J Pharm Innov
DOI 10.1007/s12247-013-9168-8

RESEARCH ARTICLE

Pharmaceutical Application of Fast Raman Hyperspectral Imaging with Compressive Detection Strategy

Derya Cebeci-Maltaş · Ryan McCann · Ping Wang · Rodolfo Pinal · Rodolfo Romañach · Dor Ben-Amotz

Abstract This article reports a new Raman imaging instrument based on liquid crystal spatial light modulator compressive detection (LC-SLM-CD) strategy that could provide a fast way of testing the composition of solid formulations. In this study, the LC-SLM-CD strategy is employed to investigate the qualitative and relative quantitative distribution of two commonly used ingredients, acetaminophen and lactose in a pharmaceutical tablet. The spatial distribution of each component is formed based on the responses of the samples to the partial least squares filters built into the instrument. This technique has proven to be a fast and feasible technique for noninvasive determination of blend quality and for determination of relative abundances of each component in a tablet.

Keywords Raman spectroscopy · Hyperspectral imaging · Compressive detection · PLS

Introduction

Throughout the formulation development of solid dosage forms such as tablets, the goal is that every unit of final drug product will have the desired performance with identical therapeutic effects. The drug's physical properties, and ultimately its performance, highly depend on the quantities of the components and how well the final product is blended. A relatively small degree of deviation from the optimum specifications may yield a widely varying therapeutic performance. Thus, determining the blend quality and generating statistical blending data quickly is imperative for production and is critical for process analytical technology (PAT) to reduce the cost and delay time in a production cycle.

High-performance liquid chromatography, UV/visible spectroscopy, or mass spectrometry are widely used tools to determine the gross composition of formulations, but none of them provides any insight into the distribution of the components within the analyzed sample. Dissolution testing, on the other hand, is employed to obtain drug release profiles, indicating the duration of component release. However, these currently used methods provide no information regarding the root cause or structural basis for changes in the dissolution profile. Thus, it is impractical and inefficient to quickly trace the failure and correct it with these traditional tools. In addition, these standard quality control techniques are destructive, laborious, and time consuming, which stalls the production awaiting test results from the quality control laboratory following the tablet compaction process. In fact, the analysis to validate the quality of blending takes more time than actual blending process.

Spectroscopic techniques, such as Raman and near infrared (NIR), can provide a rapid, nondestructive, means of monitoring a manufacturing process. Both techniques measure similar molecular vibrations. In imaging mode, they can provide an assessment of the content uniformity of a sample in terms of the spatial distribution of the ingredients. While Raman spectra is characterized by narrower peaks, giving better information on molecular level, wide, overlapping NIR bands are difficult to assign. However, NIR interrogates larger areas than Raman in shorter times. A study by Jerez-Rozo et al. evaluated the use of NIR and Raman mapping to investigate the spatial distribution of the components of polymeric films [1].

D. Cebeci-Maltaş (✉) · P. Wang · D. Ben-Amotz
Department of Chemistry, Purdue University, W. Lafayette, IN 47907, USA
e-mail: deryacebecimaltas@gmail.com

R. McCann · R. Pinal
Department of Industrial and Physical Pharmacy, Purdue University, W. Lafayette, IN 47907, USA

R. Romañach
Department of Chemistry, University of Puerto Rico, Mayagüez, PR 00680, Puerto Rico

Published online: 27 November 2013

Raman imaging is conventionally done by collecting a large number of spectra at desired spatial positions on the sample, which produces "data cubes", containing the Raman signal intensity measured as a function of x and y (spatial) and spectral dimensions [1, 2]. Subsequently, spectra are converted to chemical component information through either univariate or multivariate data analysis. The common way to generate an image of a certain component is to calculate the area of a unique spectral peak pertaining to each component of interest, and plotting the intensities obtained at each spatial location in the image. However, when it is not possible to find a unique peak for the investigated component which does not overlap with any other components, multivariate data analysis techniques [3], such as partial least squares (PLS), are typically used to extract chemical composition information from the spectra. Most importantly, collecting a full Raman spectrum typically requires at least 1 s, and so collecting a hyperspectral image can be very time consuming (for example, even a moderately small 100×100 image requires about 3 h of data collection). Hence, traditional Raman imaging requires significant expertise and time, which hinders Raman imaging applications on a routine basis.

Recently, a new way of doing fast Raman imaging based on a liquid crystal spatial light modulator compressive detection (LC-SLM-CD) strategy was reported [4]. The instrument is used with compressive detection filters loaded on programmable SLM to compress full-spectral data onto a single-channel detector. These filters can be generated by applying multivariate data analysis techniques, such as PLS. LC-SLM-CD Raman system measures the responses of Raman spectra to the filters created on SLM. This system effectively incorporates data analysis into its hardware, eliminating extensive imaging data analysis required once the necessary filters are generated [4].

In this study, we investigated a two-component pharmaceutical tablet composed of acetaminophen (APAP) and lactose to demonstrate the capability of the LC-SLM-CD Raman strategy for fast imaging of pharmaceutical composites. More specifically, we show that this method may be suited to determine the distribution of the two components as well as the gross composition of each tablet.

Methods

The Raman maps were collected with a single mode diode laser of 785 nm wavelength that delivers 80 mW of power at the sample with a ×20 (NA 0.40) NIR objective lens. The laser spot size on the sample was about 7 μm. Compressive detection filters were generated using the PLS algorithm (PLS Toolbox (Eigenvector Research Inc., WA, USA) and MATLAB (MathWorks, Inc., MA, USA)). Figure 1 illustrates how the PLS model is established to create the regression coefficient vector from lactose and APAP components to be used for compressive detection filters in the LC-SLM-CD Raman system. Because the training matrix is composed of pure components, the property vector is made of 0 s (indicating APAP) and 100 s (indicating lactose).

Spectra of pure lactose and APAP used to produce PLS regression vectors are shown in Fig. 2a. The spectra are normalized to the unit area prior to implementing PLS analysis. Figure 2b shows the output regression coefficient vector (b2 in Fig. 1) of the PLS analysis using the spectra from Fig. 2a. This vector, b2, is essentially what is used as an SLM filter function after adapting it to SLM voltage values [4]. It consists of positive and negative portions as seen in Fig. 2b. Before it is transformed to LC-SLM-CD compressive detection filters, the negative portion should be converted into positives by splitting it into two (SLM does not recognize negative vectors) and then taking the absolute values of the negative portions of each vector to convert them into a positive filter function. As a result, two LC-SLM-CD filters (Fig. 2c) are generated from one PLS regression vector (b2). The two filters are scaled to a maximum value of 1 to set them to the maximum SLM transmittance. Scaling constants are later used to regenerate the response values to form the image [4].

Results and Discussion

LC-SLM-CD measures the responses (scores in this study) generated by passing light emanating from a given point in the sample through each of the filters generated from the PLS regression coefficient vectors. Note that measuring the light transmitted (or reflected) by an optical filter is equivalent to obtaining the dot-product of the spectral vectors corresponding filter and the light from the sample. Thus, the spectrum coming from each point in the sample produces a score (dot-product) for each filter, and those scores are used to identify the composition of that point in the sample as belonging to either the APAP or lactose class.

Images (50×200 pixels) with an area of 0.5×2 mm^2 were formed by collecting a series of spectral responses from adjacent locations on the sample by moving the sample at 10 μm intervals between each measurement. The histogram plots formed in Fig. 3a show the distribution of the LC-SLM-CD responses to the digital filters at each individual pixel for each sample; pure APAP, pure lactose, and APAP–lactose mixed tablets. The abscissa represents the spectral responses coming from each location on the samples. The ordinate, on the other hand, denotes the number of pixels possessing a certain value of those responses. The spread in prediction values in the histograms may be due to instrumental fluctuations or sample inhomogeneity. The spectra necessary for the filter

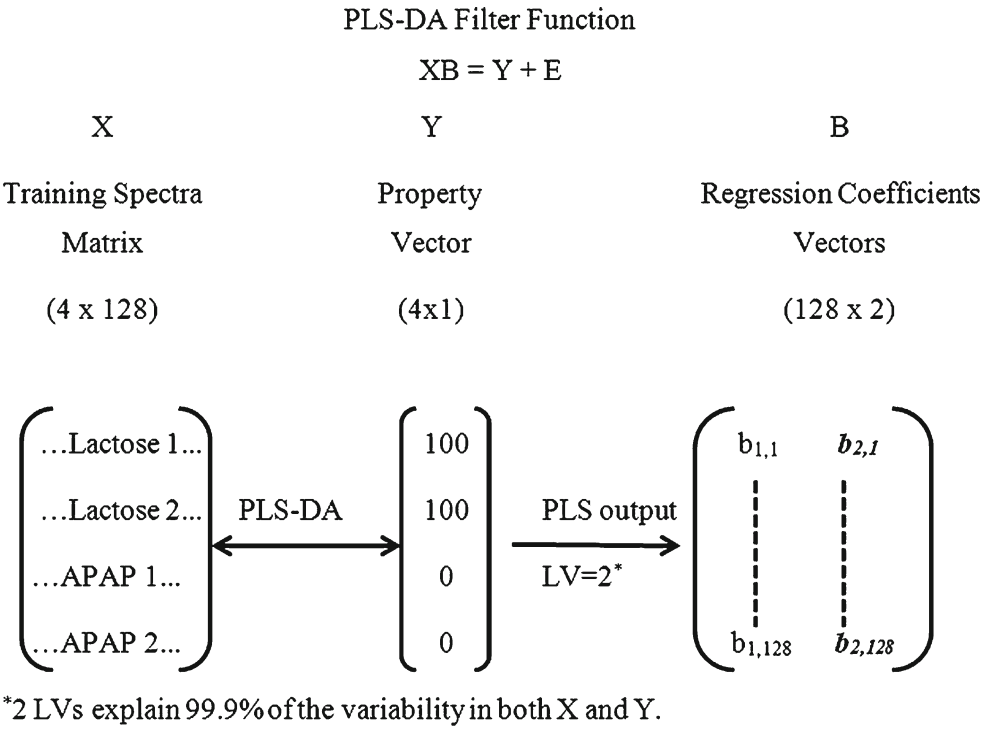

Fig. 1 Schematic of PLS-DA filter function set up to generate regression vectors. This illustration represents including only two latent vectors (LVs).

functions were collected on the limited number of points in the sample. If the sample is not chemically homogeneous, filter functions may have represented only a few points in the sample. As the heterogeneity increases, the histogram starts deviating from normal distribution. Also, since the spot size of the laser is considerably smaller than the particle size of the components, sampling volume may be another contribution in the spread. The histogram plots on pure components in Fig. 3a are produced to set a threshold range to determine which pixels are estimated to belong to the class of interest. With 99 % confidence, the threshold for APAP was set at score values of −10 to 35 and lactose is 60 to 95. Values

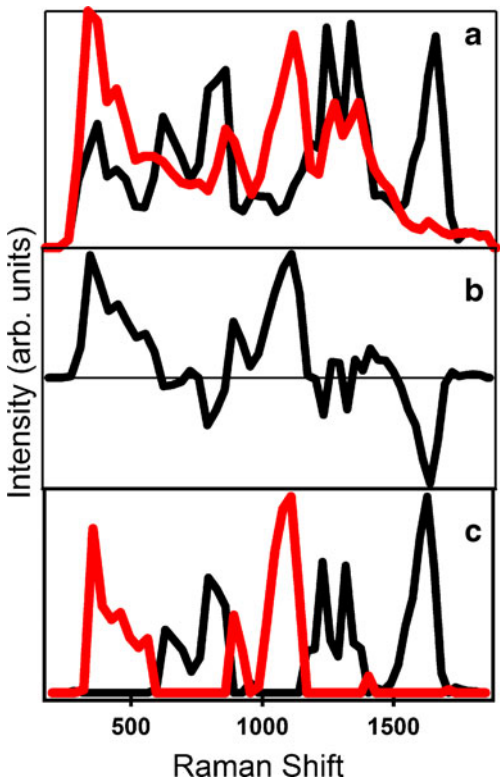

Fig. 2 PLS-derived filter generation. **a** Pure component spectra of APAP (*black*) and lactose (*red*). **b** PLS output regression vector (b2) using spectra from **a. c** Production of two PLS regression vectors by splitting the PLS regression vector in (**b**) into a positive part (*red*) and the absolute value of the negative part (*black*)

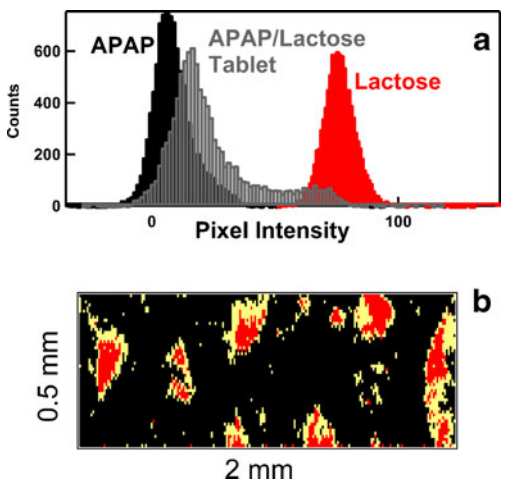

Fig. 3 **a** Histogram plots of pure APAP, pure lactose and APAP-lactose mixed tablet. **b** LC-SLM-CD response images (50×200 pixels) of APAP-Lactose mixed tablet. *Red* lactose, *black* APAP, and *light yellow* mixture of APAP and lactose

between 35 and 60 suggest that the signal is coming from a mixture of the two components.

Figure 3b is the resulting LC-SLM-CD response image of APAP–lactose mixed tablet. The image illustrates the spatial distribution of the components APAP and lactose. While black regions denote pure APAP, red regions represent pure lactose. Yellow areas indicate pixels within the image that appear to contain a mixture of the two components. Thus, each color channel in this figure contains qualitative information about the distribution of APAP and lactose in the tablet. Visual inspection of the image reveals clearly distinctive, localized lactose domains embedded in large APAP regions. Qualitatively, APAP appears to be more abundant than lactose in the image.

While the visual examination provides qualitative information, more quantitative information can be obtained by examining the statistical properties of the measured responses. Although LC-SLM-CD responses do not provide information about the absolute concentration of the components, the relative abundance information can still be reliably extracted from the image data. The density of sampling (10,000 data points in the image) offers a robust statistical analysis of sample components. The investigated volume on the sample is in microscopic range in LC-SLM-CD Raman system (~7 μm/pixel). As a result, sample components with large particle size (typically >40 μm for this study) comprise a spatially heterogeneous matrix relative to the experimental volume. Each particle typically occupies more than one pixel in the image. Hence, a single pixel can be classified as belonging either to one component or another. Accordingly, the relative quantity can be estimated without having to develop a standard concentration calibration method. The relative abundance of pixels belonging to pure APAP in the region imaged is estimated to be 80 %, by counting the number of pixels between −10 and 35 and then dividing it by the total number of pixels in the image. Classified as lactose, yielding a 10 % pixel abundance are ~1,000 pixels out of 10,000 between the values of 60 and 95. All the other pixels between 35 and 60 correspond to a mixture of the two ingredients in the volume investigated, representing 10 % of the sample.

Conclusion

Although only two-component system is investigated in this study, the possibility of expanding to matrices with multiple components has been shown in a previous work [5]. If multiple components need to be explored, then the number of necessary filter functions to define them all may also need to increase, accordingly collection time will also increase. The Raman chemical image with 10,000 points shown in Fig. 3b took 5 min to collect, while the same size image would have taken several hours to obtain using traditional CCD-based Raman instrumentation. In addition, the data analysis required to form the image is practically eliminated for LC-SLM-CD system, as the scores used to produce the image are directly detected using the instrument, rather than produced by postprocessing of conventionally measured spectra. Thus, the LC-SLM-CD Raman detection strategy shows promise as a PAT sensing technology which can rapidly and noninvasively characterize solid dosage forms. The statistical analysis of thousands of LC-SLM-CD responses or pixel/score values provides information regarding the spatial distribution and relative abundance of the components in the image. Such information is of importance for manufacturing process monitoring, as it provides statistical information regarding blend quality. Since the data collection and data analysis can be automated and displayed in real time, any deviation from ideal specifications can be detected and fixed during production.

Acknowledgments The present work was financially supported by NSF-Engineering Research Center for Structured Organic Particulate Systems (ERC-SOPS, EEC-0540855). The development of this instrument was supported by a grant from the National Science Foundation (grant IDBR 0754740).

References

1. Jérez Rozo JI, Zarow A, Zhou B, Pinal R, Iqbal Z, Romañach RJ. Complementary near-infrared and Raman chemical imaging of pharmaceutical thin films. J Pharm Sci. 2011;100:4888–95.
2. Sasic S. Raman mapping of low-content API pharmaceutical formulations. I. Mapping of Alprazolam in Alprazolam/Xanax tablets. Pharm Res. 2007;24:58–65.
3. Zhang L, Henson MJ, Sekulic SS. Multivariate data analysis for Raman imaging of a model pharmaceutical tablet. Anal Chim Acta. 2005;545:262–78.
4. Davis BM, Hemphill AJ, Cebeci Maltas D, Zipper MA, Wang P, Ben-Amotz D. Multivariate hyperspectral Raman imaging using compressive detection. Anal Chem. 2011;83:5086–92.
5. Cebeci Maltaş D, Kwok K, Wang P, Taylor LS, Ben-Amotz D. Rapid classification of pharmaceutical ingredients with Raman spectroscopy using compressive detection strategy with PLS-DA multivariate filters. J Pharm Biomed Anal. 2013;80:63–8.