Journal of Pharmaceutical and Biomedical Analysis 80 (2013) 63–68

Contents lists available at SciVerse ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Rapid classification of pharmaceutical ingredients with Raman spectroscopy using compressive detection strategy with PLS-DA multivariate filters

Derya Cebeci Maltaş[a,*], Kaho Kwok[b], Ping Wang[a], Lynne S. Taylor[b], Dor Ben-Amotz[a]

[a] Department of Chemistry, Purdue University, W. Lafayette, IN 47907, United States
[b] Department of Industrial and Physical Pharmacy, Purdue University, W. Lafayette, IN 47907, United States

ARTICLE INFO

Article history:
Received 11 December 2012
Received in revised form 14 February 2013
Accepted 20 February 2013
Available online 28 February 2013

Keywords:
Compressive detection
Raman spectroscopy
Rapid classification
PLS-DA
Process analytical technology (PAT)
Cialis®

ABSTRACT

Identifying pharmaceutical ingredients is a routine procedure required during industrial manufacturing. Here we show that a recently developed Raman compressive detection strategy can be employed to classify various widely used pharmaceutical materials using a hybrid supervised/unsupervised strategy in which only two ingredients are used for training and yet six other ingredients can also be distinguished. More specifically, our liquid crystal spatial light modulator (LC-SLM) based compressive detection instrument is trained using only the active ingredient, tadalafil, and the excipient, lactose, but is tested using these and various other excipients; microcrystalline cellulose, magnesium stearate, titanium (IV) oxide, talc, sodium lauryl sulfate and hydroxypropyl cellulose. Partial least squares discriminant analysis (PLS-DA) is used to generate the compressive detection filters necessary for fast chemical classification. Although the filters used in this study are trained on only lactose and tadalafil, we show that all the pharmaceutical ingredients mentioned above can be differentiated and classified using PLS-DA compressive detection filters with an accumulation time of 10 ms per filter.

1. Introduction

In an effort to increase the efficiency and quality of pharmaceutical manufacturing, the U.S. Food and Drug Administration (USFDA) issued a process analytical technology (PAT) initiative to encourage the industry to innovate and adopt new measurement technologies for the development, manufacturing and quality assurance of drug products [1]. The goal is to achieve increased quality and productivity through improved understanding of manufacturing processes. The PAT initiative offers a fundamental shift from current laboratory-based quality control approaches by establishing that quality should be tested frequently at earlier stages in manufacturing. It emphasizes timely measurements of quality attributes of raw and in-process materials and finished products. The PAT guidance brought a new challenge to the industry in the sense that timely measurements necessitate that testing should be brought to the production floor from the quality assurance laboratories. Traditionally, invasive instrumentations such as high performance liquid chromatography (HPLC), gas chromatography, mass spectroscopy or wet chemistry techniques are commonly employed in quality control laboratories. These techniques are generally time-consuming, labor-intensive and laboratory-based, which makes it difficult to employ them on the production floor. With the advent of PAT, the demand is now shifted to faster analytical instruments with the ability to monitor manufacturing process in real time and which can be employed along the production line [2]. The PAT initiative accordingly spurred an immense interest in vibrational spectroscopy thanks to its features such as its speed and noninvasive nature.

Although near infrared spectroscopy (NIRS) is more common as a PAT sensor, recent technological and scientific advancements have broadened the applicability of Raman spectroscopy into various areas [3]. Compared to NIRS, the Raman spectrum gives more structural information on the molecular level with a higher specificity; NIR spectral bands are generally very broad and overlapped. This attribute makes the Raman technique a valuable tool for sample identification within the pharmaceutical industry, since each chemical has a unique vibrational fingerprint. Also, Raman spectroscopy has a major advantage of being insensitive to water content, which is a significant obstacle to NIR detection. Water is a very strong IR absorber but a weak Raman scatterer. On the other hand, the cross-section of Raman is low, which makes the Raman technique quite inefficient. Thus, conventional charge-coupled device (CCD) based Raman instruments require a collection time of the order of seconds per point to produce signals with decent signal-to-noise ratio. In order to overcome the speed limitation of traditional Raman spectroscopy, recently a new Multivariate Hyperspectral Imaging (MHI) Raman method was introduced [4]. The MHI system is based on a liquid-crystal spatial-light-modulator compressive detection (LC-SLM-CD) strategy. Here we demonstrate

* Corresponding author. Tel.: +1 631 9440305.
E-mail address: deryacebecimaltas@gmail.com (D. Cebeci Maltaş).

0731-7085/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jpba.2013.02.029

that this LC-SLM-CD may also be used to implement a new hybrid supervised/unsupervised pharmaceutical chemical classification strategy.

The LC-SLM consists of an array of liquid crystal pixels whose reflectivity is computer programmable [4]. The fundamental advantage of the LC-SLM-CD detection strategy over traditional Raman instruments is that a single channel detector is used to collect all the light transmitted (reflected) by each programmable optical filter, and thus provides higher signal-to-noise than is obtainable when the same light is distributed over the many channels of an optical array detector such as a charge-coupled-device (CCD). Accordingly, the LC-SLM-CD detection strategy speeds up the collection of Raman data and makes it attractive as an analytical sensor for PAT applications by synchronizing rapid data collection and multivariate data analysis.

Prior to their use in manufacturing, all drug components are subject to identity verification testing. The tools that are typically used for this purpose generally fall into two categories: spectral correlation functions, such as hit-quality index; and factor-based multivariate algorithms, such as PLS. Spectral correlation algorithms are conventionally the more common tool in the industry mainly due to their simplicity. However this technique is not selective enough to differentiate chemical species with high chemical similarity [5]. Factor-based multivariate algorithms with higher discriminating power, on the other hand, are often very effective in detecting even subtle spectral differences [5,6]. Each chemical is identified by n-wavenumber dimensions in Raman spectrum. Multivariate algorithms basically work by compressing n-dimensional spectra onto new, fewer dimensions in a new coordinate space. They require post-processing of spectral data to produce score values from full spectra. A property of interest, such as identity, is then defined by score values on each new axis in this new compressed coordinate space. LC-SLM compressive detection strategy differs from the usual way of applying multivariate algorithms in the way that it generates the scores in its hardware, eliminating all post-processing necessary to obtain scores values from full spectra.

This study is focused on evaluating the feasibility and the implementation of the newly built LC-SLM-CD Raman instrument as a potential PAT sensor to classify various pharmaceutical ingredients using multivariate filters. The ingredients chosen are the components used in highly counterfeited Cialis® tablets where the active ingredient is tadalafil. We followed a hybrid supervised/unsupervised strategy. We built a PLS-DA multivariate identification system with a two-component training library to test the LC-SLM-CD responses obtained using those two as well as six other pharmaceutical components. More specifically, multivariate filter functions were supervised by two ingredients and then were applied on LC-SLM-CD to classify both supervised and six other unsupervised raw materials. Our goal was to test whether our LC-SLM-CD strategy could be used to identify/classify compounds that are outside of a given training library as well as the compounds in the training library.

2. Experimental

The following raw materials were used in order to evaluate the classification ability of the LC-SLM-CD Raman system: Tadalafil, lactose monohydrate, microcrystalline cellulose (MCC), magnesium stearate (MgSt), titanium (IV) oxide (TiO_2), talc, sodium lauryl sulfate (NaLS) and hydroxypropylcellulose (HPC). Lactose monohydrate, MgSt, NaLS and talc were obtained from Spectrum Chemicals (Gardena, CA). HPC was obtained from Ashland Inc. (Covington, KY). MCC was obtained from FMC Biopolymer (Philadelphia, PA). TiO_2 was obtained from the Sigma–Aldrich Company (St. Louis, MO). Genuine Cialis® tablets (20 mg) (Eli Lilly, Indianapolis, IN) were acquired from the Purdue University Pharmacy (West Lafayette, IN). Tadalafil was obtained through recrystallization following extraction from the genuine Cialis® tablets using the following procedure: The coating of three Cialis® tablets was removed using a razor blade. The uncoated tablets were ground into a fine powder using a mortar and pestle. 30 ml of ethanol (Sigma–Aldrich) was added to the powder. The solution was thoroughly mixed and insoluble material was allowed to settle by gravity overnight. The solution was filtered using filter paper (No. 1, Whatman, Maidstone, UK). The resultant solution was filtered again using a syringe filter (0.2 μm PTFE, VWR, Radnor, PA) to remove smaller particulates. The supernatant was collected and the solvent was evaporated slowly under ambient conditions yielding tadalafil crystals.

Sample powders of these drug components were placed in 96 well-plates and their spectral responses were collected on the LC-SLM-CD instrument with 785 nm excitation and with approximately 80 mW power at the sample. A 20× (NA 0.40) NIR objective lens (Olympus, LMPL20XIR) was used to focus the laser onto the sample and to collect the scattered Raman photons. Compressive detection optical filters were generated using PLS algorithm on PLS Toolbox (Eigenvector Research Inc., WA) installed in Matlab (MathWorks, Inc., MA).

3. Results and discussion

3.1. Presentation of compressive detection filter techniques

The LC-SLM based Raman instrument can function as a conventional spectrometer to acquire full Raman spectra of chemicals or as a hyperspectral imaging instrument utilizing compressive spectral detection strategies to acquire spectral responses. However, the speed advantage of this Raman instrument is realized only when it is used in hyperspectral mode, operated with compressive detection filters. The filter functions can be computed with univariate or multivariate statistical techniques. They can be represented by n-dimensional vectors ($n = 128$). LC-SLM-CD effectively measures the dot-product of the filter vector and the spectral vector. Generally, in a univariate analysis, a property of interest is calculated based on a single value. For example; area of a certain peak in the spectrum can be correlated with the quantity of the ingredient to which that certain peak corresponds. As an example, a univariate approach may be appropriate if the purpose is to investigate active pharmaceutical ingredients (API) in pharmaceutical formulations. For a large number of APIs, the Raman technique is especially sensitive because of their aromatic functional groups. Unsaturated carbon bonds and aromatic ring functional groups tend to give strong Raman signatures around 1600 cm^{-1}, whereas the majority of excipients do not tend to produce any signals due to the lack of unsaturation or aromatic ring in their molecular structure.These attributes provide an excellent means to spatially locate an API in tablets with using Raman spectroscopy. A univariate filter on the LC-SLM-CD Raman system is generated by "turning off" all the pixels on the SLM except the ones that correspond to the unique peak of a certain chemical which does not overlap with other peaks. To illustrate this, Fig. 1 shows the Raman spectra of an active pharmaceutical ingredient, tadalafil, and the majority of the excipients in Cialis® tablets. The peak boxed by black dotted-line at around 1600 cm^{-1} is unique to active ingredient tadalafil, and does not overlap with the spectral signatures from any other excipients. A univariate approach can be implemented to investigate tadalafil in a Cialis® tablet by turning on the SLM pixels associated with the peak at 1600 cm^{-1} to 100% transmittance and setting all other pixels to 0% transmittance by turning them off. As a result, only the photons that interact with the pixels that are 'on' will reach the detector and be detected. The intensity of the photons detected is directly related to the amount

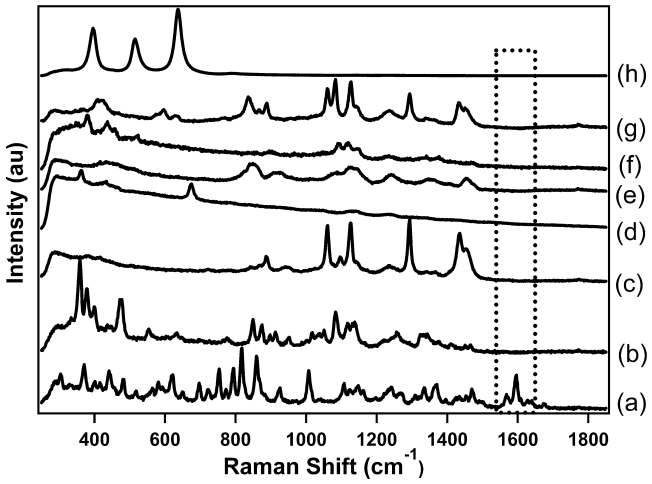


Fig. 1. Spectra of Cialis® ingredients with 785 nm laser line using a CCD-based Raman spectrometer. Each spectrum is normalized to the area and is offset on *y* axis for better illustration. (Spectra here are not corrected to the quantum efficiency of CCD detector) (a) Tadalafil, 10s (b) Lactose monohydrate, 10s (c) magnesium stearate, 60s (d) Talc, 60s, (e) hydroxypropylcellulose, 60s (f) microcrystalline cellulose, 10 s (g) sodium lauryl sulfate, 60s (h) titanium (IV) oxide, 5s.

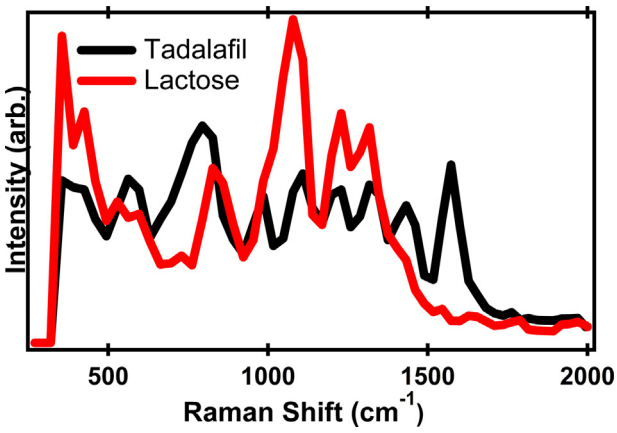
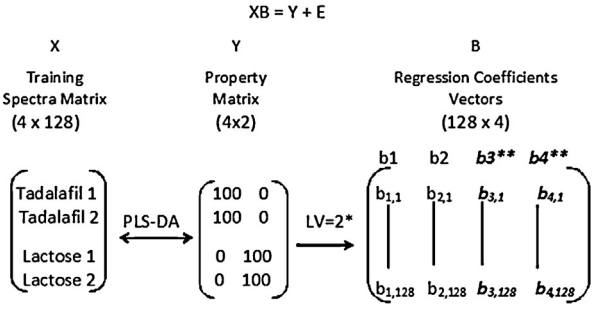


Fig. 2. Pure component spectra for PLS-DA training matrix X. Black spectrum is the Hadamard transform spectrum of tadalafil. Red is the hadamard transform spectrum of lactose. Spectra are normalized to unit area. Hadamard spectroscopy is employed with 128 resolution elements. Accumulation time per Hadamard filter was 1 s, thus total collection time of each spectrum was 128 s. Only one spectrum for each component is shown here although two spectra from each ingredient are used to be employed in PLS-DA application. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of tadalafil at a certain location on the tablet. Univariate filters may be more valuable due to their simplicity than multivariate based filters when the investigated chemical has a unique peak observable as in the tadalafil spectrum in Fig. 1. However, isolating a unique peak belonging to the query material may not always be feasible, especially considering the number of constituents typically present in a pharmaceutical solid formulation. In such cases, multivariate chemometric techniques are more appropriate than a univariate approach. While a univariate technique correlates one independent variable, such as quantity, to a single dependent variable, such as peak area or intensity, multivariate techniques take many variables into account, extracting the relevant information while disposing of information that is not correlated with the chemical variables of interest. One of the most commonly used multivariate techniques is partial least squares (PLS) [7]. This technique finds the latent vectors (LVs) describing the variance in training matrix (as in PCA and PCR) while correlating it with the property vector or matrix. It calculates scores and loadings for each latent vector computed.

3.2. Presentation of PLS-DA multivariate filters for LC-SLM compressive detection

To create compressive detection filter functions by implementing PLS-DA algorithm for the LC-SLM Raman instrument, first full Raman spectra of the chemicals are needed. The LC-SLM Raman system can be used as a conventional scanning spectrometer by utilizing Hadamard ENREF [8] filter functions to create full spectra. Fig. 2 shows the Hadamard transformed spectra of pure tadalafil and lactose, which are the two components that PLS-DA filters are trained on for this study. Prior to implementing the PLS-DA algorithm, the spectra are first normalized to unit area to minimize any variance induced by any instrumental fluctuations and sample alignment which would otherwise disturb intensity measurements. Then PLS-DA, which is demonstrated to be equivalent to performing classical least squares under certain conditions [9], is performed with a property matrix (Fig. 3). The algorithm is essentially implemented on each column of the property matrix individually, calculating specific regression vectors for each column. If the purpose was only to investigate tadalafil and lactose, a property vector would be enough to distinguish the two. However, in this hybrid supervised/unsupervised strategy we investigated not only tadalafil and lactose but also six other unsupervised ingredients as well. Hence, we needed more than one coordinate axis (Fig. 5) to be able to increase the selectivity of this model. The ultimate purpose of the PLS algorithm is to create an optimum set of regression vectors by correlating the training matrix of *X* to a property matrix of *Y*. It finds a linear relation between *X* and *Y*, using a regression coefficients matrix B and an error matrix *E* (equation in Fig. 3). Fig. 3 shows how we set the PLS-DA algorithm to generate a regression coefficients matrix B for LC-SLM-CD compressive detection filter generation. The resolution element of LC-SLM-CD instrument for this study was 128; each Raman spectrum obtained consists of 128 wavenumber units. The training matrix *X* is made from two spectra of pure tadalafil and lactose samples and the property matrix is composed of only 100s and 0s. For a two-component system such as tadalafil and lactose, the number of latent vectors (LV) to be calculated that accounts for ~100% variability in both the training matrix *X* and property matrix *Y* is typically two. The output of a PLS-DA analysis with two LVs using a 2-column property matrix *Y* is four regression vectors sorted by their contribution to the changes in *X* and *Y*. (b1, b3) and (b2, b4) correspond to the regression vector sets for the first (tadalafil = 100) and second



* The set-up is shown based on 2 LVs (although there are 4 LVs in total and 4 pairs of regression vector corresponding to each LVs).

** b3 and b4 correspond to the tadalafil and lactose regression vectors, respectively. They account for ~100% variance in both X and Y.

Fig. 3. Schematic of PLS-DA filter function set-up to generate regression coefficients vectors to be used for LC-SLM-CD filter production trained on only two components; tadalafil and lactose. *E* represents the residual matrix.

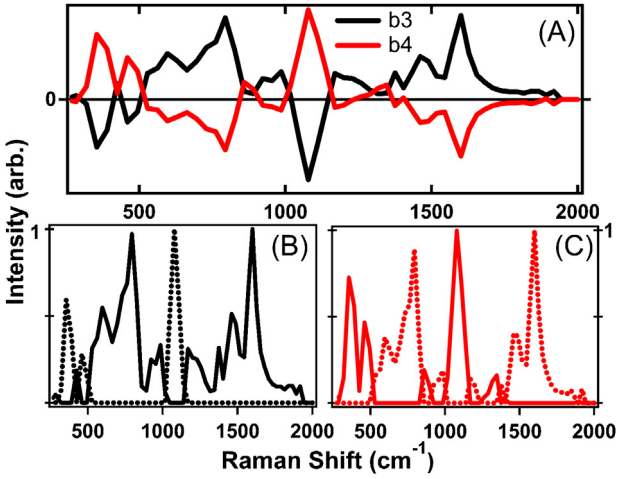


Fig. 4. PLS-DA derived compressive detection filter generation for tadalafil and lactose. (A) PLS-DA output regression vectors using spectra from Fig. 2. Black line is when tadalafil is coded 100 and red line is when lactose is coded 100. (B and C) Splitting of the vectors b3 in (B) and b4 in (C) into positive (solid) and the absolute value of the negative portion (dashed). All portions are scaled to a maximum value of 1.

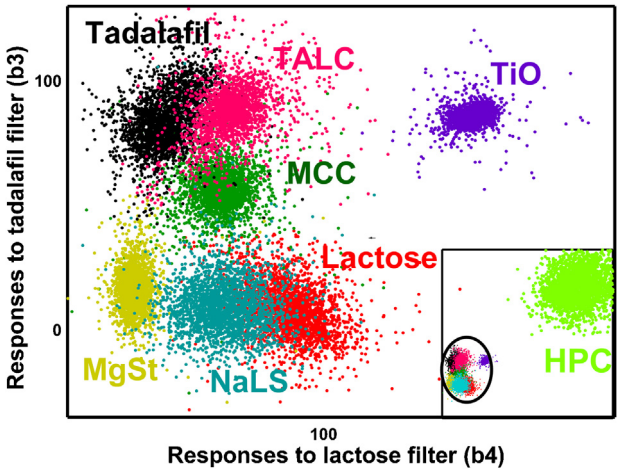


Fig. 5. LC-SLM-CD responses of raw pharmaceutical powders to the filters trained on tadalafil and lactose. The abscissa represents the filter where lactose is coded as 100 and the ordinate represents where tadalafil is coded 100. The response values of HPC are located away from those of the rest of the samples (circled in the inset). For better illustration, the HPC cluster is shown in the inset. Each cloud represents ~2500 responses of ~2500 spectra measured on each powder sample.

columns (lactose = 100) of the property matrix, respectively. The first vector pair, [b1, b2] (for LV1), accounts for only about 50% of the changes in the data set and they are not suitable for LC-SLM-CD applications. In a two component tadalafil-lactose system, the second pair, third and fourth regression vectors [b3, b4] (for LV2) is used for LC-SLM-CD filter generation (Fig. 3). Because the changes in Raman signal are uncorrelated with the noise, spectral features and noise tend to be separated into different vectors. The additional information contained in the third [b5, b6] (for LV3) and fourth [b7,b8] (for LV4) vector pairs (not shown in Fig. 3) is likely to come mostly from noise and should not be considered for filter applications.

Fig. 4(A) shows the output regression vectors of the PLS-DA analysis for tadalafil vs. lactose system. The black curve (b3) represents the regression vector where tadalafil is coded as 100 and lactose 0 into the PLS-DA algorithm, whereas the red curve (b4) is the output vector when lactose is coded as 100 and tadalafil as 0. Hence, the LC-SLM-CD response value of 100 to the regression vector b3 indicates that the signal is coming from tadalafil while the response value of 100 to b4 indicates it is lactose. Examination of the regression vectors in Fig. 4A indicate that each vector has predominantly positive features of the spectrum coded 100 and negative features of the spectrum coded 0. Before these vectors are loaded onto SLM, negative parts should be converted into positives by splitting each vector into two (SLM does not recognize negative vectors) and then taking the absolute value of the negative part to convert it to a positive value. Fig. 4B and C illustrate the way that b3 and b4 are split into two portions respectively, each of which are non-negative functions, and scaled to a maximum value of 1 which corresponds to the maximum SLM transmittance. As a result, four LC-SLM-CD filters are constructed from two PLS regression vectors shown in Fig. 4A. Later, the results obtained on these four filters are reconstructed to give the final scores values for each corresponding b3 and b4 vectors [4]. At the end, the total number of filters was five, including one more for normalization purpose in addition to four component filters. The normalization filter is formed by setting all the pixels on SLM to 100% transmittance, allowing the LC-SLM-CD system to detect all photons reaching the SLM.

When PLS-DA regression vectors are used as filters, the LC-SLM-CD produces the scores as response values. Thus, the detected signal coming from the sample is classified based on the measured score values of the spectrum with each PLS-DA filter. Fig. 5 shows the scores, which are obtained from the application of two PLS-DA regression vectors (b3 and b4) whose positive and negative components form the actual filters loaded on the SLM, plotted against each other for the PLS-DA model describing the differences between tadalafil and lactose. Before constructing Fig. 5, responses to each filter are normalized by dividing them by the responses to the normalization filter in order to minimize variance due to instrumental fluctuations and morphological differences in the sample. The abscissa of Fig. 5 denotes the responses to the filter where lactose is coded as 100 in PLS-DA algorithm. The ordinate, on the other hand, represents the responses to the filter where tadalafil is coded 100. 2500 points on each sample were collected with a collection time of 10 ms per filter. Each cluster in the figure corresponds to the filter responses of these 2500 spectra coming from various locations on each sample powder placed in 96-well plates (some obvious outliers are deleted in each cluster). The total collection time of 2500 points was about 2 min. A separation between tadalafil and lactose can clearly be seen in the figure. Also it shows the responses of the other six unsupervised pharmaceutical ingredients-MCC, MgSt, TiO, talc, NaLS and HPC-measured by the LC-SLM-CD Raman system using the same filters trained for only tadalafil and lactose. Fig. 5 illustrates that in all cases, except between lactose and NaLS, we were able to obtain good separation between the clusters. The high discrimination of tadalafil and lactose is expected since the filters are trained to describe the differences between them. It is particularly impressive that the responses of other raw ingredients to the same filters were, in fact, able to classify and discriminate them from each other as well although we see partial overlap between lactose and NaLS.

The accuary of classification on LC-SLM-CD can be improved by increasing the integration time for each filter. The powder mixture of two common excipients, lactose and magnesium stearate layered on a glass slide, are analyzed in order to evaluate how integration time affects the classification. Four sets of PLS filters that are trained on lactose and magnesium stearate (in addition to one normalization filter function) are generated to classify them. Fig. 5 contains LC-SLM-CD chemical classification plots obtained using different integration times of 500 μs, 2.5 ms, and 10 ms per point for each filter function. Even a 500 μs integration time per filter was able to generate a decent classficiation plot although better classification is achieved with increasing integration time.

Fig. 6.

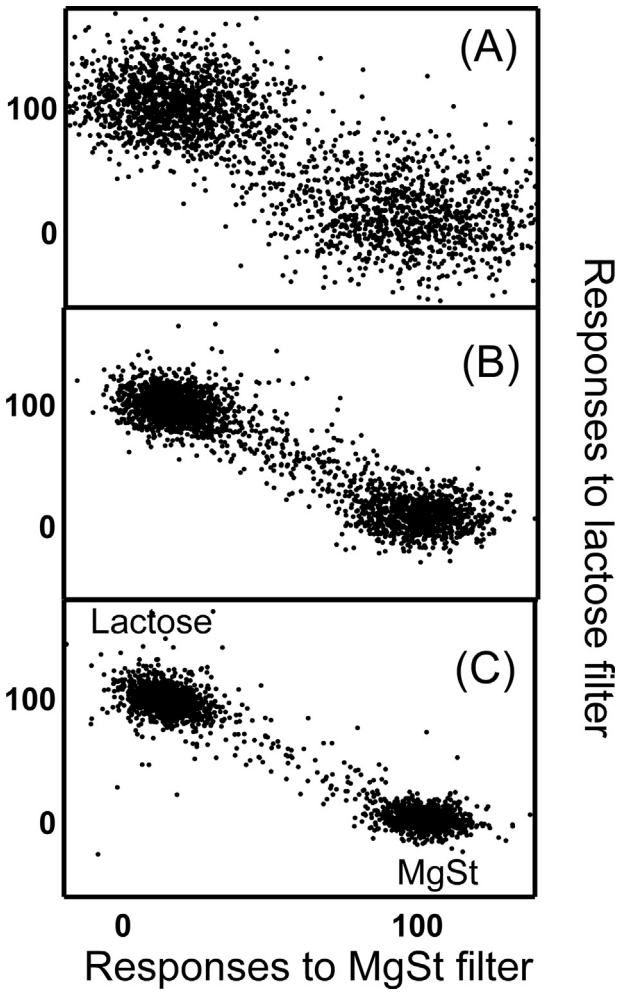


Fig. 6. LC-SLM-CD classification of magnesium stearate (MgSt) and lactose at different accumulation times. Accumulation times per filter: (A) 500 μs; (B) 2.5 ms; (C) 10 ms.

4. Conclusion

We have demonstrated the feasibility of a hybrid supervised/unsupervised hardware-based hyperspectral compressive detection strategy for rapid classification of pharmaceutical ingredients with recently designed LC-SLM-CD Raman instrument. This instrument is a near-infrared Raman spectrometer which is optimized for high speed applications using single-channel, low noise compressive detection strategy. With successful application of a multivariate projection technique PLS-DA, we report how LC-SLM-CD Raman spectroscopy can be a fast, effective analytical method for PAT applications to identify and classify various pharmaceutical raw ingredients. The method was able to discriminate several raw material components by applying PLS-DA filters trained to identify only two components. Collection of spectral responses per single location on the sample takes only 50 ms (10 ms per filter), giving almost instant results. The total collection time for 2500 points was only about 2 min, which would typically take about 42 min with a typical collection time of 1 s per spectrum using a traditional CCD-based Raman spectrometer. Although the LC-SLM-CD system was built for fast hyperspectral Raman imaging applications, this study shows that it can be readily adapted to other high speed analytical applications. Rapid and accurate techniques for classification of raw ingredients for manufacturing are an essential step in the implementation and success of the PAT program. This study is an important step toward the acceptance of LC-SLM-CD Raman spectroscopy as a valuable multivariate PAT sensor for identity testing of raw materials in the pharmaceutical industry. Although the collection time per point was 10 ms per filter, it is also shown that shorter accumulation times may be adequate. Thus, we conclude that the LC-SLM-CD Raman instrument has the potential to monitor real time manufacturing processes *in*, *on* or immediately *at* the process stream when combined with an appropriate probe. Most importantly, our implementation of this hardware-based hyperspectral compressive detection strategy demonstrates the feasibility of training a LC-SLM-CD system using a relatively small library of important pharmaceutical ingredients and then using that system not only to validate the identity of compounds in the library, but also to determine that a given sample lies outside the training library. Moreover, the response obtained from such an unknown compound could be used to assist in its subsequent chemical identification, by comparing the measured response against the response of various compounds (outside the training library). Such an LC-SLM-CD based screening procedure could greatly speed up and simplify the process of chemically identifying the unknown.

More generally, this work can be considered as a first step toward the fast imaging application of LC-SLM-CD Raman spectroscopy on solid pharmaceutical formulations. The set of digital filters used for this study can be applied to an intact Cialis® tablet to image tadalafil, lactose, MgSt, MCC, talc, TiO, NaLS and HPC to determine each component's spatial location in the tablet. The 2500 response values obtained in this study can be statistically analyzed to set the threshold values for each component in the image. Then, the values of LC-SLM-CD responses to both filters on each spatial position on the tablet would reveal the identity of the component on that specific location. Accordingly, a composite image of the tablet, where the eight components are displayed, can be formed from the LC-SLM-CD response values. For example, an image with an area of 2 mm × 2 mm with 20 μm spatial resolution takes about 3 h to collect with a typical 1 s collection time on CCD-based Raman instrument while it would take only about 8 min to collect the same area on LC-SLM-CD system with the filter set used here and at the end, an image of eight components in the Cialis tablet would be created. If the purpose, on the other hand, is to generate an image of only one component, such as the active ingredient tadalafil, then a simple univariate filter (plus an "all-on" filter) would be sufficient. In this case, imaging the same area on the tablet would only take about 3 min with 10 ms per filter collection time although shorter accumulation time may be possible (as opposed to several hours that would be required to produce a full spectral image of the same area).

Acknowledgements

This work was supported in part by a grant from the Lilly Endowment Inc.to the College of Pharmacy.

References

[1] US FDA, PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, 2004.
[2] K.A. Bakeev, Process Analytical Technology, 1st ed., Blackwell Publishing, Oxford, 2005.
[3] R.L. Mccreery, Raman Spectroscopy for Chemical Analysis, 1st ed., John Wiley & Sons, New York, 2000.
[4] B.M. Davis, A.J. Hemphill, D. Cebeci Maltas, M.A. Zipper, P. Wang, D. Ben-Amotz, Multivariate hyperspectral raman imaging using compressive detection, Anal. Chem. 83 (2011) 5086–5092.

[5] M. Blanco, M.A. Romero, Near-infrared libraries in the pharmaceutical industry: a solution for identity confirmation, Analyst 126 (2001) 2212–2217.
[6] P.J. Gemperline, L.D. Webber, F.O. Cox, Raw materials testing using soft independent modeling of class analogy analysis of near-infrared reflectance spectra, Anal. Chem. 61 (1989) 138–144.
[7] P. Geladi, B.R. Kowalski, Partial least-squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[8] P.J. Treado, M.D. Morris, A thousand points of light: the Hadamard transform in chemical analysis and instrumentation, Anal. Chem. 61 (2008) 723A–734A.
[9] J.M. Amigo, C. Ravn, N. Gallagher, R. Bro, A comparison of a common approach to partial least squares-discriminant analysis and classical least squares in hyperspectral imaging, Int. J. Pharm. 373 (2009) 179–182.