University of Mohammed V - RABAT
Faculty of Medicine and Pharmacy

Year: 2021 **THESIS N °: 26/21/CSVS**

Biopharmaceutical and Toxicological Analysis Research Team

Doctoral College: Life and Health Sciences (FMPR)

Speciality: Doctor in Drug Sciences (Analytical Chemistry/Chemometrics)

Doctoral Thesis

Presented by

Mohammed Alaoui Mansouri

Theme:

Evaluating different aspects of matrix effect on pharmaceutical drug analysis through the association of vibrational spectroscopy techniques to chemometric tools

Public defense: December 18th, 2021, Faculty of medicine and pharmacy in the presence of jury:

Pr. Yahia Cherrah	Mohammed V University in Rabat. FMPR	**Chair**
Pr. Philippe Hubert	University of Liege. Belgium	**Co-Chair**
Pr. Mustapha Bouatia	Mohammed V University in Rabat. FMPR	**Reporter**
Pr. Mohamadine ElMrabet	Institut Agronomique et Vétérinaire Hassan-II	**Reporter**
Pr. Khalid Digua	Faculté des sciences et techniques de Mohammedia	**Reporter**
Dr. Eric Ziemons	University of Liege. Belgium	**Examinator**
Pr. Roland Marini	University of Liege. Belgium	**Co-supervisor**
Pr. Abdelaziz Bouklouze	Mohammed V University in Rabat. FMPR	**Supervisor**

I am particular grateful to my supervisors ***Pr. Abdelaziz Bouklouze*** as well as ***Pr. Philippe Hubert*** and ***Pr. Roland Marini*** for giving me the opportunity to do a PhD in the laboratory of pharmacology and toxicology at University of Mohammed V in Morocco and in Laboratory of pharmaceutical analytical chemistry at university of Liege in Belgium. I am thankful for sharing their knowledge and ideas also for being available and helping me since I made my first step as a Ph. D student. My gratitude is also extended to ***Pr. Yahia Cherrah*** for the opportunity to start the Ph. D at university of Mohammed V

A special thanks to ***Dr. Eric Ziemons*** and ***Dr. Pierre-Yves Sacré*** for sharing their expertise and introducing me to vibrational spectroscopy and chemometrics as well as their scientific insight and all their wise suggestions and corrections. They have been excellent in their supervision of vibrational spectroscopy and multivariate data analysis.

I would like to show my gratitude and appreciations to my thesis committee members for their support and their valuable orientations.

I would like to thank also the reporters ***Pr. Mustapha Bouatia***, ***Pr. Khalid Digua*** and ***Pr. Mohamadine ElMrabet*** for their presence, for their careful reading of my thesis as well as for their remarks they will address to me during this defense in order to improve the quality of my manuscript.

I am so grateful to my colleagues: ***Ph. D students***, ***post-docs*** and all members of pharmaceutical and analytical chemistry laboratory at university of Liege for contributing to a perfect environment both socially and academically as well as their sympathy their great friendship.

I am also thankful to my colleagues at the laboratory of pharmacology and toxicology: ***Dr. Mourad Kharbach, Dr. Issam Barra, Mohammed Bousrabat, Dr Rabie Kamal and Mohammed Taabouz,*** for motivating and supporting me whether socially or academically. I really appreciated all the moments that we spent together.

My special thanks also to my ***Pr. Adil Touimi Benjelloun*** for inspiring and supporting me to seek the Ph.D dream since I was in a bachelor degree.

Finally, I am so grateful to my mother ***Sabah*** who passed away and my father ***Hassane*** for their love, care and support, my special thanks extended to my siblings ***Hajar, Youssef, Omar*** and all my family for their love, cheering me up and being beside me.

Mohammed

This thesis is proudly dedicated to the memory of my beloved mother ***Sabah Bouzoubaa*** *who passed away before witnessing my graduation. My mother who without her I can't be what I am today,*

I wish she was still alive to attend my defense and to be proud of me.

Titre : Evaluation des différents aspects de l'effet de matrice sur les analyses pharmaceutiques des médicaments par l'association des techniques spectroscopiques vibrationnelles aux outils chimiométriques
Auteur : Mohammed Alaoui Mansouri
Mots clés : Effet de matrice ; Chimiométrie, spectroscopie vibrationnelle ; analyses pharmaceutiques

L'objectif de cette thèse se focalise sur l'évaluation de la capacité de l'association des techniques spectroscopiques vibrationnelles à des outils chimiométriques afin de surmonter trois aspects différents de l'effet de matrice et d'interférence pour les analyses pharmaceutiques qualitative et quantitative des médicaments.
Le premier aspect était axé sur le test de la capacité d'appliquer l'analyse discriminante des moindres carrés partiels (PLS-DA) à chacune des techniques spectroscopiques infrarouge à transformée de Fourier (FT-IR) et à la spectroscopie proche infrarouge à transformée de Fourier (FT-NIR) pour discriminer entre différents produits pharmaceutiques à base de trois formes polymorphes principales du fluconazole. PLS-DA a montré sa capacité à discriminer entre les échantillons qui ne contiennent qu'une des trois formes polymorphes du fluconazole. Alors que dans le cas de l'effet matrice, qui est représenté par l'existence de deux polymorphes dans le même échantillon, PLS-DA a montré sa limite pour faire la bonne discrimination. Ainsi, une autre approche est connue par « Q residuals » et Hotteling's T^2 sont appliqués avant PLS-DA pour détecter ces échantillons avec deux formes polymorphes comme valeurs aberrantes.
Le deuxième aspect de l'effet de matrice s'est concentré principalement sur l'évaluation de l'application de la régression PLS et du l'approche « résolution des courbes multivariés » (MCR-ALS) sur les données FT-NIR pour quantifier la ciprofloxacine dans différentes marques de produits pharmaceutiques. Lorsqu'il s'agit de quantifier la ciprofloxacine dans des échantillons de même composition, les résultats obtenus basés sur l'erreur quadratique moyenne de prédiction (RMSEP) et les erreurs relatives (RE) ont prouvé que les deux approches de régression de la régression PLS et MCR-ALS sont capables quantifier la substance médicamenteuse avec précision. Cependant, la quantification de la ciprofloxacine que ce soit dans des mélanges de composition différente ou dans différentes marques de produits pharmaceutiques a clairement montré la limite de la régression PLS en raison de l'effet matrice, alors que la MCR-ALS basée sur ses faibles erreurs relatives et de prédiction, a montré sa capacité à surmonter le problème du changement dans la composition de la matrice en raison de son avantage de second ordre.
Le troisième et dernier aspect de l'effet de matrice vise à étudier une propriété particulière de la spectroscopie Raman pour quantifier l'ibuprofène dans un mélange ternaire à travers un emballage de polypropylène. Cette propriété de la spectroscopie Raman est connue sous le nom de « Spatially Offset Raman Scattering » (SORS), qui permet au laser monochromatique de traverser l'emballage et d'obtenir le spectre de ce qui se trouve à l'intérieur du contenant. Cette enquête a été menée à travers une étude de comparaison entre la rétrodiffusion (conventionnelle) et le mode SORS pour quantifier l'ibuprofène via l'interférence des emballages en polypropylène en évaluant les modèles de régression PLS avec des profils d'exactitude suivant les directives ICH Q2 (R1) sur la validation avec ±15 % comme limites d'acceptation. Sur la base des profils d'exactitudes, le mode SORS a démontré sa capacité à quantifier avec précision à travers l'interférence ce qui n'est pas possible dans le cas de mode rétrodiffusion.

Tile: Evaluating different aspects of matrix effects on pharmaceutical drug analysis through the association of vibrational spectroscopy to pharmaceutical drug analysis.
Author: Mohammed Alaoui Mansouri
Keywords: Matrix effect; Vibrational spectroscopy; chemometrics; pharmaceutical drug analysis

The aim of this thesis was to investigate the ability of the association of vibrational spectroscopic techniques to chemometric tools in order to overcome three different aspects of matrix and interference effects for qualitative and quantitative pharmaceutical drug analysis. The first aspect was focused on testing the ability of applying Partial Least Squares-Discriminant Analysis (PLS-DA) to each of Fourier transform infrared spectroscopy (FT-IR) and Fourier transform near infrared spectroscopy (FT-NIR) to discriminate between different pharmaceutical products based on three main polymorphic forms of the fluconazole. PLS-DA showed its ability to discriminate between samples that have only one of three polymorphic forms of fluconazole. While in case of the matrix effect, which is represented in the existence of two polymorphic in the same sample, PLS-DA showed its limit to do the right discrimination. Thus, another approach is known by Hotelling's T^2 and Q residuals joined PLS-DA to detect these samples with two polymorphic forms as outliers. The second aspect of matrix effect focused mainly on evaluating the application each of PLS regression and Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) model on FT-NIR data to quantify ciprofloxacin in different brands of pharmaceutical products. When it comes to quantify ciprofloxacin in samples with the same composition, the obtained results based on Root Mean Squares Error of Prediction (RMSEP) and Relative Errors (RE) proved that both of regression approaches of PLS regression and MCR-ALS are able to carry out the quantitation of the drug substance accurately. However, the quantitation of ciprofloxacin whether in mixtures of different matrix composition or in different brands of pharmaceutical products clearly showed the limit of PLS regression because of the matrix effect, whereas the MCR-ALS based on its low relative and prediction errors, has shown its ability to overcome the problem of the change in the matrix composition due to its second order advantage. The third and last aspect of interference effect aims to investigate a special property of Raman spectroscopy to quantify the ibuprofen in a ternary mixture through an interference polypropylene container. This property of Raman spectroscopy is known by Spatially Offset Raman Scattering (SORS), which allows the monochromatic laser to pass through the packaging and obtain the spectrum of what is inside the container. This investigation was carried out through a comparison study between backscattering (conventional) and SORS mode to quantify ibuprofen through the interference of polypropylene packaging by evaluating the PLS regression models with mean of accuracy profiles following the ICH Q2 (R1) guidelines on validation with ±15% as acceptance limits. Based on the obtained accuracy profiles, the SORS mode demonstrated its ability to quantify accurately through the interference which is not in case of backscattering mode.
Keywords: Matrix effect; Polymorphism; PLS-DA; Hotelling's T^2 and Q residuals; PLS regression, MCR-ALS; SORS; FT-IR; FT-NIR; Raman spectroscopy

العنوان: تقييم جوانب تأثير السواغات على التحليلات الصيدلانية للأدوية عبر ربط تقنيات التحليل الطيفي بأدوات الكيمياء الحسابية

الكاتب: محمد علوي منصوري

الكلمات الأساسية: التحليلات الصيدلانية، المقاربات الكيميائية، التحليل الطيفي، تأثير السواغات

كان الهدف من هذه الرسالة هو تقييم قدرة الربط بين التقنيات الطيفية والقياسات الكيميائية للتغلب على الجوانب المختلفة لتأثير السواغات و التحليل عبر العبوة على التحليلات الصيدلانية. يهدف الجانب الأول إلى اختبار القدرة على تطبيق التحليل الطيفي ومقاربة الكيمياء الحسابية للتمييز بين المستحضرات الصيدلانية المختلفة بناءً على ثلاثة أشكال رئيسية متعددة الأشكال الهندسية من الفلوكونازول. ركز الجانب الثاني من تأثير السواغات على تقييم تطبيق انحدار PLS ونموذج-MCR ALSباستخدام FT-NIR لتقدير سيبروفلوكساسين في العلامات التجارية الصيدلانية المختلفة. أظهر القياس الكمي للسيبروفلوكساسين في ماركات مختلفة من المستحضرات الصيدلانية بوضوح حد انحدار بسبب تأثير السواغات ، بينما أظهر MCR-ALS قدرته على التغلب على مشكلة تغيير تركيبة المصفوفة بسبب ميزتها من الدرجة الثانية. يهدف الجانب الثالث إلى دراسة خاصية مطيافية رامان لتحديد كمية الإيبوبروفين في خليط ثلاثي من خلال تداخل الحاوية. تُعرف خاصية Raman هذه من SORS ، والتي تسمح لليزر بالمرور عبر العبوة. يعتمد هذا التحقيق على دراسة مقارنة بين وضع التبعثر المرتجع ووضع SORS لتحديد كمية الإيبوبروفين من خلال البوليبروبلين من خلال تقييم نماذج الانحدار PLS من خلال الملامح مع حدود قبول ± 15٪. استنادًا إلى ملفات تعريف الدقة SORSقدرتها على التحديد الكمي من خلال التداخل لما لا يتم من خلال التشتت الخلفي.

A	
ALS	Alternating Least Squares
API	Active Pharmaceutical Ingredient
ATR	Attenuated Total Reflectance
D	
DSC	Differential Scanning Calorimetry
DRIFTS	Diffuse Reflectance Infrared Fourier Transform Spectroscopy
E	
EFA	Evolving Factor Analysis
EP	European Pharmacopeia
F	
FN	False Negative
FP	False Positive
FT	Fourier-Transform
H	
HPLC	High Performance Liquid Chromatography
I	
ICH	International Council of Harmonization
IR	Infrared
L	
LV	Latent Variables
M	
MC	Mean Centering
MCC	Microcrystalline cellulose
MCR	Multivariate Curve Resolution
MLR	Multiple Linear Regression
MSC	Multiplicative Scatter correction
N	
NIPALS	Non-Linear Iterative Partial Least Squares
NIR	Near-infrared
O	
ORS	Orbital Raster Scanning

P	
PCA	Principal Component analysis
PCR	Principal Component Regression
PLS	Partial Least Squares
PLS-DA	Partial Least Squares- Discriminant Analysis
PP	Polypropylene
PVC	Polyvinyl Chloride
R	
RE	Relative Error
RMSEC	Root Mean Squares error of Calibration
RMSECV	Root Mean Squares error of Cross Validation
RMSEP	Root Mean Squares error of Prediction
S	
SERDS	Shifted-Excitation Raman Difference Spectroscopy
SERS	Surface Enhanced Raman Spectroscopy
SIMPLISMA	self-modelling mixture analysis approach
SNV	Standard Normale Variate
SORS	Spatially Offset Raman Scattering
T	
TP	True Positive
TN	True Negative
U	
USP	United States Pharmacopeia
X	
XRPD	X-Ray Powder diffraction

General Introduction & Objectives

Vibrational spectroscopy techniques are considered more advantageous compared to other analytical techniques for different analysis fields including for instance pharmaceutical and food analysis. Nevertheless, extracting information of the analyzed sample from its spectra becomes more sophisticated due to the matrix effect. It is well known that handling matrix challenge is considered more complicated in food applications compared to pharmaceutical analysis, however; there exist different aspects where the matrix effect has an impact on the goal accomplishment whether it is qualitative or quantitative analysis. This challenge makes from combining vibrational spectroscopy with chemometric tools crucial to uncover the information hampered by the matrix effect.

This thesis focused mainly on studying three aspects that deal with different sides of matrix effect using properties of vibrational spectroscopy with chemometric tools together, and to overcome different challenges imposed by the matrix.

First aspect: Can the association of PLS-DA to FT-NIR or FT-IR discriminate between pharmaceutical samples based on polymorphic forms of the active pharmaceutical ingredient? Is it possible to base only on PLS-DA to deal with matrice effect if there is more than one polymorphic form in the sample?

To answer these questions, the first aspect was based on using fluconazole to discriminate between its pure polymorphic forms by FT-IR and FT-NIR spectroscopy combined with PLS-DA. This discrimination showed its limit by the presence of two or more polymorphic forms in one pharmaceutical product, which represent a kind of matrix effect, as in case of polymorphic conversion or the challenge of falsified medicines. These challenges make from PLS-DA unable to do the right classification of pure polymorphic form in pharmaceuticals but associating PLS-DA to the approach of hotteling's T^2 and Q residuals allows detecting these samples that could be falsified or being impacted by the presence of two or more polymorphic forms and being considered as outliers. Thus, PLS-DA can be used to classify only samples that contains pure polymorphic forms of fluconazole.

Second aspect: Can the association of FT-NIR with PLS or with MCR-ALS overcome the quantitation of active pharmaceutical ingredient in different pharmaceutical products?

The aim of this aspect can be reached by evaluating the ability of the association of FT-NIR with PLS or with MCR-ALS to quantify ciprofloxacin in different pharmaceutical products, that the excipients used with the active pharmaceutical ingredient can vary from one

pharmaceutical product to another. It is well known that the prediction error increases whenever there is a change in the matrix composition. In this case, MCR-ALS is known by its second order calibration advantage that based on correlation constraint and the optimization with alternating least squares (ALS) to allow to deal well with matrix changes between several pharmaceutical products. The analytical performances of MCR-ALS model were compared to those obtained by PLS regression.

Third aspect: Is it possible to quantify an active pharmaceutical ingredient through the packaging by combining Raman spectroscopy to PLS regression?

This question can be answered by performing the quantitation of ibuprofen in a ternary mixture with microcrystalline cellulose and mannitol through polypropylene packaging, which represents external interference, using Raman spectroscopy and PLS regression. In this case, polypropylene shows a real interference challenge from quantifying the API accurately. However; using the technology of Raman spectroscopy called spatially offset Raman spectroscopy (SORS), that is based on creating an offset between the spot of incident light and the spot of collecting the spectra, allows to acquire the spectrum of the sample through the packaging. This evaluation was carried through the comparison of different devices based on Raman conventional and SORS mode using the total error approach.

Chapter I

Literature Review

I. Pharmaceutical drug analysis

Pharmaceutical drug analysis is a procedures set of simple chemical experiments and monographs of pharmacopeias such as European Pharmacopeia (EP) and United states Pharmacopeia (USP) that includes tests as identification and determination of the pharmaceutical ingredients used in pharmaceutical products that are established for ensuring the identity and purity of different pharmaceutical ingredients including drug substances and excipients [1]

Pharmaceutical active ingredients (API) within the group of multichemical ingredients may vary in composition depending on production. Solvents that remain from the synthesis of drug substance. For example, during manufacturing process of ibuprofen, the synthesis impurities as mentioned in Figure 1, that are produced, has to be extracted, eliminated and then qualified due to their impact on the quality of final pharmaceutical products [2].

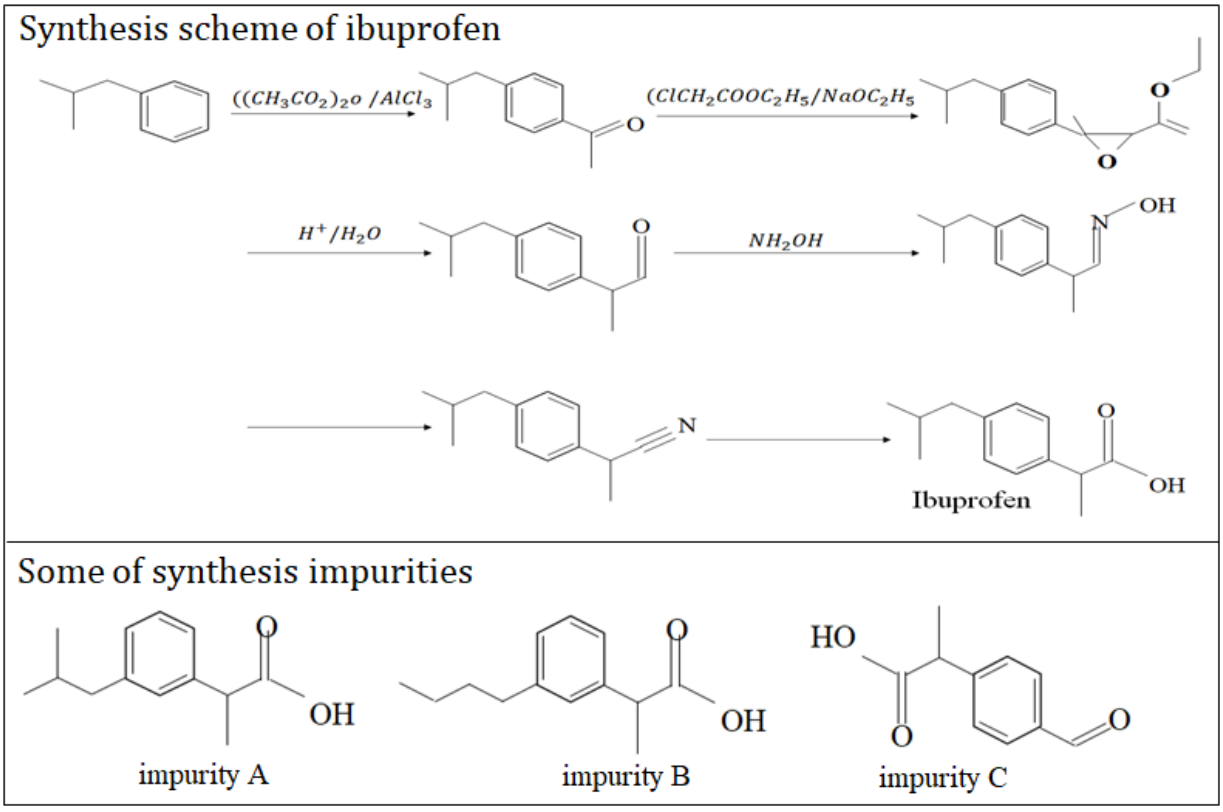

Figure 1: Synthesis scheme and synthesis impurities of ibuprofen (adapted from [3,4]).

Pharmaceutical analysis of drug is totally dependent on the nature and properties of the drug to be analyzed. Drug to be analyzed possesses two kinds of properties that are as follows:

- Physical properties are measured or observed without any change in the matter composition, but depend on the change that can occur in the crystalline structure and polymorphic forms. Physical properties include appearance, odor, color, melting point, boiling point, solubility, density, polarity, opacity, viscosity, and many others. These properties play an important role in the design of sample preparation for pharmaceutical analysis.

- Chemical properties are observed whenever a substance undergoes a chemical change. Ionization, functional groups, partition coefficient, pH, degradation and stereochemistry, heat of combustion, and enthalpy change are the important chemical properties that should be considered during the selection of an appropriate analytical method for pharmaceutical analysis.

Pharmaceutical analysis is a broader field and its applications are vast which are as follows:

- Identification and determination of APIs in pharmaceutical products and raw material.
- Identification and determination of impurities in the raw material and finished dosage forms of the pharmaceutical products.
- Identification and determination of contaminants, that can be physical as particles or fibers, chemical as moisture and biological as bacteria, in bulk drug or raw material.

The quality control of pharmaceuticals has been considered as an essential operation of the pharmaceutical industry whenever a new pharmaceutical product is produced [5]. Marketed drugs are expected to be safe and therapeutically active formulations. In order to have a consistent and predictable effectiveness, a drug must be formulated in the most appropriate dosage forms for administration to the patients for disease diagnosis and treatment.

The harmonization of analytical techniques for pharmaceutical analysis is important to accomplish legal requirements of the safety and therapy of the drug product. The first step for that purpose was the establishment of Pharmacopoeias which provided the basis for setting up national pharmacopoeias [6–9]. The second step was initiating International Conference on Harmonization (ICH) with the aim of harmonizing the efforts of registration agencies, principal pharmacopoeias, and pharmaceutical industries to improve the quality of pharmaceutical products using various analytical techniques for pharmaceutical analysis. The guidelines of ICH have been declared as authoritative worldwide with respect to drug-related

issues which make from pharmaceutical analysis an important field that increases the safety of drug therapy [10,11].

I.1. Solid-state of drug substance and polymorphism

The majority of drug compounds exist in solid state. The availability of drug compounds in the solid state is mainly refers to their high chemical stability in the solid state compared to solution. In addition, these pharmaceutical ingredients prove a higher dosage precision when they are in solid dosage forms. Besides that, they are known to be easily handled which contribute to the safety and reliability of the drug product.

In fact, the properties of solids depend on the molecular features, the intermolecular arrangement, and forces between the molecules as well as physical effects at the micro- and macroscopic level.

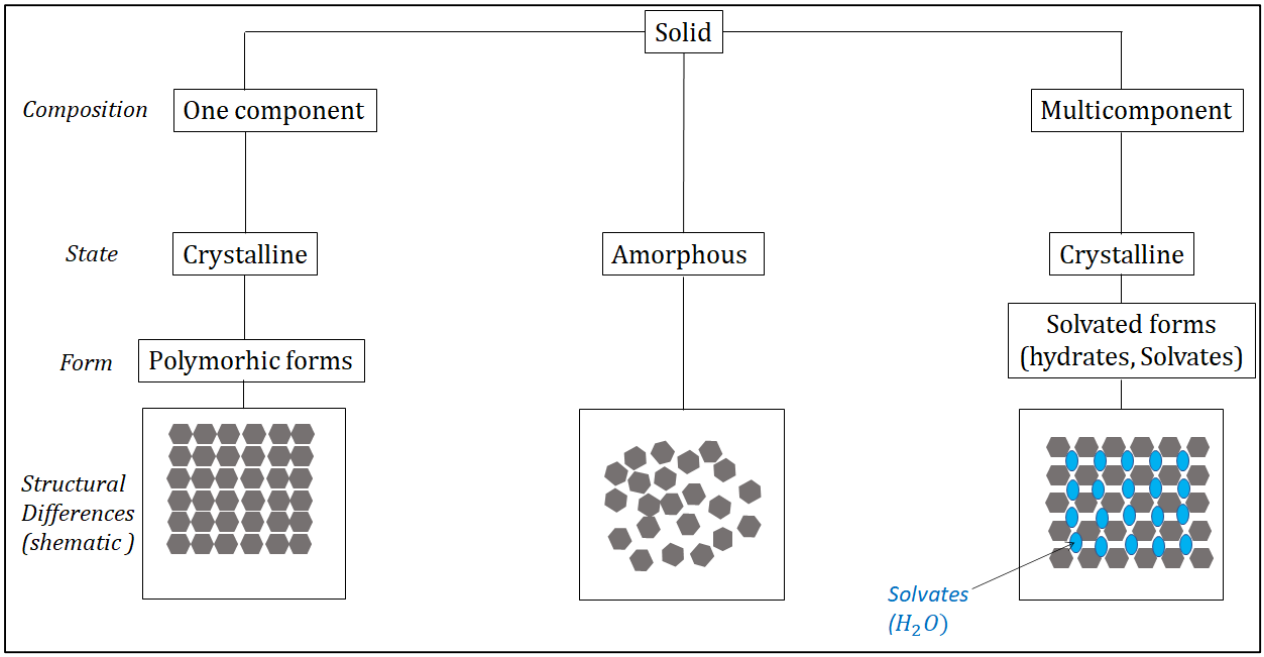

Figure 2: Classification of solid state forms (adapted from [12]).

As displayed in Figure 2, solid state can be exist in different forms:

Polymorphism is a solid-state phenomenon and is the ability of an element or compound to exist in more than one or the conformation of the molecule, or a combination of both. Crystals may also have a different shape or morphology (habit), keeping the same crystal structure and thus be of the same solid-state form [12].

Solvates are formed when spaces within the crystal lattice are occupied by solvent molecules. Due to its small size and prevalence, water is the most common solvent to be associated with

the lattice and the term hydrate is then used. Unlike hydrates, solvates are rarely used as an active pharmaceutical ingredient. However, they may be formed as precursors to the required form and so their characterization is often essential in order to understand the solid-state properties of the drug substance. The potential impact of changes in hydration state exists throughout the development process since these changes can have an impact on drug manufacturing. Substances may hydrate or dehydrate in response to changes of manufacturing and storing conditions.

Amorphous do not follow any crystalline order but since they maintain a molecular structure they still give a vibrational spectrum which may be distinct from the crystalline material. The ability of organic compounds to form crystals decreases with increasing molecular size and thus many large drug molecules, proteins, and polymers form glasses rather than crystals. However, small organic molecules also form glasses that are more or less stable. A glass is generally defined as a liquid that has no fluidity. The glassy state is of particular interest for poorly water-soluble drug compounds because this highly energetic state usually shows a higher solubility than any crystalline phase. An amorphous solid represents a metastable state that theoretically may crystallize at any time. Usually the amorphous form is also less chemically stable and more hygroscopic than a crystalline form. Amorphous forms may result from different pharmaceutical processes such as (spray drying, milling, lyophilization, granulation, fast precipitation from solvents).

I.2. Matrix and interference effect

Matrix is a set of the components exist in a sample beside the target analyte [13]. The matrix can have a significant impact on the analysis method and the accuracy of the obtained results; this significant impact is known by matrix effect [14]. For making method useful and acquiring accurate results, the elimination or reduction of the matrix effect is required. The main aspect that determine the usefulness of the method are summarized in sample preparation process and evaluate the selectivity of that method [15].

In case of spectroscopic analysis, which is relevant to the scope of the research works in this thesis, the influence of matrix effect, which are represented in pharmaceutical material properties, on optical spectroscopy techniques has been shown to be of great importance [16,17]. Dry pharmaceutical powders have significant elastic scattering and, in most wavelength bands, low absorption, resulting in potentially large effects on spectra and quantitative calibrations [18].

II. Vibrational spectroscopy

Spectroscopic methods have been designated to study molecular structure and matter. The rapid development of spectroscopic instrumentation has expanded the application of these techniques to many areas including food and pharmaceutical analysis.

In fact, spectroscopy is defined by a study of the interaction between electromagnetic radiations and the matter. The electromagnetic radiation covers a wide wavelength range, from low-energy radio waves to high-energy γ-rays as it is illustrated in Figure 3. The main characteristics of an electromagnetic radiation are wavelength λ (the length of one wave, in nm), frequency ν (the number of vibrations per unit time, in Hz), and the wavenumber (the number of waves per unit length, in cm^{-1}) [19].

The relationship between frequency ν and energy E can be given by the following equation:

$$E = h\nu$$ (h is the plank constant ($h = 6.62 * 10^{-34}$j.s))

Which

While the frequency ν is related to the wavelength λ by :

$$\nu = \frac{c}{\lambda}$$ (c is the celerity of light ($c = 3 * 10^{8} m/s$)

Thus the equation relate the energy to wavelength is :

$$E = h\frac{c}{\lambda}$$

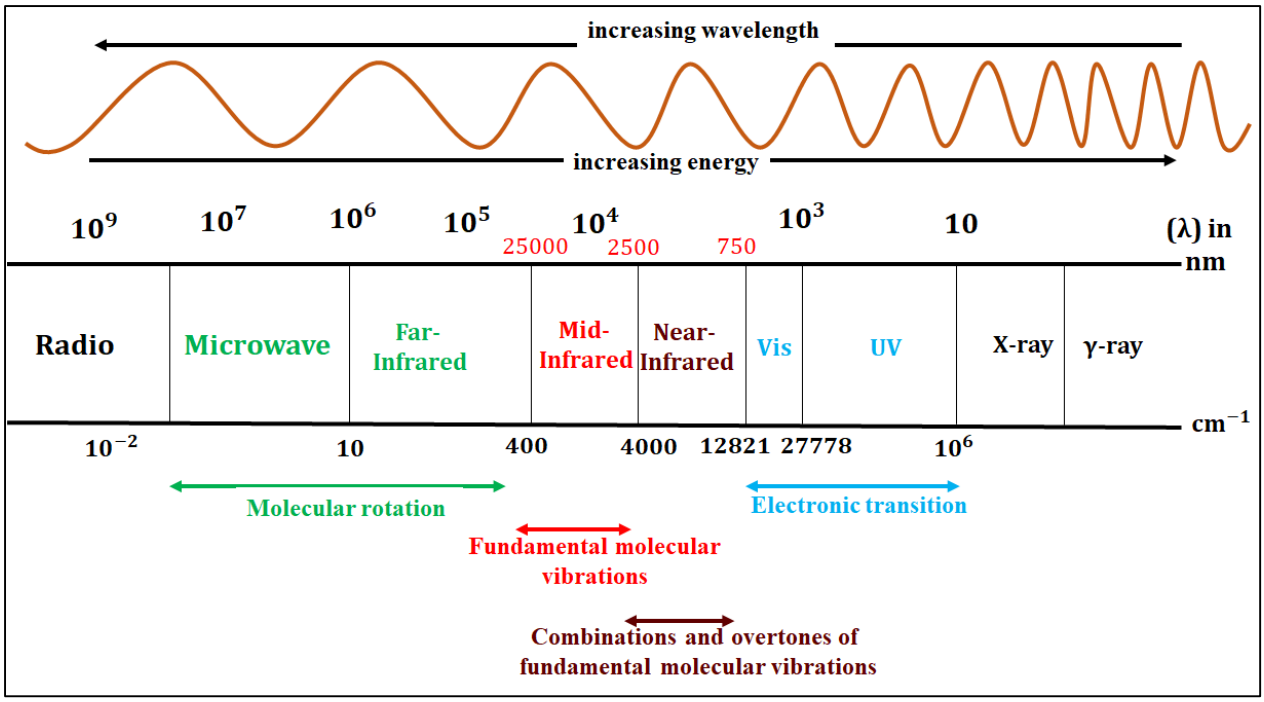

Figure 3: Regions of the electromagnetic spectrum and processes that may occur in the matter. (adapted from [20])

The main phenomenon can occur when the electromagnetic radiation interacts with the matter, it can be an absorption, transmission, reflection or scattered as illustrated in Figure 4. For example in the case of IR spectroscopy, radiation might be absorbed when the absorption of incident radiation at a particular frequency in the IR region is related to this specific vibrational excitation energy, whereas for Raman spectroscopy the radiation is scattered [21]. Both IR absorption and Raman scattering are mainly based on vibrational transitions that occur in the ground electronic state of the molecule [22].

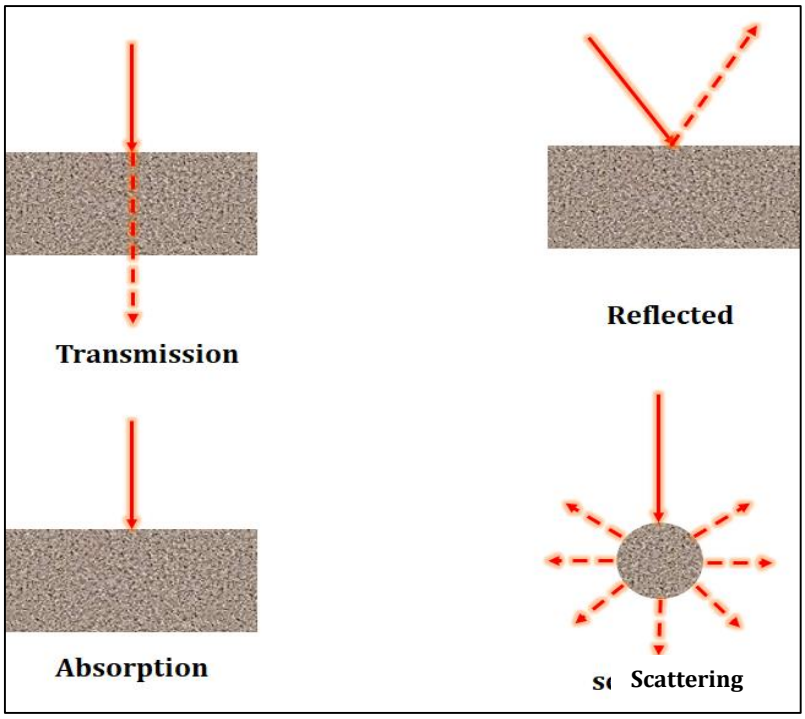

Figure 4: Phenomenon of electromagnetic radiation (adapted from [21]).

When a photon $E = h\nu$ is absorbed by a molecule, its energy is transferred to the molecule. This energy can be transformed to vibrational, rotational and electronic forms as illustrated in energy diagram of Figure 5. Rotational energy is characterized by the tumbling motion of the molecule, resulting from the absorption of energy in the microwave region. Vibrational energy corresponds to the absorption of energy by a molecule as the component atoms vibrate around the mean center of their chemical bonds. Electronic energy is linked to the transitions of electrons while they are distributed throughout the molecule, either localized within specific bonds, or delocalized over structures [23].

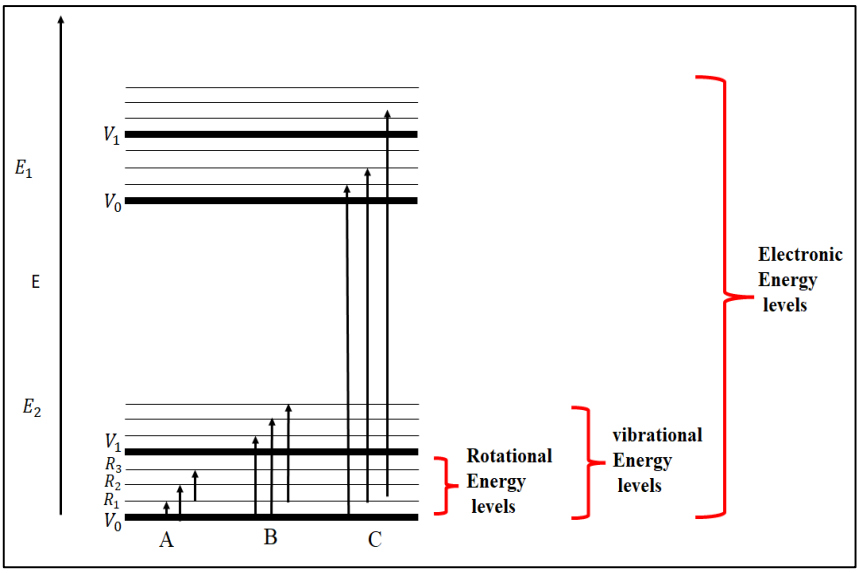

Figure 5: Energy levels for a molecule. Possible transitions: (A): Pure rotational transitions, (B) : Rotational-Vibrational transitions, (C) Rotational-Vibrational-Electronic transitions [24].

II.1. Fourier Transformed- infrared and near-infrared spectroscopy

II.1.1. Principle

Infrared (IR) spectroscopy is characterized by the feature that the atoms are in a continuous vibration and rotation. It is based on the absorption of polychromatic light by the sample and leads to vibrational transitions that occur from a ground vibrational state to an excited vibrational state as it is shown in Figure 6.

A vibrational mode cannot be observed in an infrared spectrum unless there exist a change in the molecular dipole moment. Each absorption band has an intensity that is proportional to the change in dipole moment. Thus, all vibrations of polar bonds are known to be strong in the IR spectrum such as O-H, C=O and N-H.

The IR include the spectrum between 12500 and 100 cm^{-1}, which is split into near-IR, mid-IR, and far-IR regions as follows:

- Near-infrared which range from 12500 to 4000 cm^{-1}, is used to figure out the constituents of organic matter and represents the characteristics of weak and overlapping absorption bands since these characteristics occurred by overtones and combinations, which are a vibrational mode is excited from the ground state to a higher state, and these combinations refer to the excitation of two molecular vibrations in the same time, of CH, NH, or OH stretching bands. In fact energy transitions of overtones are higher compared to fundamental vibrations [25–27].

- Mid-IR region, which range from 4000-400 cm^{-1}. The observed signals of mid-IR are due to fundamental vibrations including stretching and bending as it is seen in Figure 6. While stretching vibrations occur due to the change in bond length, bending vibrations are based on the change in bond angle [28]. Spectra can be analyzed based on two regions of mid-IR, the functional group (4000-1300 cm^{-1}) and the fingerprints (1250-600 cm^{-1}). The functional group includes molecules with simple, double and triple bonds. For molecules with simple bond, they cover the full stretching region of 4000- 400 cm^{-1}. For molecules with the triple bond ($-C \equiv C-$), they are characterized in the region of 2500–2000 cm^{-1} whereas molecules with double-bond (($-C = C-$), they cover region of 2000–1500 cm^{-1}. The region of 1250-600cm^{-1} is characterized by being a fingerprint of bending vibrations. The reason of considering the region below of 1250 cm^{-1} as a fingerprints, is due to its specificity and high number of signals appearing in this region, thus making it difficult for two samples to have identical spectral signature below 1250 cm^{-1} [29].

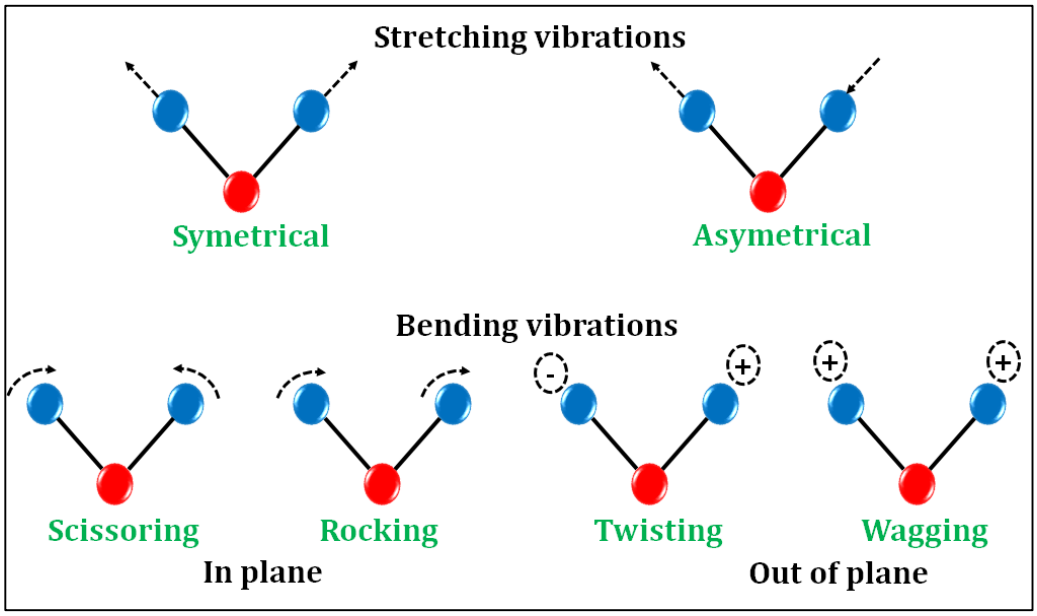

Figure 6: Major stretching and bending vibrational modes [30].

- The far-IR region that cover the region less than 400 cm^{-1} is not useful for structural elucidation as mid-IR and near-IR, but is able to elucidate the intramolecular stretching modes that bind heavy atoms, skeleton bending modes as a molecule containing heavier atoms, and vibrations of crystal lattice [31].

II.1.2. Instrumentation

There exist two kinds of infrared spectrophotometers: dispersive and Fourier transform (FT) devices. The main difference between both is summarized in using the interferometer in case

of Fourier transform as a beam splitter that will be described later in this section. The latter is known to be the most used compared to the group of dispersive equipment since the ability of the interferometer to measure the radiation from all wavelength simultaneously which is not in case of dispersive infrared spectrometers that measure all wavelengths consecutively [32,33].

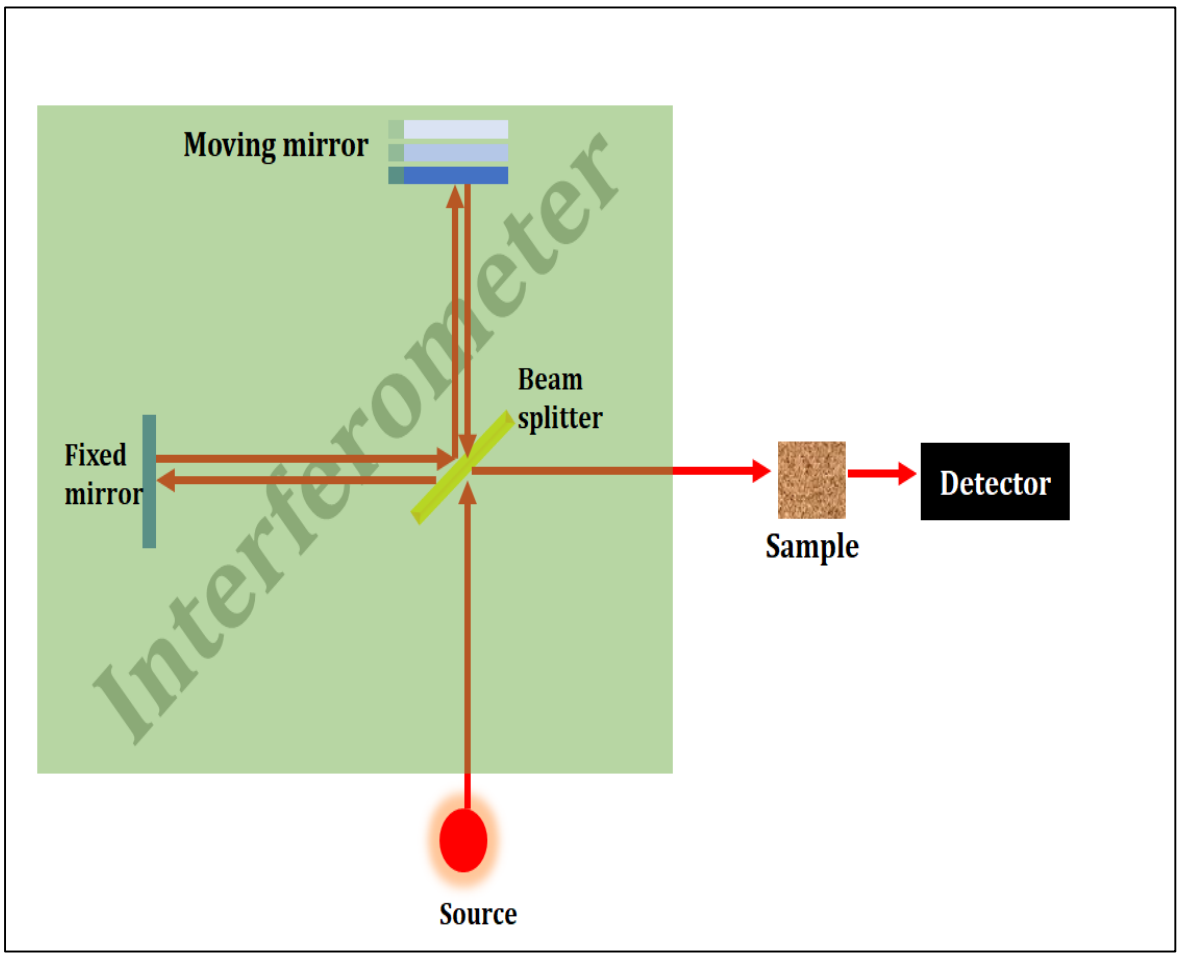

Figure 7: Principles of the Fourier-Transform Infrared Spectroscopy (FT-IR) *[34]*

FT-IR instrument is mainly composed of three components: a radiation source, an interferometer, and a detector which is illustrated in Figure 7. As cited before, the interferometer is the main component that characterize Fourier-transform among dispersive ones. It consists of a beam splitter, a fixed mirror and a moving mirror that moves in two directions back and forth. When the infrared radiation hits the beam splitter, it split into two beams, where the first is transmitted to the fixed mirror and the second is reflected to the moving mirror, then both of mirrors reflect the radiation back to the beam splitter. Because of the characteristic of the moving mirror regarding to the fixed mirror, an interference pattern is produced then goes through the sample and then recorded by the detector.

II.1.3. Sampling techniques in infrared spectroscopy

Choosing the right sampling techniques is considered critical and important issue to achieve a good qualitative and quantitative analysis whether for mid-infrared, near-infrared or even Raman spectroscopy.

Transmission, attenuated total reflection (is not used in case of FT-NIR) and diffuse reflection are the three sampling techniques commonly used in infrared as shown in Figure 8, whereas transflection (is not used in case of mid infrared), which is the combination of the reflection and transmission sampling techniques, is the least used sampling technique and it is often carried out in near-infrared spectroscopy.

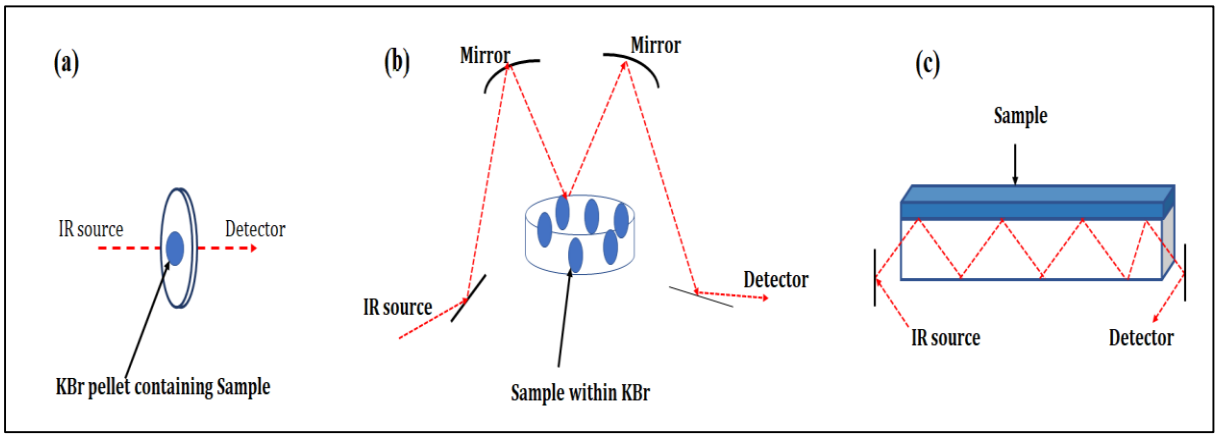

Figure 8: Most common sampling techniques for vibrational spectroscopy adapted from [35]. (a) Transmission mode, (b) Diffuse reflection infrared Fourier transform (DRIFT), (c) Attenuated total reflectance (ATR)

In transmission mode, the IR beam strikes the sample pellet and it results a transmitted IR signal that is recorded by the detector. The transparent KBr pellet is obtained by applying a high pressure on the homogeneous mixture of KBr and the sample. Unlike transmission, sample preparation of diffuse reflection infrared Fourier transform (DRIFT) mode is more simpler since it requires only mixing the sample with KBr [36]. In DRIFT, the IR beam interacts with the sample to a certain depth, and then is reflected from the sample and focused by a mirror to be oriented toward the detector. The obtained DRIFT spectrum is the same as the spectrum obtained by transmission mode [37]. Attenuated total reflectance is a recent and more flexible sampling technique. It consists of putting the sample in a close contact with a crystal of a high refraction index which is made mainly of ZnSe, Ge, ZnS, Si or diamond (that has been used in the research works of this thesis). A radiation beam enters into the crystal and undergoes total internal reflection. The beam penetrates a fraction of a wavelength beyond the reflecting surface, and then the sample in contact with the reflecting surface of

crystal absorbs radiation, the beam loses energy at the wavelength where the sample absorbs. The depth of penetration depends mainly to the wavelength, the refractive index of the crystal and the angle of incident radiation. ATR is nondestructive and allow fast and simple sampling whether the state of the sample is liquid or solid [38]. However, a good contact of sample to the crystal has to be verified which can be insured with gauge force [39].

II.1.4. Advantages and limitations of infrared

Each of mid and near-infrared spectroscopy has advantages over other analytical techniques, however; these techniques have some drawbacks. Among the most common advantages and limitations of these techniques are [24,40–43]:

Advantages of mid and near-infrared:

- Non-destructive and non-invasive method
- Rapid spectral measurements
- These techniques are low cost since sometimes they don't need reagents
- Enable the analysis through glass containers in case of near-infrared but not in case of mid-infrared.
- Infrared can carry out the analysis using optic probe which permit to analyze the sample in situ
- Sample does not often require sample preparation
- These techniques can carry out many scans of the same sample which lead to more accurate results
- Large amount of information through spectra provided from the analyzed sample as the functional group for mid-infrared.

Limitations of mid and near-infrared:

- Low sensitivity when it comes to determine analytes with concentration below 0.1% (w/w) for near-infrared whereas the sensitivity is higher for mid-infrared.

- The complexity to transfer calibration between instruments for both mid and near-infrared
- Reference samples needs accurate and chemical analysis.
- Due to the complexity to interpret the obtaining spectra especially for near-infrared (Broad and overlapping), chemometric tools are usually required for qualitative and quantitative analysis.
- Repeatability can be impacted in case of using ATR, that is used with mid-infrared, due to the small sampling volume especially if the sample is less homogeneous.

II.2. Raman spectroscopy

II.2.1. Concept of Raman spectroscopy

The Raman effect is a result of the inelastic scattering of a monochromatic light source, which is a laser with a wavelength between 200 and 1400 nm and include ultraviolet, visible or near-infrared. It has to be mentioned that 830 nm and 785 nm laser excitation are common wavelengths chosen for Raman spectroscopy and used in this thesis because they are known by their fluorescence suppression and their efficiency of good scattering [44].

In fact, the sample is irradiated by a monochromatic laser beam which interacts with the molecules of sample and originates a scattered light of Raman and Rayleigh scattering as shown in Figure 9 [45].

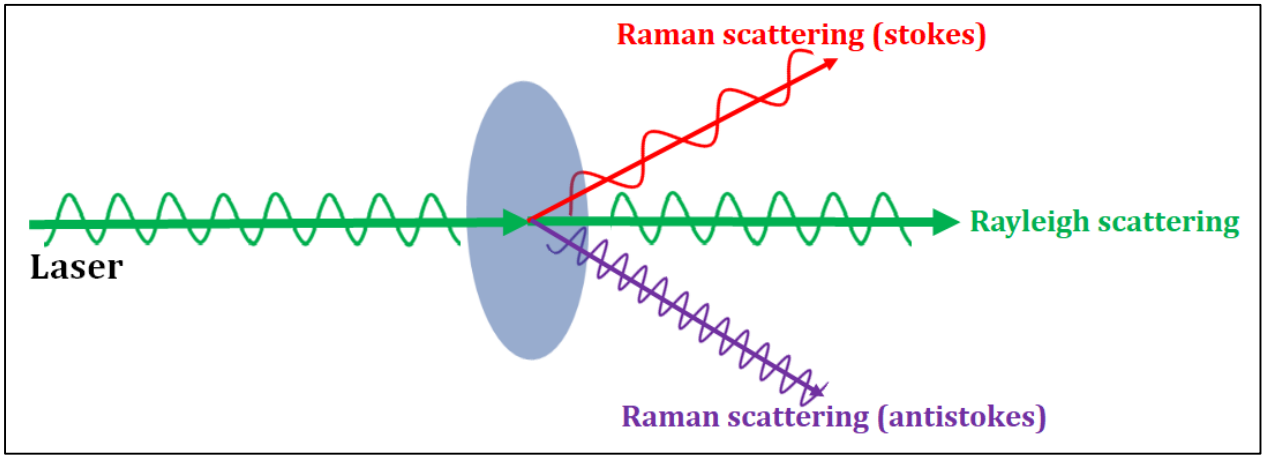

Figure 9: Raman scattering effect[46].

The scattered light has a frequency different from that of incident light (inelastic scattering) as illustrated in energy diagram of Figure 10. The Raman spectra is obtained because of the inelastic interaction between incident monochromatic light and the sample. When a monochromatic radiation strikes at sample, it scatters in all directions after its interaction with sample molecules [45,47]. The majority of the scattered radiation, which is known by Rayleigh scattering, has the same frequency as the incident radiation, whereas small fraction of scattered radiation has a frequency different from the frequency of incident radiation which is the origin of Raman scattering. In case of the frequency of incident radiation is higher than the frequency of scattered radiation, Stokes bands appear in Raman spectrum. While if the frequency of incident radiation is lower than the frequency of scattered radiation, anti-Stokes bands appear in Raman spectrum. This Raman spectrum cannot be obtained unless there exist a change in polarizability during bond vibration which is defined as the ease of distorting electrons from their original position [48–50].

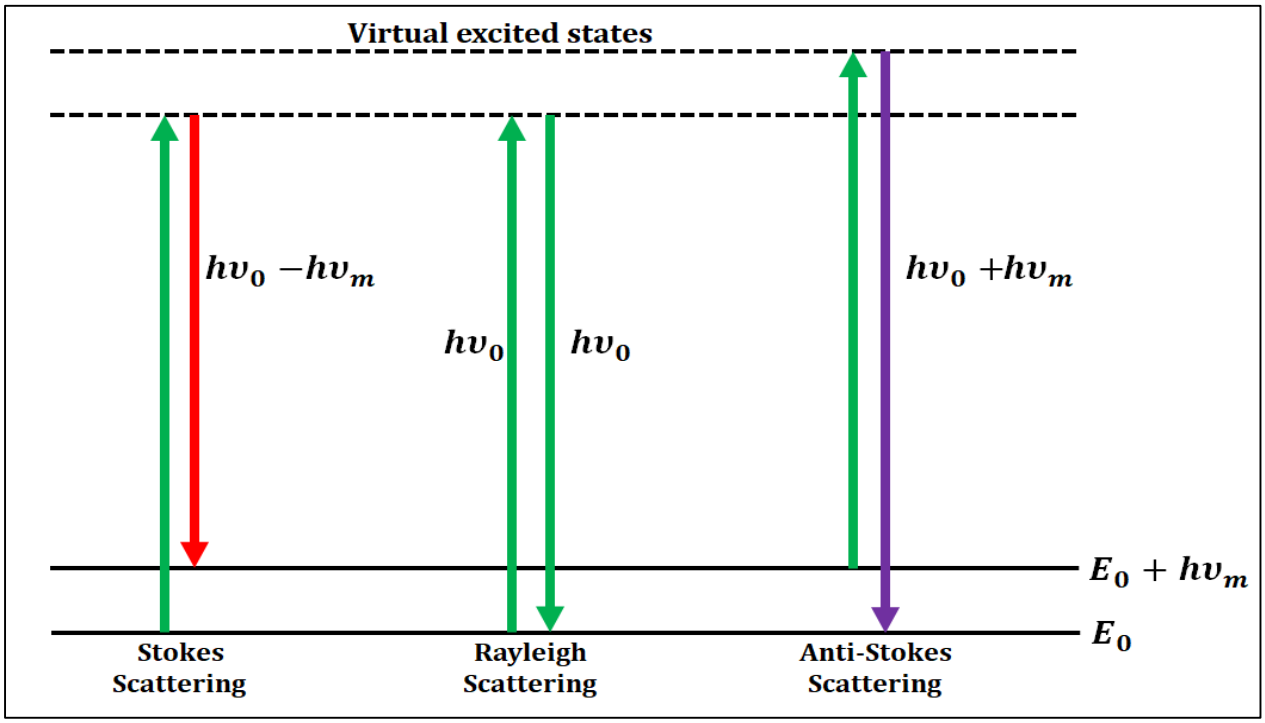

Figure 10: Energy level diagram for Rayleigh and Raman scattering, where hνm represents the difference in vibration energy levels [51].

II.2.2. Raman instrumentation

As infrared spectrophotometers, Raman instrumentation has two kinds of spectrophotometers which are dispersive and Fourier-transform. The most Raman spectrometers used in this research thesis belong to the dispersive one. The main components of the dispersive kind are simplified in Figure 11. This kind of Raman device is characterized by a diffraction grating

which works as a prism to disperse the light scattered from a sample and then is oriented toward the detector [52].

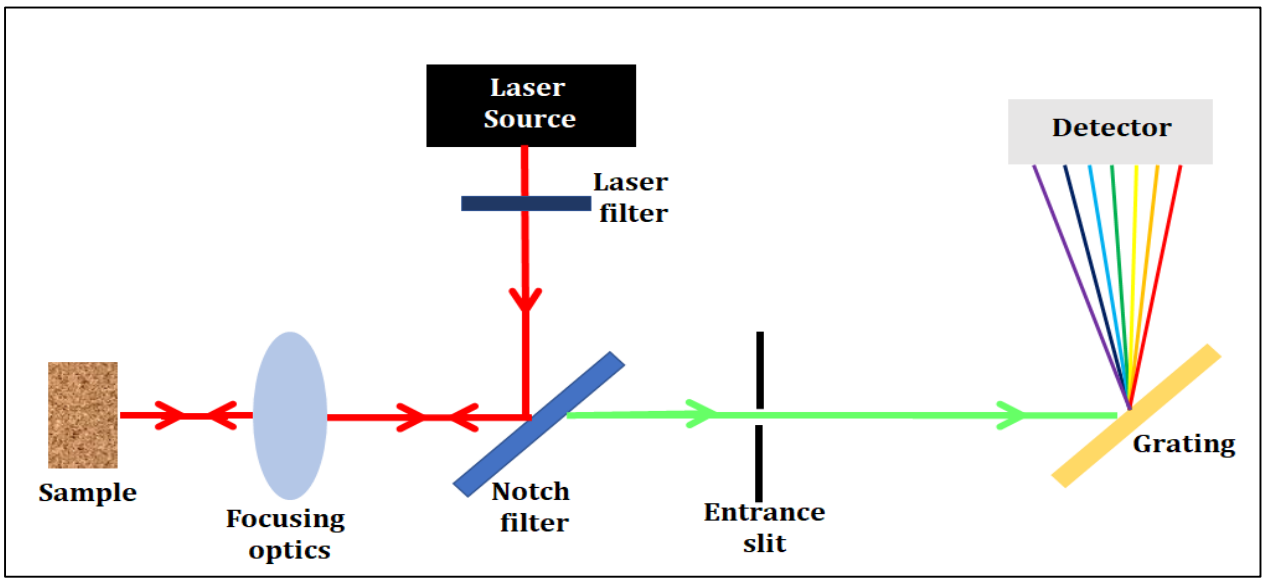

Figure 11: Schematic diagram of main components in a dispersive Raman spectrophotometer in a backscattering mode [53].

II.2.3. Spatially Offset Raman scattering (SORS)

Besides conventional Raman spectroscopy that is based on the spectra collected from backscattering, there is transmission and spatially offset Raman scattering modes [54]. Only conventional and spatially offset Raman scattering relevant to the scope of the research worked in this thesis are introduced. The main difference between conventional and SORS mode as simplified in Figure 12, is based mainly on the offset (ΔS) between the spot of the incident light and the spot of the collecting Raman signal on the sample surface for SORS mode. This offset allows the propagation of photons that emerge from deeper layers at the sample surface.

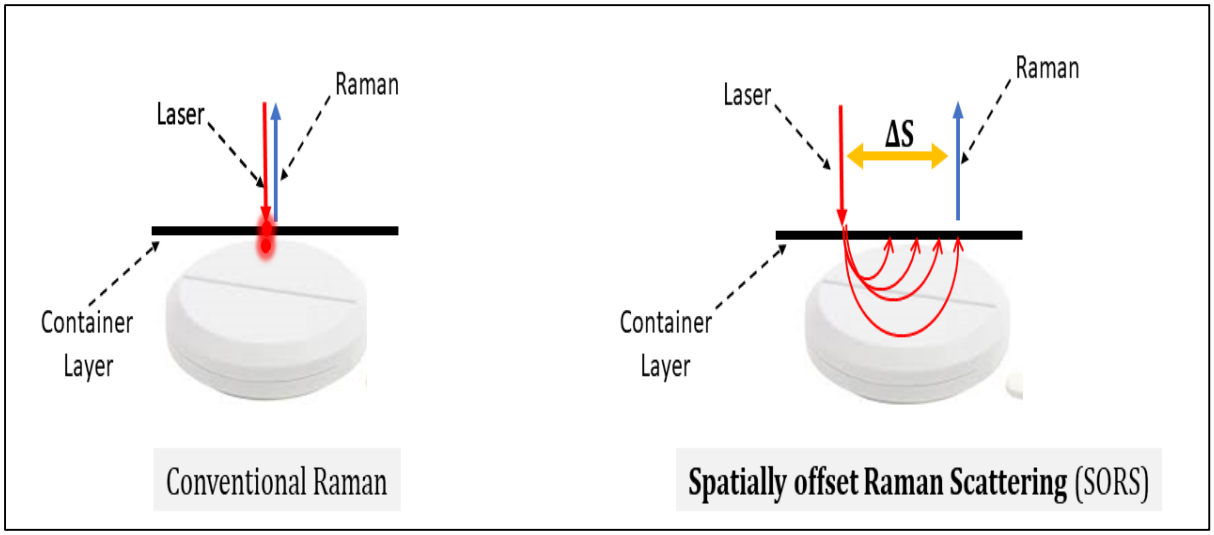

Figure 12: Schema of conventional backscattering and Raman spectroscopy and SORS mode adapted from [55] .

For example, in case of analyzing a tablet through polypropylene package using both of Raman modes, it is proven by the Figure 13 that SORS is able to obtain the spectrum of what is inside the packaging, whereas backscattering mode is impacted by the presence of polypropylene package and cannot pass through the package, thus the obtained spectrum belong to the packaging spectrum [56].

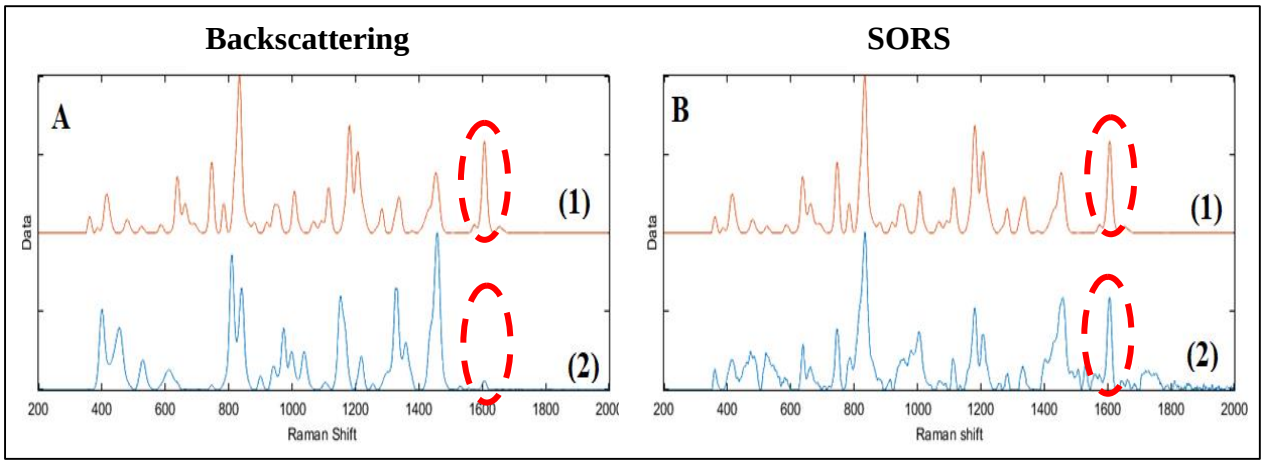

Figure 13: Raman spectra of a tablet (A): Backscattering (conventional mode); (B): SORS mode; (1): through glass vial; (2): through polypropylene packaging

II.2.4. advantages and disadvantages of Raman spectra

The main characteristics of Raman spectroscopy compared to FT-NIR and FT-MIR are summarized in Table 1. As infrared, Raman has advantages and limitations [57]:

Advantages of Raman spectroscopy:

- It is suitable for solids, liquids or gases.
- No need to sample preparation.
- Non-destructive.
- Fast obtaining Raman spectra.
- Can work with aqueous solutions unlike infrared spectroscopy has trouble with aqueous solutions because the water absorbs strongly the infrared light.
- Suitability of analyzing the sample through glass vials
- The ability to use fiber optic

Limitations of Raman spectroscopy

- Low sensitivity due the Raman effect is very weak, making from measuring low concentrations of a substance a real challenge, however, this problem can be overcome in case of using surface enhanced Raman scattering.
- Can be impacted by fluorescence.

Table 1: Comparison between MIR, NIR and Raman spectroscopy [58].

	MIR	NIR	Raman
Sample preparation	Variable	Very Simple	Very simple
Vibration	Stretching and bending	Overtones and combinations	Stretching and bending
Group frequencies	Fair	Good	Excellent
Aqueous solutions	Variable	Variable	Variable
Quantitative analysis			

III. Chemometric tools

Regarding the technological advances and especially the ability of analytical instruments, notably spectroscopic techniques, to record thousands of variables, often correlated, in a short period of time renders the use of univariate statistic tools inappropriate. Thus, it becomes important to use multivariate methods and algorithms that can handle the huge amount of the data. These multivariate data analysis methods are known by chemometric. Chemometric is a chemical discipline that involves statistical and mathematical tools to analyze chemical data. Among the most important application areas of chemometric, there exist: calibration and validation of multivariate models either for quantitation or discrimination, optimization of chemical measurements or experimental procedures and extraction of maximum physicochemical information from analytical data [59–61].

Several approaches of chemometrics relevant to the scope of the research worked in this thesis are introduced in the next sections.

III.1. Data structure

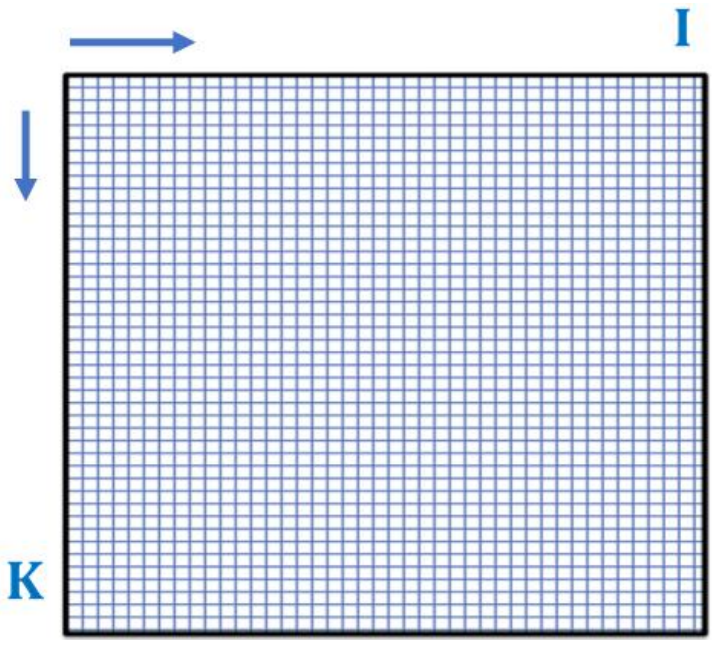

Figure 14: The data matrix or array of size I×K

Multivariate data is presented as an array (matrix) with size of I×K (I objects, K variables), whereas K variables are measured for I objects (Figure 14) that can be for instance pharmaceutical tablets. It has to be highlighted that the K variables of the data arrays could be homogeneous or a heterogeneous set. Homogeneous means that all K variables are given in the same units and generated by the same instrument. They could be for example wavelengths or wavenumber. One may also note that changing the position of these variables is not recommended since in most cases, these spectral variables have a natural order. In contrast to homogeneous variables, the heterogeneous do not have a natural order; they could be pH, thermal conductivity, electrical conductivity, temperature or blood pressure [62].

Visualization of data has always been useful in chemometrics to uncover the behavior of samples and the impact of variables, which can be achieved through using plots and figures.

III.2. Principal Component Analysis

Principal Component Analysis (PCA) is the multivariate tool used for exploratory data analysis. It is an unsupervised decomposition method that aimed at projecting the data X (I×K) from a high dimensionality space (dimension of K variables) to a lower dimensionality, which is defined by a set of new variables, called Principal Components (PCs). The first PC is the linear combination of the K original variables and explains the maximum variability of the samples (I), each following component in turn is orthogonal to the first PC, and explains the highest possible remaining variance. Each PC is composed of a vector of scores (t) and a vector of loadings (p) as it is shown in the following equation:

$$X = \sum_{f=1}^{F} t_f \, . p_f^T + E = TP^T + E$$

As it is shown in Figure 15, the original data X are modelled using f PCs after being centered, while the E represent the unmodelled part. Given that a set of PCs corresponds to the axes of the low-dimensional space the data are projected onto, the values of the score vector t_f represent the sample's coordinates on the f^{th} PC or axis, and the loadings vector p_f represents the contribution weights of the original variables to that PC [63].

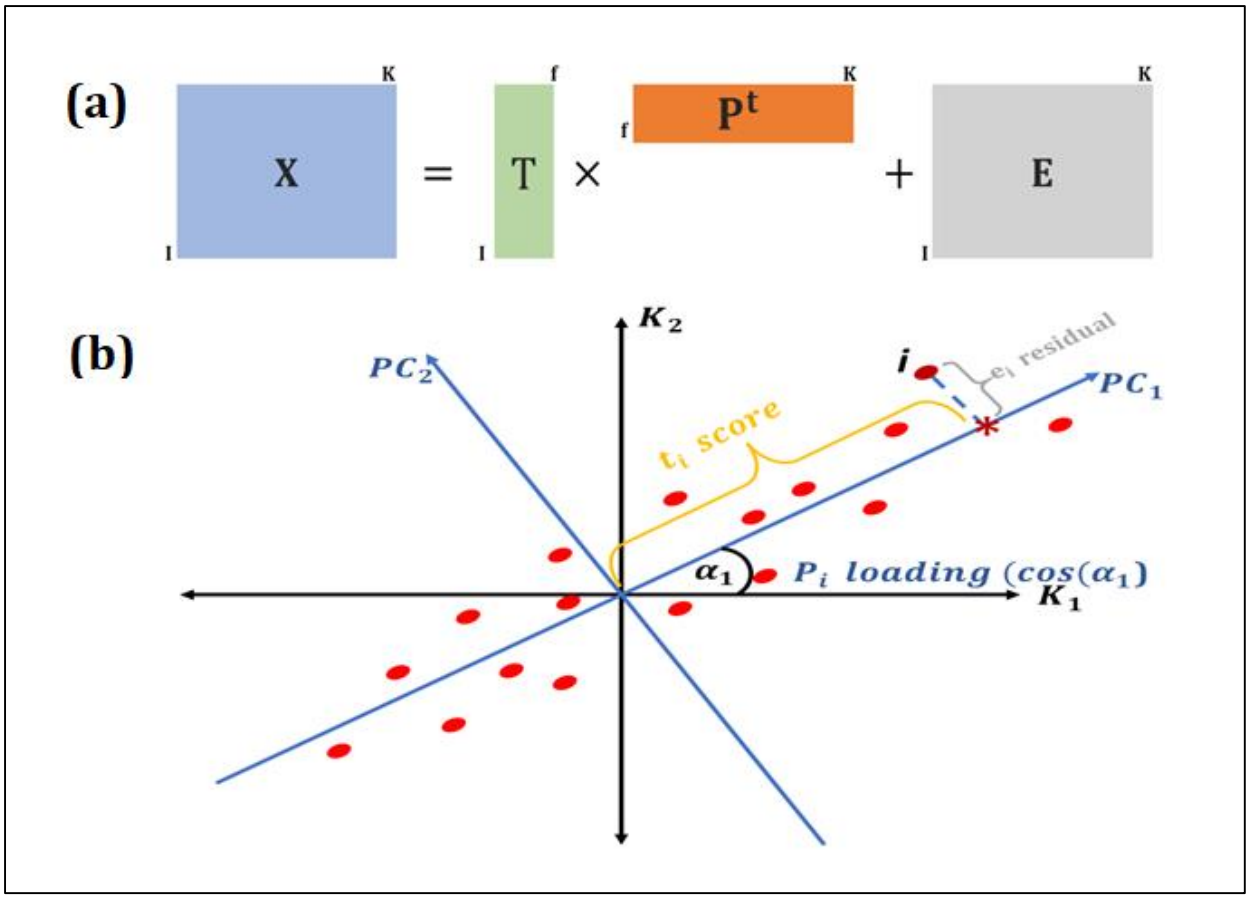

Figure 15: Graphical presentation of PCA in a space of 2D. (a): graphical of PCA equation. (b): projection of the data in the space of original variables (K1, K2) , the representation of principal components and schematic graphical elucidating the calculation of the scores and loadings. (adapted from [64])

Data explored by a PCA model can be visualized by the score plots. Distribution and behavior of samples as their similarities or outliers can be inspected with the score plots with representation of 2D or 3D of the modelled data. These score plots are visualized by plotting one score vector t against another. While the loading plots inspect the link between behavior of samples and the original variables, whose values on the different PCs (the values contained in the loading vectors p) also describe how the original variables influence each PC [63].

III.3. Outliers inspection

There are two statistics, which are known by Q statistic and Hotteling's T^2 statistic, are commonly used with PCA in evaluating the data and detecting outliers as it is shown in Figure 16.

The Q statistic measures the lack of model fit for each sample. It indicates how well each sample conforms to the PCA model through measuring the difference between the data point and its projection on the PC model. It gives the lack of fit to the model.

$$Q_i = e_i e_i^T$$

Whereas, the Hotteling's T^2 measures the variation within the PCA model. T^2 is the sum of the normalized squared scores defined as:

$$T_i^2 = t_i (T_K^T T_K)^{-1} t_i^T$$

Statistical limits can be developed whether for sample scores or hotelling's T^2 and Q residuals, and individual residuals. If a sample located outside the limits for a specific confidence level (for example 95%), this sample can be considered an outlier and thus is not representative of the data used to develop the PCA model [65,66].

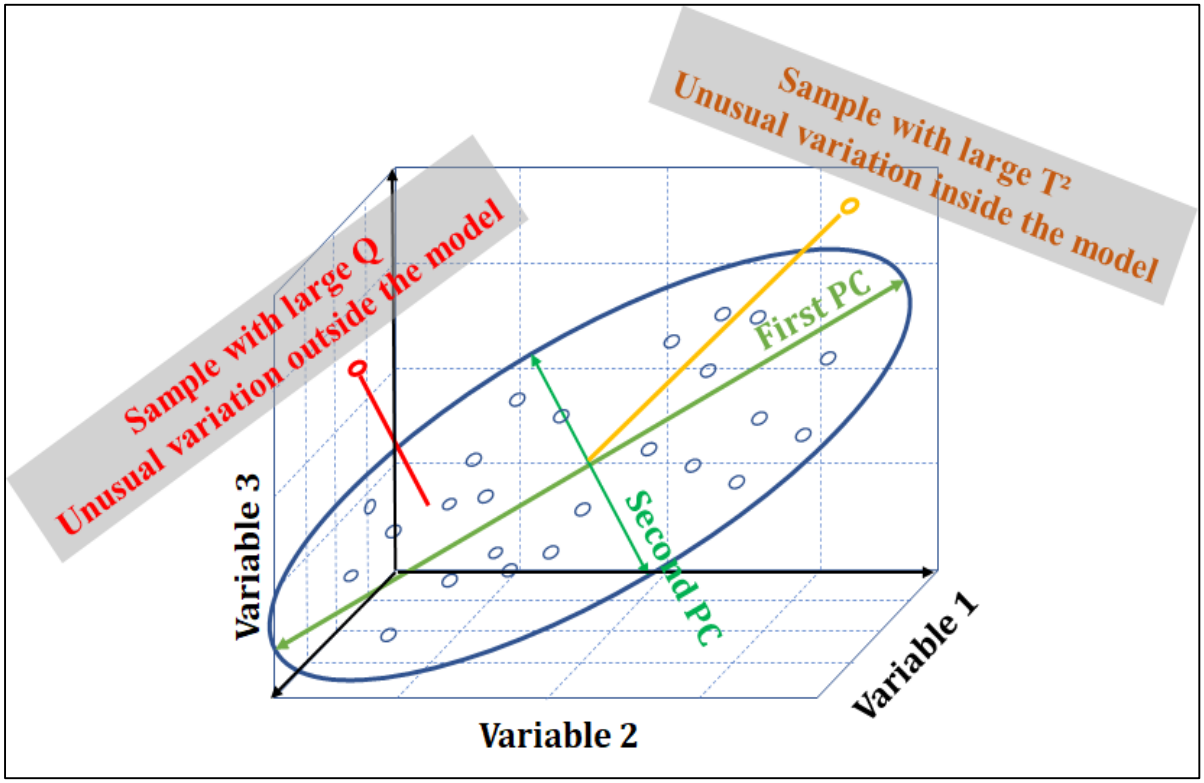

Figure 16: graphical presentation of PCA showing Q residuals and Hotelling's T^2 (adapted from [67])

III.4. Partial Least Squares (PLS)

Partial Least Squares regression (PLS-R) as Principal Component Regression (PCR) are both among multivariate regression tools that were developed to overcome the collinearity problem that exist in multilinear regression (MLR) [68]. PCR is based on extracting the PCs of X. Then, the least squares method is applied to model the responses data in Y with the PCs of X. the main limit of this multivariate regression, the extracted PCs explain the maximum

variance in X independently of the variance in Y. This independency between Y and the PCs of X may introduce errors in the estimation of the regression coefficients, mainly because useful information that correlates X and Y may be lost in the residuals, which are the PCs not included in the model, and thus leads to the problem of underfitting. Unlike PCR, in PLS, the PCs (which are known by latent variables) not only explain X, but also Y. The latent variables are calculated to explain the maximum covariance between X and Y as mentioned in following equation and shown in Figure 17 [69,70].

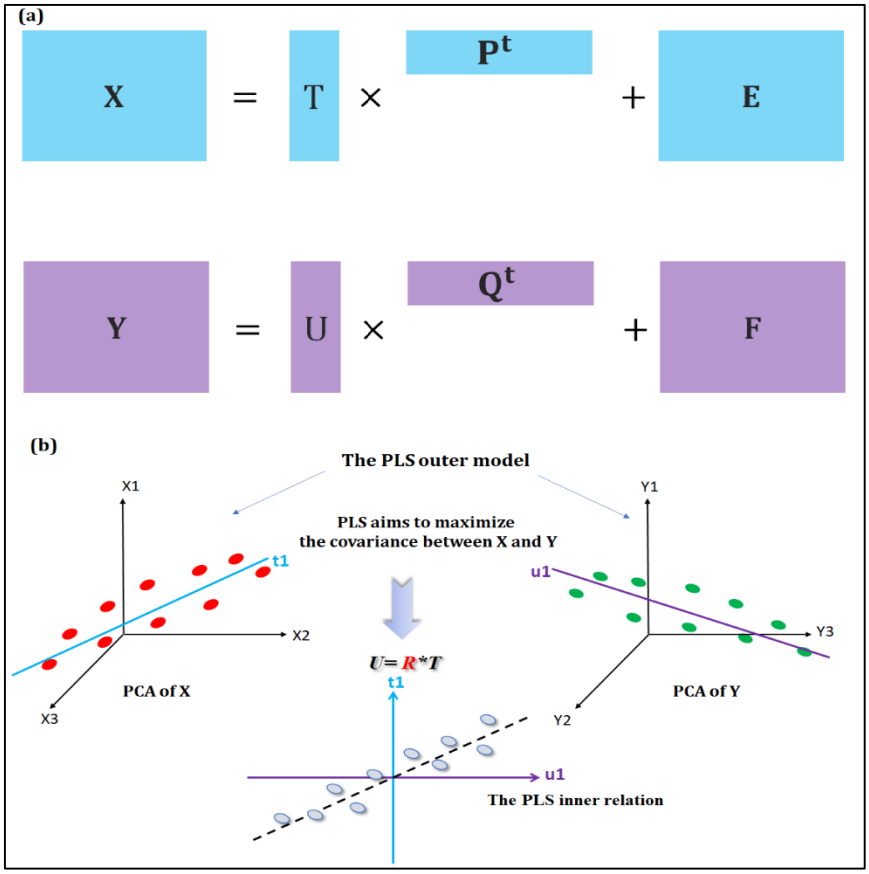

Figure 17: (a) PLS decomposition of X- and Y-blocks. (b) Illustrative presentation of the core idea of PLS(adapted from[71])

$$X = TP^T + E$$
$$Y = UQ^T + F$$

In fact, the PLS algorithm NIPALS (**N**onlinear **I**terative **P**artial **L**east **S**quares) is applied to maximize the covariance between the scores of T and U through the rotation of loadings P and Q. This algorithm is applied following several steps as summarized in Figure 18 [72,73].

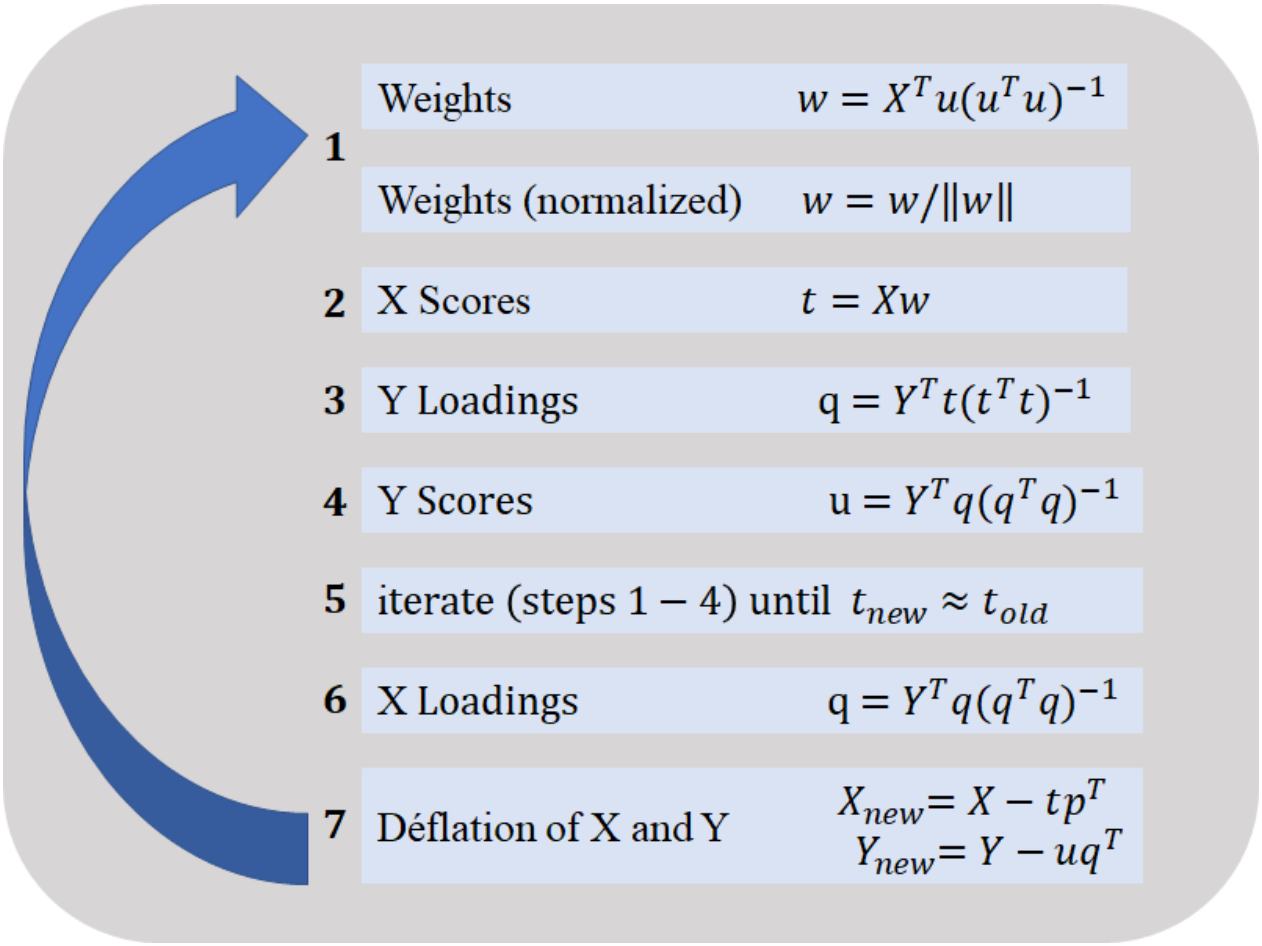

Figure 18: PLS-NIPALS algorithm (adapted from [74])

The regression coefficients B that links the predictor data in X with responses data in Y are obtained by:

$$\boldsymbol{B_{PLS} = W(P^T W)^{-1} Q^T}$$

where W is a matrix containing the weights obtained of each latent variable, P^T and Q^T are matrices of the loadings of the X- and Y-blocks respectively. Predicted responses $\mathbf{\hat{Y}}$ from a PLS model are then obtained multiplying a predictor matrix X with the PLS regressions coefficients B

$$\boldsymbol{\hat{Y} = XB_{PLS}}$$

III.5. Partial Least Squares-Discriminant Analysis (PLS-DA)

Partial least squares discriminant analysis is derived from PLS regression (PLS-R) and aims to develop a regression model between the X and Y. Whereas in PLS-R "Y" is a set of continuous numbers such as the concentration of an analyte, the Y of PLS-DA has discrete numbers [75]. in case of two levels, one level for what is sometimes called a target group (A)

takes values of +1 and the set for the rest of the data, sometimes called the alternative group (B), and have values of 0 as illustrated in Figure 19a. Partial least squares discriminant analysis can be considered as a linear two class classifier which aims to establish a line that separate the space into two regions [76,77]. Figure 19b illustrates a possible discriminant for two groups; samples to the right belong to the group A represented by red stars and samples to the left to the group B represented by yellow rhombus.

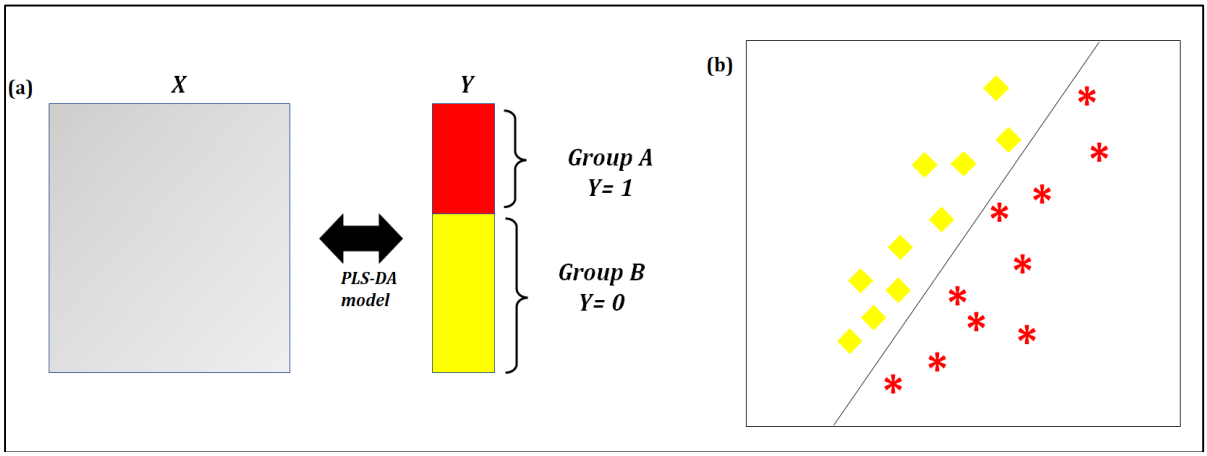

Figure 19: (a): Partial least squares discriminant analysis (PLS-DA) model for two groups, (b): a two class linear discriminator for two groups (adapted from [76].

The PLS-DA model is extended in case of there are more than two classes hence Y become a matrix of more than two columns. Each column represents a class, and each sample is considered whether to belong of the right class (Y = +1) or not (Y = 0). Thus, if there are three groups A, B, and C, and for example if the second class represents the target class B; sample set of A and C will have values of 0, and value of + 1 will be attributed to the samples of class B. the class denoted by Y= +1 will be called the target class representing one of the original groups ,while A and C denoted by value of 0 will be called alternative class as shown in figure 20.

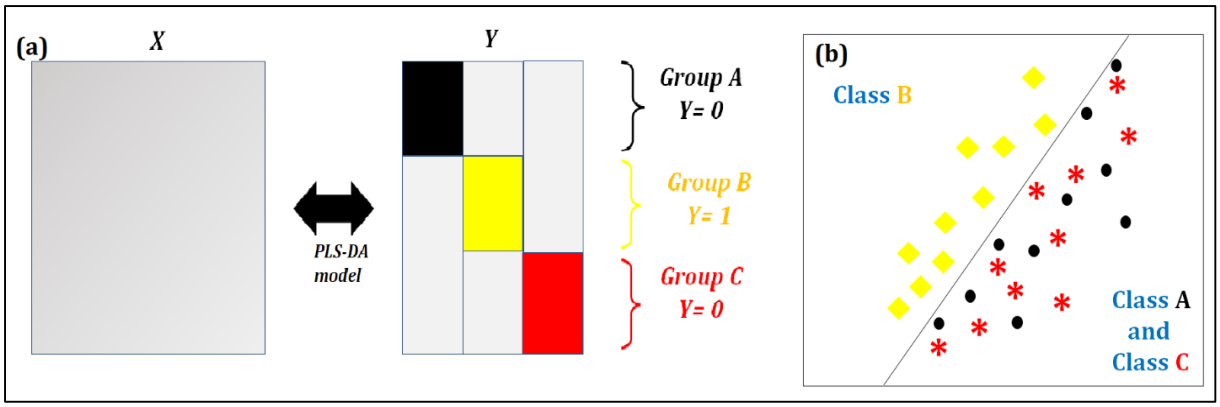

Figure 20: (a)Partial least squares discriminant analysis (PLS-DA) model for three groups, (b): a two class discriminator for three groups (adapted from [76]).

III.6. Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS)

Multivariate Curve Resolution (MCR) is a decomposition method that can extract pure components from samples , that are introduced in a given data matrix (D) [75]. Although MCR and PCA are both bilinear decomposition methods, the components of MCR are not forced to be orthogonal. This characteristic allows extracted components to have a chemical meaning, however, allowing components not to be orthogonal leads to the rotational ambiguity challenge, which does not provide a unique solution [79,80]. Hence to overcome this challenge, it is necessary to constraint the components using the approach of alternating least squares (ALS) [81]. MCR-ALS aims to obtain pure spectra and their concentration profile in each sample. Therefore PCA scores matrix (T) and loadings (P) are considered the pure concentrations matrix C and the pure resolved spectra matrix S respectively, and the MCR-ALS model can be written as follow and introduced graphically in Figure 21 [82] :

$$D = CS^T + E$$

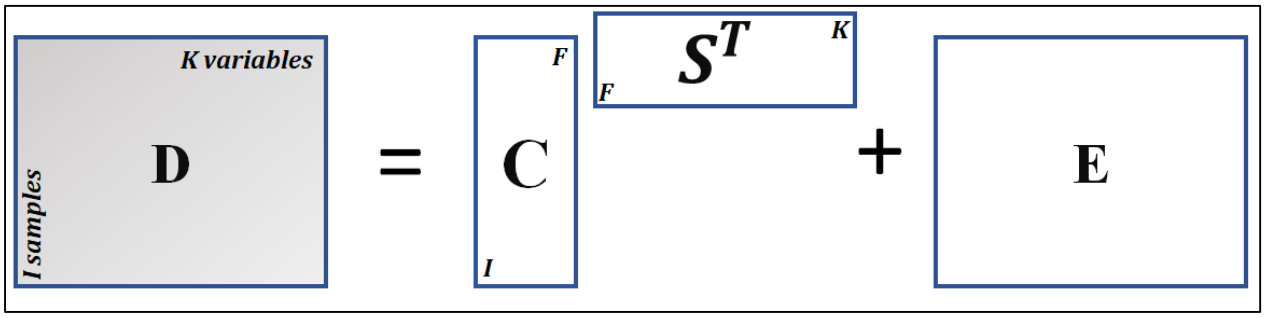

Figure 21: Graphical presentation of MCR-ALS model [79]

The process of MCR-ALS consists of several steps to achieve chemical MCR-ALS solutions as it is summarized in Figure 22.

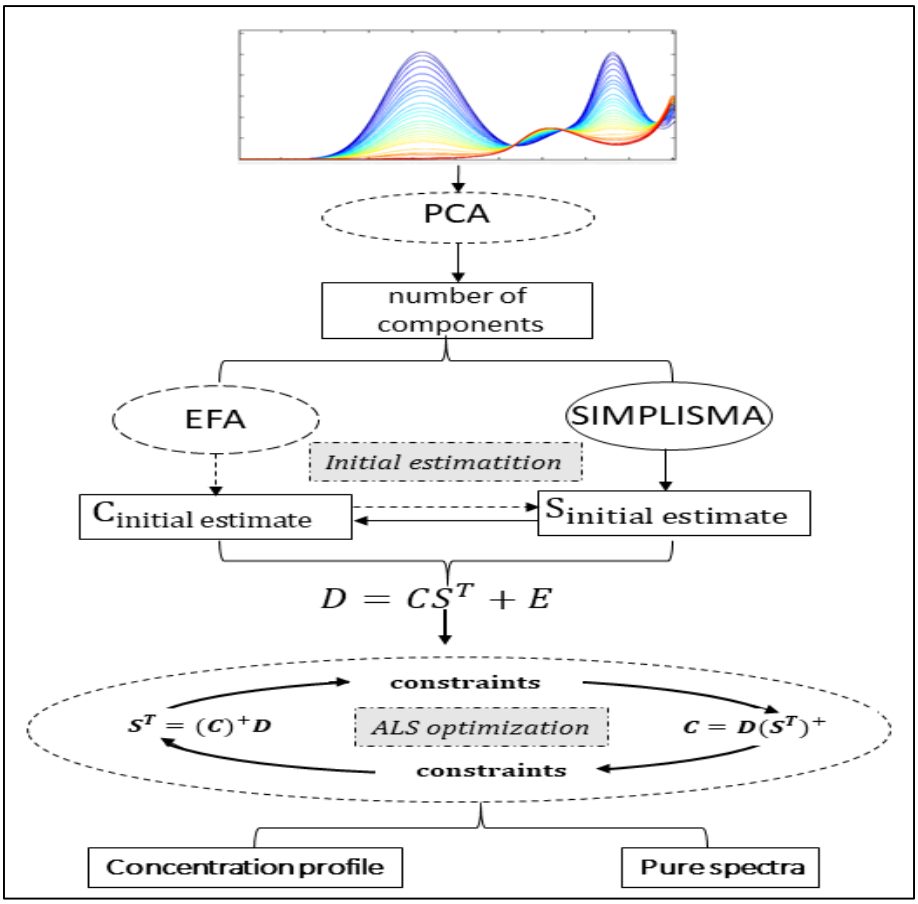

Figure 22: Overview of the steps in the multivariate curve resolution-alternating least squares (adapted from [78])

The first step of this process consists of determination the number of components present in the samples. this step can be achieved whether by the prior knowledge of the chemical compounds that exist in the samples or by using the algorithm of singular value decomposition [83].

Then the second step aims to determine the initial estimates of both concentration (C) and pure spectra profiles (S) using algorithms as evolving factor analysis (EFA) [84] or simple-to-use interactive self-modeling mixture analysis (SIMPLISMA) [85]. These initial estimates are iteratively optimized by ALS procedure using the two least squares matrix calculations:

$$C = D(S^T)^+$$

$$S^T = C^+ D$$

Where $S^+$and $C^+$are the pseudoinverse matrix of S and C respectively. These two previous equations are repeated in an iterative way until the new computed D matrix shows convergence to the experimental data matrix D. The convergence may be used whether by a

preselected number of iterative cycles or by the comparison of lack of fit values obtained in two consecutive iterations. The lack of fit is calculated according to the expression:

$$\%LOF = 100 \times \sqrt{\frac{\sum_{ik}(d_{ik} - d_{ik}^{*})^2}{\sum_{ik} d_{ik}^2}}$$

Where d_{ik} is the experimental measured value of a sample i at wavelength k and d_{ik}^{*} is the obtained value with the MCR-ALS model.

As previously mentioned and illustrated in Figure 22, constraints can be applied with ALS optimization to obtain a unique solution and overcome the problem rotational ambiguity. Among of these constraints that have been applied in this thesis are:

- Non-negativity constraint: this constraint is applied on concentration profiles due to the fact that the concentration of chemical analytes cannot be negative and admit always values equal to or larger than zero. This constraint can be applied also to spectral profiles as long as they are not preprocessed [86].

- Correlation constraint: This constraint is applied only on the concentration profile whenever MCR-ALS is used for quantitative analysis of an analyte in the presence of unknown interferences [87–89].

  This correlation constraint consists during the ALS optimization to correlate concentrations of a particular analyte in calibration samples C_{cal}^{ALS}obtained by MCR-ALS with previously known reference concentration values of the analyte $\mathrm{C_{ref}}$ in these samples. A linear model is then developed between the values C_{cal}^{ALS} and $\mathrm{C_{ref}}$.

$$\hat{\mathrm{C}}_{\mathrm{ref}} = bC_{cal}^{ALS} + b_0 + \mathrm{e_{ref}}$$

b and b_0 are the slope and offset values which better fit $C_{cal}^{ALS}$to $\mathrm{C_{ref}}$, obtained by least squares linear regression, and $\mathrm{e_{ref}}$ is the error in the reference concentrations. The corresponding concentration values of these calibration samples calculated using this model are:

$$\hat{\mathrm{C}}_{\mathrm{cal}} = bC_{cal}^{ALS} + b_0$$

And to predict the unknown concentration of the analyte in the new prediction samples $\mathrm{C_{unknown}}$, the equation used is:

$$\hat{\mathrm{C}}_{\mathrm{unknown}} = bC_{unknown}^{ALS} + b_0$$

b and b_0 are the values obtained previously in the calibration step from C_{ref}, and $C_{unknown}^{ALS}$ are the concentrations of the samples predicted by ALS. Each ALS iteration is then completed after updating the obtained values of prediction by substitution of $C_{unknown}^{ALS}$ by $\hat{C}_{unknown}$.

III.7. Validation of models

The quality of a predictive model depends on the quality of the data used in its construction, and on its representativity over the system analyzed, for example if the modelled data can represent all, or the most significant, sources of variability in a system. PCR or PLS, require that a specific number of Principal Components (PCs) or Latent Variables (LVs) are used to model the data. If too few are used, the model will not explain enough variance in the studied variables, resulting in underfit models. On the other hand, if too many are used, it will result in overfit models.

To overcome these latter problems of underfitting and overfitting The typical approaches for validation of a 'soft model' (PCR or PLS) are cross-validation, that aims to determine the optimal latent variables used to develop PLS model, based on the root mean square error of cross-validation (RMSECV) as it is shown in the Figure 23.

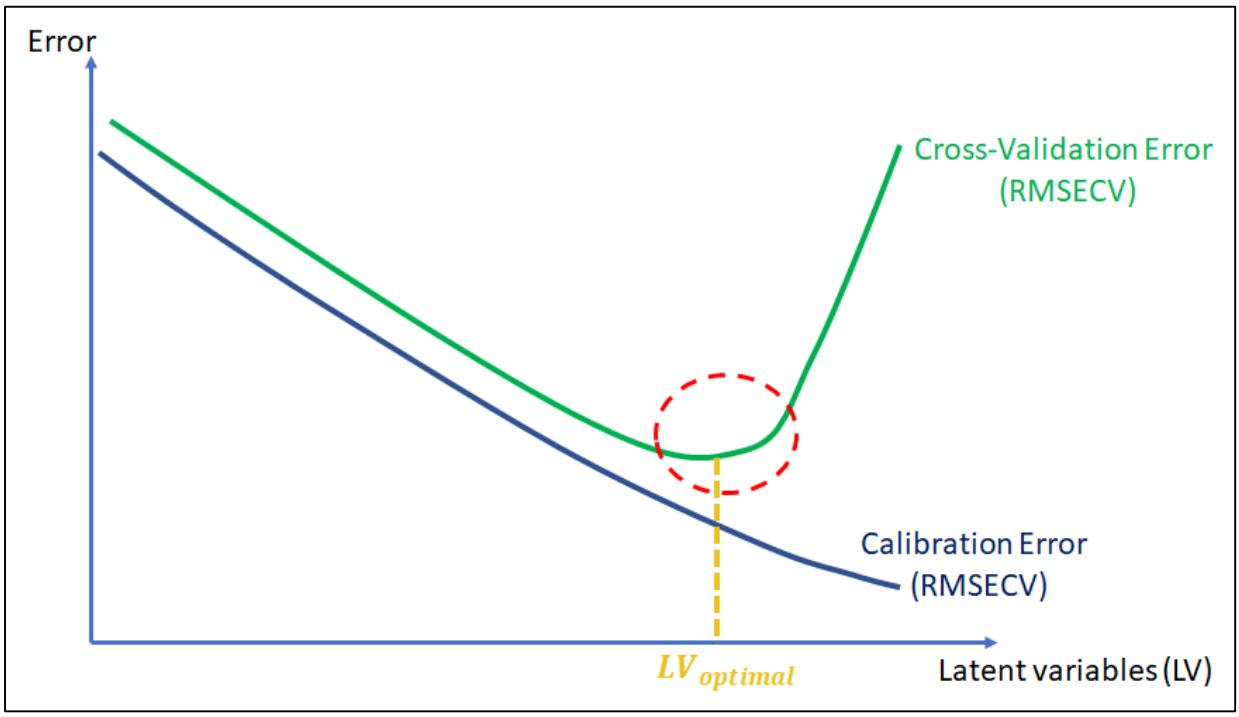

Figure 23: The scheme of determination of optimal latent variables (LV) based on the lowest error of cross validation

Cross validation can be applied by many ways, but in this thesis two kind of cross validation that have been used. As shown in figure 24:

- Leave one out: Each single object in the data set is used as a test set.

- Venetian blinds: Each test set is determined by selecting every sth object in the data set, starting at objects numbered 1 through s.

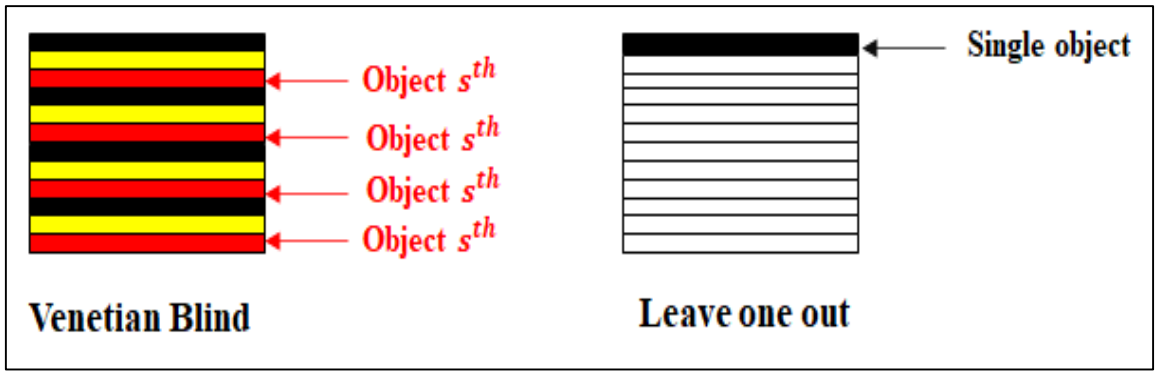

Figure 24: Cross validation methods: Venetian blinds and leave one out

Validation with external test set is considered the best method validation approach for any mathematical model as regards prediction based on experimental data [77,90]. The following expressions were used to express and evaluate the validation results:

Root mean square error of prediction: $RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(c_i-\hat{c}_i)^2}{n}}$

Standard error of prediction (SEP): $SEP = \sqrt{\frac{\sum_{i=1}^{n}(c_i-\hat{c}_i-Bias)^2}{n-1}}$

Bias (systematic error): $Bias = \frac{\sum_{i=1}^{n}(c_i-\hat{c}_i)}{n}$

Where c_i and $\hat{c}_i$arethe known and calculated analyte concentration in sample i, and n is the total number of samples considered in the validation. Also in order to evaluate the quality of the obtained results of the concentrations predicted by the application of the mathematical models, for a particular analyte using n samples, the relative error in the predicted concentrations, in percentage (RE%), is calculated as:

$$RE(\%) = 100\sqrt{\frac{\sum_{i=1}^{n}(c_i-\hat{c}_i)^2}{\sum_{i=1}^{n}{c_i}^2}}$$

III.8. Accuracy profile

Besides RMSEP, an accuracy profile is considered an approach to evaluate the accuracy of different developed model over the investigated concentration range. In this thesis, the main objective of using the accuracy profile was to evaluate the validity of spectroscopic methods to carry out the quantitation of API in the presence of matrix effect.

The accuracy profile as illustrated in Figure 25 is a graphical tool that is based on the total error including the systematic and random error. This approach of accuracy in contrast to

RMSEP, aim to plot the β-expectation tolerance interval for each concentration which represent by the blue dashed lines in Figure 23, whereas the acceptance limits represent by dashed black lines are determined according to the regulatory requirements such as ±5% for the analysis of pharmaceutical formulations and ±15% and ±10% in Figure 23 for content uniformity assessment. The systematic and the random error are represented respectively by dashed lines and dispersion of the measurement (green, blue and red points). The gray area represent the dosing range where the method is valid. The validity of the method depends on the β-expectation tolerance interval that has to fall within the acceptance limits[91,92].

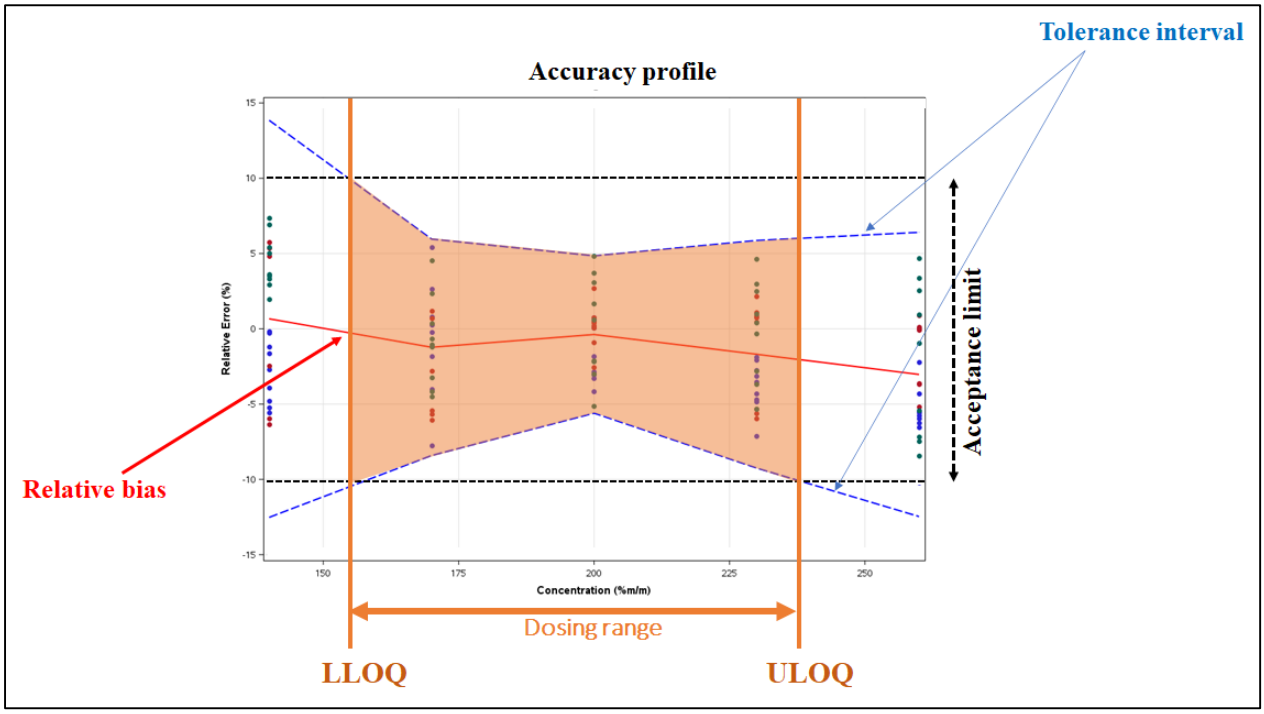

Figure 25: illustration of the graphical tool of Accuracy Profile. The red line represents the relative bias, the dashed blue lines represents β-expectation tolerance limits (β=95%) whereas the dashed black lines are the acceptance limits (±10%)

III.9. Preprocessing techniques

The goal of data preprocessing methods is to obtain the purest spectra as possible from data, which, is important to set up of correlations between dependent variables, such as concentrations of analytes, and the respective pure independent variables of analytical instruments as spectra, and to allow extracting the maximum information from the data. preprocessing of spectral data is the most important step that is usually applied before multivariate modeling as Principal Component Analysis (PCA) and Partial Least Squares (PLS). the spectra are often affected by several factors other than the chemical components of interest, and if the effects from these unwanted factors are not eliminated or at least

minimized, they can create challenges when it comes to carry out the analysis whether qualitatively or quantitatively. These effects could be for example due to the information from instrumentation, light scattering, particle-size distribution, sample density, effect of tablet face and position in relation to a probe beam [93].

Some preprocessing methods that have been used in this research work of the thesis are described briefly below.

- *Mean Centering* [77] removes the absolute intensity information from each of the variables, hence allowing the multivariate methods to focus on the response variations about the mean. This preprocessing technique calculate the mean response of each of the variables over the samples in a data matrice X.
- *Standard Normal Variate (SNV)* [94] is commonly used on spectral data to correct for multiplicative variations between spectra, SNV correction involves subtraction of the mean intensity from each of the variable intensities, followed by division of the resulting values by the standard deviation.
- *Multiplicative Signal Correction* [95] as SNV, it has been applied to deal with additive and multiplicative variations that exist in many spectroscopic applications. However, whereas SNV is carried out based on each individual spectrum, MSC is applied based on the training set. MSC correction is done by calculating the slope, which represent a multiplicative correction factor, and offset, which is an additive correction factor. Then the corrected spectra is obtained by the subtraction of the offset from each variable in all the sample spectra and divided by the slope.

  Figure 26 shows the results obtained when MSC and SNV preprocessing are applied on the same spectra of fluconazole tablets. Regardless these corrections are somewhat different, in most practical cases they perform similarly when these variations are present in the data.

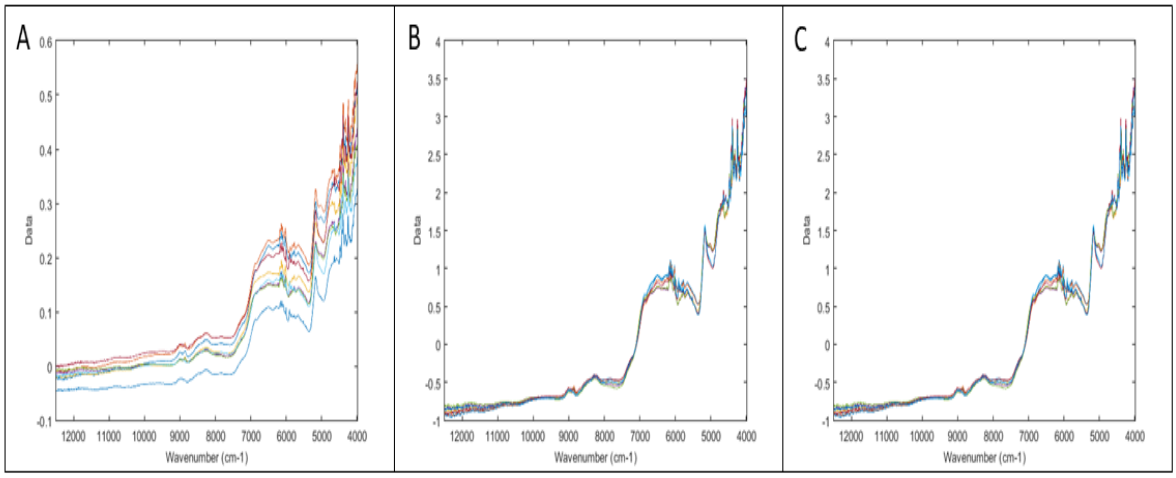

Figure 26: (A): Raw spectra of fluconazole, (B): preprocessed spectra of fluconazole by SNV, (C) : preprocessed by MSC

- *Derivatives* [96] have the ability to remove additive and multiplicative effect from the spectral data. Figure 27 illustrates an example of a Gaussian peak (blue line), with added baseline (red line), and with added baseline and multiplicative effect (black line). First derivative removes the baseline whereas second derivative removes the baseline and the linear trend.

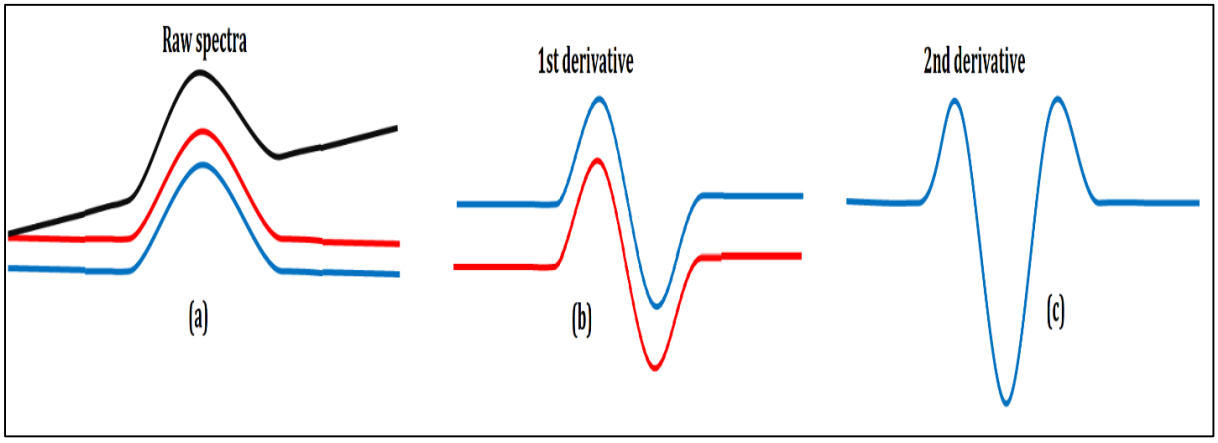

Figure 27:the effect of derivative preprocessing. (a): Raw spectra, (b): 1st derivative, (c): second derivative [97]

In practice, before applying derivatives to the data, smoothing and fitting into a low-order polynomial within a data window is often used in pre-processing methods as is the case of the algorithm of Savitzky-Golay. Since application of derivation should be done with care as this may add noise to the spectral data as it is shown in Figure 28 which elucidate the impact of 1st derivative and 2nd derivative on the ibuprofen in a ternary mixture.

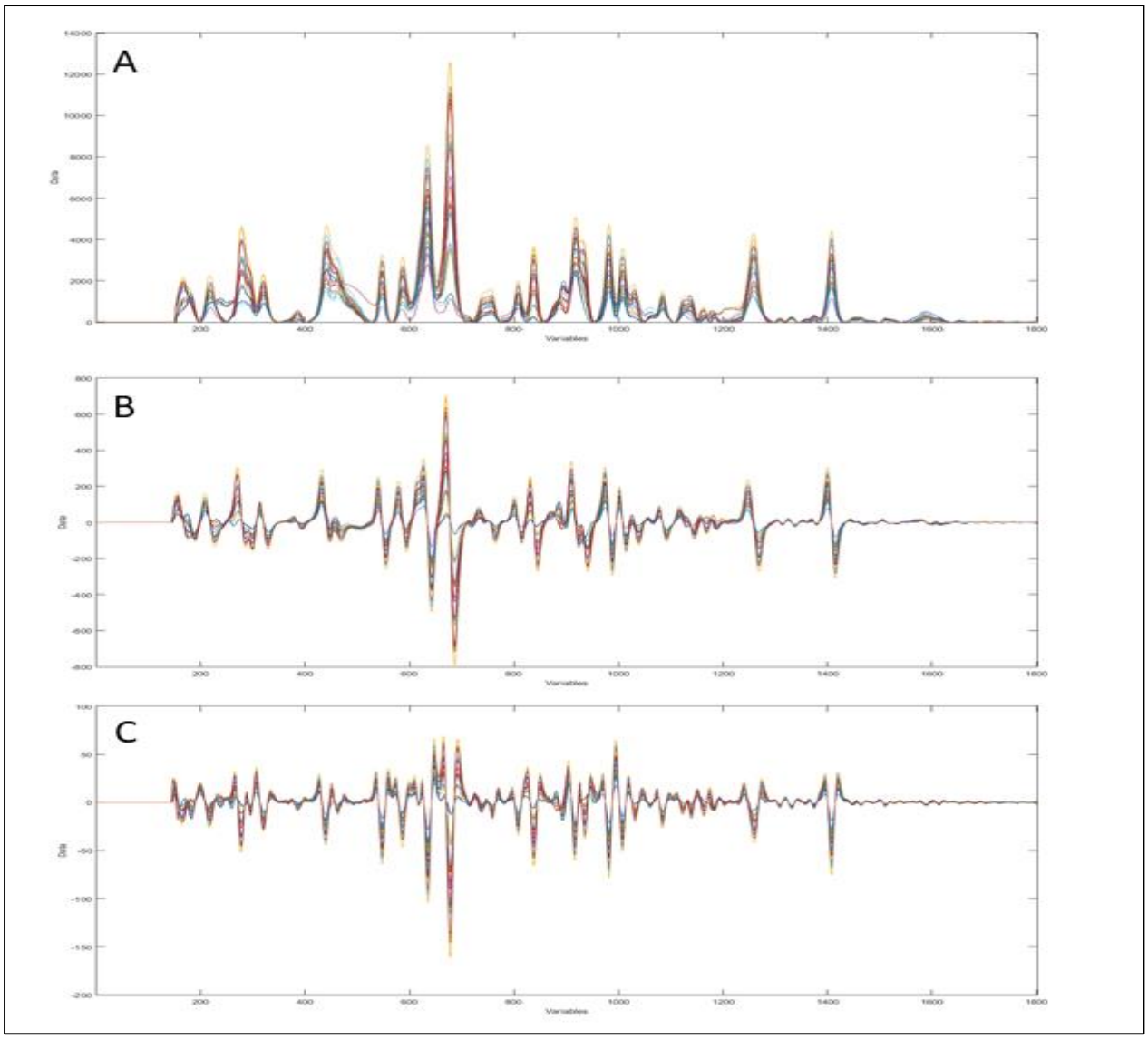

Figure 28: the effect of preprocessing of Raman spectra of ibuprofen in a ternary mixture. (a): Raw spectra, (b): 1st derivative, (c): second derivative

Chapter II

Experimental Results

1st Aspect

Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy.

Preamble

Identification and characterization of polymorphic forms of the active pharmaceutical ingredient (API) in pharmaceutical products is important to ensure their therapeutic effects. There exist many ways to accomplish these objectives. One of these ways is coupling vibrational techniques as mid-IR and near-IR to chemometric techniques such as PLS-DA for discriminating different polymorphic forms. The main challenge appears when it comes to apply PLS-DA models on new pharmaceuticals that do not have any of polymorphic forms as in case of falsified medicines or pharmaceuticals that contain two or more polymorphic forms as in case of contamination or polymorphic transformation, which can be considered as a kind of matrix effect. In these situations, PLS-DA is unable to do the right discrimination and the solution appears in associating PLS-DA with the approach of Hoteling's T^2 and Q residuals that is able to detect this kind of samples as outliers and reject them before applying PLS-DA.

Summary

The main goal of this work was to test the ability of vibrational spectroscopy techniques to differentiate between different polymorphic forms of fluconazole in pharmaceutical products. These are mostly manufactured with fluconazole as polymorphic form II and form III. These crystalline forms may undergo polymorphic transition during the manufacturing process or storage conditions. Therefore, it is important to have a method to monitor these changes to ensure the stability and efficacy of the drug.

Each of FT-IR or FT-NIR spectra were associated to partial least squares-discriminant analysis (PLS-DA) for building classification models to distinguish between form II, form III and monohydrate form. The results have shown that combining either FT-IR or FT-NIR to PLS-DA has a high efficiency to classify various fluconazole polymorphs, with a high sensitivity and specificity. Finally, the selectivity of the PLS-DA models was tested by analyzing separately each of three following samples by FT-IR and FT-NIR: lactose monohydrate, which is an excipient mostly used for manufacturing fluconazole pharmaceutical products, itraconazole and miconazole. These two last compounds mimic potential contaminants and belong to the same class as fluconazole. Based on the

plots of Hotelling's T^2 vs Q residuals, pure compounds of miconazole and itraconazole, that were analyzed separately, were significantly considered outliers and rejected. Furthermore, binary mixtures consist of fluconazole form-II and monohydrate form with different ratios were used to test the suitability of each technique FT-IR and FT-NIR with PLS-DA to detect minimum contaminant or polymorphic conversion from a polymorphic form to another using also the plots of Hotelling's T^2 vs Q residuals.

1. Introduction

Polymorphism is a characteristic where a drug substance can present one or more crystalline form due to different molecule arrangements; thus can include different solid varieties of crystalline forms. Sometimes, the crystal form is known by solvate when it has amount of solvent and it is known as hydrate if the solvent is water [98-101].

Each polymorphic form has different physico-chemical properties [102]. Variations at the level of physico-chemical properties could have an impact on dissolution rate and bioavailability, hence, the therapeutic effect of the drug substance might be influenced [103]. The manufacturing process and the storage conditions are considered as the main factors that have an impact on polymorphic transformation [104,105]. Thus, looking for a reliable analytical technique to control and monitor polymorphs of drug substance in drug products is mandatory to ensure the quality of pharmaceutical products.

The analysis of polymorphic forms of drug substance has been carried out by both destructive and non-destructive techniques. Destructive techniques are represented in differential scanning calorimetry (DSC). The main challenge of this technique is the interconversion of polymorphic forms of the drug substance that could be occurred during the analysis [9]. Non-destructive techniques are summarized in X-ray powder diffraction (XRPD) and vibrational spectroscopic techniques [106,107]. The principal advantage of these techniques is that they often do not need any sample preparation. Hence, they are fast in analyzing and acquiring results [108]. Nevertheless, these last cited techniques have some limitations. The main limitation of XRPD is that the morphology of the particle may impact the accuracy of quantitative analysis using XRPD [109]. The main challenge of vibrational spectroscopy is identifying and discriminating between polymorphic forms directly especially in the presence of the matrix that may hamper the identification of fingerprints related to the identity of the polymorphism [110,111]. Thus, associating spectroscopic techniques with chemometric tools is important to uncover more details about polymorphism.

Chemometric tools have already proven their usefulness to discriminate and quantify polymorphic forms such as PCA and PLS and reduce systematic variations by using preprocessing techniques such as standard normal variate (SNV) or multiplicative scatter

corrections (MSC) [112]. For example, polymorphic forms were quantified accurately by coupling either of FT-NIR Raman or FT-IR to Partial Least squares regression (PLS-R) [113]. Another example, is applying PCA to FT-IR data that allows detecting which of the four polymorphic forms of cimetidine was present in a pharmaceutical product [114]. In addition to PCA and PLS, another chemometric tool called multivariate curve resolution (MCR) can detect how many polymorphic forms exist in a mixture and identify their pure spectra. This has been combined to FT-IR for following the polymorphic interconversions of cimetidine [115]. Besides that, PLS-DA has been used successfully in food and other applications [116,117]. These applications of chemometric techniques nicely illustrate their efficiency to extract relevant information from the raw spectral data.

Fluconazole, 2-(2,4-difluorophenyl)-1,3-bis(1-H-1,2,4-triazol-1-yl)propan-2-ol, is an antifungal triazole. Fluconazole is used to treat superficial Candida infections. It is used for acute therapy of disseminated Candida, for systemic therapy of blast mycosis and histoplasmosis, for dermatophytic fungal infections, and for prophylaxis in neutropenic patients [118-119]. According to recent studies [120,121], fluconazole displays three main polymorphic forms: form I, II and III as well as a monohydrate form. The most stable polymorph is form III. This form is a convert form from the metastable form II, that may convert to the monohydrate form during the storage or compression under specific conditions of humidity and temperature. At the moment, the most marketed forms by Moroccan pharmaceutical industries are form II and form III while polymorphic form I is considered as unstable based on the recent study [123].

The main objective of this work was to evaluate the qualitative abilities of each vibrational techniques of FT-IR and FT-NIR to classify polymorphic forms of fluconazole in pharmaceutical products as well as investigate the suitability of both vibrational techniques to detect whether there exist any polymorphic conversion or falsified pharmaceutical product of fluconazole

2.Material and methods

2.1. Instrumentation

The FT-IR spectra were acquired in the reflectance mode in the spectral region of 4000- 650 cm-1, with an average of 32 scans at resolution of 4 cm-1, using a Frontier FT-IR spectrometer (Perkin Elmer, Waltham, USA) equipped with a diamond crystal ATR device.

For each measurement, a fraction of sample is placed onto the diamond surface. the diamond surface was cleaned with acetone and dried between each analysis. The cleaning of the diamond surface was checked spectrally.

The NIR spectra were obtained using FT-NIR spectrophotometer MPA Multi-Purpose FT-NIR Analyzer (Bruker Optics, Ettlingen, Germany) in diffuse reflectance mode in the spectral region of 12500-4000 cm-1 at resolution of 8 cm-1. The average of 32 scans was acquired for each sample by placing the optical fiber probe on the bottom of the glass vial that contains the sample.

2.2. Sample preparation

The main samples that were acquired to build a dataset were:

- Fluconazole polymorphic form II (TCI- Chemicals, Belgium), fluconazole polymorphic form III (Sigma-Aldrich, Belgium), miconazole (Sigma-Aldrich, Belgium), itraconazole (Sigma-Aldrich, Belgium) and lactose monohydrate (Sigma-Aldrich, Belgium). Fluconazole monohydrate form was obtained based on a reported recrystallization technique [124]. This recrystallization method was carried out by dissolving the fluconazole form- II in milli-Q water under constant stirring at 40° C. The saturated solution was filtered to remove all nuclei, and the filtered solution was cooled in a refrigerator at 5° C. The resulting crystals were rapidly surface dried only, then the polymorphic form of fluconazole monohydrate was checked with FT-IR and the obtained spectrum was compared to the spectrum of previous studies [122]. The pure polymorphic forms were gently mixed using pestle and mortar and transferred into vials for analysis.

- Commercial pharmaceutical products (50 mg of fluconazole) were acquired at a local drugstore. The average weight of each capsule content was 150 mg. These commercial products consist of two groups:

  - The first group is composed of 14 capsules containing polymorphic form II of fluconazole.

  - The second group is composed of 17 capsules containing polymorphic form III of fluconazole

- Eleven samples of binary mixtures were prepared. These consisted of 50mg of fluconazole containing: 1, 2.5, 5, 10, 20, 50 80, 90, 97.5 and 99% (w/w) of polymorphic form II of fluconazole with the remaining of mass balance of monohydrate form of

fluconazole. These binary mixtures were mixed gently using pestle and mortar in order to ensure their homogeneity prior to the transfer to vials for FT-IR and FT-NIR analysis.

2.3. Multivariate data analysis

2.3.1. Datasets

PLS-DA models were developed based on the partial least square algorithm – discriminant analysis using the PLS Toolbox V8.2.1 (Eigenvector research INC, USA) running on Matlab (R2018b) (The Mathworks, USA).

The dataset is composed of three main parts: training, test and suitability set. The training set was used to develop PLS-DA models for three polymorphic forms of fluconazole while test and suitability set were used to test developed models.

The approach of Kennard stone was used to split dataset into calibration and validation set. It focused mainly firstly on selecting 2 samples that are the farthest apart from each other based on their variables. These 2 samples are put into the calibration data set. This procedure is repeated until reaching the desired number of samples of calibration set

In fact, this approach was used to split the dataset in order to provide uniform coverage into Training and test set consisting of:

- samples spectra obtained from mixing 50 mg of each pure polymorphic form of fluconazole with 100 mg of lactose monohydrate in order to be closed from the drug formulation.

-samples from capsules of pharmaceutical products.

A Suitability set was used to test how the built PLS-DA models for FT-NIR and FT-IR can detect minor polymorphic transformations and prove its suitability for samples that do not contain fluconazole. Here, suitability set is composed of:

- Eleven different spectra of binary mixture of polymorphic form-II and monohydrate form from 1% to 99% (w/w)

- Three spectra of itraconazole, miconazole and monohydrate lactose were included.

2.3.2. FT-IR and FT-NIR preprocessing

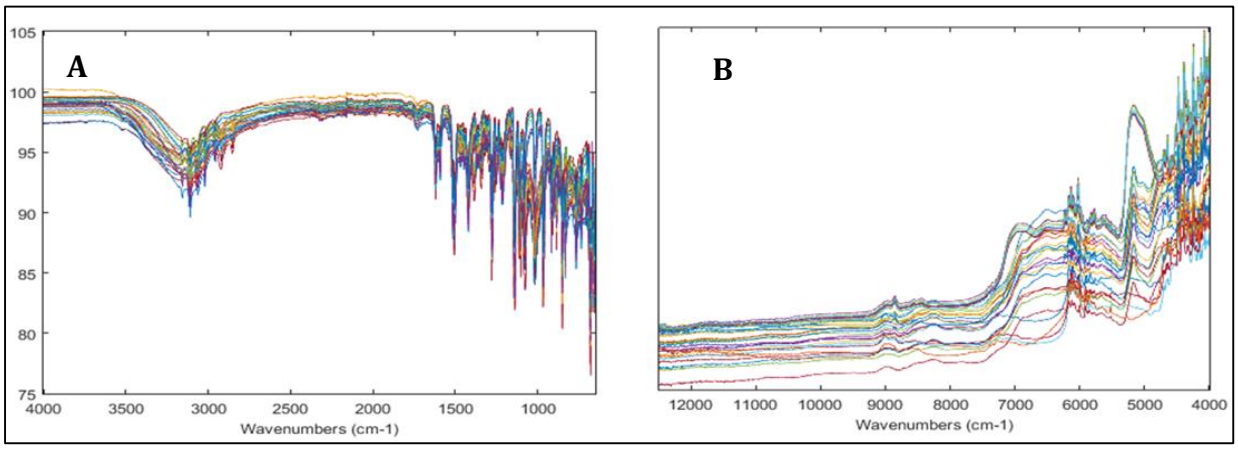

Figure 1: Raw spectra of fluconazole (A): FT-IR, (B): FT-NIR

Before developing the PLS-DA models, the raw spectra (Figure 1) need to be preprocessed to improve discrimination results. We used first derivative as preprocessing method followed by mean centering for FT-IR spectra and we applied standard normal variate (SNV) with mean centering for NIR spectra as shown in and Figure 2.

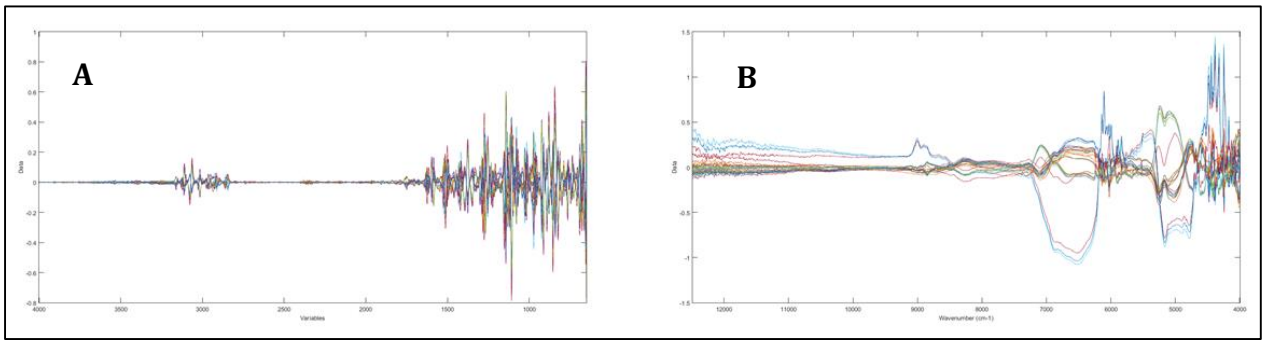

Figure 2: preprocessed spectra of fluconazole. (A): FT-IR, (B): FT-NIR

Partial least squares - discriminant analysis (PLS-DA) model was developed based on the training sets and also built only on a range of wavelengths containing variables that are considered as fingerprints of each polymorphic form.

PLS-DA is a supervised classification method that depends on X (FT-NIR or FT-IR spectra of fluconazole polymorphic forms samples) and Y (classes of different polymorphic form) matrices to develop a discriminative calibration model. It is based on reducing the data to scores and loading matrix which permit looking for the most optimal latent variables by maximizing the covariance between X and Y. In case of PLS-DA, Y contains a qualitative variable that identifies different classes of polymorphic forms of fluconazole. Therefore value

"1" is attributed to the target class that has to be discriminated from the other alternative classes that have the value of "0".

Latent variables are chosen based on the lowest error of an appropriate cross validation as it was elucidated in the Figure 23 of the first chapter. This built model is then evaluated by Root Mean Square Error of Calibration (RMSEC) and Cross-Validation (RMSECV) from Training set and Root Mean Square Error of Prediction (RMSEP) from the test set used to validate the developed model [125]. To guarantee the reliability of the model in the classification of different classes, a confusion matrix of classification parameters is used to evaluate the performance of PLS-DA models with sensitivity and specificity as shown in Table 1.

		Predicted class: Positive	Predicted class: Negative		
Actual class	**Positive**	True Positive (TP)	False Negative (FN)	**Sensitivity**	**Accuracy**
	Negative	False Positive (FP)	True Negative (TN)	**Specificity**	

Table 1: Confusion Matrix

The sensitivity is defined as the proportion of the samples of the class that are correctly attributed to the target class (true positives (TP)) while the specificity is known as the proportion of the samples that do not belong to the target class to be classified to the alternative class (true negatives (TN)) as given by following equations: [126].

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{FP + TN}$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

There are two PLS-DA approaches: the first approach is based on building PLS-DA model for each target class while the other classes are considered alternatives (so called one vs rest classification). This approach would lead us to build three PLS-DA models because of three

polymorphic forms of fluconazole. The second approach is PLS-2 regression and it is known as one vs one multiclass PLS-DA model, and it leans on building one model for all calibrated classes.

2.3.2. Suitability set

Suitability sets were also used to test the developed PLS-DA models. However, in this part we relied on Hotelling's T^2 and Q residual parameters as has been explained in Figure 16 of first chapter. These parameters helped to elucidate the behavior of the calibrated PLS-DA models towards the new integrated samples. These two parameters (Hotelling's T^2 and Q residual) were used in a plot with a threshold of 0.95 to check the homogeneity of the dataset and detect if there exist any outliers. The Q statistic is used to check how well each sample conforms to the model. The Q value is measured by the difference between the original data and the data reconstructed based on the calibrated model. They are associated to each sample of the dataset and large Q values indicate samples that have large out-of-model residuals. Hotelling's T^2 statistic represents the variation in each sample within the model; it is a measure of the sample distance from the center of the model. A sample with a large Hotelling's T^2 value means that this sample has an influence on the developed model.

3. Results and discussion

3.1 FT-IR and FT-NIR spectra of fluconazole polymorphic forms

Figure 3 reports different recorded spectra for the three pure polymorphic forms of fluconazole. The spectra agree with previously reported spectra of FT-IR and NIR that have been already analyzed and confirmed by XRPD [25,26]. Improving the discrimination ability of the PLS-DA models has been focused on the spectral regions that are responsible for the polymorphic forms. In this case, building PLS-DA models were limited to the region of 3500-2800 cm-1 and 1670- 760 cm-1 for FT-IR and from 9000 to 4500 cm-1 for FT-NIR.

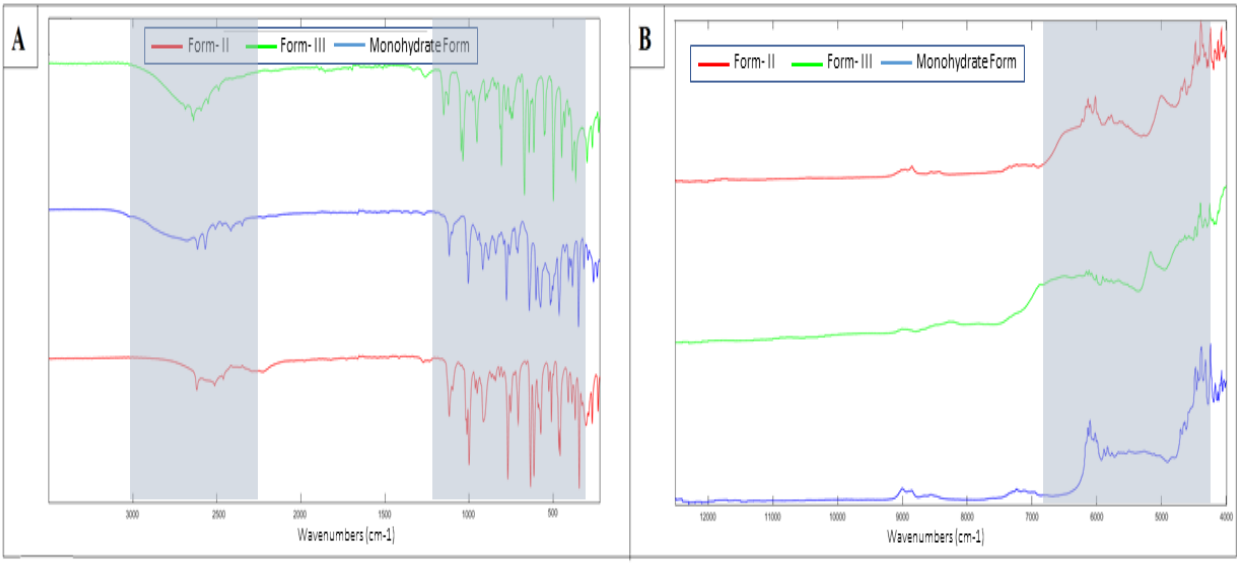

Figure 3: Spectra of Fluconazole. Pure polymorphic forms of (A): FT-IR; (B): FT-NIR

3.2. Exploring datasets

Before developing the PLS-DA models, the homogeneity of pharmaceutical products belonging to the same polymorphic form needs to be checked to ensure that there is not any difference between batches of the same polymorphic forms. This is why PCA was used in order to verify their homogeneity. Based on the PCA plot in Figure 4 using three components for both FT-NIR or FT-IR data, it is noticed that pharmaceutical products that belong to polymorphic form II or the two pharmaceutical products that are from form III are homogenous. It can be summarized that no tablet or pharmaceutical product that may have a polymorphic transformation.

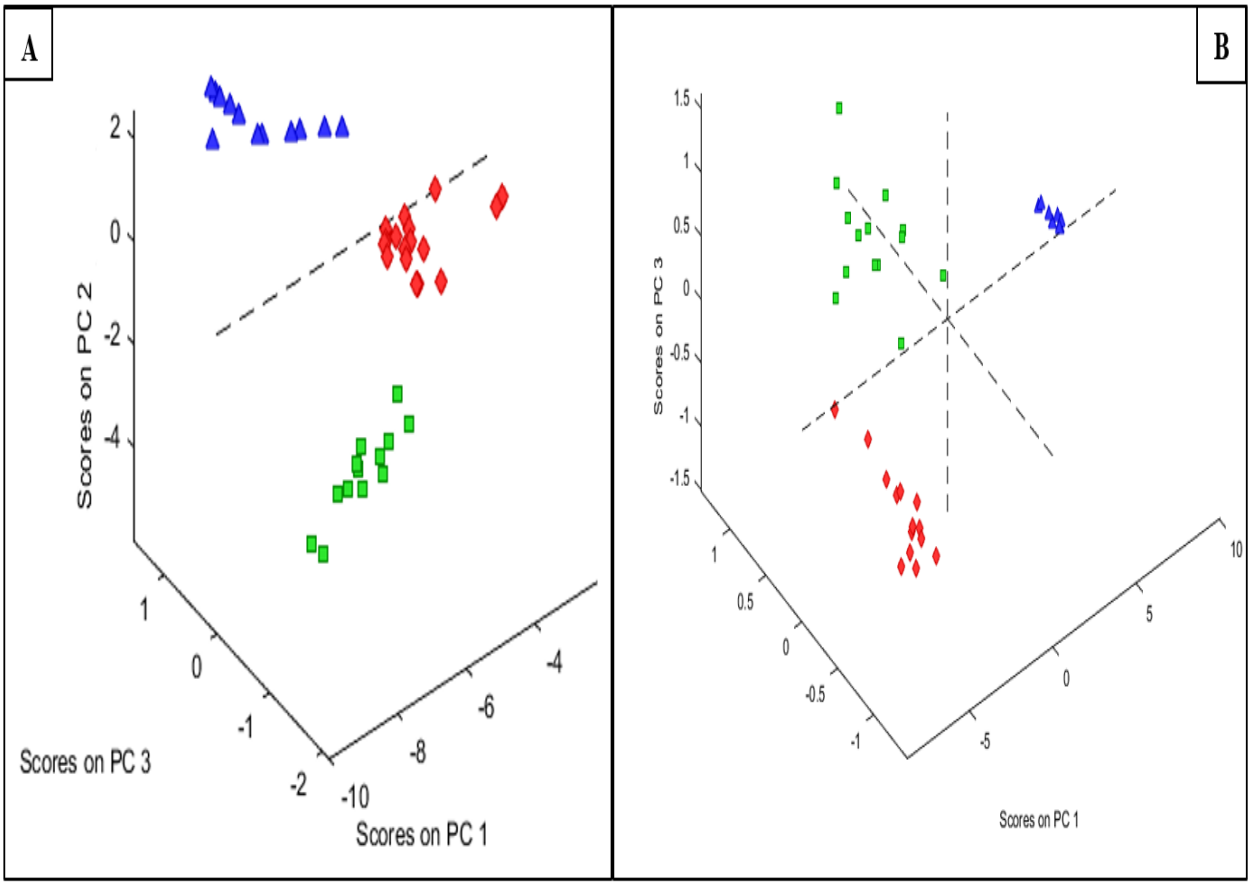

Figure 4: PCA for the data homogeneity; (A): FT- IR; (B): FT- NIR; Form- III (green square); Form-II (red diamond); Monohydrate form (blue triangle)

Another step to verify the homogeneity of the entire datasets is done based on the parameters of Hotelling's T^2 and Q residuals. According to these plotted parameters in Figure 5, it is observed that most samples have a low Hotelling's T^2 and Q residual values. Nevertheless, some samples are located above the threshold. These samples were tested by removing them and comparing RMSECV and RMSEP before and after removing these samples. Since there are no changes in the values of RMSECV and RMSEP, and thus no sample behave as an outlier whether in FT-IR or FT-NIR spectra.

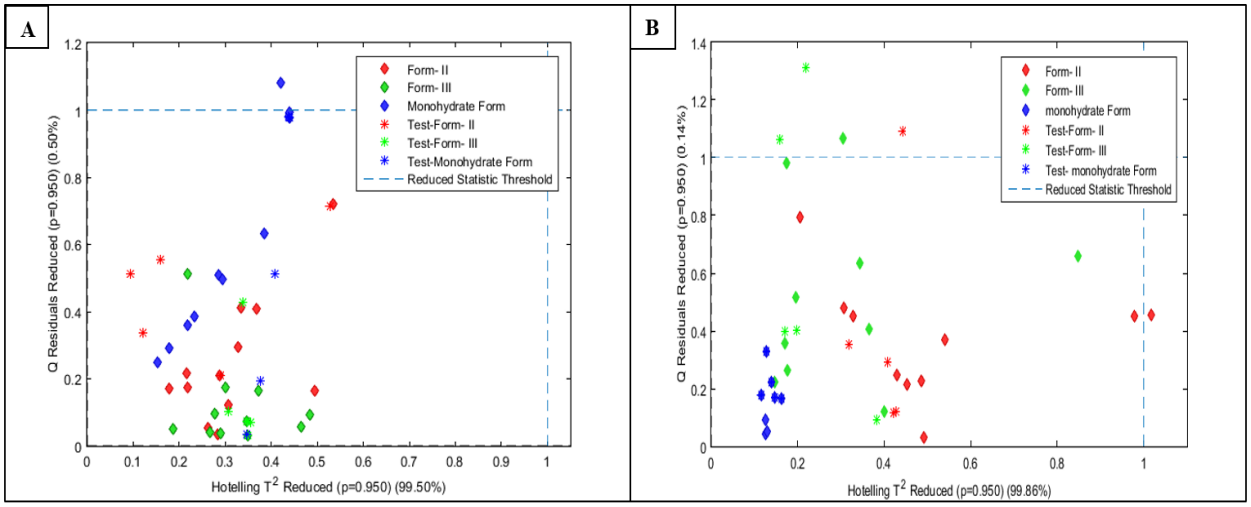

Figure 5: Hotelling T² VS Q Residuals of the training and validation set. (A): FT-IR; (B): FT-NIR

3.3. Development of PLS-DA models for FT-IR and FT-NIR spectra

A PLS-DA model was built for each instrument based on the training set. The optimal number of latent variables was chosen using a cross-validation of venetian blinds with 6 splits. We found that five latent variables for FT-NIR and three for FT-IR minimized the RMSECV.

Table 2: Classification parameters of PLS-DA

Spectroscopy	FT-IR			FT-NIR		
Polymorphic Forms	Form- II	Form- III	Monohydrate	Form- II	Form- III	monohydrate
Preprocessing	SG1D (2,15) + MC			SNV + MC		
Spectral range	3500- 2800 cm-1 & 1670- 760 cm-1			9000 to 4500 cm-1		
Cross-validation	Venetian blinds			Venetian blinds		
LV	3			5		
RMSEC (%)	11.1	6.3	10.8	4.7	3.4	2.6
RMSECV (%)	12.5	7.2	11.7	4.8	3.5	3.1
RMSEP (%)	14.1	8.1	11.9	6.5	5.5	2.3
Selectivity Cal (%)	100	100	100	100	100	100
Specificity Cal (%)	100	100	100	100	100	100
Selectivity Pred (%)	100	100	100	100	100	100
Specificity Pred (%)	100	100	100	100	100	100
Discriminant Threshold	0.4	0.3	0.4	0.6	0.3	0.3

The parameters of the PLS-DA models are shown in Table 2. These results are examined using accuracy that represents the RMSEP, specificity and selectivity. Based on the obtained results, it is noticed that there is a concordance between RMSECV and RMSEP indicating the accuracy of the PLS-DA model of FT-IR and FT-NIR. These results were confirmed by looking to the discriminant plots of the three polymorphic forms of fluconazole in Figure 6. The PLS-DA models provides 100% of specificity and selectivity for both FT-IR and FT-NIR datasets.

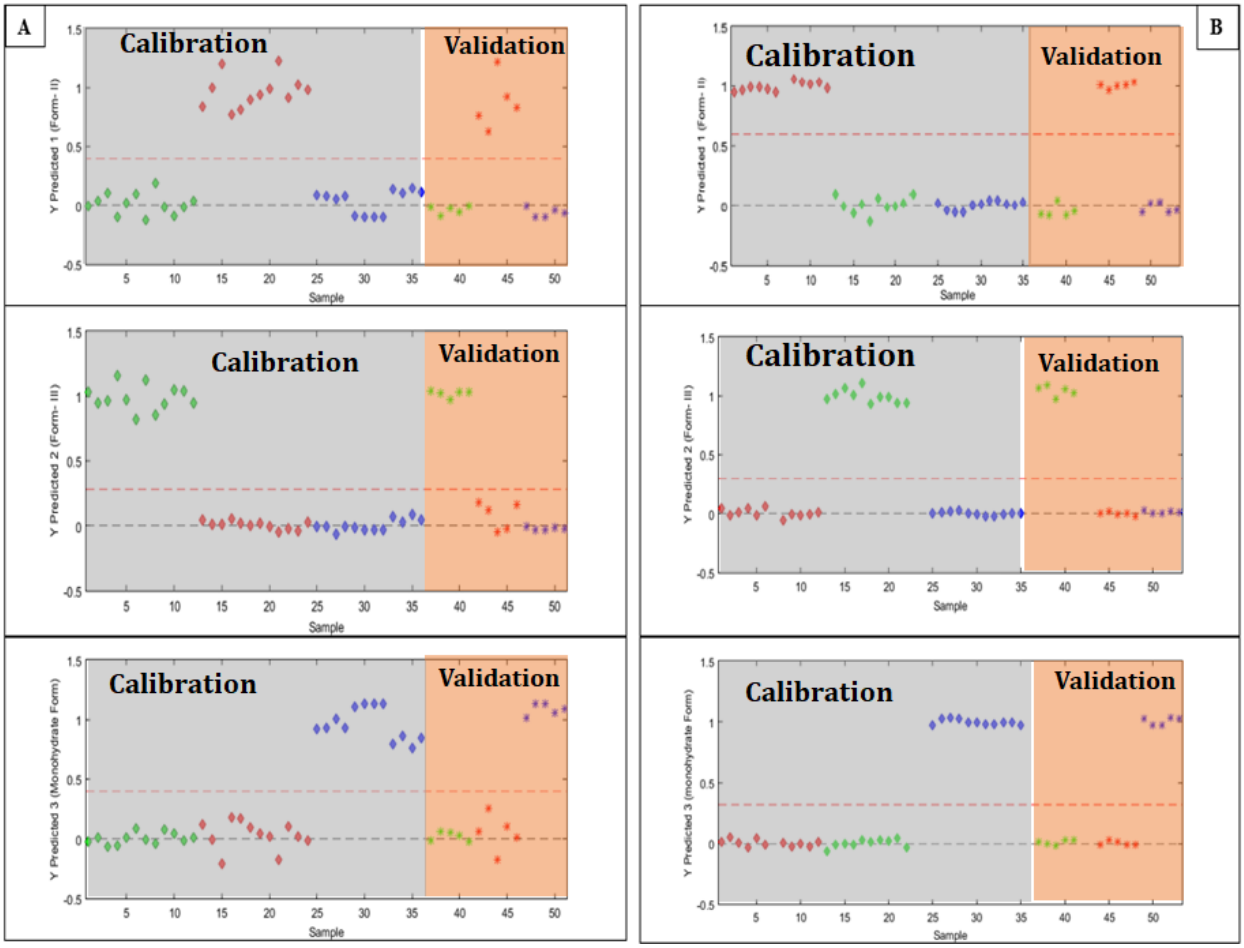

Figure 6: Discriminant plots of three polymorphic forms: (A): FT-IR; (B): FT-NIR

3.4. The suitability test of PLS- DA models

The suitability test is included in this work to investigate the robustness of the developed PLS-DA models with new samples that are different from samples that have each of the three polymorphic forms and used calibration set. Two kinds of samples were integrated: first category includes samples that do not have any of the three polymorphic forms. In this case, we included the sample lactose monohydrate as the excipient most commonly used in commercial fluconazole drug products. Each of Miconazole and itraconazole were also included since they belong to the same class of fluconazole. The second set contains fluconazole in a binary mixture: 1) fluconazole form- II, which is the metastable form and may convert to fluconazole monohydrate, and 2) the monohydrate form in different ratios from 1 to 99%. Hotelling's T^2 and Q residuals were used to ensure that these samples are considered to be outliers.

Figure 7 illustrates how the three samples including itraconazole, miconazole and monohydrate lactose are different from the training set of FT-IR and FT-NIR because they

have high values of Hotelling's T^2 and Q residuals versus the set containing the three polymorphic forms. For FT-IR, three samples: miconazole, itraconazole and lactose monohydrate are suspected to be different from the training set due to their high Q residual. In FT-NIR, the three samples were considered outliers because of their high Hotteling's T^2 and Q residual values. Thus, the plot of Hotelling's T^2 and Q residuals values could distinguish these three samples from the three polymorphic forms of fluconazole.

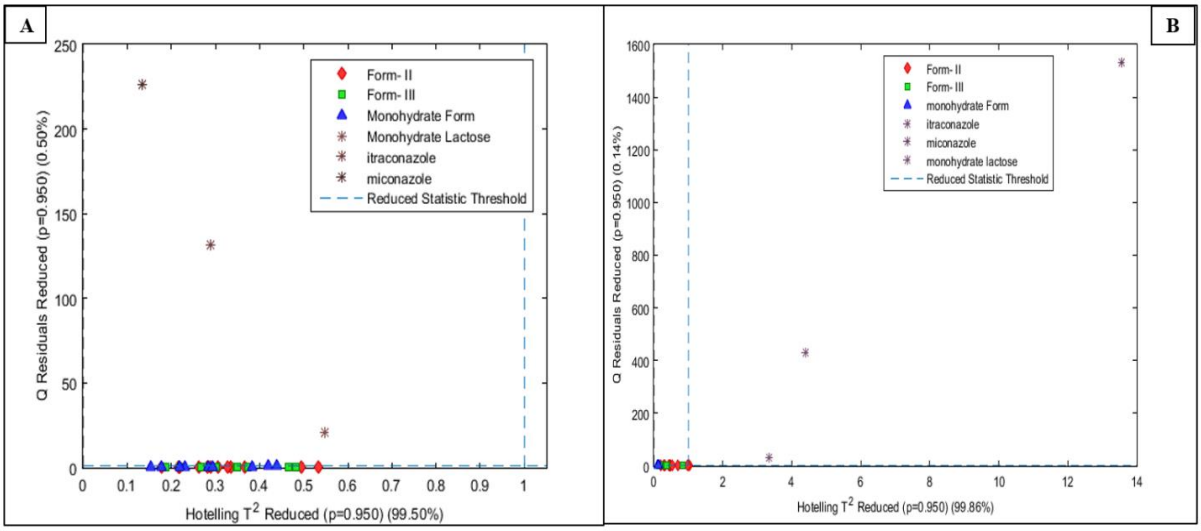

Figure 7: Hotelling's T^2 Vs Q residual plot of miconazole, itraconazole and monohydrate lactose. A: FT-IR; B: FT-NIR

Figure 8 shows the Q residual vs Hotelling's T^2 of eleven binary mixtures of fluconazole polymorphic forms composed of form- II and monohydrate form .

Figure 8A demonstrates how FT-IR combined to PLS-DA is able to reject different binary mixtures. Based on this plot, it is noticed that binary samples from the ratio of 5 to 95% can be distinguished since they have values of Q residual and Hotelling's T^2 significantly higher than training set. On the contrary, binary mixture ratios of 1 and 2.5% were considered belonging to the training set because of their low values of Q residual and Hotelling's T^2 that are similar to the values of training set of polymorphic forms.

Figure 8B shows how FT-NIR with PLS-DA can be useful to distinguish different binary mixtures from training set. Based on this plot, all binary mixtures consisted of: form II and monohydrate form are well identified as outliers versus the three calibrated polymorphic forms.

Based on these figures, it can be concluded that FT-NIR is able to detect minimum contaminant or conversion of form-II to monohydrate form because of the sensitive property of FT-NIR to hydrates.

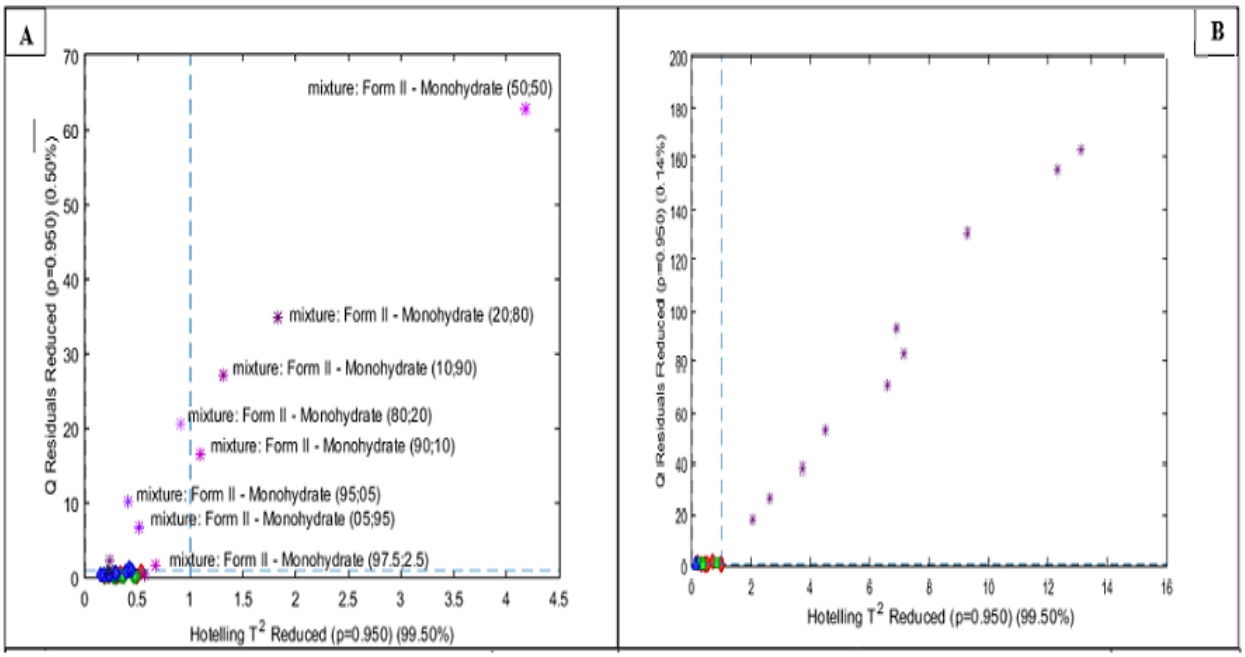

Figure 8: Hotelling's T² Vs Q residual plot of binary mixtures. A: FT-IR; B: FT-NIR

4. Conclusion

The results obtained with the PLS-DA models proved the suitability and the efficiency of combining vibrational spectroscopy whether it is FT-IR or FT-NIR with chemometric tools for identification and discrimination of fluconazole polymorphic forms. These results were performed based on several parameters summarized in: RMSEP, RMSECV, specificity and sensitivity that proved the ability of associating of FT-IR or FT-NIR to PLS-DA to discriminate between different polymorphic forms in pharmaceuticals. Finally, the suitability of the models was proven by analyzing itraconazole and miconazole as well as different binary mixtures (form II and monohydrate form) using Hotelling's T² vs Q residual plot. It has been confirmed that each of itraconazole, miconazole and lactose monohydrate are different since they have both Hotelling's T² and Q residual values significantly that are higher than the three main polymorphic forms whether using FT-IR or FT-NIR. Nevertheless, FT-NIR shows to be more suitable than FT-IR for detecting minor contamination between monohydrate form and form-II due to its high sensitivity to hydrates.

2nd Aspect

Investigation of PLS and MCR-ALS for the quantitation of ciprofloxacin in different pharmaceutical products using FT-NIR

M. Alaoui Mansouri, P.-Y. Sacré, M. Kharbach, L. Coic, C. De Bleye, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, E. Ziemons (Unpublished results)

Preamble

The advantages of quantifying drug substance using vibrational technique in different generic pharmaceuticals summarized in the accurate results obtained by one model and the ability to quantify new products without need to develop new models for each new generic sample. The main challenge represents in the excipients that can be varied from one generic to another, which is considered as matrix effect and can have an impact on the developed model. In this case two chemometric approach were combined to near-IR to be evaluated and compared for quantitation of ciprofloxacin in different generics. The first approach is PLS regression that is based on developing models based on selecting latent variables that are related to the wavelengths of the target drug substance. The second approach is MCR-ALS that is based on deconvolution of different components of a mixture to the concentration and spectra profiles. This approach can be used for quantitation purpose by using the constraint of correlation that correlate the concentration profile of drug substance of MCR to the reference concentration of the drug substance using a simple linear model. The characteristic of MCR-ALS showed its usefulness to overcome problem of matrix effect and the possibility to quantify ciprofloxacin in different generic compared to PLS.

Summary

PLS and MCR-ALS models were performed to quantify ciprofloxacin in several pharmaceutical products containing different excipients composition using FT-NIR spectroscopy.

First, both of chemometric tools (PLS and MCR-ALS) were evaluated to prove their first order advantage through the quantitation of ciprofloxacin in each of two sets independently composed of the same API but with different excipients.

Then, a prediction model of each chemometric tool was developed based on a one dataset, that was composed of samples constitutes of the same drug substance with different excipient, to evaluate the ability of both models to deal with the challenge of change in a matrix composition and quantify accurately ciprofloxacin.

Finally, PLS and MCR-ALS models were applied to quantify ciprofloxacin in different brands of commercial tablets. These commercial tablets were characterized by the presence of new excipients that were absent in the dataset composed of first and second set that have the same drug substance with different excipients.

1. Introduction

The combinason of chemometric tools with spectroscopic techniques has shown its efficiency in the pharmaceutical applications[127]. These applications vary according to the purpose whether qualitative, quantitative or unveiling the degradation process of the active pharmaceutical ingredient and reveal its multiple solid-state transitions [128,129]. Vibrational spectroscopic techniques are characterized by several advantages such as non-destructivity, simplicity and rapidity of use such as near infrared spectroscopy [130-132]. However, several challenges limit vibrational spectroscopy to be applied as in case of the presence of high overlapped absorption spectra [133]. The application of multivariate calibration methods, based on the mathematical elaboration of a high number of variables, can allow to overcome the spectral overlap of API or excipients and permit their rapid qualitative and quantitative analysis [134]. Furthermore, the chemometric tools agree with the requirements of the Green Analytical Chemistry. This area concerns the role of the analyst in developing laboratory practices more environmentally friendly, by minimizing the use of chemicals, energy use, waste and recycle.

Among of chemometric approaches that are used with vibrational spectroscopy for API quantitation in a mixture are PLS regression and MCR-ALS [135,136]. Each of these two approaches works in a different way, PLS regression aims to correlate spectral information with a dependent variable such as pharmaceutical tablet content [137]. Thus, a model is built to predict the property of interest from new sample spectra. MCR is a curve resolution method that decomposes the data matrix into its pure response profiles and their relative concentrations. MCR works in an iterative way (Alternating Least Squares–ALS) to achieve the best data decomposition. The ALS algorithm allows for several types of constraints (e.g. non-negativity in concentration/ spectral profile, correlation) to improve and reach chemical reasonable MCR solutions [138].

For the same API, the pharmaceutical product can be manufactured by many pharmaceutical companies; thus, the excipients used can vary in the different formulations [139,140]. Even though the target analyte to be analyzed with spectroscopic techniques is the same , the developed model for a specific formulation may be used only for the quantitative purpose of samples with the same composition. This constraint is due to the analysis by

spectroscopic techniques usually applied on the whole matrix and the chemometric processing is applied in the spectral data [141,142].

This work focused on performing two distinct chemometric approaches. MCR-ALS with correlation constraint and PLS regression for the quantitative analysis of ciprofloxacin in several tablets. These samples were composed of three sets. The first and second set are a Quaternary mixtures with different excipients. The third set was composed of different brands of ciprofloxacin pharmaceutical products, that were manufactured by different pharmaceutical companies. To test the ability of chemometric approaches to carry out the quantitation of ciprofloxacin in a sample with different excipients, both of first and second sets was merged and single models was developed by each chemometric approach. Both PLS and MCR-ALS models that are developed based on the merged sets, were used to quantitate ciprofloxacin in different brands of pharmaceutical products in a third set.

2. Material and methods

2.1. Sample preparation

Two sets of 16 mixtures were prepared based on a mixture design, each set consist of a quaternary mixture. Both of sets composed of ciprofloxacin (TCI-Chemicals, Belgium) with different excipients. Microcrystalline cellulose (Sigma-Aldrich, Belgium) and monohydrate lactose (Sigma-Aldrich, Belgium) are excipients of the first set, whereas povidone (Sigma-Aldrich, Belgium) and starch (Sigma-Aldrich, Belgium) are the main excipients of the second set. while stearate of magnesium (TCI-Chemicals, Belgium) is the common excipient of both sets as it is shown in Table 1A and Table 1B. Each compound was varied at five levels based on the ratio of ciprofloxacin that is existed in the pharmaceutical formulations: ciprofloxacin varied in the range between 55 and 75% (w/w), while all excipients varied between 10 and 30% (w/w) except content of stearate of magnesium kept unchanged at 5% (w/w). All mixtures range with a total weight of 200 mg.

Table 1A: Amount of ciprofloxacin in the first set

	Ciprofloxacin % w/w	MCC % w/w	Lactose % w/w	St Mg % w/w
1	75	10	10	5
2	75	10	10	5
3	65	10	20	5
4	66	19	10	5
5	70	15	10	5
6	66	11	17	5
7	60	25	10	5
8	60	10	25	5
9	55	25	15	5
10	55	15	25	5
11	55	30	10	5
12	55	10	30	5
13	60	18	17	5
14	60	17	18	5
15	70	12	13	5
16	70	13	12	5

Table 1B: Amount of ciprofloxacin in the second set

	Ciprofloxacin % w/w	Povidone % w/w	Starch % w/w	St Mg % w/w
1	75	10	10	5
2	75	10	10	5
3	65	10	20	5
4	66	19	10	5
5	70	15	10	5
6	66	11	17	5
7	60	25	10	5
8	60	10	25	5
9	55	25	15	5
10	55	15	25	5
11	55	30	10	5
12	55	10	30	5
13	60	18	17	5
14	60	17	18	5
15	70	12	13	5
16	70%	13	12	5

Each mixture was weighted and then were mixed using pestle and mortar to ensure the homogeneity. Each sample was pressed in a Specac hydraulic press under 5 ton/cm2 for 1 min to form a ciprofloxacin tablet with a diameter of 10 mm.

Besides of two sets, a set of ciprofloxacin commercial tablets, that contain 500 mg of ciprofloxacin, consists of three brands of pharmaceutical samples,. These samples were acquired from local drugstore. Each brand has different excipient as illustrated in Table 2. Each commercial tablet of ciprofloxacin was grinded and homogenized using the pestle and mortar. A weight of 200 mg of the homogenized powder was transferred to form a tablet with 10 mm of diameter using the Specac hydraulic press.

Table 2: Different commercial tablets

	Name of pharmaceutical product	Number of tablets	Manufacturer	Excipients
Brand I	Spectrum 500mg	5 tablets	Cooper pharma	Croscarmellose; microcrystalline cellulose; Povidone; magnesium stearate; Colloidal silica
Brand II	Ciproxine 500mg	5 tablets	Bayer	Sodium starch glycolate; povidone; microcrystalline cellulose; magnesium stearate
Brand III	Fleocip 250mg	5 tablets	Sothema	microcrystalline cellulose; povidone; magnesium stearate; Corn starch

2.2. Instrumentation

The samples were analyzed in transmission mode with a Fourier transform near infrared multipurpose analyzer spectrophotometer (MPA, Bruker Optics, Billerica, MA, USA). The spectra were collected with the Opus software V6.5 (Bruker Optics). Each spectrum was the average of 32 scans and the resolution was set at 8 cm-1 over the spectral range from 12500 to 4000 cm-1.

2.3. Data analysis and software

Before applying whether PLS or MCR-ALS, preprocessing of spectra were the first step to be carried out on the NIR data. The aim of this first step is to reduce the instrumental and physical artefacts as noise and light scattering that has not related to the chemical behavior of the analyzed mixtures. The performance of different preprocessing techniques including 1st and 2nd derivatives, Standard Normale variate (SNV) and Multiplicative Signal Correction (MSC) were investigated individually and in combination. Among of the preprocessing techniques, SNV and 1st derivative proved their efficiency to correct the NIR spectra based on the lowest predictive error obtained of PLS regression and MCR-ALS models.

Each of two datasets were splited into a calibration and validation sets using Kennard-Stone algorithm. The same calibration and test set were used to develop the PLS and MCR-ALS models.

In order to carry out the quantitation of ciprofloxacin using MCR-ALS, singular value decomposition (SVD) was used to determine the number of components based on the differences between eigenvalues and also the prior knowledge about one or more compounds that are present in the mixture. Then, the initial estimation was carried out using the purest variable approach to estimate the spectral profile of each compound (ST) and their profile concentration (ciprofloxacin and excipients for each set). To acquire chemically and physically meaningful solutions of both C and ST, the optimization was carried out by iterative ALS procedure. This procedure includes several constraints that were applied on the developed MCR modelto minimize the impact of rotational ambiguity issue and thus obtaining a unique solution. In this sense, the constraint of non-negativity was used only on the concentration profile whereas the same constraint was not applied on the spectral profile since the spectra were preprocessed. The correlation constraint, with the objective of performing quantitative analysis, was also used. By the application of this last constraint, it was possible to predict the concentration values in unknown samples as the concentration of a specific component, in this case, it was ciprofloxacin that was correlated with its reference concentration value. In this sense, a MCR calibration model is developed and used to predict the concentration of component of ciprofloxacin.

Both MCR-ALS and PLS models were evaluated using root mean square error of prediction (RMSEP), determination coefficient (Rp2) and relative percentage error in concentration (RE). RMSEP and RE were calculated according to the following equations:

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n}(c_i - \hat{c}_i)^2}{n}} \quad (1), \qquad RE\ (\%) = \sqrt{\frac{\sum_{i=1}^{n}(c_i - \hat{c}_i)^2}{\sum_{i=1}^{n} c_i^2}} \quad (2)$$

Where, n is the number of samples, ci is the experimental measurement for prediction samples, and $\hat{c}_i$ is the obtained value that correspond to validation.

The PLS modelling was performed using the PLS Toolbox V8.2.1 (Eigenvector Research INC, USA) running on Matlab version R2018b (MathWorks, USA) while the MCR-ALS modeling was performed through an interface graphical user [17].

3. Results and discussion

3.1. NIR spectra of each set components.

Fig. 1 shows the raw NIR spectra of compounds that form tablets of each set, where spectral feature of all compounds including the target analyte are identifiable. Whereas Fig. 1A and Fig. 1B shows spectra of each compound of the first and second set respectively. Fig. 2 represents NIR spectra of ciprofloxacin in mixtures. Regarding NIR spectra of ciprofloxacin in each mixture in Fig. 2A and Fig. 2B, it can be seen that the profile of NIR mixture spectra of the first set is different from the second one. Although the NIR spectra was recorded from 12500 to 4000 cm-1, the quantitative analysis whether by PLS or MCR-ALS was limited to 9792-7360 cm-1 in order to remove the high absorbance that were noticed at the end of spectra. Then the preprocessed techniques were applied on the selected region as it is shown in Fig. 3.

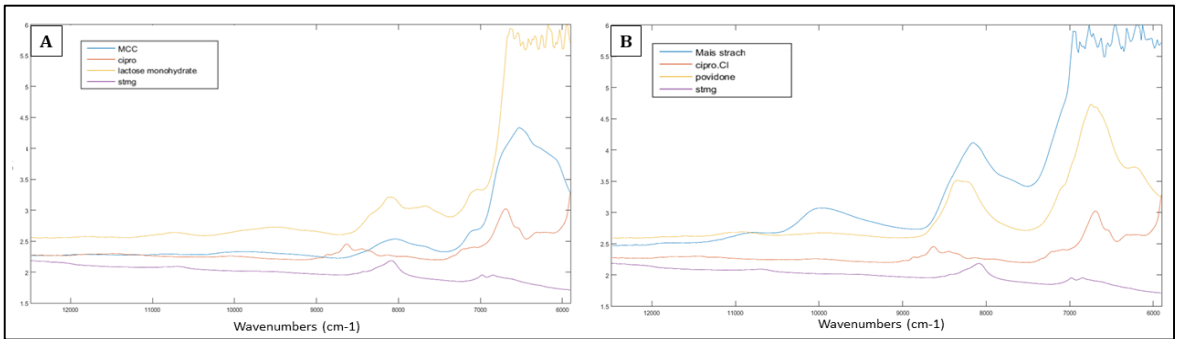

Figure 1: spectra of raw materials of each set. (from top to bottom) (A): set-1: monohydrate lactose, cellulose microcrystalline, ciprofloxacin and magnesium stearate ; (B): set-2: corn starch, povidone, ciprofloxacin and magnesium stearate

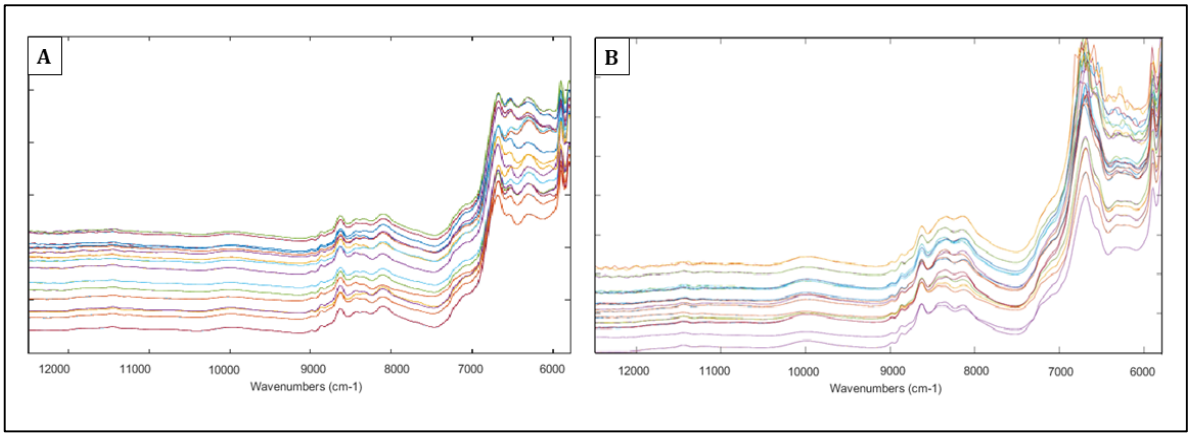

Figure 2: Near-infrared spectra of quaternary mixture of each set. (A) : set-1; (B): set-2

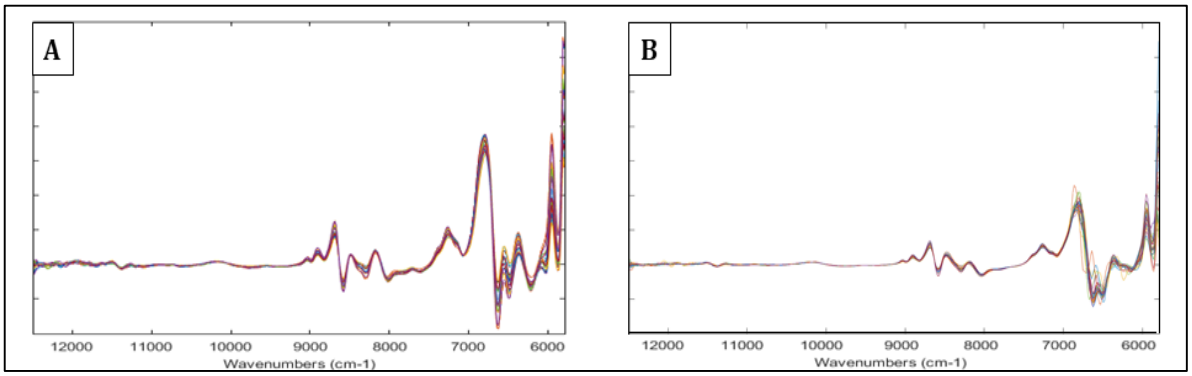

Figure 3: Preprocessed of near-infrared spectra of . (A) : set-1; (B): set-2

3.2. Quantitative analysis of ciprofloxacin.

Ciprofloxacin is the analyte that was used in this study for quantitation, where two cases were considered as illustrated in Fig. 6. Regarding to the first case, the quantitative analysis was carried out on each set independently to evaluate the quantitation by PLS and MCR-ALS without considering the matrix effect. Whereas the second case, the quantitation by PLS and MCR-ALS on the merged set based on gathering the first and second set was performed considering the matrix effect. The developed models in the second situation were tested on different brands of commercial tablets.

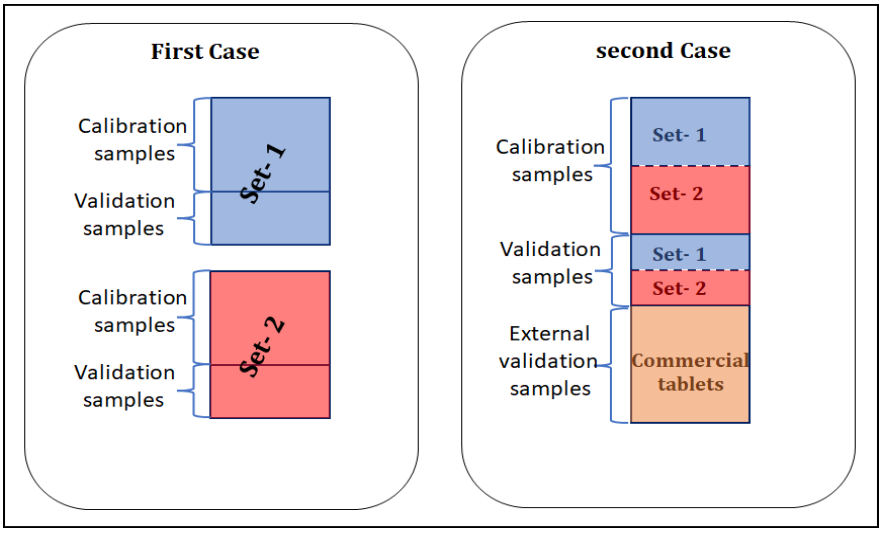

Figure 4: Schematic representation of each case. First case: models were developed on each set independently (without matrix effect). Second case. The two sets was merged in one set and one model of each of PLS and MCR was developed (with matrix effect).

3.2.1. Quantitative analysis without considering matrix effect

The concentration of ciprofloxacin was performed in mixtures of each set independently using two different multivariate regression techniques PLS and MCR-ALS to the NIR spectra. While the optimal latent variables of PLS, which used to develop the model, was obtained based on a leave-one-out cross validation, the model of MCR was based on the ALS optimization that includes constraints of non-negativity and correlation. The last cited constraint is responsible on performing the quantitative analysis by MCR. Regarding to Fig. 5A and Fig. 5C which shows the four loadings obtained by PLS in first and second set respectively, the first loading represents clearly the obtained preprocessed spectrum of ciprofloxacin. Whereas Fig. 5B and Fig. 5D correspond to spectra profiles obtained by MCR-ALS, which the main curve obtained by MCR was recognized to be the preprocessed spectrum of ciprofloxacin in both sets.

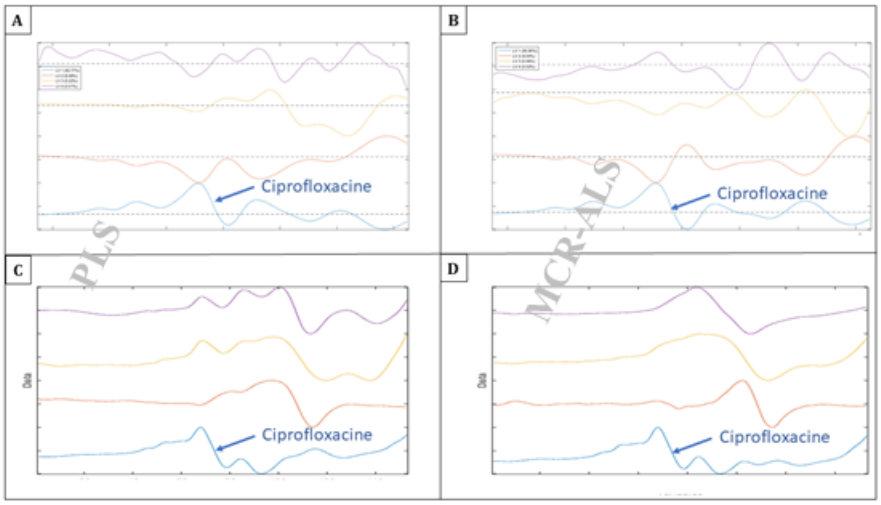

Figure 5: Loadings of PLS and recovered spectra by MCR
(A and C): loading and recovered spectra of set-1; (B and D): loading of PLS and recovered spectra of set-2

Fig. 6 shows the reference versus predicted concentration values plot for the developed models that shows clearly well distribution of predicted values around the line whether for PLS or MCR- ALS in both sets.

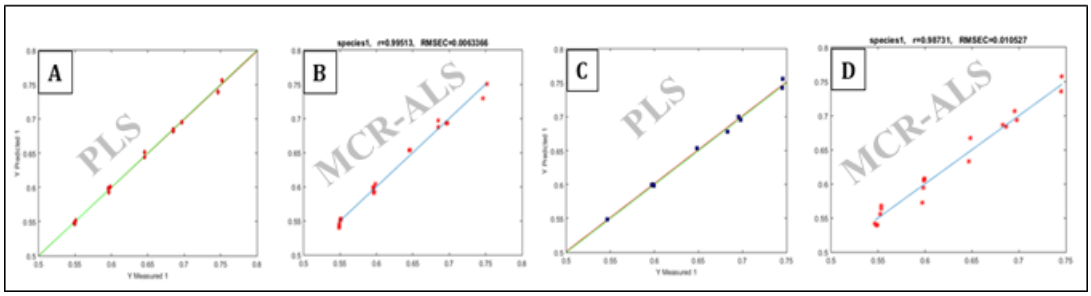

Figure 6: Plots of reference values versus predicted values for PLS and MCR for set-1 (A and B) and set-2 (C and D).

These obtained results are summarized in the Table 3. According to these results, both MCR-ALS and PLS models are able to predict the concentration values of unknown samples in the same matrix. However, in this case of quantifying ciprofloxacin in each set independently, PLS has lower RMSEP and RE compared to MCR-ALS.

Table 3: Results of PLS and MCR-ALS models for each set

	PLS				MCR-ALS			
	LV	R^2p %	RMSEP %	RE %	PCs	R^2P %	RMSEP %	RE%
Set -1	4	99.7	0.38	0.54	4	99.27	0.67	1.15
Set-2	4	99.79	0.47	0.68	4	94.59	1.76	1.98
Merged set	6	85.55	4.3	6.41	6	95.39	1.88	1.29

3.2.2. Quantitative analysis of ciprofloxacin content considering matrice effect

A model of PLS and MCR-ALS were developed on a merged set of both first and second set, in this situation the matrix effect was caused by variation in excipients from one mixture to the other one. For PLS regression model, it was developed based on six latent variables that was determined by cross validation of leave-one-out. Whereas for MCR-ALS, six principal components were determined based on singular value decomposition (SVD).

The obtained results of the merged set considering the PLS and MCR-ALS are given in Table 3. It can be seen that parameters RMSEP and RE of PLS model developed on the merged set are higher than to those parameters obtained of PLS models developed on each set independently in section 3.2.1. Whereas the MCR-ALS models displays somewhat the same lower errors whether for the merged set or for each set. These obtained results show clearly how the variation in matrices composition can impact the ability of PLS models to predict accurately the ciprofloxacin content. Unlike PLS model, MCR-ALS model proves its ability to deal with variation in matrix composition.

3.2.3. Quantitative analysis of ciprofloxacin content in different brands of pharmaceutical products

Both of PLS and MCR-ALS models that were developed using the merged set, were tested to predict the ciprofloxacin content in commercial tablets that belong to three different manufacturers. Each tablet has different excipients and some of these excipients were not included in the merged set. In Table 4, the performance of PLS and MCR-ALS are summarized. The table shows that variation in matrice composition of commercial tablets arise the error of PLS model. The commercial tablets of spectrum 500 mg are the brand that shows the highest RE and RMSEP compared to other brands of commercial tablets. This highest error is mainly due to the presence of excipients of croscarmellose and colloidal silica that they are not belong to the samples composition of merged set. Unlike PLS model results, the results of MCR-ALS show that for all brands, the errors keep similarly low no matter if there is variation in excipients composition or not. Additionally, these results proved that the MCR-ALS model can handle spectral differences due to the absence or presence of excipients in the pharmaceutical products.

Table 4: Results of the application of PLS and MCR-ALS models on the commercial tablets

	Name of pharmaceutical product	Excipients	PLS		MCR-ALS	
			RE %	RMSEP %	RE%	RMSEP %
Brand I	Spectrum 500 mg	Croscarmellose; microcrystalline cellulose; Povidone; magnesium stearate; Colloidal silica	12.82	8.39	2.66	1.74
Brand II	Ciproxine 500mg	Sodium starch glycolate; povidone; microcrystalline cellulose; magnesium stearate	9.78	6.38	2.54	1.65
Brand III	Fleocip 250mg	microcrystalline cellulose; povidone; magnesium stearate; Corn starch	8.53	6.20	1.11	0.80

4. Conclusion

A comparison between PLS and MCR-ALS were investigated to quantify ciprofloxacin using FT-NIR in different situations. First the quantitation was evaluated in the same matrice, then in a varied matrice composition and finally in different brands of pharmaceutical products that were manufactured by different pharmaceutical companies.

For the first situation, the obtained results showed the ability of both chemometric tools to quantify ciprofloxacin in the same matrice with low errors due to their first order advantage. In case of quantitation of ciprofloxacin in a dataset with different matrice composition. PLS model showed its limit due to its error that arose compared to the first situation, whereas MCR-ALS model kept its low error even if the matrice composition was changed from one sample to other, which demonstrated the ability of MCR-ALS to deal with matrice effect and to predict the content of ciprofloxacin in different matrix compositions due to its second order advantage.

3rd Aspect

Quantitation of Active Pharmaceutical Ingredient through the Packaging Using Raman Handheld Spectrophotometers: A Comparison Study

M. Alaoui Mansouri, P.-Y. Sacré, L. Coic, C. De Bleye, E. Dumont, A. Bouklouze, Ph. Hubert, R.D. Marini, E. Ziemons, Talanta. 207 (2020)

Preamble

Quantifying pharmaceuticals through their packaging represents an advantage and a challenge at the same time. The most convenient techniques for this kind of analysis are vibrational spectroscopy due to the advantage of no need of sample preparation; however, the main challenge of analysis through the packaging is the obtention of a spectrum including spectral signature of the sample beside that of packaging Thus, the spectral signature of packaging itself is considered as matrix effect. To overcome this issue, a Raman mode called spatially offset Raman scattering (SORS) allows the photons to pass through package and irradiate the sample due to the offset between illuminating and collecting spot, which allow to obtain of what is inside the packaging. Therefore, a comparative study between backscattering mode which is conventional and SORS mode has been performed in the framework of quantifying ibuprofen in a ternary mixture with mannitol and microcrystalline cellulose through polypropylene packaging using partial least squares regression models and validated by the approach of accuracy profile.

Summary

Handheld Raman spectroscopy is actually booming. Recent devices improvements aim at addressing the usual Raman spectroscopy issues: fluorescence with shifted-excitation Raman difference spectroscopy (SERDS), poor sensitivity with surface enhanced Raman scattering (SERS) and information only about the sample surface with spatially offset Raman spectroscopy (SORS). While qualitative performances of handheld devices are generally well established, the quantitative analysis of pharmaceutical samples remains challenging.

The aim of this study was to compare the quantitative performances of three commercially available handheld Raman spectroscopy devices. Two of them (TruScan and IDRaman mini) are equipped with a 785 nm laser wavelength and operate in a conventional backscattering mode. The IDRaman has the Orbital Raster Scanning (ORS) option to increase the analyzed surface. The third device (Resolve) operates with an 830 nm laser wavelength both in backscattering and in SORS modes.

The comparative study was carried out on ibuprofen-mannitol-microcrystalline cellulose ternary mixtures. The concentration of ibuprofen ranged from 24 to 52 % (w/w) while the proportions of the two excipients were varied to avoid cross-correlation as much as possible. Analyses were performed either directly through a glass vial or with the glass vial in an opaque polypropylene flask, using a validated FT-NIR spectroscopy method as a reference method. Chemometric analyses were carried out with the Partial Least Squares Regression (PLS-R) algorithm. The quantitative models were validated using the total error approach and the ICH Q2 (R1) guidelines with +/- 15% as acceptance limits.

1. Introduction

There is a growing concern toward using vibrational spectroscopy in pharmaceutical quality control [143,144]. Raman spectroscopy is considered as an important analytical tool beside Near-Infrared (NIR) and High Performance Liquid Chromatography (HPLC) [145-147]. This technique is based on the interaction between the energy of a monochromatic light and a sample inducing light scattering. It is characterized by many advantages that may be summarized in its minimal sample preparation requirement, its ability to be used on-site via handheld instrument and the possibility to analyze the sample through clear glass and bottles made from polyvinyl chloride (PVC) or polypropylene (PP). Nevertheless, there are drawbacks that present challenges using the Raman technique. Some of these drawbacks are: sample auto-fluorescence that may overwhelm the signal coming from the analyzed sample, the challenges of analyzing heterogeneous samples because of the small analyzed volume and of carrying out a qualitative or quantitative analysis of a compound in a mixture through a package material such as blisters or plastic bottles [148-150].

The last cited challenge of analyzing a drug substance through package material could be overcome using a specific measuring configuration mode called spatially offset Raman scattering (SORS). The main difference between SORS and backscattering Raman is that the scattered light is collected at a spatially offset location situated a few millimeters from the illumination site. This configuration allows collecting photons that have gone deep in the sample leading to spectra predominantly composed of content's signal [151,152]. To obtain the SORS corrected spectra, the outer (container) spectrum is scaled and removed from the offset spectrum in order to obtain a clean spectrum of the content [153]. SORS has already proved its usefulness in many sectors. For instance, in the pharmaceutical field, it was used to detect various types of raw materials in a range of non-transparent sealed containers and it further allowed detecting counterfeits [152]. For food analysis, it was demonstrated that SORS is able to detect the internal maturity of tomatoes or to detect chemical markers that are responsible for the adulteration and falsification of spirit drinks through bottles [154]. The obtained results showed that SORS is well suited to conduct analyses through different types of containers and samples.

Raman spectroscopy data must be analyzed with appropriate chemometric tools to extract relevant qualitative or quantitative information. Partial least squares regression (PLS-R) is a

multivariate data analysis method that is used to carry out the quantification of the target component in a mixture and can deal with interferences and overlapping bands [155,156].

The objective of the present study was to compare the quantitative performances of three handheld Raman spectrophotometers for the analysis a pharmaceutical powder sample directly through a glass vial and through a glass vial placed inside a polypropylene (PP) container. The pharmaceutical sample is a ternary mixture composed of ibuprofen, mannitol and microcrystalline cellulose. The prepared samples were analyzed by three handheld Raman spectrophotometers using different measurement technologies. Beside the handheld Raman devices, the samples were also analyzed with a benchtop NIR spectrophotometer used as reference equipment to check the sample preparation and detect possible outliers. The quantitative performance of the selected devices were evaluated based on accuracy profiles [157-159].

2. Material and methods

2.1. Instrumentation

The analyses were carried out on three handheld Raman instruments from different manufacturers. The first handheld Raman device is the TruScan spectrophotometer (Thermo Scientific, Waltham, MA, USA) utilizing a 785 nm excitation wavelength and covering the 250 to 2875 cm^{-1} Raman shifts range. The second device is the IDRaman mini (Ocean Optics, Largo, FL, USA) with a 785 nm excitation wavelength and covering the 400 to 2300 cm^{-1} spectral range and characterized by the option of Orbital Raster Scanning that increases the analyzed surface. The last device is the Resolve™ (Agilent Technologies, Santa Clara, CA, USA) utilizing an 830 nm excitation wavelength covering the 200 to 2000 cm^{-1} spectral region. This device can be used in two modes: conventional Raman spectroscopy and SORS. On the one hand, raw SORS data (ambient, zero and offset spectra) were extracted using the Resolve Database Data Viewer v0.0.8. Before being processed, raw offset spectra were corrected by removing the ambient spectra. However, no removal of the zero position spectra was performed. On the other hand, the backscattering spectra were pre-processed inside the device (baseline correction) and were subsequently imported directly from the latter.

The samples were also analyzed with a Fourier transform near infrared multipurpose analyzer spectrophotometer (MPA, Bruker Optics, Billerica, MA, USA). The spectra were collected with the Opus software V6.5 (Bruker Optics). Each spectrum was the average of 32 scans and the resolution was set at 8 cm^{-1} over the spectral range from 12500 to 4000 cm^{-1}.

2.2. Sample preparation

Different ternary mixtures of ibuprofen (Sigma-Aldrich, Belgium), microcrystalline cellulose (Sigma-Aldrich, Belgium) and mannitol (Sigma-Aldrich, Belgium) were realized to build both

calibration and validation sets. Ibuprofen was chosen as test molecule because of its moderate Raman scattering character. This permits obtaining a balanced signal between ibuprofen and excipients.

For the calibration set, the concentration of ibuprofen varied at five levels: 24, 32, 40, 48 and 52 % (w/w) covering the range of 60 – 130 % around the target concentration of 40 % (w/w) (equivalent to 200 mg of ibuprofen). The amount of excipients added was varied in order to keep the total sample weight constant at 500 mg in each mixture. To avoid cross correlation, the proportion between the two excipients was varied for each ibuprofen concentration level. A central mixture with each of the three components at 33.3 % w/w was also added leading to a total of 26 calibration samples.

Table 1: composition of calibration samples

Concentration level	**Ibuprofen (% w/w)**	**Excipients (% w/w)**	**Excipient proportions**									
			Proportion 10:90		**Proportion 30:70**		**Proportion 50:50**		**Proportion 70:30**		**Proportion 90:10**	
			MCC	Man	MCC	Man	MCC	Man	MCC	Man	MCC	Man
			(mg)	(mg)	(mg)	(mg)	(mg)	(mg)	(mg)	(mg)	(mg)	(mg)
60%	24	76	38	342	114	266	190	190	266	114	342	38
80%	32	68	34	306	102	238	170	170	238	102	306	34
100%	40	60	30	270	90	210	150	150	210	90	270	30
120%	48	52	26	234	78	182	130	130	182	78	234	26
130%	52	48	24	216	72	168	120	120	168	72	216	24
Central point	33.33	66.67	166.67 mg of MCC – 166.67 of Mannitol (Man)									

The validation set consisted of five concentration levels of ibuprofen (28, 34, 40, 46 and 52 % (w/w)) covering the range between 70-130 % of the target ibuprofen concentration. The ratio between the excipients was varied leading to 15 validation samples per series (see Table 1). Three series of validation were realized independently with new sample preparation and restart/recalibrate each device on each new series.

Table 2: ICH Q2 (R1) validation criteria values of the PLS models.

	Quantity of compounds (mg)				Concentration (% w/w)		
Concentration level	**Ibuprofen**	**MCC**	**Man**	**Total**	**Ibuprofen**	**MCC**	**Man**
70	140	36	324	500	28	7.2	64.8
		108	252			21.6	50.4
		180	180			36	36
85	170	231	99		34	46.2	19.8
		297	33			59.4	6.6
		33	297			6.6	59.4
100	200	90	210		40	18	42
		150	150			30	30
		210	90			42	18
115	230	243	27		46	48.6	5.4
		27	243			5.4	48.6
		81	189			16.2	37.8
130	260	120	120		52	24	24
		168	72			33.6	14.4
		216	24			43.2	4.8

Once weighted, the powders were finely grinded in a pestle and mortar to ensure homogeneous mixtures and placed in glass vials. Each mixture was analyzed in triplicate.

Both calibration and validation samples were analyzed directly through the glass vial (thickness of 1 mm) and through the glass vial placed in an opaque white PP container (thickness of 1mm). The analysis of the mixtures through the packaging was performed with all handheld Raman instruments. FT-NIR spectra were only acquired in reflectance mode on the glass vial.

2.3. Multivariate data analysis

The regression model was developed based on the partial least square (SIMPLS) algorithm using the PLS Toolbox V8.2.1 (Eigenvector Research INC, USA) running on Matlab (R2018b) (The Mathworks, USA). Different preprocessing techniques were investigated and compared based on the root mean square difference of prediction (RMSEP). The combination of standard normal variate (SNV) normalization with mean centering proved to be the most suitable for FT-NIR data, while the combination of the Savitzky-Golay 1st derivative

(polynomial order: 2, window size: 15), SNV and mean centering provided better predictions for Raman spectroscopy.

The accuracy profiles were computed using the results from the three validation series composed of three replicates at five concentration levels measured on each device using e-noval V4.0b (Pharmalex Belgium SA, Mont-Saint-Guibert, Belgium) with 95% β-expectation tolerance intervals [158,150-162]. The acceptance limits were set at +/- 15 % of relative total error following the European Pharmacopoeia general monograph 2.9.40 on the uniformity of dosage units.

3. Results and discussion

Actually, handheld Raman spectrophotometers are designed for qualitative analysis since the acquisition time is optimized at each measurement to obtain a sufficient signal to noise ratio [144]. Moreover, quantitation of active ingredients becomes a challenge when this is to be performed through thick and opaque containers. During this study, different ternary mixtures of ibuprofen/mannitol/MCC were prepared following the scheme described in section 2.2. Figure 1 shows the Raman spectra of each raw material and of the ternary mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. Once the samples prepared and placed in the glass vials, they were first analyzed by FT-NIR spectroscopy. The obtained NIR spectra were then processed by PLS-R allowing us to ensure that the mixtures were correctly prepared (the detected outliers were removed and prepared again). The computed PLS-R model based on FT-NIR data has been validated and the accuracy profile computed. The accuracy profile was well included in the previously set +/- 15 % of relative total error.

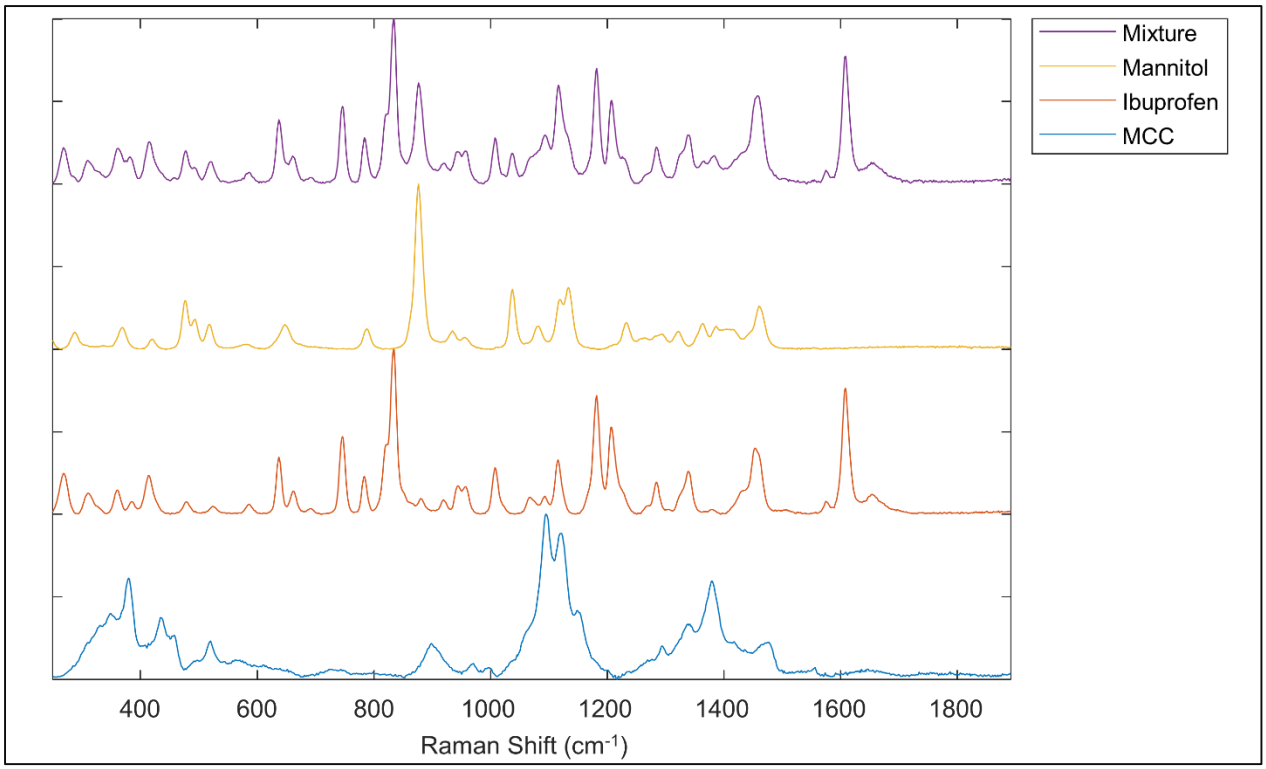

Figure 1: Raman spectra of raw materials and ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively. The spectra were acquired with the TruScan in "signature" mode through a thin plastic bag and were baseline corrected by asymmetric least squares.

Once the samples were verified by FT-NIR, they underwent analyses by the different handheld systems directly through the glass vial and through the glass vial placed in the PP container (see photos in supplementary materials).

3.1. Analysis of samples trough glass vial

Figure 2 shows the baseline corrected spectra acquired on each device in backscattering mode through the glass vial for a mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. The mixture and ibuprofen spectra were recorded through a thin plastic bag to avoid any interference from the glass. The dashed red lines indicate the main spectral features of ibuprofen and the blue dashed line indicates the main peak of mannitol. No specific peak associated to cellulose was observable because cellulose is a weak Raman scatterer and its features are masked by the other components. Nevertheless, cellulose disturbs the global signal due to a high fluorescence background. On this Figure 2, it is possible to see that the signal measured with each device exhibits a spectrum directly correlated to the mixture spectrum with the main spectral features of ibuprofen and mannitol. The spectra recorded

with the TruScan and the IDRaman mini also exhibit a spectral perturbation between 1300 and 1500 cm^{-1} due to the fluorescence of glass with the 785 nm incident laser source.

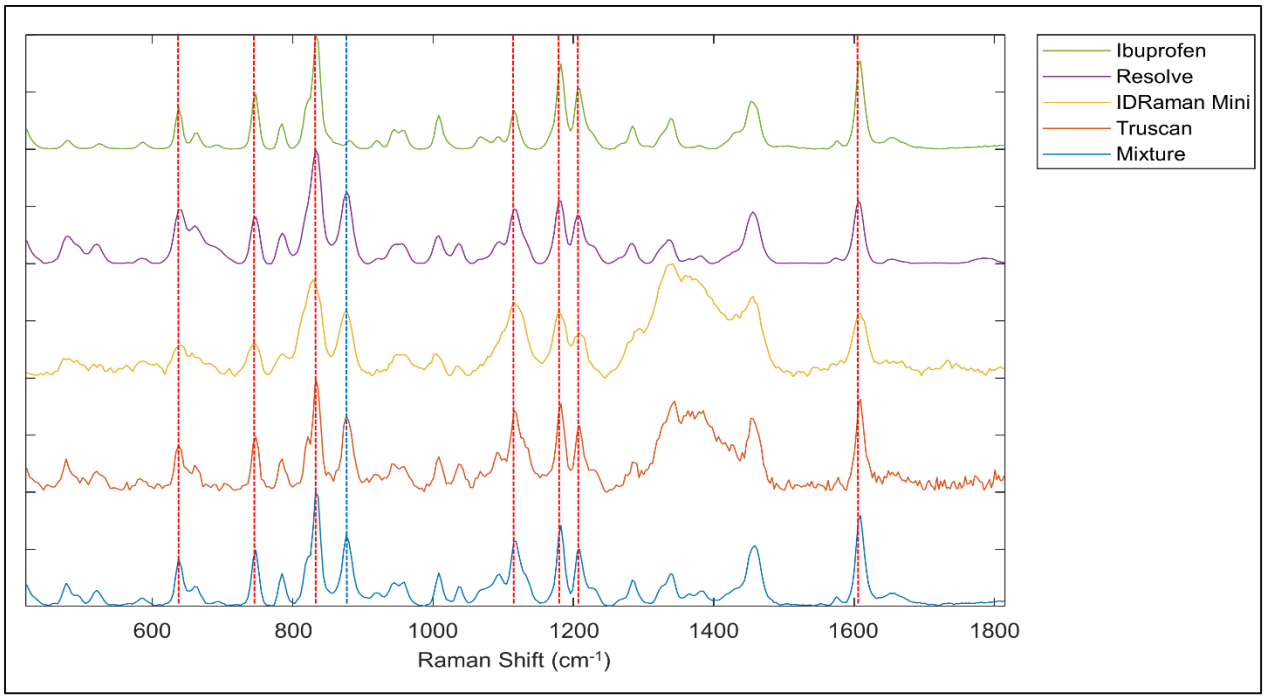

Figure 2: Raman spectra of a ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively.

The spectra obtained with each instrument (ID Raman mini, TruScan, Resolve) were modelled using PLS models. Several pre-processing and spectral ranges were tested. The best pre-processing and spectral range were selected based on the RMSEP and the bias computed on the validation set. Once the final PLS model selected, the accuracy profiles were computed for each device. Table 2 summarizes the final parameters used for the PLS models and their respective figures of merit. Each Raman spectroscopy model used the Savitzky-Golay [163] first derivative as pre-processing. Indeed, most devices use a 785 nm laser as light source. However, glass exhibits a high fluorescence background when irradiated by a 785 nm light and the use of the first derivative helped managing the fluorescence.

Table 3: Regression model parameters and figures of merit of handheld Raman devices and FT-NIR

	Resolve			TruScan RM		IDRaman mini		FT-NIR
Sample	Glass vial + PP	Glass vial + PP	Glass vial	Glass vial + PP	Glass vial	Glass vial + PP	Glass vial	Glass vial
Mode	***SORS***	***Backscattering***	***Backscattering***	***Backscattering***	***Backscattering***	***Backscattering***	***Backscattering***	***Reflection***
Spectral Range (cm^{-1})	1152-1608	400-1700	400-1700	1000-1600	400-1700	600-1700	400-1700	9000-4000
Pre-processing	SG1D (2,15); SNV; MC			SG1D (2,15); SNV; MC		SG1D (2,15); SNV; MC		SNV; MC
Latent Variables	6	7	3	6	3	5	3	5
R^2 calibration	98.7	97.2	96.0	98.7	94.0	99.8	98.7	99.0
RMSEC (mg)	5.6	8.5	10.1	5.9	12.4	2.4	5.6	5.3
R^2 Cross Validation	97.8	89.3	95.7	58.8	92.3	37.2	95.0	98.7
RMSECV (mg)	7.4	16.6	10.5	32.3	14.2	40.6	11.0	5.9
R^2 prediction	93.0	83.1	92.3	66.4	91.6	9.5	96.5	95.5
RMSEP (mg)	12.7	19.8	12.6	27.0	12.7	45.4	13.6	9.0

The computed accuracy profiles are shown in Figure 3 and the values of the validation criteria are reported in table 3. As can be noticed, all the selected devices provided good results as their 95% β-expectation tolerance intervals are included inside the acceptance limits of +/- 15%. This means that 95% of future measurements will have an accuracy (total error) of less than +/- 15%. These results indicate that for formulations with a well-balanced signal of both

excipients and API in a transparent container, satisfying quantitative performances may be obtained using Raman handheld devices in their native configuration (auto-exposure). Furthermore, no significant difference was observed between the devices and their specific acquisition modes (ORS vs single spot and 785 nm vs 830 nm).

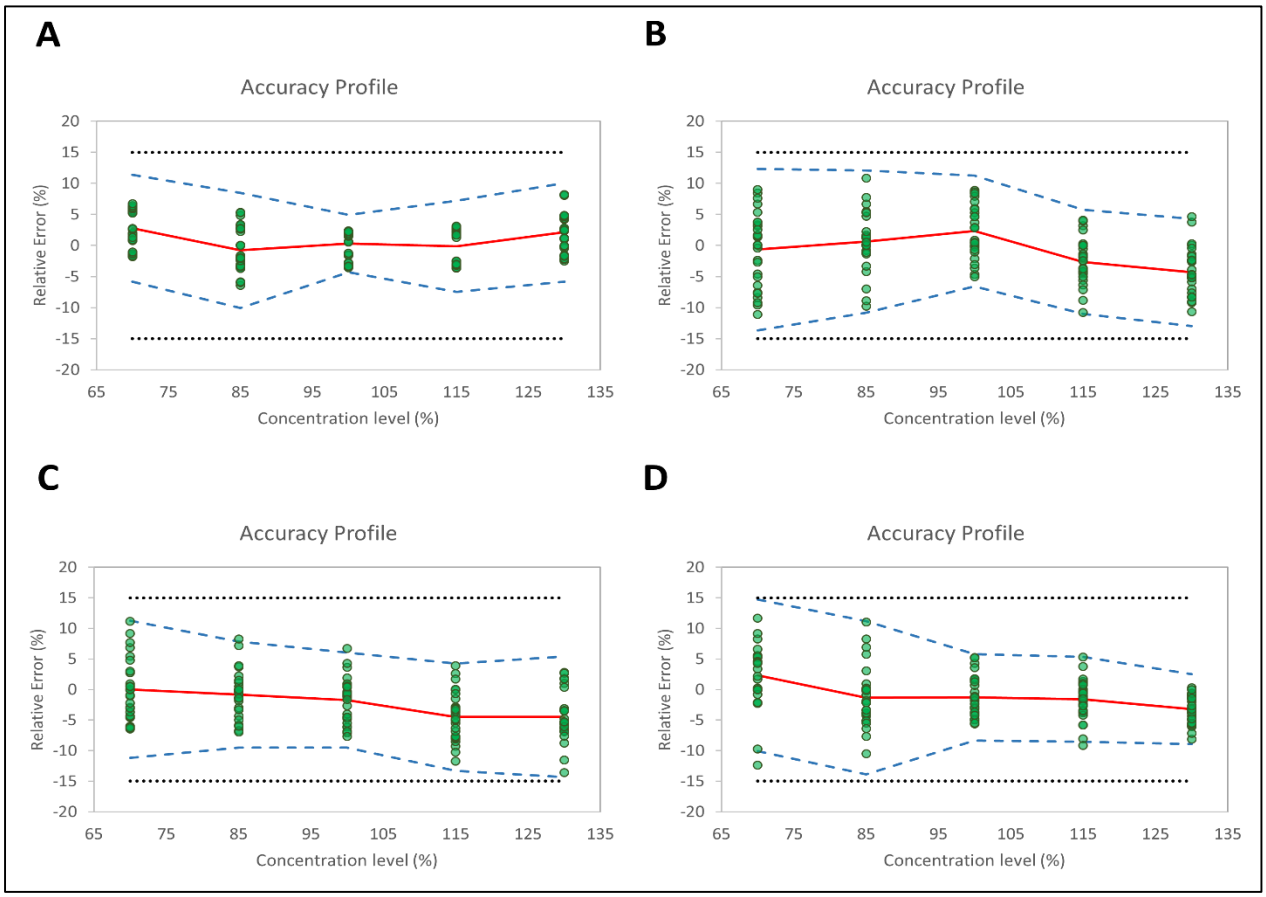

Figure 3: Accuracy profiles obtained with each handheld device (A: MPA, B: TruScan, C: IDRaman Mini, D: Resolve) in backscattering mode through the glass vial. The plain red line is the relative bias, the dashed blue lines are the β-expectation tolerance limits (β = 95%) and the dotted black lines are the acceptance limits set at 15%. The green dots represent the relative errors of each validation results.

3.2. Analysis of samples through glass vials placed in a polypropylene container:

Figure 4 shows the baseline corrected spectra acquired on each device in backscattering mode and SORS mode (for the Resolve) through the glass vial placed in the PP container for a mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. The dashed red lines indicate the main spectral features of ibuprofen and the blue dashed line indicates the main peak of mannitol.

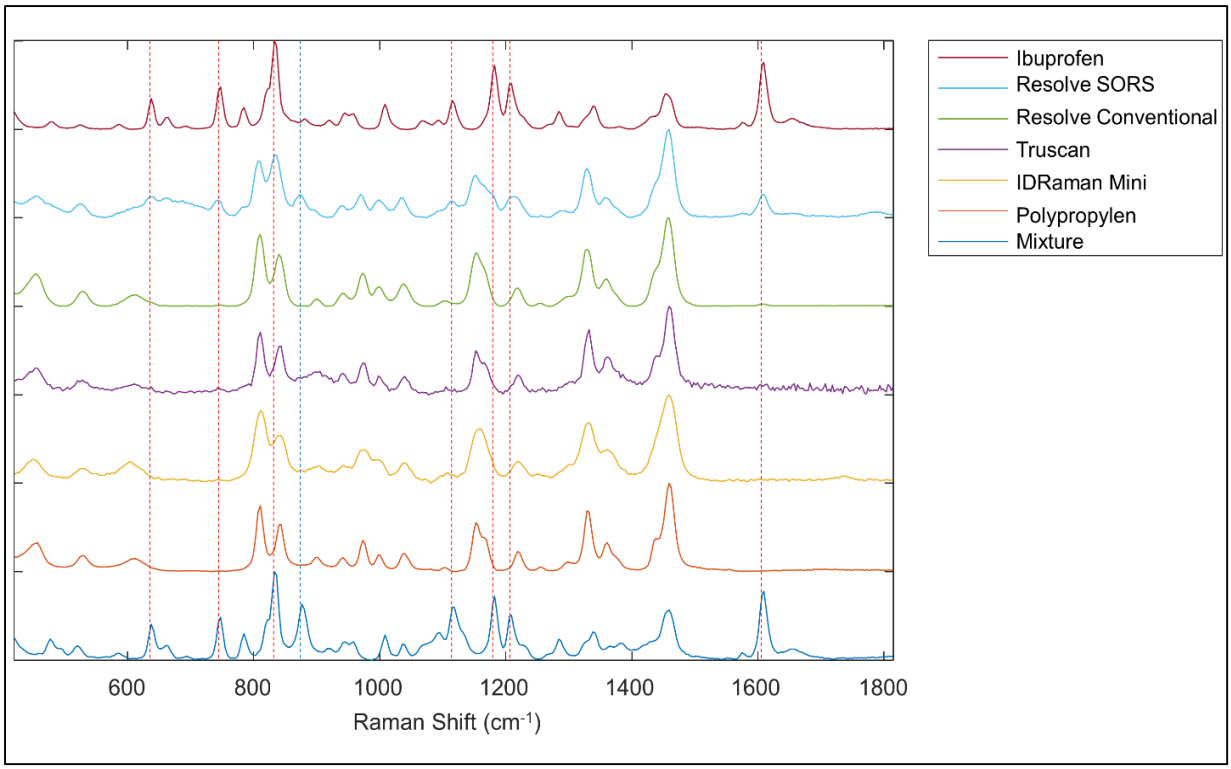

Figure 4: Raman spectra of a ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively. The spectra were acquired with each handheld device in the backscattering mode and the Resolve in SORS mode through the glass vial placed in the polypropylene container. The spectra were baseline corrected by asymmetric least squares with parameters: lambda 10^5, p: 10^{-3}. Reference spectra of ibuprofen and the ternary mixture were acquired as described in Figure 1. Ibuprofen and mannitol spectral features are marked by dashed red and blue lines respectively.

Compared to the spectra shown on Figure 2, the spectra recorded through the PP container show no spectral features associated with ibuprofen nor mannitol except for the SORS spectra. It is worth noting that the SORS spectra presented here and subsequently used in the quantitative modelling are only the offset part of the spectrum. Indeed, usually final SORS spectra are obtained after removal of the zero spectrum (equivalent to the backscattering recorded spectrum) from the offset spectrum to remove the residual container spectral features. Since the SORS correction parameters (baseline correction and removal of the scaled "zero offset" spectrum) are computed for each spectrum separately, this led to additional random error on the quantitative models. Therefore, to avoid errors when removing the zero offset, this step was skipped and the offset spectrum was directly used.

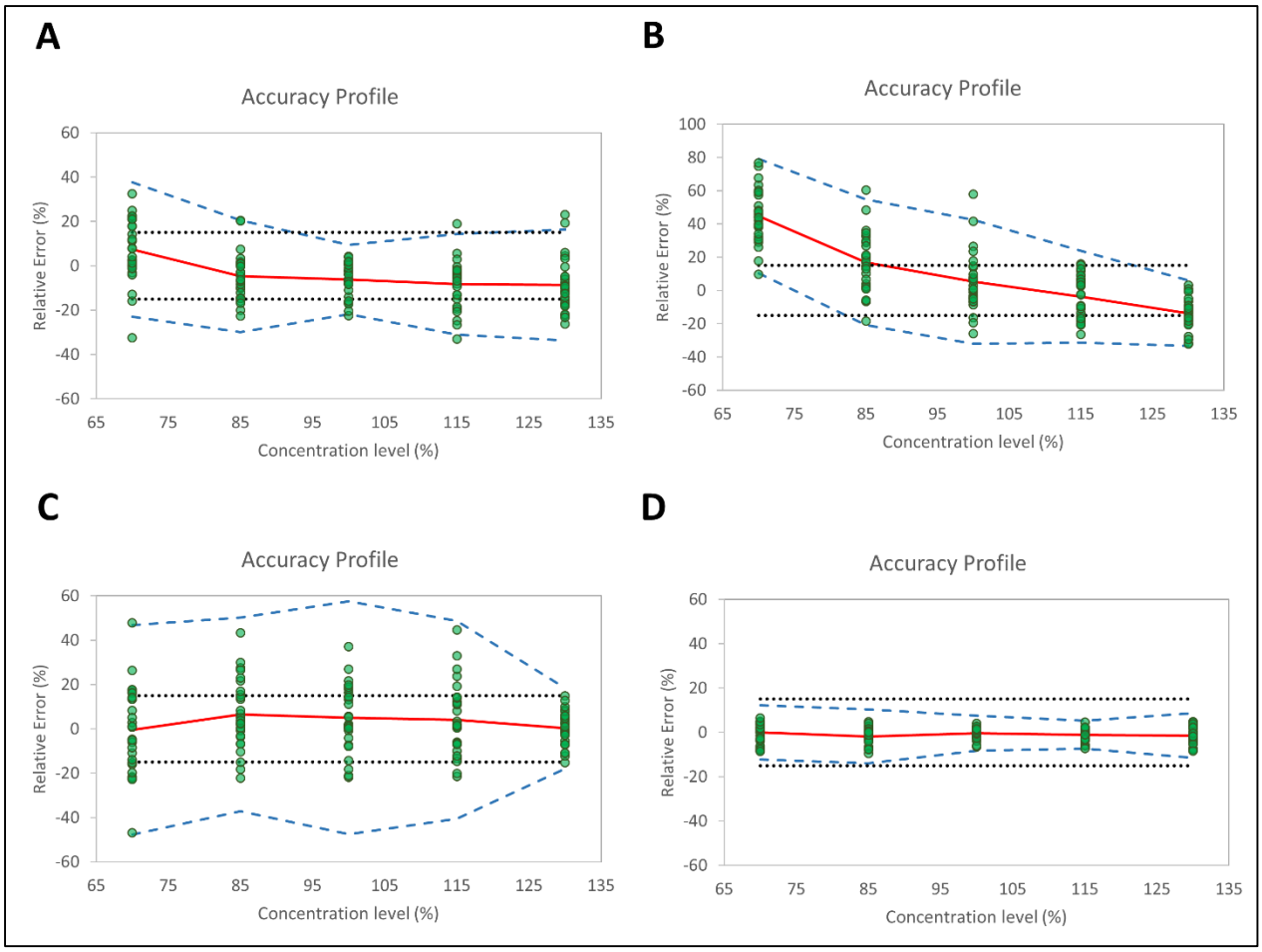

Figure 5: Accuracy profiles obtained with each handheld device (A: TruScan, B: IDRaman Mini, C: Resolve backscattering and D: Resolve SORS) mode through the glass vial placed in the opaque polypropylene container. The plain red line is the relative bias, the dashed blue lines are the β-expectation tolerance limits (β = 95%) and the dotted black lines are the acceptance limits set at 15%. The green dots represent the relative errors of each validation sample.

Once again, PLS models were built for each device and several pre-processing and spectral ranges were tested. Accuracy profiles were computed based on the predicted values from the PS models. The results are summarized in Figure 5 and the parameters values of the validated models are presented in table 3. None of the backscattering devices was able to quantify ibuprofen through the PP container because no (or very few) signal originating from the sample was measured in this configuration. Indeed, after applying the mean centering prior to the modelling, no residual signal was observed, only noise. That means that all spectra were the same for each validation and calibration sample because only the PP signal was recorded.

Table 4: ICH Q2 (R1) validation criteria values of the PLS models.

	Concentration level	**SORS (glass vial + PP)**	**Resolve (glass vial)**	**TruScan RM (glass vial)**	**IDRaman mini (glass vial)**	**FT-NIR**
Trueness Relative bias (%)	70	0.665	2.319	-0.647	0.034	2.790
	85	-1.226	-1.315	0.637	-0.848	-0.771
	100	-0.377	-1.289	2.332	-1.726	0.291
	115	-1.688	-1.587	-2.632	-4.516	-0.150
	130	-3.030	-3.197	-4.335	-4.484	2.138
Intra-assay precision Repeatability (RSD %)						
	70	3.308	5.813	5.953	5.206	2.338
	85	3.400	4.872	5.002	3.984	3.346
	100	2.406	3.375	4.244	3.707	2.205
	115	2.883	3.308	3.873	3.884	2.230
	130	3.483	1.968	3.586	3.617	2.777
Between-assay precision Intermediate precision (RSD %)						
	70	4.824	5.813	6.117	5.283	3.276
	85	3.400	5.525	5.251	4.071	3.971
	100	2.468	3.375	4.244	3.707	2.205
	115	3.331	3.308	3.922	4.089	2.932
	130	4.103	2.415	3.928	4.277	3.372
Accuracy Relative β - expectation tolerance limits (%)						
	70	[-12.52 , 13.85]	[-10.11 , 14.75]	[-13.64 , 12.35]	[-11.20 , 11.27]	[-5.82 , 11.40]
	85	[-8.41 , 5.95]	[-13.85 , 11.22]	[-10.79 , 12.06]	[-9.53 , 7.83]	[-10.04 , 8.50]
	100	[-5.60 , 4.85]	[-8.36 , 5.78]	[-6.58 , 11.25]	[-9.49 , 6.04]	[-4.33 , 4.91]
	115	[-9.25 , 5.87]	[-8.52 , 5.34]	[-10.99 , 5.73]	[-13.28 , 4.25]	[-7.48 , 7.18]
	130	[-12.47 , 6.41]	[-8.92 , 2.53]	[-12.97 , 4.30]	[-14.35 , 5.38]	[-5.79 , 10.07]

However, the Resolve operating in the SORS mode was able to achieve satisfying quantitation of the sample through the PP container since its accuracy profile was completely included in the acceptance limits. Furthermore, SORS measurements are more representative of the sample since the recorded signal has gone through a higher sample volume.

4. Conclusion

The aim of this study was to carry out a comparison study between conventional backscattering and Spatially Offset Raman Scattering (SORS) Raman handheld instruments. This comparison has been performed measuring a ternary mixture of ibuprofen/MCC/mannitol directly through a glass vial and with the glass vial placed in a PP container. PLS models were built for each device and each measurement configuration. The predicted values obtained from the PLS models on a validation set were used to compute accuracy profiles following the ICH Q2 R1 guidelines on validation with ±15% as acceptance limits.

By measuring through the glass vial, the Raman spectra showed clear features associated with both the API and the excipients leading to satisfying quantitative performances. The subsequent models and predictions were validated and their accuracy profiles were included inside the a priori defined acceptance limits. This confirms the fact that it is possible to obtain reliable quantitative information with handheld devices even with their auto-exposure default configuration. However, it is worth noting that this is only true for the studied formulation with a well-balanced signal between the API and excipients.

However, none of the backscattering Raman handheld devices was able to quantify the mixture when it was placed in an opaque 1 mm thick PP container. To be able to pass through the packaging, the SORS measurement configuration was necessary and allowed to obtain a valid PLS model.

These preliminary results pave the way to reliable quantitative Raman measurements directly in the field through opaque containers

Acknowledgments

This project has been supported by the European funds of regional development (FEDER) and by Walloon as part of the operational program “Walloon-2020.EU”.

The financial support of this research by the Walloon Region of Belgium in the framework of the Vibra4Fake project (convention n°:7517) is gratefully acknowledged.

This work was supported by the National Fund for Scientific Research, FNRS-F.R.S. (1.A030.17 - FC6921).

Chapter III

Conclusion & Perspectives

Pharmaceutical drug analysis belongs to the field of analytical chemistry that involves sets of process for the identification, determination and quantitation using analytical methods. Among of these methods, vibrational spectroscopic techniques are considered the most suitable to carry out different qualitative and quantitative drug analysis due to their fastest and simplest analysis compared to other analytical techniques, however, the usefulness of vibrational spectroscopic techniques cannot be appeared unless their association to chemometric approaches, to extract the important information from the analyzed samples and to overcome many challenges, that could be faced during the drug analysis. Matrix and interference effects are considered one of the challenges that can impact the accuracy of the results of the analyzed samples. These kinds of challenges can vary depending on the analytical technique and the sample to be analyzed itself.

This thesis was focused mainly on investigating three different aspects of matrix and interference effects that can be faced during the analysis of pharmaceutical samples whether qualitatively or quantitively and showing how using properties of vibrational spectroscopic and chemometric tools together can overcome different challenges of matrix effect.

The first aspect was investigated on the qualitative analysis and especially on discriminating between three polymorphic forms of fluconazole (Form II, Form III, and monohydrate form) through testing the ability of associating each of FT-IR and FT-NIR to PLS-DA to carry out discrimination between pharmaceutical products of fluconazole based on polymorphic forms. Firstly, the discrimination models by PLS-DA based on the data arose by FT-IR and FT-NIR were developed and evaluated on samples contained only one of three polymorphic forms. The obtained results proved the ability of PLS-DA to discriminate between different pharmaceutical products. Secondly, these developed models were tested on samples of binary mixtures containing two polymorphic forms (with different ratios of form II and monohydrate form), and samples that do not have fluconazole in case of substandard and falsified pharmaceutical products. Both last two cases represent a kind of matrix effect. In this case, PLS-DA models have shown their unusefulness to discriminate samples of binary mixtures and samples without fluconazole among samples containing one of three polymorphic forms. To overcome this problem, another chemometric approach Hotelling's T^2 and Q residuals was used successfully to detect these samples as outliers due to their high values of Hotelling's T^2 and Q residuals compared to the samples contained one polymorphic form. Thanks to this approach, it has been confirmed that FT-NIR is more suitable than FT-IR for detecting minor

contamination or conversion between monohydrate form and form-II due to its high sensitivity to hydrates.

The second aspect was carried out on the quantitative analysis through comparing two chemometric tools of MCR-ALS and PLS regression associated to FT-NIR for quantifying ciprofloxacin in different pharmaceutical formulations that do not contain the same excipients. At the first case, the quantitation was carried out on the set of formulations containing ciprofloxacin with the same excipients. In this case the obtained results showed the ability of both chemometric tools PLS and MCR-ALS. In the second case, the quantitation was carried on formulations in which the excipients are varied form one sample to another, also the quantitation was applied on different brands of pharmaceutical products. For both last cases, PLS model showed its limit due to its error that arose compared to the first case, whereas MCR-ALS model kept its low error even if the matrice composition was changed from one sample to another, which demonstrated the ability of MCR-ALS to deal with matrice effect and to predict the content of ciprofloxacin in different matrix compositions due to its second order advantage.

The last aspect was about evaluating the quantitation of ibuprofen in a ternary mixture through the interference of polypropylene container using Raman spectroscopy associated to the PLS approach. The aim was especially about comparing the ability of Spatially Offset Raman Spectroscopy (SORS) and backscattering modes to carry out the quantitation through a packaging interference. The developed PLS regression models were evaluated by mean of accuracy profiles following the ICH Q2 (R1) guidelines on validation with ±15% as acceptance limits. The analysis was carried out in the beginning of investigation only through the glass vial. The obtained results proved the quantitative abilities for both backscattering and SORS mode to quantify accurately ibuprofen in a ternary mixture. However, backscattering showed its drawback when it comes to quantify the ibuprofen through polypropylene container, whereas the SORS mode showed its ability to obtain a valid PLS model according to accuracy profile due to its offset property that allows the photons to pass through the packaging and obtaining the mixture spectra without being impacted by the presence of the interference of polypropylene container.

In conclusion, the three aspects of this thesis have proved the utility of combining the right vibrational spectroscopic technique with the suitable chemometric approach to overcome the

challenge of matrix or interference effect whenever it comes to carry out a pharmaceutical drug analysis.

Even though the satisfaction of the obtained results, each of these three aspects can be handled differently whether using different spectroscopic technique or applying new chemometric tool in order to improve the obtained results. For the first aspect, associating discrimination chemometric approach of Soft Independent Modelling of Class Analogies (SIMCA) to the data fusion of FT-IR and FT-NIR can lead to the best discrimination of binary polymorphic mixtures among the samples containing only one of three polymorphic forms without needing to the approach of Hoteling's T^2 and Q residuals. The same thing for the second approach, the data fusion between FT-NIR and Raman data can overcome the matrice effect and improve the results obtaining by MCR-ALS and especially by PLS regression, whereas the application of nonlinear models as neural network may improve the obtained results when it comes to carry out the quantitation of ibuprofen through polypropylene container using SORS mode and especially backscattering mode.

Chapter IV

References

[1] https://www.who.int/medicines/areas/quality_safety/quality_assurance/control/en/, (n.d.).

[2] S. Hansen, S. PEDERSEN-BJERGAARD, K. Rasmussen, Introduction to Pharmaceutical Analysis, second edi, 2012.

[3] S. Thakurta, Revisiting Aspirin, Paracetamol and Ibuprofen: Discovery of Synthetic Procedures and Mode of Actions, Trends Tech. Sci. Res. 04 (2020). doi:10.19080/TTSR.2020.04.555636.

[4] Z. Han, L. Lu, L. Wang, Z. Yan, X. Wang, Development and Validation of an HPLC Method for Simultaneous Determination of Ibuprofen and 17 Related Compounds, Chromatographia. 80 (2017) 1353–1360. doi:10.1007/s10337-017-3358-3.

[5] L. LEVI, G.C. WALKER, L.I. PUGSLEY, QUALITY CONTROL OF PHARMACEUTICALS, Can. Med. Assoc. J. 91 (1964) 781–785. https://pubmed.ncbi.nlm.nih.gov/14199105.

[6] USP 41—NF 36,The United States Pharmacopeia and National Formulary, United States, Pharmacopeial Convention Inc., 2018., (2018).

[7] European Pharmacopoeia (Ph. Eur.), European Directorate for the Quality of Medicines & HealthCare (EDQM), 10th ed., 2016.

[8] British pharmacopoeia, The British Pharmacopoeia Commission Secretariat of the Medicines and Healthcare Products Regulatory Agency (MHRA), TSO (The Stationery Office), (2019).

[9] R. Barret, 9 - The European Pharmacopoeia, in: R.B.T.-T.C. Barret (Ed.), Elsevier, 2018: pp. 135–144. doi:https://doi.org/10.1016/B978-1-78548-288-5.50009-3.

[10] ICH HARMONISED TRIPARTITE GUIDELINE Q4B, (2007).

[11] J.P. Boehlert, 1.5. Regulatory aspects: ICH and pharmacopoeial perspectives, in: S.B.T.-P. in P. and B.A. Görög (Ed.), Identif. Determ. Impurities Drugs, Elsevier, 2000: pp. 48–66. doi:https://doi.org/10.1016/S1464-3456(00)80007-1.

[12] M.R. Caira, Crystalline Polymorphism of Organic Compounds BT - Design of Organic Solids, in: E. Weber, Y. Aoyama, M.R. Caira, G.R. Desiraju, J.P. Glusker, A.D. Hamilton, R.E. Meléndez, A. Nangia (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 1998: pp. 163–208. doi:10.1007/3-540-69178-2_5.

[13] matrix in analysis, in: IUPAC Compend. Chem. Terminol., IUPAC, Research Triagle Park, NC, n.d. doi:10.1351/goldbook.M03758.

[14] W. Zhou, S. Yang, P.G. Wang, Matrix effects and application of matrix effect factor, Bioanalysis. 9 (2017) 1839–1844. doi:10.4155/bio-2017-0214.

[15] F. Gosetti, E. Mazzucco, D. Zampieri, M.C. Gennaro, Signal suppression/enhancement in high-performance liquid chromatography tandem mass spectrometry, J. Chromatogr. A. 1217 (2010) 3929–3937. doi:https://doi.org/10.1016/j.chroma.2009.11.060.

[16] M.D. Hargreaves, N.A. Macleod, M.R. Smith, D. Andrews, S. V Hammond, P. Matousek, Characterisation of transmission Raman spectroscopy for rapid quantitative analysis of intact multi-component pharmaceutical capsules, J. Pharm. Biomed. Anal. 54 (2011) 463–468. doi:https://doi.org/10.1016/j.jpba.2010.09.015.

[17] D.A. Gómez, J. Coello, S. Maspoch, The influence of particle size on the intensity and reproducibility of Raman spectra of compacted samples, Vib. Spectrosc. 100 (2019) 48–56. doi:https://doi.org/10.1016/j.vibspec.2018.10.011.

[18] 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. doi:10.1366/14-07645.

[19] J.M. Hollas, Modern spectroscopy, John Wiley & Sons, 2004.

[20] IUCN SSC Amphibian Specialist Group, Leptodactylus fallax: The IUCN Red List of Threatened Species 2017., XI (2017).

[21] S. Sokolov, I. Deyneka, Y. Katanaeva, N. Ugrekhelidze, E. Yatskova, N. Kulikova, The effectiveness evaluation of the carotenoids extraction from the tomatoes by means of absorption spectrophotometery method, E3S Web Conf. 175 (2020) 1011. doi:10.1051/e3sconf/202017501011.

[22] E.C.Y. Li-Chan, The applications of Raman spectroscopy in food science, Trends Food Sci. Technol. 11 (1996) 361–370.

[23] P.K. Adapa, C. Karunakaran, L.G. Tabil, G.J. Schoenau, Potential applications of infrared and Raman spectromicroscopy for agricultural biomass, Agric. Eng. Int. CIGR J. (2009).

[24] Infrared Spectroscopy, (2020). https://chem.libretexts.org/@go/page/1847.

[25] B.G. Osborne, Near-infrared spectroscopy in food analysis, Encycl. Anal. Chem. Appl. Theory Instrum. (2006).

[26] P. Larkin, Infrared and Raman spectroscopy: principles and spectral interpretation, Elsevier, 2017.

[27] L.E. Agelet, C.R. Hurburgh Jr, A tutorial on near infrared spectroscopy and its calibration, Crit. Rev. Anal. Chem. 40 (2010) 246–260.

[28] L.D.S. Yadav, Infrared (IR) spectroscopy, in: Org. Spectrosc., Springer, 2005: pp. 52–106.

[29] J. Coates, Encyclopedia of analytical chemistry, Interpret. Infrared Spectra, a Pract. Approach. Wiley, Chichester. (2000) 10815–10837.

[30] E. Mendes, N. Duarte, Mid-Infrared Spectroscopy as a Valuable Tool to Tackle Food Analysis: A Literature Review on Coffee, Dairies, Honey, Olive Oil and Wine, Foods. 10 (2021) 477. doi:10.3390/foods10020477.

[31] B.C. Smith, Quantitative spectroscopy: theory and practice, Elsevier, 2003.

[32] V. Sablinskas, Instrumentation. I: Handbook of Spectroscopy [e-bok], (2014).

[33] L.D.S. Yadav, Introduction to Spectroscopy (Spectrometry), in: Org. Spectrosc., Springer, 2005: pp. 1–6.

[34] B.C. Smith, Fundamentals of Fourier transform infrared spectroscopy, CRC press, 2011.

[35] K. Gerwert, C. Kötting, Fourier transform infrared (FTIR) spectroscopy, ELS. (2010).

[36] J.M.B.T.-R.M. in C. Chalmers Molecular Sciences and Chemical Engineering, INFRARED SPECTROSCOPY | Sample Presentation☆, in: Elsevier, 2013. doi:https://doi.org/10.1016/B978-0-12-409547-2.00254-7.

[37] S.J. Parikh, J. Chorover, FTIR Spectroscopic Study of Biogenic Mn-Oxide Formation by Pseudomonas putida GB-1, Geomicrobiol. J. 22 (2005) 207–218. doi:10.1080/01490450590947724.

[38] R. Karoui, G. Downey, C. Blecker, Mid-infrared spectroscopy coupled with chemometrics: a tool for the analysis of intact food systems and the exploration of their molecular structure− quality relationships− a review, Chem. Rev. 110 (2010) 6144–6168.

[39] I. Spectroscopy, Improved Performance with the Frontier UATR Accessory, (n.d.).

[40] Vibrational Spectroscopy, (2021). https://chem.libretexts.org/@go/page/2613.

[41] M. Jamrógiewicz, Application of the near-infrared spectroscopy in the pharmaceutical technology, J. Pharm. Biomed. Anal. 66 (2012) 1–10. doi:https://doi.org/10.1016/j.jpba.2012.03.009.

[42] A. Dutta, Fourier transform infrared spectroscopy, Spectrosc. Methods Nanomater. Charact. (2017) 73–93.

[43] V. Babushkin, A. Spiridonov, A. Kozhukhar, Application of NIR and FTIR in food analysis, J. Phys. Sci. Appl. 6 (2016) 47–50.

[44] How to Choose a Laser: How to choose a laser for Raman spectroscopy | Laser Focus World, (n.d.). https://www.laserfocusworld.com/lasers-sources/article/16555207/how-to-choose-a-laser-how-to-choose-a-laser-for-raman-spectroscopy (accessed November 18, 2021).

[45] J.M. Chalmers, H.G.M. Edwards, M.D. Hargreaves, Infrared and Raman spectroscopy in forensic science, John Wiley & Sons, 2012.

[46] S.M. Hassen, M.E.E. Mofdal, Using Raman Spectroscopy to Identify Unknown Materials, SSRG Int. J. Appl. Phys. (SSRG-IJAP). 5 (n.d.). http://www.internationaljournalssrg.org (accessed October 10, 2021).

[47] J.M. Chalmers, H.G.M. Edwards, M.D. Hargreaves, Vibrational spectroscopy techniques: basics and instrumentation, Infrared Raman Spectrosc. Forensic Sci. 2 (2012).

[48] D.A. Skoog, D.M. West, FJ Holler and SR Crouch, Princ. Instrum. Anal. 7 (2006).

[49] E. Smith, G. Dent, Modern Raman spectroscopy: a practical approach, John Wiley & Sons, 2019.

[50] Raman Spectroscopy, (2021). https://chem.libretexts.org/@go/page/111685.

[51] Introduction, Basic Theory and Principles, in: Mod. Raman Spectrosc. - A Pract. Approach, John Wiley & Sons, Ltd, Chichester, UK, 2005: pp. 1–21. doi:10.1002/0470011831.ch1.

[52] DoITPoMS - TLP Library Raman Spectroscopy - Method (dispersive Raman spectroscopy), (n.d.). https://www.doitpoms.ac.uk/tlplib/raman/method.php (accessed October 10, 2021).

[53] Raman Spectroscopy | THE DUFFY RESEARCH GROUP, (n.d.). https://duffy.princeton.edu/laboratory/raman-spectroscopy (accessed October 10, 2021).

[54] K. Buckley, P. Matousek, Recent advances in the application of transmission Raman spectroscopy to pharmaceutical analysis, J. Pharm. Biomed. Anal. 55 (2011) 645–652. doi:https://doi.org/10.1016/j.jpba.2010.10.029.

[55] A.W. Parker, Spatially Offset Raman Spectroscopy, in: Encycl. Spectrosc. Spectrom., Elsevier, 2017: pp. 143–148. doi:10.1016/B978-0-12-409547-2.12151-1.

[56] M.A. Mansouri, P.-Y. Sacré, L. Coïc, C. De Bleye, E. Dumont, A. Bouklouze, P. Hubert, R.D. Marini, E. Ziemons, Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study, Talanta. 207 (2020) 120306. doi:10.1016/J.TALANTA.2019.120306.

[57] DoITPoMS - TLP Library Raman Spectroscopy - Advantages and disadvantages, (n.d.). https://www.doitpoms.ac.uk/tlplib/raman/advantages_disadvantages.php (accessed October 11, 2021).

[58] A.P. Craig, A.S. Franca, J. Irudayaraj, Vibrational spectroscopy for food quality and safety screening, in: High Throughput Screen. Food Saf. Assess. Biosens. Technol. Hyperspectral Imaging Pract. Appl., Elsevier, 2015: pp. 165–194. doi:10.1016/B978-0-85709-801-6.00007-1.

[59] R.G. Brereton, Chemometrics in analytical chemistry. A review, Analyst. 112 (1987) 1635. doi:10.1039/an9871201635.

[60] R. Bro, Multivariate calibration: what is in chemometrics for the analytical chemist?, Anal. Chim. Acta 500,. (2003) 185–194.

[61] European Pharmacopoeia, Chemometric Methods Applied To Analytical Data, Eur. Pharmacopoeia. (2020) 5641–5658.

[62] P. Geladi, H. Grahn, The Philosophy and Fundamentals of Handling, Modeling, and Interpreting Large Data Sets—the Multivariate Chemometrics Approach, in: Multivar. Anal. Pharm. Ind., Elsevier, 2018: pp. 13–34. doi:10.1016/B978-0-12-811065-2.00003-5.

[63] D.E. Booth, Chemometrics: Data Analysis for the Laboratory and Chemical Plant, Technometrics. 46 (2004) 110–110. doi:10.1198/tech.2004.s738.

[64] W. Wu, D.L. Massart, S. de Jong, The kernel PCA algorithms for wide data. Part I: Theory and algorithms, Chemom. Intell. Lab. Syst. 36 (1997) 165–172. doi:10.1016/S0169-7439(97)00010-5.

[65] L. Mujica, J. Rodellar, A. Guemes, J. López-Diez, PCA based measures: Q-statistic and T2-statistic for assessing damages in structures, Proc. 4th Eur. Work. Struct. Heal. Monit. (2008) 1088–1095.

[66] L. Mujica, J. Rodellar, A. Fernández, A. Güemes, Q-statistic and T2-statistic PCA-based measures for damage assessment in structures, Struct. Heal. Monit. 10 (2011) 539–553. doi:10.1177/1475921710388972.

[67] R. Penha, J. Hines, Using principal component analysis modeling to monitor temperature sensors in a nuclear research reactor, Proc. Maint. Reliab. …. (2001) 6–9.

[68] O. Cloarec, Can we beat over–fitting?, J. Chemom. 28 (2014) 610–614.

[69] H. Martens, S.A. Jensen, Partial least squares regression: a new two-stage NIR calibration method, Dev. Food Sci. (1983).

[70] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130.

[71] M. Madakyaru, F. Harrou, Y. Sun, Monitoring Distillation Column Systems Using Improved Nonlinear Partial Least Squares-Based Strategies, IEEE Sens. J. 19 (2019) 11697–11705. doi:10.1109/JSEN.2019.2936520.

[72] M. Andersson, A comparison of nine PLS1 algorithms, J. Chemom. A J. Chemom. Soc. 23 (2009) 518–529.

[73] H. Abdi, Partial least square regression (PLS regression), Encycl. Res. Methods Soc. Sci. 6 (2003) 792–795.

[74] R. Bro, Multiway calibration. multilinear pls, J. Chemom. 10 (1996) 47–61.

[75] M. Barker, W. Rayens, Partial least squares for discrimination, J. Chemom. A J. Chemom. Soc. 17 (2003) 166–173.

[76] R.G. Brereton, G.R. Lloyd, Partial least squares discriminant analysis: taking the magic away, J. Chemom. 28 (2014) 213–225. doi:10.1002/cem.2609.

[77] R.G. Brereton, Applied chemometrics for scientists, John Wiley & Sons, 2007.

[78] T. Azzouz, R. Tauler, Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples, Talanta. 74 (2008) 1201–1210. doi:https://doi.org/10.1016/j.talanta.2007.08.024.

[79] J. Jaumot, R. Tauler, MCR-BANDS: A user friendly MATLAB program for the evaluation of rotation ambiguities in Multivariate Curve Resolution, Chemom. Intell. Lab. Syst. 103 (2010) 96–107.

[80] R.B. Pellegrino Vidal, A.C. Olivieri, A new parameter for measuring the prediction uncertainty produced by rotational ambiguity in second-order calibration with multivariate curve resolution, Anal. Chem. 92 (2020) 9118–9123.

[81] A.C. Olivieri, R. Tauler, The effect of data matrix augmentation and constraints in extended multivariate curve resolution–alternating least squares, J. Chemom. 31 (2017) e2875.

[82] J. Jaumot, R. Gargallo, A. de Juan, R. Tauler, A graphical user-friendly interface for MCR-ALS: a new tool for multivariate curve resolution in MATLAB, Chemom. Intell. Lab. Syst. 76 (2005) 101–110. doi:10.1016/j.chemolab.2004.12.007.

[83] E.R. Malinowski, Window factor analysis: Theoretical derivation and application to flow injection analysis data, J. Chemom. 6 (1992) 29–40.

[84] A. De Juan, J. Jaumot, R. Tauler, Multivariate Curve Resolution (MCR). Solving the mixture analysis problem, Anal. Methods. 6 (2014) 4964–4976.

[85] W. Windig, J. Guilment, Interactive self-modeling mixture analysis, Anal. Chem. 63 (1991) 1425–1432.

[86] R. Bro, S. De Jong, A fast non-negativity-constrained least squares algorithm, J. Chemom. A J. Chemom. Soc. 11 (1997) 393–401.

[87] M.C. Antunes, J.E. J. Simão, A.C. Duarte, R. Tauler, Multivariate curve resolution of overlapping voltammetric peaks: quantitative analysis of binary and quaternary metal mixtures, Analyst. 127 (2002) 809–817. doi:10.1039/b200243b.

[88] L.B. Lyndgaard, Application of Raman Spectroscopy and Multivariate Data Analysis in Food and Pharmaceutical Sciences, (2013).

[89] M. De Luca, G. Ioele, C. Spatari, G. Ragno, A single MCR-ALS model for drug analysis in different formulations: Application on diazepam commercial preparations, J. Pharm. Biomed. Anal. 134 (2017) 346–351. doi:10.1016/j.jpba.2016.10.022.

[90] R.G. Brereton, J. Jansen, J. Lopes, F. Marini, A. Pomerantsev, O. Rodionova, J.M. Roger, B. Walczak, R. Tauler, Chemometrics in analytical chemistry—part II: modeling, validation, and applications, Anal. Bioanal. Chem. 410 (2018) 6691–6704.

[91] A. Bouabidi, E. Rozet, M. Fillet, E. Ziemons, E. Chapuzet, B. Mertens, R. Klinkenberg, A. Ceccato, M. Talbi, B. Streel, A. Bouklouze, B. Boulanger, P. Hubert, Critical analysis of several analytical method validation strategies in the framework of the fit for purpose concept, J. Chromatogr. A. 1217 (2010) 3180–3192. doi:10.1016/j.chroma.2009.08.051.

[92] P. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, Harmonization of strategies for the validation of quantitative analytical procedures, J. Pharm. Biomed. Anal. 36 (2004) 579–586. doi:10.1016/j.jpba.2004.07.027.

[93] N.K. Afseth, V.H. Segtnan, J.P. Wold, Raman spectra of biological samples: A study of preprocessing methods, Appl. Spectrosc. 60 (2006) 1358–1367.

[94] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra, Appl. Spectrosc. 43 (1989) 772–777. doi:10.1366/0003702894202201.

[95] P. Geladi, D. MacDougall, H. Martens, Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat, Appl. Spectrosc. 39 (1985) 491–500. http://as.osa.org/abstract.cfm?URI=as-39-3-491.

[96] A. Savitsky, M.J. Golay, Smoothing and simplified differentiation of data by least squares procedures, Anal. Chem. 36 (1964) 1627–1639.

[97] Å. Rinnan, F. van den Berg, S.B. Engelsen, Review of the most common pre-processing techniques for near-infrared spectra, TrAC Trends Anal. Chem. 28 (2009) 1201–1222. doi:https://doi.org/10.1016/j.trac.2009.07.007.

[98] H.G. Brittain, S.R. Byrn, E. Lee, Polymorphism in pharmaceutical solids, 2009. doi:10.1016/S0168-3659(01)00252-8.

[99] J. Bernstein, J.M. Bernstein, Polymorphism in molecular crystals, Oxford University Press, 2002.

[100] X.Y. Lawrence, Pharmaceutical quality by design: product and process development, understanding, and control, Pharm. Res. 25 (2008) 781–791.

[101] H.G. Brittain, Theory and principles of polymorphic systems, Polymorph. Pharm. Solids. 192 (2009) 1.

[102] J. Aaltonen, M. Allesø, S. Mirza, V. Koradia, K.C. Gordon, J. Rantanen, Solid form screening--a review., Eur. J. Pharm. Biopharm. 71 (2009) 23–37. doi:10.1016/j.ejpb.2008.07.014.

[103] D. Singhal, W. Curatolo, Drug polymorphism and dosage form design: A practical perspective, Adv. Drug Deliv. Rev. 56 (2004) 335–347.

[104] H.K. Chan, E. Doelker, Polymorphic transformation of some drugs under compression, Drug Dev. Ind. Pharm. 11 (1985) 315–332.

[105] G.G.Z. Zhang, D. Law, E.A. Schmitt, Y. Qiu, Phase transformation considerations during process development and manufacture of solid oral dosage forms, Adv. Drug Deliv. Rev. 56 (2004) 371–390.

[106] E. Atef, H. Chauhan, D. Prasad, D. Kumari, C. Pidgeon, Quantifying solid-state mixtures of crystalline indomethacin by raman spectroscopy comparison with thermal analysis, ISRN Chromatogr. 2012 (2012).

[107] R. Suryanarayanan, Determination of the Relative Amounts of α-carbamazepine and β-carbamazepine in a Mixture by Powder X-Ray Diffractometry, Powder Diffr. 5 (1990) 155–159. doi:DOI: 10.1017/S0885715600015608.

[108] D.E. Bugay, A.W. Newman, W.P. Findlay, Quantitation of cefepime · 2HCl dihydrate in cefepime · 2HCl monohydrate by diffuse reflectance IR and powder X-ray diffraction techniques, J. Pharm. Biomed. Anal. 15 (1996) 49–61. doi:https://doi.org/10.1016/0731-7085(96)01796-7.

[109] A.A. Bunaciu, H.Y. Aboul-Enein, V.D. Hoang, Vibrational spectroscopy used in polymorphic analysis, TrAC Trends Anal. Chem. 69 (2015) 14–22.

[110] S.N.C. Roberts, A.C. Williams, I.M. Grimsey, S.W. Booth, Quantitative analysis of mannitol polymorphs. X-ray powder diffractometry—exploring preferred orientation effects, J. Pharm. Biomed. Anal. 28 (2002) 1149–1159.

[111] A. Heinz, M. Savolainen, T. Rades, C.J. Strachan, Quantifying ternary mixtures of different solid-state forms of indomethacin by Raman and near-infrared spectroscopy, Eur. J. Pharm. Sci. 32 (2007) 182–192.

[112] S. Jabeen, T.J. Dines, S.A. Leharne, B.Z. Chowdhry, Raman and IR spectroscopic studies of fenamates - Conformational differences in polymorphs of flufenamic acid, mefenamic acid and tolfenamic acid, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 96 (2012) 972–985.

[113] Å. Rinnan, Pre-processing in vibrational spectroscopy – when, why and how, Anal. Methods. 6 (2014). doi:10.1039/C3AY42270D.

[114] V. Bhavana, R.B. Chavan, M.K.C. Mannava, A. Nangia, N.R. Shastri, Quantification of niclosamide polymorphic forms – A comparative study by Raman, NIR and MIR using chemometric techniques, Talanta. 199 (2019) 679–688. doi:https://doi.org/10.1016/j.talanta.2019.03.027.

[115] N.L. Calvo, T.S. Kaufman, R.M. Maggio, A PCA-based chemometrics-assisted ATR-FTIR approach for the classification of polymorphs of cimetidine: Application to physical mixtures and tablets, J. Pharm. Biomed. Anal. 107 (2015) 419–425. doi:10.1016/j.jpba.2015.01.016.

[116] N.L. Calvo, R.M. Maggio, T.S. Kaufman, A dynamic thermal ATR-FTIR/chemometric approach to the analysis of polymorphic interconversions. Cimetidine as a model drug, J. Pharm. Biomed. Anal. 92 (2014) 90–97. doi:https://doi.org/10.1016/j.jpba.2013.12.036.

[117] Ł. Górski, W. Sordoń, F. Ciepiela, W.W. Kubiak, M. Jakubowska, Voltammetric classification of ciders with PLS-DA, Talanta. 146 (2016) 231–236. doi:https://doi.org/10.1016/j.talanta.2015.08.027.

[118] V.A.G. da Silva, M. Talhavini, I.C.F. Peixoto, J.J. Zacca, A.O. Maldaner, J.W.B. Braga, Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis, Microchem. J. 116 (2014) 235–243. doi:https://doi.org/10.1016/j.microc.2014.05.013.

[119] S.R. Byrn, J.R.R. Pfeiffer, G. Stephenson, D.J.W. Grant, W.B. Gleasonq, Solid-state Pharmaceutical Chemistry, Chem. Mater. 6 (1994) 1148–1158. doi:10.1021/cm00044a013.

[120] K. Richardson, K.W. Brammer, M.S. Marriott, P.F. Troke, Activity of UK-49,858, a bis-triazole derivative, against experimental infections with Candida albicans and Trichophyton mentagrophytes., Antimicrob. Agents Chemother. 27 (1985) 832–835. doi:10.1128/AAC.27.5.832.

[121] C. Koks, P. Meenhorst, M. Hillebrand, A. Bult, J. Beijnen, Pharmacokinetics of fluconazole in saliva and plasma after administration of an oral suspension and capsules, Antimicrob. Agents Chemother. 40 (1996) 1935–1937. http://aac.asm.org/cgi/content/long/40/8/1935 (accessed November 18, 2016).

[122] E. Ziémons, H. Bourichi, J. Mantanus, E. Rozet, P. Lebrun, E. Essassi, Y. Cherrah, A. Bouklouze, P. Hubert, Determination of binary polymorphic mixtures of fluconazole using near infrared spectroscopy and X-ray powder diffraction: A comparative study based on the pre-validation stage results, J. Pharm. Biomed. Anal. 55 (2011) 1208–1212. doi:10.1016/j.jpba.2011.02.019.

[122] H. Bourichi, Y. Brik, E. Essassi, A. Bouklouze, Y. Cherrah, Solid-state characterization of fluconazole, STP Pharma Tech. Prat. Reglementations. 17 (2007) 49.

[123] H. Bourichi, Y. Brik, P. Hubert, Y. Cherrah, A. Bouklouze, Solid-state characterization and impurities determination of fluconazol generic products marketed in Morocco, J. Pharm. Anal. 2 (2012) 412–421. doi:10.1016/j.jpha.2012.05.007.

[124] K.A. Alkhamis, M.S. Salem, R.M. Obaidat, Comparison between dehydration and desolvation kinetics of fluconazole monohydrate and fluconazole ethylacetate solvate using three different methods, J. Pharm. Sci. 95 (2006) 859–870. doi:https://doi.org/10.1002/jps.20605.

[125] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: A basic tool of chemometrics, Chemom. Intell. Lab. Syst. 58 (2001) 109–130. doi:10.1016/S0169-7439(01)00155-1.

[126] M.C. Ortiz, L. Sarabia, R. García-Rey, M.D.L. de Castro, Sensitivity and specificity of PLS-class modelling for five sensory characteristics of dry-cured ham using visible and near infrared spectroscopy, Anal. Chim. Acta. 558 (2006) 125–131. doi:https://doi.org/10.1016/j.aca.2005.11.038

[127] B. Igne, C. Airiau, S. Talwar, E. Towns, 4.02 - Chemometrics in the Pharmaceutical Industry, in: S. Brown, R. Tauler, B.B.T.-C.C. (Second E. Walczak (Eds.), Elsevier, Oxford, 2020: pp. 33–68. doi:https://doi.org/10.1016/B978-0-12-409547-2.14638-4.

[128] V.H. da Silva, J.L. Soares-Sobrinho, C.F. Pereira, Å. Rinnan, Evaluation of chemometric approaches for polymorphs quantification in tablets using near-infrared hyperspectral images, Eur. J. Pharm. Biopharm. 134 (2019) 20–28. doi:https://doi.org/10.1016/j.ejpb.2018.11.007.

[129] M. Alaoui Mansouri, E. Ziemons, P.-Y. Sacré, M. Kharbach, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy, J. Pharm. Biomed. Anal. 196 (2021). doi:10.1016/j.jpba.2021.113922.

[130] C. De Bleye, P.F. Chavez, J. Mantanus, R. Marini, P. Hubert, E. Rozet, E. Ziemons, Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications, J Pharm Biomed Anal. 69 (2012) 125–132. doi:10.1016/j.jpba.2012.02.003.

[131] J. Luypaert, D.L. Massart, Y. Vander Heyden, Near-infrared spectroscopy applications in pharmaceutical analysis, Talanta. 72 (2007) 865–883. doi:https://doi.org/10.1016/j.talanta.2006.12.023.

[132] M. Jamrógiewicz, Application of the near-infrared spectroscopy in the pharmaceutical technology, J. Pharm. Biomed. Anal. 66 (2012) 1–10. doi:https://doi.org/10.1016/j.jpba.2012.03.009.

[133] H. Moustafa, Y. Fayez, Spectrophotometric methods manipulating ratio spectra for simultaneous determination of binary mixtures with sever overlapping spectra: A comparative study, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 133 (2014) 759–766. doi:https://doi.org/10.1016/j.saa.2014.06.059.

[134] M. De Luca, G. Ioele, C. Spatari, G. Ragno, Optimization of wavelength range and data interval in chemometric analysis of complex pharmaceutical mixtures, J. Pharm. Anal. 6 (2016) 64–69. doi:https://doi.org/10.1016/j.jpha.2015.10.001.

[135] A.M. Yehia, H.M. Mohamed, Chemometrics resolution and quantification power evaluation: Application on pharmaceutical quaternary mixture of Paracetamol, Guaifenesin, Phenylephrine and p-aminophenol, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 152 (2016) 491–500. doi:https://doi.org/10.1016/j.saa.2015.07.101.

[136] M. De Luca, F. Oliverio, G. Ioele, G. Ragno, Multivariate calibration techniques applied to derivative spectroscopy data for the analysis of pharmaceutical mixtures, Chemom. Intell. Lab. Syst. 96 (2009) 14–21. doi:https://doi.org/10.1016/j.chemolab.2008.10.009.

[137] X. Wang, D.-Z. Mao, Y.-J. Yang, Calibration transfer between modelled and commercial pharmaceutical tablet for API quantification using backscattering NIR, Raman and transmission Raman spectroscopy (TRS), J. Pharm. Biomed. Anal. 194 (2021) 113766. doi:https://doi.org/10.1016/j.jpba.2020.113766.

[138] T. Rajalahti, O.M. Kvalheim, Multivariate data analysis in pharmaceutics: A tutorial review, Int. J. Pharm. 417 (2011) 280–290. doi:https://doi.org/10.1016/j.ijpharm.2011.02.019.

[139] M. De Luca, G. Ioele, C. Spatari, G. Ragno, A single MCR-ALS model for drug analysis in different formulations: Application on diazepam commercial preparations, J. Pharm. Biomed. Anal. 134 (2017) 346–351. doi:10.1016/j.jpba.2016.10.022.

[140] L. Pinto, F. Stechi, M.C. Breitkreitz, A simplified and versatile multivariate calibration procedure for multiproduct quantification of pharmaceutical drugs in the presence of interferences using first order data and chemometrics, Microchem. J. 146 (2019) 202–209. doi:10.1016/j.microc.2019.01.014.

[141] S.M. Tawakkol, M. Farouk, O. Abd Elaziz, A. Hemdan, M.A. Shehata, Comparative study between univariate spectrophotometry and multivariate calibration as analytical tools for simultaneous quantitation of Moexipril and Hydrochlorothiazide, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 133 (2014) 300–306.

[142] E. Dinc, G. Ragno, G. Ioele, D. Baleanu, DRUG FORMULATIONS AND CLINICAL METHODS-Fractional Wavelet Analysis for the Simultaneous Quantitative Analysis of Lacidipine and Its Photodegradation Product by Continuous Wavelet Transform and Multilinear Regression Calibration, J. AOAC Int. 89 (2006) 1538.

[143] D. Sorak, L. Herberholz, S. Iwascek, S. Altinpinar, F. Pfeifer, H.W. Siesler, New developments and applications of handheld raman, mid-infrared, and near-infrared spectrometers, Appl. Spectrosc. Rev. 47 (2012) 83–115. doi:10.1080/05704928.2011.625748.

[144] R. Deidda, P.-Y. Sacre, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: Critical review of innovative portable and handheld NIR and Raman spectrophotometers, TrAC Trends Anal. Chem. 114 (2019) 251–259. doi:10.1016/j.trac.2019.02.035.

[145] J.K. Mbinze, P.-Y. Sacré, A. Yemoa, J. Mavar Tayey Mbay, V. Habyalimana, N. Kalenda, P. Hubert, R.D. Marini, E. Ziemons, Development, validation and comparison of NIR and Raman methods for the identification and assay of poor-quality oral quinine drops, J. Pharm. Biomed. Anal. 111 (2015) 21–27. doi:10.1016/j.jpba.2015.02.049.

[146] C. Eliasson, P. Matousek, Noninvasive Authentication of Pharmaceutical Products through Packaging Using Spatially Offset Raman spectroscopy, Anal. Chem. 79 (2007) 1696–1701.

[147] D. Riolo, A. Piazza, C. Cottini, M. Serafini, E. Lutero, E. Cuoghi, L. Gasparini, D. Botturi, I.G. Marino, I. Aliatis, D. Bersani, P.P. Lottici, Raman spectroscopy as a PAT for pharmaceutical blending: Advantages and disadvantages, J. Pharm. Biomed. Anal. 149 (2018) 329–334. doi:10.1016/j.jpba.2017.11.030.

[148] Michael J. Pelletier, Analytical Applications of Raman Spectroscopy, Wiley-Blackwell, Oxford, UK, 1999.

[149] C.E. Miller, Process analytical technology : spectroscopic tools and implemented strategies for the chemical and pharmaceutical industries, in: Chemom. Process Anal. Chem., John Wiley & Sons, 2010: pp. 237–285. http://www.worldcat.org/isbn/047072207.

[150] L.B. Lyndgaard, F. van den Berg, A. de Juan, Quantification of paracetamol through tablet blister packages by Raman spectroscopy and multivariate curve resolution-alternating least squares, Chemom. Intell. Lab. Syst. 125 (2013) 58–66. doi:10.1016/j.chemolab.2013.03.014.

[151] A.W. Parker, Spatially Offset Raman Spectroscopy, in: Encycl. Spectrosc. Spectrom., Elsevier, 2017: pp. 143–148. doi:10.1016/B978-0-12-409547-2.12151-1.

[152] W.J. Olds, E. Jaatinen, P. Fredericks, B. Cletus, H. Panayiotou, E.L. Izake, Spatially offset Raman spectroscopy (SORS) for the analysis and detection of packaged pharmaceuticals and concealed drugs, Forensic Sci. Int. 212 (2011) 69–77. doi:10.1016/j.forsciint.2011.05.016.

[153] M. Bloomfield, D. Andrews, P. Loeffen, C. Tombling, T. York, P. Matousek, Non-invasive identification of incoming raw pharmaceutical materials using Spatially Offset Raman Spectroscopy, J. Pharm. Biomed. Anal. 76 (2013) 65–69. doi:10.1016/j.jpba.2012.11.046.

[154] D.I. Ellis, R. Eccles, Y. Xu, J. Griffen, H. Muhamadali, P. Matousek, I. Goodall, R. Goodacre, Through-container, extremely low concentration detection of multiple chemical markers of counterfeit alcohol using a handheld SORS device, Sci. Rep. 7 (2017) 1–8. doi:10.1038/s41598-017-12263-0.

[155] P. Zhang, D. Littlejohn, Interference assessment and correction in the partial least squares regression method for multicomponent determination by UV spectrophotometry, Chemom. Intell. Lab. Syst. 34 (1996) 203–215. doi:10.1016/0169-7439(96)00029-9.

[156] M. Forina, S. Lanteri, M. Casale, Multivariate calibration, J. Chromatogr. A. 1158 (2007) 61–93. doi:10.1016/j.chroma.2007.03.082.

[157] P. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, Harmonization of strategies for the validation of quantitative analytical procedures, J. Pharm. Biomed. Anal. 36 (2004) 579–586. doi:10.1016/j.jpba.2004.07.027.

[158] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures: a SFSTP proposal part IV. Examples of application, J Pharm Biomed Anal. 48 (2008) 760–771. doi:10.1016/j.jpba.2008.07.018.

[159] A. Bouabidi, E. Rozet, M. Fillet, E. Ziemons, E. Chapuzet, B. Mertens, R. Klinkenberg, A. Ceccato, M. Talbi, B. Streel, A. Bouklouze, B. Boulanger, P. Hubert, Critical analysis of several analytical method validation strategies in the framework of the fit for purpose concept, J. Chromatogr. A. 1217 (2010) 3180–3192. doi:10.1016/j.chroma.2009.08.051.

[160] P. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, The SFSTP guide on the validation of chromatographic methods for drug bioanalysis from the Washington Conference to the laboratory, Anal. Chim. Acta. 391 (1999) 135–148.

[161] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal--part III, J Pharm Biomed Anal. 45 (2007) 82–96. doi:10.1016/j.jpba.2007.06.032.

[162] C. De Bleye, P.F. Chavez, J. Mantanus, R. Marini, P. Hubert, E. Rozet, E. Ziemons, Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications, J Pharm Biomed Anal. 69 (2012) 125–132. doi:10.1016/j.jpba.2012.02.003.

[163] A. Savitzky, M.J.E. Golay, Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal Chem. 36 (1964) 1627–1639. doi:10.1021/ac60214a047.

Chapter V

Scientific production

1. M.Alaoui Mansouri, P.-Y. Sacré, L. Coïc, C. De Bleye, E. Dumont, A. Bouklouze, P. Hubert, R.D. Marini, E. Ziemons, Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study, Talanta. 207 (2020)

2. M. Alaoui Mansouri, E. Ziemons, P.-Y. Sacré, M. Kharbach, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy, J. Pharm. Biomed. Anal. 196 (2021)

3. M. Kharbach, R. Kamal, M. Bousrabat, **M. Alaoui Mansouri**, I. Barra, K. Alaoui, Y. Cherrah, Y. Vander Heyden, A. Bouklouze, Characterization and classification of PGI Moroccan Argan oils based on their FTIR fingerprints and chemical composition, Chemom. Intell. Lab. Syst. (2017)

4. M. Kharbach, R. Kamal, **M. Alaoui Mansouri**, I. Marmouzi, J. Viaene, Y. Cherrah, K. Alaoui, J. Vercammen, A. Bouklouze, Y. Vander Heyden, Selected-ion flow-tube mass-spectrometry (SIFT-MS) fingerprinting versus chemical profiling for geographic traceability of Moroccan Argan oils, Food Chem. 263 (2018)

5. I. Barra, **M.Alaoui Mansouri**, M. Bousrabat, Y. Cherrah, A. Bouklouze, M. Kharbach, Discrimination and quantification of Moroccan gasoline adulteration with diesel using Fourier transform infrared spectroscopy and chemometric tools, J. AOAC Int. 102 (2019).

6. I. Barra, **M.Alaoui Mansouri,** Y. Cherrah, M. Kharbach, A. Bouklouze, FTIR fingerprints associated to a PLS-DA model for rapid detection of smuggled non-compliant diesel marketed in Morocco, Vib. Spectrosc. 101 (2019)

7. P.-Y. Sacré, **M. Alaoui Mansouri**, C. De Bleye, L. Coïc, P. Hubert, E. Ziemons, Evaluation of distributional homogeneity of pharmaceutical formulation using laser direct infrared imaging, Int. J. Pharm. 612 (2022)

Oral Comunication

1- M. Alaoui Mansouri, E. Ziemons, P.-Y. Sacré, M. Kharbach, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy. 7ème rencontre internationale sur la chimiométrie et la qualité. Faculté des Sciences et Techniques- Fès. October 2018

2- M.Alaoui Mansouri, P.-Y. Sacré, L. Coïc, C. De Bleye, E. Dumont, A. Bouklouze, P. Hubert, R.D. Marini, E. Ziemons, Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study. Belgian society of pharmaceutical scieces May 20th , 2019

Poster

M. Alaoui Mansouri, E. Ziemons, P.-Y. Sacré, M. Kharbach, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy, conférence de la chimiométrie, Paris, January 2018

M. Alaoui Mansouri, E. Ziemons, P.-Y. Sacré, M. Kharbach, I. Barra, Y. Cherrah, P. Hubert, R.D. Marini, A. Bouklouze, Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy, the 17th Chemometrics in Aalytical Chemistry (CAC), Halifax, Canada June 2018

M.Alaoui Mansouri, P.-Y. Sacré, L. Coïc, C. De Bleye, E. Dumont, A. Bouklouze, P. Hubert, R.D. Marini, E. Ziemons, Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study. Conférence chimiométrie, Liège January 2020

Chapter V

Appendices

International Journal of Pharmaceutics 612 (2022) 121373

Contents lists available at ScienceDirect

International Journal of Pharmaceutics

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

Evaluation of distributional homogeneity of pharmaceutical formulation using laser direct infrared imaging

P.-Y. Sacré *, M. Alaoui Mansouri, C. De Bleye, L. Coïc, Ph. Hubert, E. Ziemons

University of Liege (ULiege), CIRM, Vibra-Santé Hub, Department of Pharmacy, Laboratory of Pharmaceutical Analytical Chemistry, Liege, Belgium

ARTICLE INFO

Keywords:
Laser direct infrared imaging
Raman imaging
Homogeneity
Vibrational spectroscopy
Distributional Homogeneity Index (DHI)
Pharmaceutical formulations development

ABSTRACT

The distributional homogeneity of chemicals is a key parameter of solid pharmaceutical formulations. Indeed, it may affect the efficacy of the drug and consequently its safety.

Chemical imaging offers a unique insight enabling the visualisation of the different constituents of a pharmaceutical tablet. It allows identifying ingredients poorly distributed offering the possibility to optimize the process parameters or to adapt characteristics of incoming raw materials to increase the final product quality.

Among the available chemical imaging tools, Raman imaging is one of the most widely used since it offers a high spatial resolution with well-resolved peaks resulting in a high spectral specificity. However, Raman imaging suffers from sample autofluorescence and long acquisition times. Recently commercialised, laser direct infrared reflectance imaging (LDIR) is a quantum cascade laser (QCL) based imaging technique that offers the opportunity to rapidly analyse samples.

In this study, a typical pharmaceutical formulation blend composed of two active pharmaceutical ingredients and three excipients was aliquoted at different mixing timepoints. The collected aliquots were tableted and analysed using both Raman and LDIR imaging. The distributional homogeneity indexes of one active ingredient image were then computed and compared. The results show that both techniques achieved similar conclusions. However, the analysis times were drastically different. While Raman imaging required a total analysis time of 4 h per tablet to obtain the distribution map of acetylsalicylic acid with a step size of 100 µm, it only took 7.5 min to achieve the same result with LDIR.

The results obtained in the present study show that LDIR is a promising technique for the analysis of pharmaceutical formulations and that it could be a valuable tool when developing new pharmaceutical formulations.

1. Introduction

The homogeneity of distribution of active pharmaceutical ingredients (API) but also excipients is of utmost importance for the safety, efficacy, and final quality of solid pharmaceutical formulations. Therefore, the homogeneity of distribution of the ingredients is a key element to follow during the development of a new formulation. This homogeneity is generally assessed via chapter 2.9.40 "Uniformity of dosage units" of the European Pharmacopoeia requiring the dosage of the active ingredient in several units to compute the so-called acceptance value. However, this approach only provides insight into the inter-tablet homogeneity of the active ingredient. Therefore, the European Pharmacopoeia allows the pharmaceutical analyst to use chemical imaging (chapter 5.24 "chemical imaging") to explore the homogeneity of distribution of solid dosage forms ("intra-tablet" homogeneity) since it allows to assess the spatial distribution of API but also excipients (Sacré et al., 2015; 2014a) and is therefore recommended during the development of new formulations (Bøtker et al., 2020). Several techniques may be used to perform chemical imaging analysis of pharmaceutical formulations such as vibrational spectroscopy (Sacré et al., 2014a), mass spectrometry (Belu et al., 2000; Earnshaw et al., 2010), SEM-EDX (Scoutaris et al., 2014) or UV spectroscopy (Klukkert et al., 2016). Among these, vibrational spectroscopy and more particularly near-infrared (NIR) and Raman imaging (Gowen et al., 2008) are the most frequently used.

On the one hand, NIR imaging allows the fast analysis of several samples (Prats-Montalbán et al., 2012) but suffers from a low spatial resolution and a low spectral specificity (Carruthers et al., 2021). These

* Corresponding author at: Laboratory of Pharmaceutical Analytical Chemistry, CIRM, Vibra-Santé Hub, Department of Pharmacy, University of Liege (ULiege), Avenue Hippocrate 15, B-4000 Liege, Belgium.
E-mail address: pysacre@uliege.be (P.-Y. Sacré).

https://doi.org/10.1016/j.ijpharm.2021.121373
Received 20 September 2021; Received in revised form 30 November 2021; Accepted 7 December 2021
Available online 11 December 2021

two drawbacks imply the need for chemometrics and highly trained users to set up imaging experiments and complicate the analysis of finely dispersed pharmaceutical ingredients.

On the other hand, Raman imaging has a high spatial resolution (due to the shorter wavelength used) and a high spectral specificity enabling faster development of imaging experiments. However, Raman imaging suffers from sample autofluorescence, high risk of sample burning requiring the use of low laser power, and is usually performed in a point-mapping fashion implying long acquisition times (Stewart et al., 2012). Nevertheless, the hyperspectral data cube acquired enables the analysis of unknown formulations (Coic et al., 2019) making the development of a Raman imaging method quite straightforward.

Therefore, there is a need for fast and high-resolution imaging techniques enabling the generation of results within minutes. Non-linear Raman imaging techniques (coherent anti-stokes, CARS, or stimulated Raman spectroscopy, SRS) have recently been introduced with commercial equipment (Novakovic et al., 2017). These techniques, however, still suffer from sample autofluorescence, require highly trained staff to set up the imaging experiments and are yet very expensive.

The recent technological developments of quantum cascade lasers (QCL) (Bhargava, 2012) made them available in commercial spectroscopy equipment. Recently introduced, Laser Direct Infrared Reflectance (LDIR) imaging couples a QCL source with a single diode mercury cadmium telluride (MCT) detector. The device allows the imaging of samples at different spatial resolutions (1–40 μm) at relatively high speed. The images are acquired at a single wavelength between 975 and 1800 cm^{-1}. Therefore, to ensure the specificity of the imaging method, the sample constituents must be known *a priori* and must exhibit specific peaks corresponding to the studied compound without any interfering adjacent peak related to other compounds. The software allows both mappings of single peak intensities and classification methods. LDIR imaging has been successfully used to analyse microplastic pollution in water (Scircle et al., 2020) and to detect food raw materials adulteration (da Costa Filho et al., 2020).

The objective of this study was to evaluate the suitability of LDIR imaging to assess the homogeneity of distribution in a model pharmaceutical formulation. To achieve this goal, a pharmaceutical formulation containing two painkiller pharmaceutical ingredients and three common excipients was prepared. The blend of raw materials has been progressively mixed, and aliquots were taken at several timepoints. The different aliquots were tabletted and the whole tablet surface was analysed by LDIR and Raman imaging. The homogeneity of distribution of each compound was monitored and the results of both imaging techniques were compared. To objectivise the homogeneity of distribution, the distributional homogeneity index (DHI) approach has been chosen (Farkas et al., 2017; Sacré et al., 2014b; Wahl et al., 2017) and converted into homogeneity values. Finally, the blending endpoint and the homogeneity values were compared for both imaging techniques.

2. Material and methods

2.1. Raw materials

Most raw materials (paracetamol (Compap™ PVP3), lactose monohydrate, microcrystalline cellulose (Avicel® PH102) and magnesium stearate) were kindly provided by Galephar M/F (Marche-en-Famenne, Belgium). Acetylsalicylic acid (ASA) was purchased from Fagron Belgium NV (Nazareth, Belgium).

2.2. Sample preparation

A 50 g blend of powders was realized and placed in a mortar. The blend was composed of ASA (20% w/w), paracetamol (20% w/w), lactose (25% w/w), microcrystalline cellulose (30% w/w) and magnesium stearate (5% w/w). The powders were weighed and placed in the mortar in ascending order regarding the final proportion of the component in the blend.

The powders were mixed using a pestle and three aliquots of 200 mg were collected at different places in the mortar at each timepoint. A total of seven timepoints were analysed. Each timepoint corresponded to three turns of pestle at the exception of the first timepoint that was realized before any mixing, directly after the weighing of powders.

The aliquots were placed into a 13 mm evacuable pellet die (Perkin Elmer, USA) and pressed at 1 T for 5 min. The final tablets were 2–3 mm thick and were glued (per timepoints) on a Menzel-Gläser microscope slide (Thermo Scientific, USA) using a cyanoacrylate glue.

The samples were directly analysed by LDIR after being pressed. The Raman experiments were conducted subsequently. The same total sample surface was analysed by both LDIR and Raman imaging. However, the final step size was slightly different (80 μm for LDIR and 100 μm for Raman microscopy) leading to a slightly different map size as explained below. A total of 21 tablets were analysed by both techniques (7 timepoints and 3 aliquots per timepoint)

2.3. Laser direct infrared imaging (LDIR)

The laser direct infrared imaging analyses were performed using an Agilent's 8700 LDIR imaging system (Agilent Technologies, California, USA). The LDIR system relies on a quantum cascade laser (QCL) as the source and a thermoelectrically cooled single point Mercury Cadmium Telluride (MCT) as the detector. The spectra were recorded in the 975–1800 cm^{-1} spectral range with a data point spacing of 0.5 cm^{-1}. The direct reflectance module of the equipment was used in the present study. The 8700 LDIR produces a diffraction-limited spot on the sample. Depending on IR wavelength, it could be between 5.5 and 10 μm. The step size of the imaging experiments is adjustable between 1 and 40 μm.

2.3.1. LDIR method development

The infrared chemical imaging methods were developed in the Clarity software (v.1.3.42).

The reference spectra have been recorded on small pellets (5 mm diameter) of pure material. A total of 100 direct reflectance spectra were recorded and averaged to constitute the reference spectrum of pure material, that was eventually added to the library.

Based on the acquired reference spectra, a single peak ratio analysis has been developed for each chemical present in the blend. This means that, for each chemical, two wavenumbers were chosen (peak and baseline position) based on visual inspection of the spectra at a spectral resolution of 8 cm^{-1}. The final pixel is the ratio of the baseline intensity and the peak intensity.

Eventually, the five single peak methods were concatenated in a multipeak method. The whole tablet (13 mm diameter) was analysed at a step size of 40 μm (−330 × 330 pixels). The total analysis time for the multipeak analysis was 2.5 min. Autofocus of the system was performed on each tablet to avoid low-quality results due to differences in tablet thickness.

The output of the LDIR analysis is an intensity image for each chemical. As these distribution maps cannot be exported for each chemical with the multipeak analysis, a single peak analysis has been performed for each chemical. These distribution maps were saved in the CSV format and transferred to Matlab for subsequent DHI/homogeneity analysis.

2.4. Confocal microscopy Raman imaging

Confocal microscopy Raman imaging experiments were performed on a Labram HR Evolution (Horiba Scientific) equipped with an EMCCD detector (1600 × 200-pixel sensor) (Andor Technology Ltd.), a Leica 50x Fluotar LWD objective and a 785 nm laser (XTRA II single frequency diode laser, Toptica Photonics AG) with a power reduced at 4.5 mW at a sample to avoid destruction by burning.

A circular point mapping experiment was performed with a step size

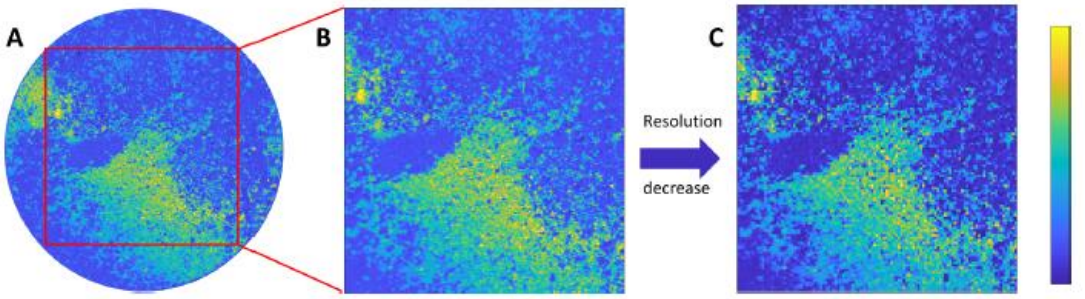

Fig. 1. Extraction of the LDIR distribution map before the computation of the homogeneity. A: Distribution map of ASA obtained by single peak analysis; B: Maximum inscribed square map extracted from A (pixels size 230 × 230); C: Maximum inscribed square map keeping one each two pixels (pixels size 115 x115).

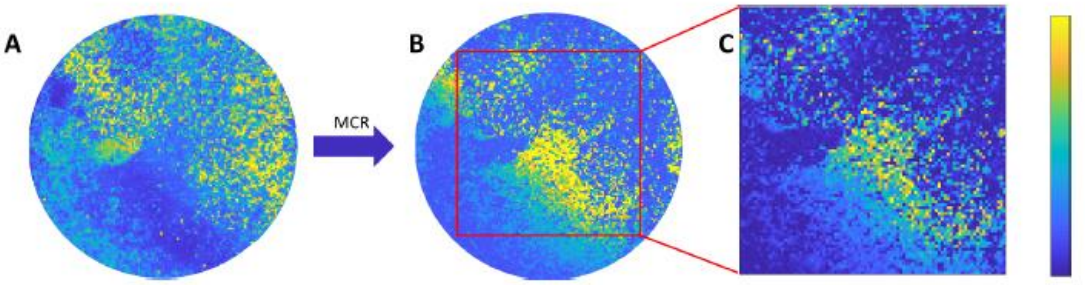

Fig. 2. Extraction of the Raman imaging distribution map before the computation of the homogeneity. A: Average intensity Raman image; B: ASA distribution map resolved by MCR-ALS analysis of A; C: Maximum inscribed square map extracted from B (pixels size 90 × 90).

of 100 µm (maps of 130 × 130 pixels) over the spectral range of 463–1853 cm^{-1}. The spectra were acquired with two accumulations of 0.1 sec exposure time and dispersed with a 300 gr/mm grating. The confocal slit hole was set at 200 µm and the electron-multiplying gain of the detector was set at 50.

2.5. Data analysis

2.5.1. LDIR data analysis

The distribution maps of the different chemicals were used to compute the homogeneity of distribution values. However, DHI cannot be computed on round images. Therefore, the biggest square inscribed in the circular distribution maps was extracted to compute DHI.

The maximum inscribed square map was computed using the formula:

$$S = \sqrt{\frac{d^2}{2}} \quad (1)$$

With S being the length of the square side (in pixels) and d being the diameter of the original distribution map. The obtained square distribution maps were of 233 × 233 pixels size.

The computation time of the DHI has an exponential relationship with the distribution map size. Therefore, to accelerate the DHI analysis, only one of each two pixels were kept reducing the distribution maps to a final size of 115 × 115 pixels. This reduced the DHI analysis time from 1244 min for 233 × 233 maps to ~ 7 min for 115 × 115 maps. The reduced maps were subsequently analyzed with the DHI algorithm. The reduction of the map also artificially increases the step size from 40 to 80 µm. This is done to lower the computation time of the homogeneity and to have more comparable results regarding the Raman experiments. Fig. 1 illustrates the analysis process.

2.5.2. Raman data analysis

The raw Raman data were pre-processed using a Savitzky Golay smoothing (polynomial order: 2, window size: 7) followed by an asymmetric least squares (AsLS) baseline correction (p: 0.001, λ: 10^5) (Eilers, 2003; Savitzky and Golay, 1964).

The pre-processed data were eventually analysed using the MCR-ALS algorithm. Four components were resolved and the distribution map corresponding to ASA was used to compute the homogeneity of distribution value. The resolved loadings exhibited a correlation coefficient above 0.9 with the corresponding reference spectrum.

The maximum inscribed square map was computed using Equation (1) leading to a final square distribution map of 90 × 90 pixels (see Fig. 2).

The pre-processing and MCR analysis were performed using the PLS Toolbox v. 8.9.2 and the MIA Toolbox v. 3.0.9 (Eigenvector Research Inc, USA) running in the Matlab R2018b (The Mathworks, USA) environment.

2.5.3. Distributional homogeneity index (DHI) analysis

The analysed tablet surface was the same for both Raman and LDIR imaging. Only the spatial resolution was different (step size of ~ 80 µm for LDIR and ~ 100 µm for Raman imaging). The square distribution maps were analysed with the DHI algorithm (Sacré et al., 2014b) with a number of simulations $n = 100$.

The DHI analysis comprises several steps among which the distribution map is first sampled by a Continuous-Level Moving Block (CLMB). For each macropixel size, the standard deviation of the macropixel value is computed. Then, the standard deviation is plotted against the macropixel size to obtain the so-called "homogeneity curve". Once the homogeneity curve is obtained for the studied distribution map, the map is randomized and the homogeneity curve of the random map is computed. The DHI value is obtained by the ratio of the area under the homogeneity curve (AUC) of the studied map and the area under the homogeneity curve of the randomized map. Because of the randomization step, many simulations are necessary to compute a mean DHI value assorted with a standard deviation value.

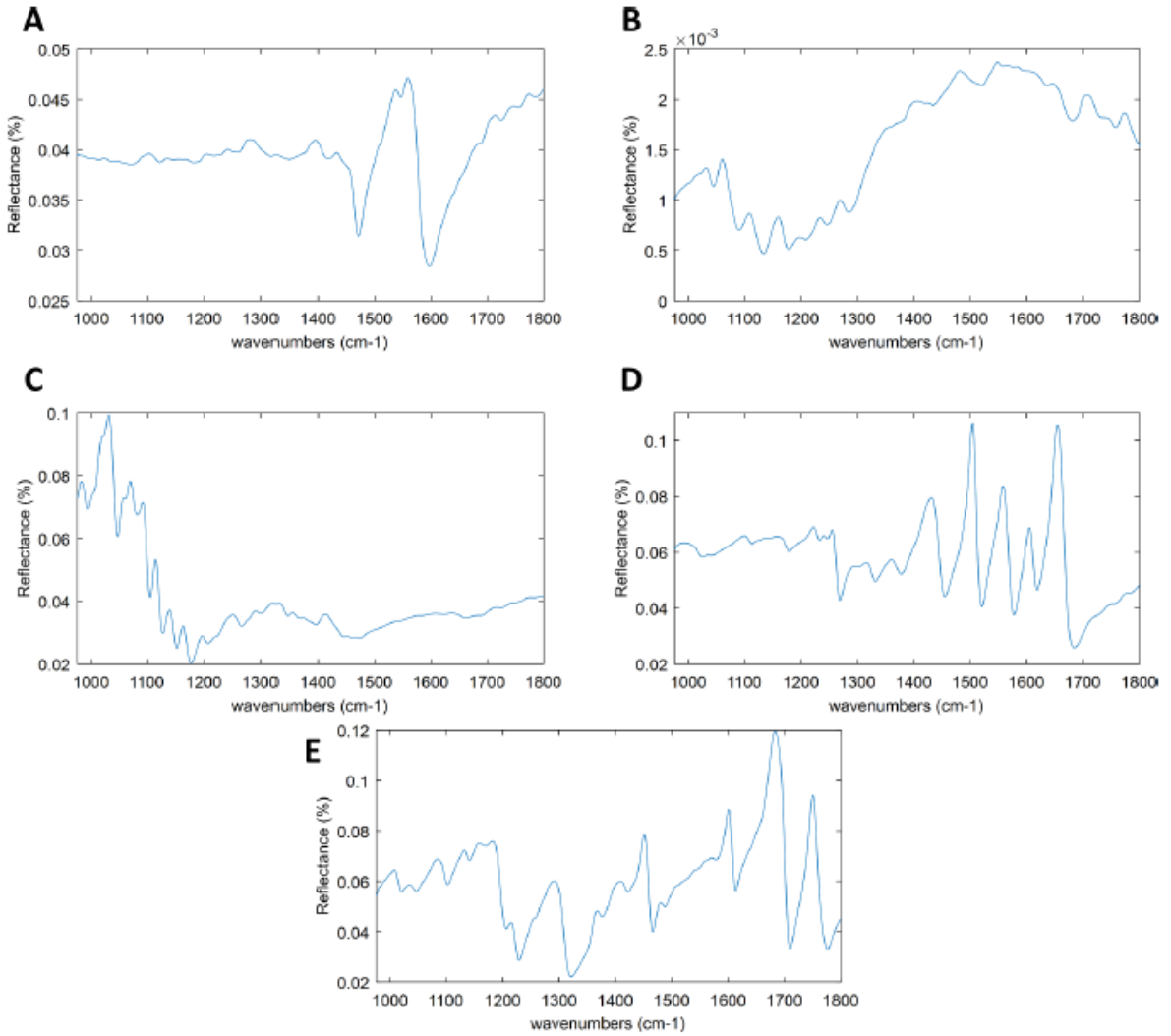

Fig. 3. LDIR spectra of the pure compounds of the pharmaceutical blend: magnesium stearate (A), microcrystalline cellulose (B), lactose (C), paracetamol (D) and acetylsalicylic acid (E).

To ease the interpretation of DHI values, the latter were converted into homogeneity values expressed as a percentage of homogeneity. The conversion into homogeneity values was realized using simulated maps (of pixel size 90 × 90 and 115 × 115 for Raman and LDIR data respectively) with increasing levels of homogeneity between 10 and 90 %. Three maps were simulated per level of homogeneity and their average DHI value was computed with a number of simulations set to $n = 100$. The ordinary least squares (OLS) regression coefficients were computed, and the coefficients of determination were above 0.99 for each map size. Figure S1 shows the OLS regression lines for the simulated maps.

2.5.4. Hardware description:

All computations were performed on a portable computer with the following specifications:

- Windows 10 Professional 64-bit
- Intel i5-6440HQ CPU 2.6 GHz Processor
- 1 TB SSD
- 16 GB RAM
- AMD Radeon R7 M370 graphics card

3. Results and discussion

3.1. LDIR analysis

Once the imaging method is developed, analysing samples with the LDIR system is very fast and straightforward. However, the development of the method may be a difficult and tedious task. Indeed, the analyst can choose different analyses options (single peak ratio with one or two baseline points or classification analysis) but all these options are based on the selection of a single wavelength (plus one or two baseline points) to represent the compound of interest. Therefore, the analysis of complex samples (more than 2 or 3 compounds) is dependent on the spectral signature and the possible overlapping of the individual spectral signals. This task is also complicated by the shape of direct reflectance IR signals exhibiting larger and less resolved peaks compared to Raman or transmission IR spectra because of the superposition of diffuse and specular reflected light (Fringeli, 1999). Fig. 3 shows the LDIR spectra acquired on pure compound pellets.

In the present work, it was decided to build single peak ratio (with one baseline point) analyses for each compound. These single peak ratio methods were eventually combined as a "multipeak analysis". This choice was guided by the will of obtaining a distribution map of intensity values for each compound that is more realistic. Indeed,

Table 1
Wavenumbers chosen as peak and baseline for each compound of the blend to build the LDIR method.

Compound name	Baseline wavenumber (cm^{-1})	Peak wavenumber (cm^{-1})
Paracetamol	1454.7	1504.0
Acetylsalicylic acid (ASA)	1709.6	1682.3
Lactose	1059.8	1029.8
Microcrystalline cellulose	1090.0	1060.0
Magnesium stearate	1597.5	1558.5

depending on the spatial resolution (step size), the spot size of the system and the particle size of the powders analysed, it is very likely that more than one compound is present in a single pixel, the measured spectrum being the weighted sum of the individual compound spectra. The classification method automatically assigns a single compound for each analysed pixel based on a predefined classification rule. The classification method works by comparing ratios of peaks, and the individual pixel gradient or brightness represents the normalized value of the ratio for the classified component. This provides some texture to the image to present a more 'natural' feel and gives the user an idea of where the ratio is strongest. Table 1 listed the wavelength values for each peak and baseline used in the single peak ratio analyses. These peak and baseline wavelengths are presented on the corresponding spectra in figures S2 – S6.

Once the method is developed, the analysis time is very short since it took 2.5 min between the start of the imaging experiment and the display of the results (31 sec for each single peak analysis). The result of the analysis is the distribution map of the five components in the tablet. Fig. 4 shows the distribution maps obtained with the LDIR system for each of the 21 tablets analysed (three tablets per timepoint and seven timepoints analysed). Shadows may be observed on some tablets (especially on tablets 1 and 3 of timepoint 4). These shadows are linked to small differences between the tablet surface and the light (out-of-focus) coming from the fact that some tablets were not lying perfectly flat on the microscope slide. This information is available in the single and multipeaks analysis because the analysis result is an intensity value for each compound in each pixel. This would have been avoided if a classification analysis model is used since the output is a categorical response. Another option, that is now commonly available on most Raman imaging systems is a topographic mapping of the sample before the Raman imaging analysis allowing an automatic adjustment of the collection optics distance to the sample keeping it constant throughout the analysis.

A straightforward visual inspection of the images shows that the different compounds are more and more homogeneously distributed as the mixing time increases. Magnesium stearate is only visible in the first timepoints T_0 and T_1. Lactose and microcrystalline cellulose are not observed in each tablet at the first timepoints because of the inhomogeneous distribution of the powders at the beginning of the mixing process. The different components are progressively blended, and their distribution becomes more and more homogeneous. However, it is difficult to visually assess the endpoint of the blending process. Therefore, a DHI analysis of the distribution maps has been realized to objectify the homogeneity of distribution (Sacré et al., 2014b).

Another question that arose from the development of the LDIR model was the confidence that we may have in the distribution maps obtained using a univariate analysis of pharmaceutical tablets. Therefore, we performed the same analyses using a Raman confocal microscope to compare the results obtained with both systems.

3.2. Raman imaging analysis

The tablets previously analysed by LDIR were eventually analysed with a Raman confocal microscope. Compared to the LDIR system, the setup of the Raman imaging analyses is very fast since the data analysis is realised afterwards. To avoid any burning of the tablet, the laser power was reduced to 4.5 mW on sample. This leads to an increase in the

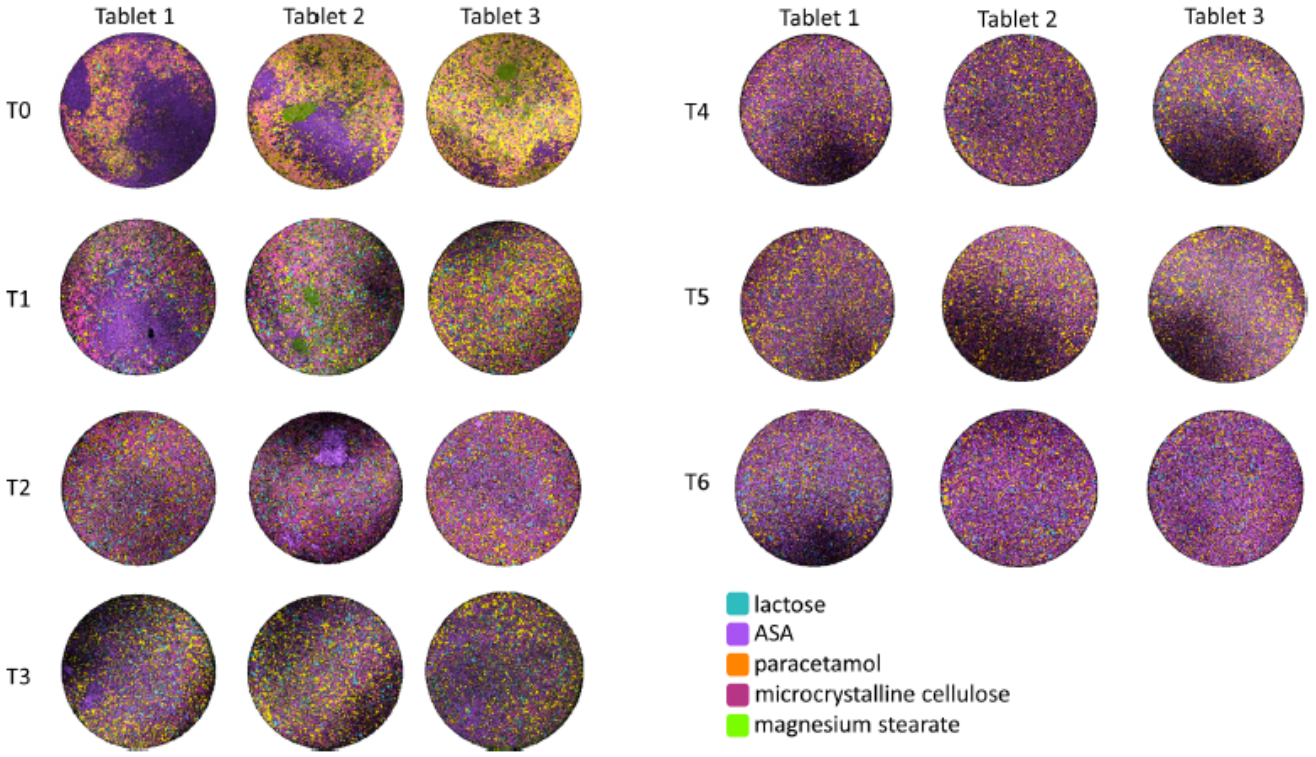

Fig. 4. Results of the LDIR multipeaks imaging analysis of the tablets. Three tablets are presented for each timepoint (T_0 – T_6).

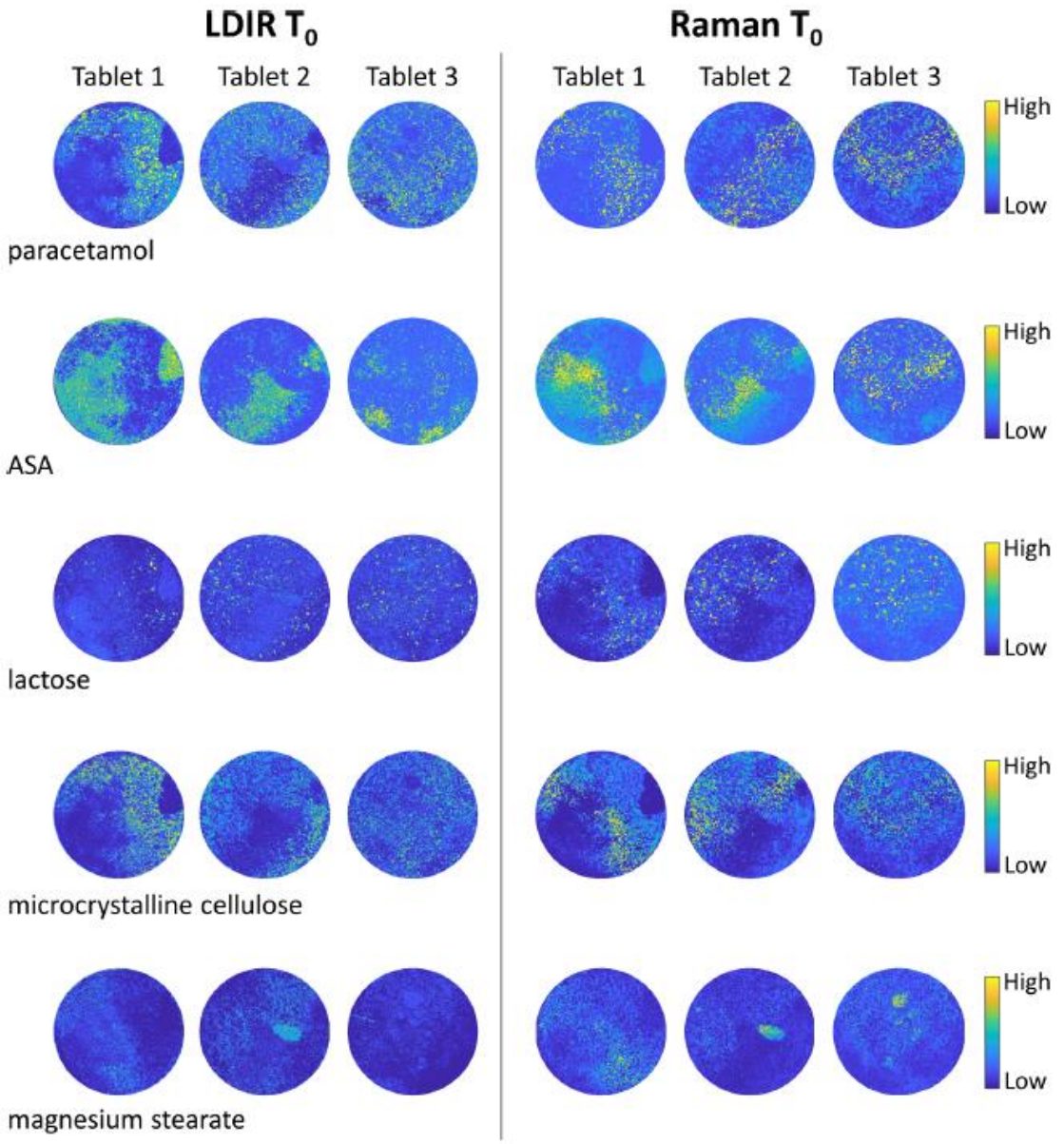

Fig. 5. Comparison of the distribution maps obtained for each compound with LDIR (single peak ratio imaging analysis) and Raman imaging (after MCR-ALS analysis) of the tablets per timepoint.

acquisition time to get a satisfying signal over noise ratio, increasing the overall analysis time to 183 min per tablet. To this time, one must still add the data processing time (−1h) leading to a total analysis time of about 4 h (−240 min) compared to the 7.5 min of LDIR (−0.5 min for the LDIR analysis and − 7 min for the DHI computation).

MCR-ALS has been used with a various number of components to extract. Three components were selected to obtain the distribution maps of the active pharmaceutical ingredients (ASA, paracetamol). To resolve the minor compounds (microcrystalline cellulose, lactose, and magnesium stearate), an MCR of 9 components was necessary. The quality of the resolution was assessed by computing the correlation coefficient between the resolved spectrum and the reference spectrum. A correlation coefficient above 0.9 was found satisfactory since a correlation coefficient of 1 describes a perfect match. Although present in each tablet, magnesium stearate was not resolved in several tablets of timepoints T_4, T_5 and T_6 because of the intimate mixture of the blend and its low concentration.

The resolved distribution maps of the Raman experiments were visually compared to the LDIR individual distribution maps (Fig. 5). This comparison was performed at T_0 since it is the timepoint with the highest inhomogeneity showing clear patterns. It was found that the LDIR distribution maps were reasonably following the Raman imaging results confirming that the developed method was able to monitor the distribution of each compound. Nevertheless, some discrepancies may be observed between the two methods (e.g. ASA in tablet 1). These discrepancies may come from the fact that the imaging setups are different along with different collection optics and the underlying physical mechanisms are also different (absorption vs scattering). In addition to these technical differences, the data analysis is completely different (peak ratio for LDIR vs multivariate data analysis for Raman).

Therefore, these differences will also be present in the DHI analysis.

3.3. Homogeneity of tablets

The homogeneity results of individual tablets are presented in supplementary table ST1.

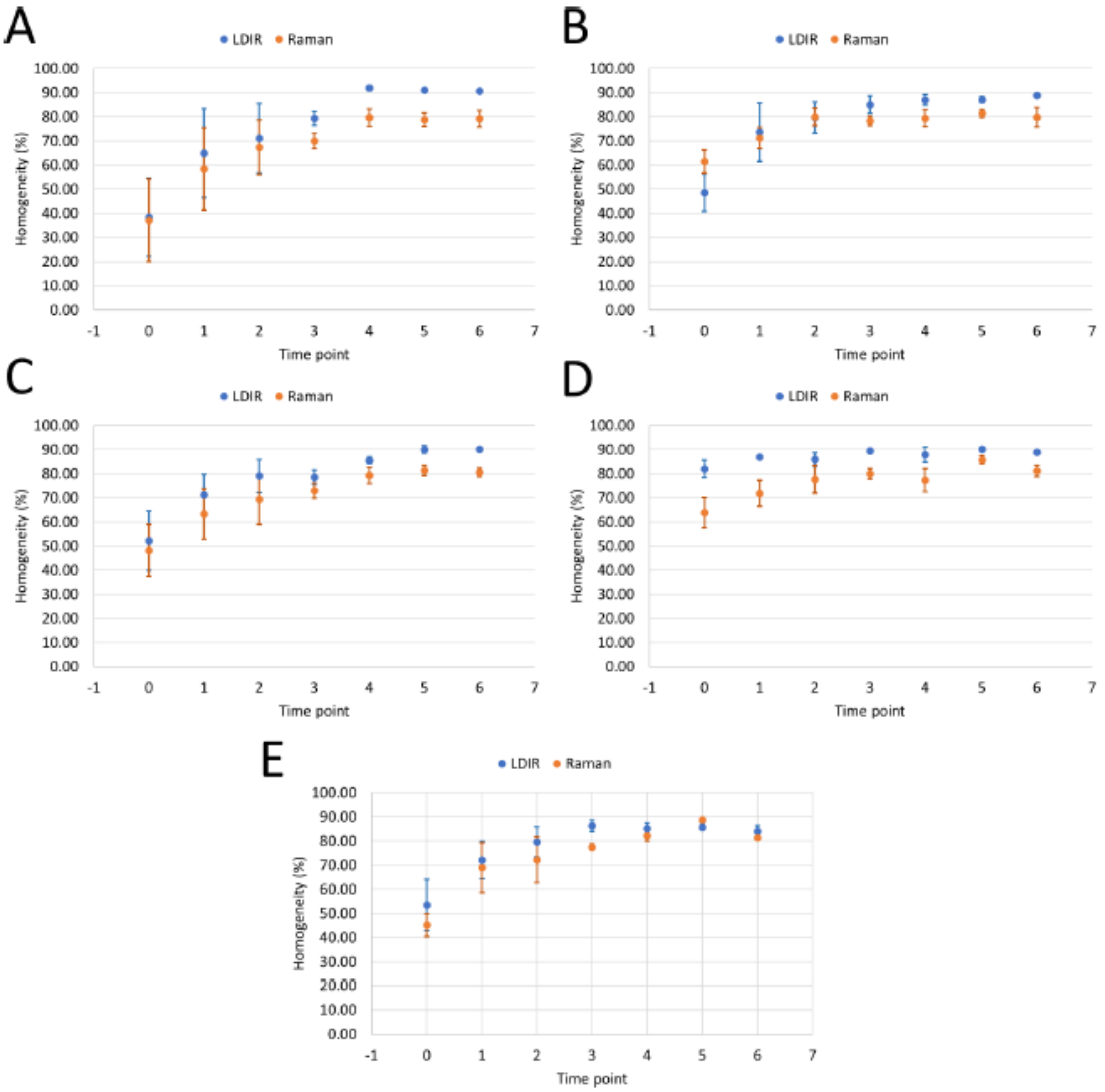

Fig. 6. Average homogeneity computed for each timepoint for LDIR and Raman distribution maps of A) ASA, B) paracetamol, C) microcrystalline cellulose, D) lactose, E) magnesium stearate.

Fig. 6 summarizes the results of the homogeneity computations at each timepoint for both Raman and LDIR experiments. The error bars were computed as the standard error (SE) using the following formula:

$$SE = \frac{\sigma_i}{\sqrt{n_i}}$$

where σ_i is the standard deviation of the computed homogeneity at timepoint i and n_i the number of replicates at timepoint i.

For the first three timepoints, there is a low average homogeneity of distribution among single tablets, but, above all, a high inhomogeneity between tablets. This was expected from the visual inspection of the LDIR multipeaks images (see Fig. 4). However, the visual appreciation is now objectivised through the computation of homogeneity. The blend appears to be homogeneous at timepoint 4 with an almost constant average homogeneity of −90% for LDIR and −80% for Raman experiments, respectively. The standard errors also become very small at this stage. The homogeneity computed from the Raman data appears to be slightly lower than the one from the LDIR experiments. This might come from the "shadow effect" observed due to the slightly non-flat surface of the tablets but also from the lower spatial resolution of the Raman experiments. Nevertheless, the endpoint of the blending is easily found and is coherent with the LDIR experiments.

4. Conclusion

Distributional homogeneity of compounds in a pharmaceutical formulation is a key parameter to be monitored since it may affect the efficacy and safety of the formulation.

Usually, this parameter is estimated by dosing the active pharmaceutical ingredient in several tablets and computing the well-known acceptance value following the pharmacopoeia rules. However, the acceptance value only informs about the inter-tablet homogeneity of the active ingredient. No information is provided regarding the intra-tablet API distribution or the distribution of excipients. This information is very important when developing a new formulation as it may influence the process parameters or properties of entry raw materials (particle size, etc.). A unique tool that provides this information is the chemical imaging technology allowing the pharmaceutical analyst to visualise the distribution of the main compounds at the surface of the tablet.

Nevertheless, when the method is successfully developed, results are obtained in a handful of seconds. Therefore, it is possible to analyse several tablets and get a unique insight into "intra-tablet" homogeneity

and "inter-tablet" homogeneity. The latter is difficult to obtain with Raman imaging because of the long acquisition times making the analysis of several tablets difficult.

Based on the presented results, LDIR seems to be a very promising tool for the analysis of the spatial distribution of chemicals in a pharmaceutical formulation. This might be very useful in routine analysis of pharmaceutical blends as quality control of the produced tablets but most of all during the development of new formulations. Indeed, when a new pharmaceutical formulation is developed, the choice of excipients, raw materials particle size or even the process parameters may greatly influence the homogeneity of distribution of chemicals (Chavez et al., 2015) and, consequently, the efficacy and safety of the formulation. Usually, this homogeneity is assessed through liquid chromatography analysis of the content of API of the different formulations. However, this conventional workflow is tedious, needs a lot of organic solvents and lab work and it does neither provide any insight into the "intra-tablet" homogeneity nor information regarding excipients. These two last points may only be assessed using hyperspectral imaging such as NIR or Raman imaging. However, NIR imaging suffers from a low spatial resolution and high penetration depth (hundreds of micrometres compared to a maximum of 15–20 μm for both LDIR and Raman) leading to blurred results (Carruthers et al., 2021). Raman imaging is relatively slow making the analysis of several tablets at a high spatial resolution hardly compatible with the industrial needs and it suffers from auto-fluorescence of chemicals (such as cellulose derivatives that are often used as excipients).

LDIR, on its side, suffers from the difficulty to setup methods when the sample is complex and does not allow a straightforward selection of unique peaks attributable to specific compounds since it images the intensity of IR reflectance at unique wavenumbers.

On the one hand, Raman spectroscopy imaging systems collecting data over a spectral range at each measured pixel have a high spectral specificity leading to a higher versatility allowing the analysis of unknown formulations or the analysis of spectra modified by the manufacturing process (chemical interactions, solid-state modifications etc.). However, this is at the cost of long analysis times. On the other hand, once the method is developed and validated or "confirmed" by another technique, LDIR enables the analysis of tablets at high speed (−31 sec per compound per tablet) which is valuable and constitutes a plus-value during the development of a new pharmaceutical formulation for routine quality control analyses.

CRediT authorship contribution statement

P.-Y. Sacré: Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing. **M. Alaoui Mansouri**: Investigation. **C. De Bleye**: Conceptualization, Writing – review & editing. **L. Coïc**: Conceptualization, Writing – review & editing. **Ph. Hubert**: Project administration, Supervision. **E. Ziemons**: Conceptualization, Methodology, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Coic Laureen reports financial support was provided by European Commission. Pierre-Yves Sacre reports financial support was provided by Fund for Scientific Research. Laureen Coic reports financial support was provided by Walloon Region. Pierre-Yves Sacre reports equipment, drugs, or supplies was provided by Agilent Technologies Inc.].

Acknowledgements

Research grants from the Walloon Region of Belgium and the EU Commission (project FEDER-PHARE) to L. Coïc are gratefully acknowledged.

FNRS and FWO through the Chimic EOS project are also acknowledged (P-Y Sacré).

Agilent is acknowledged for the lend of the LDIR system used in the present study.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijpharm.2021.121373.

References

Belu, A.M., Davies, M.C., Newton, J.M., Patel, N., 2000. TOF-SIMS characterization and imaging of controlled-release drug delivery systems. Anal. Chem. 72, 5625–5638. https://doi.org/10.1021/ac000450+.

Bhargava, R., 2012. Infrared spectroscopic imaging: the next generation. Appl. Spectrosc. 66 (10), 1091–1120. https://doi.org/10.1366/12-06801.

Bøtker, J., Wu, J.X., Rantanen, J., 2020. Hyperspectral imaging as a part of pharmaceutical product design, in: Data Handling in Science and Technology. Elsevier Ltd, pp. 567–581. https://doi.org/10.1016/B978-0-444-63977-6.00022-5.

Carruthers, H., Clark, D., Clarke, F., Faulds, K., Graham, D., 2021. Comparison of Raman and near-infrared chemical mapping for the analysis of pharmaceutical tablets. Appl. Spectrosc. 75 (2), 178–188. https://doi.org/10.1177/0003702820952440.

Chavez, P.-F., Lebrun, P., Sacré, P.-Y., De Bleye, C., Netchacovitch, L., Cuypers, S., Mantanus, J., Motte, H., Schubert, M., Evrard, B., Hubert, P., Ziemons, E., 2015. Optimization of a pharmaceutical tablet formulation based on a design space approach and using vibrational spectroscopy as PAT tool. Int. J. Pharm. 486 (1-2), 13–20. https://doi.org/10.1016/j.ijpharm.2015.03.025.

Coic, L., Sacré, P.-Y., Dispas, A., Sakira, A.K., Fillet, M., Marini, R.D., Hubert, P., Ziemons, E., 2019. Comparison of hyperspectral imaging techniques for the elucidation of falsified medicines composition. Talanta 198, 457–463. https://doi.org/10.1016/j.talanta.2019.02.032.

da Costa Filho, P.A., Cobuccio, L., Mainali, D., Rault, M., Cavin, C., 2020. Rapid analysis of food raw materials adulteration using laser direct infrared spectroscopy and imaging. Food Control 113, 107114. https://doi.org/10.1016/j.foodcont.2020.107114.

Earnshaw, C.J., Carolan, V.A., Richards, D.S., Clench, M.R., 2010. Direct analysis of pharmaceutical tablet formulations using matrix-assisted laser desorption/ionisation mass spectrometry imaging. Rapid Commun. Mass Spectrom. 24 (11), 1665–1672. https://doi.org/10.1002/rcm.4525.

Eilers, P.H.C., 2003. A Perfect Smoother. Anal. Chem. 75, 3631–3636. https://doi.org/10.1021/ac034173t.

Farkas, A., Nagy, B., Marosi, G., 2017. Quantitative evaluation of drug distribution in tablets of various structures via Raman mapping. Period. Polytech. Chem. Eng. 1–7.

Fringeli, U.P., 1999. In: Encyclopedia of Spectroscopy and Spectrometry. Elsevier, pp. 94–109. https://doi.org/10.1016/B978-0-12-374413-5.00104-4.

Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Bell, S.E.J., 2008. Recent applications of Chemical Imaging to pharmaceutical process monitoring and quality control. Eur. J. Pharm. Biopharm. 69 (1), 10–22. https://doi.org/10.1016/j.ejpb.2007.10.013.

Klukkert, M., Wu, J.X., Rantanen, J., Carstensen, J.M., Rades, T., Leopold, C.S., 2016. Multispectral UV imaging for fast and non-destructive quality control of chemical and physical tablet attributes. Eur. J. Pharm. Sci. 90, 85–95. https://doi.org/10.1016/j.ejps.2015.12.004.

Novakovic, D., Saarinen, J., Rojalin, T., Antikainen, O., Fraser-Miller, S.J., Laaksonen, T., Peltonen, L., Isomäki, A., Strachan, C.J., 2017. Multimodal nonlinear optical imaging for sensitive detection of multiple pharmaceutical solid-state forms and surface transformations. Anal. Chem. 89 (21), 11460–11467. https://doi.org/10.1021/acs.analchem.7b02639.

Prats-Montalbán, J.M., Jerez-Rozo, J.I., Romañach, R.J., Ferrer, A., 2012. MIA and NIR chemical Imaging for pharmaceutical product characterization. Chemom. Intell. Lab. Syst. 117, 240–249. https://doi.org/10.1016/j.chemolab.2012.04.002.

Sacré, P.-Y., De Bleye, C., Chavez, P.-F., Netchacovitch, L., Hubert, P., Ziemons, E., 2014a. Data processing of vibrational chemical imaging for pharmaceutical applications. J. Pharm. Biomed. Anal. 101, 123–140. https://doi.org/10.1016/j.jpba.2014.04.012.

Sacré, P.-Y., Lebrun, P., Chavez, P.-F., De Bleye, C., Netchacovitch, L., Rozet, E., Klinkenberg, R., Streel, B., Hubert, P., Ziemons, E., 2014b. A new criterion to assess distributional homogeneity in hyperspectral images of solid pharmaceutical dosage forms. Anal. Chim. Acta 818, 7–14. https://doi.org/10.1016/j.aca.2014.02.014.

Sacré, P.-Y., Netchacovitch, L., De Bleye, C., Chavez, P.-F., Servais, C., Klinkenberg, R., Streel, B., Hubert, P., Ziemons, E., 2015. Thorough characterization of a self-emulsifying drug delivery system with Raman hyperspectral imaging: a case study. Int. J. Pharm. 484 (1-2), 85–94. https://doi.org/10.1016/j.ijpharm.2015.02.052.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639. https://doi.org/10.1021/ac60214a047.

Scircle, A., Cizdziel, J.V., Tisinger, L., Anumol, T., Robey, D., 2020. Occurrence of microplastic pollution at oyster reefs and other coastal sites in the Mississippi Sound, USA: impacts of freshwater inflows from flooding. Toxics 8, 35. https://doi.org/10.3390/toxics8020035.

Scoutaris, N., Vithani, K., Slipper, I., Chowdhry, B., Douroumis, D., 2014. SEM/EDX and confocal Raman microscopy as complementary tools for the characterization of

pharmaceutical tablets. Int. J. Pharm. 470 (1-2), 88–98. https://doi.org/10.1016/j.ijpharm.2014.05.007.

Stewart, S., Priore, R.J., Nelson, M.P., Treado, P.J., 2012. Raman imaging. Annu Rev Anal Chem (Palo Alto Calif) 5, 337–360. https://doi.org/10.1146/annurev-anchem-062011-143152.

Wahl, P.R., Pucher, I., Scheibelhofer, O., Kerschhaggl, M., Sacher, S., Khinast, J.G., 2017. Continuous monitoring of API content, API distribution and crushing strength after tableting via near-infrared chemical imaging. Int. J. Pharm. 518 (1-2), 130–137. https://doi.org/10.1016/j.ijpharm.2016.12.003.

Journal of Pharmaceutical and Biomedical Analysis 196 (2021) 113922

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

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

Classification of polymorphic forms of fluconazole in pharmaceuticals by FT-IR and FT-NIR spectroscopy

Mohammed Alaoui Mansouri[a,b,*], Eric Ziemons[a], Pierre-Yves Sacré[a], Mourad Kharbach[b,c], Issam Barra[b,d], Yahia Cherrah[b], Philippe Hubert[a], Roland Djang'eing'a Marini[a], Abdelaziz Bouklouze[b]

[a] University of Liege (ULiege), CIRM, Vibra-Santé HUB, Laboratory of Pharmaceutical Analytical Chemistry, CHU, B36, B-4000, Liege, Belgium
[b] Bio-Pharmaceutical and Toxicological Analysis Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V, Rabat, Morocco
[c] Department of Analytical Chemistry, Applied Chemometrics and Molecular Modelling, CePhaR, Vrije Universiteit Brussel (VUB), Laarbeeklaan 103, B-1090, Brussels, Belgium
[d] Center of Excellence in Soil and Fertilizer Research in Africa, Mohammed VI Polytechnic University, Benguerir, Morocco

ARTICLE INFO

Article history:
Received 21 October 2020
Received in revised form 18 January 2021
Accepted 19 January 2021
Available online 26 January 2021

Keywords:
Polymorphic forms
Fluconazole
FT-IR spectroscopy
FT-NIR spectroscopy
PLS-DA
Classification

ABSTRACT

The main goal of this work was to test the ability of vibrational spectroscopy techniques to differentiate between different polymorphic forms of fluconazole in pharmaceutical products. These are mostly manufactured with fluconazole as polymorphic form II and form III. These crystalline forms may undergo polymorphic transition during the manufacturing process or storage conditions. Therefore, it is important to have a method to monitor these changes to ensure the stability and efficacy of the drug.

Each of FT-IR or FT-NIR spectra were associated to partial least squares-discriminant analysis (PLS-DA) for building classification models to distinguish between form II, form III and monohydrate form. The results has shown that combining either FT-IR or FT-NIR to PLS-DA has a high efficiency to classify various fluconazole polymorphs, with a high sensitivity and specificity. Finally, the selectivity of the PLS-DA models was tested by analyzing separately each of three following samples by FT-IR and FT-NIR: lactose monohydrate, which is an excipient mostly used for manufacturing fluconazole pharmaceutical products, itraconazole and miconazole. These two last compounds mimic potential contaminants and belong to the same class as fluconazole. Based on the plots of Hotelling's T^2 vs Q residuals, pure compounds of miconazole and itraconazole, that were analyzed separately, were significantly considered outliers and rejected. Furthermore, binary mixtures consist of fluconazole form-II and monohydrate form with different ratios were used to test the suitability of each technique FT-IR and FT-NIR with PLS-DA to detect minimum contaminant or polymorphic conversion from a polymorphic form to another using also the plots of Hotelling's T^2 vs Q residuals.

1. Introduction

Polymorphism is a characteristic where a drug substance can present one or more crystalline form due to different molecule arrangements; thus can include different solid varieties of crystalline forms. Sometimes, the crystal form is known by solvate when it has amount of solvent and it is known as hydrate if the solvent is water [1–4].

Each polymorphic form has different physico-chemical properties [5]. Variations at the level of physico-chemical properties could have an impact on dissolution rate and bioavailability, hence, the therapeutic effect of the drug substance might be influenced [6]. The manufacturing process and the storage conditions are considered as the main factors that have an impact on polymorphic transformation [7,8]. Thus, looking for a reliable analytical technique to control and monitor polymorphs of drug substance in drug products is mandatory to ensure the quality of pharmaceutical products.

The analysis of polymorphic forms of drug substance has been carried out by both destructive and non-destructive techniques. Destructive techniques are represented in differential scanning

* Corresponding author at: University of Liege (ULiege), CIRM, Vibra-Santé HUB, Laboratory of Pharmaceutical Analytical Chemistry, CHU, B36, B-4000, Liege, Belgium.
E-mail address: med.alaoui@doct.uliege.be (M. Alaoui Mansouri).

https://doi.org/10.1016/j.jpba.2021.113922
0731-7085/

calorimetry (DSC). The main challenge of this technique is the interconversion of polymorphic forms of the drug substance that could be occurred during the analysis [9]. Non-destructive techniques are summarized in X-ray powder diffraction (XRPD) and vibrational spectroscopic techniques [10,11]. The principal advantage of these techniques is that they often do not need any sample preparation. Hence, they are fast in analyzing and acquiring results [12]. Nevertheless, these last cited techniques have some limitations. The main limitation of XRPD is that the morphology of the particle may impact the accuracy of quantitative analysis using XRPD [13]. The main challenge of vibrational spectroscopy is identifying and discriminating between polymorphic forms directly especially in the presence of the matrix that may hamper the identification of fingerprints related to the identity of the polymorphism [14,15]. Thus, associating spectroscopic techniques with chemometric tools is important to uncover more details about polymorphism.

Chemometric tools have already proven their usefulness to discriminate and quantify polymorphic forms such as PCA and PLS and reduce systematic variations by using preprocessing techniques such as standard normal variate (SNV) or multiplicative scatter corrections (MSC) [16]. For example, polymorphic forms were quantified accurately by coupling either of FT-NIR Raman or FT-IR to Partial Least squares regression (PLS-R) [17]. Another example, is applying PCA to FT-IR data that allows detecting which of the four polymorphic forms of cimetidine was present in a pharmaceutical product [18]. In addition to PCA and PLS, another chemometric tool called multivariate curve resolution (MCR) can detect how many polymorphic forms exist in a mixture and identify their pure spectra. This has been combined to FT-IR for following the polymorphic interconversions of cimetidine [19]. Besides that, PLS-DA has been used successfully in food and other applications [20,21]. These applications of chemometric techniques nicely illustrate their efficiency to extract relevant information from the raw spectral data.

Fluconazole, 2-(2,4-difluorophenyl)-1,3-bis(1-H-1,2,4-triazol-1-yl)propan-2-ol, is an antifungal triazole. Fluconazole is used to treat superficial Candida infections. It is used for acute therapy of disseminated Candida, for systemic therapy of blast mycosis and histoplasmosis, for dermatophytic fungal infections, and for prophylaxis in neutropenic patients [22–24]. According to recent studies [25,26], fluconazole displays three main polymorphic forms: form I, II and III as well as a monohydrate form. The most stable polymorph is form III. This form is a convert form from the metastable form II, that may convert to the monohydrate form during the storage or compression under specific conditions of humidity and temperature. At the moment, the most marketed forms by Moroccan pharmaceutical industries are form II and form III while polymorphic form I is considered as unstable based on the recent study [27].

The main objective of this work was to evaluate the qualitative abilities of each vibrational techniques of FT-IR and FT-NIR to classify polymorphic forms of fluconazole in pharmaceutical products as well as investigate the suitability of both vibrational techniques to detect whether there exist any polymorphic conversion or falsified pharmaceutical product of fluconazole

2. Material and methods

2.1. Instrumentation

The FT-IR spectra were acquired in the reflectance mode in the spectral region of 4000−650 cm-1, with an average of 32 scans at resolution of 4 cm-1, using a Frontier FT-IR spectrometer (Perkin Elmer, Waltham, USA) equipped with a diamond crystal ATR device. For each measurement, a fraction of sample is placed onto the diamond surface. the diamond surface was cleaned with acetone and dried between each analysis. The cleaning of the diamond surface was checked spectrally.

The NIR spectra were obtained using FT-NIR spectrophotometer MPA Multi-Purpose FT-NIR Analyzer (Bruker Optics, Ettlingen, Germany) in diffuse reflectance mode in the spectral region of 12500−4000 cm-1 at resolution of 8 cm-1. The average of 32 scans was acquired for each sample by placing the optical fiber probe on the bottom of the glass vial that contains the sample.

2.2. Sample preparation

The main samples that were acquired to build a dataset were:

- Fluconazole polymorphic form II (TCI- Chemicals, Belgium), fluconazole polymorphic form III (Sigma-Aldrich, Belgium), miconazole (Sigma-Aldrich, Belgium), itraconazole (Sigma-Aldrich, Belgium) and lactose monohydrate (Sigma-Aldrich, Belgium). Fluconazole monohydrate form was obtained based on a reported recrystallization technique [28]. This recrystallization method was carried out by dissolving the fluconazole form- II in milli-Q water under constant stirring at 40 °C. The saturated solution was filtered to remove all nuclei, and the filtered solution was cooled in a refrigerator at 5 °C. The resulting crystals were rapidly surface dried only, then the polymorphic form of fluconazole monohydrate was checked with FT-IR and the obtained spectrum was compared to the spectrum of previous studies [26]. The pure polymorphic forms were gently mixed using pestle and mortar and transferred into vials for analysis.
- Commercial pharmaceutical products (50 mg of fluconazole) were acquired at a local drugstore. The average weight of each capsule content was 150 mg. These commercial products consist of two groups:
- The first group is composed of 14 capsules containing polymorphic form II of fluconazole.
- The second group is composed of 17 capsules containing polymorphic form III of fluconazole
- Eleven samples of binary mixtures were prepared. These consisted of 50 mg of fluconazole containing: 1, 2.5, 5, 10, 20, 50 80, 90, 97.5 and 99 % (w/w) of polymorphic form II of fluconazole with the remaining of mass balance of monohydrate form of fluconazole. These binary mixtures were mixed gently using pestle and mortar in order to ensure their homogeneity prior to the transfer to vials for FT-IR and FT-NIR analysis.

2.3. Multivariate data analysis

2.3.1. Datasets

PLS-DA models were developed based on the partial least square algorithm – discriminant analysis using the PLS Toolbox V8.2.1 (Eigenvector research INC, USA) running on Matlab (R2018b) (The Mathworks, USA).

The dataset is composed of three main parts: training, test and suitability set. The training set was used to develop PLS-DA models for three polymorphic forms of fluconazole while test and suitability set were used to test developed models.

A data set was splitted using Kennard-Stone algorithm to provide uniform coverage into Training and test set consisting of:

- samples spectra obtained from mixing 50 mg of each pure polymorphic form of fluconazole with 100 mg of lactose monohydrate in order to be closed from the drug formulation.
- samples from capsules of pharmaceutical products.

A Suitability set was used to test how the built PLS-DA models for FT-NIR and FT-IR can detect minor polymorphic transformations

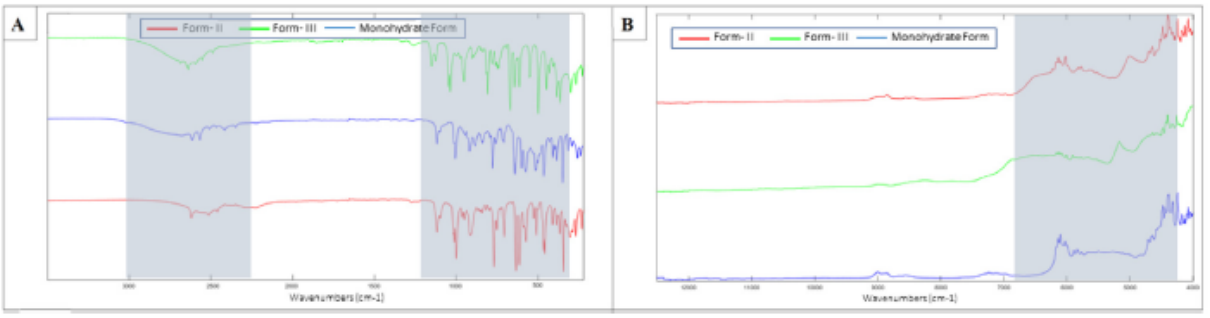

Fig. 1. Spectra of Fluconazole. Pure polymorphic forms of (A): FT-IR; (B): FT-NIR.

and prove its suitability for samples that do not contain fluconazole. Here, suitability set is composed of:

- Eleven different spectra of binary mixture of polymorphic form-II and monohydrate form from 1% to 99 % (w/w)
- Three spectra of itraconazole, miconazole and monohydrate lactose were included.

2.3.2. FT-IR and FT-NIR preprocessing

Before developing the PLS-DA models, the spectra need to be preprocessed to improve discrimination results. We used first derivative as preprocessing method followed by mean centering for FT-IR spectra and we applied standard normal variate (SNV) with mean centering for NIR spectra.

Partial least squares - discriminant analysis (PLS-DA) model were developed based on the training sets and also built only on a range of wavelengths containing the significant variables that are considered as fingerprints of each polymorphic form.

PLS-DA is a supervised classification method that depends on X (FT-NIR or FT-IR spectra of fluconazole polymorphic forms samples) and Y (classes of different polymorphic form) matrices to develop a discriminative calibration model. It is based on reducing the data to scores and loading matrix which permit looking for the most optimal latent variables by maximizing the covariance between X and Y. In case of PLS-DA, Y contains a qualitative variable that identifies different classes of polymorphic forms of fluconazole. Therefore value "1" is attributed to the target class that has to be discriminated from the other alternative classes that have the value of "0".

Latent variables are chosen based on the lowest error of an appropriate cross validation. This built model is then evaluated by Root Mean Square Error of Calibration (RMSEC) and Cross-Validation (RMSECV) from Training set and Root Mean Square Error of Prediction (RMSEP) from the test set used to validate the developed model [29]. To guarantee the reliability of the model in the classification of different classes, a confusion matrix of classification parameters is used to evaluate the performance of PLS-DA models with sensitivity and specificity. The sensitivity is defined as the proportion of the samples of the class that are correctly attributed to the target class (true positives) while the specificity is known as the proportion of the samples that do not belong to the target class to be classified to the alternative class (true negatives) [30].

There are two PLS-DA approaches: the first approach is based on building PLS-DA model for each target class while the other classes are considered alternatives (so called one vs rest classification). This approach would lead us to build three PLS-DA models because of three polymorphic forms of fluconazole. The second approach is PLS-2 regression and it is known as one vs one multiclass PLS-DA model, and it leans on building one model for all calibrated classes.

2.3.3. Suitability set

Suitability sets were also used to test the developed PLS-DA models. However, in this part we relied on Hotelling's T^2 and Q residual parameters. These parameters helped to elucidate the behavior of the calibrated PLS-DA models towards the new integrated samples. These two parameters (Hotelling's T^2 and Q residual) were used in a plot with a threshold of 0.95 to check the homogeneity of the dataset and detect if there exist any outliers. The Q statistic is used to check how well each sample conforms to the model. The Q value is measured by the difference between the original data and the data reconstructed based on the calibrated model. They are associated to each sample of the dataset and large Q values indicate samples that have large out-of-model residuals. Hotelling's T^2 statistic represents the variation in each sample within the model; it is a measure of the sample distance from the center of the model. A sample with a large Hotelling's T^2 value means that this sample has an influence on the developed model.

3. Results and discussion

3.1. FT-IR and FT-NIR spectra of fluconazole polymorphic forms

Fig. 1 reports different recorded spectra for the three pure polymorphic forms of fluconazole. The spectra agree with previously reported spectra of FT-IR and NIR that have been already analyzed and confirmed by XRPD [25,26]. Improving the discrimination ability of the PLS-DA models has been focused on the spectral regions that are responsible for the polymorphic forms. In this case, building PLS-DA models were limited to the region of 3500–2800 cm-1 and 1670–760 cm-1 for FT-IR and from 9000 to 4500 cm-1 for FT-NIR.

3.2. Exploring datasets

Before developing the PLS-DA models, the homogeneity of pharmaceutical products belonging to the same polymorphic form needs to be checked to ensure that there is not any difference between batches of the same polymorphic forms. This is why PCA was used in order to verify their homogeneity. Based on the PCA plot in Fig. 2 using three components for both FT-NIR or FT-IR data, it is noticed that pharmaceutical products that belong to polymorphic form II or the two pharmaceutical products that are from form III are homogenous. It can be summarized that no tablet or pharmaceutical product that may have a polymorphic transformation.

Another step to verify the homogeneity of the entire datasets is done based on the parameters of Hotelling's T^2 and Q residuals. According to these plotted parameters in Fig. 3, it is observed that most samples have a low Hotelling's T^2 and Q residual values. Nevertheless, some samples are located above the threshold. These samples were tested by removing them and comparing RMSECV

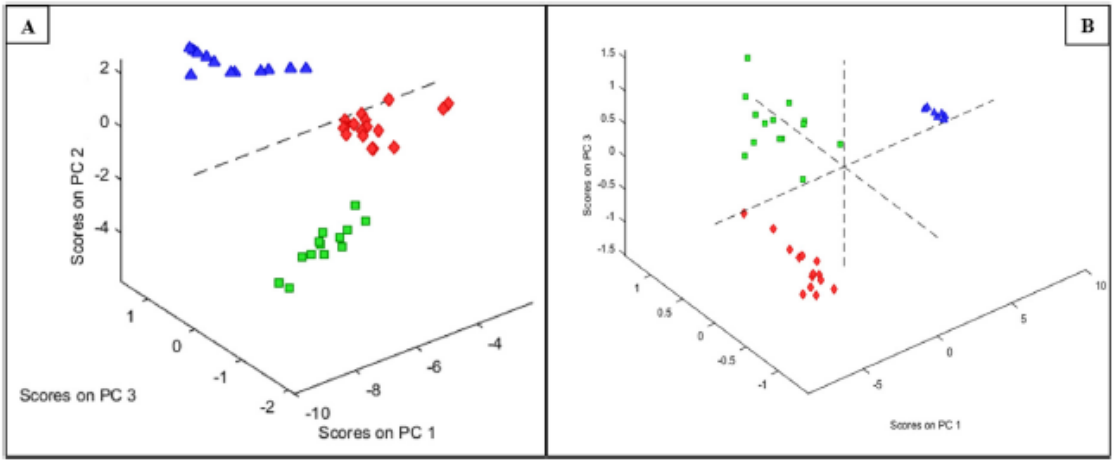

Fig. 2. PCA for the data homogeneity; (A): FT- IR; (B): FT- NIR; Form- III (green square); Form-II (red diamond); Monohydrate form (blue triangle).

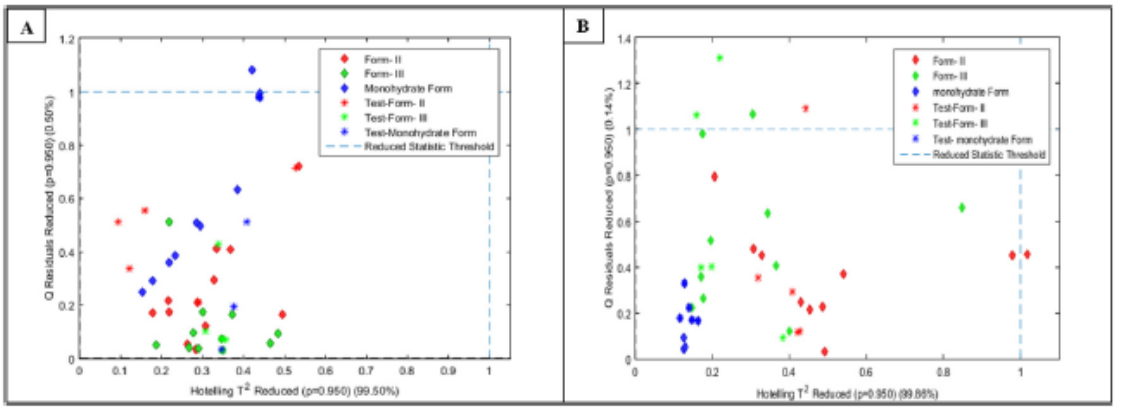

Fig. 3. Hotelling T^2 VS Q Residuals of the training and validation set. (A): FT-IR; (B): FT-NIR;

and RMSEP before and after removing these samples. Since there are no changes in the values of RMSECV and RMSEP, and thus no sample behave as an outlier whether in FT-IR or FT-NIR spectra.

3.3. Development of PLS-DA models for FT-IR and FT-NIR spectra

A PLS-DA model was built for each instrument based on the training set. The optimal number of latent variables was chosen using a cross-validation of venetian blinds with 6 splits. We found that five latent variables for FT-NIR and three for FT-IR minimized the RMSECV.

The parameters of the PLS-DA models are shown in Table 1. These results are examined using accuracy that represents the RMSEP, specificity and selectivity. Based on the obtained results, it is noticed that there is a concordance between RMSECV and RMSEP indicating the accuracy of the PLS-DA model of FT-IR and FT-NIR. These results were confirmed by looking to the discriminant plots of the three polymorphic forms of fluconazole in Fig. 4. The PLS-DA models provides 100 % of specificity and selectivity for both FT-IR and FT-NIR datasets.

3.4. The suitability test of PLS- DA models

The suitability test is included in this work to investigate the robustness of the developed PLS-DA models with new samples that are different from samples that have each of the three polymorphic forms and used calibration set. Two kinds of samples were integrated: first category includes samples that do not have any of the three polymorphic forms. In this case, we included the sample lactose monohydrate as the excipient most commonly used in commercial fluconazole drug products. Each of Miconazole and itraconazole were also included since they belong to the same class of fluconazole. The second set contains fluconazole in a binary mixture: 1) fluconazole form- II, which is the metastable form and may convert to fluconazole monohydrate, and 2) the monohydrate form in different ratios from 1 to 99 %. Hotelling's T^2 and Q residuals were used to ensure that these samples are considered to be outliers.

Fig. 5 illustrates how the three samples including itraconazole, miconazole and monohydrate lactose are different from the training set of FT-IR and FT-NIR because they have high values of Hotelling's T^2 and Q residuals versus the set containing the three polymorphic forms. For FT-IR, three samples: miconazole, itraconazole and lactose monohydrate are suspected to be different from the training set due to their high Q residual. In FT-NIR, the three samples were considered outliers because of their high Hotteling's T^2 and Q residual values. Thus, the plot of Hotelling's T^2 and Q residuals values could distinguish these three samples from the three polymorphic forms of fluconazole.

Fig. 6 shows the Q residual vs Hotelling's T^2 of eleven binary mixtures of fluconazole polymorphic forms.

Fig. 6A demonstrates how FT-IR combined to PLS-DA is able to reject different binary mixtures. Based on this plot, it is noticed that binary samples from the ratio of 5–95% can be distinguished since they have values of Q residual and Hotelling's T^2 significantly

Table 1
Classification parameters for PLS-DA.

Spectroscopy technique	FT-IR			FT-NIR		
Polymorphic Forms	Form- II	Form- III	Monohydrate	Form- II	Form- III	monohydrate
Preprocessing	SG1D (2,15) + MC			SNV + MC		
Spectral range	3500–2800 cm^{-1} & 1670–760 cm^{-1}			9000 to 4500 cm^{-1}		
Cross-validation	Venetian blinds			Venetian blinds		
LV	3			5		
RMSEC (%)	11.1	6.3	10.8	4.7	3.4	2.6
RMSECV (%)	12.5	7.2	11.7	4.8	3.5	3.1
RMSEP (%)	14.1	8.1	11.9	6.5	5.5	2.3
Selectivity Calibration (%)	1	1	1	1	1	1
Specificity Calibration (%)	1	1	1	1	1	1
Selectivity Prediction (%)	1	1	1	1	1	1
Specificity Prediction (%)	1	1	1	1	1	1
Discriminant Threshold	0.4	0.3	0.4	0.6	0.3	0.3

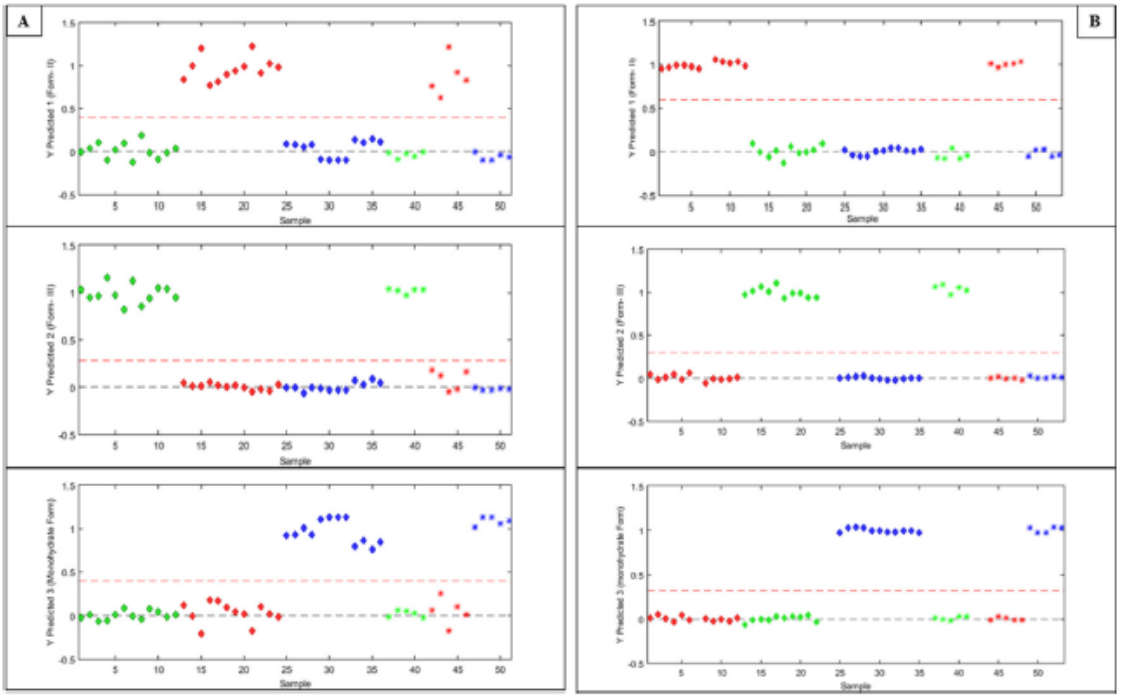

Fig. 4. Discriminant plots of three polymorphic forms: (A): FT-IR (the left); (B): FT-NIR (the right).

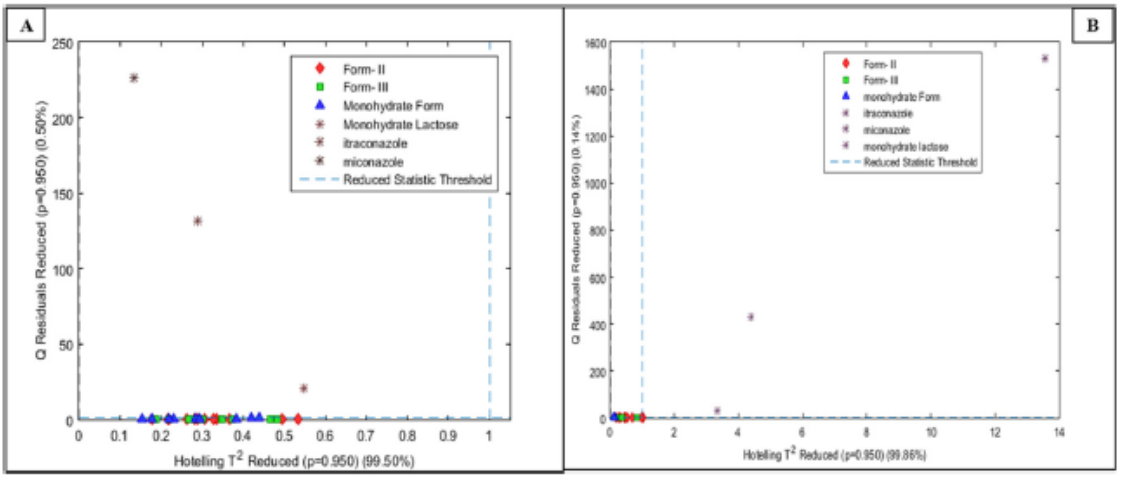

Fig. 5. Hotelling's T^2 Vs Q residual plot of miconazole, itraconazole and monohydrate lactose. A: FT-IR; B: FT-NIR.

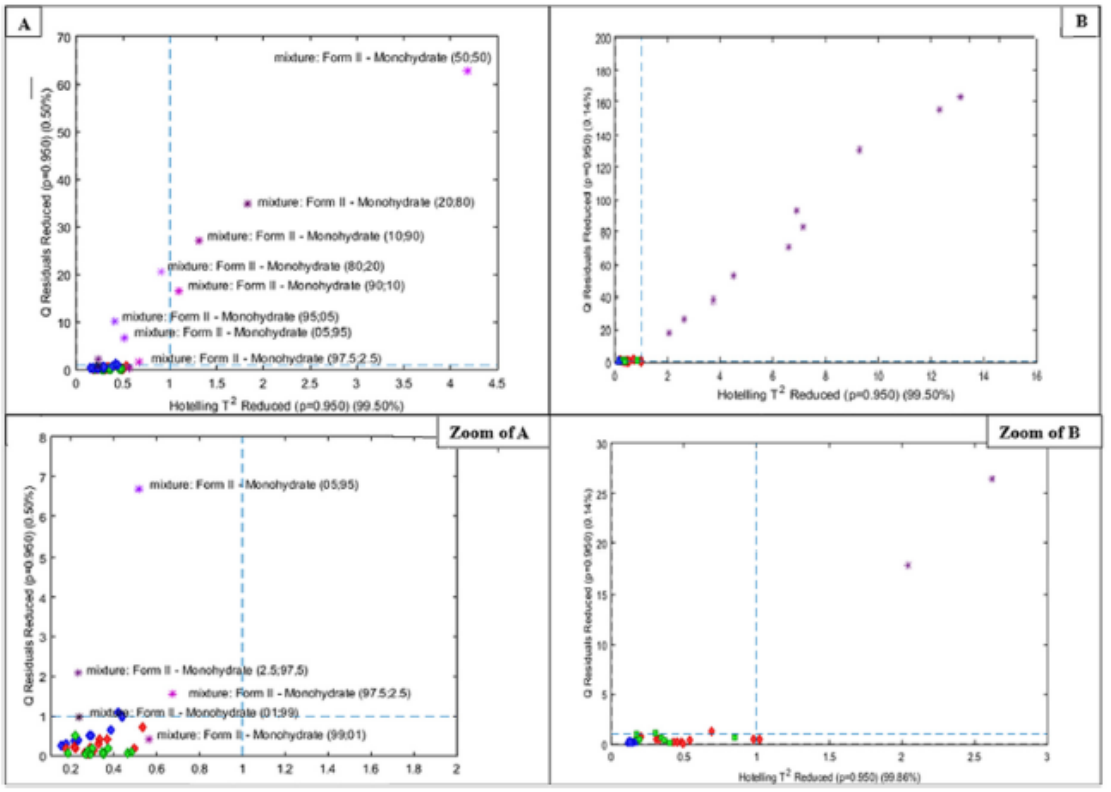

Fig. 6. Hotelling's T^2 Vs Q residual plot of binary mixtures. A: FT-IR; B: FT-NIR.

higher than training set. On the contrary, binary mixture ratios of 1 and 2.5 % were considered belonging to the training set because of their low values of Q residual and Hotelling's T^2 that are similar to the values of training set of polymorphic forms.

Fig. 6B shows how FT-NIR with PLS-DA can be useful to distinguish different binary mixtures from training set. Based on this plot, all binary mixtures consisted of: form II and monohydrate form are well identified as outliers versus the three calibrated polymorphic forms.

Based on these figures, it can be concluded that FT-NIR is able to detect minimum contaminant or conversion of form-II to monohydrate form because of the sensitive property of FT-NIR to hydrates

4. Conclusion

The results obtained with the PLS-DA models proved the suitability and the efficiency of combining vibrational spectroscopy whether it is FT-IR or FT-NIR with chemometric tools for identification and discrimination of fluconazole polymorphic forms. These results were performed based on several parameters summarized in: RMSEP, RMSECV, specificity and sensitivity that proved the ability of associating of FT-IR or FT-NIR to PLS-DA to discriminate between different polymorphic forms in pharmaceuticals. Finally, the suitability of the models was proven by analyzing itraconazole and miconazole as well as different binary mixtures (form II and monohydrate form) using Hotelling's T^2 vs Q residual plot. It has been confirmed that each of itraconazole, miconazole and lactose monohydrate are different since they have both Hotelling's T^2 and Q residual values significantly that are higher than the three main polymorphic forms whether using FT-IR or FT-NIR. Nevertheless, FT-NIR shows to be more suitable than FT-IR for detecting minor contamination between monohydrate form and form-II due to its high sensitivity to hydrates.

CRediT authorship contribution statement

Mohammed Alaoui Mansouri: Conceptualization, Methodology, Data curation, Formal analysis, Investigation, Writing - original draft, Visualization. **Eric Ziemons:** Conceptualization, Validation, Resources, Writing - review & editing, Supervision, Visualization. **Pierre-Yves Sacré:** Conceptualization, Validation, Resources, Writing - review & editing, Supervision, Visualization. **Mourad Kharbach:** Data curation, Formal analysis, Writing - review & editing. **Issam Barra:** Writing - review & editing. **Yahia Cherrah:** Supervision, Project administration, Resources. **Philippe Hubert:** Supervision, Project administration, Funding acquisition, Resources. **Roland Djang'eing'a Marini:** Conceptualization, Methodology, Validation, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Abdelaziz Bouklouze:** Conceptualization, Validation, Resources, Writing - review & editing, Supervision, Visualization, Project administration, Funding acquisition.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.jpba.2021.113922.

Declaration of Competing Interest

The authors report no declarations of interest.

References

[1] H.G. Brittain, S.R. Byrn, E. Lee, Polymorphism in Pharmaceutical Solids, 2009, http://dx.doi.org/10.1016/S0168-3659(01)00252-8.
[2] J. Bernstein, J.M. Bernstein, Polymorphism in Molecular Crystals, Oxford University Press, 2002.

[3] X.Y. Lawrence, Pharmaceutical quality by design: product and process development, understanding, and control, Pharm. Res. 25 (2008) 781–791.
[4] H.G. Brittain, Theory and principles of polymorphic systems, Polymorph. Pharm. Solids. 192 (2009) 1.
[5] J. Aaltonen, M. Allesø, S. Mirza, V. Koradia, K.C. Gordon, J. Rantanen, Solid form screening–a review, Eur. J. Pharm. Biopharm. 71 (2009) 23–37, http://dx.doi.org/10.1016/j.ejpb.2008.07.014.
[6] D. Singhal, W. Curatolo, Drug polymorphism and dosage form design: a practical perspective, Adv. Drug Deliv. Rev. 56 (2004) 335–347.
[7] H.K. Chan, E. Doelker, Polymorphic transformation of some drugs under compression, Drug Dev. Ind. Pharm. 11 (1985) 315–332.
[8] G.G.Z. Zhang, D. Law, E.A. Schmitt, Y. Qiu, Phase transformation considerations during process development and manufacture of solid oral dosage forms, Adv. Drug Deliv. Rev. 56 (2004) 371–390.
[9] E. Atef, H. Chauhan, D. Prasad, D. Kumari, C. Pidgeon, Quantifying solid-state mixtures of crystalline indomethacin by Raman spectroscopy comparison with thermal analysis, ISRN Chromatogr. 2012 (2012).
[10] R. Suryanarayanan, Determination of the relative amounts of α-carbamazepine and β-carbamazepine in a mixture by powder X-ray diffractometry, Powder Diffr. 5 (1990) 155–159, http://dx.doi.org/10.1017/S0885715600015608.
[11] D.E. Bugay, A.W. Newman, W.P. Findlay, Quantitation of cefepime · 2HCl dihydrate in cefepime · 2HCl monohydrate by diffuse reflectance IR and powder X-ray diffraction techniques, J. Pharm. Biomed. Anal. 15 (1996) 49–61, http://dx.doi.org/10.1016/0731-7085(96)01796-7.
[12] A.A. Bunaciu, H.Y. Aboul-Enein, V.D. Hoang, Vibrational spectroscopy used in polymorphic analysis, TrAC Trends Anal. Chem. 69 (2015) 14–22.
[13] S.N.C. Roberts, A.C. Williams, I.M. Grimsey, S.W. Booth, Quantitative analysis of mannitol polymorphs. X-ray powder diffractometry—exploring preferred orientation effects, J. Pharm. Biomed. Anal. 28 (2002) 1149–1159.
[14] A. Heinz, M. Savolainen, T. Rades, C.J. Strachan, Quantifying ternary mixtures of different solid-state forms of indomethacin by Raman and near-infrared spectroscopy, Eur. J. Pharm. Sci. 32 (2007) 182–192.
[15] S. Jabeen, T.J. Dines, S.A. Leharne, B.Z. Chowdhry, Raman and IR spectroscopic studies of fenamates - Conformational differences in polymorphs of flufenamic acid, mefenamic acid and tolfenamic acid, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 96 (2012) 972–985.
[16] Å. Rinnan, Pre-processing in vibrational spectroscopy – when, why and how, Anal. Methods 6 (2014), http://dx.doi.org/10.1039/C3AY42270D.
[17] V. Bhavana, R.B. Chavan, M.K.C. Mannava, A. Nangia, N.R. Shastri, Quantification of niclosamide polymorphic forms – a comparative study by Raman, NIR and MIR using chemometric techniques, Talanta 199 (2019) 679–688, http://dx.doi.org/10.1016/j.talanta.2019.03.027.
[18] N.L. Calvo, T.S. Kaufman, R.M. Maggio, A PCA-based chemometrics-assisted ATR-FTIR approach for the classification of polymorphs of cimetidine: application to physical mixtures and tablets, J. Pharm. Biomed. Anal. 107 (2015) 419–425, http://dx.doi.org/10.1016/j.jpba.2015.01.016.
[19] N.L. Calvo, R.M. Maggio, T.S. Kaufman, A dynamic thermal ATR-FTIR/chemometric approach to the analysis of polymorphic interconversions. Cimetidine as a model drug, J. Pharm. Biomed. Anal. 92 (2014) 90–97, http://dx.doi.org/10.1016/j.jpba.2013.12.036.
[20] Ł. Górski, W. Sordoń, F. Ciepiela, W.W. Kubiak, M. Jakubowska, Voltammetric classification of ciders with PLS-DA, Talanta 146 (2016) 231–236, http://dx.doi.org/10.1016/j.talanta.2015.08.027.
[21] V.A.G. da Silva, M. Talhavini, I.C.F. Peixoto, J.J. Zacca, A.O. Maldaner, J.W.B. Braga, Non-destructive identification of different types and brands of blue pen inks in cursive handwriting by visible spectroscopy and PLS-DA for forensic analysis, Microchem. J. 116 (2014) 235–243, http://dx.doi.org/10.1016/j.microc.2014.05.013.
[22] S.R. Byrn, J.R.R. Pfeiffer, G. Stephenson, D.J.W. Grant, W.B. Gleason, Solid-state pharmaceutical chemistry, Chem. Mater. 6 (1994) 1148–1158, http://dx.doi.org/10.1021/cm00044a013.
[23] K. Richardson, K.W. Brammer, M.S. Marriott, P.F. Troke, Activity of UK-49,858, a bis-triazole derivative, against experimental infections with Candida albicans and Trichophyton mentagrophytes, Antimicrob. Agents Chemother. 27 (1985) 832–835, http://dx.doi.org/10.1128/AAC.27.5.832.
[24] C. Koks, P. Meenhorst, M. Hillebrand, A. Bult, J. Beijnen, Pharmacokinetics of fluconazole in saliva and plasma after administration of an oral suspension and capsules, Antimicrob. Agents Chemother. 40 (1996) 1935–1937 (Accessed November 18, 2016) http://aac.asm.org/cgi/content/long/40/8/1935.
[25] E. Ziémons, H. Bourichi, J. Mantanus, E. Rozet, P. Lebrun, E. Essassi, Y. Cherrah, A. Bouklouze, P. Hubert, Determination of binary polymorphic mixtures of fluconazole using near infrared spectroscopy and X-ray powder diffraction: A comparative study based on the pre-validation stage results, J. Pharm. Biomed. Anal. 55 (2011) 1208–1212, http://dx.doi.org/10.1016/j.jpba.2011.02.019.
[26] H. Bourichi, Y. Brik, E. Essassi, A. Bouklouze, Y. Cherrah, Solid-state characterization of fluconazole, STP Pharma Tech. Prat. Reglementations 17 (2007) 49.
[27] H. Bourichi, Y. Brik, P. Hubert, Y. Cherrah, A. Bouklouze, Solid-state characterization and impurities determination of fluconazol generic products marketed in Morocco, J. Pharm. Anal. 2 (2012) 412–421, http://dx.doi.org/10.1016/j.jpha.2012.05.007.
[28] K.A. Alkhamis, M.S. Salem, R.M. Obaidat, Comparison between dehydration and desolvation kinetics of fluconazole monohydrate and fluconazole ethylacetate solvate using three different methods, J. Pharm. Sci. 95 (2006) 859–870, http://dx.doi.org/10.1002/jps.20605.
[29] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. 58 (2001) 109–130, http://dx.doi.org/10.1016/S0169-7439(01)00155-1.
[30] M.C. Ortiz, L. Sarabia, R. García-Rey, M.D.L. de Castro, Sensitivity and specificity of PLS-class modelling for five sensory characteristics of dry-cured ham using visible and near infrared spectroscopy, Anal. Chim. Acta 558 (2006) 125–131, http://dx.doi.org/10.1016/j.aca.2005.11.038.

Talanta 207 (2020) 120306

Contents lists available at ScienceDirect

Talanta

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

Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study

M. Alaoui Mansouri[a,b], P.-Y. Sacré[a,*], L. Coïc[a], C. De Bleye[a], E. Dumont[a], A. Bouklouze[b], Ph Hubert[a], R.D. Marini[a], E. Ziemons[a]

[a] University of Liege (ULiege), CIRM, Vibra-Santé HUB, Laboratory of Pharmaceutical Analytical Chemistry, CHU, B36, B-4000, Liege, Belgium
[b] Bio-Pharmaceutical and Toxicological Analysis Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V, Rabat, Morocco

ARTICLE INFO

Keywords:
Handheld spectrophotometers
Raman spectroscopy
Spatially offset Raman scattering comparison of quantitative performances
Quantitation through packaging

ABSTRACT

Handheld Raman spectroscopy is actually booming. Recent devices improvements aim at addressing the usual Raman spectroscopy issues: fluorescence with shifted-excitation Raman difference spectroscopy (SERDS), poor sensitivity with surface enhanced Raman scattering (SERS) and information only about the sample surface with spatially offset Raman spectroscopy (SORS). While qualitative performances of handheld devices are generally well established, the quantitative analysis of pharmaceutical samples remains challenging.

The aim of this study was to compare the quantitative performances of three commercially available handheld Raman spectroscopy devices. Two of them (TruScan and IDRaman mini) are equipped with a 785 nm laser wavelength and operate in a conventional backscattering mode. The IDRaman has the Orbital Raster Scanning (ORS) option to increase the analyzed surface. The third device (Resolve) operates with an 830 nm laser wavelength both in backscattering and in SORS modes.

The comparative study was carried out on ibuprofen-mannitol-microcrystalline cellulose ternary mixtures. The concentration of ibuprofen ranged from 24 to 52% (w/w) while the proportions of the two excipients were varied to avoid cross-correlation as much as possible. Analyses were performed either directly through a glass vial or with the glass vial in an opaque polypropylene flask, using a validated FT-NIR spectroscopy method as a reference method. Chemometric analyses were carried out with the Partial Least Squares Regression (PLS-R) algorithm. The quantitative models were validated using the total error approach and the ICH Q2 (R1) guidelines with ± 15% as acceptance limits.

1. Introduction

There is a growing concern toward using vibrational spectroscopy in pharmaceutical quality control [1,2]. Raman spectroscopy is considered as an important analytical tool beside Near-Infrared (NIR) and High Performance Liquid Chromatography (HPLC) [3–5]. This technique is based on the interaction between the energy of a monochromatic light and a sample inducing light scattering. It is characterized by many advantages that may be summarized in its minimal sample preparation requirement, its ability to be used on-site via handheld instrument and the possibility to analyze the sample through clear glass and bottles made from polyvinyl chloride (PVC) or polypropylene (PP). Nevertheless, there are drawbacks that present challenges using the Raman technique. Some of these drawbacks are: sample auto-fluorescence that may overwhelm the signal coming from the analyzed sample, the challenges of analyzing heterogeneous samples because of the small analyzed volume and of carrying out a qualitative or quantitative analysis of a compound in a mixture through a package material such as blisters or plastic bottles [6–8].

The last cited challenge of analyzing a drug substance through package material could be overcome using a specific measuring configuration mode called spatially offset Raman scattering (SORS). The main difference between SORS and backscattering Raman is that the scattered light is collected at a spatially offset location situated a few millimeters from the illumination site. This configuration allows collecting photons that have gone deep in the sample leading to spectra predominantly composed of content's signal [9,10]. To obtain the SORS corrected spectra, the outer (container) spectrum is scaled and removed from the offset spectrum in order to obtain a clean spectrum of the content [11]. SORS has already proved its usefulness in many sectors.

* Corresponding author.
E-mail address: pysacre@uliege.be (P.-Y. Sacré).

https://doi.org/10.1016/j.talanta.2019.120306
Received 4 June 2019; Received in revised form 28 August 2019; Accepted 30 August 2019
Available online 31 August 2019
0039-9140/

Talanta 207 (2020) 120306

Contents lists available at ScienceDirect

Talanta

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

Quantitation of active pharmaceutical ingredient through the packaging using Raman handheld spectrophotometers: A comparison study

M. Alaoui Mansouri[a,b], P.-Y. Sacré[a,*], L. Coïc[a], C. De Bleye[a], E. Dumont[a], A. Bouklouze[b], Ph Hubert[a], R.D. Marini[a], E. Ziemons[a]

[a] University of Liege (ULiege), CIRM, Vibra-Santé HUB, Laboratory of Pharmaceutical Analytical Chemistry, CHU, B36, B-4000, Liege, Belgium
[b] Bio-Pharmaceutical and Toxicological Analysis Research Team, Laboratory of Pharmacology and Toxicology, Faculty of Medicine and Pharmacy, University Mohammed V, Rabat, Morocco

ARTICLE INFO

Keywords:
Handheld spectrophotometers
Raman spectroscopy
Spatially offset Raman scattering comparison of quantitative performances
Quantitation through packaging

ABSTRACT

Handheld Raman spectroscopy is actually booming. Recent devices improvements aim at addressing the usual Raman spectroscopy issues: fluorescence with shifted-excitation Raman difference spectroscopy (SERDS), poor sensitivity with surface enhanced Raman scattering (SERS) and information only about the sample surface with spatially offset Raman spectroscopy (SORS). While qualitative performances of handheld devices are generally well established, the quantitative analysis of pharmaceutical samples remains challenging.

The aim of this study was to compare the quantitative performances of three commercially available handheld Raman spectroscopy devices. Two of them (TruScan and IDRaman mini) are equipped with a 785 nm laser wavelength and operate in a conventional backscattering mode. The IDRaman has the Orbital Raster Scanning (ORS) option to increase the analyzed surface. The third device (Resolve) operates with an 830 nm laser wavelength both in backscattering and in SORS modes.

The comparative study was carried out on ibuprofen-mannitol-microcrystalline cellulose ternary mixtures. The concentration of ibuprofen ranged from 24 to 52% (w/w) while the proportions of the two excipients were varied to avoid cross-correlation as much as possible. Analyses were performed either directly through a glass vial or with the glass vial in an opaque polypropylene flask, using a validated FT-NIR spectroscopy method as a reference method. Chemometric analyses were carried out with the Partial Least Squares Regression (PLS-R) algorithm. The quantitative models were validated using the total error approach and the ICH Q2 (R1) guidelines with ± 15% as acceptance limits.

1. Introduction

There is a growing concern toward using vibrational spectroscopy in pharmaceutical quality control [1,2]. Raman spectroscopy is considered as an important analytical tool beside Near-Infrared (NIR) and High Performance Liquid Chromatography (HPLC) [3–5]. This technique is based on the interaction between the energy of a monochromatic light and a sample inducing light scattering. It is characterized by many advantages that may be summarized in its minimal sample preparation requirement, its ability to be used on-site via handheld instrument and the possibility to analyze the sample through clear glass and bottles made from polyvinyl chloride (PVC) or polypropylene (PP). Nevertheless, there are drawbacks that present challenges using the Raman technique. Some of these drawbacks are: sample auto-fluorescence that may overwhelm the signal coming from the analyzed sample, the challenges of analyzing heterogeneous samples because of the small analyzed volume and of carrying out a qualitative or quantitative analysis of a compound in a mixture through a package material such as blisters or plastic bottles [6–8].

The last cited challenge of analyzing a drug substance through package material could be overcome using a specific measuring configuration mode called spatially offset Raman scattering (SORS). The main difference between SORS and backscattering Raman is that the scattered light is collected at a spatially offset location situated a few millimeters from the illumination site. This configuration allows collecting photons that have gone deep in the sample leading to spectra predominantly composed of content's signal [9,10]. To obtain the SORS corrected spectra, the outer (container) spectrum is scaled and removed from the offset spectrum in order to obtain a clean spectrum of the content [11]. SORS has already proved its usefulness in many sectors.

* Corresponding author.
E-mail address: pysacre@uliege.be (P.-Y. Sacré).

https://doi.org/10.1016/j.talanta.2019.120306
Received 4 June 2019; Received in revised form 28 August 2019; Accepted 30 August 2019
Available online 31 August 2019
0039-9140/

handheld Raman instruments. FT-NIR spectra were only acquired in reflectance mode on the glass vial.

2.3. Multivariate data analysis

The regression model was developed based on the partial least square (SIMPLS) algorithm using the PLS Toolbox V8.2.1 (Eigenvector Research INC, USA) running on Matlab (R2018b) (The Mathworks, USA). The Y block used was composed of actual weights and the same Y block was used for all instruments. Different preprocessing techniques were investigated and compared based on the root mean square difference of prediction (RMSEP). The combination of standard normal variate (SNV) normalization with mean centering proved to be the most suitable for FT-NIR data, while the combination of the Savitzky-Golay 1st derivative (polynomial order: 2, window size: 15), SNV and mean centering provided better predictions for Raman spectroscopy.

The accuracy profiles were computed using the results from the three validation series composed of three replicates at five concentration levels measured on each device using e-noval V4.0b (Pharmalex Belgium SA, Mont-Saint-Guibert, Belgium) with 95% β-expectation tolerance intervals [16,18–20]. The acceptance limits were set at ± 15% of relative total error following the European Pharmacopoeia general monograph 2.9.40 on the uniformity of dosage units.

3. Results and discussion

Actually, handheld Raman spectrophotometers are designed for qualitative analysis since the acquisition time is optimized at each measurement to obtain a sufficient signal to noise ratio [2]. Moreover, quantitation of active ingredients becomes a challenge when this is to be performed through thick and opaque containers. During this study, different ternary mixtures of ibuprofen/mannitol/MCC were prepared following the scheme described in section 2.2. shows the Raman spectra of each raw material and of the ternary mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. Once the samples were prepared and placed in the glass vials, they were first analyzed by FT-NIR spectroscopy. The obtained NIR spectra were then processed by PLS-R allowing us to ensure that the mixtures were correctly prepared (the detected outliers were removed and prepared again). The computed PLS-R model based on FT-NIR data has been validated and the accuracy profile computed. The beta-expectation tolerance intervals (β = 95%) was well included in the previously set ± 15% of relative total error.

Once the samples were verified by FT-NIR, they underwent analyses by the different handheld systems directly through the glass vial and through the glass vial placed in the PP container (see photos in supplementary materialsFig. S1).

3.1. Analysis of samples through glass vial

Fig. 2 shows the baseline corrected spectra acquired on each device in backscattering mode through the glass vial for a mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. The mixture and ibuprofen spectra were recorded through a thin plastic bag to avoid any interference from the glass. The dashed red lines indicate the main spectral features of ibuprofen and the blue dashed line indicates the main peak of mannitol. No specific peak associated to cellulose was observable because cellulose is a weak Raman scatterer and its features are masked by the other components. Nevertheless, cellulose disturbs the global signal due to a high fluorescence background. On this Fig. 2, it is possible to see that the signal measured with each device exhibits a spectrum directly correlated to the mixture spectrum with the main spectral features of ibuprofen and mannitol. The spectra recorded with the TruScan and the IDRaman mini also exhibit a spectral perturbation between 1300 and 1500 cm^{-1} due to the fluorescence of glass with the 785 nm incident laser source.

The spectra obtained with each instrument (ID Raman mini, TruScan, Resolve) were modelled using PLS models. Several pre-processing and spectral ranges were tested. The best pre-processing and spectral range were selected based on the RMSEP and the bias computed on the validation set. Once the final PLS model selected, the accuracy profiles were computed for each device. Table 2 summarizes the final parameters used for the PLS models and their respective figures of merit. Each Raman spectroscopy model used the Savitzky-Golay [21] first derivative as pre-processing. Indeed, most devices use a 785 nm laser as light source. However, glass exhibits a high fluorescence background when irradiated by a 785 nm light and the use of the first derivative helped managing the fluorescence.

The computed accuracy profiles are shown in Fig. 3 and the values of the validation criteria are reported in Table 3. As can be noticed, all the selected devices provided good results as their 95% β-expectation tolerance intervals are included inside the acceptance limits of ± 15%. This means that 95% of future measurements will have an accuracy (total error) of less than ± 15%. These results indicate that for formulations with a well-balanced signal of both excipients and API in a transparent container, satisfying quantitative performances may be

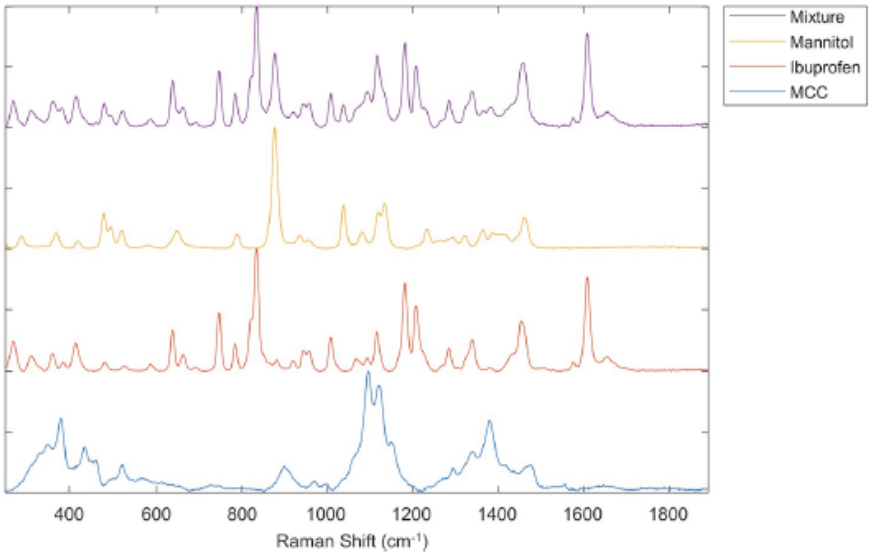

Fig. 1. Raman spectra of raw materials and ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively. The spectra were acquired with the TruScan in "signature" mode through a thin plastic bag and were baseline corrected by asymmetric least squares [22] [22] [22] [23] [23] with parameters: lambda 10^5, p: 10^{-3}.

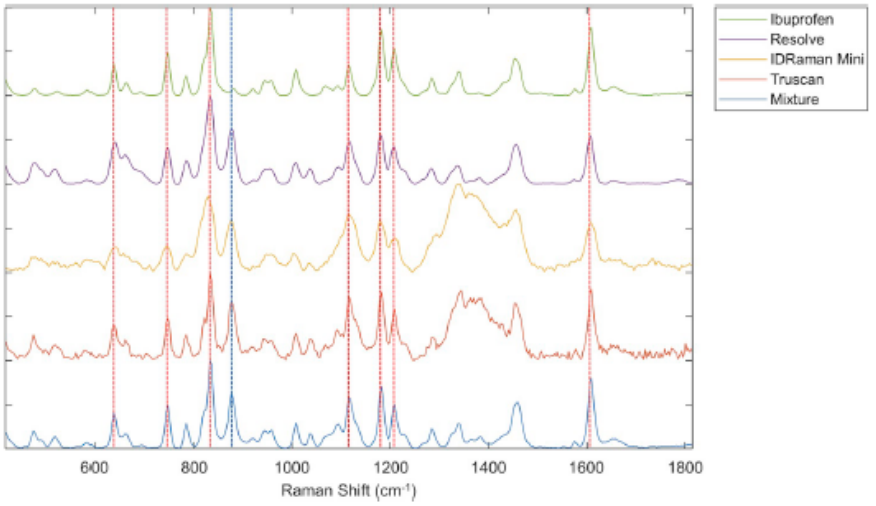

Fig. 2. Raman spectra of a ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively. The spectra were acquired with each handheld device in the backscattering mode through the glass vial. The spectra were baseline corrected by asymmetric least squares with parameters: lambda 10^5, p: 10^{-3}. Reference spectra of ibuprofen and the ternary mixture were acquired as described in Fig. 1. Ibuprofen and mannitol spectral features are marked by dashed red and blue lines respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

obtained using Raman handheld devices in their native configuration (auto-exposure).

3.2. Analysis of samples through glass vials placed in a polypropylene container

Fig. 4 shows the baseline corrected spectra acquired on each device in backscattering mode and SORS mode (for the Resolve) through the glass vial placed in the PP container for a mixture of ibuprofen/MCC/mannitol in the proportions 4/3/3, respectively. The dashed red lines indicate the main spectral features of ibuprofen and the blue dashed line indicates the main peak of mannitol. Compared to the spectra shown on Fig. 2, the spectra recorded through the PP container show no spectral features associated with ibuprofen nor mannitol except for the SORS spectra. It is worth noting that the SORS spectra presented here and subsequently used in the quantitative modelling are only the offset part of the spectrum. Indeed, usually final SORS spectra are obtained after removal of the zero spectrum (equivalent to the backscattering recorded spectrum) from the offset spectrum to remove the residual container spectral features. Since the SORS correction parameters (baseline correction and removal of the scaled "zero offset" spectrum) are computed for each spectrum separately, this led to additional random error on the quantitative models. Therefore, to avoid errors when removing the zero offset, this step was skipped and the offset spectrum was directly used.

Once again, PLS models were built for each device and several pre-processing and spectral ranges were tested. Accuracy profiles were computed based on the predicted values from the PS models. The results are summarized in Fig. 5 and the parameters values of the validated models are presented in Table 3. None of the backscattering devices was able to quantify ibuprofen through the PP container because no (or very few) signal originating from the sample was measured in this configuration. Indeed, after applying the mean centering prior to the modelling, no residual signal was observed, only noise. That means that all spectra were the same for each validation and calibration sample because only the PP signal was recorded.

However, the Resolve operating in the SORS mode was able to achieve satisfying quantitation of the sample through the PP container since its accuracy profile was completely included in the acceptance limits. Furthermore, SORS measurements are more representative of the sample since the recorded signal has gone through a higher sample volume.

4. Conclusion

The aim of this study was to carry out a comparison study between conventional backscattering and Spatially Offset Raman Scattering (SORS) Raman handheld instruments. This comparison has been performed measuring a ternary mixture of ibuprofen/MCC/mannitol directly through a glass vial and with the glass vial placed in a PP container. PLS models were built for each device and each measurement configuration. The predicted values obtained from the PLS models on a validation set were used to compute accuracy profiles following the ICH Q2 R1 guidelines on validation with ± 15% as acceptance limits.

Table 2
Regression model parameters and figures of merit of handheld Raman devices and FT-NIR.

	Resolve			TruScan RM		IDRaman mini		FT- NIR
Sample	Glass vial + PP	Glass vial + PP	Glass vial	Glass vial + PP	Glass vial	Glass vial + PP	Glass vial	Glass vial
Mode	SORS	Backscattering	Backscattering	Backscattering	Backscattering	Backscattering	Backscattering	Reflection
Spectral Range (cm^{-1})	1152–1608	400–1700	400–1700	1000–1600	400–1700	600–1700	400–1700	9000–4000
Pre-processing	SG1D (2,15); SNV; MC			SG1D (2,15); SNV; MC		SG1D (2,15); SNV; MC		SNV; MC
Latent Variables	6	7	3	6	3	5	3	5
R^2 calibration	98.7	97.2	96.0	98.7	94.0	99.8	98.7	99.0
RMSEC (mg)	5.6	8.5	10.1	5.9	12.4	2.4	5.6	5.3
R^2 Cross Validation	97.8	89.3	95.7	58.8	92.3	37.2	95.0	98.7
RMSECV (mg)	7.4	16.6	10.5	32.3	14.2	40.6	11.0	5.9
R^2 prediction	93.0	83.1	92.3	66.4	91.6	9.5	96.5	95.5
RMSEP (mg)	12.7	19.8	12.6	27.0	12.7	45.4	13.6	9.0

SG1D: Savitzky-Golay first derivative (polynomial order, window size).
SNV: standard normal variate.
MC: mean centering.

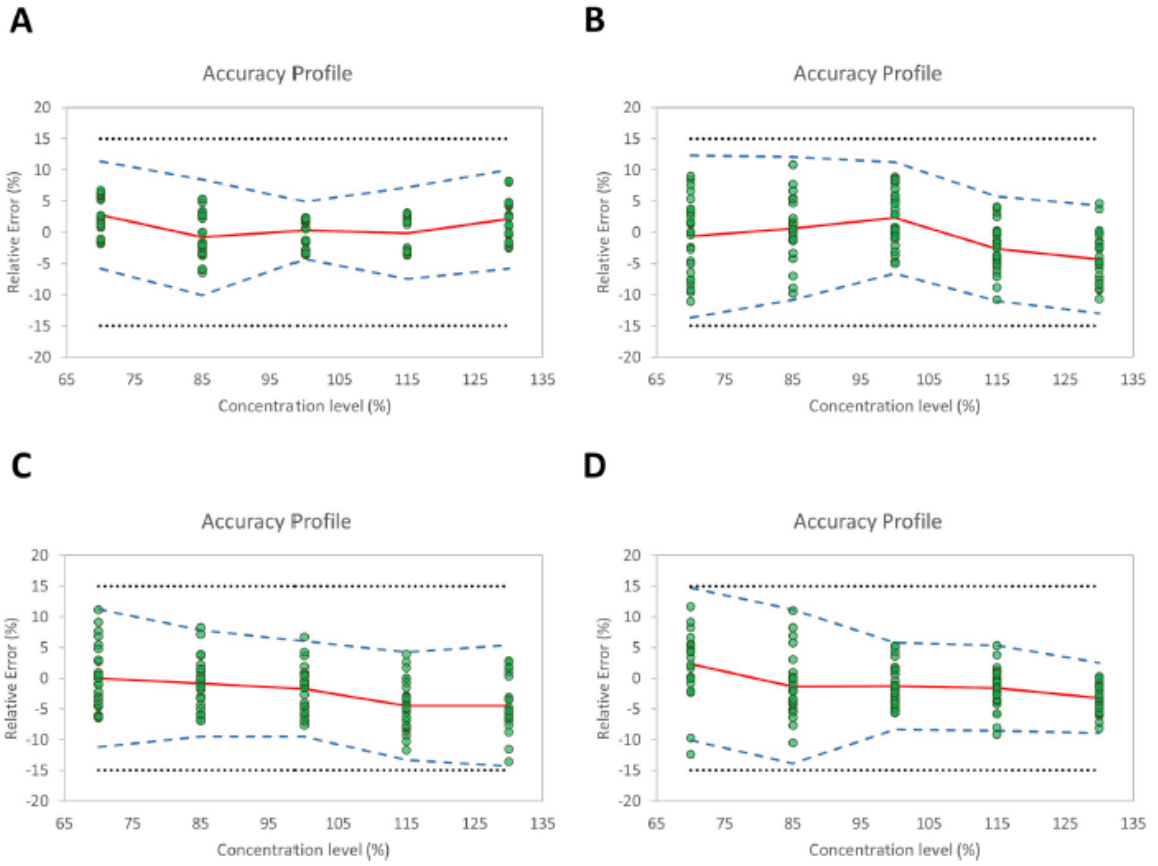

Fig. 3. Accuracy profiles obtained with each handheld device (A: MPA, B: TruScan, C: IDRaman Mini, D: Resolve) in backscattering mode through the glass vial. The plain red line is the relative bias, the dashed blue lines are the β-expectation tolerance limits (β = 95%) and the dotted black lines are the acceptance limits set at 15%. The green dots represent the relative errors of each validation results. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

By measuring through the glass vial, the Raman spectra showed clear features associated with both the API and the excipients leading to satisfying quantitative performances. The subsequent models and predictions were validated and their accuracy profiles were included inside the a priori defined acceptance limits. This confirms the fact that it is possible to obtain reliable quantitative information with handheld devices with their auto-exposure default configuration. However, it is worth noting that this is only true for the studied formulation with a

Table 3
ICH Q2 (R1) validation criteria values of the PLS models.

	Concentration level	SORS (glass vial + PP)	Resolve (glass vial)	TruScan RM (glass vial)	IDRaman mini (glass vial)	FT-NIR
Trueness	70	0.665	2.319	−0.647	0.034	2.790
Relative bias (%)	85	−1.226	−1.315	0.637	−0.848	−0.771
	100	−0.377	−1.289	2.332	−1.726	0.291
	115	−1.688	−1.587	−2.632	−4.516	−0.150
	130	−3.030	−3.197	−4.335	−4.484	2.138
Intra-assay precision	70	3.308	5.813	5.953	5.206	2.338
Repeatability (RSD %)	85	3.400	4.872	5.002	3.984	3.346
	100	2.406	3.375	4.244	3.707	2.205
	115	2.883	3.308	3.873	3.884	2.230
	130	3.483	1.968	3.586	3.617	2.777
Between-assay precision	70	4.824	5.813	6.117	5.283	3.276
Intermediate precision (RSD %)	85	3.400	5.525	5.251	4.071	3.971
	100	2.468	3.375	4.244	3.707	2.205
	115	3.331	3.308	3.922	4.089	2.932
	130	4.103	2.415	3.928	4.277	3.372
Accuracy	70	[-12.52, 13.85]	[-10.11, 14.75]	[-13.64, 12.35]	[-11.20, 11.27]	[-5.82, 11.40]
Relative β -expectation tolerance limits	85	[-8.41, 5.95]	[-13.85, 11.22]	[-10.79, 12.06]	[-9.53, 7.83]	[-10.04, 8.50]
(%)	100	[-5.60, 4.85]	[-8.36, 5.78]	[-6.58, 11.25]	[-9.49, 6.04]	[-4.33, 4.91]
	115	[-9.25, 5.87]	[-8.52, 5.34]	[-10.99, 5.73]	[-13.28, 4.25]	[-7.48, 7.18]
	130	[-12.47, 6.41]	[-8.92, 2.53]	[-12.97, 4.30]	[-14.35, 5.38]	[-5.79, 10.07]

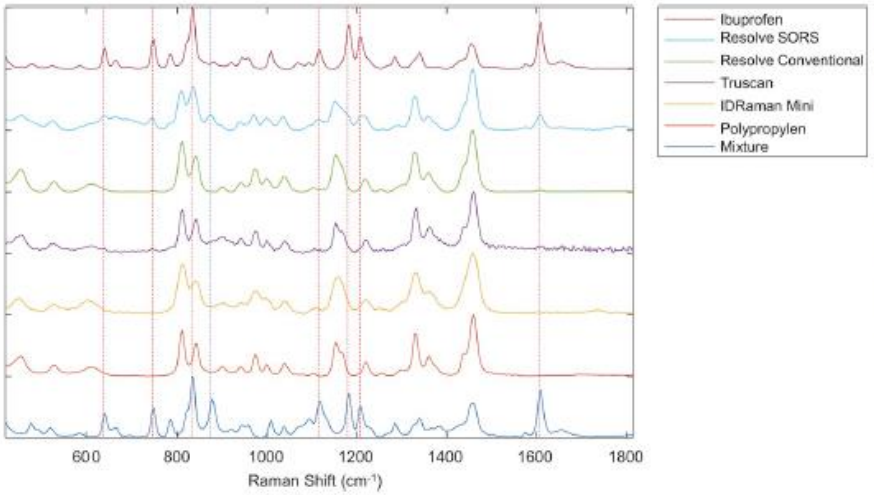

Fig. 4. Raman spectra of a ternary mixture ibuprofen/MCC/mannitol in the proportions 4/3/3 respectively. The spectra were acquired with each handheld device in the backscattering mode and the Resolve in SORS mode through the glass vial placed in the polypropylene container. The spectra were baseline corrected by asymmetric least squares with parameters: lambda 10^5, p: 10^{-3}. Reference spectra of ibuprofen and the ternary mixture were acquired as described in Fig. 1. Ibuprofen and mannitol spectral features are marked by dashed red and blue lines respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

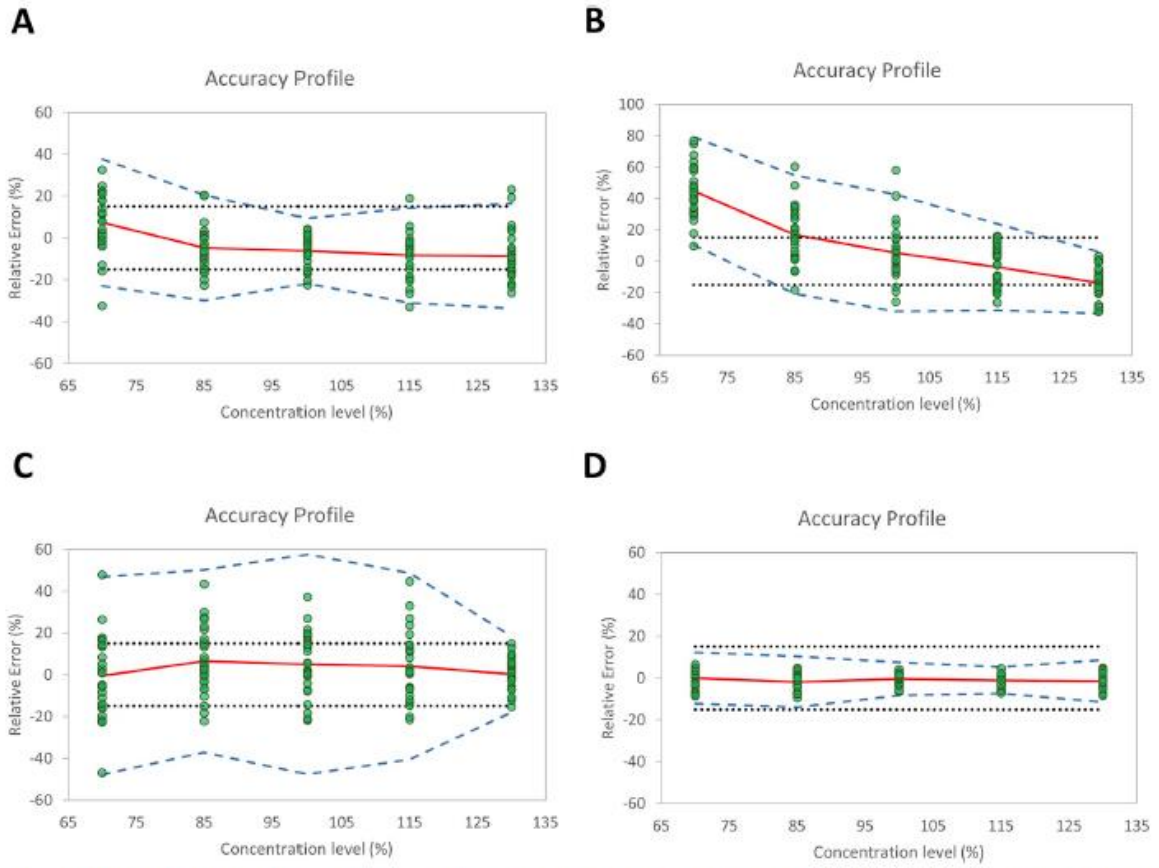

Fig. 5. Accuracy profiles obtained with each handheld device (A: TruScan, B: IDRaman Mini, C: Resolve backscattering and D: Resolve SORS) mode through the glass vial placed in the opaque polypropylene container. The plain red line is the relative bias, the dashed blue lines are the β-expectation tolerance limits (β = 95%) and the dotted black lines are the acceptance limits set at 15%. The green dots represent the relative errors of each validation sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

well-balanced signal between the API and excipients.

However, none of the backscattering Raman handheld devices was able to quantify the mixture when it was placed in an opaque 1 mm PP and 1 mm thick glass containers. To be able to pass through the packaging, the SORS measurement configuration was necessary and allowed to obtain a valid PLS model.

These preliminary results pave the way to reliable quantitative Raman measurements directly in the field through opaque containers.

Acknowledgments

This project has been supported by the European funds of regional development (FEDER) and by Walloon as part of the operational program "Walloon-2020.EU".

The financial support of this research by the Walloon Region of Belgium in the framework of the Vibra4Fake project (convention n°:7517) is gratefully acknowledged.

This work was supported by the National Fund for Scientific Research, FNRS-F.R.S. (1.A030.17 - FC6921).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.talanta.2019.120306.

References

[1] D. Sorak, L. Herberholz, S. Iwascek, S. Altinpinar, F. Pfeifer, H.W. Siesler, New developments and applications of handheld Raman, mid-infrared, and near-infrared spectrometers, Appl. Spectrosc. Rev. 47 (2012) 83–115, https://doi.org/10.1080/05704928.2011.625748.

[2] R. Deidda, P.-Y. Sacre, M. Clavaud, L. Coïc, H. Avohou, P. Hubert, E. Ziemons, Vibrational spectroscopy in analysis of pharmaceuticals: critical review of innovative portable and handheld NIR and Raman spectrophotometers, TrAC Trends Anal. Chem. 114 (2019) 251–259, https://doi.org/10.1016/j.trac.2019.02.035.

[3] J.K. Mbinze, P.-Y. Sacré, A. Yemoa, J. Mavar Tayey Mbay, V. Habyalimana, N. Kalenda, P. Hubert, R.D. Marini, E. Ziemons, Development, validation and comparison of NIR and Raman methods for the identification and assay of poor-quality oral quinine drops, J. Pharm. Biomed. Anal. 111 (2015) 21–27, https://doi.org/10.1016/j.jpba.2015.02.049.

[4] C. Eliasson, P. Matousek, Noninvasive authentication of pharmaceutical products through packaging using spatially offset Raman spectroscopy, Anal. Chem. 79 (2007) 1696–1701.

[5] D. Riolo, A. Piazza, C. Cottini, M. Serafini, E. Lutero, E. Cuoghi, L. Gasparini, D. Botturi, I.G. Marino, I. Aliatis, D. Bersani, P.P. Lottici, Raman spectroscopy as a PAT for pharmaceutical blending: advantages and disadvantages, J. Pharm. Biomed. Anal. 149 (2018) 329–334, https://doi.org/10.1016/j.jpba.2017.11.030.

[6] Michael J. Pelletier, Analytical Applications of Raman Spectroscopy, Wiley-Blackwell, Oxford, UK, 1999.

[7] C.E. Miller, Process analytical technology : spectroscopic tools and implemented strategies for the chemical and pharmaceutical industries, Chemom. Process Anal. Chem. John Wiley & Sons, 2010, pp. 237–285 http://www.worldcat.org/isbn/047072207.

[8] L.B. Lyndgaard, F. van den Berg, A. de Juan, Quantification of paracetamol through tablet blister packages by Raman spectroscopy and multivariate curve resolution-alternating least squares, Chemometr. Intell. Lab. Syst. 125 (2013) 58–66, https://doi.org/10.1016/j.chemolab.2013.03.014.

[9] A.W. Parker, Spatially offset Raman spectroscopy, Encycl. Spectrosc. Spectrom. Elsevier, 2017, pp. 143–148, https://doi.org/10.1016/B978-0-12-409547-2.12151-1.

[10] W.J. Olds, E. Jaatinen, P. Fredericks, B. Cletus, H. Panayiotou, E.L. Izake, Spatially offset Raman spectroscopy (SORS) for the analysis and detection of packaged pharmaceuticals and concealed drugs, Forensic Sci. Int. 212 (2011) 69–77, https://doi.org/10.1016/j.forsciint.2011.05.016.

[11] M. Bloomfield, D. Andrews, P. Loeffen, C. Tombling, T. York, P. Matousek, Non-invasive identification of incoming raw pharmaceutical materials using Spatially Offset Raman Spectroscopy, J. Pharm. Biomed. Anal. 76 (2013) 65–69, https://doi.org/10.1016/j.jpba.2012.11.046.

[12] D.I. Ellis, R. Eccles, Y. Xu, J. Griffen, H. Muhamadali, P. Matousek, I. Goodall, R. Goodacre, Through-container, extremely low concentration detection of multiple chemical markers of counterfeit alcohol using a handheld SORS device, Sci. Rep. 7 (2017) 1–8, https://doi.org/10.1038/s41598-017-12263-0.

[13] P. Zhang, D. Littlejohn, Interference assessment and correction in the partial least squares regression method for multicomponent determination by UV spectrophotometry, Chemometr. Intell. Lab. Syst. 34 (1996) 203–215, https://doi.org/10.1016/0169-7439(96)00029-9.

[14] M. Forina, S. Lanteri, M. Casale, Multivariate calibration, J. Chromatogr., A 1158 (2007) 61–93, https://doi.org/10.1016/j.chroma.2007.03.082.

[15] P. Hubert, J.-J. Nguyen-Huu, B. Boulanger, E. Chapuzet, P. Chiap, N. Cohen, P.-A. Compagnon, W. Dewé, M. Feinberg, M. Lallier, M. Laurentie, N. Mercier, G. Muzard, C. Nivet, L. Valat, Harmonization of strategies for the validation of quantitative analytical procedures, J. Pharm. Biomed. Anal. 36 (2004) 579–586, https://doi.org/10.1016/j.jpba.2004.07.027.

[16] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures: a SFSTP proposal part IV. Examples of application, J. Pharm. Biomed. Anal. 48 (2008) 760–771, https://doi.org/10.1016/j.jpba.2008.07.018.

[17] A. Bouabidi, E. Rozet, M. Fillet, E. Ziemons, E. Chapuzet, B. Mertens, R. Klinkenberg, A. Ceccato, M. Talbi, B. Streel, A. Bouklouze, B. Boulanger, P. Hubert, Critical analysis of several analytical method validation strategies in the framework of the fit for purpose concept, J. Chromatogr., A 1217 (2010) 3180–3192, https://doi.org/10.1016/j.chroma.2009.08.051.

[18] P. Hubert, P. Chiap, J. Crommen, B. Boulanger, E. Chapuzet, N. Mercier, S. Bervoas-Martin, P. Chevalier, D. Grandjean, P. Lagorce, M. Lallier, M.C. Laparra, M. Laurentie, J.C. Nivet, The SFSTP guide on the validation of chromatographic methods for drug bioanalysis from the Washington Conference to the laboratory, Anal. Chim. Acta 391 (1999) 135–148.

[19] P. Hubert, J.J. Nguyen-Huu, B. Boulanger, E. Chapuzet, N. Cohen, P.A. Compagnon, W. Dewe, M. Feinberg, M. Laurentie, N. Mercier, G. Muzard, L. Valat, E. Rozet, Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal–part III, J. Pharm. Biomed. Anal. 45 (2007) 82–96, https://doi.org/10.1016/j.jpba.2007.06.032.

[20] C. De Bleye, P.F. Chavez, J. Mantanus, R. Marini, P. Hubert, E. Rozet, E. Ziemons, Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications, J. Pharm. Biomed. Anal. 69 (2012) 125–132, https://doi.org/10.1016/j.jpba.2012.02.003.

[21] A. Savitzky, M.J.E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem. 36 (1964) 1627–1639, https://doi.org/10.1021/ac60214a047.

[22] P.H.C. Eilers, A perfect smoother, Anal. Chem. 75 (2003) 3631–3636, https://doi.org/10.1021/ac034173t.

SUMMARY

The aim of this thesis was to investigate the ability of the association of vibrational spectroscopic techniques to chemometric tools to overcome different aspects of matrix effects for pharmaceutical analysis. The first aspect was focused on testing the ability of applying PLS-DA to vibrational spectroscopy to discriminate between different pharmaceuticals based on three main polymorphic forms of the fluconazole. PLS-DA showed its ability to discriminate between samples have only one of three polymorphic forms of fluconazole. While in case of the matrix effect, PLS-DA showed its limit to do the right discrimination. Thus, approach of Hotelling's T^2 and Q residuals to detect samples with two polymorphic forms as outliers. The second aspect of matrix effect focused on evaluating the application each of PLS regression and MCR-ALS model on FT-NIR data to quantify ciprofloxacin in different brands of pharmaceutical products. the quantitation of ciprofloxacin in different brands of pharmaceutical products clearly showed the limit of PLS regression because of the matrix effect, whereas the MCR-ALS based on its low relative and prediction errors, has shown its ability to overcome the problem of the change in the matrix composition due to its second order advantage. The third aspect of interference effect aims to investigate a property of Raman spectroscopy to quantify the ibuprofen in a ternary mixture through a container interference. This Raman property is known by SORS, which allows laser to pass through the packaging. This investigation is based on a comparison study between backscattering and SORS mode to quantify ibuprofen through the polypropylene by evaluating the PLS regression models by accuracy profiles on validation with ±15% acceptance limits. Based on the accuracy profiles, the SORS demonstrated its ability to quantify through the interference which is not in case of backscattering.

Keywords: Matrix effect; Vibrational spectroscopy; chemometrics; pharmaceutical drug analysis

Research structure : Biopharmaceutical and toxicological analysis research team

RESUME

L'objectif de cette thèse était d'évaluer la capacité de l'association des techniques de spectroscopiques à la chimiométrie à surmonter les différents aspects des effets de matrice pour les analyses pharmaceutique. Le premier aspect visait à tester la capacité d'appliquer PLS-DA à la spectroscopie pour discriminer entre différents produits pharmaceutiques à base de trois principales formes polymorphiques du fluconazole. Le PLS-DA a montré sa capacité à discriminer entre les échantillons n'ayant qu'une des trois formes polymorphes de fluconazole. Alors que dans le cas de l'effet matrice, PLS-DA a montré sa limite pour faire la bonne discrimination. Ainsi, approche Hotteling's T^2 et Qresiduels pour détecter les échantillons avec deux formes polymorphes comme valeurs aberrantes. Le deuxième aspect de l'effet de matrice s'est concentré sur l'évaluation de l'application de la régression PLS et du modèle MCR-ALS avec FT-NIR pour quantifier le ciprofloxacine dans différentes marques pharmaceutiques. la quantification de la ciprofloxacine dans différentes marques de produits pharmaceutiques a clairement montré la limite de la régression PLS en raison de l'effet de matrice, alors que le MCR-ALS a montré sa capacité de surmonter le problème du changement de la composition de la matrice en raison de son avantage de second ordre. Le troisième aspect vise à étudier une propriété de la spectroscopie Raman pour quantifier l'ibuprofène dans un mélange ternaire à travers une interférence de conteneur. Cette propriété Raman est connue de SORS, qui permet au laser de traverser l'emballage. Cette enquête est basée sur une étude de comparaison entre la rétrodiffusion et le mode SORS pour quantifier l'ibuprofène à travers le polypropylène en évaluant les modèles de régression PLS par des profils avec des limites d'acceptation de ± 15 %. Sur la base des profils de précision, le SORS a démontré sa capacité à quantifier à travers l'interférence ce qui n'est pas en cas de rétrodiffusion.

Keywords : Effet de matrice ; Chimiométrie, spectroscopie vibrationnelle ; analyses pharmaceutiques

Structure de recherche: Equipe de recherche des analyses biopharmaceutiques et toxicologiques