Journal of Pharmaceutical and Biomedical Analysis 85 (2013) 207–212

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Short communication

Characterization of sildenafil citrate tablets of different sources by near infrared chemical imaging and chemometric tools

Guilherme P. Sabin[a], Valeria A. Lozano[b], Werickson F.C. Rocha[c], Wanderson Romão[d], Rafael S. Ortiz[e], Ronei J. Poppi[a,*]

[a] Institute of Chemistry, State University of Campinas, 13084-971 Campinas, SP, Brazil
[b] Instituto de Química Rosario (CONICET-UNR), Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina
[c] National Institute of Metrology, Quality and Technology (Inmetro), Directorate of Industrial and Scientific Metrology, Chemical Metrology Division, 25250-020, Xerém, Duque de Caxias, RJ, Brazil
[d] Federal Institute of Education, Science and Technology of Espírito Santo, 29106-010 Vila Velha, ES, Brazil
[e] Brazilian Federal Police, Ministry of Justice, Rio Grande do Sul Technical and Scientifical Division, 90160-092 Porto Alegre, RS, Brazil

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form 26 July 2013
Accepted 29 July 2013
Available online 6 August 2013

Keywords:
Sildenafil citrate
Chemical imaging
Near infrared spectroscopy
Multivariate curve resolution
Classification
Pattern recognition

ABSTRACT

The chemical imaging technique by near infrared spectroscopy was applied for characterization of formulations in tablets of sildenafil citrate of six different sources. Five formulations were provided by Brazilian Federal Police and correspond to several trademarks of prohibited marketing and one was an authentic sample of Viagra. In a first step of the study, multivariate curve resolution was properly chosen for the estimation of the distribution map of concentration of the active ingredient in tablets of different sources, where the chemical composition of all excipients constituents was not truly known. In such cases, it is very difficult to establish an appropriate calibration technique, so that only the information of sildenafil is considered independently of the excipients. This determination was possible only by reaching the second-order advantage, where the analyte quantification can be performed in the presence of unknown interferences. In a second step, the normalized histograms of images from active ingredient were grouped according to their similarities by hierarchical cluster analysis. Finally it was possible to recognize the patterns of distribution maps of concentration of sildenafil citrate, distinguishing the true formulation of Viagra. This concept can be used to improve the knowledge of industrial products and processes, as well as, for characterization of counterfeit drugs.

1. Introduction

Patents are a form of legal protection of intellectual property that provide exclusive rights to make, use, import, sell and offer for sale the invention for up to 20 years. The economic logic of this protection mechanism is that the profits provided by the production license of a patented product guarantee the patent owner the reinvestment in research and development of new products [1,2]. Social factors, however, may eventually prevail over this economic development engine aspect, discussing the possibility of patent infringement. One of these factors is the great technological discrepancy in peripheral countries in relation to developed countries, and their low purchasing power to buy the next-generation products manufactured by the major economic centers.

Viagra is the first drug approved to treat erectile dysfunction. Its mechanism is blocking the enzyme phosphodiesterase type 5 (PDE5), involved in the erection process. It has vasodilation properties and effects on blood pressure, and like nitrates, it works by the nitric oxide cyclic guanosine monophosphate pathway [3,4]. It is estimated that erectile dysfunction affects between 48% and 52% of men from 40 to 70 years. The sildenafil citrate (active ingredient of the Viagra) was registered in the European Union in 1991 by Pfizer. Nowadays, due to the expiration of the patent in Brazil, at least ten different companies are marketing this product, but in cheaper way. Additionally, the counterfeiting of Viagra tablets has become an important and dangerous problem for pharmaceutical market, where the Brazilian Federal Police has reported many seizures mainly in south region of Brazil (state of Parana). In 2007–2010 periods, the Federal Police reported a great numbers of seizures (371) being the counterfeit tablets market related to erectile dysfunction treatment responsible by 80% of the seizures. Therefore, to control the quality of new pharmaceutical formulations and distinguish between authentic and counterfeit tablets is necessary the development of powerful analytical tools. Although analytical methods such as chromatography [5], voltammetry [6] and colorimetric determination [7] were reported in Viagra tablets

* Corresponding author. Tel.: +55 19 35212134; fax: +55 19 35212134.
E-mail address: ronei@iqm.unicamp.br (R.J. Poppi).

0731-7085/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jpba.2013.07.036

analysis, they are time-consuming and require extensive sample preparation.

Due to its advantages such as non-destructive analysis, speed, and less consumption of chemicals, the near infrared spectroscopy (NIR) has been accepted in various fields of pharmaceutical industry [8,9]. NIR has the potential to provide increased process and product understanding which goes well with the process analytical technology (PAT) initiative of the Food and Drug Administration (FDA) [10].

Hyperspectral imaging shows a considerable promise for providing high-quality spectral information on active principle distribution within pharmaceutical formulations. The robust reliable combination of chemical (molecular spectroscopy) and physical (digital imaging) features have been successfully applied to diverse fields such as remote sensing [11], astronomy [12], agriculture [13], food [14] and pharmaceuticals [15].

Quantitative analysis of pharmaceutical samples using near infrared chemical imaging (NIR-CI) can be performed using partial least squares regression (PLS) [16]. However, this technique requires a complete calibration set of samples, where all constituents (analyte and interferences) must be present. In this case, the interferences do not need to be known, but present in all samples, the called "first order advantage".

Concerning the quantification purposes without needing a previous calibration model, multivariate curve resolution-alternating least squares (MCR-ALS) [17,18] may be presented as an alternative, since only initial information about pure spectra (or concentration) is need. Also, in advanced, it can present the called "second order advantage", where the analytes quantification can be performed in the presence of unknown interferences. This method decomposes the unfolded hyperspectral data, the matrix **X**, into the product of two matrices: **C** containing the concentration profiles and **S** containing the spectral profiles for each k component (Fig. 1 and Eq. (1)).

$$\mathbf{X} = \mathbf{C}\mathbf{S}^{\mathrm{T}} + \mathbf{E} \tag{1}$$

where **E** corresponds to the experimental error matrix.

To initiate the iterative MCR-ALS procedure, an initial estimation is needed for the spectral or concentration profiles. Different methods have been used for this purpose, such as evolving factor analysis [19], the determination of the purest variables [20] or the information about the sample concentrations [21].

This work aims to estimate the concentration distribution map of sildenafil citrate in tablets of different sources where the chemical composition of all excipients constituents is not truly known by using the multivariate curve resolution approach. In addition, the normalized histograms of images from active ingredient were grouped according to their similarities by hierarchical cluster analysis. This procedure make possible to recognize the patterns of distribution maps of concentration of sildenafil citrate, distinguishing the true formulation of Viagra.

2. Materials and methods

2.1. Experimental

Tablets containing sildenafil citrate as active ingredient of six different formulations from different sources were studied. These formulations were named as **A**–**F**. The formulations from **A**–**E** were provided by Brazilian Federal Police and correspond to several trademarks of prohibited marketing. The **F** formulation was an authentic sample of Viagra (Pfizer Ltda). For each formulation, it was performed an image acquisition of four tablets.

The acquisition of images was obtained by NIR-CI technique using Spotlight 400N FT-NIR Imaging by PerkinElmer. The mapping measurements were performed four times per sample type, spatial resolution of 25 μm, 16 scans and spectral range between 6500 and 4000 cm^{-1}. The data array (80 × 80 pixels and 158 wavelengths) was obtained directly on surface of the tablet (after coating removal).

2.2. Data treatment

The raw data were transformed to inverse logarithm of the reflectance values (pseudo-absorbance) and unfolded for further preprocessing by multiplicative scattering correction (MSC) [22]. Since only the sildenafil citrate spectrum is known in advance, the tool choose for construction of the distribution maps of concentration was the MCR as a quantitative way. The ALS optimization was initialized by loadings of principal component analysis (PCA). The following constraints were used as a way to minimize rotation ambiguity in MCR calculations: external spectral knowledge of the active ingredient, non-negativity and closure for concentration. Thus, the standard spectrum of sildenafil citrate was compared among all loadings per sample and substituted by its most similar loading profile and a new optimization process using ALS was performed.

The use of loadings for initialization of the MCR-ALS can be considered problematic for optimization of the **C** matrix (concentration), since the rotation ambiguity is present in this situation. In other words, the quantitative approach used in multivariate curve resolution, Eq. (1), would be confounded by qualitative information of **T** matrix (scores) and **P** matrix (loadings) obtained by PCA analysis, Eq. (2).

$$\mathbf{X} = \mathbf{T}\mathbf{P}^{\mathrm{T}} + \mathbf{E} \tag{2}$$

In this case, the **C** matrix would bring **T** matrix qualitative information because of the initializations by loadings. In this sense, the use of the purest **S** matrix and constraints are frequently a way for minimization of the ambiguities and recovery quantitative information. However, the purest spectra may present high condition number in relatively homogeneous images (i.e. high similarity among all spectral information) while loadings are always orthogonal with condition number equal to one. In MCR-ALS, **E** matrix (errors) can be seen as a lack of fit of the product of **C** matrix (concentrations) and **S** matrix (pure spectra) for recovery the **X** matrix. For good results, it is desirable decreasing the sensitivity of the **E** matrix. In this direction, a lower condition number of **S** matrix contributes to highest robustness for **C** matrix. Due to use of loadings combined with sildenafil citrate pure spectrum for initialization of MCR-ALS, a high orthogonally was reached.

After convergence of the MCR algorithm, the spectrum and concentration of the sildenafil in each pixel are obtained. Then, a distribution map of concentration can be computed, generating an image of the concentration for each sample. Based on this image the samples can be compared according their similarities. The images can be translated into histograms of frequency distribution of concentrations. This type of result analysis removes the spatial components of the acquired information, but retains the ability to study the distribution profile, i.e., the homogeneity of the active ingredient information.

The histograms were built by placing the concentration values in the abscissa axis and the concentration frequencies in the ordinate axis. After that, the histograms were grouped according to their similarities by hierarchical cluster analysis (HCA) [23]. Calculations were performed in Matlab version 7.8 using routines developed in the laboratory and the MCR Toolbox provided by Tauler [24].