JOURNAL OF RAMAN SPECTROSCOPY
J. Raman Spectrosc. 2004; **35**: 409–417
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jrs.1160

JRS

Development of sampling methods for Raman analysis of solid dosage forms of therapeutic and illicit drugs

Steven E. J. Bell,[1]* J. Renwick Beattie,[1] John J. McGarvey,[1] K. Laota Peters,[2] N. M. S. Sirimuthu[1] and S. James Speers[2]

[1] School of Chemistry, Queen's University, Belfast BT9 5AG, Northern Ireland
[2] The Forensic Science Agency of Northern Ireland, 151 Belfast Road, Carrickfergus BT38 8PL, Northern Ireland

Received 10 November 2003; Accepted 22 December 2003

The results of a study aimed at determining the most important experimental parameters for automated, quantitative analysis of solid dosage form pharmaceuticals (seized and model 'ecstasy' tablets) are reported. Data obtained with a macro-Raman spectrometer were complemented by micro-Raman measurements, which gave information on particle size and provided excellent data for developing statistical models of the sampling errors associated with collecting data as a series of grid points on the tablets' surface. Spectra recorded at single points on the surface of seized MDMA–caffeine–lactose tablets with a Raman microscope (λ_{ex} = 785 nm, 3 μm diameter spot) were typically dominated by one or other of the three components, consistent with Raman mapping data which showed the drug and caffeine microcrystals were ca 40 μm in diameter. Spectra collected with a microscope from eight points on a 200 μm grid were combined and in the resultant spectra the average value of the Raman band intensity ratio used to quantify the MDMA: caffeine ratio, μ_r, was 1.19 with an unacceptably high standard deviation, σ_r, of 1.20. In contrast, with a conventional macro-Raman system (150 μm spot diameter), combined eight grid point data gave μ_r = 1.47 with σ_r = 0.16. A simple statistical model which could be used to predict σ_r under the various conditions used was developed. The model showed that the decrease in σ_r on moving to a 150 μm spot was too large to be due entirely to the increased spot diameter but was consistent with the increased sampling volume that arose from a combination of the larger spot size and depth of focus in the macroscopic system. With the macro-Raman system, combining 64 grid points (0.5 mm spacing and 1–2 s accumulation per point) to give a single averaged spectrum for a tablet was found to be a practical balance between minimizing sampling errors and keeping overhead times at an acceptable level. The effectiveness of this sampling strategy was also tested by quantitative analysis of a set of model ecstasy tablets prepared from MDEA–sorbitol (0–30% by mass MDEA). A simple univariate calibration model of averaged 64 point data had R^2 = 0.998 and an r.m.s. standard error of prediction of 1.1% whereas data obtained by sampling just four points on the same tablet showed deviations from the calibration of up to 5%.

KEYWORDS: pharmaceuticals; drugs; tablets; ecstasy; sampling

INTRODUCTION

The advantages of using Raman spectroscopy for the analysis of the solid dosage forms of pharmaceuticals have been thoroughly documented.[1–20] The spectra can be recorded with no sample preparation and the individual components each give a unique set of bands. The richness of the spectra means that quantitative data analysis, either by manually selecting and then measuring some of the bands due to each of the components or by using multivariate data analysis techniques [e.g. partial least squares (PLS)] is straightforward.[6,10–12,21,22] Since the data processing can readily be automated and no sample preparation is required, implementing high-throughput screening or rapid analysis might appear to be a simple matter of coupling an automated sampling system, such as motorized X–Y–Z stage, to any available Raman spectrometer and then running the data acquisition and analysis under software control. However,

*Correspondence to: Steven E. J. Bell, School of Chemistry, Queen's University, Belfast BT9 5AG, Northern Ireland.
E-mail: s.bell@qub.ac.uk
Contract/grant sponsor: BBSRC; Contract/grant number: JREI18471.
Contract/grant sponsor: Recovered Asset Fund (RAF), Financial Crime Unit, UK Home Office.

in practice the precision of the quantitative analysis is very sensitive to both the spectrometer configuration and sampling errors.

This paper concentrates on sampling methods but some discussion of the absolute sensitivity of the different spectrometers used is also included, because sensitivity is particularly important in high-throughput analysis, where the total data acquisition time required per sample will typically be set by the time needed to accumulate signals with appropriate signal-to-noise (S/N) ratios. Although with a high-efficiency spectrometer these times are now typically $\leq$1 min per sample and extension to a few minutes does not seem important if single samples are being studied, when hundreds or thousands of samples are being analysed the cumulative effect of increasing acquisition times is very significant.

The main problem in measuring the bulk composition of tablets or powders from Raman data is that the signals may not be representative of the composition of the sample as a whole because sampling errors will arise if the tablets are heterogeneous on length scales greater than the size of a typical focused laser spot.[4,9] Of course, with Raman microscopes the spot size can be as small as 1 μm and this is advantageous if one wishes to characterize the microstructure of the sample; for example, it can be used to focus on and identify the individual components in complex mixtures. However, in the present context, the aim is to extract bulk data and this would normally require as large a spot size as possible. In macroscopic dispersive Raman systems the spot sizes are typically smaller than hundreds of micrometres whereas in FT-Raman spectrometers the spot is normally hundreds of micrometers but even with these much larger spot sizes there are numerous reports citing problems with sampling heterogeneous materials as the principal limitation for quantitative analysis.[4,10,17,18,23,24] Here we investigated the effect of sampling errors on the accuracy and precision of quantitative Raman measurements on tableted pharmaceuticals using a macro-Raman spectrometer (spot diameter 75 and 150 μm) and complement these studies with measurements made using a Raman microscope system (3 μm spot), which is designed primarily for obtaining highly spatially resolved spectra rather than bulk data. The data from the microscope provide an extreme case which is useful for developing models that can be extended to the less clear-cut macro-Raman situation.

The samples investigated were both street-quality seized 'ecstasy' tablets and model ecstasy tablets, which were prepared in-house and are more closely comparable to conventional pharmaceuticals. These samples were chosen because our immediate aim is to develop Raman methods for rapid, high-throughput quantitative analysis of illicit drugs, but of course any method developed will be equally valid for the analysis of mainstream pharmaceuticals.

Batches of seized ecstasy tablets from different sources typically have a different chemical composition and we are interested in exploiting the fact by using Raman composition profiling to provide intelligence information on drug manufacture and distribution.[1–3] Pilot studies carried out in a research laboratory have been very successful but moving to routine implementation of the method in a forensic drug analysis unit means that a simple automated data acquisition and processing protocol is needed. In part this will involve defining exposure times, data analysis routines, etc., but one key aspect is to understand the sampling statistics. Appropriate sampling is particularly important in this application because the tablets can exhibit very different degrees of heterogeneity, depending on the expertise of the unknown manufacturer. In our previous work on seized ecstasy tablets, the effects of spatial inhomogeneity were reduced in an entirely conventional way; the tablets were mounted in a rotating holder and the laser was line focused on to them so that scattering from a large circular region of the face (area ca 50 mm^2) was collected. It was assumed, rather than explicitly demonstrated, that the large surface area from which the signal was collected ensured that this spectrum was representative of the average composition of the sample.[1,3] Although this rotating tablet arrangement is widely used, it is typically only applied in situations where manually loading tablets one at a time is acceptable because it is difficult to combine sample rotation with automated high-throughput analysis of large numbers of tablets. In the work here, the approach taken was to mount sets of tablets on a motorized sample stage and to record data from a grid of points on each tablet (and from each tablet in the sample holder) by stepping the stage in the x and y directions. For each tablet the spectra from each of the grid points can be compared with one another (to determine the inhomogeneity) and/or summed to give an average spectrum that can be used to determine the overall sample composition.

Since, in general, it would be expected that the largest number of grid points should give the lowest sampling error for composition measurements, the obvious way to reduce sampling errors would be to select enormous numbers of sampling points. However, increasing the number of grid points also increases the overhead time required for sample positioning and detector readout cycles. In addition, the read noise, which is normally insignificant in the most common detectors [cooled charge-coupled device (CCD) detectors], may become important when large numbers of small signals are read from different points on the same tablet.[9] The optimum solution is to acquire data at the smallest number of points that are consistent with achieving an appropriately low sampling error.

EXPERIMENTAL

Seized ecstasy tablets containing MDMA (*N*-methyl-3,4-methylenedioxyamphetamine) as the psychoactive compound, along with caffeine and lactose, were obtained under

license from Forensic Science Northern Ireland (FSNI) and were used as received. Model MDEA–sorbitol ecstasy tablets (MDEA = *N*-ethyl-3,4-methylenedioxyamphetamine, the second most commonly found variant after MDMA) with known drug content were prepared by grinding weighed amounts (ca 100 mg of sorbitol and the appropriate mass of MDEA, weighing accuracy ±0.2 mg) of the constituents (Aldrich) in an agate pestle and mortar for 1 min before pressing in a conventional steel die with an hydraulic press (5000 kg cm^{-2}, 30 s).

The Raman microscope, a JY-Horiba Labram HR system, was used at 785 nm (diode laser) with a 50× NIR-optimized objective, a 600 gr grooves mm^{-1} grating and a motorized x–y stage. The confocal aperture was set very wide (1000 μm) to maximize throughput and slits were set at 150 μm. Typical exposure times were 20 s per point.

An Avalon Instruments R1 benchtop Raman spectrometer with a 785 nm diode laser, échelle spectrograph and x–y–z stage was used for the macro-Raman experiments (spot diameter 150 μm) on seized tablets. Quantitative analysis of the MDEA–sorbitol tablets was carried out on an Avalon Instruments R3 spectrometer with integral fibre-optic probe (spot diameter 75 μm) and an external x–y–z sample stage. Multivariate data analysis was carried out using PLS-IQ within GRAMS 7.0.[25]

RESULTS

Figure 1 shows a coarse Raman map taken at 785 nm with a Raman microscope (8 × 8 grid, 10 μm spacing, 50× objective) of a small randomly chosen section of a seized MDMA tablet. Although the tablet contains three major components, MDMA, caffeine and the lactose, they are not homogeneously distributed on the microscopic scale; the MDMA and caffeine are distributed predominantly as ca 40 μm particles (determined by Raman mapping). In the small region mapped in Fig. 1 there is very little caffeine and the data cover a section of the boundary between two domains which, as the spectra from each region show (Fig. 1), are predominantly composed of either lactose or MDMA. Of course, mapping the spatial distribution of all the various components has now been made straightforward by modern Raman microscopes and Raman imaging systems,[9,26,27,28,] but is not the main concern here, primarily because the main interest is in the bulk composition parameters. Generating a full high-resolution map of the entire tablet would certainly allow the bulk composition to be determined and would even provide some extra information on the way in which these components are distributed within the tablets. However, the time and expense involved in mapping numerous tablets samples chosen from every seized batch of ecstasy (typically 30 tablets from each batch) could not be justified in terms of the extra information on the composition of the tablets that it would provide.

Since full mapping is impractical, the alternative is to sample the overall composition by combining data taken at different grid points. It would not be rational to try to determine bulk composition data on the tablet by combining the spectra used to generate Fig. 1 because the total area covered by the data is only 80 × 80 μm, which is a tiny proportion of the total tablet area. Moreover, because both the laser spot diameter (or, more accurately, the sampled volume) and the spacing between grid points are smaller than the size of the domains, each of the points is dominated either by the signal from the lactose excipient or the MDMA and in regions away from the boundaries neighbouring points tend to be spectra of the same constituent.

When spectra are recorded on a coarser grid where the spacing is larger than any domains within the sample, spectra corresponding to different constituents or mixtures of constituents are randomly dispersed over all the points of the grid. For example, recording an 8 × 8 grid on a 200 μm spacing on the same sample as was used for Fig. 1 gave randomly dispersed data in which, because the spot

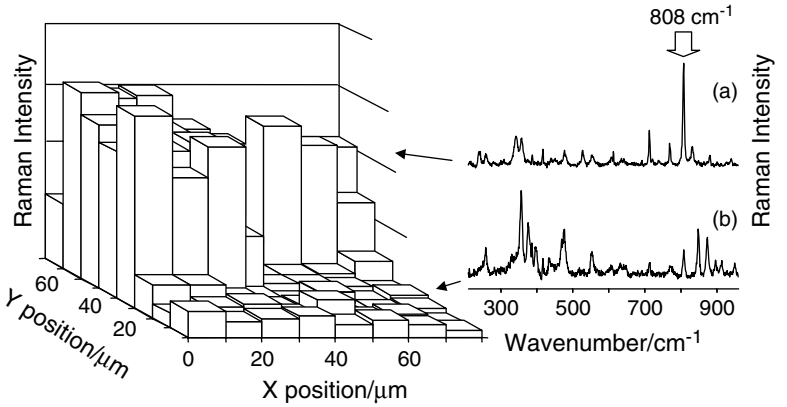

Figure 1. Plot of the peak area of the strongest MDMA band (808 cm^{-1}) in the Raman spectra of a seized ecstasy tablet recorded on an 8 × 8 grid with 10 μm spacing using a Raman microscope. The inset shows typical spectra recorded in (a) the high-drug (MDMA) region and (b) the low-drug (lactose) region.

JRS

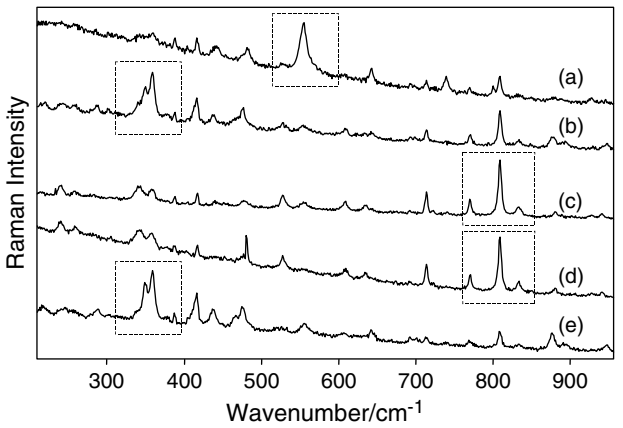

Figure 2. Raman spectra of a seized ecstasy tablet recorded with a microscope. The sequence shown is part a single row of data taken from an 8 × 8 grid on 200 μm spacing; (a) is predominantly caffeine, (b) and (e) are predominantly lactose [(b) also shows some MDMA], (c) and (d) are MDMA.

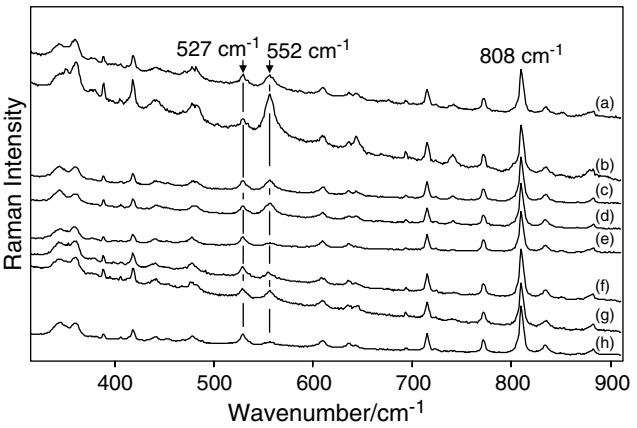

Figure 3. Raman spectra of a seized ecstasy tablet created by subdividing 64 spectra (microscope, 50× objective, 8 × 8 grid, 200 μm spacing, 20 s per point) into sets of eight [(a)–(b)] and combining them. The variation in the intensity ratio of the strongest caffeine band (552 cm^{-1}) and the closest MDMA band at (527 cm^{-1}) is readily apparent.

size had not changed, most of the spectra still showed substantial contributions of just one constituent. Figure 2 shows a sequence of spectra from a single row on the grid that shows the discrete nature of the spectra and the random occurrence of the characteristic spectra of the three components in the tablet. Under these acquisition conditions, changing the concentration of a given constituent will only change the probability of finding points where that constituent dominates the spectrum but will not dramatically change the spectra themselves.

For the purposes of composition profiling for drug intelligence work, the ratio of the strongest bands due to the MDMA and other major constituents/excipients is normally recorded. The drug : excipient ratio is then used as one of the key properties to distinguish between different batches. However, under the above conditions the data are effectively 'quantized' and caffeine, for example, dominates the spectra recorded at just seven randomly-distributed points of the 64 point (200 μm spacing) grid. This is important because, while it is essential that the relative concentrations of drug and caffeine are measured accurately, sampling errors as small as inclusion or omission of a single caffeine-dominated grid point will give changes in the drug : caffeine ratio (e.g. 6/64 vs 7/64) that may be as large as the batch-to-batch variation which is being sought.

To demonstrate the effect of sampling different numbers of points on the bulk composition statistics obtained from the same experimental data, the 64 spectra from the 8 × 8 grid were subdivided into smaller sets and then averaged. The eight averaged spectra created from each of the sets of eight is shown in Fig. 3 and it is clear that there are very significant sampling errors if only eight grid points are co-added to give a 'representative' spectrum of the tablet. Most obviously, the drug : caffeine band intensity ratio varies enormously between each of the average spectra. Normally for quantitative work the strongest drug (808 cm^{-1}) and excipient bands (552 cm^{-1} in this region for caffeine) are measured,[1] but the intensity ratios of the caffeine band (552 cm^{-1}) and the closest drug band (at 527 cm^{-1}) are easier to judge by eye and are used throughout this paper. This intensity ratio varies from ca 3 : 1 in Fig. 3(b) to ca 1 : 3 in Fig. 3(e) and (h). The extreme deviation of the spectrum in Fig. 3(b) simply reflects the accidental inclusion of several caffeine-rich points in the average, whereas the low caffeine levels in Fig. 3(e) and (h) reflect the opposite. Clearly, more than eight grid points are needed to reduce the sampling error but even averaging the 64 spectra from the grid into a pair of traces each containing data from 32 grid points does not eliminate the sampling error, as shown in Fig. 4, although the difference between the two spectra is much smaller than was apparent in the data from sets of eight points and it simply reflects additional caffeine, as shown by the difference spectrum [Fig. 4(c)]. For this sample, even if 64 sampling points are used, it is clearly the sampling error, rather than the S/N ratio, which will determine the accuracy of the measurement of the MDMA: caffeine ratio.

Of course, although the microscope is not primarily designed to be used for bulk measurements, it does have an integral x–y stage, so it is potentially possible to use it for bulk characterization and certainly if this was to be attempted the conditions could be improved over those described above. For example, recording data at even larger numbers of points or use of a lower power objective to increase the laser spot size would both reduce sampling errors. However, the former method increases the sample positioning/detector readout times and the latter option was not pursued because of the loss in absolute signal that results from use of lower power objectives. An alternative method of increasing the sampled area is to raster the beam

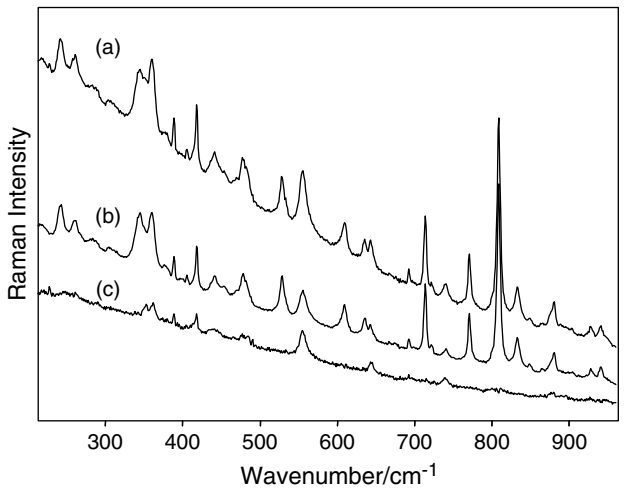

Figure 4. (a), (b) Comparison of Raman spectra of a seized ecstasy tablet created by dividing 64 spectra (microscope, 50× objective, 8 × 8 grid, 200 μm spacing, 20 s per point) into two sets of 32 and combining them. (c) Difference spectrum [(a)−(b)] scaled to remove the MDMA features and show that the residual difference is simply due to caffeine.

on the sample and the microscope does have this facility. However, all these approaches ignore the fact that the nature of the microscope means that it is excellent for characterizing small samples but cannot at the same time be optimum for bulk measurements, which are best carried out with a macro-Raman system.

Figure 5 compares spectra taken at points on the seized tablets with similar composition with the micro- and macro-Raman instruments. The S/N ratio achieved with a single 2 s accumulation on the macro system is higher than that for a 20 s accumulation with the microscope. The higher S/N ratio levels and larger spectral range covered in a single exposure mean that the macro-Raman spectrometer would be certainly be preferable for high-throughput measurements where total accumulation times are important. The difference in performance between the two spectrometers is probably not surprising since the macroscopic system has been optimized for operation at a single wavelength whereas the microscope must perform over a range of excitation wavelengths and it will invariably be necessary to compromise performance at one wavelength to achieve acceptable results at another. For example, the S/N ratio level of the microscope is much better at 633 than at 785 nm. In addition, it is generally expected that for transparent samples macro-Raman spectrometers will give higher signals than microscope-based systems because of the increased scattering volume.[9]) Although the MDMA tablets are opaque rather than transparent, they do not contain any components which absorb at 785 nm and some penetration into the body of the tablets would be expected, with a corresponding increase in signal. Diffuse reflectance measurements of microcrystalline cellulose pellets showed a depth of penetration of ≥1 mm with 700–1100 nm light.[29]

With the macroscopic system, the spot diameter of ca 150 μm is larger than the individual domains within the tablets, so that there is very much less 'quantization' of the data, although there is still detectable variation between data from different grid points. Figure 6 shows averaged (eight grid point) data for a tablet from the same seized batch as was used in the microscope experiments. In this case the differences are much less pronounced than those obtained from the microscope and even averages from just eight grid points are so similar that the differences are only apparent on close inspection. When the pair of averaged spectra from 32

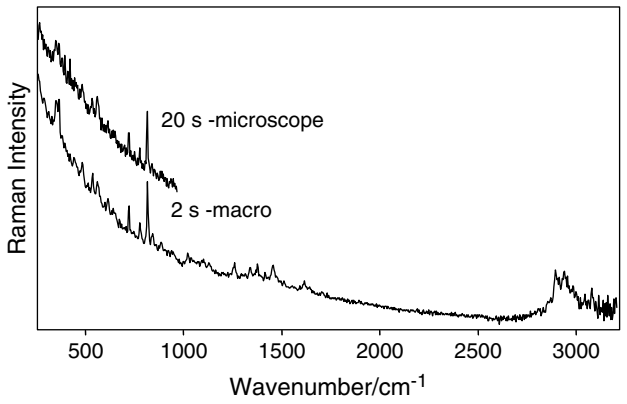

Figure 5. A comparison of the raw data obtained from ecstasy tablets using the (a) micro- and (b) macro-Raman systems (data from areas on the tablets with similar composition were chosen for this comparison). The S/N ratio is higher with the macroscopic system, despite a reduction in acquisition time from 20 to 2 s. The different spectral ranges that are covered in a single exposure by both systems are also shown.

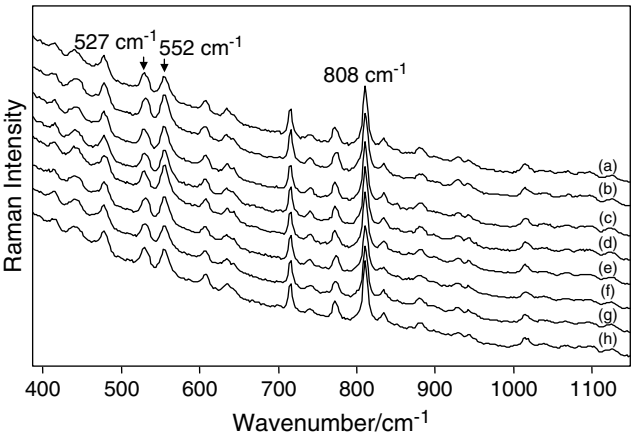

Figure 6. Raman spectra of a seized ecstasy tablet created by subdividing 64 spectra (macro Raman system, 8 × 8 grid, 0.5 mm spacing, 2 s per point) into sets of eight [(a)−(b)] and combining them. The variation in the intensity ratio of the strongest caffeine band (552 cm^{-1}) and the closest MDMA band at (527 cm^{-1}) is much smaller than that in Fig. 3.

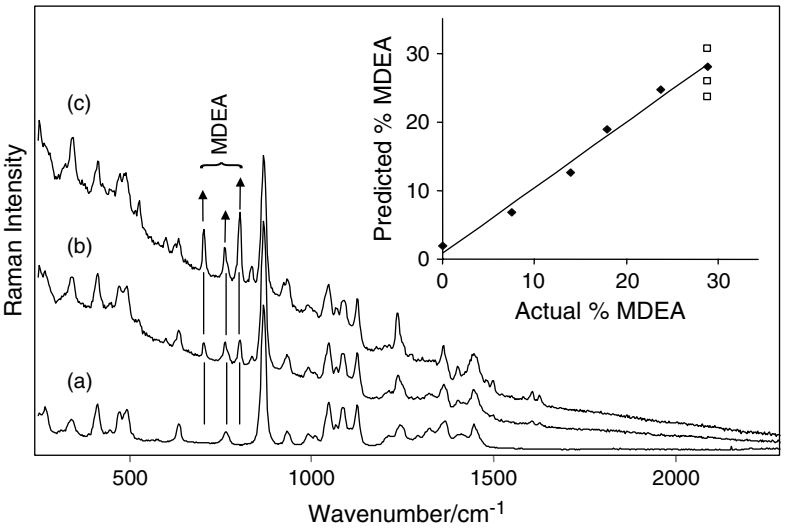

Figure 7. Representative Raman spectra (averages of all the spectra from an 8 × 8 grid on 0.5 mm spacing, 75 μm spot diameter) of a series of model MDEA–sorbitol tablets prepared with (a) 0, (b) 14 and (c) 30% by mass of MDEA. Inset: calibration plot obtained from the full set of spectra (64 point acquisition, solid points). The open squares are from experiments where the total accumulation time was unchanged but data from just four points were averaged.

grid points (not shown) are compared with each other, there is no difference large enough to detect with the naked eye. Subtraction of the two spectra with a scaling factor (0.98) that results in complete cancellation of the MDMA bands does show a small residual caffeine signal. This residual signal is three times larger than the noise but this is mainly a reflection of the high S/N ratio, since its absolute magnitude is <10% of that in the parent spectra.

Sampling errors are more straightforward to detect in quantitative analysis of known model systems. Figure 7 shows the spectra of a series of model MDEA–sorbitol tablets prepared with various drug contents in the range generally encountered in seized tablets. These tablets were prepared to test the accuracy of Raman spectroscopy in determining the drug content of ecstasy tablets. Sorbitol was arbitrarily chosen as the excipient because it is often encountered in seized ecstasy tablets but the same approach should work equally well for other excipients. Since the combination of lower sampling error, higher S/N ratio and resolution makes the macroscopic system clearly more appropriate for determining average properties of solid dosage form pharmaceuticals, these experiments were confined to the macroscopic spectrometer. The spectra shown in Fig. 7 are averages of all the spectra from an 8 × 8 grid on 0.5 mm spacing i.e. the same sampling conditions as were used for the seized tablets discussed above. The accumulation time of 1 s per point was set to give a total acquisition time of 64 s, the same total time as the 8 points ×8 s used for the each of the spectra shown in Fig. 6. The simple calibration plot obtained from PLS1 of these spectra (pre-processed by scaling on the strongest sorbitol band, mean centering, taking the first derivative (Savitsky–Golay, 15 points) and including the spectral range 678–818 cm^{-1}) is shown as an insert in Fig. 7. The model is entirely as expected with a univariate system; using just a single principal component gives a calibration plot with $R^2 = 0.988$ and an r.m.s. error of 1.1%. These tablets were carefully prepared and had a higher degree of homogeneity than typical seized tablets, but even with these tablets and a large spot diameter it is possible to introduce sampling errors. For example, averaging the spectra from four points on each tablet gives the data shown as open squares on the inset in Fig. 7, where it is clear that the sampling errors are sufficiently large to give errors up to ca 5%.

DISCUSSION

In the quantitative data of the type shown by the MDEA–sorbitol measurements, the sampling error simply introduces additional scatter in the calibration plots whose source is obvious because known samples are being used. In the model samples here, the use of 64 grid point sampling and a total acquisition time of ca 1 min was sufficient to give an acceptable calibration ($R^2 = 0.988$, r.m.s. error of 1.1%) in an entirely straightforward way but sampling four grid points would not be sufficient even if the S/N ratio was kept the same by increasing the accumulation time per point. In the case of the seized ecstasy tablets, the effect of sampling errors will inevitably be much less obvious because it is expected that there will be an unknown but possibly large variation in the composition of the tablets even within a single batch. Since sampling errors and genuine tablet-to-tablet variation would both be expected to give simple scatter about mean values for the composition parameters, the observation of a

broad distribution in the results could be due either to genuine tablet-to-tablet variation or sampling errors introduced by sampling too few points. In many cases checking that an acceptably low sampling error has been obtained is all that will be necessary for quantitative analysis but it is useful to have a conceptual and at least semi-quantitative model of the underlying statistical basis of the sampling.

With solid tablets composed of a blend of microparticulate constituents that are evenly distributed throughout the tablet, intuitively it would be expected that the sampling error will decrease as the size of the particles decreases, more points are sampled or the diameter of the laser spot is increased, but the relationship between these factors is not obvious. However, the statistical basis for determining the proportion of objects in a population with a particular property by testing a random sample is well understood and can be applied to Raman measurements on solid samples relatively easily.

The easiest data to model are those from the microscope because this approaches an ideal system, i.e. one in which the tablets are made of a compressed mixture of discrete particles of drug and a single excipient which are probed by a laser spot that is much smaller than the particles. In this case the property that is measured is essentially whether the sampled spot is drug or excipient. In standard nomenclature, the problem is to determine π, the proportion of the members of a population of objects that have a particular property (e.g. being drug rather than excipient), from a sample, size n (i.e. number of grid points probed), of that population. The proportion of objects in the sample with the property is denoted p and its value is determined by the experiments in which a series of grid points are probed. If the process is carried out numerous times, the long-run average value of p, μ_p, will be equal to the population value, π. However, as is most clearly evidenced in Fig. 3, the values of p determined in different experiments will not always be the same; it is known that they will be distributed about π and the standard deviation in the results, σ_p, will be given by[30]

$$\sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} \quad (1)$$

Equation (1) gives a statistical basis for predicting the uncertainty in the drug : excipient ratio measured in experiments which record the number of 'drug' points/total number of points sampled. Importantly, under these ideal conditions σ_p can be determined rigorously from Eqn (1) and the distribution will be approximately normal, provided that the rule of thumb conditions that np and $n(1-p)$ should both be ≥ 5 are obeyed. For example, in the case of tablets with 25% drug being probed by an infinitely small spot, the long-run average of p will always be 0.25, but if $n = 8$ (i.e. eight grid points used), σ_p will be 0.15 so the average values determined from experiments with just eight points will display considerable scatter; increasing the number of grid points to 64 reduces σ_p by a factor of $\sqrt{8}$ to 0.054.

Even in the less than ideal case of the three-component tablet shown in Fig. 3, the proportion of 'caffeine' points in the larger sample (presumably close to the population average) is 7/64 (0.109); for this tablet σ_p for experiments where $n = 8$ would be 0.110. The measured drug : caffeine peak-height ratios (μ_r) for the set of eight averaged spectra shown in Fig. 4 is 1.19 with $\sigma_r = 1.02$. To compare this value to the calculated σ_p it is necessary to correct for the relative Raman scattering cross-sections of the two components, but this is straightforward using the mean values from both sets of measurements (average probability of encountering caffeine $\mu_p = 7/64$, global average of appropriate band intensities $\mu_r = (1.19)$. Following this conversion gives an expected σ_r of 1.20, which compares well with the experimentally-determined value of 1.02. Presumably the slightly smaller than predicted experimental value arises because the spectra, although predominantly one or other of the components, do contain minor contributions from the others [see Fig. 2(b), for example]. However, the important point is that the uncertainty in the measured ratio has a purely statistical basis and it cannot be improved by, for example, increasing the S/N ratio in the data. It can be improved by sampling a larger number of grid points, but the improvement scales as the square root of the number of points, so an order of magnitude improvement would need 100× more points. Increasing the number of grid points by this huge factor is not acceptable because, even though the effect of dramatically cutting the accumulation time per point is negated by adding the data from more points together (if the increased read noise is ignored), the total acquisition time increases dramatically since it becomes controlled by the overhead time for sample positioning and detector readout.

In the case of the macroscopic spectrometer where the laser spot was larger than the individual drug particles, the observed signal even at a single grid point at least partly reflected the relative abundance of drug and excipient in the sample. Although it is slightly more difficult to treat the statistics of this situation than the simple yes/no drug/excipient signals, it is possible to make semi-quantitative estimates using a physically reasonable model as follows.

If the population of individual objects that are to be sampled are envisaged as idealized circular areas on the tablet surface whose dimension is that of the drug or excipient particles, the number of individual objects sampled will depend on the relative size of the drug particles and the laser spot diameter; for example, a spot four times the diameter of the particles in the tablet will mean that ca 14 of these areas are probed in a single accumulation at one grid point. The effective number of samples studied under those conditions would therefore be 14× the number of grid points (as opposed to one per grid point when a small spot is used) and this increase in n should reduce the uncertainty in the measured ratio, r. The experimental data for the ecstasy tablet which contains ca 40 μm particles show that the distribution of r values in the averaged spectra

from eight grid points is indeed much lower than that for the eight grid point microscope data ($\mu_r = 1.47$, $\sigma_r = 0.16$), but the improvement is very dramatic ($\sigma_r = 1.20$, micro, down to 0.16, macro). Even if we assume that the laser spot diameter has been underestimated and it is 160 rather than 150 μm, then with a particle size of 40 μm the best case improvement would be expected to be $\sqrt{14}(= 3.7)$ whereas the actual improvement is closer to 8-fold. A reduction in σ_r of this size would require an increase in sampled objects by a factor of ca 64, rather than 14. The origin in this apparent anomaly is that the simple model only includes areas while the laser probes a volume of the sample. In the microscopic system the confocality restricts probing to a surface layer of the order of tens of micrometer, so the measurements are confined to the top layer of particles and the statistics are indeed essentially those of measuring the relative areas of the surface composed of drug or excipient. With macroscopic illumination the increased spot diameter allows a larger area of the surface to be probed while the longer focal length of the objective gives increased illumination/collection depth. In the macroscopic system used here the depth of focus is >500 μm and, even though scatter effects will clearly reduce the penetration/sampling depth, it is clear that the volume of sample that is probed is large compared with the volume of the drug particles and it is this which gives the dramatic increase in n that reduces σ_p (and therefore σ_r) to acceptable levels. In the sample here, the increase from 14 probed particles from the 'area' model to 64 in the probed volume implies penetration and collection from ca 4–5 layers of particles, i.e. 160–200 μm, which is physically reasonable for these non-absorbing samples.

This model, which allows n to be estimated from the particle size and spot diameter and can be used to calculate the expected uncertainty due to sampling errors, also provides a useful framework which supports some general observations about sampling. For example, it emphasizes the fact that the number of objects sampled by instruments with different spot sizes is not a simple function of the relative spot diameters. In the case described above, the ratio of the 3 μm microscope spot to the 150 μm macro Raman spot gives an increase in area sampled of $2500\times(50^2)$, but this does not mean that the number of particles sampled, n, increases by 2500×. The fact that the particles are 40 μm in diameter means that (ignoring depth penetration) n increases from 1 per grid point on the microscope to 14 per grid point with the macro system. Moreover, changing to the microscope spot size down to 1 μm or up to 5 μm will not alter the fact that the spectra recorded at each grid point will essentially be those of a single component and so will have no effect on the sampling error. In fact, starting with a spot much smaller than the particles and then increasing its diameter will have little effect on the statistics until the diameter approaches that of the particles. Even after that point, the number of sampled objects, n, increases as the (diameter)^2 and, since σ_p falls as $\sqrt{n}$, σ_p improves only linearly with spot diameter. The general conclusion is that the increase in n on increasing the spot size will be smaller than would expected on the basis of the simple relative area ratio although, as in the tablets used here, this may be partially compensated by increased depth of focus.

Of course, in practice the particle size in the tablets will not normally be known, so it is not possible to calculate *a priori* how many grid points should be used to give a desired σ for any given sample, but it is useful to understand the reasons underlying the very dramatic difference seen in data recorded with micro- and macroscopic sampling and to see where the limits lie. In particular, the data illustrate how much more difficult it will be to generate data that properly reflect the bulk composition of a tablet with a microscope.

On a day-to-day basis, the simplest indicator that sampling errors are acceptably low is to take two or more sets of data and to compare them. This is particularly simple if a grid has been recorded because half the grid points can be summed and compared to the other half (as in Fig. 4). Provided that the difference is lower than a given threshold, the total of these can then be used as the global average spectrum in subsequent processing.

In the case of ecstasy profiling, this simple confirmation that sampling errors are lower than a specified value is sufficient because we ultimately reduce all the grid data from each tablet to a single (albeit representative) single spectrum. Since the tablets are mounted in a holder than can accommodate 30 tablets, the entire process can be automated, a grid of points is scanned for each tablet and saved before the sample stage is moved onto the next tablet. The combined data from ca 30 tablets taken from each batch are then used to characterize the batch, most obviously to obtain average composition parameters but also to confirm, for example, that it is not composed of a mixture of two or more different types of tablets.

CONCLUSIONS

It is clear that analysis of seized ecstasy tablets using a macro-Raman system in which the spectrum from each tablet is recorded as a sequence of grid points can give data with acceptably low sampling errors even when <100 grid points are sampled. Moreover, the sampling error can be determined by comparison of sub-sets of data extracted from the grid. This opens up the possibility of high-throughput automated analysis since it means that multiple tablets can be mounted in a simple holder and each can analysed in turn (i.e. a grid of sampling points on each of a grid of tablets can be used) without the need to mount each tablet in turn within a moving or rotating sample holder. The high optical efficiency of the macro-Raman system and the relatively low number of grid points that need to be probed mean that total acquisition times of ca 60 s per tablet are possible. The statistical model developed here allows the expected sampling error for any tablet to be estimated from consideration of the particle

size and beam diameter/penetration depth, but its main strength is that it shows explicitly the relationship between the various parameters that together reduce the precision of the quantitative analysis by increasing sampling error.

Finally, although the data shown here were of illicit drug samples, the same observations also apply tableted therapeutic drugs and indeed to any microheterogeneous samples (e.g. ointments, gels and transdermal patches) where it is necessary to determine bulk composition parameters from Raman spectra.

Acknowledgements

The authors are pleased to acknowledge the BBSRC's (grant No. JREI18471) support for purchase of the Raman microscope. N.M.S.S. acknowledges an EPSRC/Royal Society of Chemistry studentship. The financial support of the Recovered Asset Fund (RAF), Financial Crime Unit, UK Home Office is also gratefully acknowledged.

REFERENCES

1. Bell SEJ, Burns DT, Dennis AC, Matchett LJ, Speers JS. *Analyst* 2000; **125**: 1811.
2. Bell SEJ, Burns DT, Dennis AC, Speers JS. *Analyst* 2000; **125**: 541.
3. Bell SEJ, Barrett LJ, Burns DT, Dennis AC, Speers SJ. *Analyst* 2003; **128**: 1331.
4. Bugay DE. *Adv. Drug Deliv. Rev.* 2001; **48**: 43.
5. Jones DS, Brown AF, Woolfson AD, Dennis AC, Matchett LJ, Bell SEJ. *J. Pharm. Sci.* 2000; **89**: 563.
6. Dyrby M, Engelsen SB, Norgaard L, Bruhn M, Lundsberg-Nielsen L. *Appl. Spectrosc.* 2002; **56**: 579.
7. Kontoyannis CG, Orkoula M. *Talanta* 1994; **41**: 1981.
8. Kontoyannis CG. *J. Pharm. Biomed. Anal.* 1995; **13**: 73.
9. McCreery RL. *Raman Spectroscopy for Chemical Analysis*. Wiley: New York, 2000.
10. Niemczyk TM, Delgado-Lopez MM, Allen FS. *Anal. Chem.* 1998; **70**: 2762.
11. Pratiwi D, Fawcett JP, Gordon KC, Rades T. *Eur. J. Pharm. Biopharm.* 2002; **54**: 337.
12. Ryder AG, O'Connor GM, Glynn TJ. *J. Forensic Sci.* 1999; **44**: 1013.
13. Ryder AG. *J. Forensic Sci.* 2002; **47**: 275.
14. Severdia AG, Siek K. *Appl. Spectrosc.* 2002; **56**: 545.
15. Skoulika SG, Georgiou CA. *Appl. Spectrosc.* 2001; **55**: 1259.
16. Skoulika SG, Georgiou CA. *Appl. Spectrosc.* 2003; **57**: 407.
17. Stephenson GA, Forbes RA, Reutzel-Edens SM. *Adv. Drug Deliv. Rev.* 2001; **48**: 67.
18. Roberts SNC, Williams AC, Grimsey IM, Booth SW. *J. Pharm. Biomed. Anal.* 2002; **28**: 1135.
19. Taylor LS, Langkilde FW. *J. Pharm. Sci.* 2000; **89**: 1342.
20. Al-Zoubi N, Koundourellis JE, Malamataris S. *J. Pharm. Biomed. Anal.* 2002; **29**: 459.
21. Ryder AG, O'Connor GM, Glynn TJ. *J. Raman Spectrosc.* 2000; **31**: 221.
22. Szostak R, Mazurek S. *Analyst* 2002; **127**: 144.
23. Taylor LS, Zografi G. *Pharm. Res.* 1998; **15**: 755.
24. Vergote GJ, Vervaet C, Remon JP, Haemers T, Verpoort F. *Eur. J. Pharm. Sci.* 2002; **16**: 63.
25. *GRAMS 7.0*. ThermoGalactic: Salem, NH, 2002.
26. Breitenbach J, Schrof W, Neumann J. *Pharm. Res.* 1999; **16**: 1109.
27. Clarke FC, Jamieson MJ, Clark DA, Hammond SV, Jee RD, Moffat AC. *Anal. Chem.* 2001; **73**: 2213.
28. Schlucker S, Schaeberle MD, Huffman SW, Levin IW. *Anal. Chem.* 2003; **75**: 4312.
29. Berntsson O, Burger T, Folestad S, Danielson LG, Kuhn J, Fricke J. *Anal. Chem.* 1999; **71**: 617.
30. Devore J, Peck R. *Statistics* (3rd edn). Duxbury Press: Pacific Grove, 1996.