Journal of Pharmaceutical and Biomedical Analysis 54 (2011) 463–468

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-component pharmaceutical capsules

Michael D. Hargreaves[a], Neil A. Macleod[b], Mark R. Smith[c], Darren Andrews[a], Stephen V. Hammond[c], Pavel Matousek[a,b,*]

[a] Cobalt Light Systems Ltd, Start Electron, Fermi Avenue, Harwell Science & Innovation Campus, Didcot, Oxfordshire OX11 0QR, UK
[b] Central Laser Facility, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, Oxfordshire OX11 0 QX, UK
[c] Pfizer, Process Analytical Support Group (PASG), Pfizer Ltd, Ringaskiddy, Co. Cork, Ireland

ARTICLE INFO

Article history:
Received 20 April 2010
Received in revised form
10 September 2010
Accepted 10 September 2010
Available online 19 September 2010

Keywords:
Transmission Raman
Process control
Pharmaceutical analysis
Noninvasive

ABSTRACT

A detailed characterisation of the performance of transmission Raman spectroscopy was performed from the standpoint of rapid quantitative analysis of pharmaceutical capsules using production relevant formulations comprising of active pharmaceutical ingredient (API) and 3 common pharmaceutical excipients. This research builds on our earlier studies that identified the unique benefits of transmission Raman spectroscopy compared to conventional Raman spectroscopy. These include the ability to provide bulk information of the content of capsules, thus avoiding the sub-sampling problem, and the suppression of interference from the capsule shell. This study demonstrates, for the first time, the technique's insensitivity to the amount of material held within the capsules. Different capsules sizes with different overall fill weights (100–400 mg) and capsule shell colours were assayed with a single calibration model developed using only one weight and size sample set (100 mg) to a relative error of typically <3%. The relative root mean square error of prediction of the concentration of API for the main sample set (nominal content 75%, w/w) was 1.5% with a 5 s acquisition time. Models built using the same calibration set also predicted the 3 low level excipients with relative errors of 5–15%. The quantity of API was also predicted (with a relative error within ~3%) using the same model for capsules prepared with different generations of API (i.e. API manufactured via different processes). The study provides further foundation blocks for the establishment of this emerging technique as a routine pharmaceutical analysis tool, capitalising on the inherently high chemical specificity of Raman spectroscopy and the non-invasive nature of the measurement. Ultimately, this technique has significant promise as a Process Analytical Technology (PAT) tool for online production application.

1. Introduction

A number of analytical applications within the pharmaceutical production environment would benefit greatly from a rapid, non-invasive and, most importantly, a non-destructive method for the assay of the bulk content of pharmaceutical formulations. In recent work Matousek and Parker [1] reported previously unexploited properties of transmission Raman spectroscopy, such as gross insensitivity to sampled depth and suppression of interfering surface fluorescence and Raman signals, which are beneficial for rapid probing of pharmaceutical capsules and tablets. This development stimulated new research in this area, paving the way for the establishment of a new analytical technique well suited to a number of problems faced by the pharmaceutical industry.

In the transmission Raman geometry, the laser beam is incident upon one side of the probed capsule and the Raman light is collected from the *opposite side* (see Fig. 1). This mode of operation stems from the development of non-invasive methods such as Spatially Offset Raman Spectroscopy (SORS) [2–4] for deep probing of diffusely scattering samples [5–7]. The transmission Raman concept can be considered to be a special case of SORS with the illumination and collection points displaced to the extreme and as such it shares some common characteristics with SORS. Although the transmission Raman technique was demonstrated in the early days of Raman spectroscopy [8], its benefits for the non-invasive probing of the bulk content of pharmaceutical samples [6] had not been previously recognised or exploited. In addition to the elimination of the sub-sampling problem [1,9] the technique also effectively suppresses both fluorescence and Raman components emanating

* Corresponding author at: Central Laser Facility, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0 QX, UK. Tel.: +44 0 1235 445377; fax: +44 0 1235 445693.
E-mail address: Pavel.Matousek@stfc.ac.uk (P. Matousek).

0731-7085/$ – see front matter.
doi:10.1016/j.jpba.2010.09.015

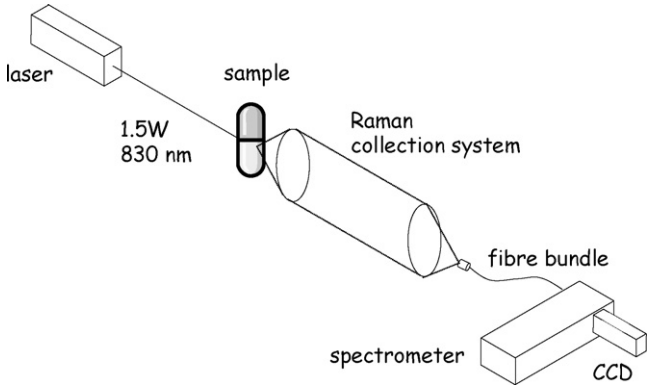

Fig. 1. Schematics of the transmission Raman experimental layout.

from coating or capsule layers [1]. Added benefits are that it does not involve sample preparation in contrast with HPLC methods, which involve dissolution of the sample.

In further research, Johansson et al. [10] established the ability of the transmission Raman technique to provide quantitative information on the binary mixture content of capsules and tablets. In this study, 20 test tablets and an unspecified number of specifically formulated capsules were used. The study achieved a relative root mean square error of prediction for active pharmaceutical ingredient (API) concentration of 2.2% and 3.6%, respectively, with a 5 s acquisition time. The investigation also demonstrated the robustness of the calibration model, which provided satisfactory prediction accuracy for a relatively low number of calibration points. A similar study was performed by Eliasson et al. [11] using transmission Raman spectroscopy on production relevant formulations prepared in a laboratory environment. The study was performed on 150 capsules and the measured relative root mean square error of prediction for the API (1.2% for a 5 s acquisition time) met the particular acceptance criterion for the uniformity of content application. In this earlier work the amount of API and the main excipient were strongly correlated making the prediction less challenging and the two other low level excipients went largely undetected, with the formulations behaving essentially as binary mixtures.

Here, further insight into transmission Raman spectroscopy and its properties is provided by performing more thorough investigations and applying the method to a more challenging formulation made up of four components and with reduced correlation between the API and the principal excipients. The robustness of the method is tested by applying a calibration model developed using a single capsule size to capsules of varying sizes with differing fill weights and to formulations with API manufactured using a different process (different API generations). Such sample property variations may be expected to pose a challenge to conventional NIR absorption methods, which would typically require calibration to be performed for each individual set separately.

2. Experimental section and methods

2.1. Instrumental

The Raman spectra were measured using a Cobalt Light System Ltd TRS100 transmission Raman optical engine. The probe beam was generated using a frequency stabilised diode laser operating at 830 nm. The laser power at the sample was 1.5 W and the laser spot diameter was ~3–4 mm. The automated sample handling system of the TRS100 was not used in this study with samples being inserted into the interaction zone manually.

Table 1
Relative concentrations (% weight) of the four ingredients of the pilot-scale produced batches.

Blend, #	API (w/w)	Talc (w/w)	Starch (w/w)	Lactose (w/w)
1	0.675	0.084	0.075	0.166
2	0.675	0.101	0.101	0.124
3	0.675	0.067	0.084	0.174
4	0.713	0.075	0.067	0.145
5	0.713	0.101	0.084	0.103
6	0.713	0.092	0.092	0.103
7	0.750	0.092	0.084	0.074
8	0.750	0.067	0.075	0.108
9	0.750	0.084	0.101	0.066
10	0.788	0.075	0.101	0.037
11	0.788	0.101	0.092	0.020
12	0.788	0.092	0.092	0.028
13	0.825	0.067	0.075	0.033
14	0.825	0.084	0.067	0.024
15	0.825	0.075	0.067	0.033

Raman light was collected from the opposite side of the sample to the illumination site using a 25 mm diameter lens with a focal length of 125 mm. The scattered light was collimated and passed through a long pass spectral filter to suppress the elastically scattered component of light. The Raman signal was focused into a bundle of optical fibres and transferred into a spectrograph. Raman spectra were collected using a NIR front illuminated deep-depletion thermoelectrically cooled CCD camera (Andor Technology) by binning the entire chip vertically. The spectra are not corrected for the variation of the detection sensitivity across the spectral range. The spectral resolution was estimated to be 5–6 cm^{-1}. The acquisition time was 5 s.

2.2. Samples

The main sample set comprised 150 soft gelatine capsules (capsule body colour white and cap red) with a target fill weight of 100 mg ± 8 %. 15 distinct production relevant blends were manufactured at pilot scale; API (67.5–82.5%, w/w) and three excipients, lactose (2.0–17.4%, w/w), maize starch (6.7–10.1%, w/w) and talcum powder (6.7–10.1%, w/w). The concentrations of individual blends are given in Table 1. The nominal fill weights of the capsules were 100 mg, with the API present at 75%, w/w (i.e. 75 mg API/capsule).

Blends were designed to minimise correlations between the API and the main excipient, lactose. A higher correlation was tolerated between starch and talc to reduce the correlation between API and lactose. A scatter plot of lactose against API concentration is shown

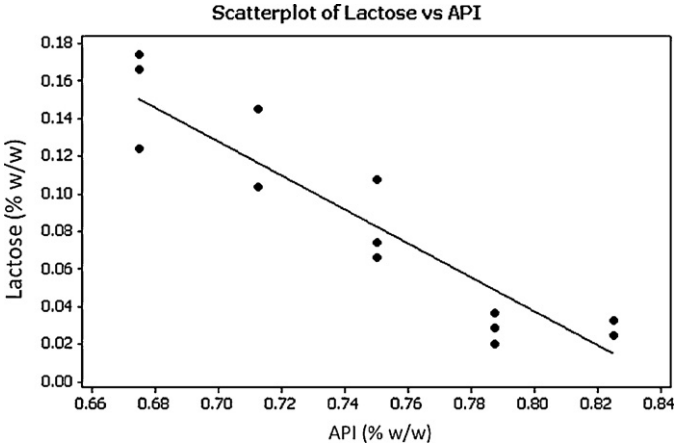

Fig. 2. A concentration scatter plot of lactose and API.

in Fig. 2. The correlation coefficient between lactose and API was −0.925 (R^2 = 0.856).

Other capsules sizes were available with the same API concentration (fill weights 133.3, 200, 266.6, 300 and 400 mg), which were used to test applicability of the model developed using the main sample set. The capsules also had different colour combinations of capsule body and cap. Both full and empty capsules were available for all fill weights.

Other samples were used to test for specificity, precision and applicability, which included empty and filled capsules for 2 production lots and individual samples of all raw materials and API. Two different API generations, manufactured by a different process, were also included in the analysis.

2.3. Numerical analysis

Data was analysed in MATLAB (version 2008b) with the PLS Toolbox (Eigenvector v5.2). The pre-processing routine comprised baseline subtraction with a non-negative spectral constraint, normalisation to unit length across the whole spectral region and mean-centring of the entire dataset. The multivariate techniques used were principal component analysis (PCA) and partial least squares (PLS1) regression [12]. In the case of PLS, cross validation was achieved using random subsets.

The main dataset comprised ~150 individual spectra representing 15 distinct chemical compositions ("blends") with 10 capsules measured from each blend. 10 clear outliers (i.e. spectra not representative of the rest of the dataset) were identified visually in the PLS analysis and excluded from further analysis. A possible reason for the outliers was manual positioning of the samples and associated error as well as a few instances of burnt capsule shell due to prolonged laser illumination during initial instrument alignment and optimisation. For the calibration and regression methods, the remaining dataset (containing ~140 samples) was randomly split into a calibration set (two thirds) and a prediction set (one third), ensuring an adequate distribution from individual blends. A model was constructed using the calibration set and subsequently applied to the prediction set. The final quality of fit was estimated from the root mean square error of prediction (RMSEP, Eq. (1)), where Y_{meas} is the reference concentration of the species of interest, Y_{pred} is the predicted concentration and n is the number of points in the prediction set.

$$\text{RMSEP} = \sqrt{\frac{\sum (Y_{meas} - Y_{pred})^2}{n}} \quad (1)$$

3. Results and discussion

For all supplied samples, transmission Raman spectroscopy provided spectra indicative of the bulk content of the capsule with minimal influence by the capsule shell. Fig. 3 shows typical spectra for both filled and empty capsules. Bands due to the capsule shell (e.g. at ~640 cm^{-1}) are largely reduced from the spectrum of the filled capsules unlike with conventional Raman spectroscopy where it was demonstrated that these bands would be exacerbated in non-invasive spectra, potentially hampering quantitative analysis [11]. No attempt was made here to remove these residual components from the spectra.

Fig. 4 shows the transmission Raman spectra of pure samples of the API, lactose, starch and talc. Comparison with the spectrum of a filled capsule confirms that transmission Raman spectroscopy successfully samples the bulk content of the capsule. A number of spectral regions (~390 cm^{-1}, ~450 cm^{-1}, ~830 cm^{-1} and ~1100 cm^{-1}) can be identified that allow the excipients (principally lactose) to be distinguished from the active ingredient.

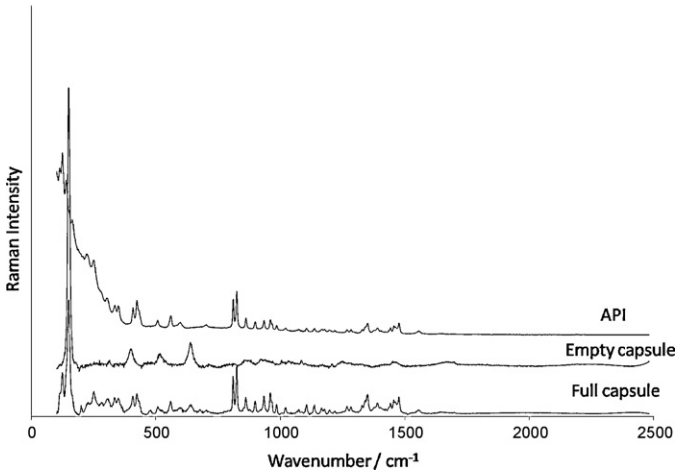

Fig. 3. Typical transmission Raman spectra for API alone, filled and empty capsules (150 mg API/capsule).

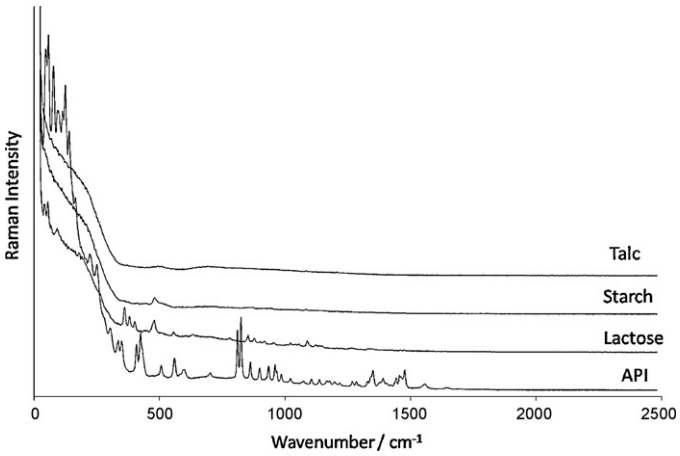

Fig. 4. Raman spectra of pure formulation components (from top to bottom: talc, starch, lactose and API).

3.1. The main sample set – 150 DOE samples

Transmission Raman spectra were obtained of 150 capsules from 15 distinct blends (10 samples from each blend). Fig. 5 shows

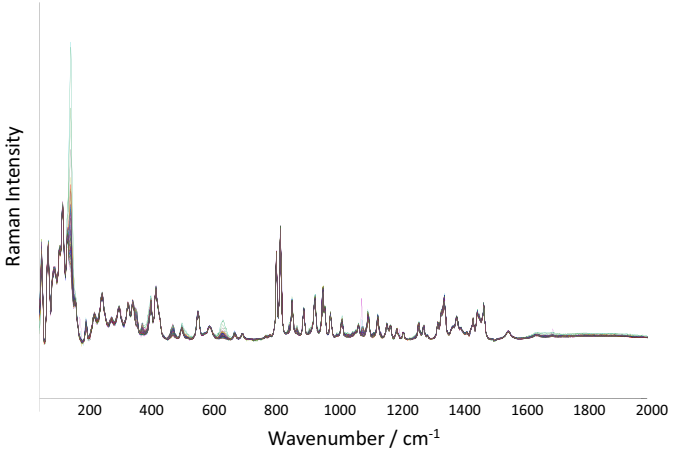

Fig. 5. Transmission Raman spectra of the 150 samples from the DOE dataset. The spectra were baseline corrected and normalised to the most intense band. (The band at ~620 cm^{-1} originates from capsule shell ($Ti:O_2$)).

Table 2
The concentration of API was successfully predicted for samples containing two different generations of API; in each case two identical samples were available.

Sample	API (w/w)	Prediction (% nominal)	Lactose (w/w)	Talc (w/w)	Starch (w/w)	Sum (w/w)
Gen1.1	0.769	102.5	0.0791	0.0940	0.0781	1.0202
Gen1.2	0.774	103.2	0.0902	0.0948	0.0785	1.0375
Gen2.1	0.738	98.4	0.0661	0.0741	0.0695	0.9477
Gen2.2	0.766	102.1	0.0759	0.0778	0.0833	1.0033
Nominal	0.750	–	0.0825	0.0838	0.0838	1.0000

the spectra of all samples; the capsules were oriented horizontally and measured through the white part of the capsule (capsule body) by illumination from below and Raman spectra collection from the top of the capsules. In this analysis the entire spectral range was used (~50–2500 cm^{-1}). Of particular note is the intense low wavenumber region of the Raman spectrum providing access to the polymorph region of the spectrum where phonon modes are present. This is a region analogous to that obtained with THz spectroscopy (albeit with different selection rules) providing additional potential capability to characterise crystalline structure of the sample in parallel with the molecular structure information, the latter available through the standard (higher frequency) molecular finger print region of the Raman spectrum.

PCA was performed on the complete data set. The first loading vector (containing the greatest variation) can be attributed to changes in the relative amount of API and lactose, which are still relatively highly (anti)correlated (see Fig. 2). Plots of the score on this vector versus API or lactose concentration show a high degree of correlation.

Subsequently, partial least squares (PLS1) regression allowed the concentration of the API and excipients to be modelled more accurately, using four latent variables (see Fig. 6). The RMSEP for

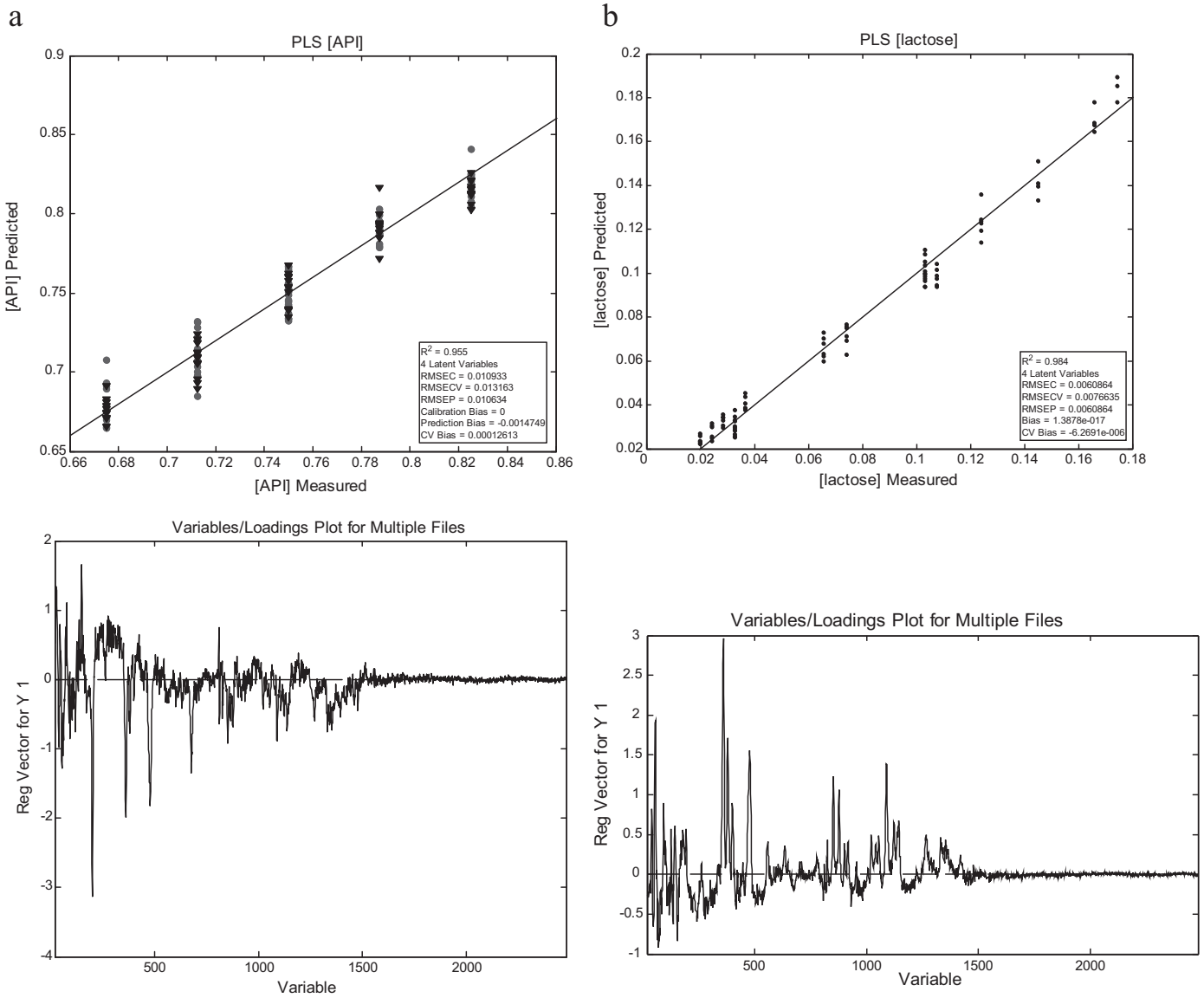

Fig. 6. The results of PLS analysis of the main DoE sample set comprising 140 capsules for individual sample components (a) API, (b) lactose, (c) talc and (d) starch.

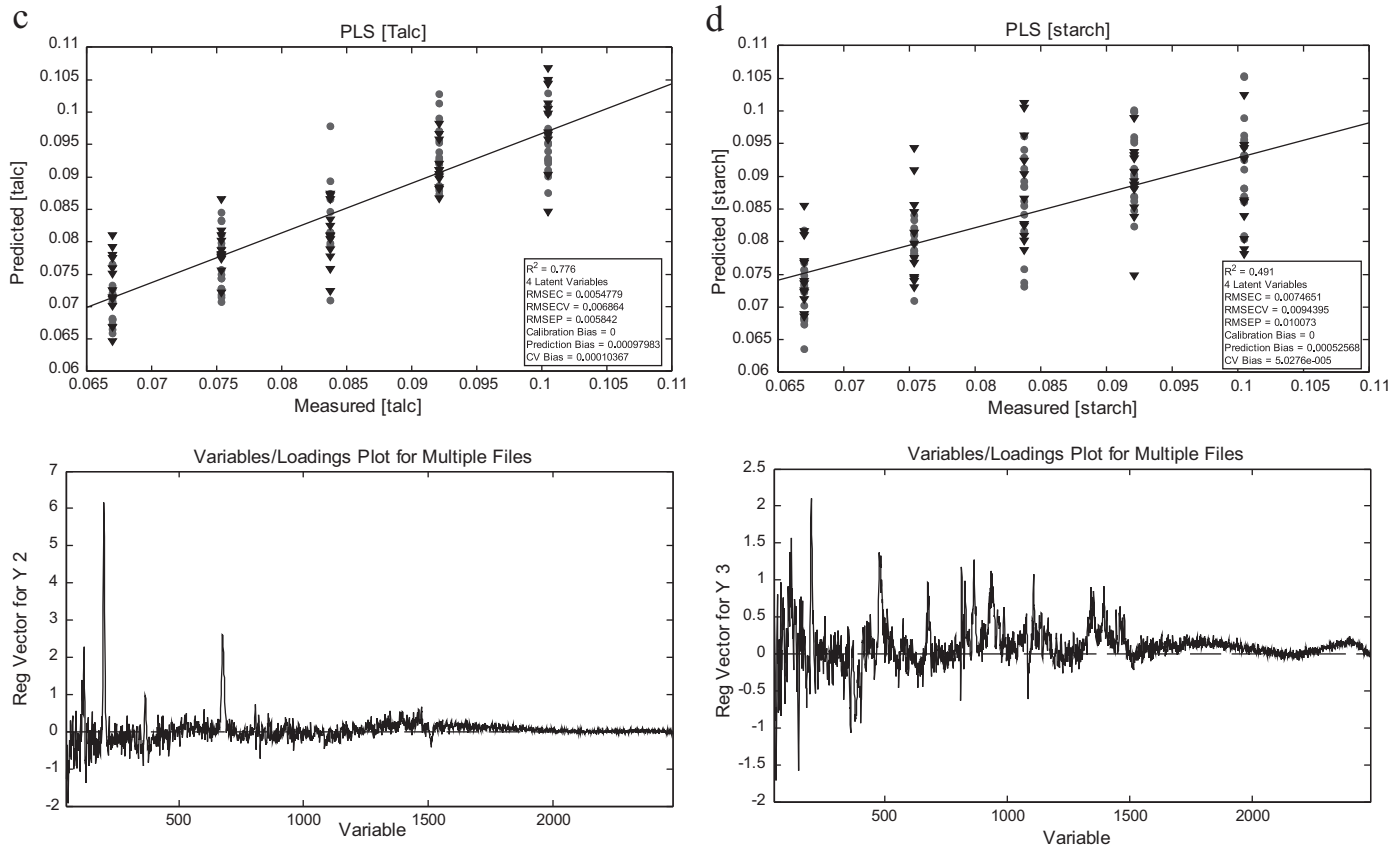

Fig. 6. (*Continued*)

Table 3
Prediction of API concentration of 6 capsule presentations with different fill weights and capsule colour combinations using the model developed on the main sample set (75 mg API/capsule).

Contents (mg API/capsule)	Predicted (w/w)	Prediction (nominal%)
75	0.767	102.3
100	0.749	99.9
150	0.771	102.8
200	0.759	101.2
225	0.750	100.0
300	0.683	91.1

the prediction set was 1.1%, w/w for the API, corresponding to 1.5% relative error. This value was similar to that found during the cross validation step (RMSECV = 1.3%, w/w). A linear fit of the data gave a gradient of 0.955. Relative prediction errors for the excipients were greater (ca. 5–15%).

3.2. Applicability

Two production samples containing different generations of API were tested using the model developed on the main sample set. Results are listed in Table 2. The concentration of API was successfully predicted in both cases; no difference was evident between different generation APIs.

3.2.1. Applicability – capsule presentations

Samples from the available 6 capsule presentations shown in Table 3 were assessed using the model developed on the main sample set (100 mg fill weight, 75 mg API/capsule). The test samples had an identical composition to that used to develop the model but with different capsule fill weights (100–400 mg) and different colour of capsule shell (capsules were either red and white or just red). The concentration of API was predicted within a relative error of typically 2–3% irrespective of their weight or capsule colour (see Table 3). However for the largest capsule (fill weight of 400 mg) a deviation from expected of ~9% was observed. This could be a result of a measurement issue, although this may also be related to reduced energy through the capsule and linearity of detector response. Overall this indicates good robustness of the approach and the potential for developing accurate calibration methods for a variety of capsule presentations based on a model developed for one single subtype, although further, more thorough, investigation is required to substantiate this result.

3.2.2. Prediction of two production samples

As a test, the PLS regression models for API, Lactose, Talc and Starch were used to predict two production samples (see Table 2). The "true" values were not known for all components of these capsules, only the expected nominal. The sum total of components in each capsule was within 5.2% of the expected figure of 100%. This also highlights the potential of predicting excipient concentrations, where important.

4. Conclusions

This study has demonstrated experimentally that the transmission geometry of Raman spectroscopy is well suited for the non-invasive probing of pharmaceutical capsules. The measured relative root mean square error of prediction of the concentration of API for the main data set was 1.5% with a 5 s acquisition time. It was also feasible to measure the concentration of the three low-level excipients in the formulation and the robustness of the model

to API prepared via different processes. The technique's insensitivity to the amount of material held within the capsules (i.e. the fill weight) was demonstrated through the capability of a single model (built using capsules at one fill weight) to accurately predict API concentration of capsules with varying fill weights (up to 4 times that of the samples in the model). There are significant benefits within the pharmaceutical industry to this approach; both in terms of resource requirements for method development and maintanance but also in terms of validation and regulatory activities.

References

[1] P. Matousek, A.W. Parker, Bulk Raman analysis of pharmaceutical tablets, Appl. Spectrosc. 60 (2006) 1353–1357.
[2] P. Matousek, I.P. Clark, E.R.C. Draper, M.D. Morris, A.E. Goodship, N. Everall, M. Towrie, W.F. Finney, A.W. Parker, Subsurface probing in diffusely scattering media using spatially offset Raman spectroscopy, Appl. Spectrosc. 59 (2005) 393–400.
[3] P. Matousek, M.D. Morris, N. Everall, I.P. Clark, M. Towrie, E. Draper, A. Goodship, A.W. Parker, Numerical simulations of subsurface probing in diffusely scattering media using spatially offset Raman spectroscopy, Appl. Spectrosc. 59 (2005) 1485–1492.
[4] M.V. Schulmerich, W.F. Finney, R.A. Fredricks, M.D. Morris, Subsurface Raman spectroscopy and mapping using a globally illuminated non-confocal fiber–optic array probe in the presence of Raman photon migration, Appl. Spectrosc. 60 (2006) 109–114.
[5] P. Matousek, Deep non-invasive Raman spectroscopy of living tissue and powders, Chem. Soc. Rev. 36 (2007) 1292–1304.
[6] N.A. Macleod, P. Matousek, Emerging non-invasive Raman methods in process control and forensic applications, Pharm. Res. 25 (2008) 2205–2215.
[7] N.A. Macleod, P. Matousek, Deep Noninvasive Raman Spectroscopy of Turbid Media, Appl. Spectrosc. 62 (2008) 291A–304A.
[8] B. Schrader, G. Bergmann, Die Intensitat Des Ramanspektrums Polykristalliner Substanzen.I. Strahlungsbilanz Von Substanz Und Probenanordnung, Zeitschrift fur Analytische Chemie Fresenius 225 (1967) 230–247.
[9] J. Johansson, S. Pettersson, S. Folestad, Characterization of different laser irradiation methods for quantitative Raman tablet assessment, J. Pharm. Biomed. Anal. 39 (2005) 510–516.
[10] J. Johansson, A. Sparen, O. Svensson, S. Folestad, M. Claybourn, Quantitative transmission Raman spectroscopy of pharmaceutical tablets and capsules, Appl. Spectrosc. 61 (2007) 1211–1218.
[11] C. Eliasson, N.A. Macleod, L. Jayes, F.C. Clarke, S. Hammond, M.R. Smith, P. Matousek, Non-invasive quantitative assessment of the content of pharmaceutical capsules using transmission Raman spectroscopy, J. Pharm. Biomed. Anal. 47 (2008) 221–229.
[12] D.M. Haaland, E.V. Thomas, Partial least-squares methods for spectral analyses 1. Relation to other quantitative calibration methods and the extraction of qualitative information, Anal. Chem. 60 (1988) 1193–1202.