Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com

Review

Characterization of pharmaceutically relevant materials at the solid state employing chemometrics methods

Natalia L. Calvo, Rubén M. Maggio, Teodoro S. Kaufman*

Instituto de Química Rosario (IQUIR, CONICET–UNR) and Área Análisis de Medicamentos, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, Rosario (S2002LRK), Argentina

ARTICLE INFO

Article history:
Received 15 May 2017
Received in revised form 8 June 2017
Accepted 12 June 2017
Available online xxx

Keywords:
Chemometrics
Characterization of the solid-state
Vibrational spectroscopies and XRPD
Chemical imaging
PAT
Quality control and stability

ABSTRACT

The understanding of materials and processes is a requirement when it comes to build quality into pharmaceutical products. This can be achieved through the development of rapid, efficient and versatile analytical methods able to perform qualification or quantification tasks along the manufacturing and control process. Process monitoring, capable of providing reliable real-time insights into the processes performance during the manufacturing of solid dosage forms, are the key to improve such understanding.

In response to these demands, in recent times multivariate chemometrics algorithms have been increasingly associated to different analytical techniques, mainly vibrational spectroscopies (Raman, MIR, NIR), but also UV–vis spectroscopy, X-ray powder diffraction and other methodologies. The resulting associations have been applied to the characterization and evaluation of different aspects of pharmaceutical materials at the solid state. This review examines the different scenarios where these methodological marriages have been successful.

The list of analytical problems and regulatory demands solved by chemometrics analysis of solid-state multivariate data covers the whole manufacturing and control processes of both, active pharmaceutical ingredients in bulk and in their drug products.

Hence, these combinations have found use in monitoring the crystallization processes of drugs and supramolecular drug associations (co-crystals, co-amorphous and salts), to access the correct crystal morphology, particle size, solubility and dissolution properties. In addition, they have been applied to identify and quantitate specific compounds, mainly active pharmaceutical ingredients in complex solid state mix-

Abbreviations: AH, anhydrate; ALS, alternating least squares; AM, amorphous; ANN, artificial neural network; API, active pharmaceutical ingredient; ATR-MIR, attenuated total reflection-mid infrared; CA, cluster analysis; CART, classification and regression trees; CI, chemical imaging; CLS, classical least squares; DH, dihydrate; DOE, design of experiments; DSC, differential scanning calorimetry; MIR, Fourier transformed mid infrared; HCA, hierarchical cluster analysis; HME, hot melt extrusion; HSI-NIR, near infrared hyperspectral images; ICA, independent component analysis; k-NN, k-nearest neighbours; LDA, linear discriminant analysis; LOD, limit of detection; MALDI-MSI, matrix-assisted laser desorption ionization mass spectral imaging; MCR, multivariate curve resolution; MH, monohydrate; MIR, mid-infrared spectroscopy; MLR, multiple linear regression; MSC, multiplicative scatter correction; NIR, near infrared spectroscopy; OPLS, orthogonal-PLS; PARAFAC, parallel factor analysis; PAT, process analytical technology; PC, principal component; PCA, principal component analysis; PCR, principal component regression; PLS, projection to latent structures regression; PLS-DA, PLS-discriminant analysis; RMSEE, root mean square error of estimation; RMSEP, root mean square error of prediction; SIMCA, soft independence modeling of class analogies; ssNMR, solid state nuclear magnetic resonance; SNV, standard normal variate; XRPD, X-ray powder diffraction

* Corresponding author.
Email address: kaufman@iquir-conicet.gov.ar (T.S. Kaufman)

http://dx.doi.org/10.1016/j.jpba.2017.06.017
0731-7085/

tures. This included drug stability against different stimuli, solid-state transformations, or detection of adulterated or fraudulent medicines.

The use of chemometrics-assisted analytical methods as part of the modern concept of process analytical technology, where every process step of every product batch from raw materials to final product must take place in a controlled manner is discussed. Finally, but no less important, the application of chemometrics methods to chemical imaging, aiming to extract spatial and compositional information is also revised.

1. Introduction

The pharmaceutical industry is perhaps the scenario where the analysis of complex mixtures founds its most critical applications. It is the place where it also becomes significantly more important. This is because the correct composition of each pharmaceutical product is essential to its quality, efficacy and safety; minor compositional variations in these complicated mixtures may have profound effects on their characteristic properties and performance [1], potentially affecting the health condition of human beings.

Due to the current stringent regulations and the increasing consumer and market awareness, pharmaceutical manufacturers are placing more and more emphasis on product quality and great efforts in its assurance. Hence, compositional profiling plays a pivotal role in guaranteeing the chemical consistency of pharmaceutically-relevant products among batches. The same holds true for other industries, including foods, flavors, agrochemicals, etc. [2]. However, the special characteristics of the medicines and their relation to human health turn knowledge of the identity and content of their active pharmaceutical ingredients, paramount to their quality assurance and safety. For the same reasons, knowledge of the formulation impurities and excipients becomes no less important.

In solution, most pharmaceutical systems can be characterized through a combination of spectroscopic (ultraviolet, mid-infrared, nuclear magnetic resonance and others), chromatographic (gas chromatography, high performance liquid chromatography, etc.) and other methods, partly because the systems are homogeneous. On the contrary, the characterization of pharmaceutical solids is somehow different, resulting in a different set of suitable methodologies [3]. Because the material is heterogeneous, consisting of particles of varying sizes and compositions, quantitative methods exhibit higher data dispersion, which has impact on method robustness and precision, among other variables [4].

Many critical measurable properties of interest, such as crystal form, structural polymorphism and drug–excipient interactions, are unique to the solid state, and disappear when the material is dissolved. In addition, the amount and nature of the information provided by different analytical techniques may differ, whether it is generated from a solid sample or a solution. This turns more necessary the use of a wider array of techniques, to better characterize a material in the solid state.

The analytical methodologies used for solid state characterization can be broadly grouped in two classes: bulk and molecular. Bulk techniques, including hot stage microscopy, thermogravimetry, differential scanning calorimetry and others, rely upon global properties (thermodynamics, particle morphology) of the system, providing information about its global state. On the other hand, molecular techniques, such as spectroscopic (Raman, near infrared, mid infrared, nuclear magnetic resonance) and X-ray diffraction methods, need to probe the molecular-level interactions of the system in order to provide information [5–10].

Steady advances in statistics, analytical methodologies, new technologies, and increased computational power at hand of the analysts have led to an enormous increase in the volume of data generated even by the most traditional analytical chemistry procedures. Recording spectra from convoluted mixtures entails the same degree of difficulty posed by their acquisition from pure compounds; however, the resulting datasets have much greater complexity, which usually hinders the extraction of useful information.

The presentation, analysis and interpretation of this huge amount of otherwise obscured or hidden chemical data needs to be properly managed, in order to yield quality information. The modern discipline of chemometrics is defined as the application of statistical and mathematical methods to analytical data to enable maximum collection and extraction of useful information. In that capacity, chemometrics can clearly potentiate the capabilities of standard analytical instrumentation and techniques through its ability to convert seemingly unrelated and incomprehensible data (involving many hundreds of variables), into valuable information on the state of a chemical system.

This is because multivariate methods are capable of performing reduction of data dimensionality. For example, Principal Components Analysis (PCA) and Projection to Latent Structures (PLS) are based on the assumption that only a few variables can account for most of the variation in the data (i.e., information). During the last two decades, this has caused a deep impact in many areas and has resulted in a marked increase in the breadth of application of even the most traditional analytical techniques.

Being one of the masterpieces of today's analytical chemistry, chemometrics is the best way to maximize the quality of the information accessible from the samples. It has also been said that it also contributes to the development of greener analytical methods [11].

The application of chemometrics methodologies to the study of chemical systems of pharmaceutical interest at the solid state still runs behind the use of chemometrics as a tool for solving chemical problems of liquid samples. However, the fast pace of the incorporation of some spectroscopic methods, especially vibrational, X-ray and solid state nuclear magnetic resonance to the study of pharmaceutical problems with the aid of chemometrics tools is rapidly bridging the gap.

In the case of solids, sample preparation and the experimental setup should be carefully planned to obtain the most reliable and consistent data. Sampling strategies and the experimental design need to be taken into account in order to reduce the sample preparation burden, and both must be related to the analytical problem aimed to be solved. Design of experiments strategies, once used only occasionally, is being employed with more frequency to achieve better results at the lowest cost.

Chemometrics algorithms have been shown to be useful aids along the whole manufacturing and control process of solid pharmaceutical formulations, from the active ingredients up to the controls of the final product [12]. Accordingly, they have been used to monitor and control drug crystallization, to obtain the most suitable crystals in terms of structural polymorph, morphology and size, as well as to monitor or predict different properties of bulk pharmaceutical solids, including particle size distribution, powder flowability and angle of repose, from their spectral data. They proved also suitable to evaluate other useful properties, such as porosity and hardness of manufactured tablets.

The chemometrics analysis of spectral data has demonstrated its aptness to detect and quantitate active pharmaceutical ingredients in their dosage forms and pinpoint counterfeit drugs and medicines. They proved also suitable to discover minor amounts of impurities and degradation products in the presence of major levels of their precursors in quality control analyses and during stability studies. In addition, the monitoring of the pharmaceutical dissolution of solids as well as thermal transformations and molecular interactions at the solid state hava also been carried out by chemometric methodologies.

Another important area of pharmaceutical solids analysis, where chemometrics plays a center stage role, is chemical imaging. This branch of the analytical science, resulting from the integration of microscopic and spectroscopic techniques, is able to generate huge amounts of complex, hard to analyze, multidimensional data. Thanks to the recent advances in chemometrics and computation, nowadays space-related chemical composition, drug distribution and performance of these systems can be examined and predicted to such a detailed extent considered impossible short time ago.

No less important is the integration of chemometrics with some vibrational spectroscopies, as part of the modern monitoring and decision-making philosophy of Process Analytical Technologies (PAT). In this framework, chemometrics enables to take decisions in real time, based on the management and analysis of the vast amount of data generated by the analytical instrumentation placed on-line, in-line or at-line. Thus, by providing a more comprehensive understanding, chemometrics makes possible a far better control of the manufacturing process.

Finally, an additional major subject relating solid state compounds with chemometrics is structural polymorphism, also termed crystal polymorphism. The impact of polymorphism in the pharmaceutical industry is of such an important magnitude and interest that its examination has been set apart, considering that it deserves a separate review. However, it is worth mentioning that chemometrics analyses have been employed as an aid to assign polymorphic identity to unknowns, to establish the content of given forms in complex mixtures, and to study polymorphic transformations including polymorph conversion, solvation and desolvation, as well as amorphization and crystallization processes.

2. The chemometrics tools. Brief theoretical background of the most used methods

Chemometrics is a young, rapidly evolving and still maturing discipline, which has caught great attention among pharmaceutical scientists, especially during the last 20 years. Although only part of the currently available chemometrics toolbox is actually used with a certain frequency, in most cases there is one or more alternative chemometrics resource that could be successfully employed to solve the same problem, and chemometrics methods can be used at different stages during the drug manufacturing and control process. It is up to the analyst to experiment and select for the most suitable option.

The basis and in-depth explanations of the different chemometrics methods are discussed elsewhere [13]; however, for the sake of self-consistency, a brief introduction to some of the most frequently used methods in the analysis of the solid state is carried out bellow. For a better understanding, the different methodologies are grouped according to their scope; these include the Design of Experiments (DOE), a very useful approach for defining the best operational conditions, data pre-treatment, which involves data conditioning for proper analysis, and multivariate data analysis, which comprises both, qualitative and quantitative methodologies.

2.1. Design of experiments

DOE is a rational approach to understand how process parameters affect the response variables, analytical signal or product properties. DOE has been used in several publications as part of the screening stage, to discover which factors may significantly affect the response of an experiment. Plackett-Burman and fractional factorial designs are the most suitable approaches to find the factors which significantly contribute to the response [14], while performing the minimum number of experiments on a maximum number of factors.

Another task of DOE is to find the values of the relevant factors that optimize a given response [15]. Statistical designs, such as factorial (fractional factorial), central composite, Box-Behnken and mixture models, are among the best choices to define the objectives and planning the experiments.

In order to visualize the variations, a response surface can be calculated for each variable to be optimized employing multiple linear regression or artificial neural networks. Finally, Derringer's desirability function [16] can be employed to calculate the optimum values if many responses are to be optimized simultaneously. Table 1 details the main objectives and some of most relevant tools of the current chemometrics arsenal.

2.2. Data pre-treatment

Data pre-treatment has substantial impact on the quality of the final models and the results of the quali- or quantitative analysis. The objective of this practice is to remove the influence of physical phenomena not related to the analytes, suppressing systematic variations, random signals and uninformative variables. This approach ensures that subsequent modeling makes focus on relevant variations in the process data, improving the quality of the results. Hence, pre-treatment or pre-processing of multivariate data is currently an integral part of chemometrics modeling.

The most widely used pre-processing techniques can be broadly divided in classical pre-processing methods, such as mean-centering and (auto)scaling, and signal correction methods, which can be further sub-divided into two groups: spectral derivatives, which includes the well-known Savitzky-Golay polynomial derivative filter, and scatter-correction methods [i.e., multiplicative scatter correction (MSC), detrending, standard normal variate (SNV) and normalization] [17].

A third group is that of dimensionality reduction methods, which aims at eliminating uninformative signals, in order to enhance the information and reduce collinearity. This includes variable selection, orthogonal signal correction and data compression methods [18]. These can also be used for quali- or quantitative purposes.

Although there are no written guidelines on what set of pre-processing algorithms should be applied in each case, a few approaches to solve the problem have been discussed [19,20]. However, most often analysts still have to rely on literature precedents, their own experience and trial and error approaches.

2.3. Qualitative methods

The major aims of chemometrics qualitative analysis are to discover groups of objects which have similar characteristics, to visualize and characterize hidden tendencies among groups, and finally to assign new samples to a known group. Supervised and unsupervised methods are two of the main categories of this field.

Table 1
Selected chemometrics resources for studying the solid state and their objectives.

Chemometrics stage	Objectives	Tools
Design of experiments	Screening of relevant factors, quantifying their effects Modelling and optimizing the processes	Factorial (full, fractional) Central Composite (CCD) Mixture model Box-Behnken Plackett-Burman Doehlert Kennard-Stone sampling
Data pre-treatment	Reduce random noise and undesired or unwanted perturbations in the signal Diminish uninformative data variation and remove background Enhance the detectability of minor features	Mean centering (MC) Baseline correction Multiplicative scatter correction (MSC) Scaling or Normalization Standard normal variate (SNV) Data alignment Detrending (DT) Savitzky-Golay smoothing and derivation (D^0, D', D") Variable selection
Classification and display methods (qualitative methods)	Discover groups of objects which have similar characteristics Discover hidden tendencies among groups Assignment of a new sample to a known group	Principal component analysis (PCA) Linear discriminant analysis (LDA) Gaussian mixture models (GMM) Hierarchical cluster analysis (HCA) Self-organizing map (SOM) K-means clustering Factor analysis PLS-DA SIMCA
Calibration (quantitative methods)	Translate data into quantitative information, mainly for predictive purposes. Perform quantitative analysis by relating two sets of variables (dependent and independent)	Multiple linear regression (MLR) Principal components regression (PCR) Projection to latent structures (PLS) Artificial neural networks (ANN) MCR-ALS PARAFAC

2.3.1. Unsupervised methods

Unsupervised methods are algorithms which require no previous knowledge about the groups present in the population. Unsupervised classification is often called clustering. In this type of classification, it is assumed that the "objects" or samples have a set of features with hidden relationships with a given class or category. Then, the purpose of the method is the discovery of groups of objects which have related characteristics; this enables their separation into different classes.

Among unsupervised methods, PCA is the most widespread used in the social and natural sciences [21]. PCA is a technique used to reduce the dimensionality of a data set, being primarily used as a display methodology in exploratory analysis. PCA involves calculating the eigenvalue decomposition of the covariance data matrix, usually after mean-centering the data. Operatively, it can be regarded as a process for selection of the best coordinate system to project the data. Fig. 1 outlines the principle of PCA. A closely related method is factor analysis (FA), where the original variables are also defined as linear combinations of the factors. However, the main objective of FA is to explain the covariances or correlations among the variables. Therefore, unlike PCA, FA can be used to understand the construct underlying the data [22].

A dendrogram is a kind of tree-shaped data chart, obtained by hierarchical clustering analysis (HCA). HCA organizes data into different sub-categories, which are then sub-divided to reach the desired level of detail [23]. Dendrograms allow the examination of grouping relations between the data, and the successive subdivisions provide an idea about the grouping criteria.

Gaussian mixture models (GMM) and K-means clustering [24], are also related methods. In the K-means clustering approach, the user needs to specify the number (K) of groups to be differentiated. All objects need to be represented as a set of numerical features (n), being all features used to describe the objects to cluster. The initial centers of the clusters are chosen at random, and each object is assigned to the closest center. Then, a new center is computed for each cluster by averaging the feature vectors of all objects assigned to it. The assignment of objects and re-computing centers is repeated until the process converges. The features requiring optimization include distance measurement, the choice of the initial center, the computation of new average centers, and the estimation of the number (K) of clusters.

Like K-means clustering, GMM uses an iterative algorithm that converges to a local optimum, being more appropriate than K-means when clusters have different sizes and correlate among them. The probabilities for each point indicate whether each data point belongs or not to a given cluster.

Kohonen neural networks or self-organizing maps (SOM) provides a way of representing multidimensional data in a two-dimensional map. A SOM consists of components called nodes or neurons, which are associated with a weight vector of the same dimension as the input data vectors and to a position in the map space. The procedure for placing a vector from the data space onto the map is to find the node with less distance. Next, the weight of the vector of this node is corrected toward object position. The procedure is iteratively repeated during a user pre-established number of times (ages) [25].

2.3.2. Supervised methods (classification)

In the supervised methods, the classes responsible of data variation are known and the information which defines each class is available to extract model information or assign a new sample to a class; these methods are also known as classification algorithms.

Soft independent modelling of class analogy (SIMCA) [26] is a supervised method which considers each class separately and performs a PCA on each one to build a class model. Any new object is challenged to every class model and it is assigned to the class that produces the smallest residue during the prediction. In this sense, SIMCA places more emphasis on similarity within a class than on discrimination among classes.

Linear discriminant analysis (LDA) is one of the most widely used supervised classification methods. Like PCA, LDA is a linear feature reduction methodology, which focuses on finding optimal boundaries (parametric) between object classes. The LDA algorithm

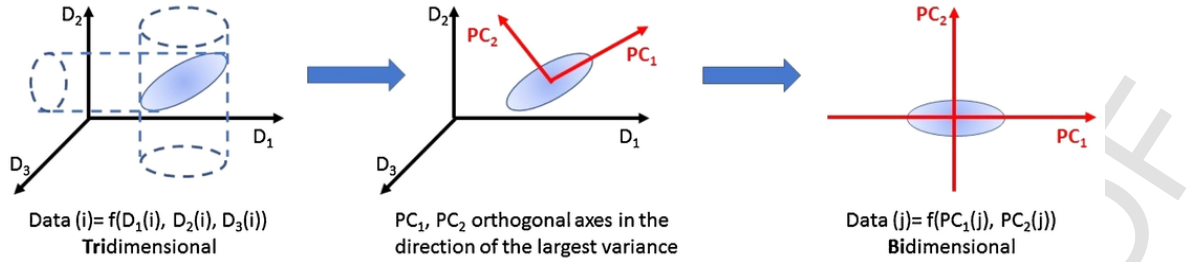

Fig. 1. Operational principle of PCA. Graphical example of data dimensionality reduction through their projection on orthogonal axes in the direction of the largest variance.

selects the space directions that achieve a maximum separation among the different classes and uses euclidean distance measurements to classify unknown samples [27,28].

Canonical discriminant analysis (CDA) is a kind of exploratory version of LDA. This is a dimensional-reduction technique, linear and parametric which, unlike LDA, aims to understand which variables are responsible for the differences among groups [29].

Another method is PLS-discriminant analysis (PLS-DA). This is a linear and parametric method which carries out a PLS analysis (see 2.4 Quantitative methods) of a Y-set of binary variables (0 or 1), which describe the categories, on a set X of predictor variables [30]. This technique is especially well-suited to deal with a small number of observations and with multi-colinearity. Orthogonal-PLS and N-way PLS can also be coupled to discriminant analysis (OPLS-DA [31] and NPLS-DA, respectively). These represent variations of PLS-DA, which improve the performance of the latter.

The k-nearest neighbours (k-NN) algorithm is a non-parametric method used for classification, where the input consists of the *k* closest training examples in the feature space. An object is classified by a majority vote of its neighbours, with the object being assigned to the most common class among its k nearest neighbours [32].

On the other hand, several types of neural networks can be used for classification. Back propagation artificial neural networks (ANN) and support vector machines (SVM) are the most frequently used for discriminant analysis [33]. As in the case of PLS-DA, the ANN-DA and the SVM-DA algorithms perform a correlation with a binary variable Y (which contains the classes). These methods are very useful when data are non-linearly distributed, even after variable transformation.

2.4. Quantitative methods

Quantitative chemometrics methods can be classified in two general groups, including first order methods (FOMs) and higher order methods (HOMs). Whereas FOMs use a data vector for each sample (such as those resulting from UV, NIR, MIR, Raman, fluorescence spectra and others) as input signal, HOMs use a data matrix. This data matrix can be a two-dimensional array (second order matrix, i.e. fluorescence matrix, or temperature-evolving infrared analysis) or a higher order one array (three-dimensional matrix, infrared image) [34].

Both FOMs and HOMs have several advantages with regard to univariate calibration [35]. FOMs can mitigate the presence of known or unknown interferences, as long as these interferences have been taken into account in the calibration procedures; this is termed the "first order advantage". This advantage can be achieved by using calibration mixtures prepared according to a calibration design, where the concentration of the analyte is already known and the interferents are present. However, when real samples are used for calibration, their analyte concentrations must be determined *a priori* by a standard method.

Within the FOMs, the standard in chemometrics is PLS [36]. This is a "full-spectrum" latent variable-based method as its ancestor, principal component regression (PCR) [37]. However, unlike PCR where data compression is made using only the instrumental response (**X** matrix), in PLS it is performed employing both, signal (**X** matrix) and analyte abundance data (**Y** matrix, reduced to a *y* vector in the PLS1 version). Therefore, the PLS model can be considered as consisting in outer relations (**X** and **Y**) and an inner relation. Both blocks can be decomposed into the corresponding products of smaller matrices of scores and loadings, and error matrices **E** and **F** ($\mathbf{X} = \mathbf{TP}^T + \mathbf{E}$ and $\mathbf{Y} = \mathbf{US}^T + \mathbf{F}$). The non-linear iterative partial least squares (NIPALS) algorithm around which the PLS is built, iteratively attempts to explain the maximum covariance between **X** (instrumental data) and the *y* vector (concentration data) by minimizing **E** and **F**, until convergence is reached The regression coefficients (*b*), calculated with the aid of the relevant scores of matrices **T** and **U** (inner relation), can then be used to predict the analyte abundance from new sample spectra. Fig. 2 depicts the operation of the PLS algorithm.

The PLS method is useful because of its ease of use and ability to handle band overlaps, colinearity, and interactions within the samples. Its success and wide applicability explain the availability of the algorithm as part of several free and commercial statistics, chemical and instrumental software packages. The most critical parameters to optimize are the number of factors (latent variables) and the spectral region to be used [38].

Back propagation artificial neural networks (ANN) [39] is a natural computing first-order method. It is just one of the possible ways to operate a non-linear mapping between an input and a target space; it operates using a large number of parallelly connected simple arithmetic units, the neurons. Mathematically, a neuron is a non-linear, parameterized and bounded function. The "neuron inputs" are the variables on which this function depends, while the calculated value of that function is called "neuron output".

ANN is very flexible; it can be custom-designed to become adaptable to different kinds of data structures, being particularly suited to handle highly non-linear problems, where the traditional statistical methods usually fail. However, ANN is prone to overfit the system when the complexity of the underlying relationship between input and output variables is poor. In addition, they require a large number of adjustable parameters, making necessary a trained operator to successfully handle the algorithm and obtain meaningful results [39].

HOMs are also known as multi-way techniques because data for a single sample are contained in a multidimensional array (a matrix, or a higher-order array). HOMs hold the ability to perform determinations in the presence of interferences, not taken into account in the calibration step. This property is known as "the second-order advantage" and turns HOMs particularly attractive for determinations in complex samples.

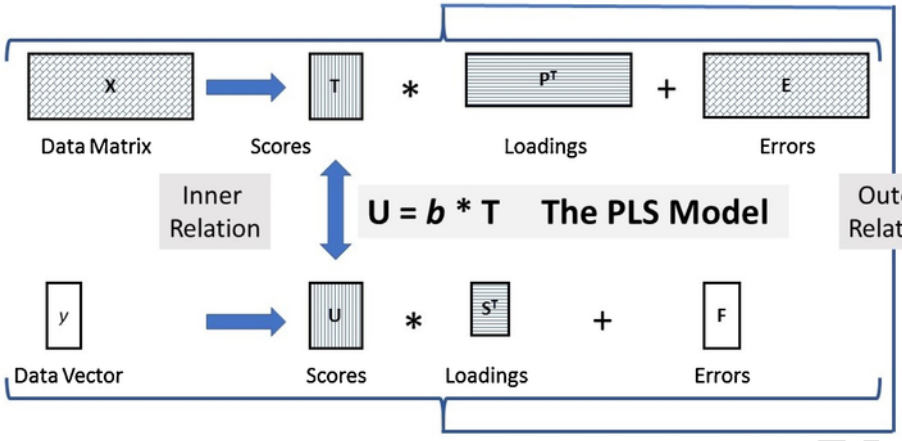

Fig. 2. Scheme of the operation of the PLS algorithm.

Only parallel factor analysis (PARAFAC) and multivariate curve resolution coupled to alternating least squares (MCR-ALS) present an inherent second-order advantage. For PARAFAC and MCR-ALS, the calibration and test (unknown) samples are placed together and decomposed by the model, so the number of factors necessary to perform the regression model is determined at once and, theoretically, equal the sum of the independent chemical species in the calibration samples and the unknown interferences in the test samples.

PARAFAC assumes that the data array for a group of samples follows a trilinear model [40], i.e., an element (i,j,k) is the sum of contributions of the form ($x_i \times y_j \times z_k$), where x_i is the relative concentration of a component in the i^{th} sample; y_j and z_k are the values of its qualitative features (for example, time and spectral profiles) at the j^{th}, k^{th} channel in each data dimension. The x_i values obtained are used for quantitative analysis using a pseudo-univariate calibration graph.

MCR-ALS places second-order data for a group of samples adjacent to each other along a dimension (usually time, or a dimension which has lost the tri-linearity), and assumes that the augmented matrix follows a bilinear model, i.e., a matrix element (m,j) is the sum of contributions of the form ($x_m \times y_j$), where x_m (m ranges from 1 to I × K) describes the profile for each sample in the augmented dimension, and y_j does the same in the other (usually spectral) dimension. For quantitative analysis, areas or heights of the sample responses (in the augmented dimension) are computed, and used to build a pseudo-univariate calibration graph [41]. Fig. 3 outlines the operation of MCR-ALS, where mixed information contained in process data, arranged in a matrix **D** can be decomposed into a product of matrices **C** and $\mathbf{S}^T$, which contain pure components information, which can be extracted after analysis.

3. Crystallization from solvents

Crystallization is one of the most important unit operations used for the preparation and purification of crystalline solids. In the pharmaceutical industry, one of its main goals is the production of a specified polymorph with a given morphology and crystal size distribution; in turn, these properties may have a large impact on the downstream unit operations, such as filtration and drying, extending its effects to the biopharmaceutical performance of the final drug products. Therefore, the appropriate design and control of the crystallization process is key to consistently obtain products with the desired properties.

PCA derived multivariate statistical process control charts were applied to the data measured from the whole crystallization process using a previously selected spectral range, to obtain the true nucleation point [42]. ATR-MIR spectroscopy was coupled to PCR and

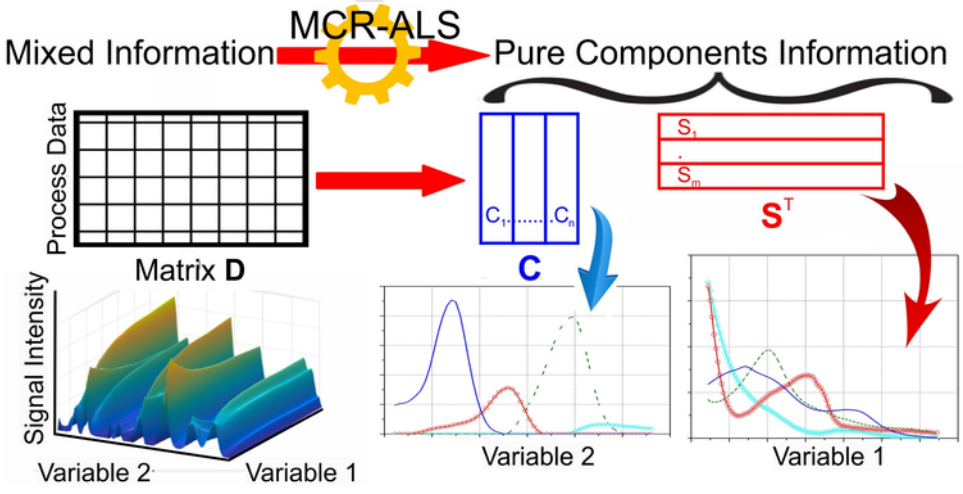

Fig. 3. Schematics of the operation of MCR-ALS.

PLS and applied to multicomponent pharmaceutical systems for the prediction of the solution concentration and the determination of the solubility curves [43]. The effects of temperature, solvent composition and impurities were modelled and the estimations of the inferential models were in close agreement with the corresponding HPLC results.

On the other hand, critical parameters of the APIs are determined by the supersaturation profile achieved during crystallization. Therefore, modelling, measurement and control of supersaturation, are critical for the crystallization process. The task of accurately determining the solution concentration in dense crystal slurries typical of industrial operations is often challenging. Vibrational spectroscopies, especially MIR and NIR, coupled to chemometrics have been proposed as modern day approaches to real-time in-line sampling without disturbing the production environment [44].

In another example, a database of 218 solvents with 24 property descriptors was explored and visualized using multivariate tools, aiming to select solvents for polymorph screening. PCA analysis exhibited 57% cumulative variance being explained by the first two PCs. A self-organizing map (SOM) was deemed more suitable for data display, being chosen as the visually most convenient way to examine solvent diversity. The strategy also demonstrated that safety aspects can be considered by labeling the solvents with toxicological information [45]. Some chemometrics applications to this problem are detailed in Table 2.

A deeper insight into the crystallization process of semi-crystalline polymers during formation of solid dispersions is always crucial to improve control of product quality in drug formulation. PLS modelling of small angle X-ray diffraction (SAXD) data of solid dispersions of lipids was used in a study where PEG 4000 with 12 different lipids were employed as a model system to examine the effect that incorporated components may have on the crystallization of the polymer [51]. The lipids were melted with PEG 4000 and the crystallization of the polymer was studied with DSC and small angle X-ray diffraction (SAXD). The PLS models indicated that small hydrophilic lipids increased the folding of PEG and that large nonpolar lipids retarded the unfolding during secondary crystallization. Thus, it was shown that the added lipid governs the behaviour of PEG solid dispersions.

Table 2
Chemometrics-assisted methods to monitor drug crystallization.

Drug	Method	Observations	Refs.
Carbamazepine	ATR-MIR/PLS	Method used to ensure the absence of metastability with respect to the undesired polymorphic form along the process.	[46]
Monosodium glutamate; L-Glutamic acid	ATR-MIR/ELSS (extended loading space standardization)	ELSS handles better than PCR/PLS temperature-induced spectral variations, scaling effects resulting from the use of physical concentration units, and variation in optical path-length during batch cooling crystallization. On-line method.	[47]
Sulfathiazole	ATR-MIR/PLS (Immersion probe)	Effect of cooling conditions on supersaturation level and process outcome were studied. Concomitant automated image analysis provided product size and shape information.	[48]
Trehalose	NIR/MLR, PCR, PLS	PCR and PLS exhibited the best performance.	[49]
Undisclosed API, MeOH	NIR/PLS (Immersion probe)	PAT tool to monitor in real-time the API (bias against HPLC = 2.88%) and the residual methanol (0.10-0.13 *w/w* with a maximum bias of 0.02%) contents. To control the seeding of an API crystallization at industrial scale.	[50]

4. Determination of physical properties of bulk pharmaceutical solids

NIR spectroscopy has been receiving an increasing attention in the pharmaceutical industry as one of the tools that fits into the PAT initiative, because it can be used to achieve rapid, real-time, non-destructive determinations with little or no sample preparation. Particle size and shape have a profound influence in the bulk properties of the powders and determine their ability to flow, mix, granulate, and dissolve properly.

NIR spectroscopy is particularly sensitive to these features; hence, not surprisingly, a number of NIR/chemometrics based methods have appeared in the literature demonstrating the distinct ability of this spectroscopy to help to predict physical properties of the bulk solid. Just a few examples of them are detailed in Table 3. Raman spectrometry has also been employed for the same purposes. Coupled to PLS, the resulting system was suggested as a suitable tool to examine the variability of tablet coatings, to predict its thickness [52].

5. Identification and assay for quality control purposes

5.1. Identification of active pharmaceutical ingredients at the solid state

In the pharmaceutical industry, quality control (QC) is a key activity within the process of ensuring that medicines have the required quality, safety and efficacy for their intended use. Hence, at the pharmaceutical companies, the QC departments are responsible for all release testing of final products but also all incoming raw materials. According to the official texts, the identification tests should establish the identity of the analyzed product and be able to discriminate between similar or related compounds that may be part of the sample. Therefore, these tests should be as specific as possible, and the lack of specificity of an identification method can be solved by combining several, preferably unrelated, alternatives.

Three different analytical techniques have been coupled to PLS-DA to identify acetaminophen present in pharmaceutical formulations. Two of these studies, employing XRPD and DSC, respectively, have been developed by the group of Komsta [53,54]. The third set of models was built from ssNMR spectroscopic data, with the aim to detect the presence of acetaminophen in over-the-counter pharmaceutical formulations [55].

In the latter case, a dataset of 11 spectra of pure substances and 21 spectra of different formulations was processed by PCA and PLS-DA. In the PCA approach, it was observed that the signal of a co-formulation ingredient strongly affected the second PC, turning it less reliable. On the other hand, the PLS-DA model was found to be more suitable, especially when after variable selection it proved to keep its performance with only 300 sensors.

Structurally related radiological contrast media (iodixanol, iohexol, caldiamide sodium and gadodiamide), were identified by NIR-chemometrics methods. The performances of classification models (SIMCA, PLS-DA), Main and Interactions of Individual Principal Components Regression (MII-PCR) and backward variable elimination-PLS (BVE-PLS) were compared. Variable selection methods were applied to optimize the classification models. BVE-PLS and MII-PCR were found to be most effective, not recognizing <1.5% of the samples [70].

Table 3
NIR/chemometrics prediction of some physical properties of bulk solids.

Drug	Chemometrics Method	Observations	Reff.
Acetaminophen and excipients	PLS	Prediction of particle size distribution, powder flowability, angle of repose, aerated and tapped bulk density, and components concentration.	[56,57]
Antipyrine	PCR	Prediction of mean particle size (MPS), angle of repose, tablet porosity and tablet hardness, in granules. MPS increased with the amount of water; larger spherical granules with narrow size distribution were made using a high-speed mixer.	[58]
Amoxycillin·3H_2O	PLS	Identification and particle size determination.	[59]
Berberine chloride	PCR	Prediction of tablet hardness and porosity.	[60]
Hydrochlorothiazide Chlorpheniramine	PLS	Determination of tablet hardness.	[61]
Lactose	PLS	Granule behaviour during fluid bed drying. Determination of water content and granule size.	[62]
Mannitol	MIIPCR	Determination of residual moisture content in freeze-dried product. Superior to PLS. MIIPCR is Main and interactions of individual PCR.	[63]
Microcrystalline cellulose	PLS	Prediction of powder density (tap, bulk, consolidated).	[64]
Orbifloxacin	PLS	Prediction of compression force, crushing strength and content uniformity. PLS performed better than MLR.	[65]
Orbifloxacin	PLS	Prediction of tablet coating thickness (with Eudragit RS/RL).	[66]
Sulpyrine	PCR	Determination of tablet hardness and weight variability.	[67,68]
Theophylline AN	PLS2	Calibration for crushing strength and relative density.	[69]

A Raman/SVM classification strategy was applied for the identification of 25 product families without error and in the absence of prior information about the sample [71]. A Raman-chemical imaging (CI) approach was used to characterize a self-emulsifying drug delivery system. This allowed the follow-up of the formulation during stability studies. A quantitative Raman-CI/PLS to assay the API in the lipid based formulation was also developed and fully validated following the "total error" approach. SNV and MC were employed as data pre-processing methods [72]. On the other hand, SIMCA and ANN were employed to classify polymers based on their thermo-mechanical properties [73].

Terahertz spectroscopy was coupled to SIMPLISMA (simple-to-use interactive self-modeling mixture analysis), a self-modeling curve resolution technique closely related to MCR-ALS, to interpret two-way THz spectral data of unknown mixtures of aminoacids and excipients for spectral resolution of single species and identification of their ingredients. An inverted second derivative was employed as an intermediate step to optimize the determination of the pure variables [74].

A combined NIR-MIR-based spectroscopic/chemometrics method was developed for the analysis of the Traditional Chinese Medicine, *Illicium verum* Hook. F., and its noxious adulterant, *Illicium lanceolatum* A.C. Smith. The spectral matrix was submitted for classification with LDA, using the successive projections algorithm (SPA) or the discrete wavelet transform as pre-treatments. The SPA-LDA and MIR approaches performed somewhat better (95–100% correct classifications) [75].

The classification of Wuyi rock tea was performed by ssNMR with the aid of PLS-DA, and quantitative descriptive analysis (QDA). The contents of caffeine, carbohydrate, polyphenol, and terpenoid were distinguished [76].

The discrimination among pharmaceutical products is an important task in the pharmaceutical industry and for safety issues. A NIR/PCA method was used for the discrimination of Chinese patent medicines, under the principal component accumulation (PCAcc) mode. In the PCAcc method, an accumulation strategy is used to combine the classification information contained in multiple PC subspaces by using a rotation, a projection and a summation operation. To improve the performance of classification, continuous wavelet transform is applied as the pretreatment method to eliminate the background [77].

Finally, it is worth mentioning that homeopathic medicines have also been examined by spectroscopic means coupled to chemometrics methods, concluding that they are "not visible" with MIR spectroscopy. Ten homeopathic remedies (organic and inorganic) as sugar granules, were daily subjected to triplicate solid-state MIR and NIR spectroscopy during 6 days, and the results were analyzed by PCA and PLS-DA. No visible differences besides atmospheric drift of spectra could be observed [78]. Table 4 compiles a selection of examples on chemometrics-assisted detection of APIs. The chemical structures of −the studied compounds participating are depicted in Fig. 4.

Table 4
Chemometrics-assisted identification of active pharmaceutical ingredients in solid-state samples.

Drug	Method	Scope and Observations	Ref.
Acetaminophen	ss-NMR/ PLS-DA ss-NMR/ PCA	Used to build discriminant models to detect the presence of the drug in over-the-counter formulations. PLS-DA with 3 factors (LV) performs better than PCA. RMSECV = 0.38; RMSEP = 0.48.	[55]
Acetaminophen	XRPD/PLS-DA	Recognize the presence/absence of the drug in a multi-component tablet. Pre-process: SNV; LV = 4.	[53]
Acetaminophen	DSC/PLS-DA	Recognize the presence of the drug in the formulations. Pre-process: SNV, MC and scaling; LV = 5. RMSECV = 0.15; RMSEP = 0.39.	[54]
Bisoprolol and hydrochlorothiazide	NIR/PCA Raman/PCA	Discriminate between different dosages (dosage form identity). NIR exhibited better results than Raman.	[79]
Bisoprolol and hydrochlorothiazide	NIR/PLS Raman/PLS NIR-Raman/ SO-PLS	Semi-quantitative prediction of drug dosage. Analysis performed for each API at a time. NIR exhibited better results than Raman. Sequential-orthogonalized PLS (SO-PLS) showed better results than PLS, when applied sequentially to NIR and Raman spectroscopic data.	[80]

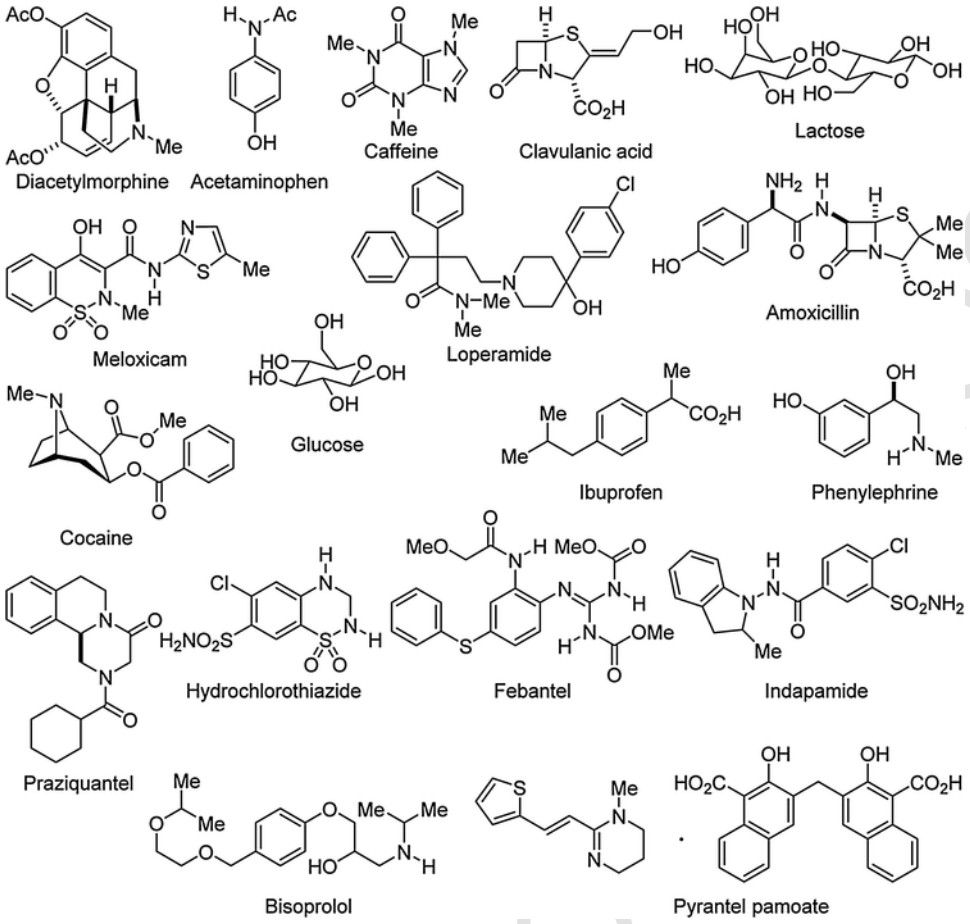

Fig. 4. Chemical structures of some of the compounds used in the development of methods for the chemometrics-assisted identification or quantification of active pharmaceutical ingredients in solid-state samples.

5.2. Assay of active pharmaceutical ingredients at the solid state

Quantitative analysis of the composition of pharmaceutical mixtures at the solid state is often used to ensure the safety and efficacy of their active pharmaceutical ingredients or to establish and validate the control of the pharmaceutical production process. Each quantitative option presents its own issues when it comes to ensuring accuracy and precision of the solid-state method. Some of these aspects have been presented and discussed in several recent reviews [4,80].

It is worth taking into account that in general, the solid-state methods suffer from the inherent difficulty of achieving sample uniformity. Hence, a well-developed and rationally optimized method will often be capable of quantifying down to 5% and exhibit relative standard deviations up to approximately 10%.

On the other hand, some problems present non-linearity effect which can be handled with variable selection before applying chemometrics methods [81]. Table 5 details selected examples of chemometrics-assisted methods for determination of APIs in mixtures at the solid state. On the other hand, the chemical structures of the compounds participating in these studies are depicted in Fig. 4.

6. Detection of adulterated drugs and medicines

As improved analytical methods are developed and more thoroughly used, an increasing number of cases of counterfeit and substandard drugs is revealed all over the world. The variety of such products is so wide that the World Health Organization (WHO) has coined the special term "SSFFC Medical product" to refer to these spurious, substandard, falsified, falsely labeled and counterfeit drugs [103].

Raman spectroscopy has been repeatedly coupled to chemometrics tools for detection and chemical profiling of counterfeit medicines [104]. Confirmation of the authenticity of an API is a challenge for the pharmaceutical industry. SIMCA classification models built on Raman spectra of samples of sorbitol have been shown to enable to distinguish between pure and impure samples of the product, including adulterations (<5%) with ethylene glycol and diethylene glycol, with no misclassifications. Adulterant concentrations as low as 2% placed the samples in a separate group. In addition, Raman spectroscopy was coupled to PLS to develop a quantitative screening tool to detect trace-level adulteration of sorbitol with ethyleneglycol. The LOD of this method was 0.9% [105].

In a recent Raman/PCA study, it was possible to discriminate between samples with different coatings, those containing varying amount of their corresponding active pharmaceutical ingredients (caffeine and quinine sulfate dihydrate, Fig. 5) and a diversity of excipients [106]. However, in the same study it was possible detected small chemical changes in the composition of acetylsalicylic acid and ascorbic acid tablets, which were caused by inappropriate storage conditions.

Table 5
Chemometrics-assisted quantification of active pharmaceutical ingredients in solid-state samples.

Drug	Method	Scope and Observations	Refs.
Acetaminophen	NIR/PLS NIR-SRS/PLS	To examine the ability of NIR-spatially resolved spectroscopy (SRS) to evaluate tablet content and sample heterogeneity. The single-point-NIR measurement spectrometer (RMSEP = 1.27%) outperformed the SRS apparatus (RMSEP = 1.71%). SRS outputs were more sensitive to tablet heterogeneity.	[82]
Acetaminophen	NIR/PLS	To compare the results obtained with those of transmittance and diffuse reflectance for a quantitative determination. Pre-process: SNV, D'. Spectral region: 9500–7500 cm^{-1}. For powder mixtures, transmittance mode gave RMSEP values 2.4–5.6 times lower than the diffuse reflectance measurements.	[83]
Acetaminophen, Caffeine, Diacetylmorphine (heroine)	DAD-NIR/PLS FT-NIR/PLS	Rapid in situ quantitative analysis of drugs with fiberoptic portable DAD-NIR. Methods exhibited similar performance to reference FT-NIR. Explained variance >90%; RMSE < 2%. r^2 = 0.909 (DAD-NIR) and 0.989 (FT-NIR).	[84]
Acetaminophen, Caffeine·H_2O, Lactose·H_2O	XRPD/Peak intensity XRPD/MCR-ALS XRPD/PLS	Simultaneous determination of multiple components in mixed powder based on XRPD (2θ = 5.00–30.0 and 35.0–45.0°). Among the three approaches, the PLS method gave the highest accuracy and precision; it can also be applied to the higher XRPD range (35.0–45.0°) where the other cannot.	[85]
Acetaminophen, Caffeine, Phenylephrine	Raman/PLS2 Raman/PLS	Quantification of the constituents in 5 components (3 APIs and 2 excipients) tablets. PLS (5 models, each optimized for one component) performed better than PLS2 for the lower concentration components. Minor improvements for optimized spectral range and number of latent variables.	[86]
Amoxicillin	NIR/PLS	For the drug in suspension. Transflectance mode was used. Pre-process: MSC. LV = 7; range: 40–65 mg mL^{-1} of amoxicillin; RMSEP = 1.6 mg mL^{-1}.	[87]
Amoxicillin, Clavulanic acid	MIR/PLS MIR/iPLS MIR/siPLS	Simultaneous determination of two APIs in commercial products. DRIFTS mode was used. Synergy interval-PLS (siPLS) gave better results than interval PLS (iPLS) and PLS. For clavulanic acid: Spectra divided in 30 intervals and combinations of 2 intervals. For amoxicillin: Spectra divided in 10 intervals and combinations of 4 intervals. Good correlation with reference HPLC method.	[88]
Amoxicillin·$3H_2O$	NIR/PLS	Development, validation and applicability of a model on real samples. LV = 4, r^2 = 0.9937, RMSEC = 2.17%, RMSEP = 2.38%. Good correlation with the reference HPLC method.	[89]
Azitromicin	Raman/SVM Raman/k-NN Raman/PLS	Classification of tablets according to manufacturer, employing SVM and k-NN and PLS-DA. PLS used for quantitative determinations. iPLS and Monte Carlo based uninformative variable elimination were used to select informative variables for improving the models. Classification and prediction were highly accurate. Pre-process: MSC.	[90]
Caffeine, Loperamide.HCl	FTIR-ATR/PLS	Quantification of two APIs at different doses in printed formulations. Pre-process: SNV. Spectral range: 400–1750 cm^{-1}. LV = 5 for both APIs. Improved predictive performance with regard to the univariate approach.	[91]
Caffeine, Cocaine, Glucose	Raman/PCA Raman/PLS	Analysis of an illegal drug and two typical diluents. PCA demonstrated that the mixtures could be discriminated according to their concentration of cocaine and either diluent. RMSEP = 4.1% (cocaine), 5.2% (caffeine) and 6.6% (glucose).	[92]
Epinephrine, Ibuprofen	NIR/PLS	Quantification of the enantiomeric excess of two APIs. Pre-process: MSC and D". In the absence of impurities or excipients, RMSEP < 2% for both substances. For ibuprofen in the finished products, RMSE = 7.0%.	[93]
Febantel, Pyrantel pamoate, Praziquantel	MIR/siPLS	Simultaneous determination of three APIs in veterinary formulations. Pre-process: Savitzky–Golay, smoothing, D', MC. LV = 5, RMSEP ≤ 0.69% for the three analytes. Good agreement with HPLC–DAD and HPLC–MS/MS results.	[94]
Hydrochlorothiazide	NIR/PLS	Determination of the API in drug products, employing a difuse reflectance approach. Pre-process: MC, smoothing (Savitzky–Golay), D' and normalization. Spectral region: 1640–1780 nm, LV = 4. RMSEP = 1.7%. NAS was used to obtain the figures of merit. Results in agreement with reference HPLC method.	[95]
Ibuprofen	Raman/OPLS NIR/OPLS	Non-destructive quantitative analysis of the API in solid pharmaceutical formulations. Pre-process: SNV and MC. Spectral region: 134.6–2238.5 cm^{-1} (Raman) and 7105–13,100 cm^{-1} (NIR). Both models predicted similar tablet contents when all matrix variations were included in the calibrations.	[96]
Ibuprofen	NIR/PCA NIR/PLS	Determination of the API in low concentration (0–5%, *w/w*) tablets. PCA was unable to discriminate between tablets below a concentration of ∼0.1% *w/w*. At low upper levels, additional PLS factors are needed to obtain useful models.	[97]

Table 5 (*Continued*)

Drug	Method	Scope and Observations	Refs.
Indapamide	NIR/PLS	High-throughput method for the determination of drug content in tablets (2–3 mg/unit; 2% *w/w*). Pre-process: SNV. LV = 6; RMSEE = 0.0311%. Results agree with HPLC method.	[98]
α-Lactose, Citric acid, Fructose	THz/PCA THz/PLS THz/ANN	Qualitative and quantitative analysis of ternary mixtures. PCA allows to place the samples in a ternary diagram. RMSE = 0.9% for the three constituents and both quantitative methods. No advantage for ANN (spectra driven by linear behaviour).	[99]
Meloxicam	NIR/PLS	Assay of API in tablets; determination of crushing strength and disintegration time. Pre-process: MSC. LV = 6; r^2 = 0.988, RMSECV = 0.22%. Good agreement with an HPLC-UV method and with European Pharmacopeia methods for crushing strength and disintegration time.	[100]
Piroxicam	MIR/PCR MIR/PLS	For assay of API in tablets and ointment. Univariate determinations seemed to give better results (RSD < 3%).	[101]
Undisclosed API	Raman-CI/ MCR-ALS SER-CI/ MCR-ALS	Quantitation of an API at low concentration (0.5–2%) in a dosage form. The Raman-CI spectra were acquired at 0.8, 3.0 and 10 s. MCR restrictions: Non-negativity for spectra and concentrations, and normalization of spectra. The drastic Raman signal enhancement in the presence of silver nanoparticles in surface-enhanced Raman (SER)-CI provided significantly improved calibration accuracy and decreased image acquisition time. r^2 = 0.9849.	[102]

The latter modifications took place prior to visual change of the tablet color. Therefore, it has been recommended to ensure that the observed changes are not a result of inadequate storage before classifying any sample as a counterfeit medicine.

A study by a Brazilian team using ATR-MIR coupled to PCA provided fast (less than 30 s) and reliable results in the forensic analyses of samples of sildenafil citrate and tadalafil. Further, it was observed that the PCA scores corresponding to many counterfeit drugs coming from unrelated seizures from different cities of Brazil were inserted in the same cluster, suggesting a common illicit source for all these medicines. The combination of PCA and similarity match (SM) techniques enabled detecting all counterfeit medicines and grouping different seizures of the illegal products [107].

In a related investigation, MIR spectroscopy was employed on a larger set of samples of sildenafil citrate and tadalafil with the aim of obtaining classification models helpful to customs officers in their effort to obtain a first evaluation of suspected samples during import inspections. Among the tested methods, models based on Classification and Regression Trees (CART) gave good performances; however, SIMCA resulted provided a 100% correct discrimination between genuine and counterfeit drugs. On the other hand, k-NN was not able to make the desired discrimination; therefore, it was considered not useful [108].

Some expired drugs are difficult to detect by conventional means. If they are repackaged and sold back into market, they will constitute a new public health challenge. A recent study for the detection of repackaged expired drugs still within drug specifications, employed commercial paracetamol tablets as a model drug. Their Raman spectra were compared against a spectral library, employing verification and classification methods. Both methods were able to confidently detect the expired drugs [109]. Table 6 details some applications of chemometrics methods to this task, which has recently been reviewed [110].

7. Solubility and dissolution studies

7.1. Drug solubility

The solubility of the active pharmaceutical ingredient of a formulation is important for its biological activity. Hence, the estimation of the solubility of a drug in the design stages, and the optimization of the formulation for proper dissolution are relevant for product development. Chemometrics strategies have been employed in order to predict drug solubility and solid state behaviour of APIs. Discussed below are a few selected examples.

PCA analysis [116] of a set of 51 compounds revealed that compounds with comparatively high molecular weight and complex molecular structure displayed increased glass-forming ability. The latter parameter has also been predicted from the molecular structure by PLS-DA [117,118]. The model suggested that molecular descriptors related to size, symmetry, branching, number of aromatic rings, and distribution of electronegative atoms impact on the glass-forming ability for 75% of the test compounds. MLR also enabled to develop a predictive model of the stability of amorphous drugs, especially poorly water-soluble compounds, employing molecular weight and enthalpy of fusion data as predictors [119]. PCA was used to ensure that the sample set was representative of the chemical and physicochemical diversity of neutral poorly soluble drugs.

On the other hand, in order to better understand the factors behind the poor solubility of drugs, the intrinsic aqueous solubility of 15 poorly soluble drugs (2.9–1100 nM) was analyzed from a physicochemical perspective, using experimentally determined solid-state properties and two-dimensional molecular descriptors. The analysis revealed that poorly soluble drugs are solubility limited by solvation rather than by their solid state. PCA was used to display the descriptors and PLS to predict de solubility based on these descriptors [120]. This analysis revealed that descriptors related to lipophilicity, size, and polarizability are important factors for restricted solubility.

In another study, PCA and PLS were employed in an attempt to gain insight into the relative importance of counterion characteristics on their corresponding salt properties, and to establish predictive models capable of describing their effect on 11 procaine derivatives. Some success was achieved in modeling the crystalline salt solubility and the glass transition temperature of the amorphous salts [121].

In a more comprehensive study, Wassvik et al. determined the solubility and solid-state of 299 compounds; in their study, PLS was used as a multivariate data analysis strategy to deduce relevant structural features of marketed drugs with limited solubility. It was found that molecules with extended ring structures and large conjugated systems were less soluble, indicating that structural features related to rigidity and aromaticity resulted in restricted solubility. Based on these findings, 2D molecular descriptors of rigidity and aromaticity were applied in a PLS analysis to predict the solubility of unknowns. The model successfully predicted the solubility of the test set and the authors suggested that such calculated molecular descriptors can be

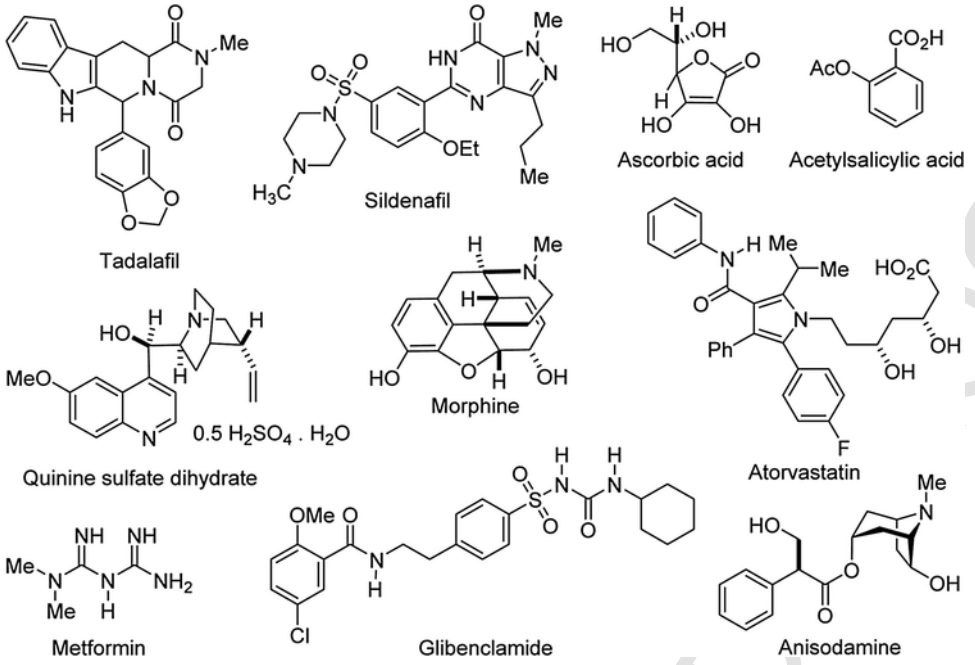

Fig. 5. Chemical structures of some of the studied compounds for the chemometrics-assisted detection of adulterated and counterfeit drugs and medicines.

used for rapid identification of synthetic target compounds with a high risk of having solid-state limited solubility [122].

Higher internal energy and molecular mobility of disordered materials can lead to increased aqueous solubility, and improved bioavailability. The potential for small molecule organic crystalline materials to become disordered as a result of high shear mechanical processing was investigated by stepwise multivariate logistic regression of different materials properties. Glass transition temperature, melting temperature, heat of fusion, crystallographic density, Young's modulus, molar volume and attachment energy were identified as having a significant correlation with the potential for material disordering [123].

7.2. Pharmaceutical dissolution

The pharmaceutical dissolution rate and the associated dissolution profiles are invariably identified as critical quality attributes for solid dosage forms, since they are related to the drug availability for absorption, a needed stage to exert its effect. Many advances have been recently made regarding the application of chemometrics tools to pharmaceutical dissolution problems, including the studies of the factors which impact on the dissolution of drug products, the quantification of the dissolved drugs during the dissolution test and the determination of the crystalline state of the API from its dissolution profile.

Analyte quantitation is another way to apply chemometrics to dissolution analysis. A demonstrative example is the study of the initial dissolution kinetics of the carbamazepine-nicotinamide co-crystal, which was investigated with the aid of the CLS method. The co-crystal was prepared by co-milling and solvent evaporation methods, and the dissolution tests of the solid were performed using an original flow cell and an UV–vis spectroscopic detector, which enabled in situ process monitoring. The acquired spectra were analyzed with CLS to separately determine the dissolved compounds. The initial dissolution profiles were interpreted using a simultaneous model of dissolution and phase changes [124].

An application of the method to a drug formulation was also presented. The tablets, obtained by direct compression under different conditions, were scanned at-line using NIR spectroscopy in transmission mode. A MLR model was built with the PCA scores of the NIR spectra and the dissolution profile parameters, obtained under two conditions, including level and shape calculations (model-independent) and after fitting a Weibull curve (model-dependent). The models successfully predicted the dissolution profiles of the individual tablets manufactured at the targeted set point (similarity factor, f_2 = 72) [125].

In related studies, we have shown that MCR-ALS can be employed as a chemometrics tool to monitor the simultaneous dissolution of pairs of drugs from fixed doses binary pharmacological associations [126,127], and to use the resulting dissolution curves to build dissolution profiles, useful for comparison between brands or lots, as well as for quality control purposes (Fig. 6).

The dissolution curves are built at once from the set of dissolution results. This approach, which uses a few calibration standards is also markedly different from the previous one, which employed PLS as the chemometrics tool to build the dissolution curves at a time-point by time-point bases, and required more calibration standards, preferably arranged according to an experimental design [128–130].

Another interesting application of chemometrics to the analysis of the dissolution is in the development of alternative statistical methods to assess similarity of dissolution profiles.

The group of Wang challenged the official method based on the f_2 factor with a set of 16 groups of dissolution profiles arranged according to a full factorial design (drug strength, tablet stability time, and dissolution testing condition). According to their values of f_2, all 16 groups were considered similar. However, a MANOVA of repeated measures suggested statistical differences. Therefore, the authors used a modified PCA to describe the dissolution curves in terms of level and shape.

This method categorized the set into three similar groups, and was consistent with the MANOVA test and its subsequent analysis using Tukey's Test. They concluded that their approach enables an im-

Table 6
Some applications of chemometrics methods to counterfeit drug detection.

Drug	Method	Scope and Observations	Refs.
A calcium channel blocker	NIR/DD-SIMCA	Counterfeit drug detection. This approach enhanced the applicability of a miniaturized NIR instrument. DD-SIMCA is Data driven SIMCA.	[111]
Acetaminophen	Raman/PLS-DA Raman/SVM Raman/k-NN	Analysis of expired drugs. Pre-process: Max-Min normalization. Average accuracy: 90.1%, 96.8% and 89.4% respectively.	[109]
Anisodamine	Raman/r^2	Identify genuine and counterfeit tablets. Counterfeit tablets identified with 100% predictive accuracy.	[112]
Anisodamine	NIR/iPCA, NIR/PLS-DA	Distinguish between manufacturing plants. The rejection rate and recognition rate were both 100%.	[112]
Atorvastatin	NIR/PLS-DA Raman/PLS-DA	Accurate identification of genuine and counterfeit tablets. PCA revealed that storage conditions, affect the NIR data, due to the adsorption of water after unpacking from the blister.	[113]
Acetylsalicylic acid and Ascorbic acid	Raman/PCA	Detection of small chemical changes in counterfeit tablet composition, caused by inappropriate storage conditions before visually observing their effects.	[106]
Caffeine and Quinine sulfate dihydrate	Raman/PCA	Detection of a wide range of counterfeits. Samples with several coatings, varying amounts of the API and even different excipients were distinguished.	[106]
Glibenclamide, Gliclazide, Glimepiride, Glipizide, Gliquidone, Metformin, Pioglitazone	Raman/PCA Raman/LSLS (Local Straight-Line Screening)	Identify 3 types of counterfeits: Types I (no API, only excipients); Type II (different, unrelated API); Type III (structurally related less expensive API). Types I and II are much easier to discriminate. The correct rates of the 3 types were all >95%. Total sensitivity, specificity and accuracy are 96.8%, 97.5% and 96.4%, respectively.	[114]
Morphine derivates (H_2SO_4, HCl and base)	DSC-DTG/ HCA	Differentiation the origin of the samples. Classification model was built from structural fingerprints.	[115]
Sildenafil citrate And Tadalafil	ATR-MIR/ PCA	Distinction between authentic and counterfeit samples in forensic routine. Grouped samples according to their chemical profiles, distinguishing successfully between authentic and counterfeits.	[107]
Sildenafil citrate, Tadalafil and placebo	ATR-MIR/ PCA, k-NN, CART, SIMCA	Intended for customs, to obtain a first evaluation of suspected samples. SIMCA gave the best model, with a 100% correct discrimination between genuine and counterfeit drugs. k-NN was less satisfactory, being unable to make the desired discrimination.	[108]

proved statistical analysis, allowing to better ascertain both, the statistical significance and the clinical relevance of the results; hence, supporting regulatory decisions in a more objective form [131].

In the same way, Maggio et al. developed PCA-CR, a methodology based on PCA and Hotelling tests for the establishment of a confidence region [132], as a new approach to assess the dissolution similarity of a drug product, employing furosemide and acetaminophen tablets as models. Reference and test data from multiple pre-specified time points dissolution measurements were simultaneously subjected to PCA and pairwise comparisons, which involved plotting the weighed scores of the first two principal components of reference and test lots. The decision about "similarity" was taken by checking the inclusion of more than 80% of the test lot units in the 95% confidence ellipse of the reference samples.

Unlike the f_2 criterion, the proposed method reflects variability within the individual dissolution curves, being also highly sensitive to their shape and size variations (profile). Comparison between the area enclosed by the confidence ellipses and the region obtained from the bootstrap-calculated acceptable values of the corresponding f_2 tests suggested that PCA-CR represents, in general, a more discriminating standard.

This approach was also applied to differentiate among the three polymorphic forms of furosemide in a capsule formulation. In this approach, the calibration space was created from the different pure polymorphs of furosemide (I–III), formulated as the commercial product. The commercial samples were displayed into a two-dimensional PC space and the Hotelling test was carried on the entire group. Then, the polymorphic identity of the unknowns was assigned by using simple pre-established PCA-CR rules [133].

In another case, the preparation parameters of a dried meloxicam nanosuspension were studied with the aim to improve its dissolution. For improving the dissolution rate, the drug was formulated as a nanosuspension using different methods like emulsion diffusion, high-pressure homogenization, and sonication. A self-modeling curve resolution (SMCR) study on the XRPD patterns of the nanosuspensions revealed the strong interaction between the stabilizer and meloxicam and the crystalline form of the drug. This provided a better understanding of the formulation [134], which enabled to increase the rate of dissolution of the dried meloxicam nanosuspension.

Raman spectroscopy was used as a PAT tool for in-line measurement of API content during continuous manufacturing of strip films containing nanoparticles of fenofibrate and naproxen as model of poorly water-soluble APIs [135]. Their concentrations ranged from 3% to 26% *w/w* in the calibration model. PLS was employed for in-line and off-line measurements, yielding r^2 values > 0.9946 and RMSEC = 0.44%. Prediction errors were 1.3% and method robustness was established by considering sensing location, substrate speed and film thickness. PCA was used to explain the relations between processing variables and calibration models, suggesting that film thickness could also be monitored using Raman spectroscopy.

Theophylline crystal forms are AH and MH, the former displaying the higher dissolution rate. THz spectra of theophylline tablets containing AH, MH, microcrystalline cellulose and magnesium stearate exhibited a specific absorption peak at 0.96 THz, related to AH. PLS analyses were performed to correlate the tablet spectra with both, their content of AH (RMSECV = 2.89%; r^2 = 0.9927) and their dissolution rate, which was gradually delayed as the proportion of AH decreased (RMSECV = 3.29%; r^2 = 0.9423). No significant differences between predicted and measured amounts of dissolved drug in tablets stored for 12 or 36 h at 25 °C and 84% RH [136].

The amounts of riboflavin sodium phosphate and excipients were predicted from ATR-FTIR spectra using PARAFAC and N-PLS as multi-way modelling techniques. Data matrices consisted of dissolved and undissolved parallel samples with different drug content (2–16%) and spectra, collected at axially cut surface of the flat-faced matrix tablets. The N-PLS method was more robust for accurate quantification of the amounts of the components in the sample whereas the PARAFAC model gave approximate relative amounts of components [137].

With the aim to improve the rate of dissolution of the poorly soluble meloxicam in capsule form, mannitol was used as a carrier in different ratios, in physical mixtures and melted forms. MCR was

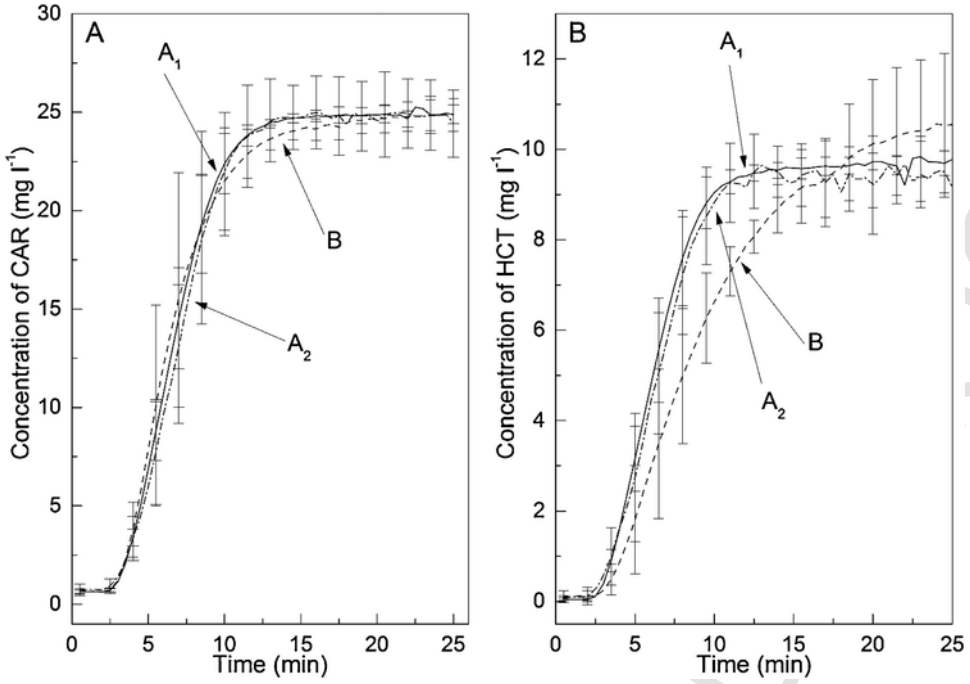

Fig. 6. Dissolution profiles of three lots of the hydrochlorothiazide-carvedilol association. Lots A_1 and A_2 are similar with regard to both APIs; however, lot B does not comply with the f_2 similarity factor criterion for hydrochlorothiazide. Taken from Ref. [126].

used as a chemometric method to interpret X-ray diffractograms of their binary mixtures, revealing that the amount of mannitol and the particle size of ME were important factors in the rate of dissolution [138]. A short summary of selected achievements is displayed in Table 7.

8. Stability testing

Stability studies designed to ensure the maintenance of product quality, safety and efficacy throughout the shelf life, are currently considered a pre-requisite for the acceptance and approval of any pharmaceutical product.

The purpose of stability testing is to provide evidence on how the quality of a drug substance or drug product changes with time under the influence of a variety of environmental factors such as temperature, oxygen, humidity and light. The tests allow the rational establishment of recommended storage conditions, re-test periods and shelf lives [144].

These studies are required to be conducted under the guidelines issued by the International Conference on Harmonization (ICH), WHO and or other agencies, which established different test conditions and requirements for active pharmaceutical ingredients and their drug products. The codes and subjects covered by the ICH guides are outlined in Table 8.

Chemometrics strategies have been employed in order to predict the stability of pharmaceutically relevant systems in the solid state. It has been shown that electron paramagnetic resonance (EPR) spectroscopy can be used to monitor the extent of oxidative degradation in solid-state samples of an active pharmaceutical ingredient. Moreover, compared to the traditional methods of peak height (PH) and double-integral (DI) measurements, the quantitation of the EPR response can be improved by the use of PLS. Q-band EPR and X-band Electron nuclear double resonance (ENDOR) spectroscopy also gave additional information to differentiate the organic radical species involved in the oxidation process [154].

The stability of a pharmaceutical product may also be altered by incompatibility with some excipients of the formulation. The studies of binary physical mixtures of atenolol and selected excipients by MIR spectroscopy under chemometrics assistance by PCA and CA have shown that β-cyclodextrin is incompatible with the drug. This finding was confirmed by complementary methods, such as DSC, thermogravimetry and XRPD [149]. Other applications of MIR spectroscopy to the characterization of pharmaceutical systems have been reviewed [145].

Amorphization is one of the techniques to enhance the dissolution rate and/or bioavailability of sparingly soluble drug substances. Generally, however, amorphized drug substances recrystallize easily, since they have a higher energy state and are physically more unstable. Therefore, in order to improve the physical stability of the amorphous drug substances, they are prepared as solid dispersions with a polymer. ^{13}C ssNMR has been coupled to chemometrics (PLS) in order to successfully predict the recrystallization behaviour of troglitazone during storage, in its solid dispersion with polyvinylpyrrolidone [146]. Interestingly, this strategy proved better than examination of XRPD, which exhibited no changes.

The decomposition of lisinopril dihydrate upon submission to heat (24–170 °C) was studied employing a MIR/MCR-ALS monitoring approach. The drug decomposes to the monohydrate, the anhydrate and the diketopiperazine. The pure spectra and the relative thermal-dependent contributions of each component were obtained, as well as the critical temperature for each transformation [147]. Table 9 provides some examples on the use of chemometrics methods in this field, whereas the chemical structures of the studied APIs are depicted in Fig. 7.

One of the causes of instability of multicomponent amorphous systems is their tendency to phase separate during storage. The characterization of amorphous–amorphous systems is often harder than in amorphous–crystalline ones, because it requires detection of phase separation. The PCA multivariate data analysis for X-ray powder diffraction-pair-wise distribution function (XRPD-PDF) data was em-

Table 7
Selected examples of chemometrics-assisted studies of drug solubility and pharmaceutical dissolution, and their applications.

Target Drug	Coupled method or data input	Scope, Predicted parameters and Observations	Refs.
Acetaminophen	NIR/PCA Dissolution/ PCA NIR/PCR	Display data dispersion and predict dissolution based on level and shape modeling or a Weibull curve. Calibration data constructed by a 3^{4-1} factorial design.	[125]
Acetaminophen	NIR/PCA NIR/PLS	Exploratory data analysis with PCA employed to study the spectral changes arising from the shear differences. PLS-2 used to develop a calibration model for% drug released vs. applied strain and predict tablet dissolution profiles.	[139]
API for asthma and COPD (chronic obstructive-pulmonary-disease)	PCA PLS	PCA was used to assess differences in powder properties and in vitro performance of batches of the drug. Variable importance in projection (VIP) was used in order to assess the most influential variables for powder characterization. Particle size, density and rate of flowability are significant for modeling the Delivered Dose of the API and the total quantity of powder related to each dose.	[140]
Carbamazepine–cinnamic acid co-crystal	UV/NAS-SAM	Standard addition method and net analyte signal (NAS-SAM) used for prediction of product solubility and its dissolution.	[141]
Fenofibrate and naproxen	Raman/PCA Raman/PLS	PCA was used to analyze the impact of process parameters on Raman spectra. PLS was used for in-line and off-line quantification of these poorly soluble drugs in strip films.	[135]
Furosemide Forms I, II and III)	PCA-CR	Assignment of polymorphic identity to an unknown in a capsules dosage form based on a principal component representation of its dissolution profile and its associated confidence region.	[133]
Furosemide, Acetaminophen	PCA-CR	Test dissolution profile similarity based on a principal component representation of its dissolution profile and its associated confidence region (CR).	[132]
Griseofulvin blends	PLS	Correlation between selected descriptors (lipophilicity, size, cohesive energy density, hydrogen bonding capacity) and dissolution efficiency (DE) was established by PLS (log DE_{blend}/DE_{pure}). It may be used to semi-quantitatively predict the dissolution behaviour of drug blends.	[142]
Hydrochlorothiazide and Bisoprolol	UV/MCR-ALS	Monitoring of the dissolution profile. Simultaneous construction of two dissolution profiles of concomitantly dissolving drugs from the drug association.	[127]
Hydrochlorothiazide and Carvedilol	UV/MCR-ALS	Monitoring of the dissolution profile. Simultaneous construction of two dissolution profiles of concomitantly dissolving drugs from.	[126]
Theophylline	THz/PLS	Prediction of the level of the anhydrous form and the dissolution profile of the product.	[136]
Undisclosed API	O2PLS	The influence of granule and compression variability (introduced by a DOE) on the entire dissolution profile was studied with bi-directional projection to orthogonal structures (O2PLS) as the chemometric tool. It was shown that the disintegration phase (10–15 min) was controlled by granule attributes and tablet hardness, while the later phase (15–30 min) was solely controlled by granule attributes.	[143]

Table 8
ICH Guidelines referred to stability.

Code	Document Title
Q1A	Stability Testing of New Drug Substances and Products
Q1B	Stability Testing: Photostability Testing of New Drug Substances and Products
Q1C	Stability Testing for New Dosage Forms
Q1D	Bracketing and Matrixing Designs for Stability Testing of New Drug Substances and Products
Q1E	Evaluation of Stability Data
Q1F	Stability Data Package for Registration Applications in Climatic Zones III and IV

ployed to detect phase separation in freeze-dried binary amorphous polymer-sugar mixtures. The analysis provided a more clear 'miscible' or 'phase separated' interpretation through the distribution pattern of samples on a score plot presentation compared to the residual plot method, being in agreement with DSC results.

Table 9
Examples of chemometrics-assisted stability studies of selected drugs.

Drug	Scope	Method	Observations	Refs.
Acetaminophen, Azithromycin, Lidocaine, Epinephrine	Quantitative analysis of known primary degradation products.	Raman/CLS	Average LOD = 3.9%, based on *y*-intercepts for the four analytes. Need a significant percentage of API by mass.	[148]
Atenolol	Incompatibility with excipients.	MIR/PCA MIR/CA	Incompatibilities only with β-cyclodextrin. Method reliability confirmed by DSC and XRPD.	[149]
Captopril	Evaluation of the sample degradation in relation to the date of manufacture.	NIR-CI/MCR-ALS	Enables visualization of the degradation process in different layers, especially top and bottom surfaces of the tablet. More informative than bulk analysis.	[150]
Cimetidine, Ranitidine.HCl, Famotidine	Photostability	NIR/SIMCA	Differences detected between the irradiated, zero and dark samples of ranitidine.	[151]
Ezetimibe	Hydration kinetics (23 °C/75% RH)	Raman/MCR-ALS	Hydration almost complete within 30 min.	[152]
Ezetimibe	To quantify drug hydration due to excipients humidity.	Raman/PLS	Cross validation errors ~0.6% (*w/w*), for both crystalline forms, and $r^2 > 0.96$.	[152]
Risperidone	Drug-excipient compatibility	MIR/PCA	Incompatibility with Mg stearate, lactose and microcrystalline cellulose. Verified using a stability-indicating LC method	[153]
Undisclosed API	Study of oxidative degradation, after 12 months of exposure to 25 °C/60% RH and 40 °C/75% RH.	EPR/PLS	Method was compared with HPLC ($r^2 = 0.966$) and with traditional options based on peak high and peak area.	[154]
Vaccine freeze-dried with attenuated virus	Normal and accelerated stability studied (4 weeks at 4 °C and 37 °C)	NIR/PCA MIR/PCA	Titer decreased at 37 °C. NIR found changes in hydrogen bond strength between the stabilizer and the virus proteins. MIR revealed decrease of β turn and increase of α helix.	[155]

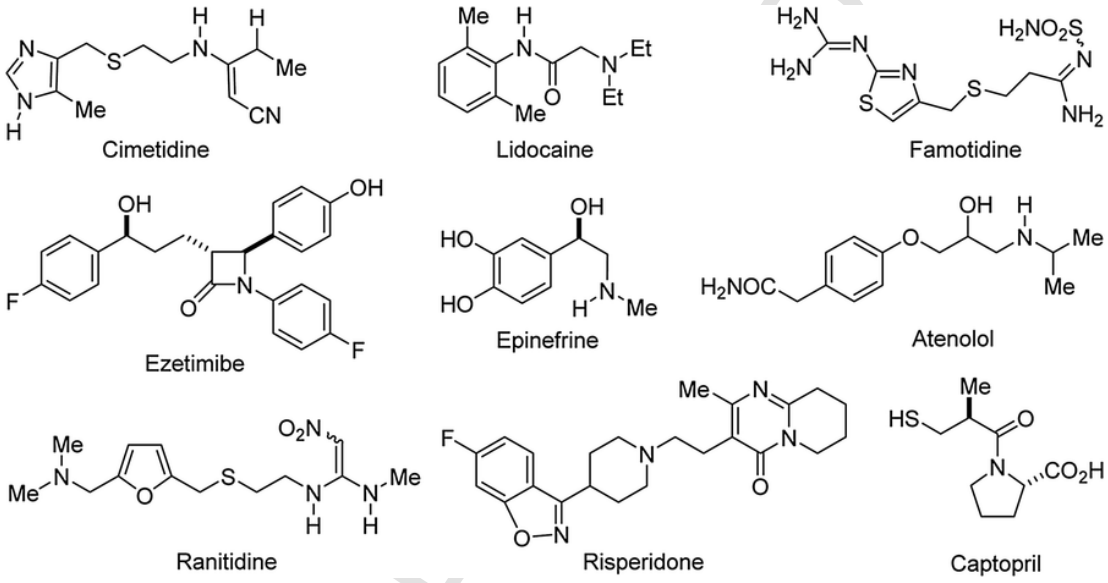

Fig. 7. Chemical structures of some of the compounds employed for the development of chemometrics-assisted analytical methods for stability studies.

This chemometrics-assisted approach improves the clarity of the PXRD-PDF results and was suggested as an alternative explorative data analytical tool in detecting phase separation in freeze-dried binary amorphous systems [156].

9. Thermal transformations

Solid-state transformations may take place during any stage of pharmaceutical processing and upon storage of a solid dosage form. Early detection and quantification of these transformations during the manufacture of solid dosage forms is important, since the physical form of an active pharmaceutical ingredient can significantly influence its processing behaviour, including powder flow and compressibility, as well as the biopharmaceutical properties of the final product such as solubility, dissolution rate and bioavailability. Polymorphic transformations under thermal stimuli, including generation of an alternate form, desolvation, crystallization and changes in crystallinity, have been detailed in a companion review [157].

Hot melt extrusion (HME) is a continuous process that can give rise to solid state transformations. The application of this method within the pharmaceutical industry is steadily increasing, considering that it can improve the solubility and bioavailability of poorly soluble drugs by encompassing them in a polymeric carrier and by forming solid dispersions. In HME, the use of high shearing forces, increase the blending rate and thus ensure the homogeneity and uniformity of the melt; In addition, high temperatures are used to produce and keep the melted intermediate. However, these forces may cause the polymer as well as the API to suffer changes, including degradation.

Different methods have been developed to study this process. For example, the solid state of celecoxib formulations prepared by HME was examined by PCA of the DSC thermograms of a single heating cycle. A narrow range of temperatures (100–190 °C) was used, be-

cause it was sufficient to capture the information related to the melting peak of the drug. The PC1 ($r^2 = 0.93$) versus PC2 ($r^2 = 0.06$) scores plot was able to cluster the end products in an identical manner as their XRPD analysis [158].

However, the off-line analysis with DSC and XRPD appeared not to be sensitive enough to detect small fractions of crystalline celecoxib that later on induced recrystallization of the drug in the extrudates during storage. This small fraction was noticed by an analogous on-line Raman/PCA approach, which allowed the differentiation of celecoxib as glassy solid solutions and crystalline dispersions, but also provided information on the stability of the extrudates.

The thermal decomposition of magnesium salts of organic acids used in pharmacy, such as valproic, lactic, citric and aspartic was studied by chemometric analysis of calorimetric and thermoanalytical methods [159]. The chemometric evaluation of the thermoanalytical results was performed by PCA. The PC1 versus PC2 plots of the thermal decomposition data suggested structural similarity among the magnesium salts of the organic acids.

Partial transformation (~70%) of cimetidine takes place above the melting point of the drug, upon heating between 160 and 180 °C. A multispectroscopic-chemometrics (UV-ATR-MIR-^{1}H NMR-^{13}CNMR/MCR-ALS) strategy enabled the characterization of the product as the stable N_3-enamino tautomer of the drug (Fig. 8).

This methodology also revealed that the transformation follows first order kinetics and provided parameters of its formation [160]. In this case, the use of a single spectroscopy was unable to provide the detailed quali- and quantitative information required to characterize the process and its final product.

A transflectance NIR spectroscopy approach was used to simultaneously measure in real time the drug and plasticizer content of polymer melts with varying opacity, during the HME process of carbamazepine in a polyvinyl pyrrolidone-vinyl acetate co-polymer (PVP-VA) matrix, employing polyethylene glycol (PEG) as plasticizer [161]. Calibration and validation of PLS models for the analytes were performed using a wide range of drug and plasticizer loadings. Once calibrated and validated (RMSEC = 0.79%; RMSEP = 0.67%), the technique was used to simultaneously quantitate ($r^2 > 0.99$) both widely different analytes.

10. Co-crystals, co-amorphous and salts

Pharmaceutical co-crystallization is a promising alternative to improve the solubility and dissolution rate of APIs, and to modulate other key physical properties based on crystal engineering approaches. The co-crystallization process involves the interaction among several molecular species and its control requires precise knowledge of the different phases that might appear during their production and the ability to monitor their abundance.

On the other hand, salts are usually prepared as a means to obtain highly crystalline, easily purifiable compounds, and used as such for the preparation of drug products. Their interaction with the excipients matrix may cause instability, chemical reactions and even disproportionation.

Raman spectroscopy has been associated with chemometrics means (PCA and PLS) to quali- and quantitatively study the disproportionation of pioglitazone hydrochloride in low loading (5% *w/w*) tablets [162]. A two-step Raman mapping approach (1500–1800 cm^{-1}) with suitable sensitivity was developed. The first stage was to locate the area of interest where the drug particles reside throughout the tablet surface employing a deliberate sub-sampling strategy. The second stage was a step by step mapping of the selected area to examine more in detail the degree disproportionation of the salt. This may help formulation scientists to better understand *in situ* drug-excipients compatibility. Table 10 provides some examples on the use of chemometrics methods in this field.

11. Chemical imaging

Hyperspectral chemical imaging is the analytical capability to create a visual image of the components distribution in a sample by integration of imaging and spectroscopy through the simultaneous acquisition of spectral and spatial/time information.

In recent times, this technique has acquired great importance in pharmaceutical analysis because it offers the possibility of accessing to visual representations of the spatial distribution of the analytes and enables their simultaneous identification and quantitation. Currently, chemical imaging is mostly based on vibrational spectroscopies [173], but other technologies are emerging at a fast pace. Due to the

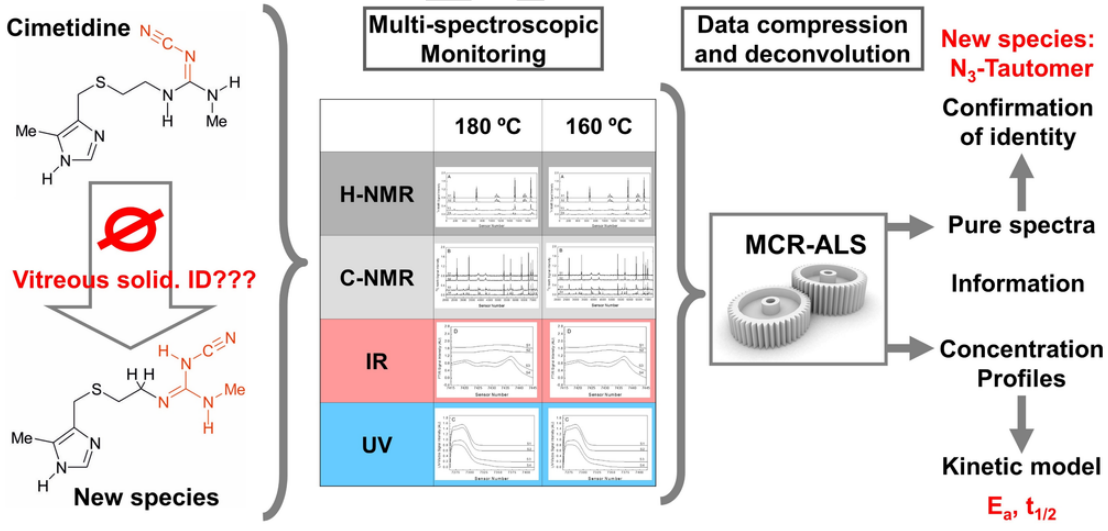

Fig. 8. Multispectroscopic/chemometrics (UV-ATR-MIR-^{1}H NMR-^{13}C NMR/MCR-ALS) approach toward the characterization of the stable N_3-enamino tautomer of the drug as the main product formed upon heating cimetidine above its melting point. Taken from Ref. [160].

Table 10
Characterization of co-amorphous and co-crystals. Use of chemometrics methods.

Co-crystal	Coupled method	Observations	Refs.
Carbamazepine:saccharin (1:1 co-crystal)	MIR/PCA MIR/PLS	System suitability examined, for the quantification of the co-crystals with carbamazepine form I in mixtures with forms I and III and saccharin. PCA used for phase classification; PLS for their quantification. Pre-processing with the SNIP (sensitive nonlinear iterative peak) clipping algorithm was effective for background removal, enhancing the informative peaks [163].	[164]
Diclofenac:ranitidine (2:1 co-crystal)	UV/MCR-ALS	Functional characterization of the co-crystal with an on-line UV/MCR-ALS approach to dissolution monitoring. It was demonstrated that co-crystallization enhances the solubility of diclofenac.	[165]
Furosemide:nicotinamide (1:1 co-crystal)	NIR/PCA	Batch statistical process monitoring was used to create control charts to perceive the process trajectory and define control limits.	[166]
Ibuprofen:nicotinamide (1:1 co-crystal)	Raman/PLS	For co-crystal yield and purity after synthesis. Raman/PLS (mean errors < 5% for all components) was superior to ATR-MIR, DSC and XRPD. MIR and DSC were unsuitable for solving the ternary mixture; XRPD only quantified satisfactorily the co-formers.	[167]
Ibuprofen:nicotinamide carbamazepine:nicotinamide (1:1 co-crystals)	NIR/PLS	Prediction of co-crystal concentration in powder mixtures of the co-crystals and the corresponding APIs.	[168]
Indomethacin:tryptophan Furosemide:tryptophan (co-amorphous)	ssNMR/PCA	Study of ease of amorphization under ball-milling and mechanism of co-amorphization. The degree of amorphization was quantified from the PCA scores of the normalized [169] mean centered ^{13}C ssNMR spectra.	[170]
Indomethacin:saccharin Indomethacin:aspartic acid (1:1 co-crystals)	NIR/PCA	Successful use of NIR/PCA was possible through the inclusion of a set of reference mixtures of parent and guest molecules, representing possible solid-state outcomes from the co-crystal screening. Pre-treatment: SNV baseline correction and scaling.	[171]
Naproxen:indomethacin (co-amorphous)	XRPD/PLS	Simultaneous determination of up to four solid state fractions. RMSEP = 5.57% for the co-amorphous molar fraction. Co-amorphization is employed to stabilize amorphous phases.	[172]

complexity of the data, the most successful applications of vibrational spectroscopic imaging require the combination of spectral, chemometrics, and digital image analysis.

The data structure of these techniques is three dimensional, with two spatial and one wavelength dimension. Behind each image pixel, which exhibits the spatial information, there is spectral information, where the scanned wavelengths can be used as variables and their intensity as responses.

Not all techniques are equally suitable for solving a given problem. For example, the theoretical spatial resolution, which is diffraction-limited, is a weakness of THz spectroscopy, varying with the wavelength in the case of MIR, whereas it depends on the laser used in the case of Raman chemical imaging.

After an appropriate chemometrics treatment (MCR-ALS, PLS, etc.) the image display is changed and new colors are assigned, to exhibit the relative distribution of the components in the sample. This information can be used for component identity assignment and for quantification purposes, according to the intensity of such colors [174]. Recently, time series hyperspectral chemical imaging has surged as a sophisticated evolution that offers the possibility of more comprehensively understanding complicated systems or their process dynamics, by adding the temporal dimension and generating a hyper-cube of data [175].

The main advantages of these techniques, when used for the characterization of pharmaceutical solids, are that they require minimal sample preparation and do not use contaminating solvents, being non-destructive and also non-invasive. However, they usually lack sensitivity and the acquired data are still complex to analyze for the purpose of extracting numerical information with high degree of accuracy and precision. Therefore, they may be time consuming and also may face problems of reproducibility. Typical sources of error in the quantitative analysis by vibrational spectroscopies-mediated mapping include sample packing and its positioning.

The amount and complexity of the data generated by such techniques turn their association with chemometrics methods almost natural [176]. Hence, univariate analyses in this area became currently rather infrequent [177] and the advances regarding the development of methods for fast analysis of multivariate data are becoming rapid and widely adopted within the pharmaceutical field. In order to further develop chemical imaging into a fully applicable PAT, there is a need to increase the data processing speed for enhanced process controllability.

Besides increasing computer velocities, minimizing the number of data processing steps is a much-sought alternative. The repertory of the most relevant chemometrics tools for image analysis [178,179] and the impact of coupling chemical imaging and chemometrics on PAT [180,181] have been reviewed.

11.1. Raman-mapping

Raman spectroscopy may be implemented through a microscope; this technique yields fine scale axial and lateral chemical maps. Both univariate and multivariate methods, applied to Raman spectral data, can be used to obtain the Raman distribution maps. Originally, spatially resolved Raman data were analyzed with univariate techniques, which require the analysis of a constituent-specific spectral band, free of interference from other constituents in the system; however, increasing sample complexity has moved attention toward multivariate approaches [182].

The molecular structure of many drugs turns Raman spectroscopy into a fit for the purpose approach to interrogate pharmaceutical systems, especially tablets and solid dispersions. Hence, during the last decade, Raman imaging has evolved to become a highly useful tool for the analysis of pharmaceutical relevant samples [9,183] enabling the assessment of the quantitative distribution and solid state characteristics of APIs in pharmaceutical formulations [184,185], and even to detect low doses [186] and counterfeit medicines [187].

The combination of Raman chemical imaging and chemometrics strategies to solve pharmaceutically relevant problems has been reviewed [9,188]. Some pharmaceutical applications of chemometrics-assisted Raman-Mapping are found in Table 11.

11.2. NIR-mapping

NIR-Mapping is a powerful technique, which is widely used in the pharmaceutical industry for detecting counterfeit products, quan-

Table 11
Selected pharmaceutical applications of chemometrics-assisted Raman-Mapping.

Drug	Chemometrics Algorithm	Observations	Refs.
Acetaminophen (Forms I, II, III)	PCA MCR-ALS	The crystallization of the amorphous drug in both covered and uncovered geometries was studied. Surface crystallization prevails in uncovered samples, leading to forms I and II, whereas in covered samples bulk crystallization dominates and leads to form III.	[190]
Carbamazepine	MCR-ALS	Used to describe polymorphic transformations. Multiset MCR-ALS analysis provided global (image) and local (pixel) process profiles. Spatial image information discriminates compounds with identical process profiles.	[191]
Drug-Cyclodextrin	MCR-ALS	Analysis of API-cyclodextrin interactions in formulations and sample heterogeneity were studied.	[192]
Indomethacin (α, γ and AM)	PLS	The distribution of the different solid forms loaded into porous silica particles to enhance drug dissolution, and analysis of the heterogeneity of recrystallized drug samples were investigated. PCA revealed the presence of non-modelled forms. Fluorescence and sample burning are drawbacks of the method.	[193]
Piracetam	PLS	Proline used as excipient. A five-steps robust and accurate methodology for low-content (<0.1%) quantification of the drug was developed. The method has LOD comparable to HPLC and its prediction accuracy is ∼2.4% for a 0.05−1.0% concentration range.	[194]
Verapamil.HCl	MCR-ALS	The spatial distribution of the drug in extruded tablets was examined. The spatial homogeneity of two formulations was compared and differences assignable to the manufacturing process were identified.	[195]

tifying polymorphs, and examining the components distribution in complex matrices, such as pharmaceutical dosage forms. Although the NIR-mapping technology cannot overcome all the drawbacks of chemical imaging, it can achieve their reduction and get high S/N ratios by spectral averaging at the expense of loss of spatial information. Signals are usually pre-processed in order to enable the most efficient extraction of the relevant information [189]. Examples of NIR-Mapping cases under chemometrics-assistance are found in Table 12. The chemical structures of the active pharmaceutical ingredients examined in these studies are shown in Fig. 9.

11.3. Terahertz spectroscopy-mapping

Terahertz time-domain spectroscopy is a non-destructive spectroscopy, useful for imaging, operating in the 0.1–10 THz range. Coupled to chemometrics, the method can detect structural defects in component distribution such as pharmaceutical products' coatings. It provides broadband spectral signatures for dielectric materials and fine time resolution [196].

Terahertz pulsed imaging (TPI) technology is gaining acceptance in 3D chemical imaging and tablet coating thickness measurement. However, 3D terahertz chemical imaging is still in its infancy and needs much further development. The technique has many potential advantages over NIR-CI; especially, spectral data can be acquired from within the sample without destroying it [197,198]. In principle, with the aid of suitable chemometric models it would allow visualizing the distribution of the API within the matrix before dissolution testing.

Analogously, it could provide information regarding the degradation rate and spatial distribution of the degradants during stability testing and help solving pharmaceutical solid state characterization problems [210]. Despite the power it has gained thanks to its coupling to computational techniques, the technique is still maturing and has great potential for solving pharmaceutical solid state problems.

11.4. Mass spectrometry chemical imaging

ToF-SIMS is a powerful characterization technique. When used as a spectroscopy, its high surface sensitivity enables identification of a specific molecule's orientation at individual crystal faces, whereas its surface chemical imaging capabilities may provide insights into drug distribution within solid dosage forms, being complementary with existing technologies, and offering significant improvements in spatial and elemental resolution.

A significant advantage of ToF-SIMS imaging is the acquisition of full mass spectra, which enables the study of any fragment within the mass spectrum, or the extraction of a spectrum from any region of interest from an image.

Table 12
Selected pharmaceutical applications of chemometrics-assisted NIR-Mapping.

Drug	Chemometrics Algorithm	Observations	Ref.
Acetaminophen	PLS, PCA	Component distribution in blends with microcrystalline cellulose and lactose MH.	[199]
Carbamazepine (Forms I, III)	MCR-ALS PARAFAC	Time-dependent changes in the distribution of carbamazepine forms on the tablet surface upon transformation of form III to form I, by heating the samples was studied. by NIR-CI coupled to multi-way methods (MCR or PARAFAC). PARAFAC furnished global system information being unable to resolve the spectral components, because lack or trilinearity. MCR provided information about the dynamic process, generating distribution maps for each acquisition time.	[200]
Diclofenac Na	CLS	For a dissolution release study of pellets. Revealed distribution of the formulation components in the coating and inner layers. Complementary to SEM, which furnishes morphological and physical information.	[201]
Fexofenadine.HCl (Forms I, II)	PLS MCR-ALS	Analysis of polymorphic distribution on a tablet surface and their quantification. MCR-ALS quantified the forms (RMSEP < 6% *w/w*) efficiently and generated distribution maps. The PLS model exhibited better recovery of the concentrations. Pre-process: SNV.	[202]
Furosemide (Forms I, II, III)	PLS	Results of solid form quantification in ternary powder mixtures were compared with PLS coupled to NIR, Raman and ATR-MIR spectroscopy. NIR-CI, NIR, and Raman are suited for the purpose; however, ATR-MIR is less appropriate for an accurate quantification because of the pressure-dependent conversion of Form II to Form I.	[203]
Indomethacin (AM → crystalline)	MCR-ALS	Monitoring of the crystallization of the amorphous in dispersion of PVP, by imaging system. PARAFAC and PARAFAC2 were unable to solve the system due to the lack of strict trilinearity.	[204]
Nimodipine (Forms I and II)	PLS	The drug crystallized in formulations stored at 15 °C and 25 °C, contained significant proportion of form I. Formulations stored at 25 °C/60% RH were primarily modification II. Water content of the formulations is implied.	[205]
Nimodipine	PLS, PCR	Model performed form discrimination. Corroborated predictions made by other chemometric methods on form distribution in different binary mixtures resulting from recrystallization from co-solvent formulations.	[206]
Piroxicam (MH → AH)	PARAFAC2	Monitoring of the dehydration of the monohydrate in the surface if tablets with lactose, by imaging system. MCR-ALS and PARAFAC were unable to solve the system due to the lack of strict trilinearity of the pixels among the unfolded series NIR-images and due to rotational ambiguity, respectively,	[204]
Piroxicam (Forms I, II)	PLS	Form distribution in pharmaceutical formulations. RMSEP < 4% *w/w* for both entities.	[207]
Piroxicam MH	MCR-ALS	The dynamics of the heat-mediated (23–120 °C) solid-state transformations were studied in tablets containing PVP and lactose (MH or AH). The dehydration of piroxicam and lactose could be mapped separately despite being transformed simultaneously (80–120 °C). Results well correlated with thermogravimetric analysis. Lack of chemical selectivity in the pixels made PCA unreliable; the obtained PCs describe a mixed combination of effects in the sample, hindering the pure chemical interpretation of the images.	[208]
Tacrolimus (AM and crystalline)	PLS2	Pre-process: Transformation of reflectance to absorbance, which was masked, truncated, and normalized by mean centering and scaling to unit variance by spectrum. Libraries were built for each set of binary samples. PLS2 fitting was employed to obtain the PLS concentration scores and images.	[209]

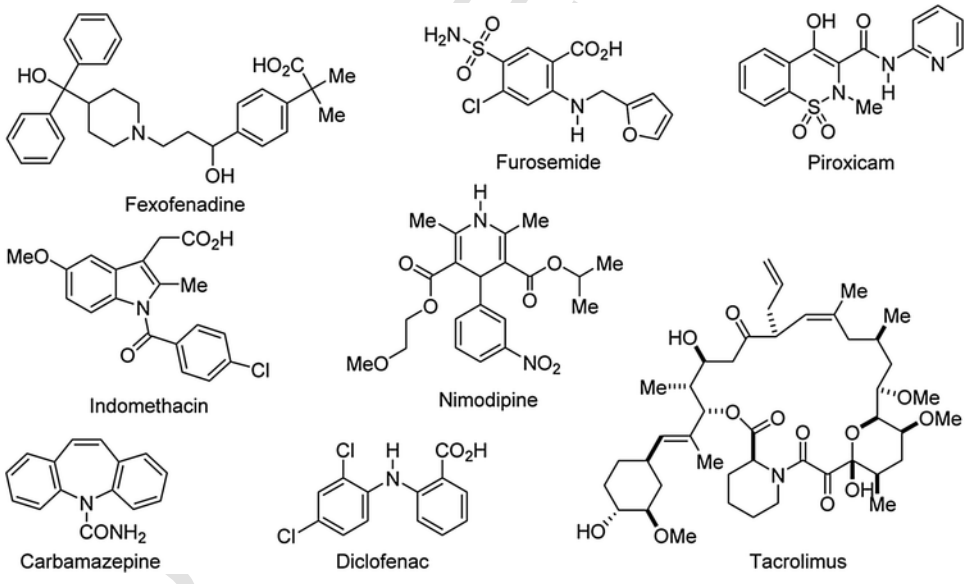

Fig. 9. Chemical structures of some of the compounds employed in the development of NIR-mapping studies.

Calcium carbonate is a commonly used excipient, also useful as a biomaterial with potential in biomineralization. The solid has three polymorphs, the surfaces of which were differentiated by time-of-flight secondary ion mass spectroscopy (ToF-SIMS) coupled to PCA [211]. For analysis, the intensity of each selected peak was normalized to the total intensity of all the selected peaks to eliminate

systematic differences between the spectra and spectra were mean-centered to reduce data scatter. The polymorphs could be clustered into three different groups by PCA scores.

Since ToF-SIMS is still under vigorous development, it is likely that future improvements will make possible to achieve image resolution of a few nanometers, enabling to visualize drug distribution within pharmaceutical nanoparticles. Likewise, the association with PCA will further enhance the understanding of the relationship between spectral intensities and molecular orientation of macromolecular drugs.

In the MALDI-MSI technique, the image is based on the characteristic *m*/*z* ratios of the detected molecules. Its main advantages are mass accuracy, mass resolution and spatial resolution. The method is capable of imaging small and large molecules, making it a very versatile technique for pharmaceutical drugs. Chemometric methods turn easier the management of the high volume of data generated.

Hence, MALDI–MSI was coupled to PCA, ICA, MCR–ALS and NMF (non-negative matrix factorization) in order to extract spatial and spectral information about perindopril tablets [212]. The statistical analysis comprised three steps. The first one was pre-processing; this was followed by estimation of the number of statistical components, where PCA gave the best performance and a final stage of multivariate statistical analysis. In this case, ICA was able to extract the appropriate contributions of the components in homogeneous and heterogeneous datasets. On the contrary, NMF and MCR–ALS were less accurate in obtaining the right contribution in a homogeneous sample, but they were better at distinguishing the semi-quantitative information in a heterogeneous MALDI dataset.

11.5. X-ray image analysis

The dissolution behaviour of drugs from capsules depends on a number of factors, including the distribution of the powder within the capsules. The group of Gosselin investigated how to distinguish un-agglomerated from agglomerated powders in capsules by X-ray image analysis. Capsules were identified a priori based on human opinion and expertise before being used to train SIMCA models. The chemometric models were based on the notion that X-ray images of agglomerated and un-agglomerated capsules produce different, and characteristic, grayscale intensity histograms. The results showed that X-ray imaging can automatically detect and classify powder agglomerates within pharmaceutical capsules, thus reducing reliance on operator depending inspection while increasing the online potential of capsule imaging [213].

12. Process analytical technology

PAT is one of the cornerstones of the Quality by design (QbD) paradigm. These are based on the concept that the thorough understanding of a process minimizes the risk of delivering a poor-quality product [214]. A process is considered well understood when it is designed to consistently meet product critical quality attributes, when its critical sources of variation are identified and controlled, and when it is continually monitored and updated to allow for uniform quality. On-line analytics aims to ensure consistent product quality along the manufacturing process and decrease the burden of finished product testing. It also may reduce or eliminate the need of batch reworking, saving time and increasing manufacturing efficiency.

PAT has been emphasized as one of key elements for the full implementation of the quality by design paradigm in the pharmaceutical area. Among the different spectroscopies, NIR [215] and Raman are particularly suited for PAT applications. These methods can be implemented in-line, are non-destructive, have short measurement times and proceed without the need for sample preparation. Especially when combined with traditional and less frequently used chemometric tools [216], they also have the ability to differentiate among polymorphs and pseudo-polymorphs, being amenable for qualitative and quantitative analysis of their mixtures. The increased emphasis on process monitoring, will inevitably mean that quantitative analysis in the solid state by these spectroscopies will increase.

The use of PAT strategies, including the aspects related to the association of chemometrics methods to aid process interpretation has been recently reviewed [217]. A few examples on the application of chemometrics tools to PAT are mentioned below.

The crystallization of etoricoxib was optimized and controlled by seeding with the desired polymorph at a moderate supersaturation condition. To enhance the process robustness, a NIR/PCA approach was used to obtain a discriminant method able to detect the presence of solids produced by premature crystallization [218]. Once a spectrum was qualified as that of a clear solution, the concentration of the drug was calculated by a NIR/PLS model. The method was applied at a pilot plant level, demonstrating its capability of detecting the presence of solids produced by premature crystallization before seeding.

In another example, the crystallization process of indomethacin following solvent to anti-solvent and anti-solvent to solvent schemes was monitored using an in-line NIR/PCA method [219]. Integration of the PCA results with off-line characterization (SEM, XRD, DSC) enabled the elucidation of the crystallization process under each scheme as composed by three distinct consecutive steps. The high reproducibility of the PCA plots ensured proper and real time process control.

An in-line NIR/PCA strategy was employed to monitor the phase transformations of erythromycin DH in pellets during a miniaturized fluid bed drying process. Transformation to erythromycin DH was observed at 45 and 60 °C, at a moisture content 1.4% *w/w* [220].

A Raman/PLS calibration strategy was developed for the continuous monitoring of solvent-mediated phase transition of citric acid, estimating in-line the overall solid concentration in suspension, the composition of the solid phase and supersaturation [221].

NIR/PCA was employed to monitor the freeze-drying process of a multicomponent formulation [222]. On the other hand, improved process understanding of fluid bed granulation was achieved by performing multivariate (PCA and PLS) analysis of data [223] provided by an in-line particle size analyzer [224] based on a spatial filtering technique, that converts light obscuration signals from individual particles into size information [225].

The polymer–drug solid-state behaviour and molecular interactions during hot-melt extrusion of metoprolol tartrate were evaluated by a NIR/PLS assembly for an in-line polymer–drug solid-state characterization system. The NIR spectra indicated the presence in the extrudates of amorphous drug and hydrogen bonds between the drug and polymeric matrix. The amount of the amorphous phase increased with the extrudate temperature, and at room temperature loss of hydrogen bond interactions between the drug and the polymer take place, with separation of polymer and drug phases, while metoprolol recrystallizes [226].

Heat and mass transfer during fluidization creates the risk for solid-state changes, which may impact on the therapeutic and manufacturing behaviour of pharmaceutically relevant materials. In-line NIR and Raman spectroscopies were coupled to PLS regression and used to monitor and quantify the solid-state form of theophylline, as a model substance, in granules [227]. The conversion of the drug from the MH to the AH form was assessed in real time at 328 and 333 °K,

and absolute humidity and pressure measurements indicated that dehydration had occurred.

A specially designed probe tip was designed to allow for robust in-situ spectral acquisition, and an in-line NIR/PLS approaches were described for monitoring of pharmaceutical powder moisture in fluidized-bed dryers [228,229]. The method proved to be as efficient as traditional off-line analyses; however, since the moisture content of the powder is known in real-time, the risk of over-drying the product is reduced.

NIR and Raman spectroscopies, were also coupled to PCR and PLS to allow the rapid and accurate determination of polymorphic changes in extended release formulations of theophylline. Data pre-processing included MSC, SNV and second derivative (D") [230].

NIR spectroscopy also proved able to differentiate anhydrous theophylline from its hydrate [231]. A NIR/PLS approach enabled to measure in-line and in situ the exact composition of mixtures of the different forms of theophylline with water during the process of drying theophylline monohydrate in a stirred bed vacuum contact drying [232]. This approach allowed studying the influence of operating parameters (temperature and water activity) on the kinetics of solid state transformations. It was shown that the dehydration proceeded by formation of a metastable anhydrate, yielding later the stable form [233].

The mechanism of the solvent (ACN-H_2O)-mediated transformation of sulfamerazine Form I into its enantiotropic Form II was studied by an in-line NIR-chemometrics approach. PCA was employed to detect the presence of solids due to premature crystallization, whereas a PLS calibration was applied to quantify in real time the form transformation; the end-point was determined by clear differentiation between crystal forms based on PLS analysis [234].

Acoustic chemometrics is an emerging method for on-line process monitoring, which is rapidly evolving into a proven PAT technology [235]. It is based on the analysis of system vibrations, usually generated by manufacturing processes or transportation flows. It can be applied for the quantitative analysis of constituents during process monitoring (composition, mixing fractions, mixing progress) and for the physical characterization of the state of the process (moisture, density, particle size, temperature, flow) and equipment.

Sound in the ultrasonic range (20–1000 kHz) emitted during high-shear granulation was analyzed using multivariate techniques [236] and correlated with variations in the physical properties of the obtained granules and the evolution of acoustic emissions taking place during their formation. The multivariate model was capable of predicting the particle size distribution of the granules with prediction error <2%; it also proved useful to predict their moisture content (RMSEP = 1.9%) [237]. On the other hand, process monitoring and the end-point of the heated fluidized bed drying of silica gel was investigated by PLS analysis of acoustic data [238]. Further examples are summarized in Table 13.

13. Conclusions and perspectives

The main objective of this review was to examine the application of chemometrics methods to solve different aspects of the most relevant quali- and quantitative issues of pharmaceutical interest, which involve bulk drugs and excipients as well as their drug products, at the solid state. Because of its importance and special significance to the pharmaceutical industry, the topic of structural polymorphism was specifically excluded to enable its more in-depth discussion in a separate companion review. Without being fully comprehensive, em-

Table 13
Examples on the association of vibrational spectroscopies and chemometrics to solve pharmaceutically relevant problems, complying with the PAT initiative.

Drug	Method	Scope and observations	Ref.
Acetaminophen	NIR/PLS	Monitoring of the powder density. Pre-process: SNV, D', D".	[239]
Acetaminophen	NIR-CI/PLS	Macropixel (pixel cluster) analysis used as a measure of image heterogeneity within the chemical images. Provides quantitative information about the heterogeneity of pharmaceutical products.	[240]
Cetrimonium bromide	Process variables/ DOE PCA PLS	DOE was used to evaluate the interactions and effects of design factors (water amount, wet massing time and lubrication time), on response variables (blend flow, compressibility and tablet dissolution). PCA was employed to examine both batch/sample and variable relationships. PLS was used to predict dissolution profile using 70 process variables (granulation, blending, compression, particle size/distribution, bulk/tapped density, hardness, dissolution and others).	[241]
Cyclosporine A	NIR/PLS	Assessment of the concentration of a phospholipid within the product during liposome formation. Diffuse reflectance mode and spectral imaging techniques were employed.	[242]
Diltiazem.HCl	Raman/SIMCA	Monitoring of blend homogeneity and in . A fiber optical immersion probe was used.	[243]
Excipients	NIR/PLS	Monitoring of powder bulk density and drug concentration at a pilot plant level.	[244]
Ibuprofen	NIR/PLS, PCR, MLR	Monitoring the contents of all components, including the API during the mixing process. Content. Transflectance mode was used. Pre-process: Savitsky-Golay, D'. PLS exhibited the best performance.	[245]
Indomethacin	NIR/PCA NIR/PLS	Diffuse reflectance and transmission modes used during HME to monitor drug transformation of the solid solution (PCA) and API concentration (PLS). Pre-process: SNV.	[246]
Itraconazole	Raman/PLS	Determine API concentration during an HME process. Pre-process: Baseline correction, Whittaker filter and MC.	[247]
Metoprolol tartrate	NIR/PLS NIR/PCA	Drug quantification during HME.	[226]
Mosapride citrate·$2H_2O$	NIR/PLS	Monitoring drug content, moisture, compression force, mean particle size and tablet hardness during tablet production. Diffuse reflectance mode was used. Pre-process: MSC, D'; SNV, D', normalization, D".	[248]
Multi-APIs	NIR/PLS	Monitoring of water content range (2–13% *w/w*) during the entire process. Diffuse reflectance mode was used. Pre-process: Savitzky-Golay smoothing, D', D".	[249]
Phenylpropanolamine.HCl	Process DOE	Studies related to HME process of the drug. Examination of die temperature, shear rate, die length, drug particle size and drug release profile. Dull-factorial and central composite designs were employed, together with response surface methodology.	[250]

Table 13 (*Continued*)

Drug	Method	Scope and observations	Ref.
Spironolactone	Process DOE	Statistical tools and a factorial experimental design employed to optimize HME parameters.	[251]
Sucrose	NIR/PLS	Monitoring of the moisture content during freeze-drying. A fiber optics probe was employed. Pre-process: D' and SNV	[252]
Tablets	NIR/PLS	Monitoring of content uniformity of the product. Transmission mode was employed.	[253]
Theophylline	NIR/PLS	Monitoring of water content upon drying. Diffuse reflectance mode was used. The API has polymorphism. Pre-process: MC and D".	[233]
Theophylline	NIR/PCA NIR/PLS	Monitoring of moisture content, tapped and bulk density. A fiber optic diffuse reflectance probe was used.	[254]
Theophylline	NIR-CI/PCA NIR-CI/PLS	Monitoring of drug flowability. Several PCA and PLS models were developed using a multivariate data analysis software package. Noncontact Raman fiber optics.	[254]
Tolmetin and acetaminophen	Process/PLS	Determination of roller compaction process, ribbon porosity, post-milled particle size and tablet tensile strength. Pre-process: SNV.	[255]
Undisclosed API	NIR/PLS	API content in non-coated tablets. Transmission and reflection modes were used.	[256]
Undisclosed API	DOE NIR/PCA Raman/PCA	Mixture design for optimization of a pharmaceutical tablet formulation. Spectroscopy/PCA employed for determination of particle size distribution.	[256]
Warfarin	Raman/PLS	Monitoring of content uniformity. Pre-process: SNV.	[257]

phasis has been placed in compiling mainly the most significant advances of the last decade.

Chemometrics attempts to ascertain information by the interpretation of data originated in chemical systems. Chemometric methodologies are currently used in all branches of Chemistry, where they have proved very useful, and sometimes even necessary to provide amazingly simple solutions to otherwise difficult analytical problems, bordering impossibility. In recent times, the arsenal of multivariate methods has evolved into a powerful toolbox of alternatives for separating the signals of individual components in complex pharmaceutical mixtures and for extracting useful information that resides within the data generated by different analytical methodologies, enabling the analysis of these systems for understanding the systems and their evolution or just for quality control purposes.

Different multivariate exploration/classification, regression and resolution methods have been applied to various challenging chemical problems of pharmaceutical interest, including the identification of specific components in mixtures containing many ingredients, which sometimes comprise unknown chemicals, such in cases of adulteration.

Other problems were those associated with the quantification of a specific chemical entity often widely interfered by the many other constituents of its matrix, without physical separation of the analytes from the matrix and essentially without sample pre-treatment, and the comprehension of evolutive phenomena, including drug crystallization to obtain the correct solid form and particle size, pharmaceutical (un)stability and drug degradation under different stimuli, pharmaceutical dissolution and drug-drug or drug-excipient interactions, among others.

Established algorithms, such as PCA, PLS and multivariate curve resolution with alternating least squares (MCR-ALS), have become three of the most popular chemometrics methodologies of choice to solve the ample variety of analytical problems associated to the solid state of pharmaceutical principles.

Many significant advances have taken place during the last two decades. Some of them comprised significant improvements in analytical instrumentation, especially in terms of equipment costs and new technologies, as well as better sensitivity and signal to noise ratio of the determinations; also, the advent of faster and more powerful computers, the availability of suitable software. No less important was the popularization of chemometrics as a still young discipline and a better understanding of the inner-works of the multivariate black box tools, which helped to bridge the gap between chemical analysts and chemometricians.

The numerous examples discussed in this review summarized the state of the art, and made clear that chemometrics became and will keep being a critical aid for the quali- and quantitative characterization of static and dynamic pharmaceutical systems at the solid state. Hence, the marriage between chemometrics and analytical chemistry, especially vibrational, X-ray and solid state nuclear magnetic resonance spectroscopies, may serve as a scaffold to undertake more daring problems and to advance our understanding of relevant pharmaceutical materials, systems and processes. After witnessing the current impact of these advances, the future looks very exciting; it can be envisioned that the next decade will bring a more generalized use of chemometrics methods by the academic and industrial community of analytical chemists and they will provide novel and imaginative ways of their application to more demanding pharmaceutical problems.

Acknowledgements

The authors gratefully acknowledge CONICET (PIPN° 2012-0471) and ANPCyT (PICT 2014-0455) for financial support. NLC is thankful to CONICET for her fellowship.

References

[1] S. Trefi, V. Gilard, S. Balayssac, M. Malet-Martino, R. Martino, The usefulness of 2D DOSY and 3D DOSY-COSY 1H NMR for mixture analysis: application to genuine and fake formulations of sildenafil (Viagra), Magn. Reson. Chem. 47 (2009) S163–S173.
[2] A.J. Charlton, W.H. Farrington, P. Brereton, Application of 1H NMR and multivariate statistics for screening complex mixtures: quality control and authenticity of instant coffee, J. Agric. Food Chem. 50 (2002) 3098–3103.
[3] S.N.A. Bukhari, N.S. Hwei, I. Jantan, Recent advances in solid-state analysis of pharmaceuticals, Open Pharm. Sci. J. 2 (2015) 13–20.
[4] G.A. Stephenson, R.A. Forbes, S.M. Reutzel-Edens, Characterization of the solid state: quantitative issues, Adv. Drug Deliv. Rev. 48 (2001) 67–90.
[5] S. Wartewiga, R.H.H. Neubert, Pharmaceutical applications of Mid-IR and Raman spectroscopy, Adv. Drug Deliv. Rev. 57 (2005) 1144–1170.
[6] A.W. Newman, S.R. Byrn, Solid-state analysis of the active pharmaceutical ingredient in drug products, Drug Discov. Today 8 (2003) 898–905.
[7] A. Paudel, D. Raijada, J. Rantanen, Raman spectroscopy in pharmaceutical product design, Adv. Drug Deliv. Rev. 89 (2015) 3–20.
[8] B. Swarbrick, The current state of near infrared spectroscopy application in the pharmaceutical industry, J. Near Infrared Spectrosc. 22 (2014) 153–156.
[9] K.C. Gordon, C.M. McGoverin, Raman mapping of pharmaceuticals, Int. J. Pharm. 417 (2011) 151–162.
[10] Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, N. Jent, A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies, J. Pharm. Biomed. Anal. 44 (2007) 683–700.

[11] A.B. Eldin, O.A. Ismaiel, W.E. Hassan, A.A. Shalaby, Green analytical chemistry: opportunities for pharmaceutical quality control, J. Anal. Chem. 71 (2016) 861–871.
[12] S. Matero, Chemometrics Methods in Pharmaceutical Tablet Development and Manufacturing Unit Operations. Publications of the University of Eastern Finland. Dissertations in Health Sciences, Ph.D. Dissertation, 2010. (Accessed April 26, 2017) http://epublications.uef.fi/pub/urn_isbn_978-952-61-0143-9/urn_isbn_978-952-61-0143-9. pdf.
[13] R.G. Brereton (Ed.), Applied Chemometrics for Scientists, Wiley Chichester, UK, 2007.
[14] C. Coscollà, S. Navarro-Olivares, P. Martí, V. Yusà, Application of the experimental design of experiments (DoE) for the determination of organotin compounds in water samples using HS-SPME and GC–MS/MS, Talanta 119 (2014) 544–552.
[15] J.R. Wagner Jr., M.I.M. Eldridge, H.F. Giles Jr., Design of experiments, Extrusion. The Definitive Processing Guide and Handbook. A Volume in Plastics Design Library, 2nd ed., Elsevier Oxford, England, 2014291–308.
[16] G. Derringer, R. Suich, Simultaneous optimization of several response variables, J. Qual. Technol. 12 (1980) 214–219.
[17] A. Rinnan, F.V.D. Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near infrared spectra, Trends Anal. Chem. 28 (2009) 1201–1222.
[18] S. Roussel, S. Preys, F. Chauchard, J. Lallemand, Multivariate data analysis (chemometrics), in: C.P. O'Donnell, C. Fagan, P.J. Cullen (Eds.), Process Analytical Technology for the Food Industry, Springer, New York, USA, 2014, pp. 7–60. (Chapter 2).
[19] J. Sun, Statistical analysis of NIR data: data pretreatment, J. Chemometr. 11 (1997) 525–532.
[20] F. Allegrini, A.C. Olivieri, An integrated approach to the simultaneous selection of variables, mathematical pre-processing and calibration samples in partial least-squares multivariate calibration, Talanta 115 (2013) 755–760.
[21] H. Abdi, L.J. Williams, Principal component analysis, WIRE Comput. Stat. 2 (2010) 433–459.
[22] F.A.N. Alhija, Factor analysis: an overview and some contemporary advances, in: P. Peterson, E. Baker, B. McGaw (Eds.), International Encyclopedia of Education, 3rd ed., Elsevier, Amsterdam, Netherlands, 2010, pp. 162–170.
[23] I. Lee, J. Yang, Common clustering algorithms, in: S.D. Brown, R. Tauler, B. Walczak (Eds.), Comprehensive Chemometrics, Elsevier Oxford, England, 2009, pp. 577–618.
[24] Gaussian mixture models and k-means clustering, in: W. Press, S. Teukolsky, W. Vetterling, B. Flannery (Eds.), Numerical Recipes: The Art of Scientific Computing, Cambridge University Press, New York, USA, 2007.
[25] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybern. 43 (1982) 59–69.
[26] S. Wold, M. Sjostrom, SIMCA. A method for analyzing chemical data in terms of similarity and analogy, in: B.R. Kowalski (Ed.), Chemometrics Theory and Application, American Chemical Society, Washington, DC, USA, 1977, pp. 243–282.
[27] K. Varmuza, P. Filzmoser, Introduction to Multivariate Statistical Analysis in Chemometrics, 1st ed., CRC Press, Boca Raton, USA, 2008.
[28] B.G.M. Vandeginste, D.L. Massart, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics, Part B, Chapter 33, Elsevier, Amsterdam, Netherlands, 1998.
[29] J.F. Hair, R.E. Anderson, R. Tatham, W.C. Black (Eds.), Multivariate Data Analysis, Prentice Hall, New Jersey, USA, 1995.
[30] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. 17 (2003) 166–173.
[31] M. Bylesjo, M. Rantalainen, O. Cloarec, J.K. Nicholson, E. Holmes, J. Trygg, OPLS discriminant analysis: combining the strengths of PLS-DA and SIMCA classification, J. Chemom. 20 (2006) 341–351.
[32] D. Coomans, D.L. Massart, Alternative k-nearest neighbour rules in supervised pattern recognition: part 1. k-Nearest neighbour classification by using alternative voting rules, Anal. Chim. Acta 136 (1982) 15–27.
[33] K. Crammer, Y. Singer, On the algorithmic implementation of multiclass kernel-based vector machines, J. Mach. Learn. Res. 2 (2001) 265–292.
[34] G. Sabina, V. Lozano, W. Rocha, W. Romão, R. Ortiz, R.J. Poppi, Characterization of sildenafil citrate tablets of different sources by near infrared chemical imaging and chemometric tools, J. Pharm. Biomed. Anal. 85 (2013) 207–212.
[35] A.C. Olivieri, Analytical advantages of multivariate data processing. One two, three, infinity?, Anal. Chem. 80 (2008) 5713–5720.
[36] S. Wold, M. Sjostrom, PLS-regression. A basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130.
[37] I.T. Jolliffe, A note on the use of principal components in regression, J. R. Stat. Soc. 31 (1982) 300–303.
[38] D.M. Haaland, E.V. Thomas, Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information, Anal. Chem. 60 (1988) 1193–1202.
[39] F. Marini, R. Bucci, A.L. Magrì, A.D. Magrì, Artificial neural networks in chemometrics: history, examples and perspectives, Microchem. J. 88 (2008) 178–185.
[40] R. Bro, PARAFAC. Tutorial and applications, Intell. Lab. Syst. 38 (1997) 149–171.
[41] R. Tauler, B. Kowalski, Multivariate curve resolution applied to spectral data from multiple runs of an industrial process, Anal. Chem. 65 (1993) 2040–2047.
[42] K. Pöllänen, A. Häkkinen, S.-P. Reinikainen, J. Rantanen, P. Minkkinen, Dynamic PCA-based MSPC charts for nucleation prediction in batch cooling crystallization processes, Chemometr. Intell. Lab. Syst. 84 (2006) 126–133.
[43] T. Togkalidou, H.-H. Tung, Y. Sun, A. Andrews, R.D. Braatz, Solution concentration prediction for pharmaceutical crystallization processes using robust chemometrics and ATR MIR spectroscopy, Org. Proc. Res. Dev. 6 (2002) 317–322.
[44] Z.K. Nagy, G. Fevotte, H. Kramer, L.L. Simon, Recent advances in the monitoring, modelling and control of crystallization systems, Chem. Eng. Res. Des. 91 (2013) 1903–1922.
[45] M. Allesø, F. Van Den Berg, C. Cornett, F.S. Jørgensen, B. Halling-Sørensen, H. Lopez De Diego, L. Hovgaard, J. Aaltonen, J. Rantanen, Solvent diversity in polymorph screening, J. Pharm. Sci. 97 (2008) 2145–2159.
[46] M. Müller, U. Meier, D. Wieckhusen, R. Beck, S. Pfeffer-Hennig, R. Schneeberger, Process development strategy to ascertain reproducible API polymorph manufacture, Cryst. Growth Des. 6 (2006) 946–954.
[47] Z.-P. Chen, J. Morris, A. Borissova, S. Khan, T. Mahmud, R. Penchev, K.J. Roberts, On-line monitoring of batch cooling crystallization of organic compounds using ATR-MIR spectroscopy coupled with an advanced calibration method, Chemometr. Intell. Lab. Syst. 96 (2009) 49–58.
[48] K. Pöllänen, A. Häkkinen, S.-P. Reinikainen, M. Louhi-Kultanen, L. Nyström, A study on batch cooling crystallization of sulphathiazole process monitoring using ATR-MIR and product characterization by automated image analysis, Chem. Eng. Res. Des. 84 (2006) 47–59.
[49] B. Connolly, T.W. Patapoff, Y.J. Wang, J.M. Moore, T.J. Kamerzell, Vibrational spectroscopy and chemometrics to characterize and quantitate trehalose crystallization, Anal. Biochem. 399 (2010) 48–57.
[50] C. Schaefer, C. Lecomte, D. Clicq, A. Merschaert, E. Norrant, F. Fotiadu, On-line near infrared spectroscopy as a Process Analytical Technology (PAT) tool to control an industrial seeded API crystallization, J. Pharm. Biomed. Anal. 83 (2013) 194–201.
[51] J. Unga, P. Matsson, D. Mahlin, Understanding polymer–lipid solid dispersions – the properties of incorporated lipids govern the crystallisation behaviour of PEG, Int. J. Pharm. 386 (2010) 61–70.
[52] S. Romero-Torres, J.D. Perez-Ramos, K.R. Morris, E.R. Grant, Raman spectroscopic measurement of tablet-to-tablet coating variability, J. Pharm. Biomed. Anal. 38 (2005) 270–274.
[53] Ł. Komsta, J.K. Maurin, Recognition of active ingredients in tablets by chemometric processing of X-ray diffractometric data, Talanta 82 (2010) 850–853.
[54] W. Miltyk, E. Antonowicz, Ł. Komsta, Recognition of tablet content by chemometric processing of differential scanning calorimetry curves—an acetaminophen example, Thermochim. Acta 507–508 (2010) 146–149.
[55] K. Paradowska, M.K. Jamróz, M. Kobyłka, E. Gowin, P. Mączka, R. Skibiński, L. Komsta, Detection of drug active ingredients by chemometric processing of solid-state NMR spectrometry data—the case of acetaminophen, J.AOAC Int. 95 (2012) 704–707.
[56] M.C. Sarraguça, A.V. Cruz, H.R. Amaral, P.C. Costa, J.A. Lopes, Comparison of different chemometric and analytical methods for the prediction of particle size distribution in pharmaceutical powders, Anal. Bioanal. Chem. 399 (2011) 2137–2147.
[57] M.C. Sarraguca, A.V. Cruz, S.O. Soares, H.R. Amaral, P.C. Costa, João A. Lopes, Determination of flow properties of pharmaceutical powders by near infrared spectroscopy, J. Pharm. Biomed. Anal. 52 (2010) 484–492.
[58] M. Otsuka, Y. Mouri, Y. Matsuda, Chemometric evaluation of pharmaceutical properties of antipyrine granules by near-infrared spectroscopy, AAPS PharmSciTech 4 (2003) 334–340.
[59] L.K.H. Bittner, N. Heigl, C.H. Petter, M.F. Noisternig, U.J. Griesser, G.K. Bonn, C.W. Huck, Near-infrared reflection spectroscopy (NIRS) as a successful tool for simultaneous identification and particle size determination of amoxicillin trihydrate, J. Pharm. Biomed. Anal. 54 (2011) 1059–1064.
[60] H. Tanabe, K. Otsuka, M. Otsuka, Theoretical analysis of tablet hardness prediction using chemoinformetric near infrared spectroscopy, Anal. Sci. 23 (2007) 857–862.
[61] K.M. Morisseau, C.T. Rhodes, Near-infrared spectroscopy as a nondestructive alternative to conventional tablet hardness testing, Pharm. Res. 14 (1997) 108–111.

[62] F.J.S. Nieuwmeyer, M. Damen, A. Gerich, F. Rusmini, K. van der Voort, H. Maarschalk, Vromans, granule characterization during fluid bed drying by development of a near infrared method to determine water content and median granule size, Pharm. Res. 24 (2007) 1854–1861.
[63] W.L. Yip, I. Gausemel, S.A. Sande, K. Dyrstad, Strategies for multivariate modeling of moisture content in freeze-dried mannitol-containing products by near-infrared spectroscopy, J. Pharm. Biomed. Anal. 70 (2012) 202–211.
[64] A.D. Román-Ospino, R. Singh, M. Ierapetritou, R. Ramachandran, R. Méndez, C. Ortega-Zuñiga, F.J. Muzzio, R.J. Romañach, Near infrared spectroscopic calibration models for real time monitoring of powder density, Int. J. Pharm. 512 (2016) 61–74.
[65] S.H. Tabasi, R. Fahmy, D. Bensley, C. O'Brien, S.W. Hoag, Quality by design, Part I: application of NIR spectroscopy to monitor tablet manufacturing process, J. Pharm. Sci. 97 (2008) 4040–4051.
[66] S.H. Tabasi, R. Fahmy, D. Bensley, C. O'Brien, S.W. Hoag, Quality by design, Part II: application of NIR spectroscopy to monitor the coating process for a pharmaceutical sustained release product, J. Pharm. Sci. 97 (2008) 4052–4066.
[67] M. Otsuka, I. Yamane, Prediction of tablet hardness based on near infrared spectra of raw mixed powders by chemometrics, J. Pharm. Sci. 95 (2006) 1425–1433.
[68] M. Otsuka, I. Yamane, Prediction of tablet properties based on near infrared spectra of raw mixed powders by chemometrics: scale-up factor of blending and tableting processes, J. Pharm. Sci. 98 (2009) 4296–4305.
[69] S.M. Short, R.P. Cogdill, P.L.D. Wildfong, J.K. Drennen III, C.A. Anderson, A near-infrared spectroscopic investigation of relative density and crushing strength in four-component compacts, J. Pharm. Sci. 98 (2009) 1095–1109.
[70] W.L. Yip, T.C. Soosainather, K. Dyrstad, S.A. Sande, Classification of structurally related commercial contrast media bynear infrared spectroscopy, J. Pharm. Biomed. Anal. 90 (2014) 148–160.
[71] Y. Roggo, K. Degardin, P. Margot, Identification of pharmaceutical tablets by Raman spectroscopy and chemometrics, Talanta 81 (2010) 988–995.
[72] P.-Y. Sacré, L. Netchacovitch, C. De Bleye, P.-F. Chavez, C. Servais, R. Klinkenberg, B. Streel, P. Hubert, E. Ziemons, Thorough characterization of a self-emulsifying drug delivery system with Raman hyperspectral imaging: a case study, Int. J. Pharm. 484 (2015) 85–94.
[73] B.M. Lukasiak, J.C. Duncan, Classification of polymer groups by means of a new polymer testing instrument the identiPol QA, coupled with pattern recognition techniques, Anal. Meth. 2 (2010) 1948–1957.
[74] Y. Ma, P. Huang, D. Hou, J. Cai, Q. Wang, G. Zhang, The spectral resolution of unknown mixture based on THz spectroscopy with self-modeling technique, Chemometr. Intell. Lab. Sys. 150 (2016) 65–73.
[75] Y. Wang, M. Mei, Y. Ni, S. Kokot, Combined NIR/MIR analysis: a novel method for the classification of complex substances such as Illicium verum Hook, F. and its adulterants, Spectrochim. Acta A 130 (2014) 539–545.
[76] H. Cai, R. Cheng, Y. Jin, S. Ding, Z. Chen, Evaluation of Oolong teas using 1H and 13C solid-state NMR sensory analysis, and multivariate statistics, J. Chin. Chem. Soc. 63 (2016) 792–799.
[77] R. Shan, Z. Mao, L. Yin, W. Cai, X. Shao, Discrimination of Chinese patent medicines using near-infrared spectroscopy and principal component accumulation method, Anal. Meth. 6 (2014) 4692–4697.
[78] K. Gorlowska, J. Gorlowska, R. Skibinski, Ł. Komsta, Chemometrics meets homeopathy-an exploratory analysis of infrared spectra of homeopathic granules, J. Pharm. Biomed. Anal. 115 (2015) 36–38.
[79] A. Hertrampf, R.M. Sousa, J.C. Menezes, T. Herdling, Semi-quantitative prediction of a multiple API solid dosage form with a combination of vibrational spectroscopy methods, J. Pharm. Biomed. Anal. 124 (2016) 246–253.
[80] A.B. Eldina, O.A. Ismaiel, W.E. Hassan, A.A. Shalaby, Green analytical chemistry: opportunities for pharmaceutical quality control, J. Anal. Chem. 71 (2016) 861–871.
[81] B. Nagy, A. Farkas, A. Balogh, H. Pataki, B. Vajna, Z.K. Nagy, G. Marosi, Quantification and handling of nonlinearity in Raman micro-spectrometry of pharmaceuticals, J. Pharm. Biomed. Anal. 128 (2016) 236–246.
[82] B. Igne, S. Talwar, H. Feng, J.K. Drennen, C.A. Anderson, Near-infrared spatially resolved spectroscopy for tablet quality determination, J. Pharm. Sci. 104 (2015) 4074–4081.
[83] A. Sánchez-Paternina, A.D. Román-Ospino, M. Martínez, J. Mercado, C. Alonso, R.J. Romañach, Near infrared spectroscopic transmittance measurements for pharmaceutical powder mixtures, J. Pharm. Biomed. Anal. 123 (2016) 120–127.
[84] D. Melucci, D. Monti, M. D'Elia, G. Luciano, Rapid in situ repeatable analysis of drugs in powder form using reflectance near-infrared spectroscopy and multivariate calibration, J. Forens. Sci. 57 (2012) 86–92.
[85] Y. Otsuka, A. Ito, S. Matsumura, M. Takeuchi, S. Pal, H. Tanaka, Quantification of pharmaceutical compounds based on powder X-ray diffraction with chemometrics, Chem. Pharm. Bull. 64 (2016) 1129–1135.
[86] J. Griffen, A. Owen, P. Matousek, Comprehensive quantification of tablets with multiple active pharmaceutical ingredients using transmission Raman spectroscopy—a proof of concept study, J. Pharm. Biomed. Anal. 115 (2015) 277–282.
[87] M.A.M. Silva, M.H. Ferreira, J.W.B. Braga, M.M. Sena, Development and analytical validation of a multivariate calibration method for determination of amoxicillin in suspension formulations by near infrared spectroscopy, Talanta 89 (2012) 342–351.
[88] A.L. Hermes Müller, R.S. Picoloto, M.F. Ferrão, F.E. Barcellos da Silva, E.I. Müller, E.M. de Moraes Flores, Simultaneous diffuse reflectance infrared determination of clavulanic acid and amoxicillin using multivariate calibration techniques, Drug Test. Anal. 4 (2012) 500–506.
[89] A.N. Khan, R.K. Khar, P.V. Ajayakumar, Diffuse reflectance near infrared-chemometric methods development and validation of amoxicillin capsule formulations, J. Pharm. Bioall. Sci. 8 (2016) 152–610.
[90] Y. Li, G. Du, W. Cai, X. Shao, Classification and quantitative analysis of azithromycin tablets by Raman spectroscopy and chemometrics, Am. J. Anal. Chem. 2 (2011) 135–141.
[91] M. Palo, K. Kogermann, N. Genina, D. Fors, J. Peltonen, J. Heinämäki, N. Sandler, Quantification of caffeine and loperamide in printed formulations by infrared spectroscopy, J. Drug Deliv. Sci. Technol. 34 (2016) 60–70.
[92] A.G. Ryder, G.M. O'Connor, T.J. Glynn, Quantitative analysis of cocaine in solid mixtures using Raman spectroscopy and chemometric methods, J.Raman Spectrosc. 31 (2000) 221–227.
[93] L. Rigoni, S. Venti, M. Bevilacqua, R. Bucci, A.D. Magrì, A.L. Magrì, F. Marini, Quantification of the enantiomeric excess of two APIs by means of near infrared spectroscopy and chemometrics, Chemometr. Intell. Lab. Syst. 133 (2014) 149–156.
[94] M.S. Piantavini, F.L.D. Pontes, C.P. Uber, D.P. Stremel, M.M. Sena, R. Pontarolo, Chemometric quality inspection control of pyrantel pamoate, febantel and praziquantel in veterinary tablets by mid infrared spectroscopy, Spectrochim. Acta A 125 (2014) 396–403.
[95] M.H. Ferreira, J.W.B. Braga, M.M. Sena, Development and validation of a chemometric method for direct determination of hydrochlorothiazide in pharmaceutical samples by diffuse reflectance near infrared spectroscopy, Microchem. J. 109 (2013) 158–164.
[96] A. Sparén, M. Hartman, M. Fransson, J. Johansson, O. Svensson, Matrix effects in quantitative assessment of pharmaceutical tablets using transmission Raman and near-infrared (NIR) spectroscopy, Appl. Spectrosc. 69 (2015) 580–589.
[97] M. Alcalà, J. León, J. Ropero, M. Blanco, R.J. Romañach, Analysis of low content drug tablets by transmission near infrared spectroscopy: selection of calibration ranges according to multivariate detection and quantitation limits of PLS models, J. Pharm. Sci. 97 (2008) 5318–5327.
[98] I. Tomuta, L. Rus, R. Iovanov, L.L. Rus, High-throughput NIR-chemometric methods for determination of drug content and pharmaceutical properties of indapamide tablets, J. Pharm. Biomed. Anal. 84 (2013) 285–292.
[99] J. El Haddad, F. de Miollis, J. Bou Sleiman, L. Canioni, P. Mounaix, B. Bousquet, Chemometrics applied to quantitative analysis of ternary mixtures by terahertz spectroscopy, Anal. Chem. 86 (2014) 4927–4933.
[100] I. Tomuta, R. Iovanov, E. Bodoki, L. Vonica, Development and validation of NIR-chemometric methods for chemical and pharmaceutical characterization of meloxicam tablets, Drug Dev. Ind. Pharm. 40 (2014) 549–559.
[101] A.A. Bunaciu, S. Fleschin, H.Y. Aboul-Enein, A new method for a quantitative determination of piroxicam in pharmaceutical formulations using FT-IR spectrometry, Anti-Inflamm. Anti-Aller. Agents Med. Chem. 11 (2012) 262–266.
[102] T. Firkala, A. Farkas, B. Vajna, Z.K. Nagy, G. Pokol, G. Marosi, I.M. Szilágyi, Quantification of low drug concentration in model formulations with multivariate analysis using surface enhanced Raman chemical imaging, J. Pharm. Biomed Anal. 107 (2015) 318–324.
[103] World Health Organization, Definitions of SSFFC Medical Products, http://www.who.int/medicines/regulation/ssffc/definitions/en/. (Accessed (leave a space)on April 20, 2017).
[104] K. Dégardina, Y. Roggo, F. Been, P. Margot, Detection and chemical profiling of medicine counterfeits by Raman spectroscopy and chemometrics, Anal. Chim. Acta 705 (2011) 334–341.
[105] C.M. Gryniewicz-Ruzicka, J.D. Rodriguez, S. Arzhantsev, L.F. Buhse, J.F. Kauffman, Libraries, classifiers, and quantifiers: a comparison of chemometric methods for the analysis of Raman spectra of contaminated pharmaceutical materials, J. Pharm. Biomed. Anal. 61 (2012) 191–198.
[106] S. Neuberger, C. Neusüß, Determination of counterfeit medicines by Raman spectroscopy: systematic study based on a large set of model tablets, J. Pharm. Biomed. Anal. 112 (2015) 70–78.
[107] R.S. Ortiz, K. de Cássia Mariotti, B. Fank, R.P. Limberger, M.J. Anzanello, P. Mayorga, Counterfeit Cialis and Viagra fingerprinting by ATR-MIR spec-

troscopy with chemometry: can the same pharmaceutical powder mixture be used to falsify two medicines?, Forensic Sci. Int. 226 (2013) 282–289.
[108] D. Custers, T. Cauwenbergh, J.L. Bothy, P. Courselle, J.O. De Beer, S. Apers, E. Deconinck, ATR-MIR spectroscopy and chemometrics: an interesting tool to discriminate and characterize counterfeit medicines, J. Pharm. Biomed. Anal. 112 (2015) 181–189.
[109] Q. Gao, Y. Liu, H. Li, H. Chen, Y. Chai, F. Lu, Comparison of several chemometric methods of libraries and classifiers for the analysis of expired drugs based on Raman spectra, J. Pharm. Biomed. Anal. 94 (2014) 58–64.
[110] B. Krakowska, D. Custers, E. Deconinck, M. Daszykowski, Chemometrics and the identification of counterfeit medicines − a review, J. Pharm. Biomed. Anal. 127 (2016) 112–122.
[111] Y.V. Zontov, K.S. Balyklova, A.V. Titova, O.Y. Rodionova, A.L. Pomerantsev, Chemometric aided NIR portable instrument for rapid assessment of medicine quality, J. Pharm. Biomed. Anal. 131 (2016) 87–93.
[112] L. Li, H. Zang, J. Li, D. Chen, T. Li, F. Wang, Identification of anisodamine tablets by Raman and near-infrared spectroscopy with chemometrics, Spectrochim. Acta A 127 (2014) 91–97.
[113] P. de Peinder, M.J. Vredenbregt, T. Visser, D. de Kaste, Detection of Lipitor® counterfeits: a comparison of NIR and Raman spectroscopy in combination with chemometrics, J. Pharm. Biomed. Anal. 47 (2008) 688–694.
[114] L. Feng, W. Xinxin, C. Yifeng, Y. Yongjian, Y. Yinjia, D. Gengli, A novel identification system for counterfeit drugs based on portable Raman spectroscopy, Chemometr. Intell. Lab. Syst. 127 (2013) 63–69.
[115] G. Petrusevski, J. Acevska, G. Stefkov, A. Poceva Panovska, I. Micovski, R. Petkovska, A. Dimitrovska, S. Ugarkovic, Characterization and origin differentiation of morphine derivatives by DSC/TG and FTIR analysis using pattern recognition techniques, J. Therm. Anal. Calorim. 123 (2016) 2561–2571.
[116] J.A. Baird, B. Van Eerdenbrugh, L.S. Taylor, A classification system to assess the crystallization tendency of organic molecules from undercooled melts, J. Pharm. Sci. 99 (2010) 3787–3806.
[117] B. Van Eerdenbrugh, J.A. Baird, L.S. Taylor, Crystallization tendency of active pharmaceutical ingredients following rapid solvent evaporation-classification and comparison with crystallization tendency from undercooled melts, J. Pharm. Sci. 99 (2010) 3826–3838.
[118] D. Mahlin, S. Ponnambalam, M. Heidarian Höckerfelt, C.A. Bergström, Toward in silico prediction of glass-forming ability from molecular structure alone: a screening tool in early drug development, Mol. Pharm. 8 (2011) 498–506.
[119] K. Nurzyńska, J. Booth, C.J. Roberts, J. McCabe, I. Dryden, P.M. Fischer, Long-term amorphous drug stability predictions using easily calculated predicted, and measured parameters, Mol. Pharm. 12 (2015) 3389–3398.
[120] C.A.S. Bergström, C.M. Wassvik, K. Johansson, I. Hubatsch, Poorly soluble marketed drugs display solvation limited solubility, J. Med. Chem. 50 (2007) 5858–5862.
[121] P. Guerrieri, A.C.F. Rumondor, T. Li, L.S. Taylor, Analysis of relationships between solid-state properties, counterion, and developability of pharmaceutical salts, AAPS PharmSciTech 11 (2010) 1212–1222.
[122] C.M. Wassvik, A.G. Holmén, R. Draheim, P. Artursson, C.A.S. Bergström, Molecular characteristics for solid-state limited solubility, J. Med. Chem. 51 (2008) 3035–3039.
[123] Y. Lin, R.P. Cogdill, P.L.D. Wildfong, Informatic calibration of a materials properties database for predictive assessment of mechanically activated disordering potential for small molecule organic solids, J. Pharm. Sci. 98 (2009) 2696–2708.
[124] Y. Hattori, M. Sato, M. Otsuka, Initial dissolution kinetics of cocrystal of carbamazepine with nicotinamide, J. Pharm. Pharmacol. 67 (2015) 1512–1518.
[125] P. Pawar, Y. Wang, G. Keyvan, G. Callegari, A. Cuitino, F. Muzzio, Enabling real time release testing by NIR prediction of dissolution of tablets made by continuous direct compression (CDC), Int. J. Pharm. 512 (2016) 96–107.
[126] N.L. Calvo, R.M. Maggio, T.S. Kaufman, An eco-friendly strategy using on-line monitoring and dilution coupled to a second-order chemometric method, for the construction of dissolution curves of combined pharmaceutical associations, J. Pharm. Biomed. Anal. 89 (2014) 213–220.
[127] R.M. Maggio, M.A. Rivero, T.S. Kaufman, Simultaneous acquisition of the dissolution curves of two active ingredients in a binary pharmaceutical association, employing an on-line circulation system and chemometrics-assistance, J. Pharm. Biomed. Anal. 72 (2013) 51–58.
[128] R.M. Maggio, P.M. Castellano, T.S. Kaufman, A multivariate approach for the simultaneous determination of losartan potassium and hydrochlorothiazide in a combined pharmaceutical tablet formulation, Anal. Bioanal. Chem. 391 (2008) 2949–2955.
[129] S.E. Vignaduzzo, P.M. Castellano, R.M. Maggio, T.S. Kaufman, Application of a chemometric method for the simultaneous determination of acetaminophen and diclofenac in content-uniformity and drug −dissolution studies, Anal. Bioanal. Chem. 382 (2005) 1711–1714.
[130] S.E. Vignaduzzo, R.M. Maggio, P.M. Castellano, T.S. Kaufman, PLS and first derivative of ratio spectra mehtods for determination of hydrochlorothiazide and propranolol hydrochloride in tablets, Anal. Bioanal. Chem. 386 (2006) 2239–2244.
[131] Y. Wang, R.D. Snee, G. Keyvan, F.J. Muzzio, Statistical comparison of dissolution profiles, Drug Dev. Ind. Pharm. 42 (2016) 796–807.
[132] R.M. Maggio, P.M. Castellano, T.S. Kaufman, A new principal component analysis-based approach for testing similarity of drug dissolution profiles, Eur. J. Pharm. Sci. 34 (2008) 66–77.
[133] R.M. Maggio, P.M. Castellano, T.S. Kaufman, PCA-CR analysis of dissolution profiles. A chemometric approach to probe the polymorphic form of the active pharmaceutical ingredient in a drug product, Int. J. Pharm. 378 (2009) 187–193.
[134] R. Ambrus, P. Kocbek, J. Kristl, R. Sibanc, R. Rajko, P. Szabó-Révész, Investigation of preparation parameters to improve the dissolution of poorly water-soluble meloxicam, Int. J. Pharm. 381 (2009) 153–159.
[135] J. Zhang, Y. Ying, B. Pielecha-Safira, E. Bilgili, R. Ramachandran, R. Romañach, R.N. Davé, Z. Iqbal, Raman spectroscopy for in-line and off-line quantification of poorly soluble drugs in strip films, Int. J. Pharm. 475 (2014) 428–437.
[136] J. Hisazumi, T. Suzuki, H. Nakagami, K. Terada, Quantification of pharmaceutical polymorphs and prediction of dissolution rate using theophylline tablet by terahertz spectroscopy, Chem. Pharm. Bull. 59 (2011) 442–446.
[137] S. Matero, J. Pajander, A.-M. Soikkeli, S.-P. Reinikainen, M. Lahtela-Kakkonen, O. Korhonen, J. Ketolainen, A. Poso, Predicting the drug concentration in starch acetate matrix tablets from ATR-FTIR spectra using multi-way methods, Anal. Chim. Acta 595 (2007) 190–197.
[138] P.R. Nassab, R. Rajko, P. Szabó-Révész, Physicochemical characterization of meloxicam–mannitol binary systems, J. Pharm. Biomed. Anal. 41 (2006) 1191–1197.
[139] E. Hernandez, P. Pawar, G. Keyvan, Y. Wang, N. Velez, G. Callegari, A. Cuitino, B. Michniak-Kohn, F.J. Muzzio, R.J. Romanach, Prediction of dissolution profiles by non-destructive near infrared spectroscopy in tablets subjected to different levels of strain, J. Pharm. Biomed. Anal. 117 (2016) 568–576.
[140] A. Elia, M. Cocchi, C. Cottini, D. Riolo, C. Cafiero, R. Bosi, E. Lutero, Multivariate data analysis to assess dry powder inhalers performance from powder properties, Powder Technol. 301 (2016) 830–838.
[141] A. Shayanfar, K. Asadpour-Zeynali, A. Jouyban, Solubility and dissolution rate of a carbamazepine–cinnamic acid cocrystal, J. Mol. Liq. 187 (2013) 171–176.
[142] A. Cutrignelli, A. Trapani, A. Lopedota, M. Franco, D. Mandracchia, N. Denora, V. Laquintana, G. Trapani, Griseofulvin/carrier blends: application of partial least squares (PLS) regression analysis for estimating the factors affecting the dissolution efficiency, AAPS PharmSciTech 12 (2011) 1019–1030.
[143] M. Dumarey, D.J. Goodwin, C. Davison, Multivariate modelling to study the effect of the manufacturing process on the complete tablet dissolution profile, Int. J. Pharm. 486 (2015) 112–120.
[144] International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), Q1A (R2). Specifications: Stability Testing of New Drug Substances and Products, 2003. (Accessed April 20, 2017.) http://www.ich.org/fileadmin/Public_Web_ Site/ICH_ Products/Guidelines/Quality/Q1A_R2/Step4/Q1A_R2__Guideline.pdf.
[145] B. Van Eerdenbrugh, L.S. Taylor, Application of mid-IR spectroscopy for the characterization of pharmaceutical systems, Int. J. Pharm. 417 (2011) 3–16.
[146] A. Ito, T. Watanabe, S. Yada, T. Hamaura, H. Nakagami, K. Higashi, K. Moribe, K. Yamamoto, Prediction of recrystallization behaviour of troglitazone/polyvinylpyrrolidone solid dispersion by solid-state NMR, Int. J. Pharm. 383 (2010) 18–23.
[147] W.J. Tan, E. Widjaja, Multivariate chemometric approach to thermal solid-state FT-IR monitoring of pharmaceutical drug compound, J. Pharm. Sci. 97 (2008) 3379–3387.
[148] C. Shende, W. Smith, C. Brouillette, S. Farquharson, Drug stability analysis by raman spectroscopy, Pharmaceutics 6 (2014) 651–662.
[149] B. Rojek, M. Wesolowski, Fourier transform infrared spectroscopy supported by multivariate statistics in compatibility study of atenolol with excipients, Vibrat. Spectrosc. 86 (2016) 190–197.
[150] L. de Moura França, M.F. Pimentel, S.S. Simões, S. Grangeiro Jr., J.M. Prats-Montalbán, A. Ferrer, NIR hyperspectral imaging to evaluate degradation in captopril commercial tablets, Eur. J. Pharm. Biopharm. 104 (2016) 180–188.
[151] M. Jamrógiewicz, A. Ciesielski, Application of vibrational spectroscopy thermal analyses and X-ray diffraction in the rapid evaluation of the stability in solid-state of ranitidine, famotidine and cimetidine, J. Pharm. Biomed. Anal. 107 (2015) 236–243.
[152] M.A.D.S. Farias, F.L.F. Soares, R.L. Carneiro, Crystalline phase transition of ezetimibe in final product, after packing, promoted by the humidity of excip-

ients: monitoring and quantification by Raman spectroscopy, J. Pharm. Biomed. Anal. 121 (2016) 209–214.
[153] J.S.P. Daniel, I.P. Veronez, L.L. Rodrigues, M.G. Trevisan, J.S. Garcia, Risperidone −solid-state characterization and pharmaceutical compatibility using thermal and non-thermal techniques, Thermochim. Acta 568 (2013) 148–155.
[154] H.E. Williams, V.C. Loades, M. Claybourn, D.M. Murphy, electron paramagnetic resonance spectroscopy studies of oxidative degradation of an active pharmaceutical ingredient and quantitative analysis of the organic radical intermediates using partial least-squares regression, Anal. Chem. 78 (2006) 604–608.
[155] L. Hansen, J. Van Renterghem, R. Daoussi, C. Vervaet, J.P. Remon, T. De Beer, Spectroscopic evaluation of a freeze-dried vaccine during an accelerated stability study, Eur. J. Pharm. Biopharm. 104 (2016) 89–100.
[156] N. Chieng, H. Trnka, J. Boetker, M. Pikal, J. Rantanen, H. Grohganz, Detecting phase separation of freeze-dried binary amorphous systems using pair-wise distribution function and multivariate data analysis, Int. J. Pharm. 454 (2013) 167–173.
[157] N.L. Calvo, R.M. Maggio, T.S. Kaufman, Chemometrics-assisted solid-state characterization of pharmaceutically relevant materials, J. Pharm. Biomed. Anal. (2017), Please, replace the words "companion review" with the page numbers of our companion review in the same issue of this Journal.(Companion review).
[158] L. Saerens, D. Ghanam, C. Raemdonck, K. Francois, J. Manz, R. Krüger, S. Krüger, C. Vervaet, J.P. Remon, T. De Beer, In-line solid state prediction during pharmaceutical hot-melt extrusion in a 12-mm twin screw extruder using Raman spectroscopy, Eur. J. Pharm. Biopharm. 87 (2014) 606–615.
[159] P. Szynkaruk, M. Wesolowski, M. Samson-Rosa, Principal component analysis of thermal decomposition of magnesium salts used as drugs, J. Therm. Anal. Calorim. 101 (2010) 505–512.
[160] N.L. Calvo, S.O. Simonetti, R.M. Maggio, T.S. Kaufman, Thermally-induced solid-state transformation of cimetidine. A multi-spectroscopic/chemometrics determination of the kinetics of the process and structural elucidation of one of the products as a stable N3-enamino tautomer, Anal. Chim. Acta 875 (2015) 22–32.
[161] A.L. Kelly, S.A. Halsey, R.A. Bottom, S. Korde, T. Gough, A. Paradkar, A novel transflectance near infrared spectroscopy technique for monitoring hot melt extrusion, Int. J. Pharm. 496 (2015) 117–123.
[162] H. Nie, Z. Liu, B.C. Marks, L.S. Taylor, S.R. Byrn, P.J. Marsac, Analytical approaches to investigate salt disproportionation in tablet matrices by Raman spectroscopy and Raman mapping, J. Pharm. Biomed. Anal. 118 (2016) 328–337.
[163] M. Morhac, V. Matousek, Peak clipping algorithms of background estimation in spectroscopic data, Appl. Spectrosc. 62 (2008) 91–106.
[164] R. Caliandro, G. Di Profio, O. Nicolotti, Multivariate analysis of quaternary carbamazepine–saccharin mixtures by X-ray diffraction and infrared spectroscopy, J. Pharm. Biomed. Anal. 78–79 (2013) 269–279.
[165] R.O. Gaitano, N.L. Calvo, G.E. Narda, T.S. Kaufman, R.M. Maggio, E.V. Brusau, Preparation and physical characterization of a diclofenac-ranitidine co-precipitate for improving the dissolution of diclofenac, J. Pharm. Biomed. Anal. 105 (2016) 1258–1268.
[166] M.C. Sarraguca, P.R.S. Ribeiro, A.O. Dos Santos, J.A. Lopes, Batch statistical process monitoring approach to a cocrystallization process, J. Pharm. Sci. 104 (2015) 4099–4108.
[167] F.L.F. Soares, R.L. Carneiro, Evaluation of analytical tools and multivariate methods for quantification of co-former crystals in ibuprofen-nicotinamide co-crystals, J. Pharm. Biomed. Anal. 89 (2014) 166–175.
[168] C. Wood, A. Alwati, S. Halsey, T. Gough, E. Brown, A. Kelly, A. Paradkar, Near infrared spectroscopy as a multivariate process analytical tool for predicting pharmaceutical co-crystal concentration, J. Pharm. Biomed. Anal. 129 (2016) 172–181.
[169] H. Winning, F.H. Larsen, R. Bro, S.B. Engelsen, Quantitative analysis of NMR spectra with chemometrics, J. Magn. Reson. 190 (2008) 26–32.
[170] K.T. Jensen, F.H. Larsen, C. Cornett, K. Löbmann, H. Grohganz, T. Rades, Formation mechanism of coamorphous drug-amino acid mixtures, Mol. Pharm. 12 (2015) 2484–2492.
[171] M. Allesø, S. Velaga, A. Alhalaweh, C. Cornett, M.A. Rasmussen, F. van den Berg, H. Lopez de Diego, J. Rantanen, Near-infrared spectroscopy for cocrystal screening. a comparative study with Raman spectroscopy, Anal. Chem. 80 (2008) 7755–7764.
[172] A. Beyer, H. Grohganz, K. Löbmann, T. Rades, C.S. Leopold, Multivariate quantification of the solid-state phase composition of co-amorphous naproxen-indomethacin, Molecules 20 (2015) 19571–19587.
[173] S. Sasic, Y. Ozaki, Raman, Infrared, and Near-Infrared Chemical Imaging, Wiley, New York, USA, 2010.
[174] S. Piqueras, J. Burger, R. Tauler, A. de Juan, Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution, Chemometr. Intell. Lab. Syst. 117 (2012) 169–182.
[175] A.A. Gowen, F. Marini, C. Esquerre, C. O'Donnell, G. Downey, J. Burger, Time series hyperspectral chemical imaging data: challenges, solutions and applications, Anal. Chim. Acta 705 (2011) 272–282.
[176] J.M. Amigo, Study of pharmaceutical samples by NIR chemical-image and multivariate analysis, Trends Anal. Chem. 27 (2008) 696–713.
[177] J. Wu, Linear quantification calibration of crystallinity based on a simple univariate analysis of binary chemical images by selecting a threshold to define the sampling depth in Raman mapping, J. Raman Spectrosc. 44 (2013) 1578–1586.
[178] A. de Juan, M. Maeder, T. Hancewicz, L. Duponchel, R. Tauler, Chemometric tools for image analysis, in: R. Salzer, H.W. Siesler (Eds.), Infrared and Raman Spectroscopic Imaging, Wiley Weinheim, Germany, 2009.
[179] J.M. Amigo, Practical issues of hyperspectral imaging analysis of solid dosage forms, Anal. Bioanal. Chem. 398 (2010) 93–109.
[180] Y. Roggo, M. Ulmschneider, Chemical imaging and chemometrics useful tools for process analytical technology, in: S.C. Gad (Ed.), Pharmaceutical Manufacturing Handbook: Regulations and Quality, Wiley, New York, USA, 2008, pp. 411–431.
[181] J. Djuris, S. Ibric, Z. Djuric, Chemometric methods application in pharmaceutical products and processes analysis and control, in: J. Djuris (Ed.), Computer-Aided Applications in Pharmaceutical Technology, Woodhead Publishing, Cambridge, UK, 2013, pp. 57–90.
[182] B. Vajna, G. Patyi, Z. Nagy, A. Bódis, A. Farkas, G. Marosi, Comparison of chemometric methods in the analysis of pharmaceuticals with hyperspectral Raman imaging, J. Raman Spectrosc. 42 (2011) 1977–1986.
[183] A.A. Gowen, C.P. O'Donnell, P.J. Cullen, S.E.J. Bell, Recent application of chemical imaging to pharmaceutical process monitoring and quality control, Eur. J. Pharm. Biopharm. 69 (2008) 10–22.
[184] W.-Q. Lin, J.-H. Jiang, H.-F. Yang, Y. Ozaki, G.-L. Shen, R.-Q. Yu, Characterization of chloramphenicol palmitate drug polymorphs by Raman mapping with multivariate image segmentation using spatial directed agglomeration clustering method, Anal. Chem. 78 (2006) 6003–6011.
[185] A. Docoslis, K.L. Huszarik, G.Z. Papageorgiou, D. Bikiaris, A. Stergiou, E. Georgarakis, Characterization of the distribution, polymorphism, and stability of nimodipine in its solid dispersions in polyethylene glycol by micro-Raman spectroscopy and powder X-ray diffraction, AAPS J. 43 (2007) 361–370.
[186] S. Sasic, Raman mapping of low-content API pharmaceutical formulations. I. Mapping of alprazolam in alprazolam/xanax tablets, Pharm. Res. 24 (2007) 58–65.
[187] P.-Y. Sacré, E. Deconinck, L. Saerens, T. De Beer, P. Courselle, R. Vancauwen-berghe, P. Chiap, J. Crommen, J.O. De Beer, Detection of counterfeit Viagra® by Raman microspectroscopy imaging and multivariate analysis, J. Pharm. Biomed. Anal. 56 (2011) 454–461.
[188] C. Gendrin, Y. Roggo, C. Collet, Pharmaceutical applications of vibrational chemical imaging and chemometrics: a review, J. Pharm. Biomed. Anal. 48 (2008) 533–553.
[189] P.-Y. Sacré, C. De Bleye, P.-F. Chavez, L. Netchacovitch, Ph. Hubert, E. Ziemons, Data processing of vibrational chemical imaging for pharmaceutical applications, J. Pharm. Biomed. Anal. 101 (2014) 123–140.
[190] J.B. Nanubolu, J.C. Burley, Investigating the recrystallization behavior of amorphous paracetamol by variable temperature Raman studies and surface Raman mapping, Mol. Pharm. 9 (2012) 1544–1558.
[191] S. Piqueras, L. Duponchel, R. Tauler, A. de Juan, Monitoring polymorphic transformations by using in situ Raman hyperspectral imaging and image multiset analysis, Anal. Chim. Acta 819 (2014) 15–25.
[192] B. Vajna, I. Farkas, A. Farkas, H. Pataki, Z. Nagy, J. Madarász, G. Marosi, Characterization of drug–cyclodextrin formulations using Raman mapping and multivariate curve resolution, J. Pharm. Biomed. Anal. 56 (2011) 38–44.
[193] S. Hellstén, H. Qu, T. Heikkilä, J. Kohonen, S.-P. Reinikainen, M. Louhi-Kultanen, Raman spectroscopic imaging of indomethacin loaded in porous silica, CrystEngComm 14 (2012) 1582–1587.
[194] B. Li, A. Calvet, Y. Casamayou-Boucau, C. Morris, A.G. Ryder, Low-content quantification in powders using Raman spectroscopy: a facile chemometric approach to sub 0.1% limits of detection, Anal. Chem. 87 (2015) 3419–3428.
[195] B. Vajna, H. Pataki, Z. Nagy, I. Farkas, G. Marosi, Characterization of melt extruded and conventional Isoptin formulations using Raman chemical imaging and chemometrics, Int. J. Pharm. 419 (2011) 107–113.
[196] J. El Haddad, B. Bousquet, L. Canioni, P. Mounaix, Review in terahertz spectral analysis, Trends Anal. Chem. 44 (2013) 98–105.
[197] J.A. Zeitler, P.F. Taday, D.A. Newnham, M. Pepper, K.C. Gordon, T. Rades, Terahertz pulsed spectroscopy and imaging in the pharmaceutical setting −a review, J. Pharm. Pharmacol. 59 (2007) 209–223.

[198] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, R. Raskar, Terahertz time-gated spectral imaging for content extraction through layered structures, Nat. Commun. 7 (2016) 12665.
[199] H. Ma, C.A. Anderson, Characterization of pharmaceutical powder blends by NIR chemical imaging, J. Pharm. Sci. 97 (2008) 3305–3320.
[200] L.A. Terra, R.J. Poppi, Monitoring the polymorphic transformation on the surface of carbamazepine tablets generated by heating using near-infrared chemical imaging and chemometric methodologies, Chemometr. Intell. Lab. Syst. 130 (2014) 91–97.
[201] G.P. Sabin, M.C. Breitkreitz, A.M. Souza, P. Fonseca, L. Calefe, M. Moffab, R.J. Poppi, Analysis of pharmaceutical pellets: an approach using near-infrared chemical imaging, Anal. Chim. Acta 706 (2011) 113–119.
[202] A.M. Brondi, L.A. Terra, G.P. Sabin, J.S. Garcia, R.J. Poppi, M.G. Trevisan, Mapping the polymorphic forms of fexofenadine in pharmaceutical tablets using near infrared chemical imaging, J. Near Infrared Spectrosc. 22 (2014) 211–220.
[203] S.A. Schönbichler, L.K.H. Bittner, A.K.H. Weiss, U.J. Griesser, J.D. Pallua, C.W. Huck, Comparison of NIR chemical imaging with conventional NIR, Raman and ATR-MIR spectroscopy for quantification of furosemide crystal polymorphs in ternary powder mixtures, Eur. J. Pharm. Biopharm. 84 (2013) 616–625.
[204] G.L. Alexandrino, J.M. Amigo, M.R. Khorasani, J. Rantanen, A.V. Friderichsen, R.J. Poppi, Unveiling multiple solid-state transitions in pharmaceutical solid dosage forms using multi-series hyperspectral imaging and different curve resolution approaches, Chemometr. Intell. Lab. Syst. 161 (2017) 136–144.
[205] Z. Rahman, A. Siddiqui, M.A. Khan, Assessing the impact of nimodipine devitrification in the ternary cosolvent system through quality by design approach, Int. J. Pharm. 455 (2013) 113–123.
[206] A. Siddiqui, Z. Rahman, V.A. Sayeed, M.A. Khan, Chemometric evaluation of near infrared Fourier transform infrared, and Raman spectroscopic models for the prediction of nimodipine polymorphs, J. Pharm. Sci. 102 (2013) 4024–4035.
[207] W.F. de Carvalho Rocha, G.P. Sabin, P.H. Março, R.J. Poppi, Quantitative analysis of piroxicam polymorphs pharmaceutical mixtures by hyperspectral imaging and chemometrics, Chemometr. Intell. Lab. Syst. 106 (2011) 198–204.
[208] G.L. Alexandrino, M.R. Khorasani, J.M. Amigo, J. Rantanen, R.J. Poppi, Monitoring of multiple solid-state transformations at tablet surfaces using multi-series near-infrared hyperspectral imaging and multivariate curve resolution, Eur. J. Pharm. Biopharm. 93 (2015) 224–230.
[209] A. Siddiqui, Z. Rahman, S. Bykadi, M.A. Khan, Chemometric methods for the quantification of crystalline tacrolimus in solid dispersion by powder X-ray diffractrometry, J. Pharm. Sci. 103 (2014) 2819–2828.
[210] Y.C. Shen, Terahertz pulsed spectroscopy and imaging for pharmaceutical applications: a review, Int. J. Pharm. 417 (2011) 48–60.
[211] M. Nia, B.D. Ratner, Differentiating calcium carbonate polymorphs by surface analysis techniques −an XPS and TOF-SIMS study, Surf. Interface Anal. 40 (2008) 1356–1361.
[212] Y. Guta, M. Boiret, L. Bultel, T. Renaud, A. Chetouani, A. Hafiane, Y.-M. Ginot, R. Jennane, Application of chemometric algorithms to MALDI mass spectrometry imaging of pharmaceutical tablets, J. Pharm. Biomed. Anal. 105 (2015) 91–100.
[213] R. Gosselin, E. Vachon, L.A. Cournoyer, F. Clarke, Classifying pharmaceutical capsules through X-ray image analysis based on the agglomeration of their contents, J. Pharm. Innov. 11 (2016) 92–101.
[214] S. Matero, F. Van Den Berg, S. Poutiainen, J. Rantanen, J. Pajander, Towards better process understanding: chemometrics and multivariate measurements in manufacturing of solid dosage forms, J. Pharm. Sci. 102 (2013) 1385–1403.
[215] J. Rantanen, H. Wikström, R. Turner, L.S. Taylor, Use of in-line near-infrared spectroscopy in combination with chemometrics for improved understanding of pharmaceutical processes, Anal. Chem. 77 (2005) 556–563.
[216] Z. Chen, D. Lovett, J. Morris, Process analytical technologies and real time process control a review of some spectroscopic issues and challenges, J. Proc. Control 21 (2011) 1467–1482.
[217] S. Laske, A. Paudel, O. Scheibelhofer, S. Sacher, T. Hoermann, A. Kelly, J. Rantannen, O. Korhonen, F. Stauffer, F. De Leersnyder, T. De Beer, J. Mantanus, P.-F. Chavez, B. Thoorens, P. Ghiotti, M. Schubert, P. Tajarobi, G. Haeffler, S. Lakio, M. Fransson, A. Sparén, S. Abrahmsén-Alami, S. Folestad, A. Funke, I. Backx, B. Kavsek, F. Kjell, M. Michaelis, T. Page, J. Palmer, A. Schaepman, S. Sekulic, S. Hammond, B. Braun, B. Colegrove, J. Khinast, A review of PAT strategies in secondary solid oral dosage manufacturing of small molecules, J. Pharm. Sci. 106 (2017) 667–712.
[218] G.X. Zhou, L. Crocker, J. Xu, J. Tabora, Z. Ge, In-line measurement of a drug substance via near infrared spectroscopy to ensure a robust crystallization process, J. Pharm. Sci. 95 (2006) 2337–2347.
[219] H.-E. Lee, M.-J. Lee, W.-S. Kim, M.-Y. Jeong, Y.-S. Cho, G.J. Choi, In-line monitoring and interpretation of an indomethacin anti-solvent crystallization process by near-infrared spectroscopy (NIRS), Int. J. Pharm. 420 (2011) 274–281.
[220] M. Römer, J. Heinämäki, I. Miroshnyk, N. Kivikero, N. Sandler, J. Rantanen, J. Yliruusi, Phase transformation of erythromycin a dihydrate during fluid bed drying, J. Pharm. Sci. 97 (2008) 4020–4029.
[221] A. Caillet, F. Puel, G. Fevotte, In-line monitoring of partial and overall solid concentration during solvent-mediated phase transition using Raman spectroscopy, Int. J. Pharm. 307 (2006) 201–208.
[222] J.G. Rosas, H. de Waard, T. De Beer, C. Vervaet, J.P. Remon, W.L.J. Hinrichs, H.W. Frijlink, M. Blanco, NIR spectroscopy for the in-line monitoring of a multicomponent formulation during the entire freeze-drying process, J. Pharm. Biomed. Anal. 97 (2014) 39–46.
[223] S. Matero, S. Poutiainen, J. Leskinen, K. Järvinen, J. Ketolainen, S. Reinikainen, M. Hakulinen, R. Lappalainen, A. Poso, The feasibility of using acoustic emissions for monitoring of fluidized bed granulation, Chemometr. Intell. Lab. Syst. 97 (2009) 75–81.
[224] T. Närvänen, T. Lipsanen, O. Antikainen, H. Räikkönen, J. Heinämäki, J. Yliruusi, Gaining fluid bed process understanding by in-line particle size analysis, J. Pharm. Sci. 98 (2009) 1110–1117.
[225] J. Huang, C. Goolcharran, J. Utz, P. Hernandez-Abad, K. Ghosh, A. Nagi, PAT approach to enhance process understanding of fluid bed granulation using in-line particle size characterization and multivariate analysis, J. Pharm. Innov. 5 (2010) 58–68.
[226] L. Saerens, L. Dierickx, T. Quinten, P. Adriaensens, R. Carleer, C. Vervaet, J.P. Remon, T. De Beer, In-line NIR spectroscopy for the understanding of polymer–drug interaction during pharmaceutical hot-melt extrusion, Eur. J. Pharm. Biopharm. 81 (2012) 230–237.
[227] J. Aaltonen, K. Kogermann, C.J. Strachan, J. Rantanen, In-line monitoring of solid-state transitions during fluidization, Chem. Eng. Sci. 62 (2007) 408–415.
[228] A.-M. Demers, R. Gosselin, J.-S. Simard, N. Abatzoglou, In-line near infrared spectroscopy monitoring of pharmaceutical powder moisture in a fluidised bed dryer: an efficient methodology for chemometric model development, Can. J. Chem. Eng. 90 (2012) 299–303.
[229] V.S. Barla, R. Kumar, V.R. Nalluri, R.R. Gandhi, K. Venkatesh, A practical evaluation of qualitative and quantitative chemometric models for real-time monitoring of moisture content in a fluidized bed dryer using NIR technology, J. Near Infrared Spectrosc. 22 (2014) 221–228.
[230] M. Korang-Yeboah, Z. Rahman, D.A. Shah, M.A. Khan, Spectroscopic-based chemometric models for quantifying low levels of solid-state transitions in extended release theophylline formulations, J. Pharm. Biomed. Anal. 105 (2016) 97–105.
[231] K.L. Vora, G. Buckton, D. Clapham, The use of dynamic vapour sorption and near infra-red spectroscopy (DVS-NIR) to study the crystal transitions of theophylline and the report of a new solid-state transition, Eur. J. Pharm. Sci. 22 (2004) 97–105.
[232] A. Touil, R. Peczalski, F. Zagrouba, Monitoring of theophylline dehydration in a vacuum contact dryer by near-infrared spectroscopy, Chem. Eng. Res. Des. 91 (2013) 1063–1070.
[233] A. Touil, R. Peczalski, F. Zagrouba, On-line monitoring of vacuum drying of theophylline using NIR spectroscopy: solid-state transitions, water content and semi-empirical modeling, Drug Dev. Ind. Pharm. 42 (2016) 1825–1832.
[234] M.-J. Lee, D.-Y. Seo, I.-C. Wang, N.-H. Chun, H.-E. Lee, M.-Y. Jeong, W.-S. Kim, G.J. Choi, Quantitative in-line monitoring of solvent-mediated polymorphic transformation of sulfamerazine by near-infrared spectroscopy, J. Pharm. Sci. 101 (2012) 1578–1586.
[235] M. Halstensen, K.H. Esbensen, Acoustic chemometric monitoring of industrial production processes, in: K.A. Bakeev (Ed.), Process Analytical Technology: Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries, 2nd ed., Wiley, New York, USA, 2010, pp. 281–302.
[236] D. Daniher, L. Briens, A. Tallevi, End-point detection in high-shear granulation using sound and vibration signal analysis, Powder Technol. 181 (2008) 130–136.
[237] M.K. Papp, C.P. Pujara, R. Pinal, Monitoring of high-shear granulation using acoustic emission: predicting granule properties, J. Pharm. Innov. 3 (2008) 113–122.
[238] F.N. Ihunegbo, C. Ratnayake, M. Halstensen, Acoustic chemometrics for on-line monitoring and end-point determination of fluidised bed drying, Powder Technol. 247 (2013) 69–75.
[239] T. Feng, F. Wang, R. Pinal, C. Wassgren, M.T. Carvajal, Investigation of the variability of NIR in-line monitoring of roller compaction process by using fast Fourier transform (FFT) analysis, AAPS PharmSciTech 9 (2008) 419–424.
[240] M.L. Hamad, C.D. Ellison, M.A. Khan, R.C. Lyon, Drug product characterization by macropixel analysis of chemical images, J. Pharm. Sci. 96 (2007) 3390–3401.

[241] J. Huang, G. Kaul, C. Cai, R. Chatlapalli, P. Hernandez-Abad, K. Ghosh, A. Nagi, Quality by design case study: an integrated multivariate approach to drug product and process development, Int. J. Pharm. 382 (2009) 23–32.
[242] A.S. Zidan, M.J. Habib, M.A. Khan, Process analytical technology: nondestructive evaluation of cyclosporine a and phospholipid solid dispersions by near infrared spectroscopy and imaging, J. Pharm. Sci. 97 (2008) 3388–3399.
[243] T.R.M. De Beera, C. Bodson, B. Dejaegher, B. Walczak, P. Vercruysse, A. Burggraeve, A. Lemos, L. Delattre, Y. Vander Heyden, J.P. Remon, C. Vervaet, W.R.G. Baeyens, Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process, J. Pharm. Biomed. Anal. 48 (2008) 772–779.
[244] R. Singh, A.D. Román-Ospino, R.J. Romañach, M. Ierapetritou, R. Ramachandran, Real time monitoring of powder blend bulk density for coupled feed-forward/feed-back control of a continuous direct compaction tablet manufacturing process, Int. J. Pharm. 495 (2015) 612–625.
[245] H. Wu, M. Tawakkul, M. White, M.A. Khan, Quality-by-Design (QbD): An integrated multivariate approach for the component quantification in powder blends, Int. J. Pharm. 372 (2009) 39–48.
[246] M.T. Islam, N. Scoutaris, M. Maniruzzaman, H.G. Moradiya, S.A. Halsey, M.S.A. Bradley, B.Z. Chowdhry, M.J. Snowden, D. Douroumis, Implementation of transmission NIR as a PAT tool for monitoring drug transformation during HME processing, Eur. J. Pharm. Biopharm. 96 (2015) 106–116.
[247] L. Netchacovitch, J. Thiry, C. De Bleye, E. Dumont, A. Dispas, C. Hubert, F. Krier, P.-Y. Sacré, B. Evrard, P. Hubert, E. Ziemons, A simple calibration approach based on film-casting for confocal Raman microscopy to support the development of a hot-melt extrusion process, Talanta 154 (2016) 392–399.
[248] F. Peng, W. Wang, W. Luo, Y. Liu, H. Li, Application of near-infrared spectroscopy for monitoring the formulation process of low-dose tablets, Anal. Meth. 6 (2014) 1905–1913.
[249] D. Mainali, J. Li, P. Yehl, N. Chetwyn, Development of a comprehensive near infrared spectroscopy calibration model for rapid measurements of moisture content in multiple pharmaceutical products, J. Pharm. Biomed. Anal. 95 (2014) 169–175.
[250] A.G. Sarraf, S. Cherkaoui, O. Jordan, R. Gurny, E. Doelker, Controlled drug release from melt-extrudates through processing parameters: a chemometric approach, Int. J. Pharm. 481 (2015) 9–17.
[251] T. Vigh, G. Drávavölgyi, P.L. Sóti, H. Pataki, T. Igricz, I. Wagner, B. Vajna, J. Madarász, G. Marosi, Z.K. Nagy, Predicting final product properties of melt extruded solid dispersions from process parameters using Raman spectrometry, J. Pharm. Biomed. Anal. 98 (2014) 166–177.
[252] A. Kauppinen, M. Toiviainen, M. Lehtonen, K. Järvinen, J. Paaso, M. Juuti, J. Ketolainen, Validation of a multipoint near-infrared spectroscopy method for in-line moisture content analysis during freeze-drying, J. Pharm. Biomed. Anal. 95 (2014) 229–237.
[253] D. Xiang, M. Konigsberger, B. Wabuyele, K. Hornung, J. Cheney, Development of robust quantitative methods by near-infrared spectroscopy for rapid pharmaceutical determination of content uniformity in complex tablet matrix, Analyst 134 (2009) 1405–1415.
[254] M. Fonteyne, S. Soares, J. Vercruysse, E. Peeters, A. Burggraeve, C. Vervaet, J.P. Remon, N. Sandler, T. De Beer, Prediction of quality attributes of continuously produced granules using complementary pat tools, Eur. J. Pharm. Biopharm. 82 (2012) 429–436.
[255] N. Boersen, M.T. Carvajal, K.R. Morris, G.E. Peck, R. Pinal, The influence of API concentration on the roller compaction process: modeling and prediction of the post compacted ribbon, granule and tablet properties using multivariate data analysis, Drug Dev. Ind. Pharm. 41 (2015) 1470–1478.
[256] P.-F. Chavez, P.-Y. Sacré, C. De Bleye, L. Netchacovitch, J. Mantanus, H. Motte, M. Schubert, P. Hubert, E. Ziemons, Active content determination of pharmaceutical tablets using near infrared spectroscopy as Process Analytical Technology tool, Talanta 144 (2015) 1352–1359.
[257] J. Arruabarrena, J. Coello, S. Maspoch, Raman spectroscopy as a complementary tool to assess the content uniformity of dosage units in break-scored warfarin tablets, Int. J. Pharm. 465 (2014) 299–305.