International Journal of Pharmaceutics 536 (2018) 251–260

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

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

Multivariate modeling of diffuse reflectance infrared fourier transform (DRIFT) spectra of mixtures with low-content polymorphic impurities with analysis of outliers

Slobodan Šašić[a,*], Andrew Gilkison[a], Mark Henson[b]

[a] SSCI/AMRI, 3065 Kent Ave. West Lafayette, IN, 47906, United States
[b] Shire, Analytical Technology, 730 Stockton Dr. Exton, PA, United States

ARTICLE INFO

Keywords:
DRIFTS
Infrared spectroscopy
Chemometrics
PLS regression
Poly
morphs
Outliers

ABSTRACT

Diffuse Reflectance Fourier Transform Infrared Spectroscopy (DRIFTS)-based multivariate models were developed to quantify the content of two polymorphic impurities in mixtures with the desired active pharmaceutical ingredient (API) form, with the impurities not exceeding 2% wt/wt. In addition, close attention was paid to the outlier detection criteria: Q residuals; Hotelling T^2; and score bi-plot. While reasonably accurate results were obtained for the relatively simple calibration models for both forms of the impurity, the predictions for "blank" samples (separately verified to be impurity-free) were apparently biased. Thus, the model training sets were augmented with spectra from calibration mixtures incorporating some of the API from batches used in the prediction. The performance of the updated models as assessed by cross-validation was somewhat degraded as a result, while predictions against independent batches of API showed a decrease in bias indicating robustness had improved. Nevertheless, the Q residuals criterion disqualified a large number of prediction samples as outliers in contrast to the other two criteria that reported no issues at all. The results here demonstrated the effectiveness of DRIFTS for quantifying low concentration polymorphic impurities, while simultaneously highlighting the variability issues that can be encountered in practice and which need to be understood and managed appropriately to ensure the success of any automated or Good Manufacturing Practice (GMP) compliant application of multivariate modeling.

1. Introduction

Vibrational spectroscopy has long been in use for analyzing polymorphism of active pharmaceutical ingredients (APIs). Raman spectroscopy has been particularly useful due to the typically strong Raman response of APIs and good differentiation of Raman bands. Near-infrared (NIR) spectroscopy on the other hand has often been coupled with multivariate calibration for quantitative analyses of drug product and drug substance (API) mixtures in pharmaceutical industry and it is the most frequently used spectroscopic technology. Diffuse Reflectance Fourier Transform Infrared Spectroscopy (DRIFTS) has gained much attention for characterizing polymorphs qualitatively for legal purposes (patents) but it is not frequently used for quantifying polymorphic impurities. This is despite DRIFTS spectral resolution and information content far exceeding that of NIR as well as DRIFTS being less prone to intensity variations seen in Raman spectra. This is likely due to sample preparation requirements for DRIFTS analysis (mixing / diluting the sample with KBr) which are mostly unnecessary for NIR or Raman analysis and which may carry a risk of solid state transformation for unstable forms. Nevertheless, DRIFTS offers several key advantages for implementation as a routine quality control (QC) test. Among them, FTIR instrumentation is a standard and familiar element in the vast majority of pharmaceutical analysis laboratories around the world, both at innovator R&D and manufacturing sites as well as contractor facilities. This is due to the fact that, thanks to USP < 197 > and EP, the vast majority of material identifications performed as part of pharmaceutical QC testing is done via FTIR. Based on this combination of scientific and practical/business considerations, the study herafter described was conducted to develop a multivariate DRIFTS-based method for quantification of low levels of two polymorphic impurities in mostly binary powder API samples.

The mixtures addressed here are intended for calibration of an impurity test, thus the concentrations of polymorphic impurities do not exceed 2% wt/wt. Since only API mixtures are analyzed, the majority of the mixtures in this study are binary. These two impurities are treated as unlikely to be both present at the same time so there are only a few

* Corresponding author.
E-mail address: slsasic@yahoo.com (S. Šašić).

https://doi.org/10.1016/j.ijpharm.2017.11.058
Received 18 September 2017; Received in revised form 3 November 2017; Accepted 26 November 2017
0378-5173/ © 2017 Elsevier B.V. All rights reserved.

ternary mixtures in the models below.

The literature lists several studies that employ DRIFTS for similar purposes (Calvo et al., 2016; Bunaciu et al., 2001; Guo et al., 2017; Atici and Karliga, 2015; Nemet et al., 2009; Kachrimanis et al., 2010; Kipouros et al., 2006; Siddiqui et al., 2013) but the concentration variation of the polymorphs in some of those studies is quite broad and it may range from 0–100% wt/wt, i.e. from one pure polymorph to another pure polymorph. It is not easy to see where exactly the outcomes of those studies can be practically applied, except for analyzing products of unrefined API synthesis, or during very preliminary polymorph screening. On the contrary, we target a commercial API substance of high purity in which only minor amounts of polymorphic impurity may be present. As the specification to be developed from the final method was expected to fall well below 1% wt/wt, the upper impurity concentration limit for the calibration samples was set at 2%.

Due to the relatively low signal-to-noise levels realized via DRIFTS analysis of solid state impurities at these concentration levels, multivariate modeling is a necessity to effectively leverage the spectral data to the maximum extent. Thus, partial-least-square (PLS) regression was employed for building separate multivariate models for each of the two impurities. However, the aim of the project was not only to produce quantitative models but also to ensure they were suitable for incorporation into a routine QC method run in a lab controlled per Good Manufacturing Practice (GMP) regulations. Thus, at each stage in model development and testing, consideration was paid to outlier detection as a means for assessing applicability of the model to new data/samples. While in this example we do not necessarily describe the final, most robust version of the models capable of accurately predicting future samples, procedures for initial assessment of the boundaries of applicability are described which involve multiple commonly-used criteria for outlier detection: Q residuals, Hotelling T^2, and score bi-plots (Brereton, 2007; Naes, 1989; De Maesschalck et al., 2000; Anon, 2017). All of these are conveniently produced by the software used for multivariate calibration and thus easy to discuss. The outlier detection criteria are crucial for any automated / GMP application of multivariate models as they mathematically (i.e. objectively) indicate the agreement between the prediction spectra and the model space. In long-term use of multivariate prediction in an industrial environment issues may occur with samples drifting or being out of the model space which may lead to erroneous predictions (e.g., out-of-trend [OOT] and/or out-of-specification [OOS] results) which in turn may result in significant business / financial / quality / regulatory issues due to investigations and delay in release or failure of batches. In theory, samples not fitting the model space should be disqualified thereby preventing any multivariate prediction-caused false alarms. Yet, these criteria and their effectiveness are seldom addressed in multivariate studies. Several recent studies mention the three criteria above in various context none of which is in line with the intent described here, or with comparable samples (Bu et al., 2013; Kuligowski et al., 2012; Pöllänen et al., 2005; Nagy et al., 2017).

2. Experimental

2.1. Sample preparation

For legal/proprietary reasons we are unable to disclose the chemical identity of the substances in this study. The two polymorphs of interest are thus simply addressed throughout as ‘Form I’ and ‘Form II’. Being purely analytical, the results of the study are by no means affected by using these notations. Pure reference samples of both Form I and Form II of the impurity species were synthesized at laboratory scale. The phase purity of each sample was verified by Powder X-Ray Diffraction (PXRD). Similarly, a bulk amount of API for use in the study was obtained by PXRD analysis of a commercial lot of API to verify that neither Form I nor Form II of the impurity were present at detectable levels.

The first set of calibration mixtures consisted of the total of six samples (mixtures) for each impurity polymorph, of which five were binary and one was ternary. The concentrations of both polymorphs varied from 0 to 1.8% wt/wt. The mixtures were obtained by using mortar and pestle with moderate pressure applied. Between 1 and 2 g of mixtures were prepared for each concentration of the polymorph. For the prediction study, various commercial lots of the desired API were used.

2.2. Spectroscopy

IR spectra were acquired using a Nicolet 6700 FTIR spectrophotometer (Thermo Nicolet) equipped with an Ever-Glo mid/far IR source, an extended range KBr beamsplitter, and a DTGS detector. A diffuse reflectance accessory (Collector™ II, Spectra-Tech), equipped with a 13 mm sampling cup, was used for analysis. Calibration mixtures were diluted to approximately 2.5% sample concentration, by weight, in spectroscopic grade potassium bromide (KBr) that had been ground and dried prior to the experiment. Three samples of each calibration mixture were analyzed. The triplicate sampling / dilution / scanning approach was intended partially to address the risk of inhomogeneity in the calibration mixtures, but also to gain information regarding the repeatability of the procedure. DRIFT spectra were acquired using 128 co-added scans and $4\,cm^{-1}$ resolution. Background spectra were acquired using KBr, and were typically obtained at the start of each day only. Total time from the start of mixing the samples and KBr diluent to completion of spectral data acquisition was approximately six minutes for all samples. A couple of months passed between the acquisition of calibration and prediction DRIFT spectra.

2.3. Chemometrics

Solo v.7.5.2 (Eigenvector Inc., Wenatchee WA) was used for all calculations. Based on preliminary optimization studies, the DRIFT spectra of Form I were pre-treated with standard normal variate (SNV) normalization followed by Savitzky-Golay (SG) 1st derivatives (2nd order) with 21 points. Form II was also treated with SNV first followed by SG 2nd derivatives (2nd order) with 21 points. The spectra were truncated to the $1275–1575\,cm^{-1}$ region observed to encompass the key bands of interest of the desired API form and the two impurity polymorphs. All calibration models were leave-one-out (LOO) cross-validated. However, it is noteworthy that venetian blinds cross validation with four to five spectra out of 18 left out in each validation turn produced only slightly worse results than LOO cross validation, indicative of robust results not unique to the specific choice of cross-validation pattern. In addition to predictive performance metrics, each PLS model was also scrutinized using a series of diagnostic/data distribution metrics: Q residuals, Hotelling T^2, and score bi-plots. Q residuals are a sum of the squares of each row of the regression error matrix and as such are indicators of the spectral fit of the regression (or the lack of fit). The smaller the Q residuals, the better the spectral fit of the model. Being sums of squares, Q residuals are normally reported as numbers but one can also plot individual Q residual plots per sample and learn on the distribution of the spectral errors. Hotelling T^2 are sums of the squared (normalized) scores and as such indicate the distance between a sample’s projection on the model space and the center of the model. Large Hotelling T^2 values are indicative of samples with greater potential leverage upon the model. It is often instructive to plot Hotelling T^2 and Q residuals together, as these two criteria are based on very different calculations. Score bi-plot is a well-known and often used tool for outlier detection. In this study we only provide PC1 vs. PC2 score bi-plots since these two scores account for substantial majority of variation (normally $\geq 95\%$) in all calibrations and predictions. The contribution of lesser PC scores is somewhat reflected in Hotelling T^2 values. Besides, it would be impractical and of perhaps questionable justification to produce numerous other bi-plots that might indicate

Download English Version:

https://daneshyari.com/en/article/8520471

Download Persian Version:

https://daneshyari.com/article/8520471

Daneshyari.com