a model  
35 pharmaceutical system containing low levels of a polymorphic contaminant. It was possible to  
36 generate quantitative models using PLS to predict low content FII polymorph contaminant in these  
37 tablets with an accuracy of ~0.6% by total weight (or a 6% API contamination level). In each case,  
38 some spectral pre-processing was necessary for the development of robust quantitative models as it  
39 amplified the chemical information in the spectra. Models built using Backscattering Raman data

1 were the poorest due to the inherent sub-sampling issue associated with the particular system used.  
2 Quantitative models built using Transmission Raman and NIR spectra were comparable in terms of  
3 limits of detection with TRS slightly better with a LOD of 0.6 %FII with NIR having a LOD of  
4 0.7 %FII. This small improvement may be due to the fact that the spectral differences between the  
5 polymorphs were more readily observable in the Raman spectra because of the sharp well resolved  
6 spectral features compared with the broad featureless NIR bands. Another consideration is the fact  
7 that the NIR method does not effectively sample the complete thickness of the tablet whereas TRS  
8 does so very effectively. In terms of analysis time, TRS was by far the quickest means of analysis,  
9 taking approximately 6 seconds per tablet which contrasts to the 2.5 and 3 minutes per sample  
10 necessary for NIR and BR respectively. It is also important to note that the improvement in the  
11 quality of optical filters used for Rayleigh scattering rejection have improved dramatically,  
12 enabling this TRS system to access the low wavenumber region ( $44\text{-}300\text{ cm}^{-1}$ ) which is not readily  
13 available to NIR or FT-IR spectroscopic methods. In the near future we expect that most benchtop  
14 Raman systems (BR and TRS) will fitted with these types of filter thus providing routine and  
15 inexpensive access to this spectral range which is of course important for the analysis of lattice and  
16 phonon modes of the solid state. This study validates the potential of TRS for the quantification of  
17 low levels ( $\sim 1\%$ ) of polymorphic contaminant in pharmaceutical formulations. These LODs are  
18 reasonable considering the experimental design and instrumentation employed. To improve the  
19 LODs below this level, will entail refining the instrumentation to reduce/eliminate the  
20 diffusely/Mie scattered light to reduce the shot noise which obscures the Raman signals from very  
21 low levels of polymorph contaminants.

## 22 23 24 **5. ACKNOWLEDGEMENTS**

25 The authors are grateful for funding received from Science Foundation Ireland (SFI) for the  
26 funding of the Solid State Pharmaceutical Cluster (SSPC). Prof. Pavel Matousek, Dr. Darren  
27 Andrews and Dr. Andrew Owen of Cobalt Light Systems Ltd. are thanked for the use of the Cobalt  
28 Light Systems Ltd. TRS100 Transmission Raman Optical Engine. Dermot McGrath is thanked for  
29 DSC analysis.

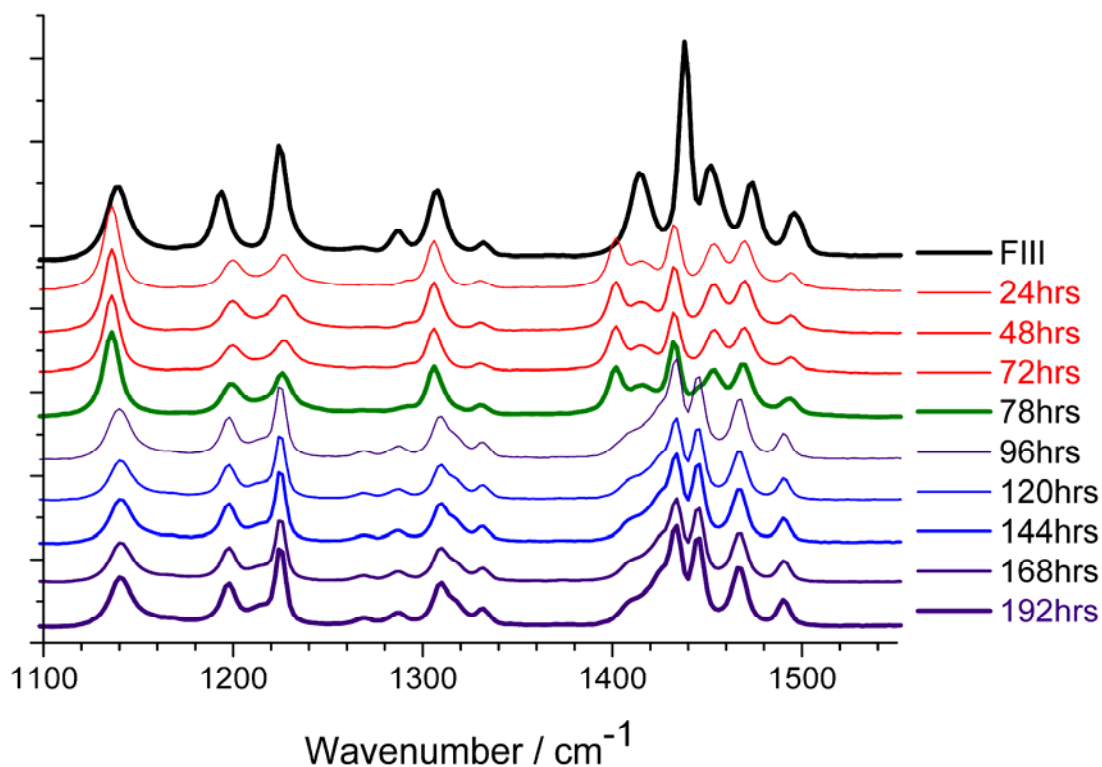
## 1 REFERENCES

- 2 [1] S. R. Vippagunta, H. G. Brittain, D. J. W. Grant, Crystalline solids, *Advanced Drug Delivery Reviews*,  
3 48 (2001) 3-26.
- 4 [2] J. Haleblan, W. McCrone, Pharmaceutical applications of polymorphism, *Journal of Pharmaceutical*  
5 *Science*, 58 (1969) 911-929.
- 6 [3] S. R. Chemburkar, J. Bauer, K. Deming, H. Spiwek, K. Patel, J. Morris, R. Henry, S. Spanton, W. Dziki,  
7 W. Porter, J. Quick, P. Bauer, J. Donaubauer, B. A. Narayanan, M. Soldani, D. Riley, K. McFarland,  
8 Dealing with the Impact of Ritonavir Polymorphs on the Late Stages of Bulk Drug Process Development,  
9 *Organic Process Research & Development*, 4 (2000) 413-417.
- 10 [4] J. Bauer, S. Spanton, R. Henry, J. Quick, W. Dziki, W. Porter, J. Morris, Ritonavir: An Extraordinary  
11 Example of Conformational Polymorphism, *Pharmaceutical Research*, 18 (2001) 859-866.
- 12 [5] P. McArdle, K. Gilligan, D. Cunningham, A. Ryder, Determination of the polymorphic forms of  
13 bicifadine hydrochloride by differential scanning calorimetry-thermogravimetric analysis, X-ray powder  
14 diffraction, attenuated total reflectance-infrared spectroscopy, and attenuated total reflectance-near-infrared  
15 spectroscopy, *Applied Spectroscopy*, 59 (2005) 1365-1371.
- 16 [6] Y. Hu, A. Erxleben, A. G. Ryder, P. McArdle, Quantitative analysis of sulfathiazole polymorphs in  
17 ternary mixtures by attenuated total reflectance infrared, near-infrared and Raman spectroscopy, *J. Pharm.*  
18 *Biomed. Anal.*, 53 (2010) 412-420.
- 19 [7] D. A. Burns, E. W. Ciurczak, *Handbook of Near-Infrared Analysis*, Second ed., 2001.
- 20 [8] P. Chalus, Y. Roggo, S. Walter, M. Ulmschneider, Near-infrared determination of active substance  
21 content in intact low-dosage tablets, *Talanta*, 66 (2005) 1294-1302.
- 22 [9] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, A review of near infrared  
23 spectroscopy and chemometrics in pharmaceutical technologies, *J. Pharm. Biomed. Anal.*, 44 (2007) 683-  
24 700.
- 25 [10] C. J. Strachan, T. Rades, K. C. Gordon, J. Rantanen, Raman spectroscopy for quantitative analysis of  
26 pharmaceutical solids, *J. Pharm. Pharmacol.*, 59 (2007) 179-192.
- 27 [11] K. Kachrimanis, D. E. Braun, U. J. Griesser, Quantitative analysis of paracetamol polymorphs in  
28 powder mixtures by FT-Raman spectroscopy and PLS regression, *J. Pharm. Biomed. Anal.*, 43 (2007) 407-  
29 412.
- 30 [12] J. Johansson, A. Sparen, O. Svensson, S. Folestad, M. Claybourn, Quantitative transmission Raman  
31 spectroscopy of pharmaceutical tablets and capsules, *Applied Spectroscopy*, 61 (2007) 1211-1218.
- 32 [13] S. Mazurek, R. Szostak, Quantification of atorvastatin calcium in tablets by FT-Raman spectroscopy,  
33 *Journal of pharmaceutical and biomedical analysis*, 49 (2009) 168-172.
- 34 [14] A. G. Ryder, G. M. O'Connor, T. J. Glynn, Identifications and quantitative measurements of narcotics  
35 in solid mixtures using near-IR Raman spectroscopy and multivariate analysis, *J. Forensic Sci.*, 44 (1999)  
36 1013-1019.
- 37 [15] P. Matousek, A. W. Parker, Bulk Raman Analysis of Pharmaceutical Tablets, *Applied Spectroscopy*,  
38 60 (2006) 1353-1357.
- 39 [16] J. Johansson, S. Pettersson, S. Folestad, Characterization of different laser irradiation methods for  
40 quantitative Raman tablet assessment, *J. Pharm. Biomed. Anal.*, 39 (2005) 510-516.
- 41 [17] M. V. Schulmerich, W. F. Finney, R. A. Fredricks, M. D. Morris, Subsurface Raman spectroscopy and  
42 mapping using a globally illuminated non-confocal fiber-optic array probe in the presence of Raman photon  
43 migration, *Applied Spectroscopy*, 60 (2006) 109-114.

- 1 [18] C. Eliasson, N. A. Macleod, L. C. Jayes, F. C. Clarke, S. V. Hammond, M. R. Smith, P. Matousek,  
2 Non-invasive quantitative assessment of the content of pharmaceutical capsules using transmission Raman  
3 spectroscopy, *J. Pharm. Biomed. Anal.*, 47 (2008) 221-229.
- 4 [19] M. Fransson, J. Johansson, A. Sparén, O. Svensson, Comparison of multivariate methods for  
5 quantitative determination with transmission Raman spectroscopy in pharmaceutical formulations, *J.*  
6 *Chemometr.*, 24 (2010) 674-680.
- 7 [20] A. Aina, M. D. Hargreaves, P. Matousek, J. C. Burley, Transmission Raman spectroscopy as a tool for  
8 quantifying polymorphic content of pharmaceutical formulations, *Analyst*, 135 (2010) 2328-2333.
- 9 [21] M. D. Hargreaves, N. A. Macleod, M. R. Smith, D. Andrews, S. V. Hammond, P. Matousek,  
10 Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-  
11 component pharmaceutical capsules, *Journal of pharmaceutical and biomedical analysis*, 54 (2011) 463-468.
- 12 [22] M. D. Hargreaves, N. A. Macleod, M. R. Smith, D. Andrews, S. V. Hammond, P. Matousek,  
13 Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-  
14 component pharmaceutical capsules, *J. Pharm. Biomed. Anal.*, 54 (2011) 463-468.
- 15 [23] C. M. McGoverin, M. D. Hargreaves, P. Matousek, K. C. Gordon, Pharmaceutical polymorphs  
16 quantified with transmission Raman spectroscopy, *J. Raman Spectrosc.*, 43 (2012) 280-285.
- 17 [24] C. M. McGoverin, L. C. H. Ho, J. A. Zeitler, C. J. Strachan, K. C. Gordon, T. Rades, Quantification of  
18 binary polymorphic mixtures of ranitidine hydrochloride using NIR spectroscopy, *Vibrational Spectroscopy*,  
19 41 (2006) 225-231.
- 20 [25] A. W. Pavlova, Polymorphic Forms of Piracetam, *Pharmazie*, 34 (1979) 449-450.
- 21 [26] M. Kuhnert-Brandstaetter, A. Burger, R. Voellenklee, Stability behavior of piracetam polymorphs,  
22 *Scientifica Pharmaceutica*, 62 (1994) 307-316.
- 23 [27] F. P. A. Fabbiani, D. R. Allan, W. I. F. David, A. J. Davidson, A. R. Lennie, S. Parsons, C. R. Pulham,  
24 J. E. Warren, High-Pressure Studies of Pharmaceuticals: An Exploration of the Behavior of Piracetam,  
25 *Crystal Growth and Design*, 7 (2007) 1115-1124.
- 26 [28] R. Ceolin, V. Agafonov, D. Louer, V. A. Dzyabchenko, S. Toscani, J. M. Cense, Phenomenology of  
27 polymorphism. III. The p,T diagram and stability of piracetam polymorphs, *Journal of Solid State*  
28 *Chemistry*, 122 (1996) 186-194.
- 29 [29] G. Admiraal, J. C. Eikelenboom, A. Vos, Structures of the triclinic and monoclinic modifications of (2-  
30 oxo-1-pyrrolidinyl)acetamide, *Acta Crystallographica Section B*, 38 (1982) 2600-2605.
- 31 [30] D. M. Croker, M. C. Hennigan, A. Maher, Y. Hu, A. G. Ryder, B. K. Hodnett, A comparative study of  
32 the use of powder X-ray diffraction, Raman and NIR spectroscopy for quantification of binary polymorphic  
33 mixtures of Piracetam. , *J. Pharm. Biomed. Anal.*, 63 (2012) 80-86.
- 34 [31] A. Pavlova, K. Konstantinova, H. Daskalov, A. Georgiev, A study of crystalline forms of piracetam,  
35 *Pharmazie*, 38 (1983) 634-635.
- 36 [32] H. Martens, T. Naes, *Multivariate Calibration*, 2nd Edition ed., Wiley, New York, 1991.
- 37 [33] Y. D. Khamchukov, S. N. Shashkov, I. V. Lukomskii, Vibrational Spectra of 2-Oxo-  
38 Pyrrolidineacetamide, *Journal of Applied Spectroscopy*, 53 (1990) 1281-1286.
- 39 [34] R. Picciochi, H. P. Diogo, M. E. da Piedade, Thermodynamic characterization of three polymorphic  
40 forms of piracetam, *Journal of Pharmaceutical Science*, 100 (2011) 594-603.
- 41 [35] F. Bonnier, A. Mehmood, P. Knief, A. D. Meade, W. Hornebeck, H. Lambkin, K. Flynn, V.  
42 McDonagh, C. Healy, T. C. Lee, F. M. Lyng, H. J. Byrne, In vitro analysis of immersed human tissues by  
43 Raman microspectroscopy, *J. Raman Spectrosc.*, 42 (2010) 888-896.

1  
2  
3

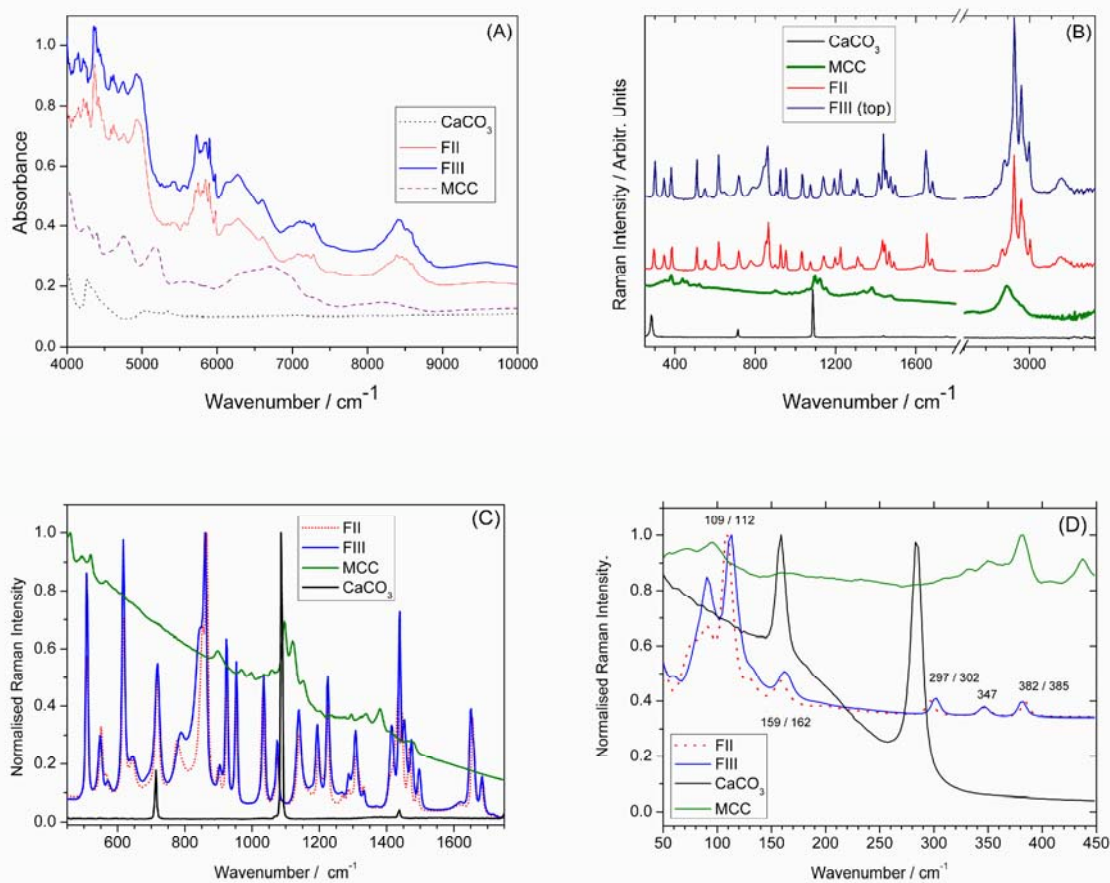
## Figures



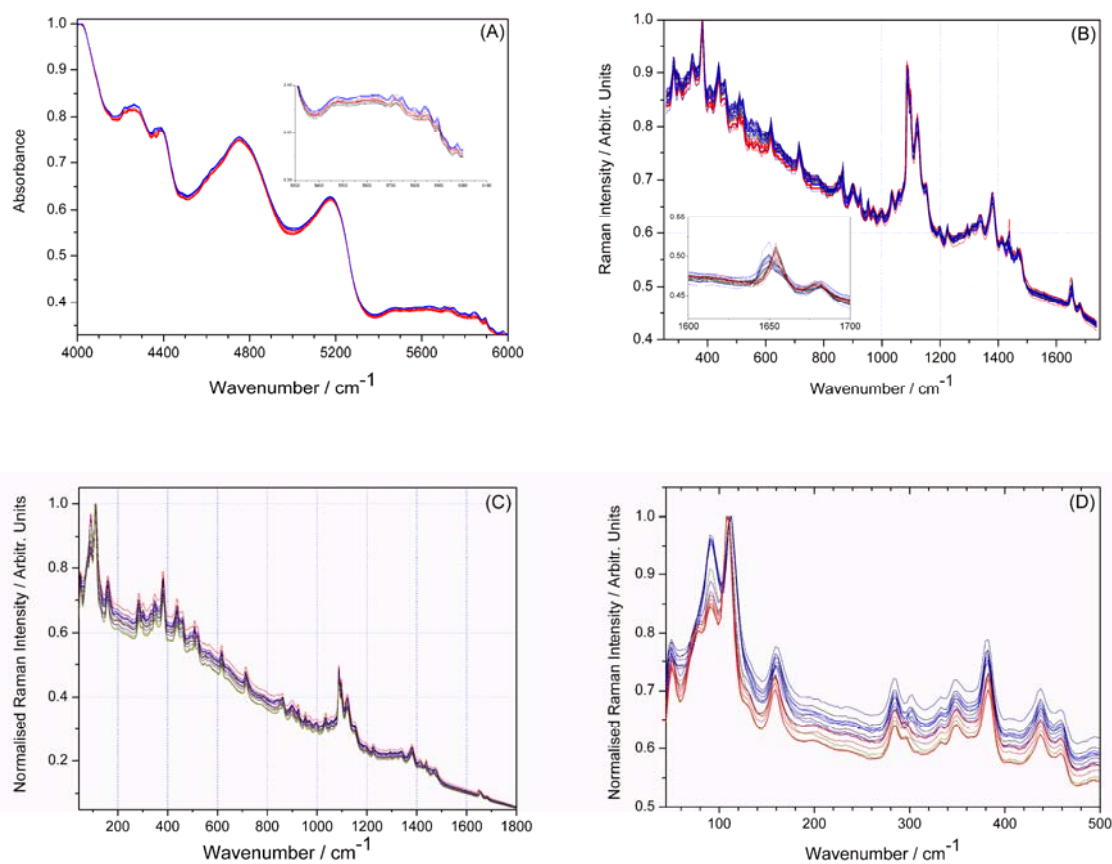
4  
5  
6  
7  
8  
9

**Figure 1:** Raman spectra (Backscattered) showing the conversion of Piracetam Form III to Form II via transformation through FI. The top trace represents starting (FIII) material at room temperature, the 24-72 hr. traces are the sample at 140 °C (FI form), and the 78-192 hr traces the samples at room temperature after heating.

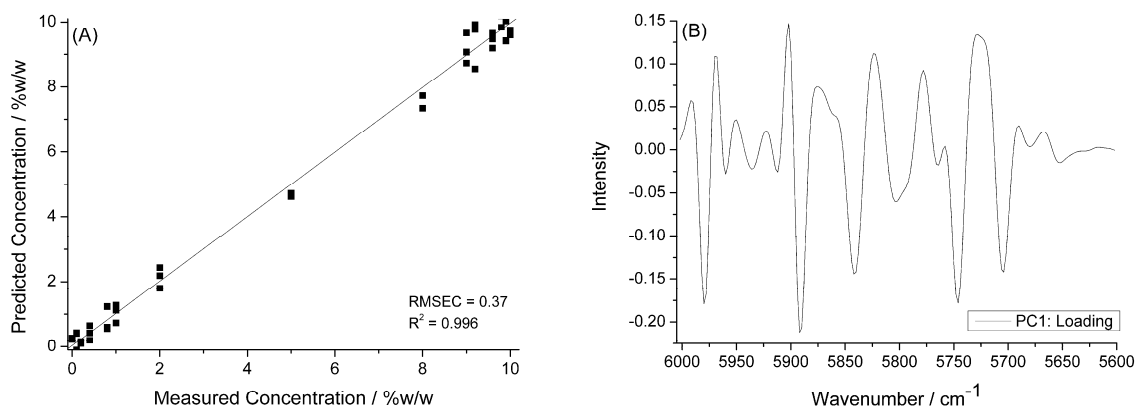




1  
2 **Figure 2:** Spectral analysis of the FII and FIII Piracetam polymorphs and excipient materials used for tablet  
3 fabrication. (A). NIR spectra, (B). Backscattering Raman using 785 nm excitation, (C). Transmission  
4 Raman spectra using 830 nm excitation, and (D). Transmission Raman spectra of the low wavenumber  
5 region.

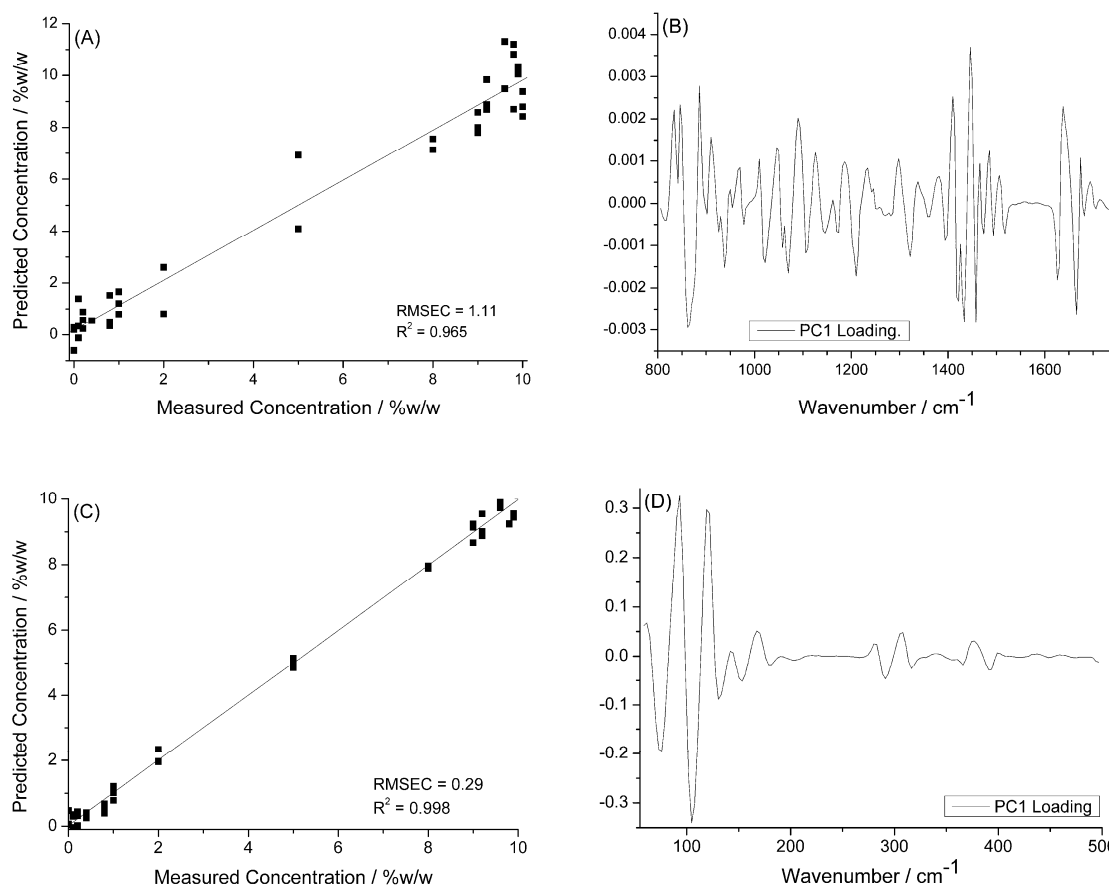


1  
2 **Figure 3:** Normalised but otherwise unprocessed Spectra collected from 1 mm thick Piracetam tablets: (A).  
3 NIR, (B). BR using 785 nm excitation, (C). TRS using 830 nm excitation, and (D). TRS data from the sub  
4 500 cm<sup>-1</sup> region. All spectra have been normalised to the peak of maximum intensity without any  
5 background correction.  
6

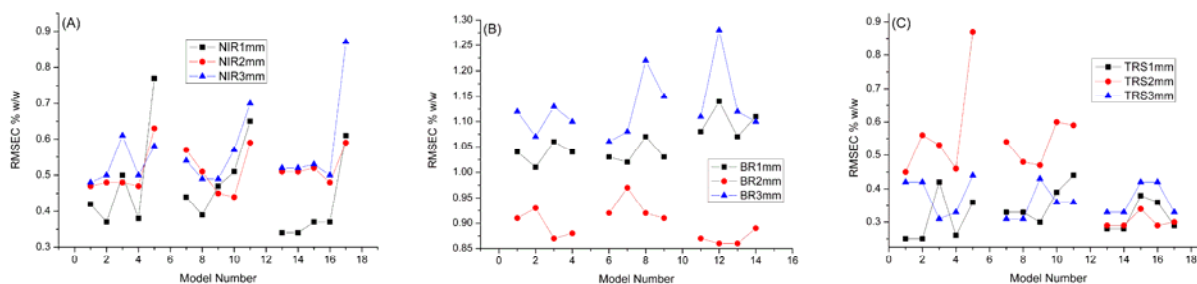


7

1 **Figure 4:** Best Quantitative model for predicting FII concentration in 1 mm thick tablets using NIR  
2 spectroscopy. (A). Measured vs. predicted FII Piracetam (calibration & prediction data) and (B). PC1  
3 loading plots from the best performing quantitative PLS model indicating the most significant spectral  
4 features.



5  
6 **Figure 5:** Best Quantitative models for predicting FII concentration in 1 mm thick tablets as determined by  
7 Raman spectroscopy. (A). Measured vs. predicted FII Piracetam concentration plot generated from BR data  
8 (calibration & prediction data), (B). PC1 loadings plot from the best performing PLS model using BR data.  
9 (C). Measured vs. predicted FII Piracetam concentration plot generated from the low wavenumber region  
10 TRS data (calibration & prediction data), (D). PC1 loadings plot from the best performing PLS model  
11 using the low wavenumber region TRS data.



1  
 2 **Figure 6:** Comparison of model performance in terms of RMSEC (% w/w) for the different tablet  
 3 thicknesses according to: (A). NIR analysis, (B). Backscattered Raman, and (C). Transmission Raman.  
 4 The model number (reading from left to right) correspond to the models as listed in Tables 2-4 (reading  
 5 from top to bottom).  
 6  
 7

8 **TABLES**

9

Technique	Name/Identifier	Wavenumber region (cm <sup>-1</sup> )
<i>NIR</i>	Full	10000 – 4000
	Half	6000 – 4000
	A and B	6000 – 5600, 4314 - 4080
	A	6000 – 5600
	B	4314 - 4080
<i>Backscattered Raman</i>	Full	250 – 3300
	Half	250 – 1730
	A	1350 – 1730
	B	802 – 1730
<i>Transmission Raman</i>	Full	44 – 2450
	Half	44 – 678
	A	1370 – 1730
	B	801 – 1730
	C	44 – 250

10  
 11 **Table 1:** Regions of interest selected for model generation using NIR and Raman data.  
 12

1

Pre-processing	Region	LV	RMSEC	R <sup>2</sup>	RMSECV	RMSEP
Raw data	Full	5	0.42	0.995	0.52	0.96
	Half	5	0.37	0.996	0.44	1.19
	A and B	4	0.50	0.993	0.57	1.50
	A	3	0.38	0.996	0.43	0.87
	B	2	0.77	0.983	0.83	2.18
MSC	Full	4	0.44	0.995	0.51	0.94
	Half	3	0.39	0.996	0.52	1.01
	A and B	2	0.47	0.994	0.54	1.44
	A	2	0.51	0.993	0.56	0.72
	B	2	0.65	0.988	0.70	1.24
MSC & 2 <sup>nd</sup> deriv.	Full	1	0.34	0.997	0.36	0.55
	Half	1	0.34	0.997	0.36	0.56
	A and B	1	0.37	0.996	0.39	0.59
	A	1	0.37	0.996	0.38	0.54
	B	1	0.61	0.990	0.63	1.22

2

3

**Table 2:** Performance of the regression models using different pre-processing methods. This data was the NIR spectra collected from the 1 mm thick tablets. RMSE values in % w/w.

4

5

6

Pre-processing	Region	LV	RMSEC	r <sup>2</sup>	RMSECV	RMSEP
Raw data ( <i>norm.</i> )	Full	4	1.04	0.969	1.29	2.50
	Half	3	1.01	0.970	1.20	2.75
	A	3	1.06	0.968	1.20	2.62
	B	3	1.04	0.969	1.21	2.37
MSC	Full	4	1.03	0.969	1.27	2.62
	Half	3	1.02	0.970	1.18	2.59
	A	3	1.07	0.967	1.21	2.75
	B	3	1.03	0.969	1.18	2.52
MSC & 2 <sup>nd</sup> deriv.	Full	1	1.08	0.967	1.17	2.40
	Half	1	1.14	0.963	1.19	1.92
	A	2	1.07	0.967	1.16	2.65
	B	1	1.11	0.965	1.16	2.11

7

8

**Table 3:** Performance of the regression models using different pre-processing methods for 1 mm BR spectra. RMSE values in % w/w.

9

10

1

Pre-processing	Region	LV	RMSEC	r <sup>2</sup>	RMSECV	RMSEP
Norm data	Full	3	0.25	0.998	0.28	0.37
	Half	3	0.25	0.998	0.28	0.36
	A	2	0.42	0.952	0.46	0.25
	<b>B</b>	<b>3</b>	<b>0.26</b>	<b>0.998</b>	<b>0.29</b>	<b>0.31</b>
	C	2	0.36	0.997	0.40	0.39
MSC	Full	2	0.33	0.997	0.37	0.44
	Half	2	0.33	0.997	0.36	0.44
	A	2	0.30	0.998	0.35	0.39
	B	2	0.39	0.996	0.44	0.33
	C	2	0.44	0.995	0.51	0.37
MSC & 2 <sup>nd</sup> deriv.	Full	1	0.28	0.998	0.30	0.61
	Half	1	0.28	0.998	0.30	0.61
	A	1	0.38	0.996	0.41	0.30
	<b>B</b>	<b>1</b>	<b>0.36</b>	<b>0.997</b>	<b>0.37</b>	<b>0.29</b>
	C	1	0.29	0.998	0.30	0.64

2

3

4

5

6

**Table 4:** Performance of the regression models using different pre-processing methods. This data was the TRS spectra collected from the 1 mm thick tablets. RMSE values in % w/w.