Journal of Pharmaceutical and Biomedical Analysis 103 (2015) 35–43

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Distribution of a low dose compound within pharmaceutical tablet by using multivariate curve resolution on Raman hyperspectral images

Mathieu Boiret[a,*], Anna de Juan[b], Nathalie Gorretta[c], Yves-Michel Ginot[a], Jean-Michel Roger[c]

[a] Technologie Servier, Orléans, France
[b] Grup de Quimiometria, Dept. Química Analítica, Universitat de Barcelona, Spain
[c] IRSTEA, UMR ITAP 361, Montpellier, France

ARTICLE INFO

Article history:
Received 14 August 2014
Received in revised form 15 October 2014
Accepted 23 October 2014
Available online 31 October 2014

Keywords:
Multivariate curve resolution
Alternating least squares
Raman hyperspectral imaging
Active and excipient distributions
Low dose compound

ABSTRACT

In this work, Raman hyperspectral images and multivariate curve resolution-alternating least squares (MCR-ALS) are used to study the distribution of actives and excipients within a pharmaceutical drug product. This article is mainly focused on the distribution of a low dose constituent. Different approaches are compared, using initially filtered or non-filtered data, or using a column-wise augmented dataset before starting the MCR-ALS iterative process including appended information on the low dose component. In the studied formulation, magnesium stearate is used as a lubricant to improve powder flowability. With a theoretical concentration of 0.5% (w/w) in the drug product, the spectral variance contained in the data is weak. By using a principal component analysis (PCA) filtered dataset as a first step of the MCR-ALS approach, the lubricant information is lost in the non-explained variance and its associated distribution in the tablet cannot be highlighted. A sufficient number of components to generate the PCA noise-filtered matrix has to be used in order to keep the lubricant variability within the data set analyzed or, otherwise, work with the raw non-filtered data. Different models are built using an increasing number of components to perform the PCA reduction. It is shown that the magnesium stearate information can be extracted from a PCA model using a minimum of 20 components. In the last part, a column-wise augmented matrix, including a reference spectrum of the lubricant, is used before starting MCR-ALS process. PCA reduction is performed on the augmented matrix, so the magnesium stearate contribution is included within the MCR-ALS calculations. By using an appropriate PCA reduction, with a sufficient number of components, or by using an augmented dataset including appended information on the low dose component, the distribution of the two actives, the two main excipients and the low dose lubricant are correctly recovered.

1. Introduction

In the last decade, the use of imaging coupled with vibrational spectroscopies (near infrared, mid infrared, fluorescence and Raman) has grown quickly in research and development environments. The spatial and spectral information contained in hyperspectral images can be associated with the distribution of the different constituents within the sample. Different areas such as polymer research [1], biomedical analysis [2], environment field [3] and pharmaceutical development [4] are using these new analytical tools based on vibrational hyperspectral imaging. During the analytical lifecycle of a pharmaceutical drug product, hyperspectral imaging became a very powerful technique to explore the compound distributions on the tablet surface or within a powder mixture [5]. This technology appeared as innovative and promising to ensure the final quality of the drug product [6] from the development to the production.

Because of the huge amount of data contained in hyperspectral images, a direct interpretation of the acquired images is often not possible. Therefore, several chemometric tools have previously been applied [7,8]. Qualitative analyses such as principal component analysis (PCA) have already been used with near infrared [9] and Raman [10] chemical imaging in order to study the compound distribution in a sample. Since PCA is mainly linked to the dataset variability and as calculated loadings do not have chemical meaning, this approach is used as a descriptive method. To extract

* Corresponding author. Tel.: +33 238 238 175.
E-mail address: mathieu.boiret@fr.netgrs.com (M. Boiret).

http://dx.doi.org/10.1016/j.jpba.2014.10.024
0731-7085/

quantitative information at a global and pixel level, principal component regression (PCR) and partial least squares regression (PLS-R) have already demonstrated through several studies that they were powerful chemometric techniques [11,12]. However, these methods can be time consuming and difficult to implement as they usually require a calibration step to develop predictive models. To overcome this problem, resolution methods seem to be a good alternative.

The aim of resolution methods is to provide the distribution maps and pure spectra related to the image constituents of a sample from the information contained in the raw image [13]. Multivariate curve resolution-alternating least squares (MCR-ALS) is one of the most famous tools applied on hyperspectral images [14,15]. MCR-ALS decomposes the initial data in a bilinear model, assuming that the observed spectra (i.e. each pixel of the image) are a linear combination of the spectra of the pure components in the system. In order to ensure an accurate resolution, constraints have to be used during the optimization process. Indeed, due to rotational or intensity ambiguities, resolution of a multicomponent hyperspectral image might not be unique [16]. Different constraints were established and tested [17,18]. In image resolution, non-negativity, spectral normalization and local rank analysis are generally the most successful tools. Local rank analysis describes the spatial complexity of an image by identifying the rank of a pixel neighbourhood area. Combined with reference spectra of the image constituents, the absence of one or more specific constituent in a pixel can be highlighted. Some constraints used for the resolution of a chemical process, such as unimodality, closure or hard-modelling should not be used to analyze hyperspectral images because concentration profiles in the pixels of an image do not present the global continuous evolution that process profiles have [19].

Raman chemical imaging, because of its advantages such as negligible sample preparation, high chemical specificity and high spatial resolution, emerges as a new analytical tool in the quality control process of a solid drug product [20]. Final drug products are usually manufactured by using at least one active pharmaceutical ingredient (API) and several excipients. To improve powder flowability, most of the pharmaceutical manufacturing process includes a lubricant in the final drug formulation [21]. This compound is commonly present in a very low concentration in the powder blend and a spectroscopic bulk analysis will not be able to extract its contribution. Indeed, the corresponding variance of this constituent is very weak comparing with the other compounds of the sample. PCA, which aims at describing the directions of maximum global variance in the data, may have difficulties in retrieving information linked to a low dose constituent when the variance allocated to this component is similar in level to noise, which is often large in hyperspectral images. By offering the possibility to acquire images with a high spatial resolution, Raman chemical imaging coupled with appropriate chemometric methods appears as a promising technique to detect a low dose compound within a solid drug formulation.

In this work, MCR-ALS was applied on Raman chemical imaging data in order to provide the distribution of actives and excipients in a commercialized tablet. MCR-ALS was challenged by trying to identify the low dose lubricant in the hyperspectral image. The effect of using algorithms driven by finding directions of maximum variance explained is studied. In this sense, the effect linked to the first step of noise-filtering based on PCA, which is often used in MCR-ALS to remove noise and non-useful spectral information, is studied. By applying MCR-ALS on a noise-filtered PCA matrix, it is shown that the information of the low dose constituent may be lost during data reduction. The comparison between the MCR-ALS decomposition on a filtered and a non-filtered PCA matrix is presented. Moreover, to keep the low dose constituent information during the PCA reduction, calculations are performed on an augmented matrix including the low dose constituent spectrum. The necessity of using appropriate pre-processing methods and constraints to find out the correct information linked to these low dose constituents is emphasized. This article shows the strategies to be followed in MCR-ALS analysis to retrieve correct information for low dose image constituents, from pre-processing, conditions to drive the iterative optimization to proper inclusion of constraints.

2. Materials and methods

2.1. Samples

A commercial coated tablet of Bipreterax®, prescribed for arterial hypertension treatment and commercialized by "Les Laboratoires Servier", was used for the study. It is also known as Perindopril/Indapamide association. Final drug product contains respectively 4 mg of Perindopril (API1) and 1.25 mg of Indapamide (API2). Actives are known to have several solid state forms, but only one of them is present in this formulation. Major core excipients are lactose monohydrate, microcrystalline cellulose (Avicel). Magnesium stearate (MgSt), which is used as a lubricant, was added to the blend before compression with a theoretical mass concentration corresponding to 0.5% (w/w). In order to analyze the tablet core, the coating was removed by eroding the sample with a Leica EM Rapid system (Leica, Wetzlar, Germany). A visual examination of the tablet did not provide any information concerning the distribution of the different compounds within the tablet.

2.2. Raman imaging system

The image was collected using a RM300 PerkinElmer system (PerkinElmer, Waltham, MA) and the Spectrum Image version 6.1 software. The microscope was coupled to the spectrometer and spectra were acquired through it with a spatial resolution of 10 μm in a Raman diffuse reflection mode. Wavenumber range was 3200–100 cm^{-1} with a resolution of 2 cm^{-1}. Spectra were acquired at a single point on the sample, then the sample was moved and another spectrum was taken. This process was repeated until spectra of points covering the region of interest were obtained. A 785 nm laser with a power of 400 mW was used. Two scans of 2 s were accumulated for each spectrum. An image of 70 pixels per 70 pixels corresponding to 4900 spectra was acquired for a surface of 700 μm by 700 μm.

2.3. Pre-processing

Data were preprocessed in order to remove non-chemical biases from the spectra (scattering effect due to non-homogeneity of the surface, interference from external light source, spikes due to cosmic rays, random noise). First of all, data were spike-corrected in order to reduce the effect of cosmic rays [22]. The spectral range was reduced in order to focus only on the region of interest, corresponding to a Raman shift from 1800 cm^{-1} to 200 cm^{-1}. Reduced spectra were preprocessed by asymmetric least squares (AsLS) to correct baseline variations due to fluorescence contributions [23]. Finally, to enhance slight spectral variations, a Savitzky–Golay first derivative with a 2nd order polynomial smoothing on a 9 points window [24] was applied.

2.4. Multivariate curve resolution-alternating least squares (MCR-ALS)

A brief description of the MCR-ALS algorithm is given here. The algorithm was previously described in detail in Refs. [17,18]. As any resolution methods, the main goal of MCR-ALS is decomposing the original matrix $\mathbf{D}_{(\mathbf{n},\mathbf{p})}$ (**n** samples or rows and **p** variables or