Non-invasive, quantitative analysis of drug mixtures in containers using spatially offset Raman spectroscopy (SORS) and multivariate statistical analysis

Applied Spectroscopy (2012). *Accepted version of manuscript.*

William J. Olds[a]*, Shankaran Sundarajoo[a], Mark Selby[a], Biju Cletus[a], Peter M. Fredericks[a], Emad L. Izake[a].

[a]Discipline of Chemistry, Faculty of Science and Technology, Queensland University of Technology, 2 George Street, Brisbane, Queensland 4001, Australia. *Corresponding author: w.olds@qut.edu.au; Ph +61 7 3138 9076; Fax: +61 7 3138 9079.

Abstract

In this paper, spatially offset Raman spectroscopy (SORS) is demonstrated for non-invasively investigating the composition of drug mixtures inside an opaque plastic container. The mixtures consisted of three components including a target drug (acetaminophen or phenylephrine hydrochloride) and two diluents (glucose and caffeine). The target drug concentrations ranged from 5% to 100%. After conducting SORS analysis to ascertain the Raman spectra of the concealed mixtures, principal component analysis (PCA) was performed on the SORS spectra to reveal trends within the data. Partial least squares (PLS) regression was used to construct models that predicted the concentration of each target drug, in the presence of the other two diluents. The PLS models were able to predict the concentration of acetaminophen in the validation samples with a root-mean-square error of prediction (RMSEP) of 3.8% and the concentration of phenylephrine hydrochloride with an RMSEP of 4.6%. This work demonstrates the potential of SORS, used in conjunction with

multivariate statistical techniques, to perform non-invasive, quantitative analysis on mixtures inside opaque containers. This has applications for pharmaceutical analysis, such as monitoring the degradation of pharmaceutical products on the shelf, in forensic investigations of counterfeit drugs, and for the analysis of illicit drug mixtures which may contain multiple components.

Index Headings

Raman spectroscopy, SORS, PCA, PLS, pharmaceutical analysis, drug mixtures, forensic analysis, quantitative analysis.

Introduction

In recent times, there has been a marked increase in the use of Raman spectroscopy for forensic and national security applications.[1] To a large extent, this has resulted from the advent of new compact Raman instruments that can be used in the field to analyze a wide variety of target compounds. Such instruments have provided law enforcement officers and first-defenders with a reliable and rapid means of identifying drugs, explosives and many other hazardous substances.[1]

There has also been an increased use of Raman spectroscopy to combat the proliferation of counterfeit pharmaceutical products.[2-5] Counterfeit medicines generally pose as expensive hormones, steroids and anti-cancer drugs or, more commonly in developing nations, as treatments for life-threatening diseases such as malaria, tuberculosis and HIV/AIDS.[3, 4, 6] Raman spectroscopy can be used to screen pharmaceutical products at all stages of their production and distribution, to provide qualitative and quantitative information about the samples.[7, 8]

In many cases, the samples requiring analysis are mixtures. In pharmaceutical products, the active pharmaceutical ingredient (API) is usually combined with other compounds such

as binders, preservatives, colorants and other excipients.[7] In the case of illicit drugs, these are often diluted with cutting agents, which are common household products that appear similar to the drug, in order to increase drug dealers' profits.[9] The Raman spectra of mixtures are often very complex since they contain the spectra of all ingredients present in the sample. For this reason, it is necessary to use chemometric techniques to extract quantitative information from the mixture spectra. Multivariate methods can utilize information from the entire spectrum, rather than just a single characteristic peak, and are well suited to analysis of cluttered Raman spectra with many overlapping features.[7, 10]

Many groups have combined conventional Raman spectroscopy and multivariate analytical techniques. Ryder et al. analyzed two-part mixtures of cocaine and reported a partial least squares (PLS) regression model for predicting the concentration of cocaine in glucose.[11] The samples ranged in concentration from 0% to 100% of cocaine (by weight) and they achieved a root-mean-square error of prediction (RMSEP) of 2.3%. They later added caffeine, achieving an RMSEP of approximately 4% for these 3-part mixtures.[12] Fenton et al. created four separate PLS models for predicting the concentrations of four different illicit drug surrogates in a large set of sample mixtures.[13] The surrogates occupied between 20% and 100% of each sample, with the remainder comprising random proportions of up to 3 different cutting agents. RMSEP values of approximately 4% were attained for each of the surrogates. Eliasson et al. examined pharmaceutical capsules using transmission Raman spectroscopy.[10] Their PLS model was designed to use a much smaller range of API concentrations (27.0% to 33.0%) that would be expected in QA/QC applications, and they achieved prediction errors as low as ~1%.[10]

In this research, we demonstrate that the spatially offset Raman spectroscopy (SORS) technique enables quantitative analysis of drug mixtures while retained in their packaging. SORS was first described by Matousek et al.,[14] and is aimed at identifying compounds that

are concealed inside diffusely scattering and opaque (non-metallic) containers. Using SORS, the spectral features and fluorescence from the container (surface layer) can be suppressed, enabling the Raman spectrum of the contents (sub-layer) to be collected.[14-16] This is achieved by collecting Raman spectra from one or more locations that are spatially offset from the point of illumination on the container surface. Spectra captured at larger offsets will contain, in relative terms, higher contributions of Raman photons originating from deep within the sample. This effect arises from the wider lateral diffusion of photons emerging from greater depths in the sample, due to scattering of the photons within the sample.[14, 17] Raman signals contributed by the surface layer can then be removed by performing a scaled subtraction of a spectrum measured without a spatial offset. The resulting spectrum is a pure Raman spectrum of the concealed sample that is free from signals from the surface layer. SORS has been demonstrated for numerous applications, including the identification of active ingredients in tablets, analysis of pharmaceutical tablets inside plastic containers and for the identification of capsules through their shells and in blister packs,[3-5, 18, 19] as well as for several forensic and national security applications, including detection of liquid explosives and illicit drugs inside diffusely scattering containers.[2, 18, 20-23]

To date, the foregoing SORS studies have mainly focused on qualitative analysis, i.e. identification, of the concealed compounds. Further work is needed to clarify whether the technique may also be employed for quantitative analysis and to understand the levels of accuracy that it can be expected to deliver. In this paper, samples comprising a target drug mixed with two common diluents and concealed inside an opaque plastic container were screened using SORS. The target drugs included a relatively strong (acetaminophen) as well as a weaker (phenylephrine hydrochloride) Raman scatterer. The SORS spectra were analyzed using principal component analysis (PCA) and PLS to extract qualitative and quantitative information regarding the composition of the mixtures. These results are a

preliminary demonstration of the quantitative capability of the SORS technique, establishing its suitability for many of the applications mentioned above.

Materials and Methods

Sample preparation

Pure standards of acetaminophen (paracetamol), phenylephrine hydrochloride, and caffeine were purchased from Sigma-Aldrich. Anhydrous D-glucose was purchased from Chem-Supply.

For both drugs, 13 different mixtures were prepared, in which the drug concentrations ranged from 5% to 100% by weight (Table I). The remainder was composed of equal parts of caffeine and glucose by weight. The mixtures were made by weighing appropriate amounts of each component followed by thorough mixing with a spatula. Local regulations precluded the use of illicit drugs (e.g. cocaine, heroin). However a successful demonstration of the technique for these substances would indicate that the approach is viable for most other compounds, since they will have a unique Raman spectrum for their identification.

The mixtures were packed in a white opaque plastic polypropylene container of ~ 2 mm wall thickness. Each sample was measured in 3 replicates (at the same position) and the replicate measurements were used in the PCA and PLS models.

Reference Raman spectra were obtained for each pure compound by screening a small amount with the SORS instrument.

SORS instrumentation

The SORS instrument (Figure 1) was built in-house and is described in detail elsewhere.[18] An NIR laser source at 785 nm with ~ 400 mW of power and a ≈ 4 mm beam

diameter was used for excitation. To conduct SORS, each sample was placed into the plastic container and spectra were captured for spot and ring-shaped illumination. Ring illumination provided the 'spatially offset' measurement: the Raman signal was collected from a point in the central non-illuminated region of the ring, so that the offset distance was given by the ring radius (≈ 8 mm).[15, 16] The ring illumination was produced using an axicon lens. Subsequently, the illumination optics were adjusted to produce a spot and a second measurement was taken (zero offset). The spot measurement was akin to conventional backscattered Raman measurement since the illumination and collection zones overlapped. A scaled subtraction was then performed on the two spectra (ring spectrum minus spot spectrum) to obtain a spectrum of the concealed mixture.

The container with the concealed mixture was placed 60 mm in front of the collection system, equal to the front objective lens focal length (Figure 1). The collected light traversed several filters to suppress the residual scattered laser light and was directed into a 900 μm-diameter optic fiber bundle by the rear lens of the system (60 mm focal length). The light was dispersed in the spectrograph and measured with a thermoelectrically cooled CCD camera (PIXIS 256; Princeton Instruments). All spectra were recorded in the range from 634 to 1330 cm^{-1} using 2 s exposures with 30 accumulations (1 min per spectrum). Tests were conducted in a dark laboratory, and all spectra were corrected with a background spectrum captured while the laser was switched off.

Direct Raman measurements were also obtained on the samples without concealment in the opaque container. The samples were placed in an open metallic dish on the bench with the illumination and collection optics situated above in a vertical orientation. The same laser spot size used in the SORS measurements was used for excitation. Three replicates were acquired for each sample and used in the PLS modeling.

Data processing

Scaled subtraction (SORS analysis) of all spectra was performed using Matlab R2009b (The Mathworks). Preprocessing and multivariate analyses of the resultant mixture spectra were performed using PLS_toolbox 6.2.1 (Eigenvector Research Inc.).

Results and Discussion

Reference spectra for pure standards of acetaminophen, phenylephrine hydrochloride, caffeine and glucose are shown in Figure 2. There was considerable overlap of the spectral features between all four compounds, highlighting that multivariate analysis would be needed to extract information from mixture spectra.

SORS data treatment

The SORS data processing is demonstrated for the 100% (pure) acetaminophen sample in Figure 3. The ring and spot spectra were background corrected by subtracting the minimum intensity. The spot measurement was dominated by spectral lines from the container material, while in the ring spectrum there was a relative enhancement of the Raman signals from the sub-layer (concealed sample). The spectra were normalized by dividing by the intensity of the most prominent container peak (807 cm^{-1}). Subtracting the normalized spectra (ring minus spot) gave the Raman spectrum of the contents. The SORS result compared favorably with the reference Raman spectrum for acetaminophen (Figure 3) and the spectral lines from the container material were effectively removed from the spectrum. The same processing procedure was applied for analyzing all of the mixtures. The resulting SORS spectra were then analyzed using multivariate methods to quantify the drug concentrations within the mixtures.

Principal component analysis (PCA)

Prior to conducting PCA, data are generally subjected to several preprocessing steps. Features such as sloping baselines, backgrounds, noise, or differences in overall spectral intensity can cause unwanted variability to be incorporated into the PCA model, instead of the variations of interest (Raman spectral bands) and these features should be corrected with preprocessing. Since the spectra had already been normalized during SORS analysis, they did not require re-normalization. A baseline correction was performed on each spectrum using weighted least-squares with a second-order polynomial. The final preprocessing step was mean centering, which ensured that the first principal component (PC) described the direction of maximum variance of the data and not the mean. It was not necessary to smooth the data as the spectra had relatively low noise. Once these preprocessing steps had been performed, it was possible to conduct PCA using the PLS_toolbox software.

The data from sets A and B were run together in the PCA model and the variance described by each PC was examined. Figure 4a shows the scores plot for the first two PCs (PC2 vs. PC1), which indicates good separation of the two datasets according to drug type. Evidently, PC1 (82.95% variance) tends to relate to the differences between the two datasets. Along PC1, there is a negative correlation between samples that contain acetaminophen and phenylephrine. Accordingly, samples with higher concentrations of acetaminophen and phenylephrine are further apart on the graph; the 5% samples are closest together since they are the most similar samples in the two sets (i.e. caffeine and glucose were their main ingredients). PC2 (13.93% variance) tends to relate to increasing concentration, for both acetaminophen and phenylephrine. There is also a marked trend present in both datasets: the data seem to be distributed along two lines and, with the exception of a few datapoints, are arranged in order of their concentration (Figure 4).

This trend was also investigated in the scores plots for higher PCs. There is some residual variance accounted for by PC3 (2.34% variance), which is mostly linear and hence consistent with the first PC plot (Figure 4b). Again, the data in this plot tend to be arranged in order of their concentration, although the effect is weaker than in the earlier PCs. It was found that PC4 (0.27%) and higher PCs accounted for minimal variance, which could likely be attributed to noise. Confirming this, the scores plots for PC4 (PC4 vs. PC1, PC2 and PC3, respectively) showed that the data were randomly distributed along the PC4 axis (plots not shown).

These results indicate that the drug concentrations were well explained by simple linear relationships using the PCs, and that no further complex trends were present in the datasets.[24] Good agreement was observed between the three replicates measured on each sample, confirming the reproducibility of the results (Figure 4a).

The loadings plots for each PC were also investigated, to examine which spectral features were important for discrimination of the mixtures amongst the various PCs. Loading 1 closely resembles the reference spectrum for acetaminophen (Figure 5), with the added appearance of small peaks at approximately 740, 1025 and 1080 cm^{-1} (likely caffeine) and approximately 1000 cm^{-1} (likely phenylephrine). Loading 2 still retains a clear contribution from the acetaminophen spectrum, but the contribution of phenylephrine is much stronger in this PC compared to PC1 (Figure 6). This is particularly evident from the emergence of the dominant phenylephrine peak at 1000 cm^{-1} and other features at approximately 770, 940 and 1090 cm^{-1}. Finally, loading 3 is the most feature-rich plot, although it explains only 2.34% variance. Loading 3 is plotted with the glucose and caffeine reference spectra in Figure 7. It contains contributions from acetaminophen and phenylephrine as well as features due to caffeine (e.g. 740 cm^{-1}) and glucose (e.g. 840 cm^{-1}).

In summary, the results from the PCA confirmed that simple multivariate models should be suitable for the quantitative prediction of the drug concentrations of the mixtures.

Quantitative analysis using partial least squares (PLS) regression

Quantitative prediction models for the target drug concentrations were developed using PLS algorithms. Separate models were created for the acetaminophen and phenylephrine sets (sets A and B, respectively). The same preprocessing steps (baseline correction and mean centering) that were used for PCA were also used here.

The 13 mixtures containing acetaminophen were divided into a calibration set (8 mixtures) and a prediction set (5 'unknown' mixtures). The PLS model was developed using the 8 calibration samples including all 3 replicates of each (n = 24 measurements).

Cross-validation was used during model development to provide guidance on the appropriate number of latent variables (LVs) to use in the PLS model. The error obtained during cross-validation will generally decrease and plateau as further LVs are incorporated into the model, becoming representative of the prediction error for 'unknown' samples. At the same time, it is important to use as few LVs as are necessary to avoid overfitting the data and incorporating noise into the model, which would compromise its quality. Here, since the data contained replicate measurements, the 'random subsets' method was used in PLS_toolbox for cross-validation. In this method, the 24 measurements were divided into s subsets (here, s = 6) that contained n/s measurements, so that each measurement appeared in only one subset. The PLS model was applied to each subset to predict the acetaminophen concentrations, treating the samples as if they were unknown mixtures. The error was calculated and the results were averaged across the s tests. Finally, the process was iterated 20 times. The root-mean-square error of cross-validation (RMSECV) was calculated as the average of the iterations. The RMSECV was found to be unacceptably high using only one

LV; however, it decreased substantially to ~ 2.8% and began to plateau after including two LVs. Additional LVs provided only moderate benefit to the RMSECV. This indicated that two LVs would be optimal for the model.

The two loadings, which correspond to the two LVs, are plotted in Figure 8 (preprocessing algorithms removed). The first LV accounted for 98.80% of the variance in the data while the second accounted for 0.63% of the variance. Both plots showed a close resemblance to the reference spectrum for acetaminophen, confirming the model was targeting the spectral variations that were relevant in the data.

Figure 9 shows the predicted versus measured acetaminophen concentrations for the calibration samples and the prediction samples. The calibration mixtures followed a clear linear relationship (R^2 = 0.992) and the best-fit line of the data was close to unit slope. The predicted concentrations for the five unknown samples were in close agreement with the true values (Figure 9). The RMSEP was calculated as 3.8%, which represents an acceptable error.

The second PLS model, for the phenylephrine mixtures, followed the same division of calibration and prediction samples as used for the acetaminophen mixtures, and the random subsets method was again used for cross-validation. The RMSECV reached ~ 2.4% after including two LVs and remained stable as additional LVs were included. Thus, two LVs were also considered optimal for this model. The corresponding loadings for the two LVs (preprocessing algorithms removed) are plotted in Figure 10 and these closely resemble the Raman spectrum of phenylephrine hydrochloride. The first LV accounted for 94.95% of the variance and the second accounted for 2.67%.

A plot of predicted versus measured values (Figure 11) revealed that the calibration data followed a linear trend (R^2 = 0.995) with the best-fit line being situated close to unit slope. Finally, the 5 prediction samples were used to calculate the RMSEP. The predicted values

agreed well with the true values (Figure 11) and the RMSEP was 4.6%, which is acceptable for this application.

Direct Raman measurements

To evaluate the impact of the container on the quantitative accuracy, all samples were reanalyzed using direct Raman. The experimental procedures and multivariate analysis methods were kept consistent with the prior SORS measurements. The unconcealed acetaminophen samples gave a lower RMSEP of 3.2% while the phenylephrine samples gave a slightly higher RMSEP of 5.2%. These results seem to suggest that the presence of an opaque container does not adversely affect the quantification accuracy, when the concealed substances are analyzed using SORS.

General Discussion

These results show the capability of SORS combined with multivariate analysis to quantitatively assess mixtures inside opaque containers. While other studies have reported similar or superior levels of accuracy for Raman analysis of mixtures,[9, 12, 13] these have been for unconcealed samples and some level of sample preparation and handling was required. This may not be satisfactory if rapid, in-field screening is required for packaged products or concealed hazardous materials, and where non-contact and non-invasive techniques would be favored.

As indicated by the above results, the SORS technique succeeded in removing the Raman contributions from the concealing container and provided a non-invasive and non-contact means of screening mixtures in their packaging. Screening of concealed mixtures using SORS provided RMSEP values less than 5%, for samples with target drug

concentrations that ranged from 5% to 100%, and this value might be improved with optimization of the equipment and procedures.

When using SORS for the quantitative analysis of mixtures, the accuracy is expected to somewhat depend on the properties of the container. Containers with thicker walls or higher density polymer would be less penetrable to incoming excitation light and outgoing Raman photons. Therefore, the spatially offset measurements acquired on these containers would contain a weaker contribution of signals from the concealed substance, closer to the uncertainty and noise levels in the measurements, compared with thinner more light-penetrable containers. The uncertainty and noise in these measurements might be transferred into the resulting SORS spectra and could affect the accuracy of the quantifications. In some cases, it would be important to generate individual PLS models for different types of containers to ensure the most accurate results. For example, large variations in the size or shape of the sub-layer (e.g. due to different container sizes) may affect the amount of concealed substance in the Raman sampling volume. This should be avoided since it would alter the signal intensity from the sub-layer. Individualized models may also be necessary where the absorbance of the container material is significant and varies across the spectral region of interest. Otherwise, using a different container, or a container with a greater wall thickness, could cause inaccuracies in the quantification. Finally, the light scattering properties of the drug and diluents may affect the obtained results. For example, the differences in RMSEP between acetaminophen and phenylephrine might be due to slightly different scattering properties of the powders as well as their intrinsic Raman scattering properties.

There is a notable advantage of using SORS for mixture analysis, namely, that it should reduce the effect of sample inhomogeneity on the quantification accuracy.[7, 9, 13] Mixtures are often poorly homogenized, so acquiring spectra from different regions of a sample may

produce different quantitative results. Many studies reported the need to homogenize samples in order to build accurate multivariate models, particularly when using micro-Raman instruments.[7, 9, 11, 12] In this work, acceptable results were achieved without requiring such sample treatment (although sample inhomogeneity may still partly contribute to the attained RMSEP values). The reason SORS is less prone to sample inhomogeneity is due to its greater illumination and sampling volumes.[14, 18] This is an inherent feature of SORS since it relies on the diffusion of photons deeply into the medium and probes sample volumes considerably larger than many other Raman instruments and handheld portable devices allow. The use of larger illumination areas also has the advantage of reducing the potential for photo-damage and degradation of the samples.[15]

Conclusion

The interrogation and quantitative analysis of mixtures inside opaque containers has considerable practical benefit in many situations. This work has shown that the SORS technique can be combined with relatively straightforward chemometric methods to attain quantitative information regarding concealed mixtures. Analysis using SORS did not cause a substantial compromise to the prediction accuracy when compared with results from direct Raman. Further work should investigate the ultimate accuracies that can be expected when performing the analysis through an overlying layer (i.e. the container wall) and focus on optimization of the technique for particular applications. SORS provides the advantage of being less affected by sample inhomogeneity, due to its wider illumination and sampling areas, compared with many common Raman spectroscopy techniques. SORS has considerable potential as a portable technique and could be incorporated into a mobile instrument for conducting non-invasive analysis of substances in the field. This would be particularly beneficial in security and forensic applications where non-destructive analytical

methods are preferred, as well as for the screening and authentication of pharmaceutical products throughout transportation and distribution.

Acknowledgements

We gratefully acknowledge the support of the National Security Science and Technology scheme (Prime Minister and Cabinet, Australian Government), the Queensland Government (National and International Research Alliance Partnerships scheme), Australian Future Forensics Innovation Network, the Singapore Government and the Queensland University of Technology.

References

1. E. Izake, Forensic Sci. Int. **202**, 1 (2010).
2. P. Matousek, Chem. Soc. Rev. **36**, 8, 1292 (2007).
3. C. Ricci, C. Eliasson, N. A. Macleod, P. N. Newton, P. Matousek, and S. G. Kazarian, Anal. Bioanal. Chem. **389**, 5, 1525 (2007).
4. C. Eliasson and P. Matousek, Anal. Chem. **79**, 4, 1696 (2007).
5. P. Matousek and A. W. Parker, J. Raman Spectrosc. **38**, 5, 563 (2007).
6. P. N. Newton, M. D. Green, F. M. Fernández, N. P. J. Day, and N. J. White, Lancet. Infect. Dis. **6**, 9, 602 (2006).
7. C. J. Strachan, T. Rades, K. C. Gordon, and J. Rantanen, J. Pharm. Pharmacol. **59**, 2, 179 (2007).
8. M. Pelletier, *Analytical applications of Raman spectroscopy* (Wiley-Blackwell, 1999).
9. K. Y. Noonan, L. A. Tonge, O. S. Fenton, D. B. Damiano, and K. A. Frederick, Appl. Spectrosc. **63**, 7, 742 (2009).
10. C. Eliasson, N. A. Macleod, L. C. Jayes, F. C. Clarke, S. V. Hammond, M. R. Smith, and P. Matousek, J. Pharm. Biomed. Anal. **47**, 2, 221 (2008).
11. A. G. Ryder, G. O'Connor, and T. J. Glynn, Journal of Forensic Sciences **44**, 5, 1013 (1999).
12. A. G. Ryder, G. M. O'Connor, and T. J. Glynn, J. Raman Spectrosc. **31**, 3, 221 (2000).
13. O. S. Fenton, L. A. Tonge, T. H. Moot, and K. A. Frederick, Spectrosc. Lett. **44**, 4, 229 (2011).
14. P. Matousek, I. P. Clark, E. R. C. Draper, M. D. Morris, A. E. Goodship, N. Everall, M. Towrie, W. F. Finney, and A. W. Parker, Appl. Spectrosc. **59**, 4, 393 (2005).

15. P. Matousek, Appl. Spectrosc. **60**, 11, 1341 (2006).

16. M. V. Schulmerich, K. A. Dooley, M. D. Morris, T. M. Vanasse, and S. A. Goldstein, J. Biomed. Opt. **11**, 6, 0605021 (2006).

17. B. B. Das, F. Liu, and R. R. Alfano, Rep. Prog. Phys. **60**, 2, 227 (1997).

18. W. J. Olds, E. Jaatinen, P. Fredericks, B. Cletus, H. Panayiotou, and E. L. Izake, Forensic Sci. Int. **212**, 69 (2011).

19. P. Matousek and A. W. Parker, Appl. Spectrosc. **60**, 12, 1353 (2006).

20. B. Cletus, W. Olds, E. L. Izake, P. M. Fredericks, H. Panayiotou, and E. Jaatinen, "Toward non-invasive detection of concealed energetic materials in-field under ambient-light conditions", in *Next-Generation Spectroscopic Technologies IV*, M. A. Druy and R. A. Crocombe, Eds. (SPIE, Orlando, USA, 2011), Vol. 8032, p. 80320I.

21. M. Bloomfield, P. W. Loeffen, and P. Matousek, "Detection of concealed substances in sealed opaque plastic and coloured glass containers using SORS", in *Optics and Photonics for Counterterrorism and Crime Fighting VI,* and *Optical Materials in Defence Systems Technology VII*. C. Lewis, D. Burgess, R. Zamboni, F. Kajzar, and E. M. Heckman, Eds. (SPIE, Toulouse, France, 2010), Vol. 7838, p. 783808-1.

22. M. D. Hargreaves and P. Matousek, "Threat detection of liquid explosive precursor mixtures by Spatially Offset Raman Spectroscopy (SORS)", in *Optics and Photonics for Counterterrorism and Crime Fighting V*, C. Lewis, Ed. (SPIE, Berlin, Germany, 2009), Vol. 7486, p. 74860B.

23. P. W. Loeffen, G. Maskall, S. Bonthron, M. Bloomfield, C. Tombling, and P. Matousek, "Chemical and explosives point detection through opaque containers using spatially offset Raman spectroscopy (SORS)", in *Chemical, Biological, Radiological, Nuclear, and Explosives (CBRNE) Sensing XII*, A. W. Fountain III and P. J. Gardner, Eds. (SPIE, Orlando, USA, 2011), Vol. 8018, p. 80181E-1.

24. P. J. Gemperline, "Principal Component Analysis", in *Practical Guide to Chemometrics,* P. J. Gemperline, Ed. (CRC/Taylor & Francis, Boca Raton, 2006), p. 86.

Figure captions

Figure 1. Schematic diagram of the SORS instrument and experimental setup.[18]

Figure 2. Reference Raman spectra for the four mixture ingredients: acetaminophen, phenylephrine hydrochloride, caffeine and glucose.

Figure 3. SORS data processing, illustrated with the 100% (pure) acetaminophen sample inside the polypropylene container. (a) Spectra acquired with spot and ring-shaped illumination. (b) Baseline-corrected and normalized spectra. (c) Subtraction of the scaled spectra (ring minus spot) gave the spectrum of the concealed acetaminophen.

Figure 4. Scores plots for the PCA, presented as (a) PC2 vs. PC1 and (b) PC3 vs. PC1. 'P' denotes phenylephrine hydrochloride, 'A' denotes acetaminophen, and the concentration by weight is given in percent next to each data point (data labels removed in (b) for clarity).

Figure 5. (a) Loadings plot for PC1 (82.95% variance) and (b) the reference spectrum for acetaminophen.

Figure 6. (a) Loadings plot for PC2 (13.93% variance) and (b) the reference spectrum for phenylephrine hydrochloride.

Figure 7. (a) Loadings plot for PC3 (2.34% variance) and the reference spectra for (b) caffeine and (c) glucose.

Figure 8. Loadings plots for (a) LV1 and (b) LV2 of the acetaminophen PLS model.

Figure 9. PLS regression model for predicting the percentage of acetaminophen (by weight) in the mixtures in set A (■ calibration mixtures; ○ validation mixtures).

Figure 10. Loadings plots for (a) LV1 and (b) LV2 of the phenylephrine hydrochloride PLS model.

Figure 11. PLS regression model for predicting the percentage of phenylephrine hydrochloride (by weight) in the samples in set B (■ calibration mixtures; ○ validation mixtures).

Figures

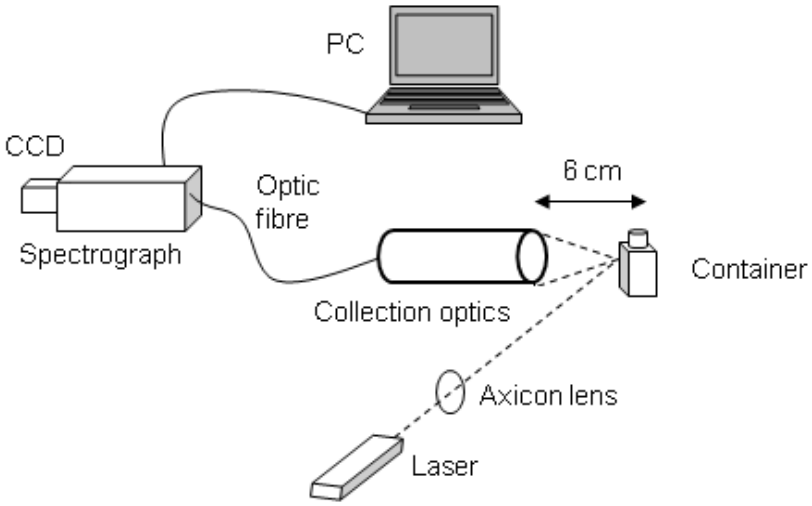


Figure 1.

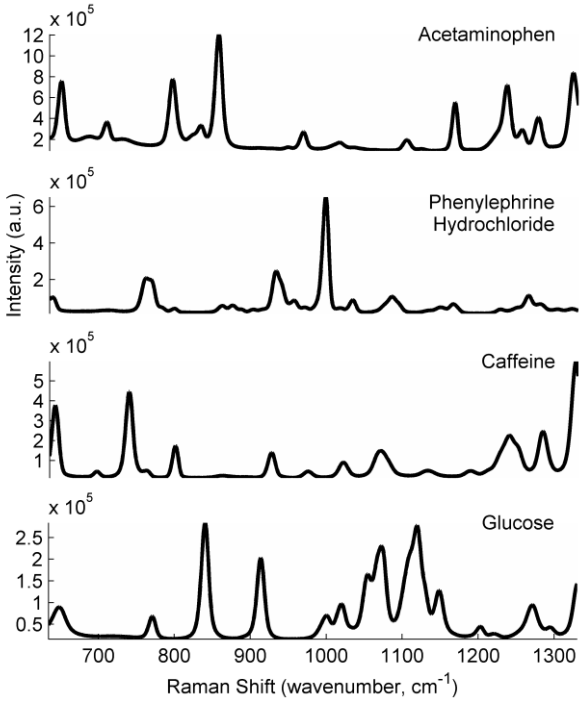


Figure 2.

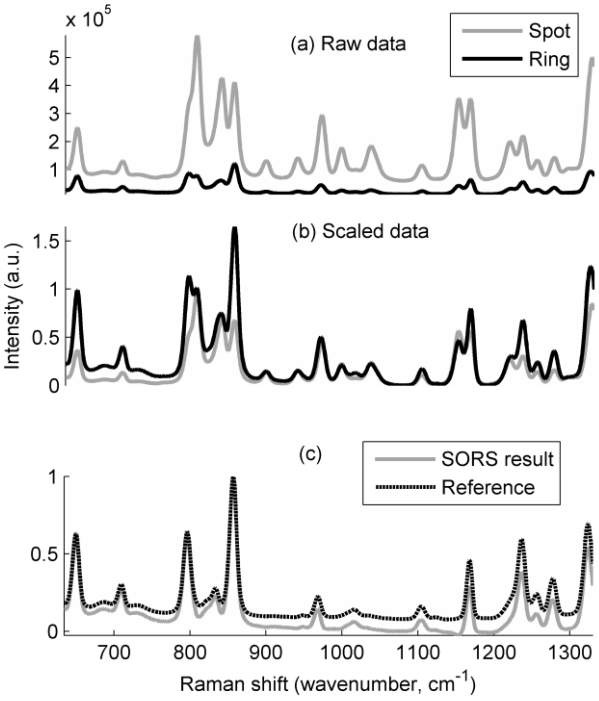
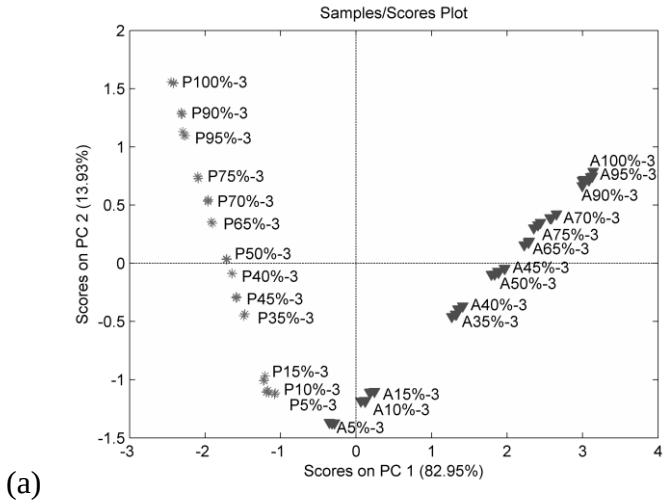
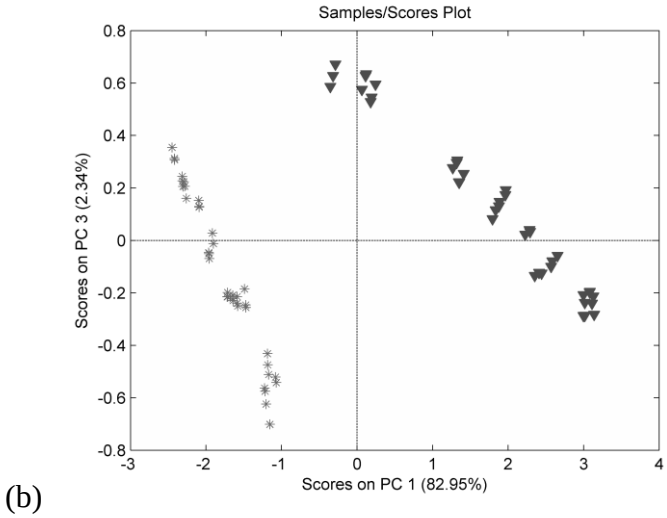

x 10^5
(a) Raw data
Spot
Ring
(b) Scaled data
Intensity (a.u.)
(c)
SORS result
Reference
Raman shift (wavenumber, cm^-1)

Figure 3.


Samples/Scores Plot
Scores on PC 2 (13.93%)
Scores on PC 1 (82.95%)
P100%-3
P90%-3
P95%-3
P75%-3
P70%-3
P65%-3
P50%-3
P40%-3
P45%-3
P35%-3
P15%-3
P10%-3
P5%-3
A100%-3
A95%-3
A90%-3
A70%-3
A75%-3
A65%-3
A45%-3
A50%-3
A40%-3
A35%-3
A15%-3
A10%-3
A5%-3

(a)


Samples/Scores Plot
Scores on PC 3 (2.34%)
Scores on PC 1 (82.95%)

(b)

Figure 4.

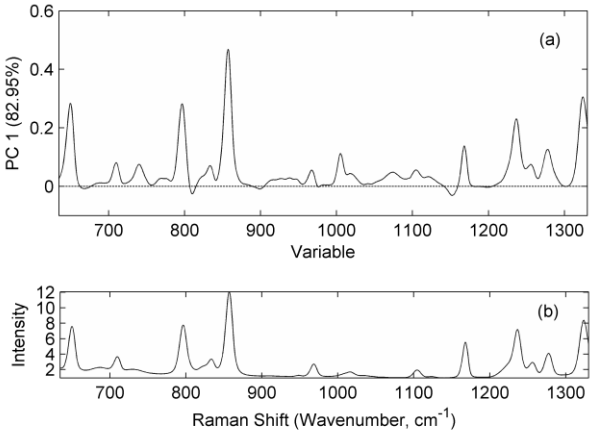


Figure 5.

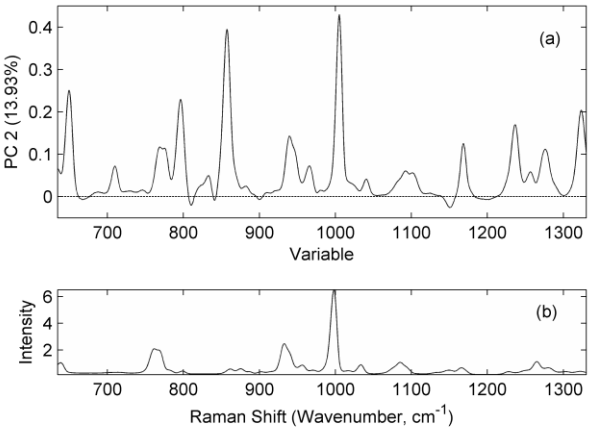


Figure 6.

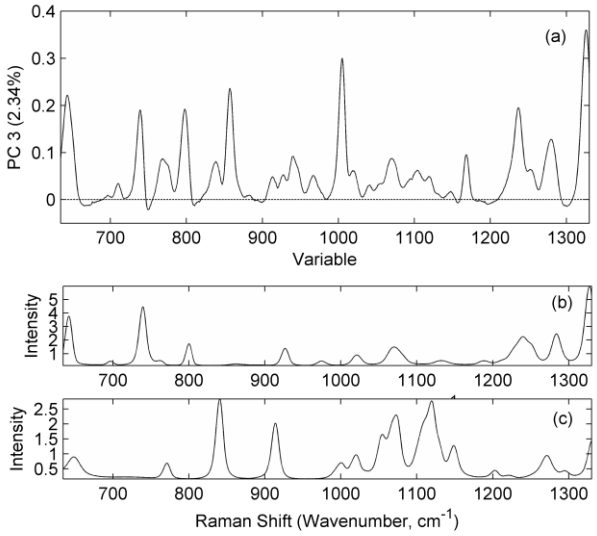


Figure 7.

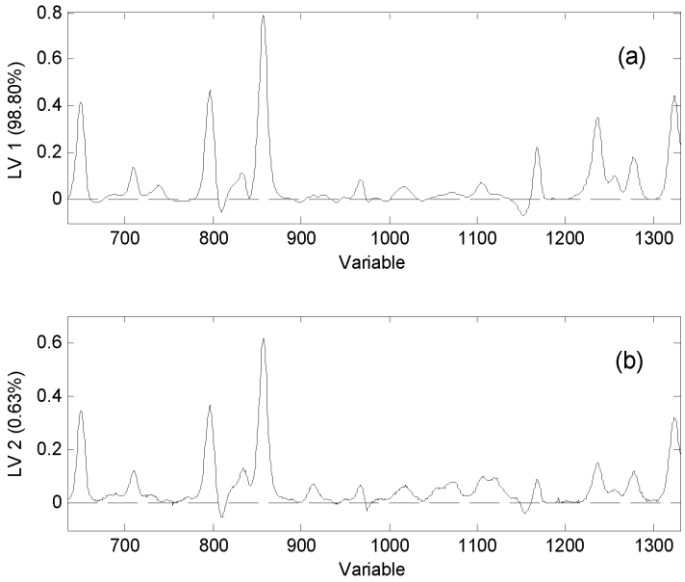


Figure 8.

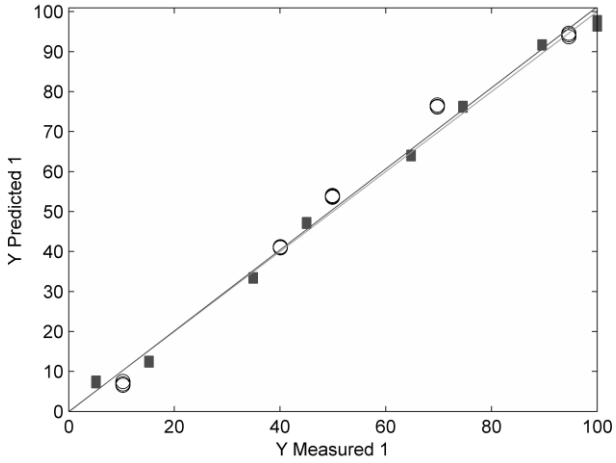


Figure 9.

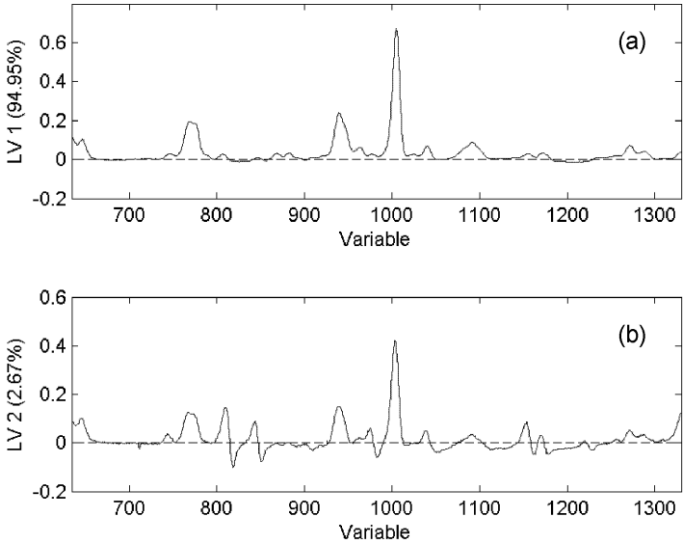


Figure 10.

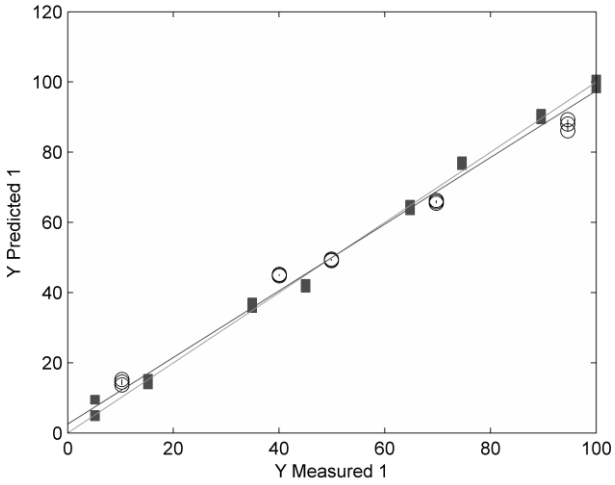


Figure 11.

Table

Table I. Compositions of the two sets of mixtures containing acetaminophen (set A) and phenylephrine hydrochloride (set B) as the target drugs.

Set A: Acetaminophen				**Set B: Phenylephrine hydrochloride**			
Sample #	Acetaminophen (%w/w)	Caffeine (%w/w)	Glucose (%w/w)	Sample #	Phenylephrine hydrochloride (%w/w)	Caffeine (%w/w)	Glucose (%w/w)
A1	5.16	47.50	47.34	P1	5.10	47.40	47.49
A2	10.24	44.96	44.80	P2	10.02	45.06	44.92
A3	15.19	42.42	42.38	P3	15.02	42.32	42.66
A4	34.92	32.67	32.41	P4	34.93	32.45	32.62
A5	40.02	30.18	29.81	P5	39.98	30.00	30.02
A6	45.01	27.70	27.28	P6	44.85	27.61	27.54
A7	49.88	25.25	24.87	P7	49.64	25.27	25.09
A8	64.78	17.69	17.53	P8	64.76	17.80	17.45
A9	69.73	15.30	14.97	P9	69.72	15.21	15.07
A10	74.59	13.04	12.36	P10	74.62	12.66	12.71
A11	89.60	5.45	4.95	P11	89.47	5.23	5.31
A12	94.61	2.90	2.49	P12	94.42	2.72	2.85
A13	100.00	-	-	P13	100.00	-	-