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 xxx

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].

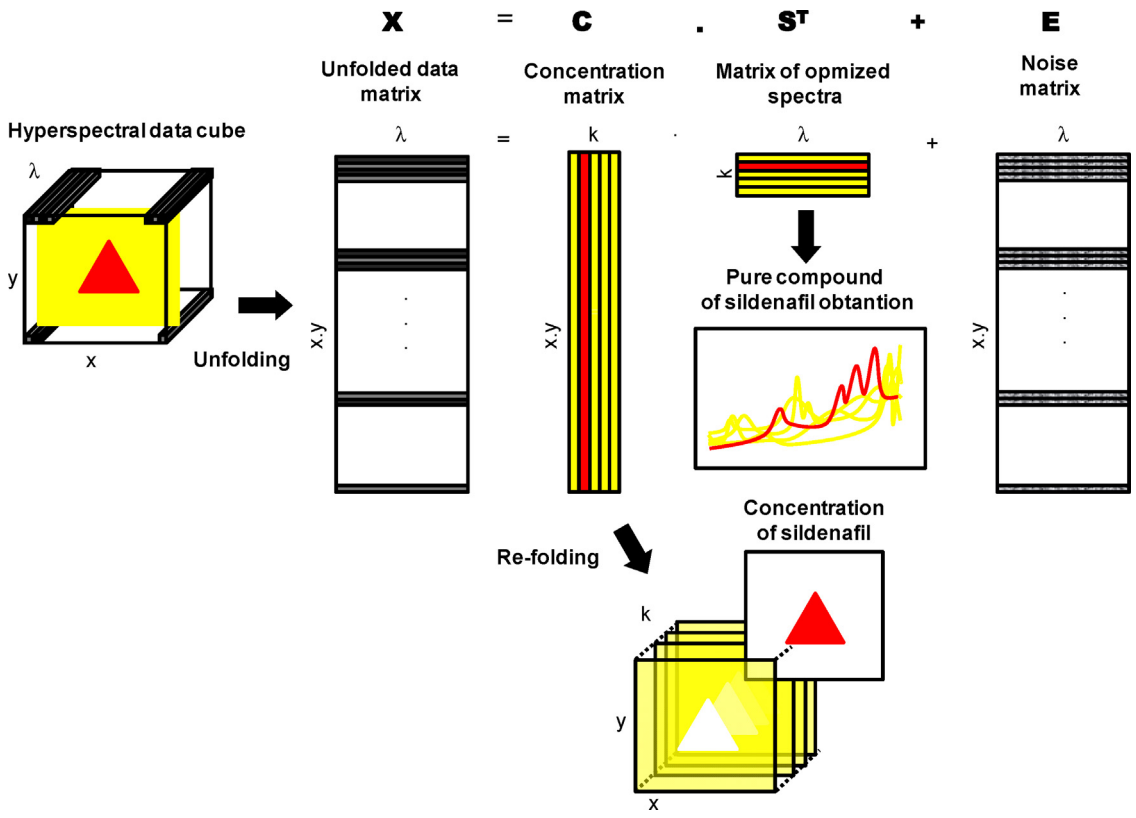

Fig. 1. Scheme of MCR-ALS analysis application to the hyperspectral cube.

3. Results and discussion

The results of MCR-ALS optimization are presented in Fig. 2, where is possible to observe excellent recoveries for the active ingredient spectra for all formulations samples. In Fig. 2, the dotted spectra were obtained directly from a standard of sildenafil citrate, while the line spectra were obtained by MCR-ALS. Note that the information recovered from optimization process in MCR is highly correlated with the known spectrum of the active ingredient standard, reaching selectivity for quantitative approach. Moreover, loadings assurance a lower similarity among all excipients and sildenafil citrate profiles, which help to increase the sensitivity of the results.

This procedure is especially interesting for analysis where the components are not totally known, but only part of the information is available. The lack of fit lower than 1% between the MCR-ALS resolution results and the original **X** matrices added to good recovery of the sildenafil citrate spectrum assurance reliable images. The calculation of the lack of fit was performed by using the Eq. (3):

$$\text{Lack of fit}(\%) = \sqrt{\frac{\sum (x_{ij}^{*} - x_{ij})^2}{\sum x_{ij}^2}} \times 100 \quad (3)$$

where x_{ij} is an element of the experimental matrix **X** and x_{ij}^{*} the element of the MCR-ALS reproduced matrix $\mathbf{X}^{*}$.

In Fig. 3, concentration maps of four different tablets for each source of the drug formulation (**A**–**F**) are showed. Note that for the drug **A**, there are a lot of big heterogeneous regions (i.e. excipient has big particle size) also observed on tablets through to optical microscopic view. In second line **B** a slightly different profile arises, smaller and denser point of active may be seen. In next two samples, respectively **C** and **D**, more homogeneous images are brought. Those drugs have very similar characteristics; maybe by represent the same brand, but different packaging, batch and seizure. Likewise **A**, the drug formulation **E** is recognized by low homogeneity that divides its pixels at nearly absence or presence of the sildenafil information. The drug formulation **F**, the original drug formulation, presents a very homogeneous profile of concentration. Moreover, similar pattern of distributions among image replicates is possible to observe in this figure.

Fig. 4 shows the normalized histograms in order that the mean concentration is equal to one. Thus, it becomes possible to analyze profile of histograms independently of their mass fraction on the tablets surface.

The histogram is a graphical technique that provides information both the skewness and kurtosis of the data set. In this case, it is represented by the frequency distribution of concentrations. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. Negative kurtosis would indicate a flat distribution, which is said to be platykurtic. Positive kurtosis would indicate a peaked distribution, which is said to be leptokurtic. Finally, the normal distribution has zero kurtosis, and it is said to be mesokurtic. In Fig. 4, it is possible to classify the kurtosis from the concentration distribution of **D** and **F** drugs in mesokurtic, because among the six drugs analyzed they have a more homogeneous concentration distribution on the sample surface. For **A**, **B** and **E** drugs the kurtosis can be classified as platykurtic, because they are less homogeneous. For **C** drugs the kurtosis can be classified as leptokurtic.

Skewness is a measure of the distribution symmetry, or more precisely, the lack of symmetry. From observation of the Fig. 4, it can be inferred that in **A** and **E** formulations the sildenafil citrate concentration distribution on the surface is less symmetrical due to the particle size heterogeneity in those formulations. By the other hand, **B**, **C**, **D** and **F** formulations have more symmetrical distribution among the formulations studied, since the size of the particles are more homogeneous. Therefore, it can be concluded that by kurtosis and skewness histogram classification it is possible to study

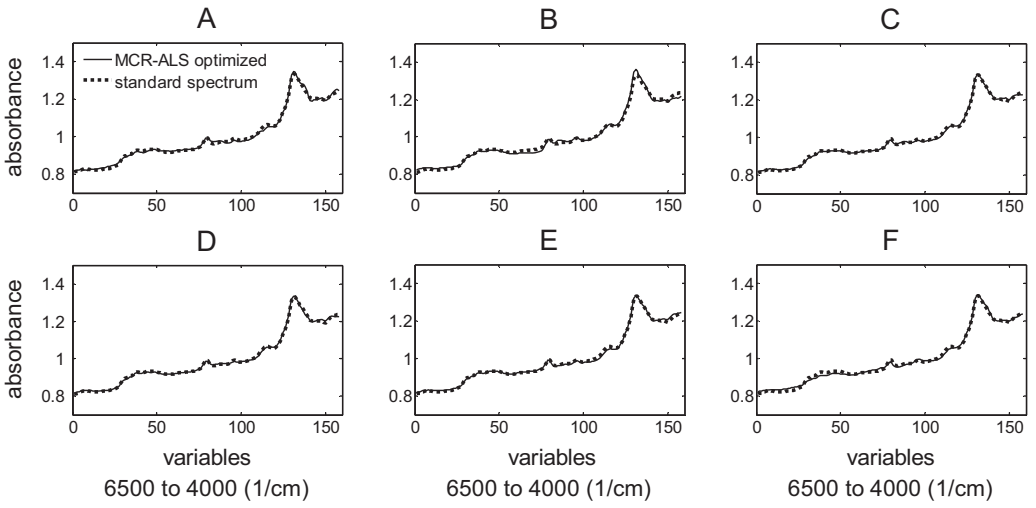

Fig. 2. Overlap of the spectra of sildenafil citrate: standard and optimized by MCR-ALS for the drug formulations **A–F**.

drug formulation homogeneity. Also they can be used as an indicative to characterize the formulations.

In order to provide a cluster analysis of the distribution of sildenafil citrate concentrations in tablets, it was used an established chemometric method called Hierarchical Cluster Analysis (HCA). It is based on the multivariate distances among the samples by using an agglomerative procedure. The objective has been to find groups of the different kind of formulations by similarity of concentration distribution map of the active. Thus, it is possible to obtain a way of recognition of images and its variability is an important parameter to study the similarities within and between samples. The dendrogram, Fig. 5, describes different clusters for each formulation and it suggests an effective method to identify the product or process of manufacture. This principle may be useful to identify different type

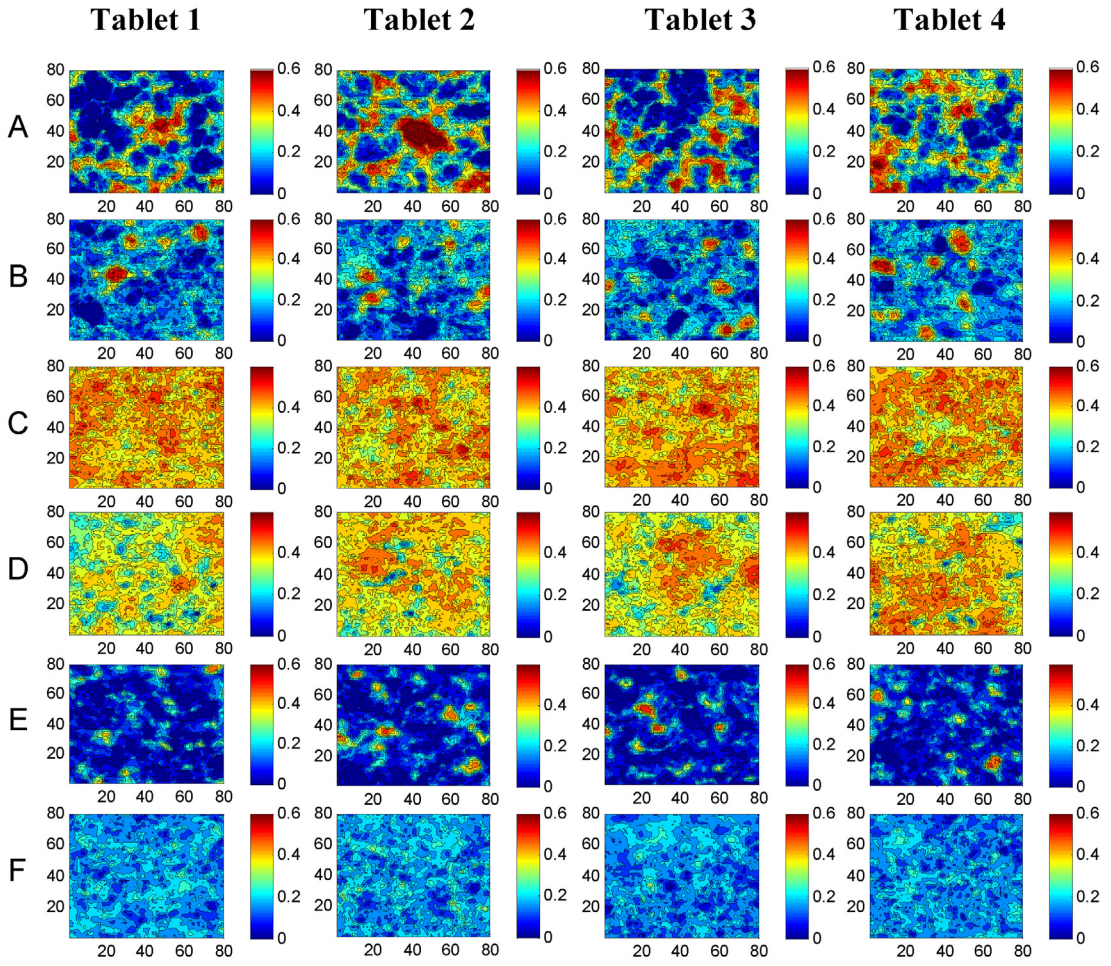

Fig. 3. Distribution maps of concentration to six different drugs (A–F) obtained by MCR-ALS.

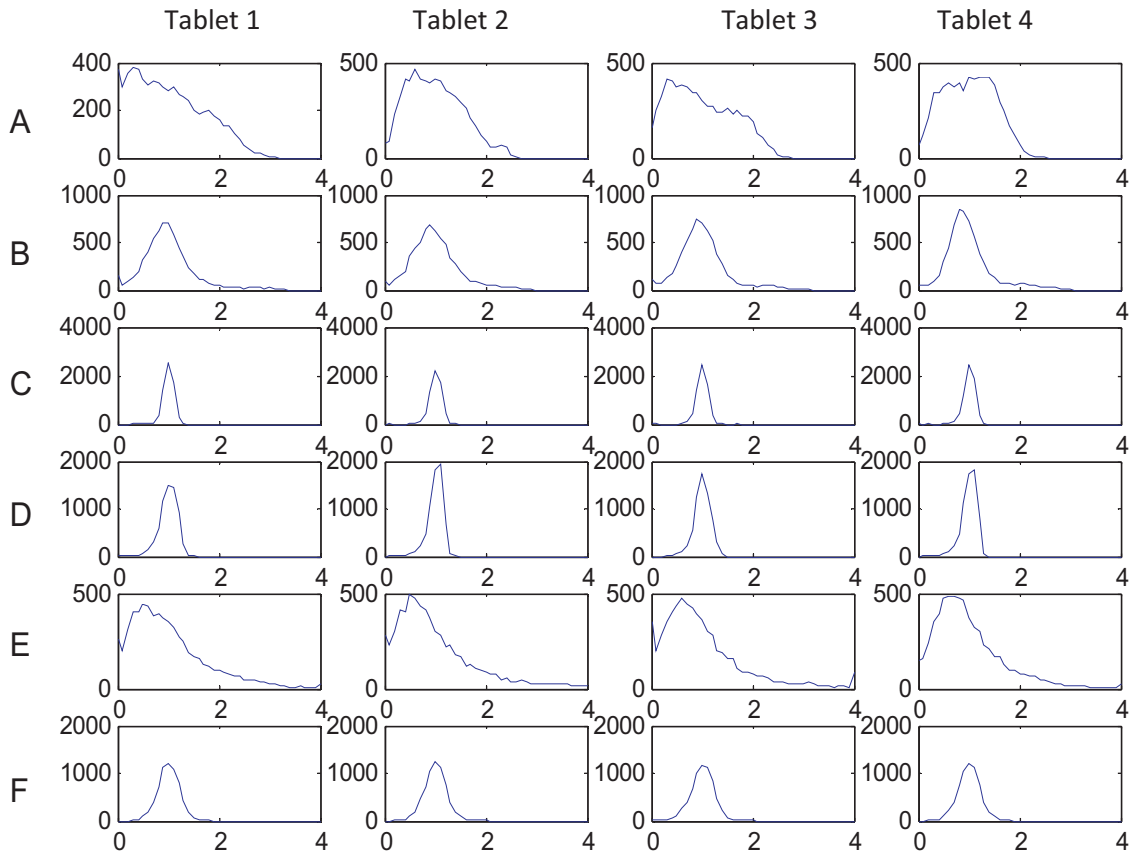

Fig. 4. Histogram of the images centered at unit for the drug formulations A–F.

of counterfeits including drugs that present the same composition, but different process of homogenization, particle size among other.

In Fig. 5, the images are identified 1–24 divided in 6 groups of 4 replicates. The images of the same formulation have been clustered due the similarity of the profiles of histograms for each product. This pattern is repeated in all images analyzed. However, it is less evident for formulations **C** (Fig. 5, samples 9–12) and **D** (Fig. 5, samples 13–16). Those formulations have the same specification and brand but different package characteristics.

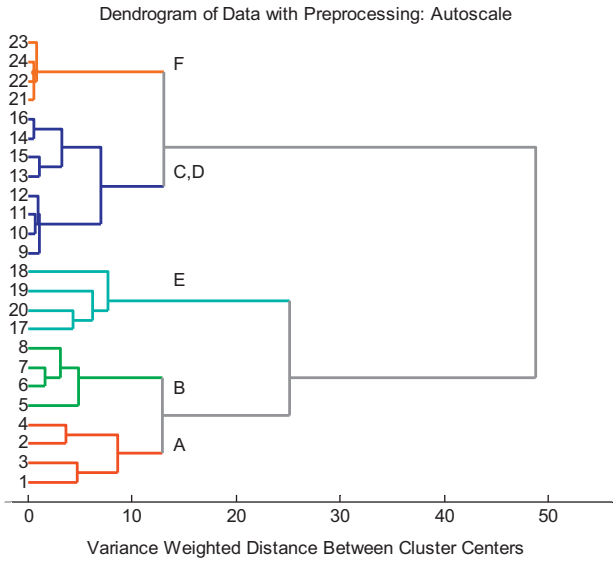

Fig. 5. HCA of histograms for drugs A–F.

In this direction, the dendrogram obtained by HCA shows that this is an appropriate procedure for clustering by multivariate similarity. However, more heterogeneous samples (flat histograms) are also less reproducible for the same image size. It is observed through the node of link of the clusters. In contrast, images of the original product **F** (Fig. 5, samples 21–24) showed a much better homogeneity and repeatability among images.

4. Conclusions

In this work, it was explored the technique of chemical mapping by near infrared spectroscopy in formulations with tablets of different sources in situations where the whole composition is not 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. The results obtained suggest an important way in which multivariate curve resolution – alternating least squares (MCR-ALS), by using the advantage of second order, can be successfully applied. Furthermore, the images of the active ingredient have been adequately grouped by their histograms, distinguishing the true formulation of Viagra. Thus, the presented methodology may be useful both for development of product and process, as well as, identification of the counterfeit in final products.

Acknowledgments

The authors thank Fapesp and CNPq for providing fellowships and financial support for doing this research work at the Institute of Chemistry of Unicamp, Brazil.

References

[1] B.H. Tsai, Do the infringement litigations of intellectual property, rights hinder enterprise innovation? An Empirical Analysis of the Taiwan IC Industry, vol. 46, 2010, pp. 173–203.
[2] G.S. Erickson, Inventive behavior and patent protection International Journal of Technology Management 18 (1999) 510–519.
[3] G. Pomara, G. Morelli, S. Pomara, S. Taddei, L. Ghiadoni, N. Dinelli, F. Travaglini, M. Dicuio, N. Mondaini, A. Salvetti, C. Selli, Cardiovascular parameter changes in patients with erectile dysfunction using Pde-5 inhibitors: a study with sildenafil and vardenafil, Journal of Andrology 25 (2004) 625–629.
[4] A.E. Gousse, M. Lambert, R. Kester, Oral pharmacotherapy to manage erectile dysfunction in spinal cord-injured men, Topics in Spinal Cord Injury Rehabilitation 8 (2002) 51–62.
[5] J.J. Berzas, J. Rodríguez, M.J. Villaseñor, A.M. Contento, M.P. Cabello, Validation of a capillary gas chromatographic method for the determination of sildenafil citrate in its pharmaceutical formulations (Viagra). Experimental design for evaluating the ruggedness of the method, Chromatographia 55 (2002) 601–606.
[6] K. Tyszczuk, M. Korolczuk, Voltammetric method for the determination of sildenafil citrate (Viagra) in pure form and in pharmaceutical formulations, Bioelectrochemistry 78 (2010) 113–117.
[7] Y.M. Issa, W.F. El-Hawary, A.F. Youssef, A.R. Senosy, Spectrophotometric determination of sildenafil citrate in pure form and in pharmaceutical formulation using some chromotropic acid azo dyes, Spectrochimica Acta A 75 (2010) 1297–1303.
[8] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies, Journal of Pharmaceutical and Biomedical Analysis 44 (2007) 683–700.
[9] T. De Beer, A. Burggraeve, M. Fonteyne, L. Saerens, J.P. Remon, C. Vervaet, Near infrared and Raman spectroscopy for the in-process monitoring of pharmaceutical production processes, International Journal of Pharmaceutics 417 (2011) 32–47.
[10] Food and Drug Administration Guidance for industry, PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, US Department of Health and Human Services Food and Drug Administration – Center for Biologics Evaluation and Research, Rockville, 2004.
[11] J. Verrelst, M.E. Schaepman, Z. Malenovsk, J.G.P.W. Clevers, Effects of woody elements on simulated canopy reflectance: implications for forest chlorophyll content retrieval, Remote Sensing of Environment 114 (2010) 647–656.
[12] C. Pina, L.H. Kidder, I.W. Levin, J.C. Fraser, J.F. Arens, E.N. Lewis, Infrared spectroscopic imaging: from planetary to cellular systems, Applied Spectroscopy 52 (1998) 106A–120A.
[13] S. Sommer, J. Hill, J. Mégier, The potential of remote sensing for monitoring rural land use changes and their effects on soil conditions, Journal of Agriculture, Ecosystems & Environment 67 (1998) 197–209.
[14] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, G. Downey, J.M. Frias, Hyperspectral imaging – an emerging process analytical tool for food quality and safety control, Trends in Food Science & Technology 18 (2007) 590–598.
[15] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, S.E. Bell, Recent applications of chemical imaging to pharmaceutical process monitoring and quality control, European Journal of Pharmaceutics and Biopharmaceutics 69 (2008) 10–22.
[16] W.F.C. Rocha, G.P. Sabin, P.H. Março, R.J. Poppi, Quantitative analysis of piroxicam polymorphs pharmaceutical mixtures by hyperspectral imaging and chemometrics, Chemometrics and Intelligent Laboratory Systems 106 (2011) 198–204.
[17] A. de Juan, R. Tauler, R. Dyson, C. Marcolli, M. Rault, M. Maeder, Spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis, Trends in Analytical Chemistry 23 (2004) 70–79.
[18] J.M. Amigo, J. Cruz, M. Bautista, S. Maspoch, J. Coello, M. Blanco, Study of pharmaceutical samples by NIR chemical-image and multivariate analysis, Trends in Analytical Chemistry 27 (2008) 696–713.
[19] H. Gampp, M. Maeder, C.J. Meyer, A.D. Zuberbuhler, Calculation of equilibrium constants from multiwavelength, spectroscopic data—IV: Model-free least-squares refinement by use of evolving factor analysis, Talanta 33 (1986) 943–951.
[20] W. Windig, J. Guilment, Interactive self-modeling mixture analysis, Analytical Chemistry 63 (1991) 1425–1432.
[21] G.P. Sabin, W.F.C. Rocha, R.J. Poppi, Study of the similarity between distribution maps of concentration in near-infrared spectroscopy chemical imaging obtained by different multivariate calibration approaches, Microchemical Journal 99 (2011) 542–547.
[22] H. Martens, J.P. Nielsen, S.B. Engelsen, Light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures, Analytical Chemistry 75 (2003) 394–404.
[23] N. Bratchell, Cluster analysis, Chemometrics and Intelligent Laboratory Systems 6 (1989) 105–125.
[24] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, Chemometrics and Intelligent Laboratory Systems 76 (2005) 101–110.