Study of the Homogeneity of Drug Loaded in Polymeric Films Using Near-Infrared Chemical Imaging and Split-Plot Design

GUILHERME L. ALEXANDRINO, RONEI J. POPPI

Institute of Chemistry, State University of Campinas – UNICAMP, Campinas, São Paulo 13084-971, Brazil

Received 2 April 2014; revised 20 May 2014; accepted 21 May 2014

Published online in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/jps.24051

ABSTRACT: Split-plot design (SPD) and near-infrared chemical imaging were used to study the homogeneity of the drug paracetamol loaded in films and prepared from mixtures of the biocompatible polymers hydroxypropyl methylcellulose, polyvinylpyrrolidone, and polyethyleneglycol. The study was split into two parts: a partial least-squares (PLS) model was developed for a pixel-to-pixel quantification of the drug loaded into films. Afterwards, a SPD was developed to study the influence of the polymeric composition of films and the two process conditions related to their preparation (percentage of the drug in the formulations and curing temperature) on the homogeneity of the drug dispersed in the polymeric matrix. Chemical images of each formulation of the SPD were obtained by pixel-to-pixel predictions of the drug using the PLS model of the first part, and macropixel analyses were performed for each image to obtain the *y*-responses (homogeneity parameter). The design was modeled using PLS regression, allowing only the most relevant factors to remain in the final model. The interpretation of the SPD was enhanced by utilizing the orthogonal PLS algorithm, where the *y*-orthogonal variations in the design were separated from the y-correlated variation. © 2014 Wiley Periodicals, Inc. and the American Pharmacists Association J Pharm Sci

Keywords: imaging spectrosocopy; factorial design; macropixel analysis; multivariate analysis; near-infrared spectroscopy; polymers; partial least squares.

INTRODUCTION

Chemical imaging (CI) coupled with vibrational spectroscopy [essentially Mid-infrared, near-infrared (NIR), and Raman] allows further analytical characterization of solid dosage forms of pharmaceuticals, and it has been encouraged by US FDA's process analytical technology to ensure quality control over pharmaceutical processes.[1,2] CI provides unique spectral information with spatial resolution (i.e., the sample region of interest is split in regular squares, denominated pixels, that possess their own spectra), and this analytical methodology has been exploited during numerous pharmaceutical process applications, such as the chemical distribution of the ingredients,[3] blend homogeneity,[4] coating,[5] polymorphism,[6] authentication,[7] and dissolution.[8]

In NIR-CI, because of the overlap of the spectral bands in the pixels, multivariate methods are usually required to extract chemical information from the samples. During the quantitative analyses, partial least-squares (PLS) regression has been used to obtain the concentration of the constituents in the pixels. PLS models are developed from CI data because the bulk concentrations of the analyte in the entire samples are known. Although the concentration of the analyte cannot be distinguished between the pixels in a sample, its respective bulk concentration can be correlated with the mean spectrum of that sample. After developing a PLS model, it can be applied in the spectra of the pixels, generating concentration maps.[6,9]

Transdermal drug delivery systems are currently being developed for existing drugs because of the advantages of these matrices, including their noninvasive application, the avoidance of first-pass hepatic metabolism for drugs with poor oral bioavailability, sustained delivery to achieve a steady plasma profile, and patient-friendly flexibility (i.e., reduction of dose schedules and easy withdrawal if there is any inconvenience).[10,11] Moreover, different associations and/or proportions among the polymers in films can be adjusted to achieve the desired drug release profile.[12]

The homogeneity of active pharmaceutical ingredient and excipients in pharmaceuticals is an important parameter to be monitored when developing pharmaceutical formulations to guarantee well-defined dissolution profiles and, consequently, satisfactory drug pharmacological properties in the organism.[13] In CI, for images generated from the quantification methods, the distribution of the constituents in a sample can be analyzed using the frequency histograms plotted from their respective concentration values in the pixels.[3,14,15] However, the resulting analysis will be subjective and must be followed by the respective images, whereas images with different levels of homogeneity can generate the same histogram.[16] In this context, Hamad et al.[17] developed the concept of macropixel analysis for chemical images to generate reasonable and quantitative analyses of image homogeneity. Macropixel analysis divides an image into equally sized blocks of pixels and calculates statistical parameters between the blocks. Further information regarding the different forms of macropixel analysis and the statistics between the macropixels used to quantify the homogeneity of the chemical images are described in the section *Macropixel Analysis of Chemical Images*.

In this work, a new strategy for a quantitative evaluation of the homogeneity of a drug loaded in a polymeric film is presented using NIR-CI and experimental design. The pharmaceutical matrix model used for simulating polymeric films loaded with a drug for transdermal applications was composed

Correspondence to: Ronei J. Poppi (Telephone: +55-19-35213126; Fax: +55-19-35213023; E-mail: ronei@iqm.unicamp.br)

Journal of Pharmaceutical Sciences

by blends of the biocompatible polymers hydroxypropilmetilcellulose (HPMC), polyvinylpyrrolidone (PVP), and polyethyleneglycol (PEG), loaded with paracetamol (PAR). In the first part of the study, a PLS model was developed for the pixel-to-pixel predictions of the drug in this type of film, so that the chemical images of the drug in future samples could be obtained. Next, a split-plot design (SPD) was elaborated to stress how the variability of different proportions of the polymers (mixture variables) and process conditions (process variables) can affect the homogeneity of the drug dispersed throughout films. The process variables chosen for this study were: (1) drug content in the formulations (DRUG) and (2) curing temperature (TEMP). The numerical values representing the drug homogeneity in the samples from the SPD (*y*-responses) were extracted for each sample through macropixel analysis of their respective chemical images obtained from the PLS model computed in the first part. The experimental design (containing main and interaction terms of process and mixture variables) was also modeled using PLS regression, and the importance of each term for the design matrix *X* for the *y*-explained variance was analyzed using the significance of the regression coefficients and by calculating the variable importance on projection (VIP) parameters. The robustness of the PLS regressions while excluding nonsignificant terms was evaluated using the evolution of the cross-validation error and visual inspection of the residues. When the design matrix contained only significant terms (*X*sig), a multivariate regression was also performed using orthogonal projections to latent structures (OPLS) algorithm. This method has an advantage over PLS; the *y*-orthogonal variations in *X*sig can be separated from the y-correlated variation, enriching the model interpretation.

THEORY

Macropixel Analysis of Chemical Images

Macropixel analysis requires chemical images to be characterized using a defined score value for each pixel, including images built for quantification (e.g., concentration), exploratory data analysis (e.g., PCA, unsupervised clustering methods), and others. Macropixel analysis can be conducted using nonoverlapped macropixels (Discrete-Level Tiling method), or with all possible macropixel sizes, through the method designated Continuous-Level Moving Block (CLMB). In the first method, the macropixels must be equally sized so that they can be fitted adjacently throughout the image; in the CLMB method, one macropixel systematically scrolls through the entire image row-to-row and column-to-column, and this procedure is performed repeatedly until the macropixel has the same dimensions as the entire image. In both methods, the variance between the subregions of the images is minimized while varying the sizes of the macropixels; however, during the CLMB, all possible macropixel sizes can be analyzed. To comprehend how the CLMB method works, the reader is referred to Ref. 17. The macropixel statistics are usually calculated as the score average value of the included pixels, and the standard deviation of the averaged macropixels in an image can be used as a numeric parameter for homogeneity. Hamad et al.[17] defined the specific concentration ranges of PAR in the macropixels using the chemical images of tablets and used a minimum macropixel size that included all of the macropixels in this range as a quantitative criterion of homogeneity. Wu et al.[18] analyzed the macropixels in binary images to study the homogeneity excipients in tablets via NIR-CI. Rosas and Blanco[19,20] employed a macropixel analysis for the chemical images of colored sand samples and pharmaceutical solid dosage forms by adjusting known mixing indexes from the macropixel statistics to study the homogeneity of the constituents in these samples.

Split-Plot Experimental Design with PLS Regression

Split-plot designs are formed by blocked experiments that randomly vary easy-to-change factors within the blocks, whereas the hard-to-change factors are randomly varied only when moving from one block to another, resulting in two experimental units.[21,22] The hard-to-change factors that define the blocks are the main plots, whereas the subsets of those are the subplots. The split-plot approach can be exploited in mixture designs submitted to different process conditions; the mixture and process variables are treated as subplots and main plots, respectively.[23] For a design with two main plots and three subplots, the influence of these variables on an experimental response can be analyzed via multivariate regression according to a general mixed model equation presented in Eq. (1).

$$Y_{\mathrm{ijklm}} = \mu + z_i + z_j + z_{ij} + x_k + x_l + x_m + W + e_{ijklm} \qquad (1)$$

where Y_{ijklm} is the model response; z_i and z_j are the main plot effects (process variables); z_{ij} is their interaction term; x_k, x_l, and x_m are the subplot effects (mixture variables); W comprises all of the main subplot and subplot–subplot interaction terms; μ is the model intercept; and e_{ijklm} is the regression error. Blocking the experiments results in two error sources, the main plot error, which is the same for experiments within the blocks, and the subplot error, which is uncorrelated within and among the blocks.[24] Equation (1) can also be rewritten in matrix notation, as shown in Eq. (2):

$$y = bX + \delta + \varepsilon \qquad (2)$$

where X is the design matrix (composed by the main and interaction fixed effects), b is the regression coefficients vector, δ and ε are the main plot and subplot error vectors, respectively, and y is the responses vector. For balanced designs, b is usually calculated from Eq. (2) through the generalized least-squares (GLS) regression: $b_{\mathrm{GLS}} = (X^{\mathrm{T}}V^{-1}X)^{-1}X^{\mathrm{T}}V^{-1}y$, where V is a diagonal matrix that accounts for the main plot and subplot error variances, ${\sigma^2}_{\mathrm{MP}}$ and ${\sigma^2}_{\mathrm{SP}}$, respectively. An appropriate ANOVA can be used to calculate the variances of the errors in fully replicated SPD,[25] but this method becomes impracticable if numerous experiments are necessary. Subsequently, nonfully replicated SPD have also been proposed,[26] and the error variances of the main plot and subplot can be obtained using additional replicates only in the central point.

Generalized least-squares is a suitable method for regression in which two independent error variances must be accounted during the uncertainty estimation of the regression coefficients; however, the covariance matrix must not be ill-conditioned, generating good estimations of the regression coefficients. In mixture designs, the constraint $\sum^{x}(i) = 1$ results in nonindependent and highly collinear variables that may lead to ill-conditioned covariance matrices. Afterward, projection methods, such as PLS regression, become more suitable for modeling mixture designs after they can handle