

**DEVELOPMENT OF A METHODOLOGY TO DETERMINE DRUG DISTRIBUTION IN
POLYMERIC THIN FILM FORMULATIONS USING HYPERSPECTRAL IMAGE
ANALYSIS**

By

Jackeline I. Jerez Rozo

A dissertation submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
Applied Chemistry

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS
2013

Rodolfo J. Romañach, PhD
President, Graduate Committee

Date

Jessica Torres, PhD
Member, Graduate Committee

Date

Aldo Acevedo, PhD
Member, Graduate Committee

Date

Oscar Perales, PhD
Member, Graduate Committee

Date

Paul Sundaram, PhD
Representative, Graduate Studies

Date

René Vieta, PhD
Chairperson, Chemistry Department

Date

ABSTRACT

This dissertation is focused on the development of analytical methods to determine drug distribution throughout polymeric films using hyperspectral image analysis. The NIR-CI and Raman mapping techniques have been used to analyze the distribution and quantification of drug in a novel pharmaceutical formulation. This pharmaceutical formulation was developed with the goal of maintaining the drug with a specific particle size in a non agglomerated form and to satisfy two commonly encountered pharmaceutical needs: enhanced dissolution rate of poorly soluble drugs and the content uniformity of drugs administered in low doses. In these films, the active ingredient is a poorly soluble drug, which is dispersed in the polymer and additionally; surfactant and lubricant are added. The water poorly soluble drugs have to be uniformly distributed in a film formulation to have an acceptable drug content uniformity. This is the reason why drug distribution is an important factor in these type of pharmaceutical formulations. In order to obtain the desired distribution it is necessary to find those areas of the film where the drug is agglomerated, and use this information to improve the process.

Chapter 3 presents the results of the first methodology developed. The first methodology was based on determining the pixels of maximum intensity value at 2080 nm. At this wavelength, pixels with positive values correspond to the drug. A film with large agglomerates was used to develop this methodology. Results show that the drug should be highly agglomerated to find pixels of pure drug. Therefore, the methodology was developed to evaluate the distribution of those pixels that are composed mainly of drug but also containing others components. Agglomerations of these pixels are called drug rich areas in this dissertation. The procedure followed to evaluate these drug rich areas is: to Identify of drug rich areas, observe the distribution of drug rich areas throughout the film surface, and do a visual comparison between different areas or films. This methodology was applied in the images analysis of the chapter 5-7.

Chapter 4 compares the previously developed methodology with a new approach using Multivariate Image Analysis (MIA). The previously developed methodology is based on observing the distribution of drug rich areas at 2080 nm. The new methodology is based on the Bharati and MacGregor approach for incorporating the textural information of the image. MIA was used to explore the spectral and spatial relationship between the API and the different excipients. Results obtained with the two methodologies are comparable in terms of drug distribution.

Chapters 5 -7 show the results obtained by applying the methodologies developed. These results are based on analysis of high intensity values at 2080 nm and score images. Nanosuspensions produced from wet stirred media milling (WSMM) were used to prepare these polymeric films. In these chapters the effect of stabilizers on controlling growth and agglomeration of the drug, the influence of the drug molecule on the distribution of drug rich areas and the impact of the drying process in the agglomeration of these drug rich areas, were evaluated.

Chapter 8 shows the results obtained during the internship. The objective of this internship was to develop skills in solving challenging problems in both fundamental and applied research. This internship was conducted in a Pharmaceutical Chemical Plant during a period of 8 months. During this time, two NIR methods were developed in an effort by the Pharmaceutical Chemical Plant to implement techniques of analysis faster and cost effective. A feasibility study using Raman Spectroscopy for the ID of raw materials also was completed.

Chapter 9 summarizes the scientific contribution of this dissertation.

Copyright ©2013

By

Jackeline I. Jerez Rozo

Dedication

To God

My family

ACKNOWLEDGEMENTS

During the development of my doctoral studies at the University of Puerto Rico at Mayagüez campus, several persons and institutions collaborated directly and indirectly with my research. I want to express my sincere gratitude to my advisor, Rodolfo J. Romañach Ph. D. for his continuous support during my Ph.D study and research, and for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me all the time in research and in the writing of this dissertation. I also have to thank the members of my Ph.D. committee, Professors Aldo Acevedo, Jessica Torres and Oscar Perales for their helpful career advice and suggestions in general.

I am most grateful to Luigi Guariniello Ph. D., and Edgardo Hernandez Ph. D., for offering me the internship opportunity at Lilly and Lucas Sievens-Figueroa Ph.D., for his great contribution to this thesis.

Also I thank to Dianelys, Jenny, Yleana, Miriam and Tatiana for helping me get through the difficult times, and for all the emotional support, entertainment, and caring they provided.

I would also like to thank my family for the support they provided me through my entire life.

I recognize that this research would not have been possible without the financial assistance of the Engineering Research Center for Structure Organic Particulate System (C-SOPS).

Contents

ABSTRACT	ii
ACKNOWLEDGEMENTS	vi
List of original publications.....	i
List of Tables.....	ii
List of Figures.....	iv
List of Abbreviations and Symbols	vii
CHAPTER 1. Overview	1
1.1 Motivation	1
1.2 Scope	2
1.3 Hypothesis.....	2
1.4 Goals of the dissertation	2
1.5 Dissertation overview.....	3
Chapter 2 - Background.....	6
2.1 Summary.....	6
2.2 Near Infrared Spectroscopy.....	7
2.3 NIR-CI Literature Review	7
2.4 Principles of NIR-CI	9
2.5 NIR-CI Instrumentation.....	11
2.6 Spectral correction	12
2.7 Chemometrics for the analysis of Hyperspectral Images.....	13
2.7.1 Data Pretreatments.....	13
2.7.2 Reduction of variables by principal component analysis (PCA)	14
2.7.3 Multivariate calibration for quantitative analysis	14
2.7.4 PLS-DA classification.....	15
2.8 Concatenation	16
2.9 Histogram	17
CHAPTER 3. Near-Infrared Chemical Imaging of Pharmaceutical Thin Films: Maximum Intensity Values Approach.	19
3.1 Summary.....	19
3.2 Materials and Instrumentation	20

3.2.1 NIR Chemical Imaging.....	21
3.2.2 Data treatment for NIR spectra.....	23
3.2.3 Raman Mapping	23
3.3 Methodology	24
3.4 Results and discussion.....	24
3.5 Conclusions.....	34
Chapter 4. MIA and NIR Chemical Imaging for pharmaceutical product characterization.....	35
4.1 Analysis of hyperspectral images of polymeric films; comparison between Multivariate Image Analysis (MIA) and Partial Least Square Discriminant Analysis (PLS-DA).....	35
4.2 Reagents, Instrumentation and Data Pretreatment	36
4.3 Methodology	36
4.3.1 Multivariate Image Analysis (MIA)	37
4.4 Results and discussion.....	38
4.4.1 Films with larger GF agglomerated	38
4.4.2 Films with a better drug distribution.....	42
4.5 Conclusions.....	50
CHAPTER 5. Hyperspectral image analysis to evaluate the effects of stabilizers on particle dispersion from polymer films.....	51
5.1 Summary.....	51
5.2 Materials and Methods	52
5.2.1 Polymeric films	52
5.2.2 NIR chemical imaging	53
5.3 Results and Discussion.....	54
5.4 Conclusions.....	62
CHAPTER 6. Hyperspectral image analysis of polymeric films containing different BCS Class II drugs	63
6.1 Summary.....	63
6.2 Samples	63
6.3 Instrumentation.....	64
6.4 Data treatment.....	65
6.5 Results	66
6.6 Conclusions.....	73
CHAPTER 7. Hyperspectral Image Analysis to evaluate the effect of drying on drug distribution.	74

7.1 Summary.....	74
7.2 Samples	74
7.3 Instrumentation.....	76
7.4 Data treatment.....	77
7.5 Results	77
7.6 Conclusion	85
Chapter 8. Development of Raman and Near Infrared (NIR) Methods in a pharmaceutical chemical plant.....	86
8.1 Objective.....	86
8.2 Summary.....	86
8.3 Perform a preliminary feasibility study to use portable Raman technology for the ID testing of raw and starting material used at manufacturing site.....	86
8.3.1 Problem	86
8.3.2 Background.....	87
8.3.3 Methodology	87
8.3.4 Reagents used in this report.....	88
8.3.5 Results	89
8.3.6 Raman Spectra.....	91
8.3.7 NIR Spectra	98
8.3.8 Future Studies and Tasks.....	100
8.4 Perform a feasibility study for use NIR to monitor process reaction completion.....	101
8.4.1 Problem	101
8.4.2 Background.....	101
8.4.3 Methodology	101
8.4.4 Instrumentation.....	102
8.4.5 Results	102
8.4.6 Next Steps.....	107
8.5 Perform a feasibility study for implementation of NIR for solvent recovery process.....	110
8.5.1 Problem	110
8.5.2 Methodology	110
8.5.3 Instrumentation.....	110
8.5.4 Results	110

8.5.5 Next Steps.....	114
9. Research Contribution.....	115
9.1 Future Work	116
10. REFERENCES	117

List of original publications

This dissertation is based on the following original papers.

- I. Beck, C.; Sievens-Figueroa, L.; Gärtner, K.; Jerez-Rozo, J. I.; Romañach, R. J.; Bilgili, E.; Davé, R. N., Effects of stabilizers on particle redispersion and dissolution from polymer strip films containing liquid antisolvent precipitated GF particles. *Powder Technology* **2013**, 236, 37-51. doi:10.1016/j.powtec.2012.05.047.
- II. Prats-Montalbán, J. M.; Jerez-Rozo, J. I.; Romañach, R. J.; Ferrer, A., MIA and NIR Chemical Imaging for pharmaceutical product characterization. *Chemometrics and Intelligent Laboratory Systems* **2012**,117, 240-249. doi:10.1016/j.chemolab.2012.04.002.
- III. Sievens-Figueroa, L.; Bhakay, A.; Jerez-Rozo, J. I.; Pandya, N.; Romanach, R. J.; Michniak-Kohn, B.; Iqbal, Z.; Bilgili, E.; Dave, R. N., Preparation and characterization of hydroxypropyl methyl cellulose films containing stable BCS Class II drug nanoparticles for pharmaceutical applications. *International Journal of Pharmaceutical* **2012**,423 (2), 496-508. doi:10.1016/j.ijpharm.2011.12.001.
- IV. Jerez-Rozo, J. I.; Zarow, A.; Zhou, B.; Pinal, R.; Iqbal, Z.; Romanach, R. J., Complementary near-infrared and Raman chemical imaging of pharmaceutical thin films. *Journal of pharmaceutical sciences* **2011**,100 (11), 4888-95. doi: 10.1002/jps.22653.
- V. Court, K.A.; Jerez-Rozo, J.I.; Romañach, R. J., Torres-Lugo, M.; Particle Encapsulation in Crosslinked Hydrogel Networks: Particle Distribution Optimization. *Journal of Materials Science and Engineering B*; **2012**; 10, 539-550.

List of Tables

Table 3.1 Composition of the films of other batch.....	21
Table 3.2 Domain Size Statistic results of API Based on Binary Images Generated from Intensity Value at 2080nm	32
Table 4.1 Statistical Analysis of PLS Score values derived from the analysis of the two polymeric films analyzed.....	48
Table 5.1. Composition of the three films analyzed.....	53
Table 5.2 Intensity value at 2080 nm of the different films analyzed.	55
Table 5.3 Domain size statistic results of API based on binary images generated.....	59
Table 6.1 Nominal concentrations of the polymeric films.....	64
Table 6.2 Distribution of APIs in thin films according to drug score image using NIR-CI with a 10µm/pixel objective.	70
Table 6.3 Distribution of APIs in thin films according to the drug score images using NIR-CI with a 131 µm/pixel objective.	71
Table 7.1 Factors included in the D-Optimal design and Nominal concentration of the polymeric films.	75
Table 7.2 Composition A: Mixed in (2:1) ratio (12% HPMC E15LV+5% Glycerin: API nanosuspension).....	75
Table 7.3 Composition B: Mixed in (2:1) ratio (15% HPMC E15LV+5% Glycerin: API nanosuspension).....	76
Table 7.4 Nominal and predicted values of the different films analyzed.	81
Table 7 5 Domain Size Statistic results of API based on binary images generated from scores images.....	81
Table 8.1 Raw materials of the pharmaceutical chemical plant.....	90
Table 8.2 Correlation between Raman spectra.	99
Table 8.3 Correlation between NIR spectra.....	99
Table 8.4 MMAA Model.	109
Table 8.5 Calibration and Validation values* obtained with the NIR model using plant samples.	109

Table 8.6 EtOAcModel.	112
Table 8.7 Calibration and Validation values* obtained with the NIR model using solvent recovery samples.....	113

List of Figures

Figure 2.1 Three-dimensional hyperspectral data cube (x×y×z).	10
Figure 2.2 Two methods of acquiring a hyperspectral imaging data cube: Point mapping and Global imaging- Focal plane array (FPA).	11
Figure 2.3 Classifying images: PLS-DA approach	15
Figure 2.4 Concatenation of NIR-CI images	16
Figure 2.5 Histogram plot.....	17
Figure 3.1 Chemical structure of pure components.....	21
Figure 3.2 Films prepared with larger drug agglomerates. a. Image obtained with a microscope coupled to an NIR-CI using a magnification of 10 µm per pixel. b. Image obtained with the SEM	25
Figure 3.3 Spectra of HPMC E15LV and GF. a. Raw spectra of pure components. b. Second derivative spectra of pure components.	25
Figure 3.4 Image and histograms for the different components within film at 2080 nm.	27
Figure 3.5 Images obtained with a microscope coupled to an NIR-CI using a magnification of 10 µm per pixels Image at 2080 nm. a. Average spectra of HPMC E15LV and GF. b. Spectra of different points of the gel strip formed by GF and HPMC E15LV.....	28
Figure 3.6 Binary Image of films with larger agglomerates of drug obtained with different threshold values.....	30
Figure 3.7 GF rich areas observed at 2080 nm.....	31
Figure 3.8 Histograms providing quantitative descriptions of the API and HPMC contents for Films 1, 2, 3, and 4. These histograms are based on Raman scores images.	33
Figure 4.1 Procedure to obtain the data structure used to perform the colour-textural MIA. 3×3 neighborhood window.....	38
Figure 4.2 Scores images from PCA. Scatter plot displaying the distribution of the first and second Principal Component.	40
Figure 4.3 Score image of the first six Principal Components Analyses.	41
Figure 4.4 Pixels used to create the reference library.	43

Figure 4.5 a. Red-green (RG) color-coded NIR images wherein red represents the distribution of the API, GF and green that of the excipient, HPMC E15LV. b. Image obtained using the microscope coupled to an NIR-CI.	44
Figure 4.6 Scores images from PCA. Scatter plot displaying the distribution of the first and second Principal Components.	45
Figure 4.7 Score image of the first six Principal Components Analyses.	46
Figure 4.8 Score image of PLS-DA. Concatenated images show the distribution of HPMC E15LV and GF through of film surface.	47
Figure 4.9 Red-green (RG) color-coded NIR images wherein red represents the distribution of the API, GF and green that of the excipient, HPMC E15LV.....	49
Figure 5.1 Chemical structure of the PF-127 and SDS	52
Figure 5.2 Image and histograms for the HPMC E15LV stabilized film at 2080 nm. This film does not contain stabilizers.	55
Figure 5.3 Image and histograms for the HPMC E15LV stabilized film at 2080 nm. This film does not contain stabilizers.	56
Figure 5.4 Image and histograms for the HPMC E15LV- PF-127 stabilized film at 2080 nm.....	56
Figure 5.5 GF drug rich areas observed at 2080 nm.....	58
Table 5.3 Domain size statistic results of API based on binary images generated.....	59
Figure 5.6 Adsorption model of the two surfactants used. Adapted from Changhoon Chai, “Hierarchical structures of micro and nano API particulates on gelling polymeric matrix”, ERC-SOP.....	61
Figure 6.1 Structural formula of different API and average spectra of pure components.	67
Figure 6.2 API binary images using 10 $\mu\text{m}/\text{pixel}$ objective. Domain threshold is set to include only those pixels of highest abundance of API according to the maximum value of each colorbar.	68
Figure 6.3 API Score image using 131 $\mu\text{m}/\text{pixel}$ objective.	69
Figure 6.4 FT-IR spectra of NPX and NPX film.....	72
Figure 7.1 PLS predicted API images are displayed on the same scale color using the 131 $\mu\text{m}/\text{pixel}$ objective. The images show how the variation in air flow (m/s), viscosity	

(CPS), and temperature ($^{\circ}$ C) affects the API distribution. Pixel values used to create binary images are shown on each colorbar. Binary images were created for pixels where the intensity is greater than the mean value + 2 standard deviations..... 79

Figure 7.2 Binary image generated from the scores images using the 131 μ m/pixel objective. A threshold was established such that only those pixels with values beyond + 2 standard deviations of the mean are shown. 80

Figure 8.1 Structures of the eight reagents used in this report..... 88

Figure 8.2 Structures and spectra of sodium bicarbonate, sodium carbonate, sodium carbonate monohydrate and sodium acetate..... 91

Figure 8.3 Structures and spectra of Methyl Phenyl sulfoxide and DimethylSulfoxide 92

Figure 8.4 Structures and spectra of Ethyl acetate and Methyl acetate 93

Figure 8.5 Structures and spectra of 1-propanol and Isopropanol 94

Figure 8.6 Spectrum of KOH..... 95

Figure 8.7 Structures and spectra of 2-methyltetrahydrofuran and Tetrahydrofuran... 96

Figure 8.8 Structures and spectra of Toluene, Bromobenzene, Phenol and Benzene 97

Figure 8.9 NIR Spectra of Ethyl acetate (EtOAc), DimethylSulfoxide (DMSO), Isopropanol, Tetrahydrofuran (THF) and Toluene..... 98

Figure 8.10 Score plot using Raman Spectra. 100

Figure 8.11 Set up of the arylation reaction. 102

Figure 8.12 Interference of KOH in excess. 103

Figure 8.13 Spectra of an arylation reaction. 103

Figure 8.14 Spectra of an arylation reaction. Subtle differences are observed in this spectral range. These differences are increased with the pretreatment..... 104

Figure 8.15 PCA of the laboratory and plant samples..... 105

Figure 8.16 PCA of the laboratory samples. 106

Figure 8.17 Spectra andPCA of the plant samples. 108

Figure 8.18 Selected spectral range. 111

List of Abbreviations and Symbols

API	Active Pharmaceutical ingredient
APIs	Active Pharmaceutical ingredients
BCS	Biopharmaceutical Classification System
NIR-CI	Near Infrared Chemical Imaging
ERC	Engineering Research Center
HPMC E15LV	Hydroxy Propyl Methyl Cellulose
PLS-DA	Partial Least Square Discriminant Analysis
MCR	Multivariate Curve Resolution
CLS	Classical Least Square
CDMs	Chemical Distribution maps
MIA	Multivariate Image Analysis
LAS	Liquid Antisolvent
NPX	Naproxen
GF	Griseofulvin
FNB	Fenofibrate
WSMM	Wet stirred media milling
HPLC	High Performance Liquid Chromatography
MS	Mass Spectroscopy
NIR	Near Infrared
GC	Gas Chromatography

CHAPTER 1. Overview

1.1 Motivation

Analytical techniques are being developed to understand pharmaceutical processes. Process understanding allows the identification of parameters that can affect the quality of the final product. High quality products require quality by design, and production processes that can be monitored with faster and more efficient techniques.^{1,2} Analytical techniques capable of monitoring processes facilitate achieving operational excellence, which is the goal of the pharmaceutical industry. Process understanding has become an important aim for the pharmaceutical industry.³

The ability to visualize and assess the compositional heterogeneity and structure of the end products is extremely important for both the development and manufacture of solid dosage forms in the pharmaceutical industry.⁴ A basic problem in pharmaceutical manufacturing is that even a relatively simple formulation may produce widely varying therapeutic performance depending on the distribution of ingredients in the final matrix.^{5,6}

The component distribution may impact bioavailability, dissolution, or other product performance attributes. The optimal determination of the distribution of the drug and excipients affect blend homogeneity, content uniformity, and may also affect dissolution. These issues are related, not only to the manufacturing process, but also to the solubility of the Active Pharmaceutical Ingredient (API). The solubility of drug molecules is an important factor to take into account during the development and design of new drug products. According to the Biopharmaceutical Classification System (BCS), Class II drugs are characterized by their poor solubility and high permeation in the human body.⁷ Most recently, the buccal route is getting more and more awareness for the application of API. The application via the buccal route offers different advantages: an easy application, no degradation of API by gastrointestinal fluids, bypassing the first hepatic metabolism and potentially improved bioavailability in

order to ensure rapid invasion and fast onset. Many advantages of this route have been recently recognized and various dosage forms are under development.⁸ It is estimated that over 40% of all possible new active drug candidates have very low solubility.^{7,9-11} A number of techniques were recently developed for online measurement of specific film properties, but those are mostly based on single-point probes (or probe arrays). Therefore, they have a limited spatial coverage of the film surface; spatial measurements are key in identifying local composition gradients and flow patterns that may compromise global quality of the films.^{12,13}

1.2 Scope

The scope of this dissertation is to assess the compositional uniformity of organic composites using hyperspectral imaging, with an emphasis in films formed by excipients and Active Pharmaceutical Ingredient (API) dispersed in polymeric matrices.

1.3 Hypothesis

One possible approach for improving the solubility, and hence dissolution, of poorly soluble drugs is to disperse them in a polymeric film, reducing surface tension, and thus, preventing them from aggregating.^{7,9-11}

A critical evaluation of the film formulation requires information on the spatial distribution of the active ingredient. Near Infrared Chemical Imaging (NIR-CI) can be used to identify the presence of drug-agglomerated clusters. The distribution of these drug agglomerates may be related with the results of the dissolution test.

1.4 Goals of the dissertation

The objective of this dissertation is the development of analytical methods to determine the distribution of drug throughout the surface of polymeric films using chemical imaging techniques. This research was performed within the Engineering Research Center for Structured Organic Particulate Systems (ERC-CSOPS) a center focused on

the scientific-based development of structured organic particle-based products and their manufacturing processes.

Polymeric thin films are a novel pharmaceutical formulation; therefore new analytical methods are needed to characterize these new drug delivery forms. The approach used in this work was to relate the spatial distribution of the drug rich areas (drug clusters) within these films with the results of the dissolution test. It has been shown that by increasing/controlling specific surface area, the dissolution rate may be significantly improved. A commonly used approach to poor water solubility is drug particle size reduction, thus increasing its surface area, through either top-down or bottom-up techniques. The dissolution tests were performed at NJIT and Rutgers. The best results of the dissolution test were obtained in the films where the drug was more uniformly distributed.

The research also required defining a drug rich area. This term was difficult to define because the poorly soluble drug is found on every area of the film; additionally, these drug rich areas depend on the particle size and polymer-API miscibility. In this dissertation, drug rich areas are those areas with the greatest concentration of API. These drug rich areas are found by analyzing the histograms of each score or intensity images. The number of standard deviations selected to establish the threshold value was determined as a function of the film and the problem to be resolved; based on the results obtained, two standard deviations from the mean was a good choice.

The procedure followed to analyze the different hyperspectral images were: Identify the drug rich areas, determine the distribution of these drug rich areas throughout the film surface and compare them.

1.5 Dissertation overview

Chapter 2 presents the introduction and background describing the basic principles of NIR-CI, instrumentation, chemometrics and different tools used for image processing.

Chapter 3 presents the development of the methodology of the first approach used to analyze the NIR hyperspectral images. The first approach used was to observe the maximum intensity values of the second derivative at 2080 nm on films that contain only GF and Hydroxy Propyl Methyl Cellulose (HPMC E15LV). This chapter introduces the term drug rich areas.¹⁰

Chapter 4 presents the second approach used in this dissertation; to develop a methodology where partial least squares discriminant analysis (PLS-DA) method determined drug or polymer distribution throughout the film. These images also were analyzed using a three step methodology based on the use of chemical oriented models using Multivariate Curve Resolution (MCR) and Classical Least Square (CLS) for extracting the chemical distribution maps (CDMs) from hyperspectral images, and afterwards performing multivariate image analysis (MIA) on the CDMs. This approach transforms the spectral information into separated concentration maps; and MIA improves the interpretation when separating the joint variability into orthogonal information maps (the *Score Images*) with a clear physical interpretation (drug-HPMC blend and separation zones). The results obtained with both approaches are compared in terms of drug distribution.⁹

Chapters 5 - 7 show the results obtained by applying the new methods developed. One method consists in observing the distribution of pixels of maximum intensity value at 2080 nm. The other method is based on observing the distribution of the pixel with the highest drug abundance based on the analysis of scores images obtained from PLS-DA. Chapter 5 presents the analysis of polymeric films produced via liquid antisolvent (LAS) method. These films contain all ingredients of the pharmaceutical formulation. The results of the analysis of the hyperspectral images evidence the dual role of stabilizers on controlling growth and agglomeration of griseofulvin (GF) particles.¹¹ Chapter 6 presents the analysis of the distribution of three different API through the film surface. The three drugs analyzed were: naproxen (NPX), GF and

fenofibrate (FNB). The particle sizes of the different API were reduced using wet stirred media milling (WSMM); the GF, FNB and NPX suspensions exhibited a median particle size of 163 nm, 201 nm, and 144 nm, respectively. Results obtained confirm the relationship between the distribution of the drug rich areas and the chemical structure of the drug molecule.⁷ Chapter 7 presents the impact of drying method over the distribution of API. The films were dried using a convective heating unit; this unit allows the variation of the temperature and air velocity. The analysis of the distribution of the drug rich areas allows finding the optimal values of these variables. The drug used in this study was GF.¹⁴ In summary, hyperspectral images analysis provided significant information on the distribution of drug particles in films prepared under different conditions. The chemical imaging information contributed to the optimization of the film formulation.

Chapter 8 presents the research performed during the internship in an active pharmaceutical ingredient (API) manufacturing site. NIR calibration models were developed to determine the amounts of toluene and ethyl acetate in recovered toluene and the amount of *the component X* in reaction completion. A feasibility study was conducted using Raman Spectroscopy as an alternate technique for the ID analysis of raw material.

Chapter 9 summarizes the research contributions and future work.

Chapter 2 - Background

2.1 Summary

Traditional quality control methods such as High Performance Liquid Chromatography (HPLC) and Mass Spectroscopy (MS) are time consuming, destructive, expensive, require lengthy sample preparation and do not provide information about the distribution of components within a sample. Due to the destructive and time consuming nature of these methods, only small samples of drugs may be tested from given production batches. Vibrational spectroscopic techniques such as Near Infrared (NIR) and Raman spectroscopy have emerged as alternative techniques due to the speed of analysis and because sample preparation is not required.³

Hyperspectral imaging is being introduced in pharmaceutical research laboratories to overcome the above drawbacks of classical techniques and to increase the knowledge and understanding of many commercially important materials.¹⁵ The additional information provided by the spatial perspective of NIR-CI offers access to greater understanding and therefore control of the manufacturing process of complex composite materials and products. In the pharmaceutical industry, solid dosage forms are used to confer physical stability to drug formulations and are considered more attractive due to their convenience and consumer preference aspects. Solid dosage forms are not composed of a single material; rather they are carefully designed mixtures to form a pharmaceutical formulation. A typical formulation may include one or more active pharmaceutical ingredients (APIs), fillers, binders, disintegrants, lubricants, and other materials. Each of these materials is chosen to provide desirable characteristics for manufacturing, storage, handling, and eventual release of the therapeutic agent.¹⁶ Throughout formulation development and the scale-up process, the primary goal is to produce a formulation and a manufacturing procedure that is robust and consistent.¹⁷⁻¹⁹ Modern formulations are further increasing the demands on product uniformity. Products containing more potent APIs may be formulated at quantities ranging from a few to even less than one milligram per dose, yet the finished

product must still be large enough for convenient handling. Therefore, the therapeutic agent may represent significantly less than 1% (w/w) of the bulk form, and maintaining content uniformity is absolutely crucial.^{5,6} Hyperspectral imaging techniques may be used to understand content uniformity and drug distribution within these formulations.

2.2 Near Infrared Spectroscopy

Near-infrared (NIR) spectroscopy is a very rugged and flexible technique that can be generally adopted for use in a wide range of chemical analyses in many different research and industrial applications.^{5,20-22} NIR spectroscopy has a number of advantages that have contributed to its wide adoption in process analytical applications. The principal advantage is the possibility of acquiring spectra without sample preparation, simultaneously obtaining, chemical and physical information; the dual dependence of the analytical signal on the physical and chemical nature of the sample facilitates both its identification and the determination of physical and chemical parameters.^{20,23,24} NIR spectral range extends from 750 to 2500 nm, where most organic compounds absorb radiation with CH-, OH-, and NH- vibrations.²⁵⁻²⁸ NIR absorption bands are typically broad, overlapping and 10–100 times weaker than their corresponding fundamental mid-IR absorption bands. Absorption in this spectral region is due to overtone and combination bands of molecular vibrations that modulate the dipole moment of the molecule.

2.3 NIR-CI Literature Review

Polymers are excellent materials for NIR analysis, since most polymer residues are rich in -CH, -NH, and -OH intramolecular bonds. The critical performance parameters of polymers, such as viscosity, yield strength, or glass transition, can be correlated to the long-range bonding structure of the material (e.g., chain length, cross-linking, crystallinity), residual monomer concentration, or the degree of substitution, which are readily detected using NIR spectroscopy.²⁹ Polymeric materials, especially cellulose

derivatives, have served an important role in manufacture as either fillers or coating agents. Hydrophilic polymers, such as poly (ethylene glycol) (PEG), poly (lactic acid)(PLA), or poly (DL-lactic-co-glycolide) (PLGA), are increasingly utilized for sustained-release implant dosage forms.²⁶ This section summarizes a number of previous NIR and NIR-CI methods for polymeric materials.

Svensson et al.³⁰ demonstrated the combination of NIR spectroscopy and chemometrics methods to classify 11 cellulose derivatives among more than 400 batches of pharmaceutical excipients with a high degree of accuracy. Gustaffson et al.³¹ used NIR and FT-IR spectroscopy and chemometrics to accurately predict tablet performance quality parameters from the spectra of raw hydroxyl propyl methyl cellulose (HPMC E15LV). This study analyzed 12 grades of HPMC E15LV in terms of methoxy- and hydroxylpropyl concentrations, total degree of substitution, methoxy/hydroxypropyl ratio, apparent density and specific surface area.

Furukawa et al.³² studied four kinds of PHB/PLLA blends with the PLLA content ranging from 20 to 80 wt% using NIR-CI to elucidate the blend quality of PHB/PLLA blends. This work discusses the possible interactions that exist between the excipients and the API. The domains for each blend component were dispersed as smaller crystallized or amorphous particles. The interfacial energy between PHB and PLLA may be so small that they make extremely fine dispersions within each other. This small difference in interfacial energy produces PHB/PLLA blends with a high degree of homogeneity.

Polymer hydrogels are increasingly being considered for sustained release of injectable biotechnological therapeutics and oral delivery of gastric labile protein or peptide therapeutics. Early work in this area by Nerella and Drennen³³ utilized dispersive NIR spectroscopy with a novel controlled aperture to develop a method for depth-resolved determination of drug content in a polymer hydrogels. The researchers utilized the NIR method to solve the diffusion constants of salicylic acid through a hydrogel matrix, which served as a practical in vitro model for a transdermal drug

delivery system. In much later work, Blanco and Romero³⁴ demonstrated the use of NIR transfectance spectroscopy for the determination of dexketoprofen in a hydrogel.

One of the problems in the design of new delivery drug is the lack of analytical techniques to characterize these new products. Garsuch⁸ used Scanning Electron Microscopy (SEM) and Near Infrared chemical Imaging for the morphologic characterization of fast-dissolving buccal wafers, using caffeine as model drug because is well absorbed buccally.

Gendrin et al.³⁵ compared Partial Least Squares (PLS) with at least two response variables (PLS2) and the Classical Least Squares (CLS) algorithm in predicting API and excipient content by NIR hyperspectral imaging in simple binary mixtures of API and cellulose. This is a pharmaceutical product quantification strategy using hyperspectral imaging with and without calibration.

2.4 Principles of NIR-CI

In conventional NIR spectroscopy a bulk NIR spectrum is measured providing an average composition of the measured sample. Recent advances in spectroscopic detector technologies have enabled spectral as well as spatial information to be recorded simultaneously. NIR Chemical Imaging (NIR-CI) offers the opportunity to explore not only what chemical species are present at a micro-scale level but also provides spatial information on their distribution within a sample.⁴ A hyperspectral image consists of a three-dimensional data cube (Figure 2.1), where two dimensions are spatial coordinates (pixels) and the third contains spectral information (wavelengths). NIR-CI provides one spectrum per pixel and hence a vast amount of both spatial and spectral information (usually more than 80,000 spectra per sample) that can be acquired in a very short time (1–2 min).^{3,22,36,37} Each element within the cube contains the intensity-response information measured at that spatial and spectral index. The hypercube can be treated as a series of spatially resolved spectra (called pixels) or, alternatively, as a series of spectrally resolved images (called image planes or channels).⁵ The hyperspectral data cube can either be viewed as an array of spectra

(one for each pixel) from which images can be created or as a row of images from which spectra can be extracted. Selecting a single pixel (xy-coordinate) through the z-plane will show the spectrum recorded at this particular spatial location, which provides the spectral signature of chemical components present in that exact part of the sample. Selecting an image plane (xy-plane) at a specific wavenumber (z-value) will show the intensity values for all pixels at that wavenumber also called a single wavenumber image. The single wavenumber image contains spatial distribution information about a chemical component with a distinct spectral characteristic at that particular wavelength.

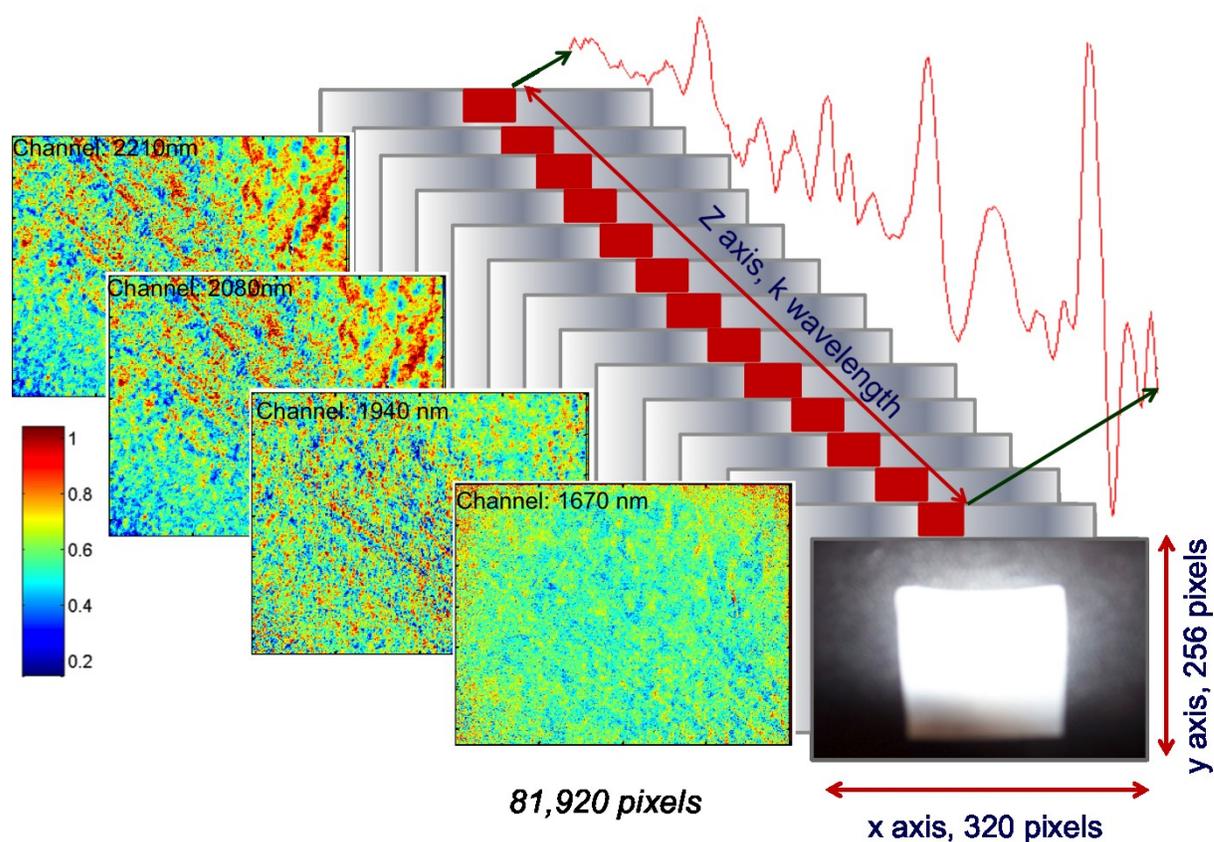


Figure 2.1 Three-dimensional hyperspectral data cube (x*y*z).

All images require some type of contrast to differentiate regions of interest in a field of view. The most common source of image contrast is variation in the intensity of diffuse reflectance.²⁸

The expectations on Chemical Imaging are focused on obtaining quantitative information about the content of each component and provide reliable information about the distribution of the component.³⁵ Spatial information provides a better understanding of the sample or as a problem-solving tool, and also statistical and quantitative parameters to evaluate the quality of the final product.³⁸ In recent years, NIR-CI has been used in different fields of the pharmaceutical industry.^{32,39,40} This diverse applicability is due to its capability of providing robust and reliable chemical and spatial information on the distribution of components in pharmaceutical solid dosage forms.⁴¹⁻⁴⁴

The acquisition of extensive data could be a limiting factor in applications of Process Analytical Technology (PAT). However, this extensive data can be reduced with the elimination of background pixels, image and spectral data compression, and data processing. Spectral image data do not have an intrinsically high information density because much of the data are correlated in all three dimensions.²⁶

2.5 NIR-CI Instrumentation

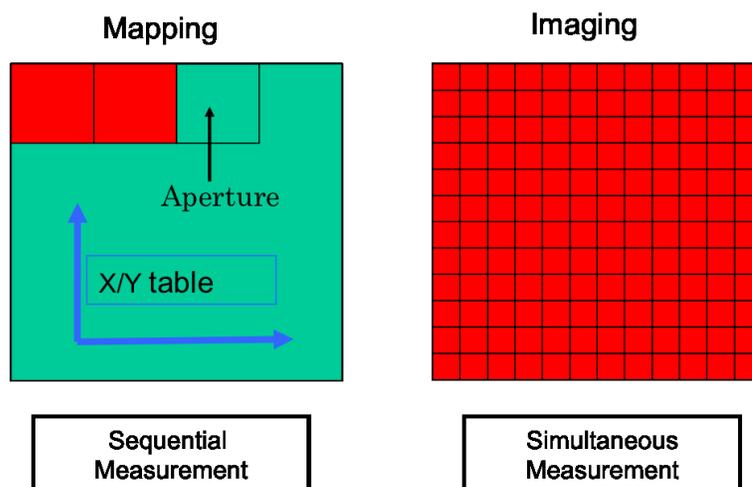


Figure 2.2 Two methods of acquiring a hyperspectral imaging data cube: Point mapping and Global imaging- Focal plane array (FPA).

In classical spectroscopy a spectrum includes the integrated spectral information of the

sample surface, which depends on the spot size generated by the beam of light. There are two main types of NIR-CI systems to obtain the raw data hyperspectral data cube: line mapping and global imaging.⁴¹ Only the global imaging system has been used in the experimental studies in this dissertation.

The mapping system consists of a classic spectrometer combined with a moving stage. In the mapping system regular spatial positions are defined above the sample surface. A spectrum is measured at one position, the sample moves to the next measurement point on the grid, a further spectrum is recorded, and so on for all positions in the area defining the image. In this way a grid of spectral information is created from the lines of spectra until spectra from all pixels in the defined sample area are obtained to constitute the hyperspectral data cube (Figure 2.2). Global imaging measures the NIR absorption intensity values in each pixel of the defined sample area at one particular wavelength at a time. The imaging technique uses two-dimensional focal plane arrays (FPA), where these optical detectors are composed of several thousand elements forming a matrix of pixels. FPA enable thousands of spectra to be acquired simultaneously. The number of pixels in an image is thus fixed for global imaging systems and the pixel size and sample areas are defined by the magnification optics (e.g. 10 μm /pixel objective leading to a 2.8 \times 3.2 mm image using a 320 \times 256 pixels FPA). The NIR absorption intensity values are measured in every pixel of the FPA at each individual wavelength, which is sequentially changed by the tunable filter.

The major advantage of global imaging is the faster acquisition time. The larger field of view is also an advantage, and the possibility of reducing the volume of data by choosing a specific wavelength. In global imaging, due to its smaller area, a single pixel element of a 2D detector integrates less signal than a single-element detector used in a mapping study, therefore, the signal to noise ratio (SNR) is lower in global imaging than in mapping when the same detector illumination is used.³⁸ During a global imaging experiment, other sources of noise due to the optics and the nonuniform pixel illumination must also be taken into account.

2.6 Spectral correction

In the absorbance-based spectral measurements, the intensity data represented in a raw (single beam) chemical image is a combination of the spectral response of both the instrument and the sample. In order to remove the instrument response component, it is necessary to ratio the data to a background reference. For diffuse reflectance measurements, the background is a separate data cube typically acquired from a uniform, high-reflectance standard such as a white ceramic. In addition to a background correction, dark current from the detector must be subtracted from both, the sample and background. The dark scans are collected by using a mirrored surface in the place of the sample. The correction to reflectance (R) is therefore:

$$R = (\text{Sample} - \text{Dark}) / (\text{Background} - \text{Dark}) \quad (2.1)$$

Further processing is also usually performed to transform the reflectance image cube to its $\log_{10}(1/R)$ form, which is the sample absorbance. After the correct pretreatment, the chemical image brightness is often related with the analyte concentration. The brightness in the images is useful for comparative as well as quantitative purposes.⁵

2.7 Chemometrics for the analysis of Hyperspectral Images

2.7.1 Data Pretreatments

Interference of spectral parameters, such as light scattering, path length variations and random noise, generate variations in the sample properties or instrumental effects. Mathematical correction, so-called data pre-processing is necessary to eliminate or standardize their impact on the spectra.

The most widely used pre-processing techniques in NIR-CI spectroscopy (in both reflectance and transmittance mode) can be divided into two categories: scatter-correction methods and spectral derivatives. Mathematical treatments used to compensate for scatter-induced baseline offsets include multiplicative scatter correction (MSC) and standard normal variate (SNV). Baseline shifts and intensity differences resulting from variable positioning or pathlength variations may be reduced or eliminated by normalization algorithms. Derivatives can be applied to improve the

resolution of overlapping bands, and eliminate baseline offsets. Since spectral noise is also amplified by the use of derivatives, derivatives are usually combined with Taylor or Savitzky-Golay smoothing algorithms.⁴⁵⁻⁴⁷

2.7.2 Reduction of variables by principal component analysis (PCA)

Principal component analysis (PCA) is widely used variable reduction method. PCA is a mathematical procedure that resolves the spectral data into orthogonal components whose linear combinations approximate the original data. The new variables, called principal components (PC), eigenvectors or factors, correspond to the largest eigenvalues of the covariance matrix, thus, accounting for the largest possible variance in the data set.⁴⁸ The first PC represents maximum variance amongst all linear combinations and each successive Principal Component accounts for as much of the remaining variability as possible. The first step is to do mean centering:

$$X_{\text{centered } i,j} = X_{i,j} - X_{.j}, \quad (2. 2)$$

where $x_{\text{centered } i,j}$ was the corrected absorbance (sample i and wavelength j), $x_{i,j}$ the raw data and $x_{.j}$ the mean absorbance at wavelength j . The new coordinates were computed as follows: $T=X_c.P$, where T was the scores matrix, P the loadings matrix and X_c the mean centered spectral matrix.⁴⁹

2.7.3 Multivariate calibration for quantitative analysis

The spectrum obtained at a pixel is represented by the concentration-weighted sum of the contributions of the pure spectra of the image constituents. These concentration weights vary from pixel to pixel, depending on the composition of the material, but the pure spectra of the constituents are the same along the whole image. Raw images can be decomposed into the product of a matrix of pure spectra by the concentration weights of these pure components in each pixel, $\mathbf{D} = \mathbf{CS}^T + \mathbf{E}$. The matrix \mathbf{S}^T contains the pure spectra of the image constituents and the matrix \mathbf{C} the concentration values of

these constituents in every pixel, \mathbf{E} contains the experimental error due to signal variation not due to the signal provided by the chemical compounds.⁵⁰

The Multivariate Curve Resolution (MCR) method allows for the resolution of individual contributions when the spectra of pure components are not available, under certain specific constraints that can be introduced in the model.⁵¹ Different algorithms can be used for obtaining \mathbf{C} and \mathbf{S} . When the original constituents of the mixture and their spectra are known a priori, the Classical Least Square (CLS) model may be used. CLS regression consists of projecting each sample spectrum forming an \mathbf{X} matrix on the pure spectra, hence obtaining the concentration directly related to the chemical compounds in it.^{52,53}

2.7.4 PLS-DA classification

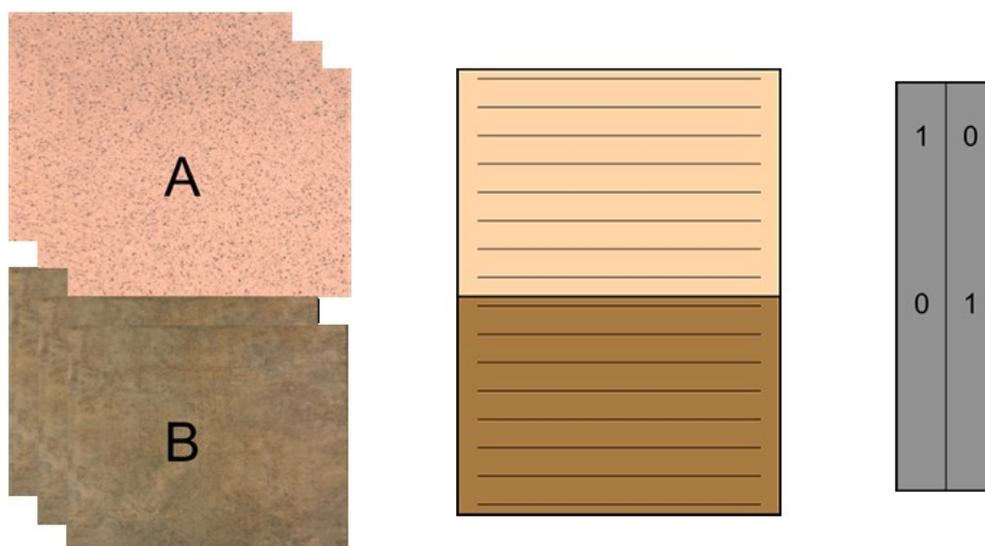


Figure 2.3 Classifying images: PLS-DA approach

Partial least squares discriminant analysis (PLS-DA) is a parametric and linear regression method that is used for classification. The purpose of PLS-DA is to achieve maximum discrimination among classes of objects on the basis of their X -variables by means of a model, providing a quantitative estimation of the discriminatory power of each descriptor.^{54,55}

In this method, \mathbf{X} contains the input information about the objects to be classified and \mathbf{Y} the class membership information. The regression coefficient matrix \mathbf{B} is calculated with the training set: $\mathbf{Y}_{\text{training}} = \mathbf{X}_{\text{training}} \cdot \mathbf{B}$, where $\mathbf{Y}_{\text{training}}$ is constructed with ones and zeros in each column as shown in Figure 2.3.⁴⁹ The calibration set is created with the spectra of pure components; PLS-DA1⁵⁶ is calculated using only one pure component and PLS-DA2 is calculated using all pure components. As a result of the prediction, scores images with values between 0 and 1 are obtained for each component included in the model. These maps will indicate the XY spatial distribution of each pure component in the mixture analyzed.

The score images analyzed in this dissertation are based on applying Partial Least Square Discriminant Analysis (PLS-DA). Images obtained by applying Multivariate Image analysis (MIA) and PLS-DA are compared in the chapter 4 of this dissertation.

2.8 Concatenation

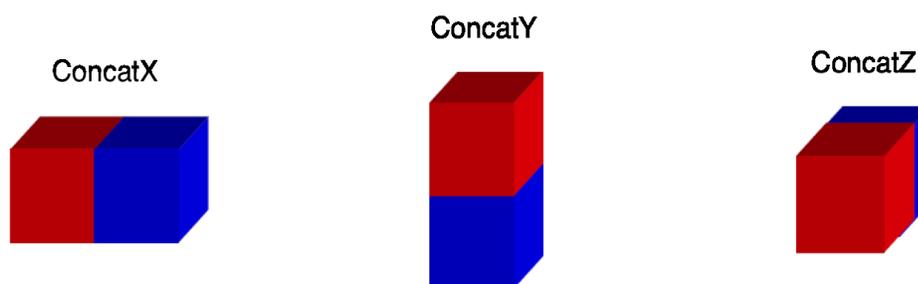


Figure 2.4 Concatenation of NIR-Cl images

Concatenation may be used to compare different areas of the same film. Images are acquired and concatenated and the distributions of the pixels of drug rich areas are then observed. Concatenation is useful to link two or more data sets so that they are processed in exactly the same way. An $A \times B$ image means an image of A pixels in x -axis and B pixels in y -axis. The data cubes ($A \times B \times x$ wavelengths) of the samples were grouped together in order to help image interpretation. Concatenation has been performed along the x and y axis to create a larger data cube (Figure 2.4).^{57,58}

2.9 Histogram

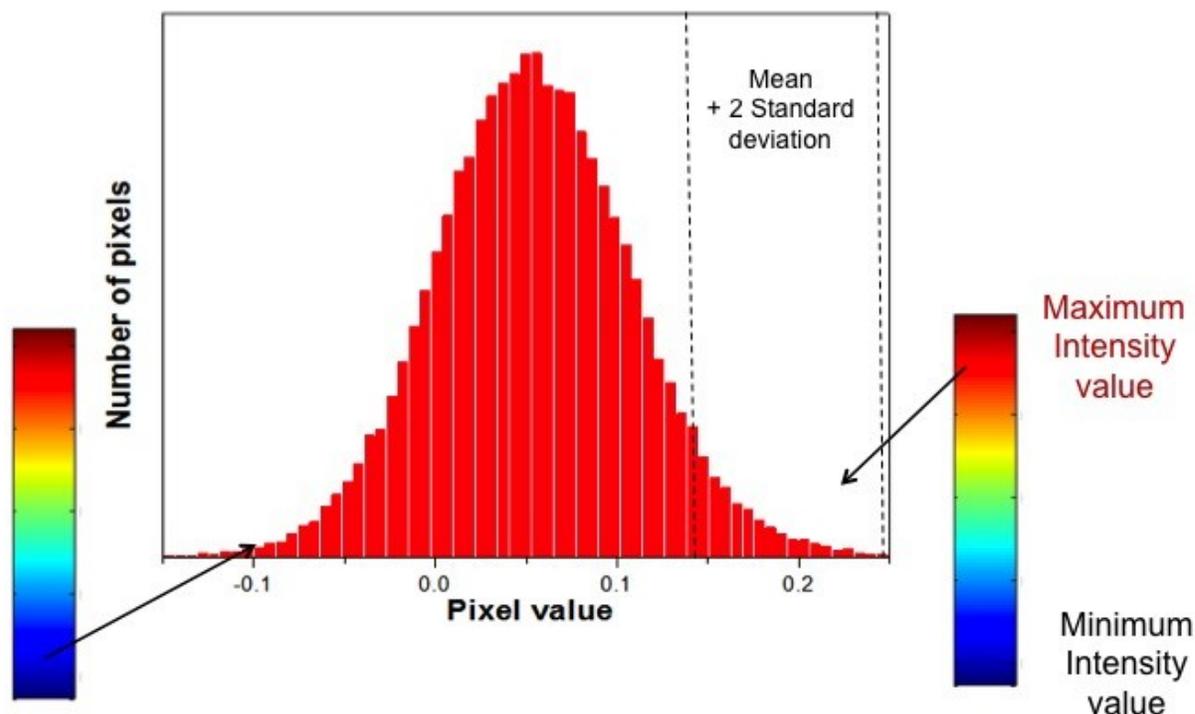


Figure 2.5 Histogram plot.

A histogram is a bar graph where different values of intensity of all the pixels can be seen as shown in Figure 2.5. The intensity values are on the x-axis and the numbers of pixels associated with each of these intensity values are on the y-axis. Figure 2.5 shows that the cursors truncate only the pixels of higher intensity; in this case, these pixels are directly related to the presence of API. Binary images may be obtained by assigning the pixels that are within the cursors a value of one and the remaining have a value of zero. This is an excellent way of isolating the pixels of interest, those that may be associated spectrally with API. One disadvantage of using histograms to generate binary images is that practically the same number of pixels are always observed.

Also the histograms can be used to evaluate the grade of distribution of the different components in the polymeric film. Three statistical parameters were computed to quantify their differences. First the mean value of the distribution is estimated:

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.3)$$

In equation 2.3 n represents the total number of pixels within the image and X_i the value of the pixel i . The variance will also be calculated:

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \mu)^2 \quad (2.4)$$

The kurtosis gives information about the shape of the histogram peak (equation 2.5).

$$k = \frac{\sum_{i=1}^n (X_i - \mu)^4}{\sigma^4} - 3 \quad (2.5)$$

Kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution. Positive kurtosis indicates a relatively peaked distribution. Negative kurtosis indicates a relatively flat distribution. Normal distributions produce a kurtosis statistic of about 0. The statistics of the intensity distributions from image to image can be compared to evaluate the degree of uniformity of the samples. Many other statistical characteristics can be calculated, including the center of mass, and number of domains of each component. The same image can be used to look at proximity of ingredients and reveal if two ingredients tend to co-agglomerate or disperse normally amongst the other components.

CHAPTER 3. Near-Infrared Chemical Imaging of Pharmaceutical Thin Films: Maximum Intensity Values Approach.

Published in *Journal of Pharmaceutical Science*, 2011, Volume 5, 4888-4895. Jerez - Rozo, J. I.; Zarow, A.; Zhou, B.; Pinal, R.; Iqbal, Z.; Romanach, R. J.

3.1 Summary

The drugs analyzed in this dissertation are poorly soluble. It is very important to determine the distribution of drug in formulations that use these type of drugs. The drug distribution in a polymeric film can impact many properties of the formulation such as the dissolution test and potency, among others. These polymeric films are a novel pharmaceutical formulation developed to achieve and maintain the drug dispersed, thus preventing agglomeration. It is necessary to develop analytical imaging techniques to assess the distribution of poorly soluble drugs in such novel pharmaceutical formulations. This chapter describes the first approach developed, characterized by the rapid collection of information from a hyperspectral image. Therefore, this approach can be used for real time analyses.

In this dissertation, a methodology was developed to evaluate drug distribution. This methodology was based on the analysis of hyperspectral images. In order to develop the technique, a film with highly agglomerated drug was analyzed. Results show that although the drug is highly agglomerated, pixels of pure drug are difficult to find. The term drug rich areas were introduced in the analysis, which allows analyzing films with more uniform drug distribution. This chapter contributes to hyperspectral images analysis with the introduction and definition of the term drug rich areas. Defining this term is difficult since small agglomerates will always be found through the sample because the active pharmaceutical ingredient (API) is poorly soluble. The scope of this work is to develop the methodology to find those areas with the highest griseofulvin (GF) concentration per pixel. One of the challenges was to establish the threshold value to find these areas of interest. It is very difficult to find pixels of pure drug in a film, and thus, if the drug is distributed uniformly, most of the pixels will have the

contribution of the polymer and GF. Chemical composition of each pixel will be affected by the fact that the drug is poorly soluble and therefore will be dispersed throughout the film surface and the scattering of the NIR radiation. The NIR radiation passes through the film, this radiation is reflected by the ceramic disk and again passes through the film. Part of this radiation will be absorbed by the film, whereas other photons will be reflected specularly and will not be detected. The radiation that arrives at the detector will contain mixed information from various locations in the polymeric film. The spectral information at each pixel will be influenced not only by the components of the measured pixel, but also by neighboring pixels.

NIR and Raman mapping techniques have been used to study the distribution of drug particles suspended in a polymeric film.¹⁰ A total of four film batches were prepared for this study using Hydroxyl Propyl Methyl Cellulose (HPMC) with griseofulvin as the active pharmaceutical ingredient. The NIR method analyzed a film area of 3 mm × 2.6 mm, whereas Raman mapping analyzed an area of 10 μm × 10 μm. Every sample was analyzed by the two methods.

3.2 Materials and Instrumentation

GF ((2*S*,6'*R*)- 7-chloro- 2',4,6-trimethoxy- 6'-methyl- 3*H*,4'*H*-spiro [1-benzofuran- 2,1'-cyclohex[2]ene]- 3,4'-dione)(Figure 3.1), is a poorly soluble drug used as an active pharmaceutical ingredient (API). A micronized drug, with a particle size below 10 μm was used in the study. Hydroxy propyl methyl cellulose (HPMC E15LV) was chosen as the film-forming excipient. These materials were purchased from Sigma-Aldrich Corporation (Milwaukee, WI). Blends of API and HPMC E15LV were prepared by first dispersing the two components together in hot water (80°C), followed by stirring for 12 hours and drying at room temperature.

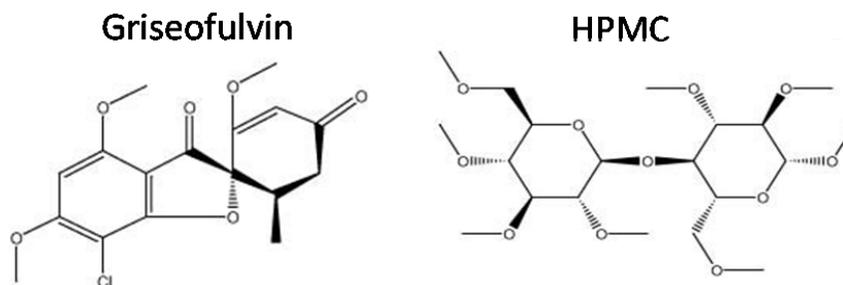


Figure 3.1 Chemical structure of pure components. Images were obtained from Fisher Scientific.

The film strips produced had a thickness of 100 μm . The four film batches used for this study were prepared at the New Jersey Institute of Technology (Table 3.1).

Table 3.1 Composition of the films of other batch.

	GF (% wt/wt)	HPMC E15LV (%wt/wt)
Film 1	57	43
Film 2	50	50
Film 3	44	56
Film 4	36.4	63.6

3.2.1 NIR Chemical Imaging

Near infrared hyperspectral images were acquired using the Malvern SyNIRgi Near Infrared Chemical Imaging System (Malvern, UK). The images were obtained in transmittance mode by placing the gel strips over a white ceramic disk with a diameter of 28 mm. NIR was chosen to be used since the NIR radiation easily penetrates these gel strips. The spectra of the HPMC E15LV and GF were obtained as pure powders in the diffuse reflectance mode. The spectra were collected with the system's focal plane array detector that has 320 X 256 pixel elements, with a total collection time of about 2 minutes. The optical magnification was 10 μm providing images of an area of approximately 3.2 x 2.6 mm. The spectra were obtained with 5 scans using a spectral

range of 1200 - 2400 nm, analyzing two different areas within the same gel strip. The data collected was analyzed using the ISys software (Version 5.0.0.14).

3.2.1.1 Spatial resolution issues

When the incident radiation reaches the sample surface, it may be immediately reflected (specular reflectance) or enter the sample. The photons that arrive at the detector will contain mixed information about the sample components at various depths and locations in the sample. The final result is a spectral pattern influenced not only by the components of the measured pixel, but also by neighboring pixels and by the depth of penetration.^{59,60} As the contrast in NIR chemical images is based on NIR spectral signatures, it therefore implies that the depth of penetration, will limit the highest useful lateral magnification; depending on wavelength, approximately the 70% of the photons coming from the sample in a diffuse reflectance measurement will have interacted with a spherical region with a diameter of approximately 60 - 90 μ m.⁶¹

The images of the polymeric films were acquired in transmittance mode, the radiation passes through the film. The hyperspectral image provides chemical information of the bulk and of the surface of the polymeric film, therefore, special care must be taken in the interpretation of the chemical image obtained. Several areas of the same polymeric films were analyzed using different optical magnifications to guarantee the results obtained.

The acquisition of high quality spectra require consideration of the intensity of the light, since, the illumination sources generate high levels of energy that induce strong heating of the samples. Therefore the polymeric films were placed under a transparent window to keep the samples flat and reduce any drying effects due to heat from the lamps. It is preferable to analyze the flat surface to minimize specular reflection of the light source owing to the roughness of the surfaces.⁴¹ Pyrex windows were also placed over the lamps to absorb any mid-infrared radiation as an additional precaution to reduce any heating effects. In imaging, it is very difficult to eliminate completely the variation in illumination, but this effect can be minimized. It is better to use lamps with

polarizers, and ensure that the lamps are illuminating the center of the film and preferably work the lamps at low intensities.

3.2.2 Data treatment for NIR spectra

NIR spectra were obtained in the reflectance mode. The logarithm, $\log_{10}(1/R)$, was first applied to the data cube to convert the spectra to absorbance units. Pixel correction was applied in order to remove areas of non-uniform illumination and to remove the effect of unresponsive pixels on the detector. A low-pass Fourier filter was applied to reduce spectral noise caused by the use of transmittance for a semi-transparent matrix. The spectra were normalized using the Standard Normal Variate and Savitzky-Golay second derivative (filter order 3, filter width 9).

3.2.3 Raman Mapping

Raman mapping spectra were collected using a Jobin Yvon/Horiba LabRAM micro-Raman spectrometer, with 632.81 nm He/Ne laser excitation and a 10x objective, producing a laser spot size of 3 μm . The Raman spectra were recorded in the 800-1650 cm^{-1} range. A typical chemical map consisted of 30x30 pixels, with a step size of 0.3 μm . These Raman images were acquired at New Jersey Institute of Technology.

3.2.3.1 Data treatment for Raman spectra

A spatial mask was initially applied in order to eliminate background pixels, followed by a low pass filter, baseline correction, standard normal variate (SNV) transformation and Savitzky-Golay second derivative (filter order 3, filter width 13). Spectra from film 3 were used to create a reference library based on partial least squares discriminant analysis PLS-DA. Spectra were obtained from pixel with maximum intensity value at 1620 cm^{-1} . At this wavelength, griseofulvin has a strong vibrational mode. Partial least squares discriminant analysis (PLS-DA) was performed to provide a chemical image based on the distribution of the drug throughout the gel strip.

3.3 Methodology

The methodology used was as follows: the drug rich areas were identified in the films, after the distribution of GF was observed throughout the film surface and finally, a visual comparison was carried out amongst the different areas chosen. This visual analysis is based on comparison of the drug rich areas in term of: their number, their mean value, and their standard deviations (STD).

Images were acquired and concatenated and the distributions of the pixels of drug rich areas are then observed and compared for different areas of the film.

3.4 Results and discussion

The best approach is from simplest to more complex observation. In this case, the first film analyzed was one that exhibited a high agglomeration of drug visible to the naked eye as shown in Figure 3.2. This film was prepared in this manner, to facilitate method development, by having one film with significant large agglomerates. This polymeric film was composed of drug (GF) and the main excipient (HPMC E15LV). A GF suspension was mixed with the HPMC E15LV solution without proper agitation to facilitate the agglomeration of the drug within this film. The diameter of the drug agglomerate obtained was approximately 0.3 mm.

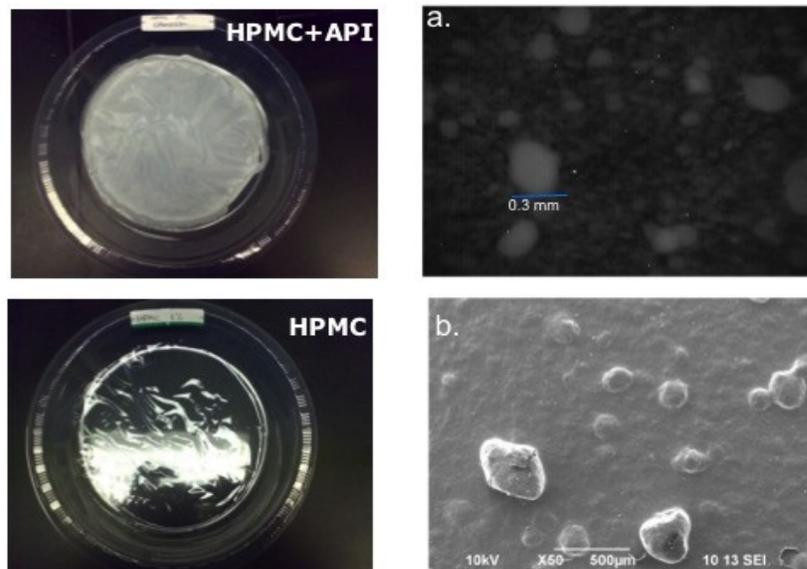


Figure 3.2 Films prepared with larger drug agglomerates. a. Image obtained with a microscope coupled to an NIR-Cl using a magnification of 10 μm per pixel. b. Image obtained with the SEM

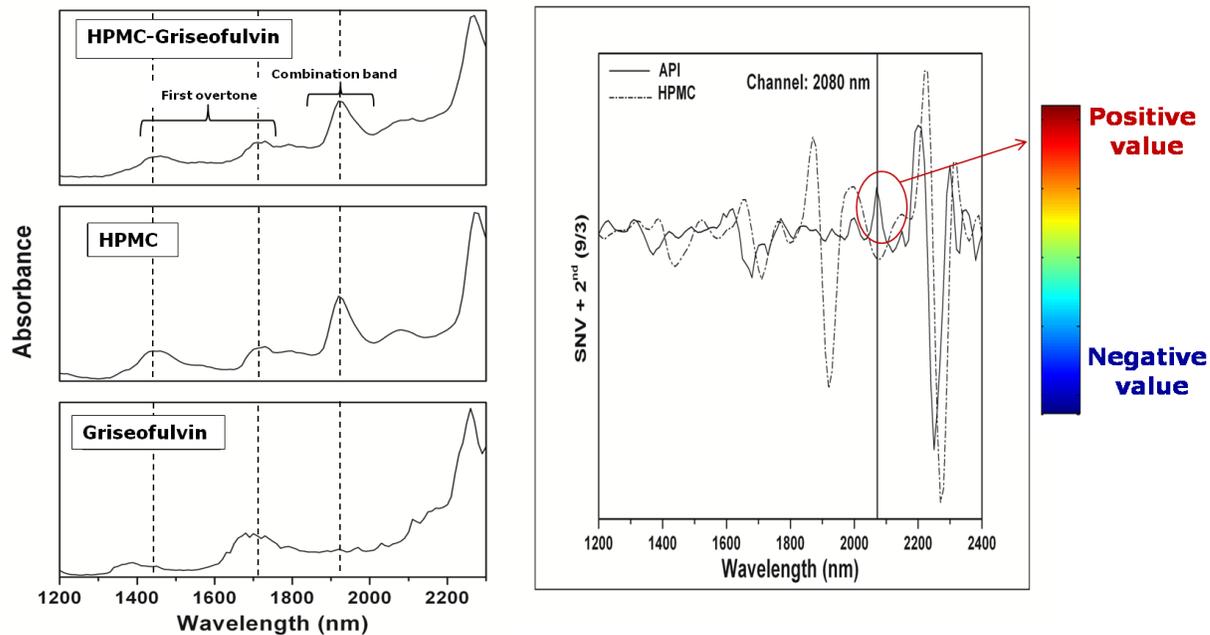


Figure 3.3 Spectra of HPMC E15LV and GF. a. Raw spectra of pure components. b. Second derivative spectra of pure components.

Figure 3.3 shows that it was not possible to identify a characteristic absorption band for the pure components due to the large band overlapping. A series of pretreatments were applied to the spectral data to differentiate between these two components (GF and HPMC E15LV). The spectral differences between API and HPMC E15LV are more clearly observed in the second derivative spectra. Differences were observed at several wavelengths, but the 2080 nm band was used since it represents an absorption maximum for GF and the minimum absorption for the HPMC E15LV. Therefore, positive values at 2080 nm indicate the presence of GF, while negative values indicate the presence of HPMC E15LV.

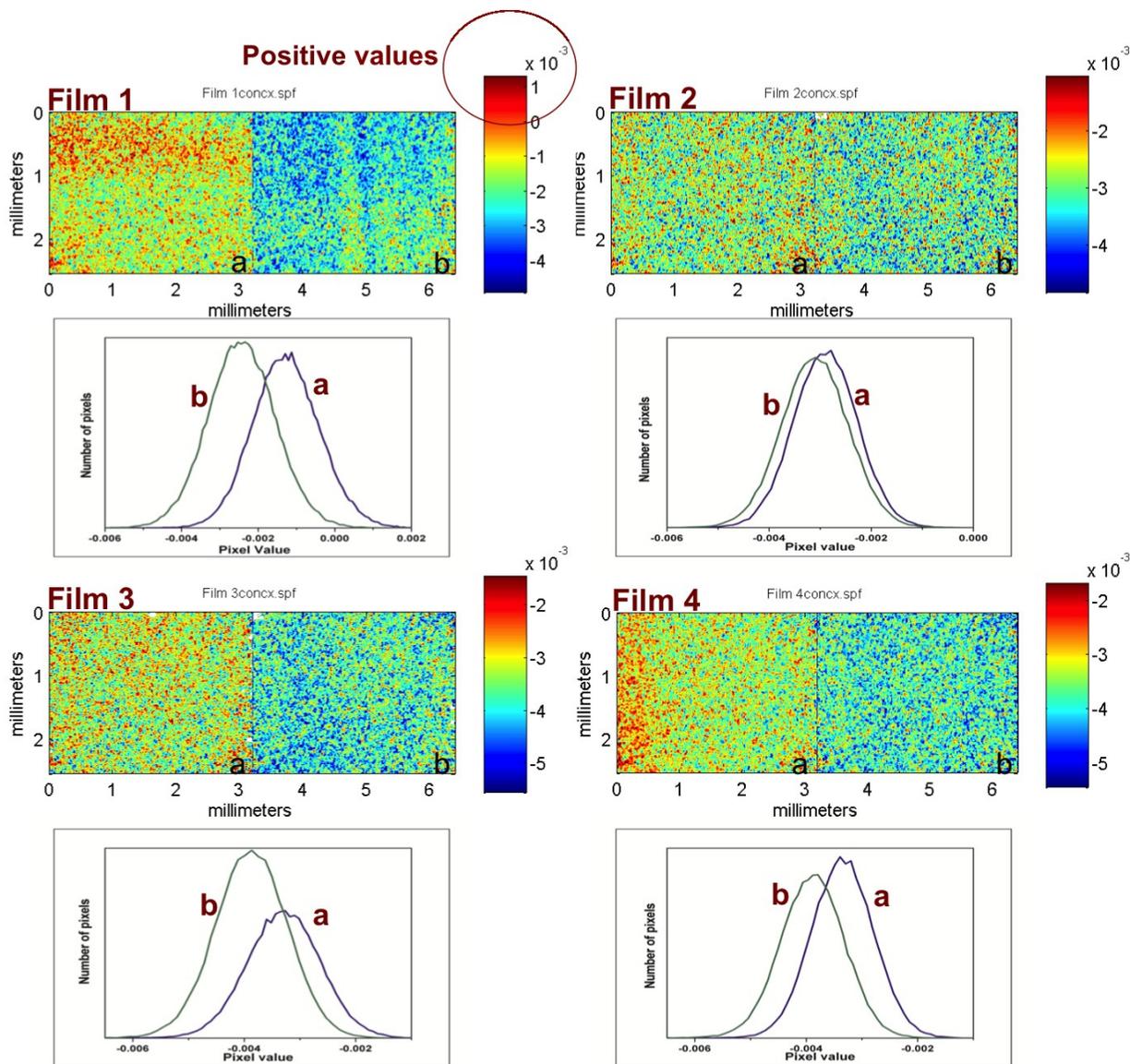


Figure 3.4 Image and histograms for the different components within film at 2080 nm.

Figure 3.4 shows a second approach to provide an easy detection of drug clusters, based on observing only the pixels with the highest second derivative intensity at 2080 nm. Figure 3.4 provides important information on the distribution of GF and HPMC E15LV throughout the gel strip; however it is still difficult to observe clusters of the drug. The pixels with the highest intensity correspond to the sections of the gel strip that are richest in GF. These images provide a qualitative assessment of the distribution of different components within the films. Assuming the histogram is a

normal distribution, these histograms show a symmetric distribution and low heterogeneity of the components within the film. Under the criteria that only pixels with a positive value correspond to drug, film 1 is the only film in which areas of pure drug would be observed because the drug is more agglomerated.

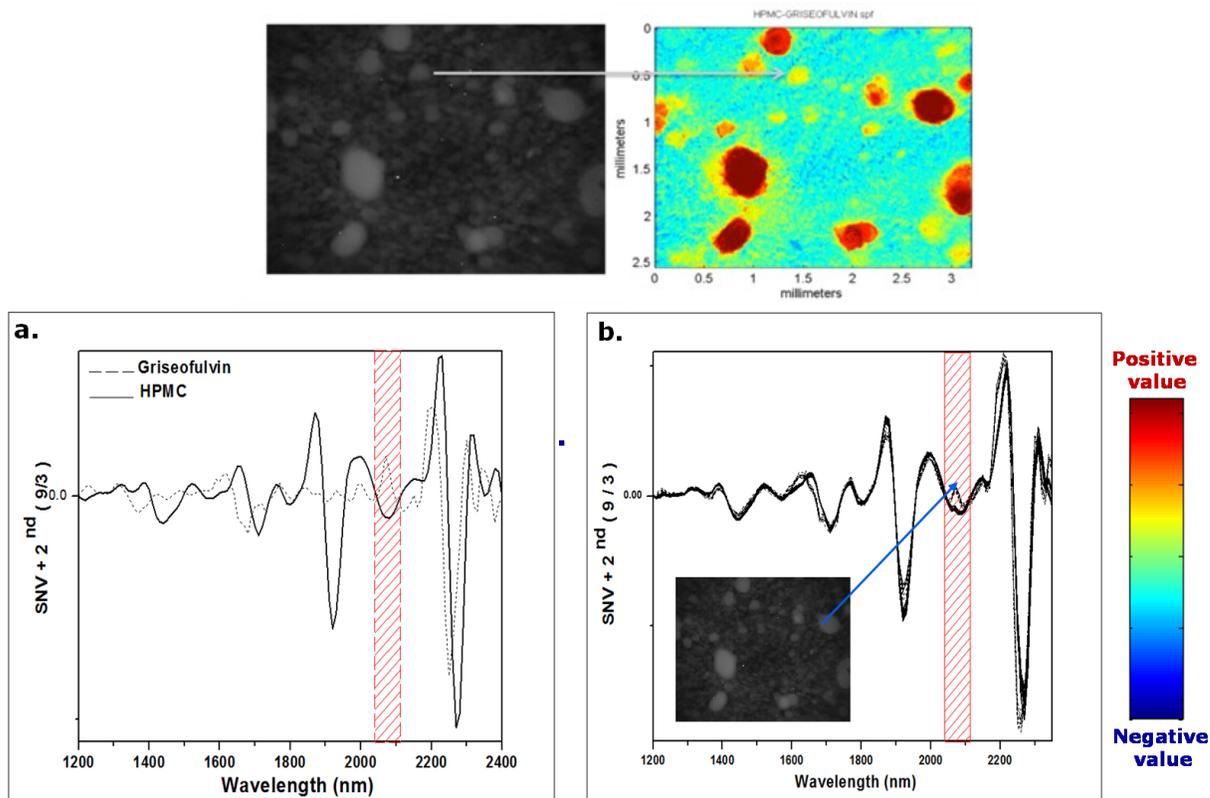


Figure 3.5 Images obtained with a microscope coupled to an NIR-CI using a magnification of 10 μm per pixels Image at 2080 nm. a. Average spectra of HPMC E15LV and GF. b. Spectra of different points of the gel strip formed by GF and HPMC E15LV.

The red areas in Figure 3.5 do not correspond to all of the areas of drug observed under the microscope. Some drug agglomerates are observed in yellow at 2080 nm. These areas represent the drug but they are not red because they also have HPMC E15LV. In others words, in these areas the drug is not pure. If there are no pure drug pixels, what is seen is the distribution of those pixels that have the maximum drug contribution. This is the definition for drug rich pixels or drug rich areas. For this type of pharmaceutical formulation, drug rich pixels will be identified as the ones that have the maximum signal intensity at 2080 nm. The selection of these pixels is based on

histograms. Histograms are used to impose the threshold value. A histogram is a bar graph where different values of intensity of all the pixels can be seen. The intensity values are on the x-axis and the numbers of pixels associated with each of these intensity values are shown on the y-axis.

Figure 3.6 explains the importance of establishing the threshold value to generate binary images. Clusters observed in the binary images vary depending on the selected threshold value. Binary images are created by setting a threshold value in the histograms. Histogram cursors truncate only the pixels of high intensity with maximum intensity values in second derivatives. The pixels that are within the cursors have a value of one and the remaining have a value of zero; this is a way of isolating the pixels of interest. The number of standard deviations selected to establish the threshold should be determined as a function of the sample and the problem to be resolved. In this case, the areas of interest are the ones of greater concentration of GF, so the values higher than two standard deviations from the mean are a good choice. Once the clusters were identified, a binary image was generated in which the clusters were isolated from the surrounding matrix.

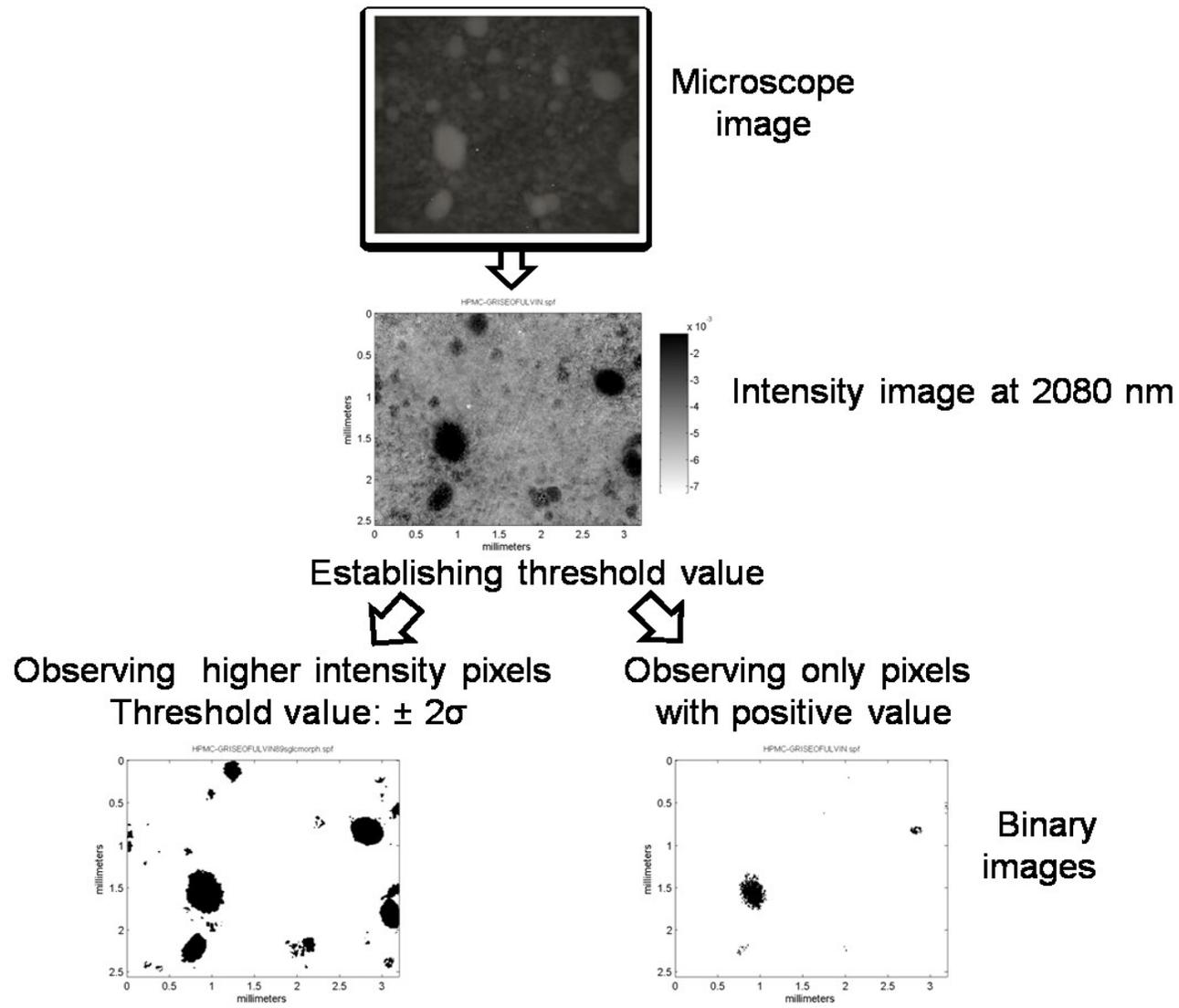


Figure 3.6 Binary Image of films with larger agglomerates of drug obtained with different threshold values.

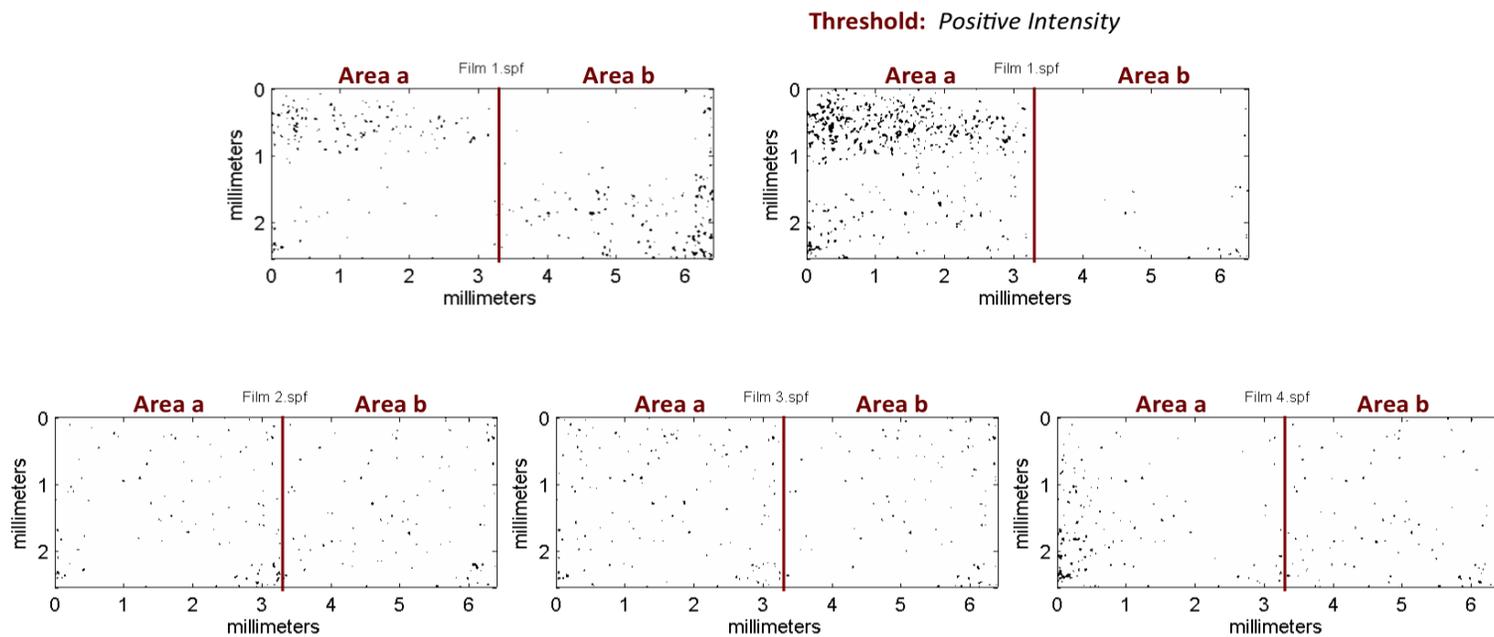


Figure 3.7 GF rich areas observed at 2080 nm.

Image analysis was then used to evaluate the number of particles observed in each image using the ISys software (Version 5.0.0.14). Figure 3.7 shows the binary image obtained to apply the threshold value. Visual examination of the binary images also indicates that in Films 1 and 4 the particles are not as widely dispersed as in the other films. The threshold value used in the analysis of these images was the mean + 2 times standard deviation. In film 1, two of the approaches were used.

Table 3.2 Domain Size Statistic Results of API Based on Binary Images Generated from Intensity Value at 2080nm

	Number of domains	Largest domain (μm^2)	Mean area (μm^2)	Area STD (μm^2)
Threshold (Positive Intensity) Avg. Film 1 (57% wt/wt API)	373	12000	1034	1386
Avg. Film 1 (57% wt/wt API)	283	5100	672	698
Avg. Film 2 (50% wt/wt API)	211	3500	444	412
Avg. Film 3 (44% wt/wt API)	248	1900	304	411
Avg. Film 4 (36.4% wt/wt API)	229	12800	558	926

Image analysis was then used to evaluate the number of particles observed from the previous binary images as reported in Table 3.2. All films show drug clusters when this approach is followed. This evaluation confirms the absence of large clusters in the film. However in films 1 and 4, the particles are not as widely dispersed as in the other films; the mean area of the clusters varies from about 300 to 670 μm^2 . Films 1 and 4 show the largest domains found and they have the largest standard deviations. The clusters that are identified by this approach have a high drug concentration but include the presence of HPMC E15LV.

The advantage of observing only pixels with positive values at 2080 nm is that this strictly defines drug rich areas, provided that the drug is highly agglomerated within the polymeric film. The advantage of observing drug rich areas is that it allows the visualization of drug rich areas in optimized processes. The disadvantage is that using this threshold value (2σ of the mean), approximately the same numbers of pixels are observed.

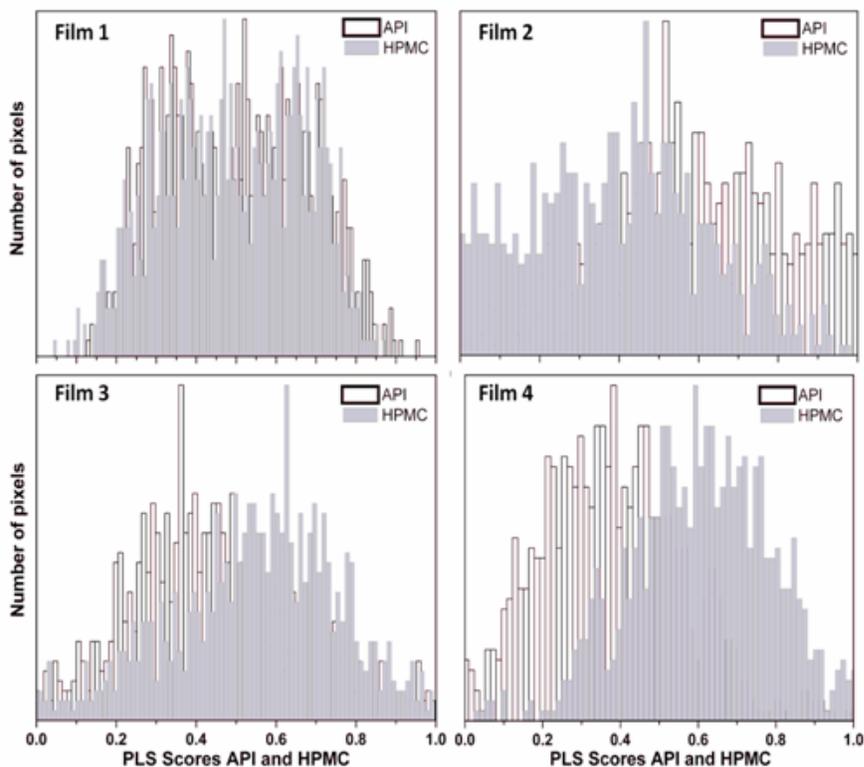


Figure 3.8 Histograms providing quantitative descriptions of the API and HPMC contents for Films 1, 2, 3, and 4. These histograms are based on Raman scores images.

Figure 3.8 shows the histograms based on PLS-DA scores of the Raman images for the different films. The histograms show a very low number of areas with high scores. The high scores are indicative of high concentrations of HPMC or griseofulvin. However, a large number of areas are observed where the scores range from 0.4 – 0.6 indicating a lower concentration of the griseofulvin or polymer. These intermediate concentration values indicate that the Raman method is detecting both griseofulvin and polymer, and that these are mixed throughout the film strip. This type of information is unique to hyperspectral images and is not available from spectra recorded from sample averaging or integrating spectrometers.

NIR-CI and Raman are complementary techniques. The two techniques are also complementary in that water and O-H bonds are poor Raman scatterers, while providing strong absorbance bands in the NIR spectral region. There are also complementary differences in the scale of scrutiny of the two methods. The NIR-CI

method provides a larger scale of scrutiny than the Raman method and hence an ability to rapidly analyze larger areas of a sample.

3.5 Conclusions

Pure drug pixels were observed only in film 1, since in this film, the drug is more agglomerated. Drug rich areas were observed in all films when was used as threshold value the mean + 2 times standard deviation. The advantage of applying this approach is that it allows to find drug rich areas in films with better drug distribution. Each approach has its advantages and disadvantages. NIR-CI can be used to find and analyze drug rich areas distribution. This is an important consideration because drug load uniformity is particularly a critical quality attributes of products where the dose of the API is low. Independently of how low is the dose of API, the dosage form must be of “regular” size if handled by the patient. Even though the sample analyzed in this case consisted of a gel strip covering an area of approximately 3.2 mm x 2.6 mm, NIR-CI could also be used at a lower magnification covering a larger area. The results obtained indicate that the NIR-CI approach could be particularly helpful to determine distribution uniformity in low dose products.

Polymer film strips loaded with small (submicron) API particles represent API composite materials with many possibilities for functionalization and use as drug delivery matrices in different types of dosage forms. Trapping fine API particles in a polymeric matrix offers an advantageous way of harvesting engineered particles. However, to fully take advantage of API particles immobilized in film strips, it is necessary to develop the ability to assess the distribution of particles in the film.

Chapter 4. MIA and NIR Chemical Imaging for pharmaceutical product characterization.

Published in *Chemometrics and Intelligent Laboratory Systems*, 2012, *Volume 117*, Pages 240-249. Prats-Montalbán, J. M.; Jerez-Rozo, J. I.; Romañach, R. J.; Ferrer, A.

4.1 Analysis of hyperspectral images of polymeric films; comparison between Multivariate Image Analysis (MIA) and Partial Least Square Discriminant Analysis (PLS-DA).

This chapter describes the effort to introduce MIA in the hyperspectral image analysis of polymeric films. This research began as a collaborative work with the Universidad Politécnica de Valencia Departamento de Estadística e I.O. Aplicadas y Calidad in Spain, with a desire to learn about MIA and its possible benefits to extract this information from NIR-CI. Part of this work was performed in Spain. This work is pioneering effort in using multivariate image analysis to evaluate drug distribution in this type of pharmaceutical formulation. The results of this collaboration were rather important in the development of the analytical technique. The use of MIA made it possible to obtain spatial distribution information of each of the components of the formulation, and understand how these components are mixed.

The scientific contribution of this chapter was to obtain the spatial distribution information between HPMC E15LV - GF; the spatial distribution of the chemical segregation zones, and the spatial distribution of the chemical mixing zones. This spatial distribution was obtained using MIA on hyperspectral images.

Chemometric tools have to be applied to hyperspectral images because they contain large quantities of spectral information, and it is impossible to extract this information without a well-structured scientific approach. Different chemometric techniques have been developed to analyze hyperspectral images.⁶²⁻⁶⁵ The majority of these techniques are based on algorithms that require a deep knowledge in programming and chemometrics. Additionally, these techniques require a large time of analysis.^{42,66,67}

Simpler hyperspectral analyses must be developed to extract information in a fast and effective way.

This chapter describes the methodology for the use of partial least squares Discriminant analysis (PLS-DA) to analyze the hyperspectral images. PLS-DA is a multivariate inverse least squares discrimination method where the response variable is a categorical one, useful for classification tasks.

In this chapter, PLS-DA was used to generate models to distinguish between GF and HPMC E15LV. The images were analyzed using two methodologies: MIA and PLS-DA.⁹ Results obtained with both methodologies are comparable in terms of drug distribution.

4.2 Reagents, Instrumentation and Data Pretreatment

During the development of this collaborative work several aspects in relation with the acquisition of the hyperspectral image were optimized. One of the improvements in the image acquisition process was verifying the quality of the image by performing a PCA to the data row. Another improvement was to wait one hour before acquiring the images when turning on the NIR equipment. In some images, when all the wavelengths were used, the MCR model generated poor results. Residuals of the MCR model were larger at the beginning and end of the spectral range. Therefore it was decided to acquire the hyperspectral images in the spectral range of 1300-2300 nm.

This chapter analyzes the same films as in Chapter 3. The reagents, instrumentation and data pretreatment followed are described in Chapter 3.

4.3 Methodology

The films were analyzed using two different approaches. The first approach applied was MCR and CLS to extract the chemical distribution maps (CDMs) from hyperspectral images, and thereafter performing multivariate image analysis (MIA) on the CDMs. The Bharati and MacGregor approach was applied to avoid textural

(spatial) information loss during the unfolding stage. Texture describes the distribution of the chemical compounds in the image. This analysis was done at the Universidad Politécnica de Valencia Departamento de Estadística e I.O. Aplicadas y Calidad in Spain.⁹

The other approach used was to apply a supervised chemometrics method called PLS-DA using the ISys software (Version 5.0.0.14). A library was generated with the films with the largest GF agglomerates. The abundance⁶⁸ (an approximation of concentration) of GF was obtained without requiring a calibration set, thus, the resulting model is then applied to each image to produce score predictions related to the abundance of GF in each image. The term abundance refers to the presence of API in each pixel. The intensity of each pixel is determined by the quantity of GF predicted at that spatial location, based on the reference spectra. Higher scores indicate more similarity to the pure component spectra. The variation in relative abundance of GF across of the polymeric film is then visualized.

4.3.1 Multivariate Image Analysis (MIA)

Once the hyperspectral images have been translated into real multivariate images, they can be analyzed by traditional Multivariate Image Analysis (MIA) techniques. MIA unfolds the images in the same fashion as in the previous MCR; it obtains a matrix where the rows are linked to the pixels and the columns to the channel.

During the unfolding stage, the textural (spatial) information is lost. In this variable reduction stage, each pixel position is changed. Therefore, the new low dimensional matrix does not represent exactly the original matrix. These two matrices can be equivalent if the information of its nine immediate neighbors is incorporated to each pixel before performing the unfolding. This way, the information is maintained during the entire process. One simple and effective way to do this is to apply the Bharati and MacGregor approach to incorporate the textural information in the unfolding process. This procedure is illustrated in the scheme presented in Figure 4.1.

From this matrix, a PCA model can be built, obtaining the scores matrix T and loadings matrix P , which can be used to obtain the Score Images (by re-folding the pixels into their original 2D dimension) and to interpret the relationship that exists within and between the chemical compounds.^{9,69}

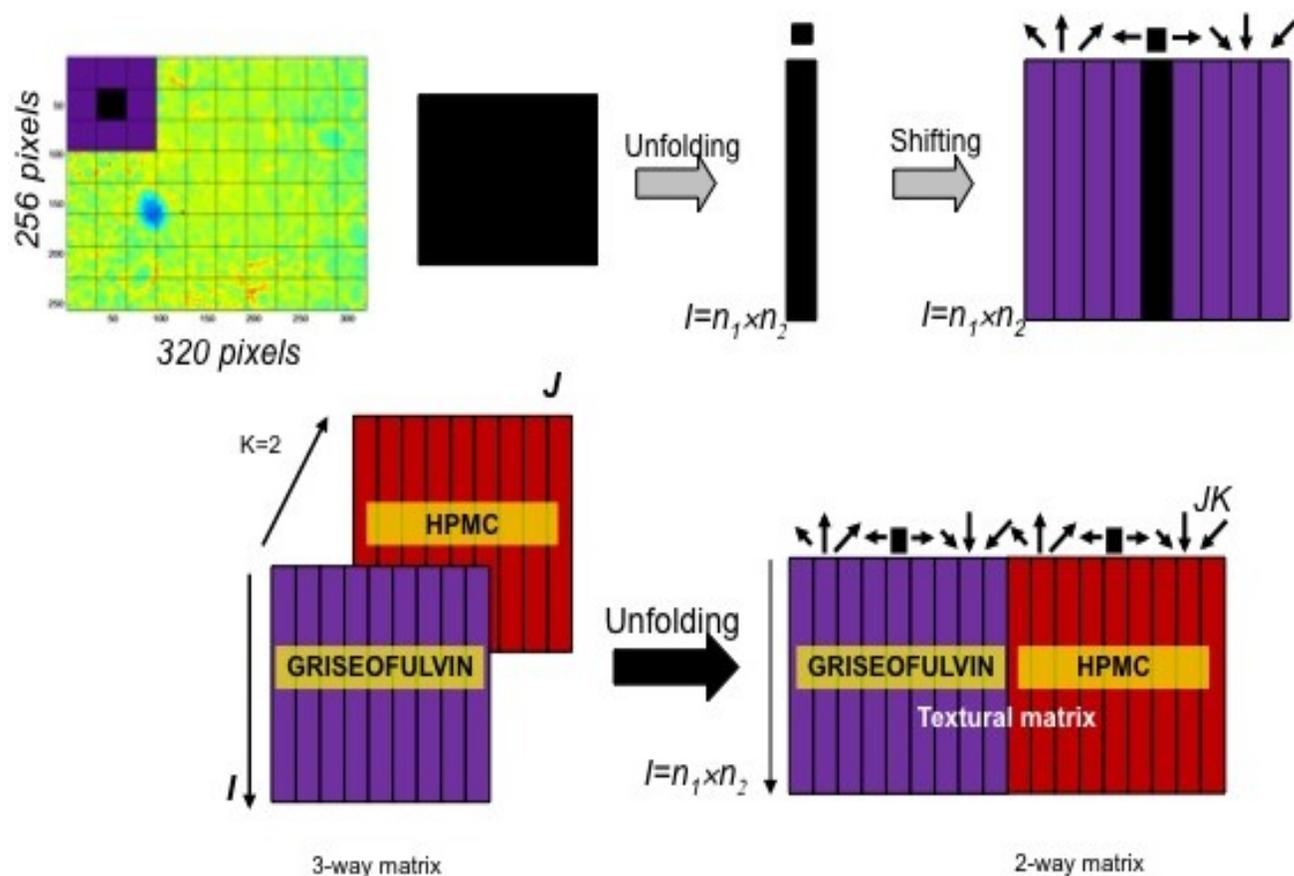


Figure 4.1 Procedure to obtain the data structure used to perform the color-textural MIA. 3×3 neighborhood window.

4.4 Results and discussion

4.4.1 Films with larger GF agglomerated

A PCA model was built, thus obtaining the different loadings and score images. The score image information was used to interpret the relation that exists within and between HPMC E15LV and GF.

The scatter plot from Figure 4.2 shows the distribution of the first two Principal Components from the PCA. These two Principal Components explain 99.89% of the variability in the film. In this scatter plot, the pixels that remain tightly clustered possess similar variability, while pixels that are separated represent the main source of variation. The spatial location of this main source of variation can be determined if these separate pixels are selected; in other words, select the pixels that do not seem to behave as a normal distribution. The main source of variation corresponds to the areas where GF is highly agglomerated. Based on the results of the previous chapter, these areas contain pixels with 100% GF and pixels that are composed mostly of GF but that also contain HPMC E15LV.

Figure 4.3 shows the score images of the first six Principal Components. The first PC is mainly showing areas around the larger cluster that contains pixels with average contribution of GF and HPMC E15LV. The second PC shows the areas of the pixels that are constituted mainly of GF or of HPMC E15LV. The rest of the Principal Components provide spatial information of each chemical compound separately. Principal Components 3 and 5 gather the textural aspects of the GF whereas Principal Components 4 and 6 explain textural aspects of the HPMC E15LV. It is very difficult to find pixels exclusively of HPMC E15LV or GF in this film, which represents the worst case of GF distribution. This information helps to understand how the drug is dispersed throughout the film surface. The film was prepared via casting; the surface tension between HPMC E15LV and GF is reduced causing the GF to be dispersed. Drug dispersion is an important factor to take in account in formulations with poorly soluble drugs.

GF is an poorly soluble drug, therefore, the only two ways to obtain pure GF for the acquisition of its image is as a powder or tablet. In both cases, specular reflection is a problem. Consequently, the film with large clusters was used to generate the reference library. The approach used to create the library with two classes, GF and HPMC E15LV, was the following: observe the maximum and minimum values of the second derivative intensity at 2080 nm. Figure 4.4 shows the pixels used to create a reference

library; 2816 pixels were used to create the GF class and 2204 pixels for the HPMC E15LV class. These two classes have pixels with pure GF or pure HPMC E15LV, but also have GF/HPMC E15LV rich pixels. A GF rich pixel is a pixel that is composed mainly of GF but that also contains HPMC E15LV. A HPMC E15LV rich pixel is a pixel that is composed mainly of HPMC E15LV but that also contain GF.

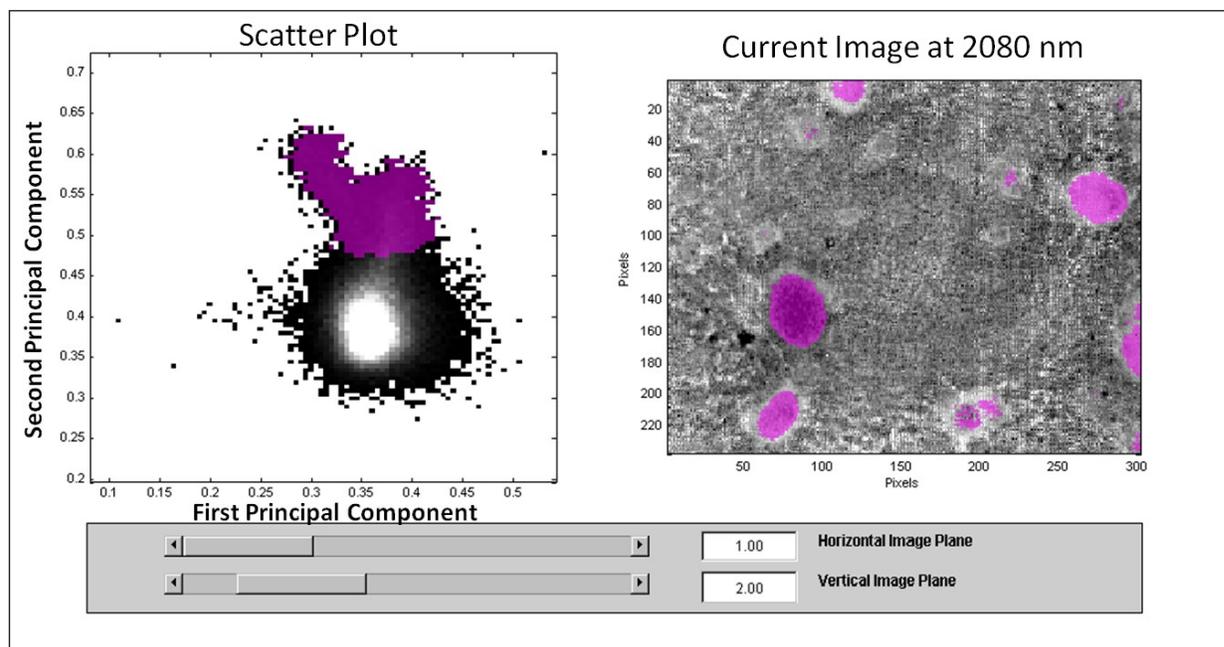


Figure 4.2 Score images from PCA. Scatter plot displaying the distribution of the first and second Principal Components.

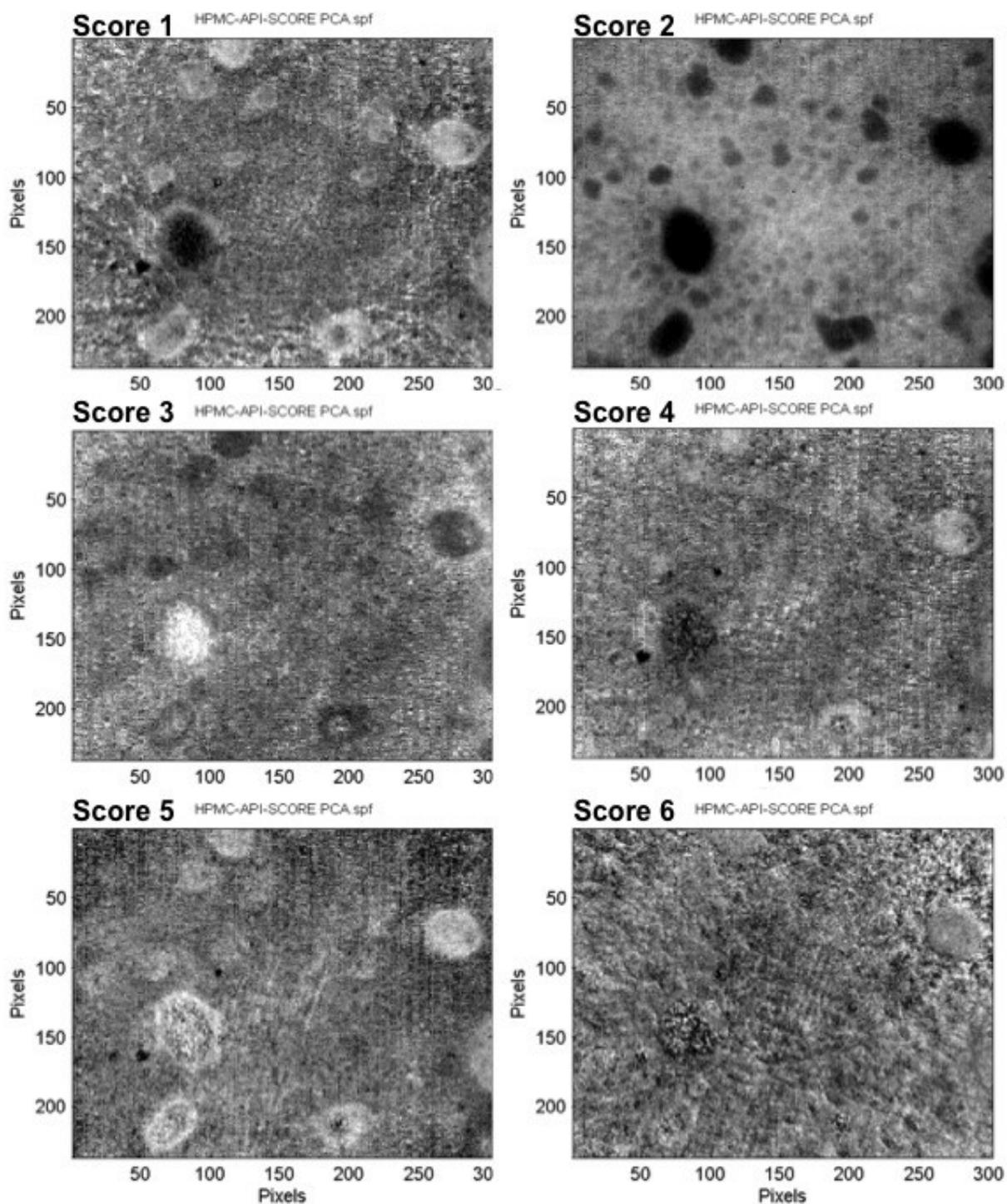


Figure 4.3 Score image of the first six Principal Components Analyses.

PLS-DA analysis based on a reference library was applied to calculate the contribution of GF at every pixel. The contribution or abundance of GF in a pixel is given by a score value. A score value of 1 or 100% means that the pixel is of pure drug. The abundance of GF in this film is of 21.7% (wt/wt). This value is obtained from an area of approximately 2.8 x 3.2 mm by pixels that are composed mainly of GF but also contain HPMC E15LV. A way to observe the distribution of each component on the film is by generating composite images.

Composite images are a way to visualize, simultaneously, the distribution of different components in a polymeric film. Composite images provide a visualization of the distribution of the different components and are very useful because the different components are shown together. This image is created by overlaying two single channels, each assigned the colors, red or green. RG image was obtained from the score image. The color intensity of each pixel in a red or green image is proportional to the intensity values from the score image it is generated from.

Figure 4.5 shows the false color spatial mapping for the predicted distribution of API (red) and HPMC E15LV (green). A red-green color-coded composite image was developed to show the localization of two components simultaneously. The RG image provided an assessment of drug distribution throughout the film. This image is easy to interpret; the drug clusters can be easily observed. This RG image obtained is consistent with the images obtained using the microscope coupled to an NIR-CI and with the image obtained with the SEM, mentioned in the previous chapter.

4.4.2 Films with a better drug distribution

Two films with a better distribution of GF were analyzed using the same methodology applied to films with larger clusters. These films contain a 44% and 50% of (wt/wt) GF.

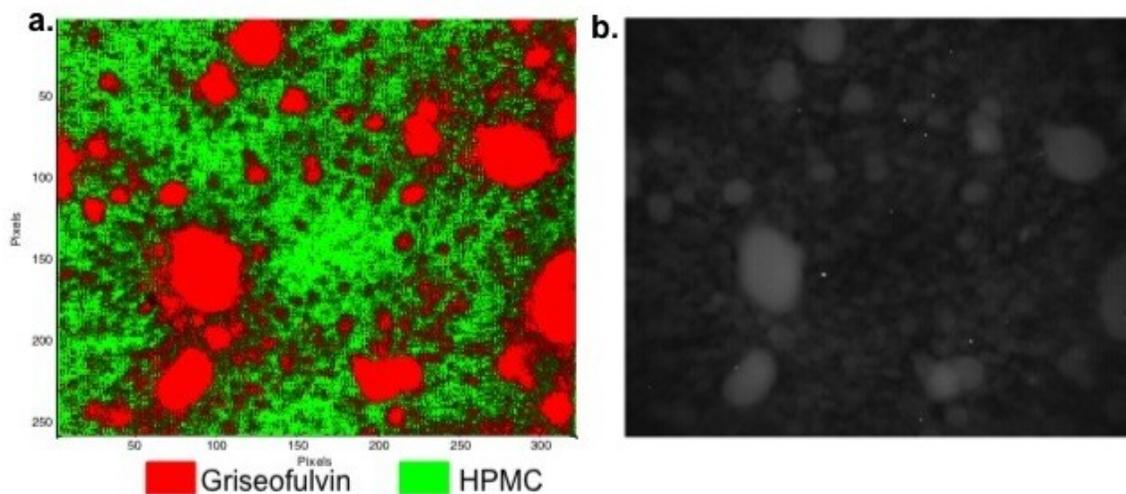


Figure 4.5 a. Red-green (RG) color-coded NIR images wherein red represents the distribution of the API, GF and green that of the excipient, HPMC E15LV. b. Image obtained using the microscope coupled to an NIR-Cl.

Two areas were analyzed in each film; these areas were concatenated for a better visualization of the results. The abundance results reported are based on the analysis of these concatenated images.

Figure 4.6 shows the distribution of the pixels when analyzing the scatter plot of the first vs. the second Principal Component of the PCA of the film with 50% (wt/wt) of GF. These two Principal Components explain the 99.86% of the variability in this film. Figure 4.6 shows that most of the pixels are together; and few pixels are separated. The next step is to observe the spatial distribution of these separate pixels. The spatial distribution of these pixels is localized on the edges of the film. The variability explained by these pixels is probably due to an issue related with the illumination. Illumination problems are presented in the edge of the images, where the lighting is not similar as in the center of the image. In imaging, it is very difficult to eliminate completely the variation in illumination, but this effect can be minimized. It is better to use lamps with polarizers, and make sure that the lamps are illuminating the center of the film and preferably work the lamps at low intensities.

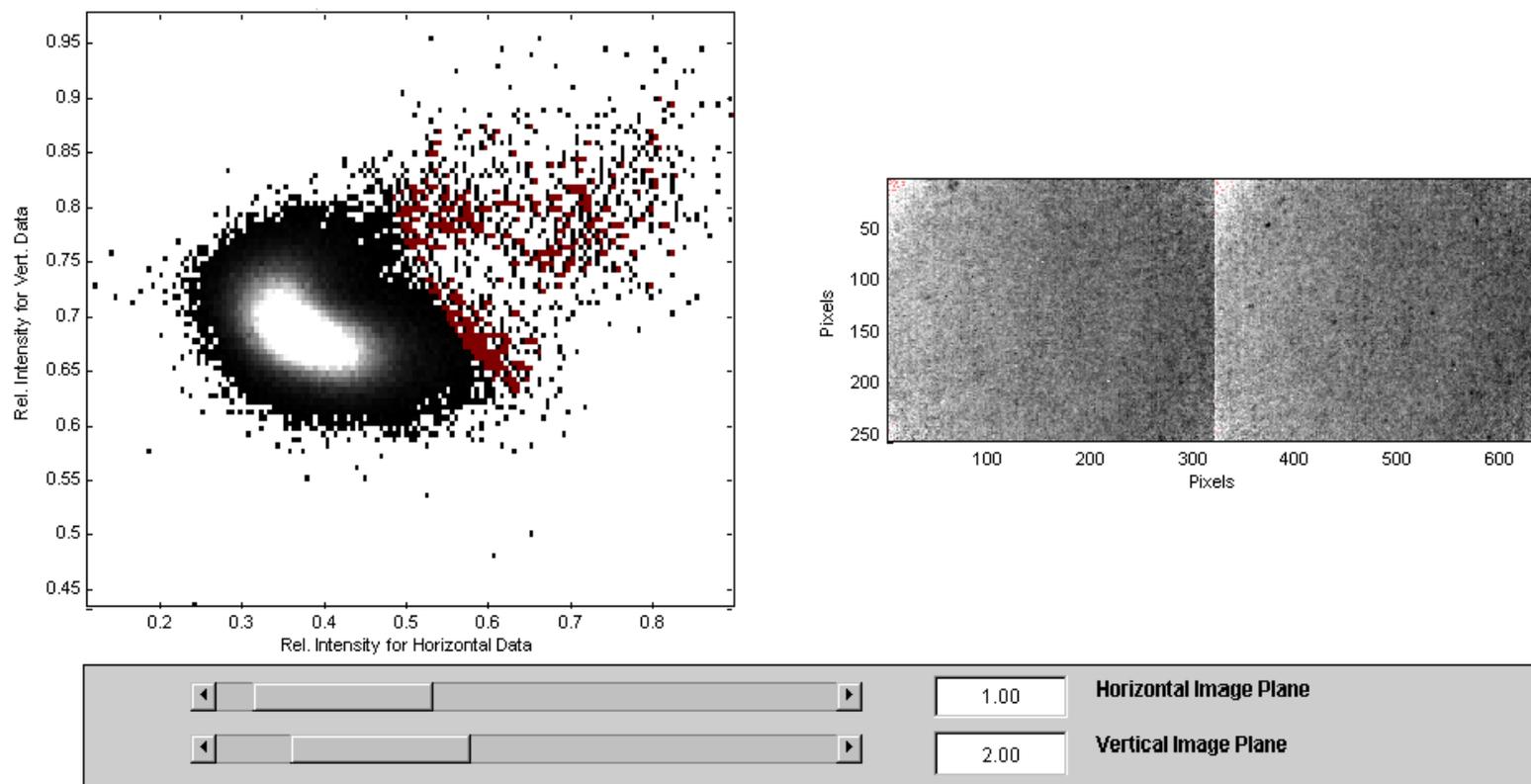


Figure 4.6 Scores images from PCA. Scatter plot displaying the distribution of the first and second Principal Component.

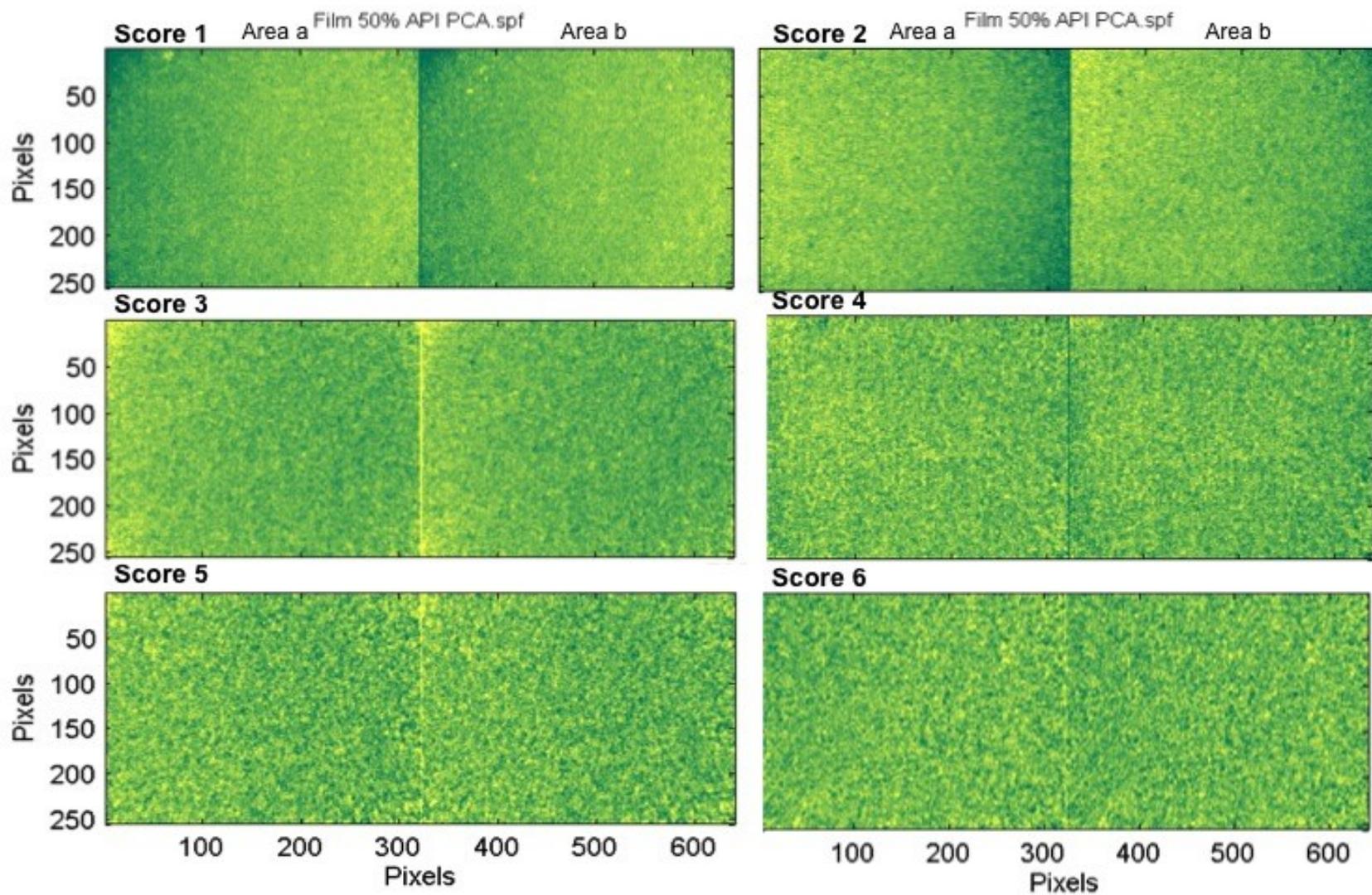


Figure 4.7 Score image of the first six Principal Components Analyses.

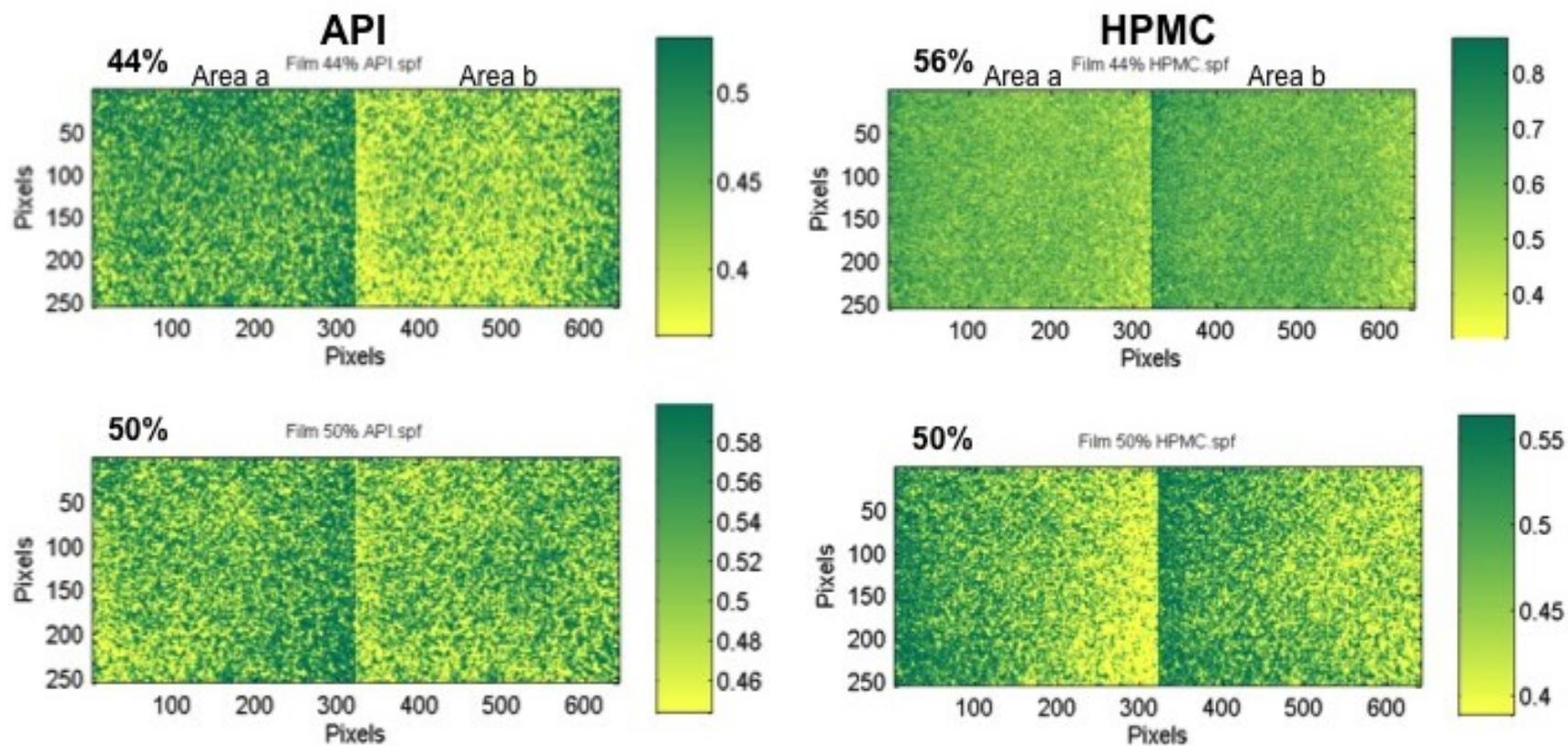


Figure 4.8 Score image of PLS-DA. Concatenated images show the distribution of HPMC E15LV and GF throughout the film surface.

The film in Figure 4.7 has a 50% (wt/wt) content of GF; it has greater drug content than the film with the larger cluster. Figure 4.7 confirms that regardless of drug content, when the drug is uniformly distributed, it is more difficult to find the relationship between drug-excipient. The graph of the first six Principal Components shows some contrast at the edges of the films, but no evidence of agglomerates.

Table 4.1 Statistical analysis of PLS Score values derived from the analysis of the two polymeric films analyzed.

Nominal value (% w/w)	Predicted	Statistical Analysis			
		Number of pixels	1 STD % (w/w)	Number of pixels	3 STD % (w/w)
44	44.6	105,436	4.6	162,862	9.0
50	52.4	104,768	4.2	162,884	8.4

Figure 4.8 shows the score images from PLS-DA. The score value scale is an indication of the spatial distribution and relative abundance of the drug and HPMC E15LV. Abundance refers to the drug distribution by a pixel. The top of the scale (green color in the color bar) indicates the highest contribution of drug or HPMC E15LV to any pixel. NIR-CI is used to evaluate the overall drug distribution throughout the film.

PLS-DA allows the comparison of the abundance results obtained throughout the surface of the films. Table 4.1 provides a summary of the abundance values for the two films. The predicted values shown in Table 4.1 are based on the abundance values of more than 100,000 pixels. The mean value is therefore comparable to the concentration value that could be obtained with a single standard NIR spectrum of the film. The nominal and predicted values are comparable. Based on the STD results, the films have a good GF distribution. The STD values of the score images are the same as the standard errors of prediction (SEP) for the response variable in the results of the ISys software (Version 5.0.0.14). Therefore, based on the STD results, NIR-CI can give quantitative results with small standard errors of prediction.

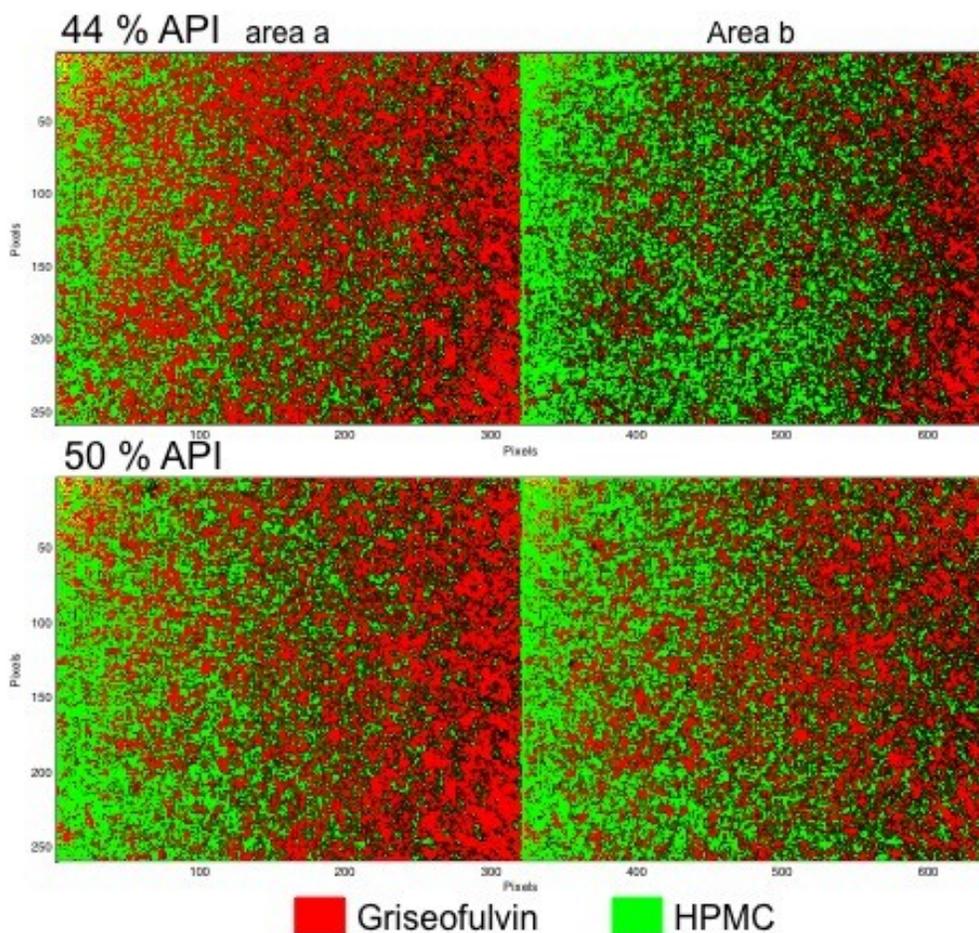


Figure 4.9 Red-green (RG) color-coded NIR images wherein red represents the distribution of the API, GF and green that of the excipient, HPMC E15LV.

Based on scores images from PLS-DA, RG images were generated. Figure 4.9 shows the RG images of the two films analyzed. In these images, each component is assigned to a separate color channel. In this case, the distribution of GF is red and the distribution of HPMC E15LV is green. In these images drug agglomerates are not observed, but the drug rich areas are not uniformly distributed.

The PLS-DA results are similar to those reported in reference 9. This reference uses a methodology based on the use of chemical oriented models (MCR and CLS) to extract the chemical distribution maps (CDMs) from the hyperspectral images, and afterwards

performing multivariate image analysis (MIA) on the CDMs, and finally extracting “channel” and textural features from the score images related to quality characteristics. PLS-DA and MIA can be used to evaluate the drug distribution throughout the surface of the polymeric film.

4.5 Conclusions

The present work demonstrated that PLS-DA can perfectly analyze this novel pharmaceutical formulation. The analysis of the film with larger GF agglomerates permitted the extraction of spectral and spatial information about the relationship between HPMC E15LV and GF. In films with better GF distribution, the relationship between HPMC E15LV and GF are not easy to visualize using PCA; it is necessary to work with scores images from PLS-DA. Results obtained are analogous with the results obtained in Reference 9 except in explaining the mixing of the materials. MIA-texture approach used in Reference 9 provided information about the spatial distribution of the chemical segregation zones, and the spatial distribution of the chemical mixing zones. This work also helped improve the image acquisition conditions.

CHAPTER 5. Hyperspectral image analysis to evaluate the effects of stabilizers on particle dispersion from polymer films.

Published in *Powder Technology*, 2013, *Volume 236*, Pages 37-51. Beck, C.; Sievens-Figueroa, L.; Gärtner, K.; Jerez-Rozo, J. I.; Romañach, R. J.; Bilgili, E.; Davé, R. N.

5.1 Summary

This dissertation is focused on the development of an analytical technique to evaluate the distribution of a poorly soluble drug on the surface of polymeric films. It is very important to recognize and understand how each component affects the distribution of the drug. The ultimate goal is to achieve a formulation of the film with a uniform drug distribution. In this chapter the effect of the surfactant on dispersion of the drug was evaluated.

The scientific contribution of this chapter was to understand the role of stabilizers to control growth and agglomeration of particles and study the formation of agglomeration of drug rich areas formed under these conditions using the methodology developed in this dissertation. The methodology identified the drug rich areas, determined the distribution of these drug rich areas throughout of the film surface and compared the distribution of drug rich areas between films.

This chapter evidences the dual role of stabilizers to control growth and agglomeration of particles formed via a liquid anti-solvent (LAS) process. Stabilization was examined using two surfactants: the non-ionic surfactant PLURONICF127 (PF- 127) and the anionic surfactant Sodium Dodecyl Sulfate (SDS). The chemical structures of these surfactants are shown in Figure 5.1. Reduction in agglomeration and increase in growth of GF were observed when using the PF127.

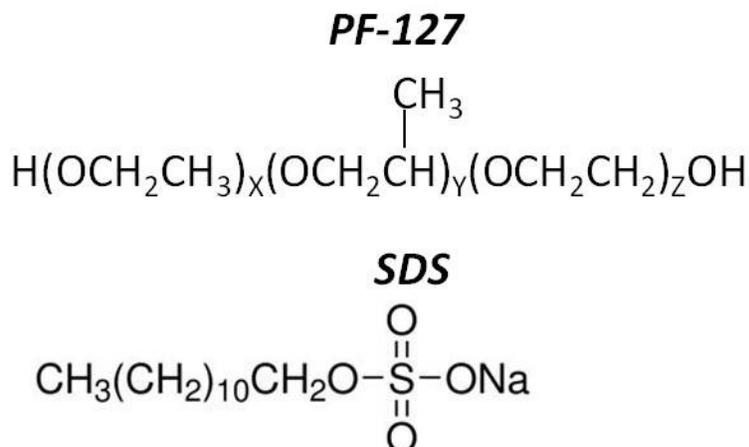


Figure 5.1 Chemical structure of the PF-127 and SDS

5.2 Materials and Methods

Griseofulvin was purchased from Letco Medical (Livonia, MI, USA). Low and high molecular weight hydroxyl propyl methyl cellulose (HPMC LV 15 Methocel, 80–120 cp) was provided by DOW Chemical (Newark, DE, USA). The wetting agent, poly (ethylene oxide)–poly (propylene oxide)–poly (ethylene oxide) (PF 127) was obtained from Sigma-Aldrich (Saint Louis, MO, USA) and glycerin (Gly) was purchased from Sigma-Aldrich (Saint Louis, MO, USA). All these materials were used as received.

5.2.1 Polymeric films

An appropriate amount of GF (GF) was briefly dissolved in acetone with some surfactant, while Hydroxy Propyl Methyl Cellulose (HPMC E15LV) was dissolved in deionized water. Following precipitation, the suspensions were centrifuged to remove most of the organic solvent, to increase the compounds concentrations. A small amount of water was added to disperse the particles and the resultant suspensions were then mixed with a high viscosity polymer solution followed by film casting and drying.

Reference 11 explains in detail the methodology used to prepare polymeric films. Table 5.1 describes the composition of the three films analyzed.

Table 5.1. Composition of the three films analyzed.

	HPMC E15LV	SDS	PF-127
HPMC E15LV (wt %)	1.49	1.49	1.49
Glycerin (wt %)	1.49	1.49	1.49
GF (wt %)	0.498	0.498	0.498
SDS (wt %)	-	0.01	-
PF 127 (wt %)	-	-	0.05

GF, HPMC E15LV and glycerin contents are the same for all films. Four different areas of the films were analyzed independently with the 10 μm /pixel objective, and for better visualization of the results, the images were concatenated.

5.2.2 NIR chemical imaging

Near infrared hyper spectral images were acquired using the Malvern SyNIRgi NIR-CI System (Malvern, UK). These images were obtained in transfectance (transmission) mode by placing the film strips over a white ceramic disk of 28 mm diameter. The 10 μm /pixel objective used provided images of an area of approximately 2.8 \times 3.2 mm. Spectra were obtained with only 1 scan in the spectral range of 1300–2300 nm. The data collected was analyzed using the ISys software (Version 5.0.0.14). The data cube was converted to absorbance units. Pixel correction was applied in order to remove areas of non-uniform illumination and to remove the effect of un-responsive pixels on the detector. A low-pass Fourier filter was also used and spectra were normalized using mean center unit variance by spectrum and a Savitzky–Golay second derivative (filter order 3, filter width 7) in order to eliminate variations in slope.

5.3 Results and Discussion

The objective is to produce a polymeric film with GF particles widely dispersed. In a supersaturated medium, GF particle coarsening due to variations in radii of curvature results in the growth of GF particles to reduce Gibbs free energy and surface area. It is necessary that GF particles be kept separate from each other, in order to prevent aggregation. An important factor to take in account is to stabilize GF suspensions. The surfactant reduces hydrophobicity of GF particles by reducing interfacial or surface tension producing more stable suspensions due to electrosteric stabilization. It is necessary to examine the use of surfactants to prevent that the growth dominates the precipitation of GF particles. The effect of surfactant on the agglomeration of the GF particles was studied using NIR-CI.

The approach followed for the NIR image analysis of these films was to identify those pixels with the highest second derivative intensity at 2080 nm. Positive values at 2080 nm are indicative of presence of pure GF. Concatenation was used to link the four areas analyzed of each polymeric film so that they are processed in exactly the same way, facilitating the visualization as well as comparison of the results.

The second derivative intensity recorded at the different pixels of the data cube at a wavelength of 2080 nm was used to generate the chemical images and histograms shown in Figures 5.2, 5.3 and 5.4. The histograms of the four areas of each film show subtle differences in the symmetry of their distributions. Differences in symmetry can be related with differences in the distribution of GF rich areas. These differences are based on the intensity distribution of all pixels on each area; if the GF rich areas are uniformly distributed, the histograms are symmetric. The HPMC E15LV stabilized film and the HPMC E15LV-SDS stabilized film present the greatest differences in the pixel distribution. The distribution of pixels corresponding to the HPMC E15LV- PF-127 stabilized film is better; their histograms are centrally located. These images provide a qualitative assessment of the heterogeneity of the polymeric thin films.

Table 5.2 Intensity values at 2080 nm of the different films analyzed.

	High Intensity Value	Low Intensity Value
<i>HPMC E15LV</i>	-0.006786	-0.02753
<i>SDS</i>	-0.005185	-0.02533
<i>PF-127</i>	-0.003479	-0.02945

Table 5.2 shows the high and low signal intensity value of pixels different at 2080 nm.

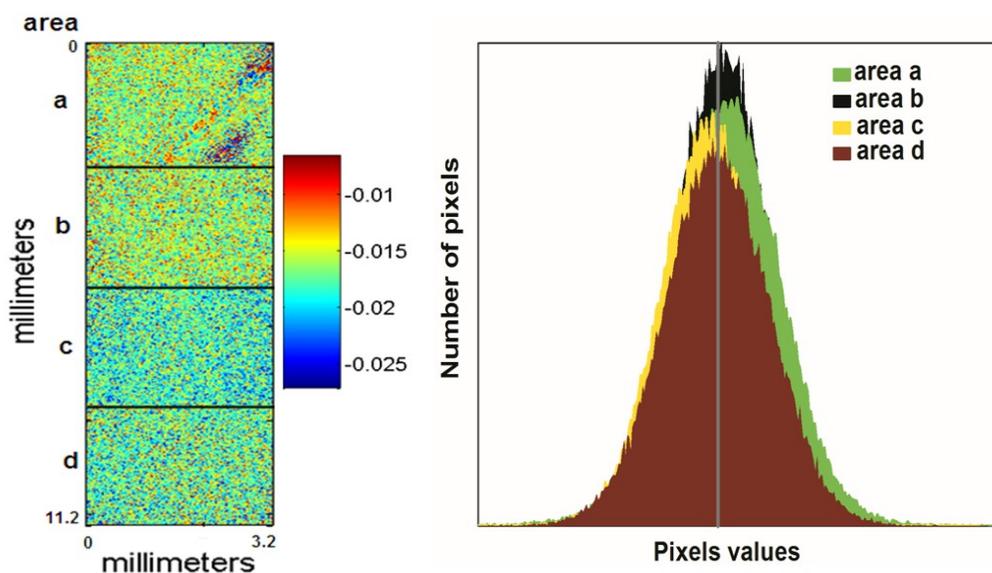


Figure 5.2 Image and histograms for the HPMC E15LV stabilized film at 2080 nm. This film does not contain stabilizers.

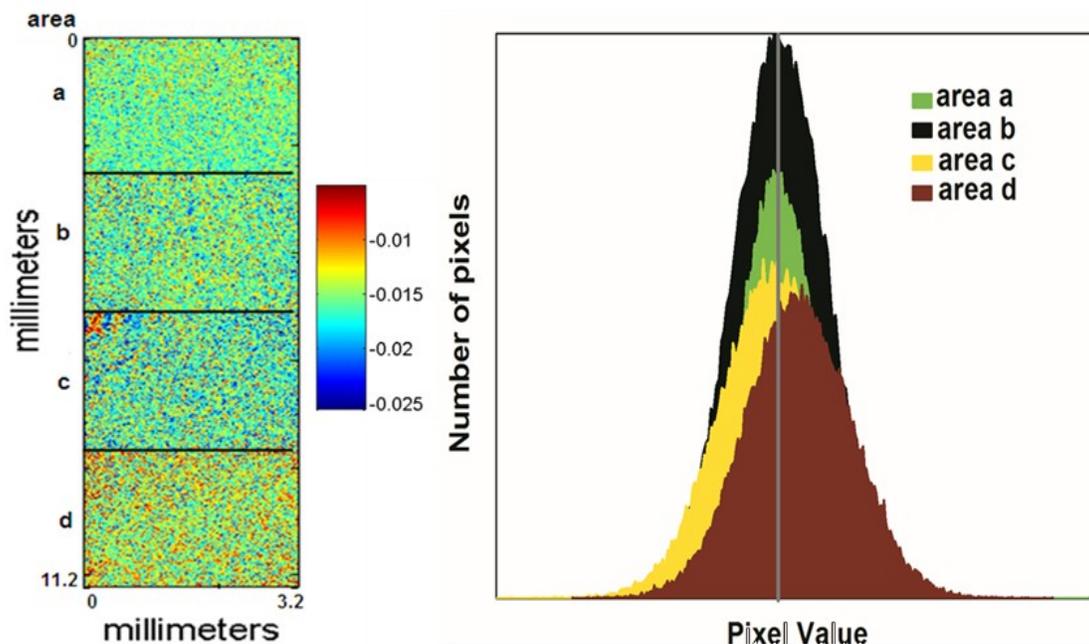


Figure 5.3 Image and histograms for the HPMC E15LV-SDS stabilized film at 2080 nm. This film does not contain stabilizers.

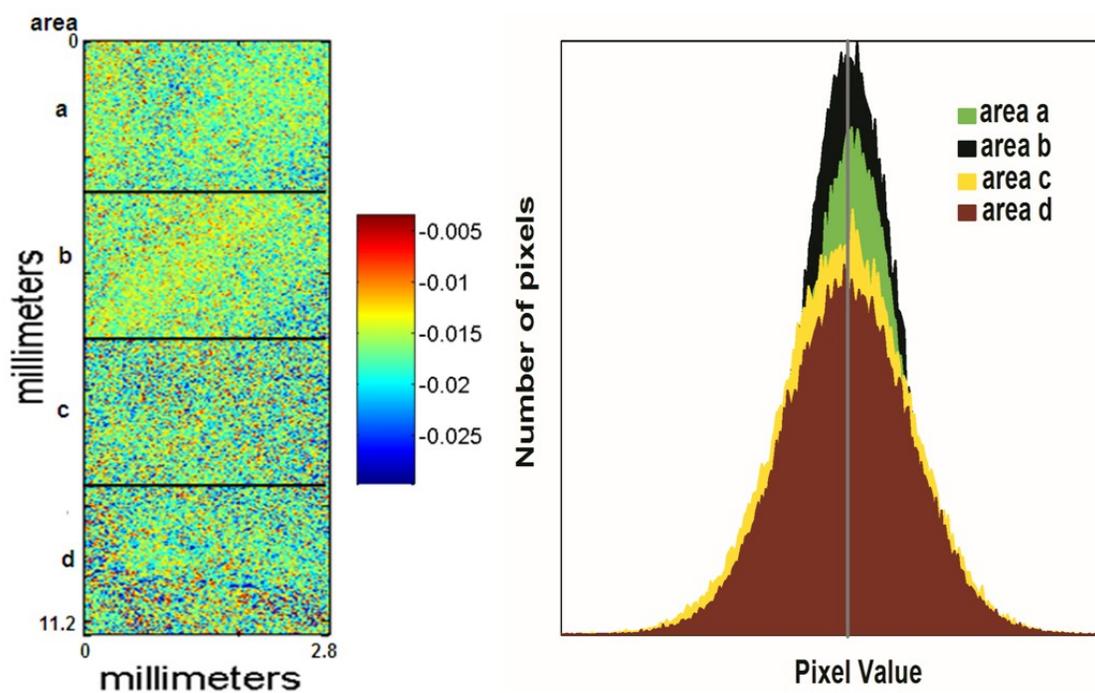


Figure 5.4 Image and histograms for the HPMC E15LV- PF-127 stabilized film at 2080 nm.

These films did not have pixels of pure GF. The pixels with maximum intensity value observed at 2080 nm correspond to those pixels with the highest concentration of GF, but, these pixels also have the contribution of the different excipients. Pixels associated with a greater contribution of GF are those that have the highest signal intensity or those which are closer to zero. The pixels of the HPMC E15LV- PF-127 stabilized film have the highest contribution of GF per pixel: the GF is more agglomerated per pixel. This agglomeration of GF may be due to the Ostwald ripening effects.⁷⁰ With progress of time, the larger particles in the dispersion grow in size at the expense of the finer particles.

The addition of stabilizers impacts the GF particle distribution and different surfactants lead to different degrees of agglomeration of the GF rich areas. The decrease in agglomeration is due to the suppression of GF particle interactions via surface shielding. In the absence of stabilizers, significant rapid agglomeration occurs.

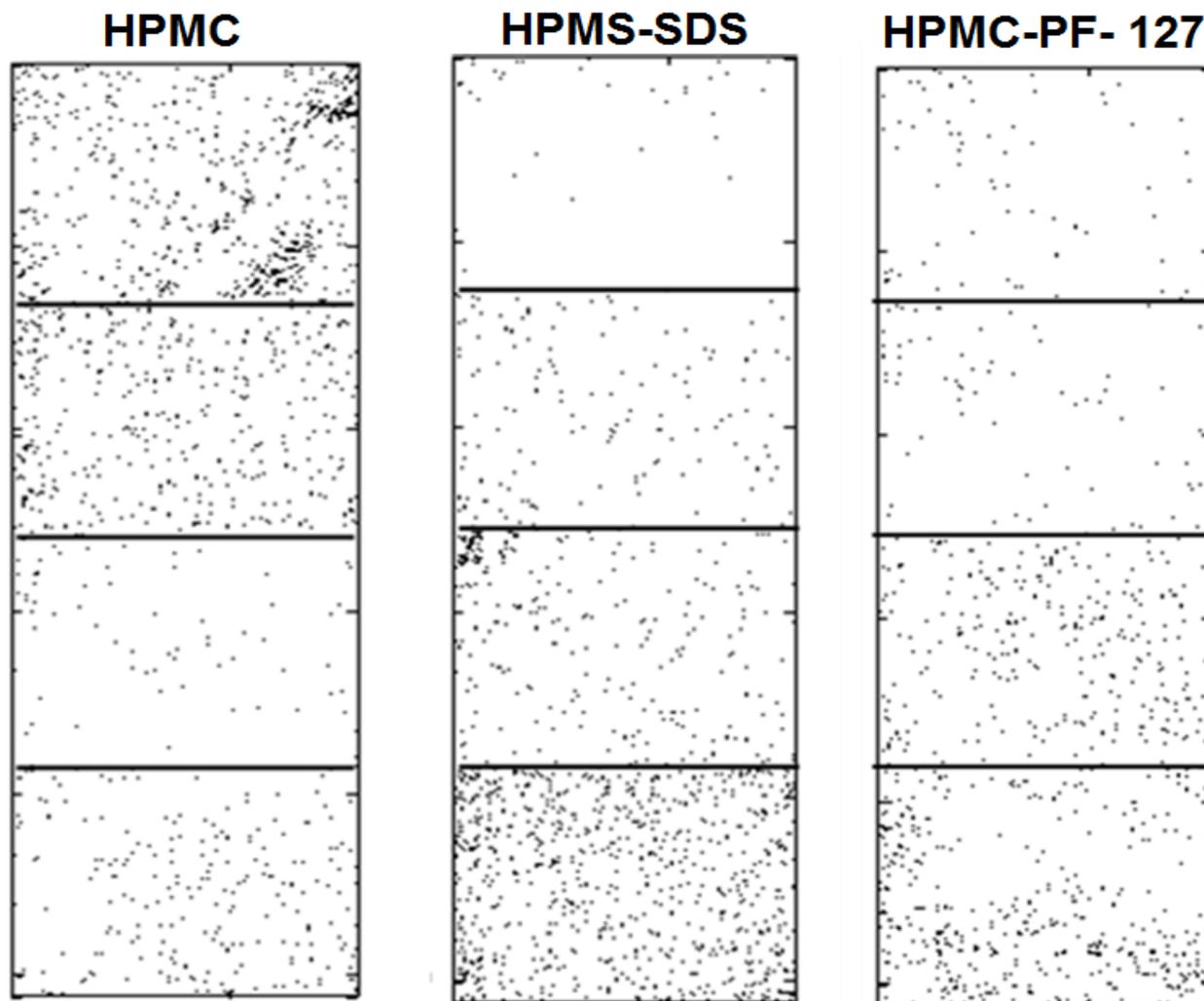


Figure 5.5 GF drug rich areas observed at 2080 nm.

Table 5.3 Domain size statistic results of API based on binary images generated from intensity value at 2080 nm.

	Number of domains	Largest domain (μm^2)	Mean area (μm^2)	Area STD (μm^2)
HPMC E15LV	2255	35300	410	867
HPMC E15LV-SDS	2193	19700	392	545
HPMC E15LV-PF-127	1665	4200	358	308

Areas of high contrast seen in the intensity images are due to the chemical differences in the constituents. However, chemical differences and drug clusters are somewhat difficult to observe in intensity images. Thus, binary images were used to observe the drug clusters more clearly. Binary images were created with those pixels with values greater than 2 standard deviations of the mean of the second derivative intensities observed at 2080 nm. These pixels with the highest intensity are considered GF rich areas. The pixels that are within two standard deviations of the mean are assigned a value of zero and those that exceed are given a value of one. Consequently, GF rich areas are isolated and binary images are obtained, also shown in Figure 5.5. Visual examination of these binary images indicates that the GF rich areas are not widely dispersed in the HPMC E15LV stabilized film. Image analysis was then used to evaluate the number of GF rich areas observed as reported in Table 5.3.

Table 5.3 shows the statistics from the binary image analysis, where the pixels exceeding two standard deviations indicate the presence of GF rich areas. The HPMC E15LV stabilized film has the highest standard deviation, the highest number of GF rich domains, and the largest domain. All of this indicates that the GF particles are not uniformly distributed in the HPMC E15LV stabilized film as compared to the HPMC E15LV-PF127 and HPMC E15LV-SDS stabilized films. The best dispersion was observed for the HPMC E15LV-PF127 case.

The area covered is the number of pixels associated with GF rich areas in binary image (dark pixels) as a percent of the total number of image pixels. The percent area covered gives a rough estimate of the API concentration. As the three films have the same concentration of API, it is expected that the percentage of the area is similar in the three films. The percentage of the area covered is lower for the HPMC E15LV-PF-127 stabilized film (1.82%) compared with the other two films (2.82% for the HPMC E15LV stabilized film and 2.62% for HPMC E15LV-SDS stabilized film). These GF rich areas are uniformly distributed in the four areas of the HPMC E15LV-PF127 stabilized film in comparison with the others two films. The discrepancy in absolute values is attributed to the fact that the drug may not distribute in the same way along the depth of the films and that the API signal does not necessarily translate to the exact amount of drug but rather to the 2-D distribution and the extent of the agglomeration.

Results suggest that the GF rich areas in the HPMC E15LV-PF127 stabilized films are more dispersed, are smaller, and show less variation in their size. These results should be interpreted together. The film with better drug distribution will be the film that has the smallest area STD and domains of all the films that are being compared. The percentage of the area covered is important when the films have a similar number of GF rich areas. Films with lowest percentage of the area covered show the better drug distribution. Another important point is the comparison of areas within each film. Number and distribution of GF rich areas should be consistent through the different areas analyzed for each film. The information obtained confirms that the films with the best distribution of GF rich areas is the HPMC E15LV-PF127 stabilized film.

Additionally, the dissolution rate was substantially higher and more distinct for the films where the suspension was stabilized using HPMC E15LV-PF127 compared to films where the suspensions were stabilized using HPMC and SDS.¹¹

Figure 5.6 shows the main difference in the mechanism of stabilization of these two surfactants; SDS and PF-127. In the adsorption process, HPMC E15LV displaces SDS, eliminating their stabilization effect. When PF-127 is used, on the other hand, the adsorption of the HPMC E15LV occurs without replacing it. The mechanism of stability of the GF/HPMC E15LV/PF-127 system may occur in the following manner: the hydrophobic core of PF-127 anchors onto the surface of the GF particle leaving its hydrophilic side chain protruding in solution, and the HPMC E15LV backbone attaches itself on the particle surface. NIR-CI results indicated that the combination of PF-127 and HPMC E15LV provides stability against agglomeration.

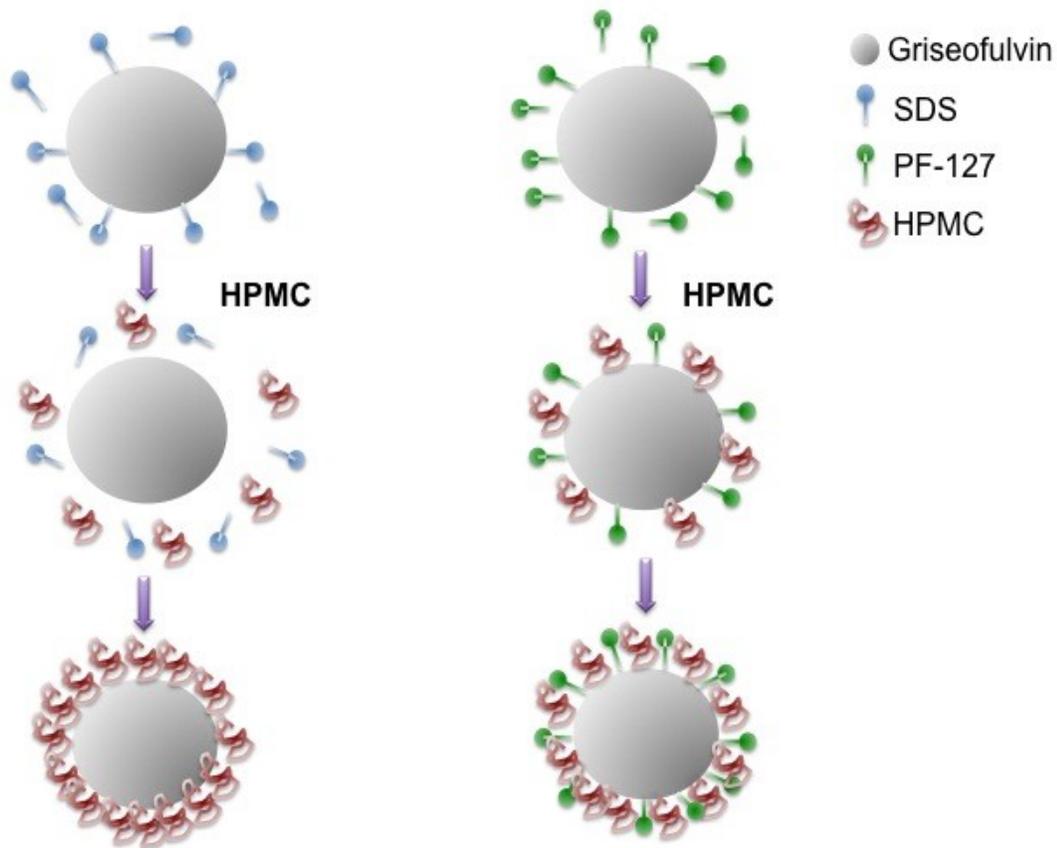


Figure 5.6 Adsorption model of the two surfactants used. Adapted from Changhoon Chai, “Hierarchical structures of micro and nano API particulates on gelling polymeric matrix”, Presented at Rutgers University, New Jersey- ERC-SOP.

5.4 Conclusions

The addition of surfactant and polymer resulted in the reduction of particle size and the minimization of agglomeration. NIR-CI helped assess agglomeration behavior in the films. NIR-CI provided important information about the distribution of GF rich areas when two different surfactants were used. The HPMC E15LV- PF-127 stabilized film has the highest contribution of GF per pixel; the GF is more agglomerated per pixel.

These GF rich areas are uniformly distributed throughout the four areas analyzed. NIR-CI results enable concluding that the combination of PF-127 and HPMC E15LV provides stability against agglomeration. These results are consistent with the results obtained using other characterization techniques.¹¹

CHAPTER 6. Hyperspectral image analysis of polymeric films containing different BCS Class II drugs

Published in *International Journal of Pharmaceutics*, 2012, *Volume 423*, Pages 496-508. Sievens-Figueroa, L.; Bhakay, A.; Jerez-Rozo, J. I.; Pandya, N.; Romanach, R. J.; Michniak-Kohn, B.; Iqbal, Z.; Bilgili, E.; Dave, R. N.

6.1 Summary

Over 40% of the new pharmaceutical formulations contain poorly soluble drugs as API.⁷ Griseofulvin was the first poorly soluble drug analyzed with the analytical methods developed and described in this dissertation. The methods were also used to study formulations of polymeric films containing other poorly soluble APIs.

Three BCS Class II drugs were studied in this chapter. This investigation was based on the analysis of their API distribution rich areas. Results demonstrate that the methodology developed can be applied to formulations with other very poorly soluble drugs. This increases the capability of the analytical method developed. This is the scientific contribution of this chapter.

This chapter is based on the analysis of polymeric films that contain three different BCS Class II drugs: naproxen (NPX), fenofibrate (FNB), and griseofulvin (GF). Differences in aggregation behavior of APIs in these films were observed through NIR chemical imaging analysis. NPX exhibited the strongest aggregation compared to the other drugs. The aggregation had a direct effect on drug content uniformity in the film.

6.2 Samples

The drug molecules utilized were griseofulvin (Sigma–Aldrich), naproxen (Medisia), and fenofibrate (Ja Radhe Sales). The particle sizes of these drugs were reduced using a wet stirred media mill (WSMM).

Sodium Dodecyl Sulfate (SDS) (Sigma–Aldrich), low molecular weight Hydroxyl Propyl Methyl Cellulose (HPMC E15LV) (Dow Chemical) and Glycerin (Sigma) were used without further processing. In this work, each one of these three films (NPX, FNB, GF) drugs were dispersed in HPMC E15LV using SDS as a surfactant and glycerin as a moisturizer. The concentrations of the films are shown in Table 6.1.

Table 6.1 Nominal concentrations of the polymeric films.

API	28% (wt/wt)
HPMC E15LV Methocel E-15LV	39% (wt/wt)
Sodium Dodecyl Sulfate (SDS)	1.0% (wt/wt)
Glycerin	32% (wt/wt)

The procedure of preparation of all the suspensions was identical.⁷ A solution containing HPMC and glycerin was prepared by adding the glycerin to water and heating to 80°C. At this temperature, the polymer was then added until well dispersed and the temperature was decreased to room temperature to dissolve the polymer completely. The components were mixed until a clear solution was obtained. The resulting solution was then let to rest until no bubbles were seen. The polymer solution was then added to the drug nanosuspension produced from WSMM and then mixed for 3 h and left to rest until no bubbles were observed. The GF, FNB and NPX nanosuspensions had a median particle size of 163 nm, 201 nm, and 144 nm, respectively. The final viscous suspension was then casted onto a stainless steel plate. The film was then dried overnight in an oven at 42°C.

Different areas of the films were analyzed independently with a 10µm/pixel objective, and the images were concatenated for better visualization of the results. The whole film was analyzed using the 131 µm/pixel objective which has a field view of 34 x 42 mm capturing the whole film.

6.3 Instrumentation

Near infrared Hyperspectral images were acquired using the Malvern SyNIRgi NIR-CI System (Malvern, UK). The images were obtained in transreflectance mode by placing the film or gel strip over a white ceramic disk of 28 mm diameter. The optical magnifications used were 10 μm and 131 μm which provided images of an area of approximately 2.8 x 3.2 mm and 34 x 42 mm, respectively. The spectra were obtained with 1 scan in the spectral range of 1300-2300 nm. The data collected was analyzed using the ISys software (Version 5.0.0.14).

6.4 Data treatment

The spectra acquired by the NIR-CI System are reflectance spectra. Therefore, the logarithm, $\log_{10}(1/R)$, was first applied to the data cube to convert the spectra to absorbance units. Pixel correction was applied in order to remove areas of non-uniform illumination and to remove the effect of unresponsive pixels on the detector. A low-pass Fourier filter was also used to reduce spectral noise. The spectra were normalized using Standard Normal Variation (SNV) and a Savitzky-Golay second derivative (filter order 3, filter width 7) in order to eliminate variations in slope. The spectral range was restricted to 2000-2300 nm as shown in Figure 6.1. The greatest differences between excipients and different APIs were observed in this region of the spectrum. Pure component imaging data was collected from powder samples of the APIs and excipients and processed in the same manner. A reference library with the pure components spectra was built. These training spectra were used as predictors to build a Partial Least Squares Discriminant Analysis (PLS-DA) in the classification model. The intensity of each pixel in the resultant scores image is determined by the degree of membership (scaled from 0 to 1) predicted for the spectrum at that spatial location. Based on reference spectra, brighter pixels (red in the color bar), are the ones that have the stronger degree of membership with regard to the API classes predicted at that location. The variation in pixel intensity reflects the variation in concentration across the sample. These scores images were evaluated morphologically using a particle statistics tool.

6.5 Results

Figure 6.1 shows the spectral range selected. Some of the bands observed in this spectral range correspond to the second overtone of the carboxylate anion stretching and the stretch combination of a double bond. These types of vibrations are not present in the excipients. Additionally, in this region, the greatest difference between the APIs and the excipients spectra can be observed. Therefore, the spectral changes observed in this range can be related to variations in API concentration. This is important because it helps to develop a more selective PLS-DA model.

Figure 6.2 shows the distribution of NPX, GF and FNB throughout the surface of the films using the 10 μ m/pixel objective. Twelve areas of each film were analyzed with this objective; each area generating a score image. These scores images were concatenated for a better visualization of the results. The score value scale is an indication of the spatial distribution and relative abundance of API. Abundance refers to the API distribution at a particular sampling point which is the film area analyzed by a pixel from the NIR chemical imaging system. The top of the scale (red color in the color bar) indicates the highest contribution of API to any pixel. Pixels rich in API are observed in each of the analyzed areas. Areas of pure API are evidenced by an abundance value of 1.0 or 100%. Pixels that show only the spectrum of the API are not observed when using the 10 μ m/ pixel objective (the largest magnification objective used). Pixels rich in API(0.25-0.6 or 25%(wt/wt)-60%(wt/wt)) are observed in these polymeric films. This confirms that each spectrum acquired by a pixel is a mixture of excipients and API. Therefore, this indicates that the API is not highly agglomerated in these films. These results are consistent with results of past analyses using different approaches.^{9,10}

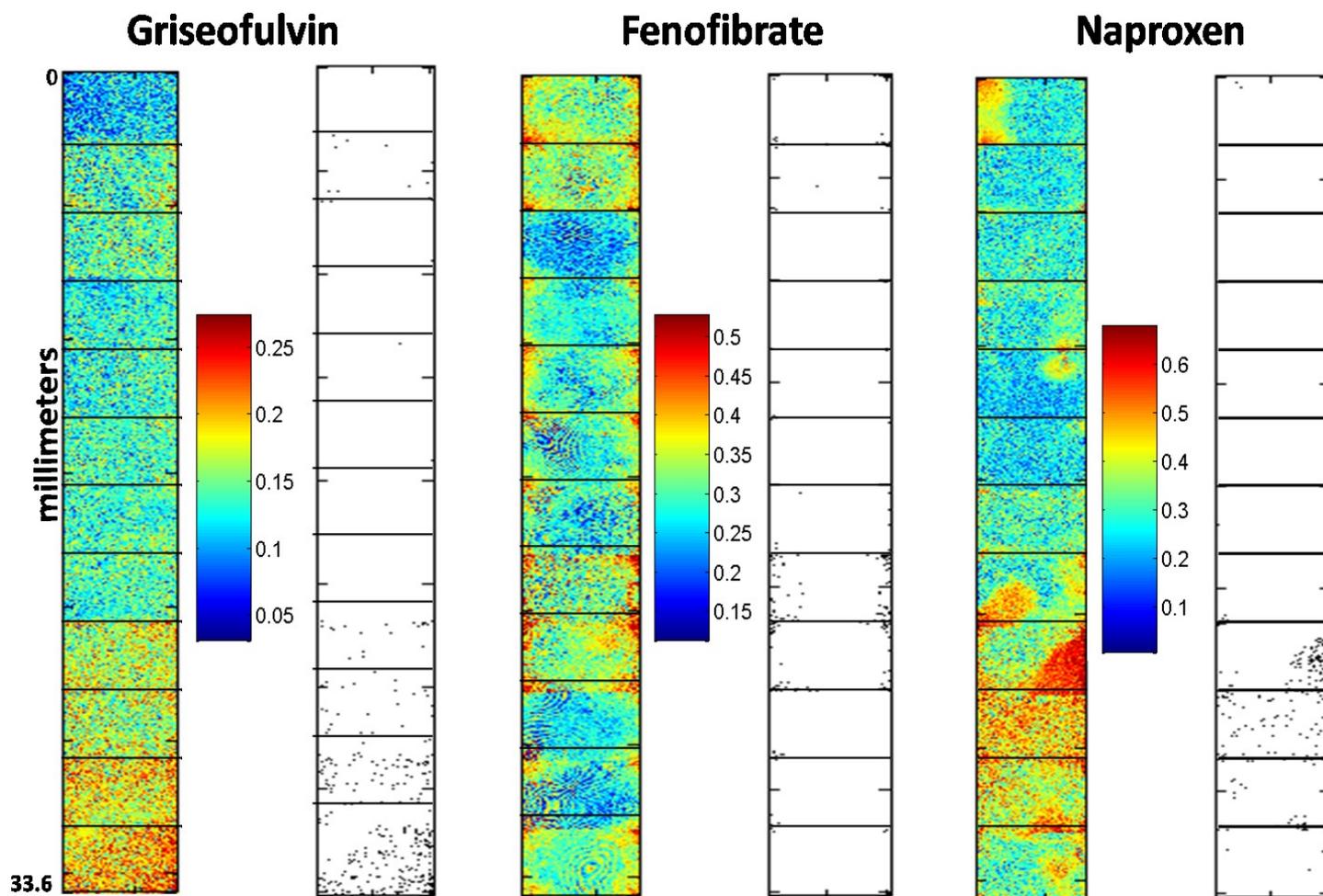


Figure 6.2 API binary images using 10 μm /pixel objective. Domain threshold is set to include only those pixels of highest abundance of API according to the maximum value of each color bar.

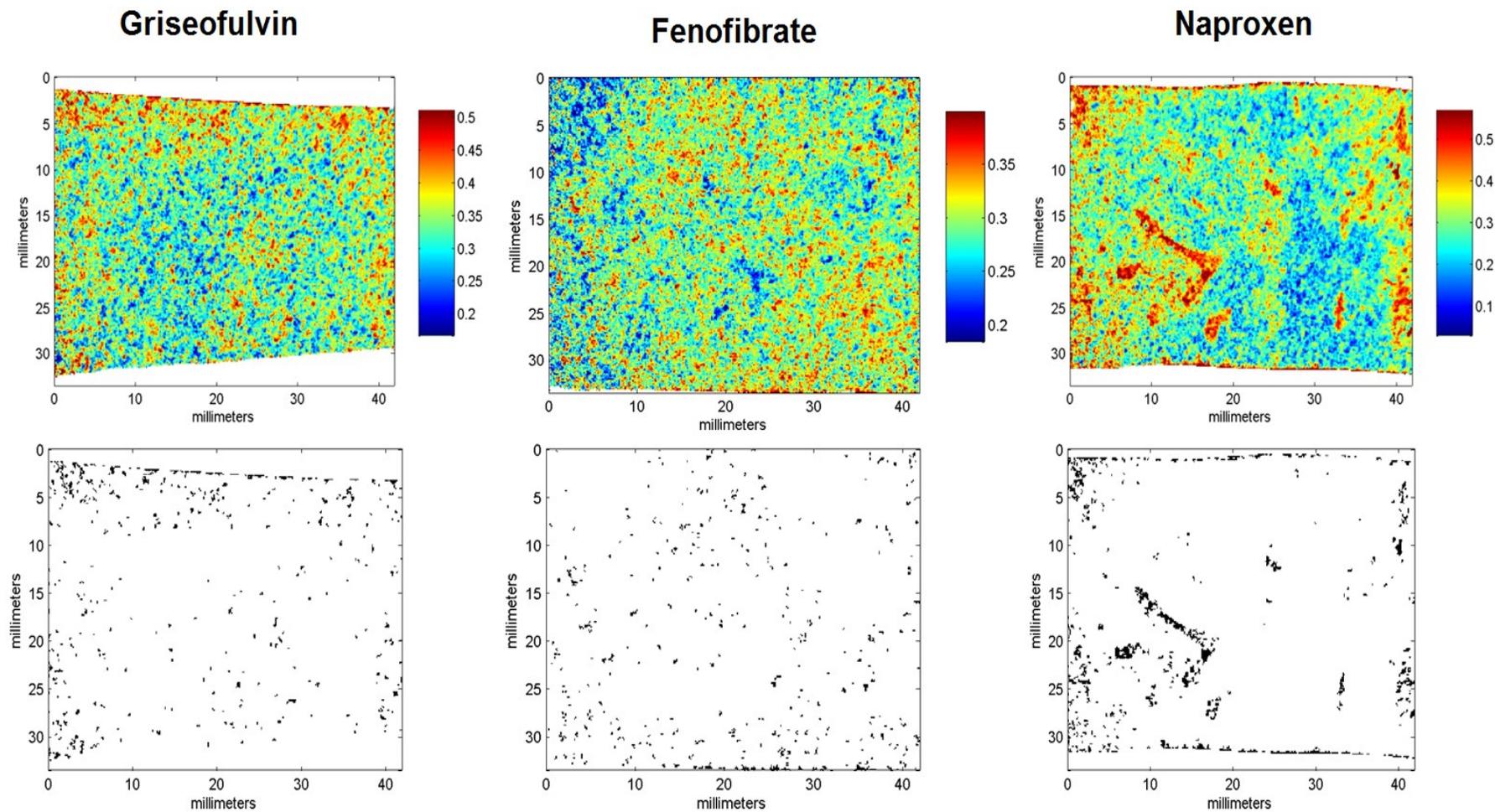


Figure 6.3 API Score images using 131 μm /pixel objective.

The maximum score value observed is about 0.6 (60%) in the NPX film. This means that NPX is more agglomerated by pixel in comparison with the other APIs. The mean value of the scores of API in the entire set of pixels provides the bulk abundance, which estimates the API concentration in the bulk (Lewis et al., 2005). The mean value obtained was 15.0% for GF, 31.3% for FNB, and 23.4% for NPX, as shown in Table 6.2. The lowest mean value obtained was for GF (15%) using the 10 μm / pixel objective. Differences in film thickness and in surface roughness of the films are observed with this objective. This type of physical information affects the results obtained. This was probably the cause for the GF film to exhibit deviations from the nominal value. The other mean values obtained are near the nominal value (27% w/w).

Table 6.2 Distribution of APIs in thin films according to drug score image using NIR-CI with a 10 μm /pixel objective.

	GF		FNB		NPX	
	Mean value	15.0%	Mean value	31.3%	Mean value	23.4%
No. of domains:	7046		3913		4681	
Mean area (mm ²):	2.2 x10 ⁻⁴		4.1x10 ⁻⁴		3.2x10 ⁻⁴	
SD area (mm ²):	6.1x10 ⁻⁴		2.1 x10 ⁻³		4.3x10 ⁻³	
Largest domain (mm ²):	3.3x10 ⁻²		7.5x10 ⁻²		0.265	
Lowest domain (mm ²):	1.0x10 ⁻⁴		1.0x10 ⁻⁴		1.0x10 ⁻⁴	
Area covered:	1.60%		1.65%		1.51%	

Figure 6.2 also shows binary images generated with the pixels of highest abundance of drug, as a way to facilitate the visualization of areas rich in API and to estimate the size of these areas. The pixels with scores that exceeded the mean value + 2 standard deviations were given a value of 1 (dark), and all others a value of zero. Visual examination of these images shows that the NPX film presents the greatest agglomerations of API rich areas. Table 6.2 shows that this film also exhibits the

largest differences in standard deviation and the largest domain found in the three films. Films that have GF and FNB as active ingredients show better distribution of pixels rich in drug.

Table 6.3 Distribution of APIs in thin films according to the drug score images using NIR-CI with a 131 μm /pixel objective.

	GF		FNB		NPX	
	<i>Mean value</i>	34.0%	<i>Mean value</i>	29.3%	<i>Mean value</i>	30.0%
No. of domains:	241		115		464	
Mean area (mm^2):	3.5×10^{-2}		2.7×10^{-2}		8.8×10^{-2}	
SD area (mm^2):	3.4×10^{-2}		2.8×10^{-2}		1.8×10^{-1}	
Largest domain (mm^2):	2.2×10^{-1}		2.1×10^{-1}		1.6	
Smallest domain (mm^2):	1.7×10^{-2}		1.7×10^{-2}		1.7×10^{-2}	
Area covered:	0.71%		0.23%		3.16%	

Figure 6.3 shows the distribution of NPX, GF and FNB throughout the surface of the films using the 131 μm /pixel objective. The maximum score observed is about 0.6 (60%) in the NPX film. The mean value obtained was 34% for GF, 29% for FNB, and 30% for NPX, as shown in Table 6.3. The mean values of the bulk abundance are all very close to the targeted 27% (w/w) film composition. Table 6.3 also reveals differences in the standard deviation of the score values. NPX had the highest standard deviation in the score values, and thus the greatest variation in API distribution throughout the film. The FNB film has the mean value closest to the nominal value, as indicated in Table 6.3, and the lowest standard deviation. Additionally, this film shows the lowest difference between the mean value and the minimum and maximum values observed in the PLS-DA classification scores. These results indicate that this film has the best API distribution through the analyzed surface in comparison with the other two films.

Agglomeration in NPX was also observed in SEM images.⁷ This aggregation of particles might have occurred during film drying. This behavior could be attributed to

the strong hydrogen bonding of the carboxylic group in NPX.^{71,72}

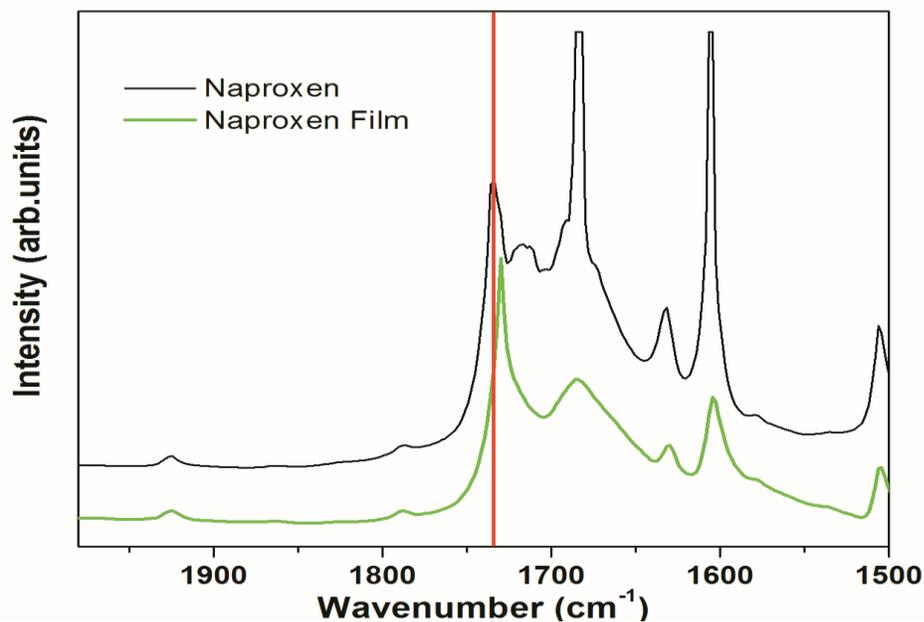


Figure 6.4 FT-IR spectra of NPX and NPX film.

Figure 6.4 shows the spectra for NPX and NPX film. The top spectrum of pure NPX exhibits an infrared band at 1730 cm^{-1} attributed to the free or non hydrogen bonded carboxylic group(monomer) and a band at 1684 cm^{-1} corresponding to the hydrogen bonded carboxylic group (dimer). In the film, the monomer peak decreases in intensity. The reduction in intensity of the monomer peak suggests the involvement of the unassociated NPX in hydrogen bonding with the other components of the pharmaceutical formulation. Additionally, the band at 1730 cm^{-1} shifts to a lower wavenumber, indicating the presence of intermolecular hydrogen bonds between NPX and the other components. This observation confirms the effect of the hydrogen bonding in the NPX film.

6.6 Conclusions

Results showed that the aggregation behavior is dependent on the drug used. The aggregation could be due to strong hydrophobic interactions of drug particles. The difference in aggregation also has a strong effect on the drug content in the films. The strong hydrogen bonding interactions between NPX and other excipients were evidenced with the presence of larger agglomerations of API rich areas.

The analytical method developed can be used to investigate film formulations containing different poorly soluble drugs. In this type of formulations it is very important evaluate the drug distribution through of film surface. Drug distribution impacts drug content of the polymeric film formulation. Polymeric films have a potential applicability on pharmaceutical industry. Film formation may offer a more economical and scalable option if compared to several ways of making solid dosage forms from drug nano-suspensions such as freeze drying, lyophilization, and spray drying.

The drug loading achieved in this type of formulation is higher than for current products that contain drug nanoparticles. The amount of drug in the film could be increased by increasing the film thickness. These results set the foundation for process development of films containing poor water-soluble drugs, for different drug delivery applications.

CHAPTER 7. Hyperspectral Image Analysis to evaluate the effect of drying on drug distribution.

Submitted to *International Journal of Pharmaceutics*, April 2013.

7.1 Summary

Drying is one of the most important processes in the production of polymeric films. The drying temperature affects the thickness and roughness of the film; therefore it is essential to maintain a uniform drying temperature throughout the film surface. A NIR-CI method was developed to assess the manner in which these parameters affect drug distribution. This chapter describes film image analyses where several parameters related to the drying process, such as drying temperature and air flow, were varied. The results show that the parameters studied do affect drug distribution.

Some of the parameters evaluated were: temperature, air velocity, and film-precursor viscosity. Films were produced using Griseofulvin (GF) as the active component. . Convective drying was carried out in a custom made unit designed to provide uniform heating of polymer-based suspensions. Convective drying helped reduce drying times from about 12 hours to an hour or less, minimizing the aggregation of GF. NIR-CI results indicate that a better API rich area distribution was obtained for films formed using higher viscosity film-precursor.

7.2 Samples

Griseofulvin (Sigma-Aldrich) was utilized as active component. Sodium Dodecyl Sulfate (SDS) (Sigma- Aldrich) and low molecular weight Hydroxyl Propyl Methyl Cellulose (HPMC E15LV) (Dow Chemical) were used as stabilizers. HPMC E15LV was also used as a film former. Glycerin (Sigma-Aldrich) was used as a plasticizer. GF particle size was reduced using a wet stirred media mill (WSMM). All other materials were used without further processing.

In this work, eight films with two different GF concentrations were analyzed. These films were prepared at New Jersey Institute of Technology (NJIT) using a D-Optimal design. The GF was dispersed in HPMC E15LV using SLS as a surfactant and glycerin as a moisturizer. The concentrations and the factors included in the D-Optimal design are shown in Table 7.1, 7.2 and 7.3.

Table 7.1 Factors included in the D-Optimal design and Nominal concentration of the polymeric films.

	Factor 1	Factor 2	Factor 3	Factor 4
	Wet Thickness (micro)	Viscosity (cP)	Air Temperature (° C)	Air flow (m/s)
A1	1000	2400	40	0.6
A3	1000	2400	40	1.5
A7	1000	2400	60	0.6
A9	1000	2400	60	1.5
B1	1000	6200	40	0.6
B3	1000	6200	40	1.5
B7	1000	6200	60	0.6
B9	1000	6200	60	1.5

The polymer solution was prepared by adding weighed amount of HPMC and glycerin to water at 90 °C. The solution was then allowed to cool down to room temperature while being stirred continuously. The resulting solution was then let to rest overnight until no air bubbles were seen. This polymer solution was then added to the nanosuspension produced from WSMM in 2:1 ratio and mixed for 4 h. The resulting suspension was left to rest at room temperature for a period of 30 min until no bubbles were observed. Two compositions, A and B, were formed by mixing about 12% HPMC E15LV and 15% HPMC E15LV solutions with GF nanosuspension. The resulting viscosities for the two suspensions used in this study were 2400 cP and 6200 cP respectively. Approximately 6 grams of the final viscous suspensions were then cast manually onto a stainless steel substrate. The casting thickness was set at 1000 µm. The dimensions of wet films cast were kept at about 8cm x 9cm. The film was dried using a convective drying unit.

Table 7.2 Composition A: Mixed in (2:1) ratio (12% HPMC E15LV+5% Glycerin: API nanosuspension)

Film composition	
	w/w
GF	19
HPMC E15LV	58
SDS	1
Glycerin	22

Table 7.3 Composition B: Mixed in (2:1) ratio (15% HPMC E15LV+5% Glycerin: API nanosuspension)

Film composition	
	(w/w) %
GF	17
HPMC E15LV	62
SDS	2
Glycerin	19

Convective drying was carried out in a custom made unit designed at NJIT. The unit is equipped with a heating element, an air blower with a heating coil, a rotating vane anemometer, and an automated balance connected to the sample stage. It is controlled by a computer-based acquisition system which records the sample weight every 15 seconds during drying. The drying process was reduced to an hour or less using this convective drying unit.

7.3 Instrumentation

Near infrared Hyperspectral images were acquired using the Malvern SyNIRgi NIR-CI System (Malvern, UK). The images were obtained in transreflectance mode by placing the film or gel strips over a white ceramic disk of a 28 mm diameter. The entire film was also analyzed using the 131 μ m/pixel objective which has a field view of 34 x 42 mm. For better visualization of the results, the images are shown concatenated. Spectra were obtained with 1 scan in the spectral range of 1300-

2300 nm. The data collected was analyzed using the ISys software (Version 5.0.0.14).

7.4 Data treatment

The spectra acquired by the NIR-CI System are reflectance spectra. Therefore, the logarithm, $\log(1/R)$ was first applied to the data cube to convert the spectra to absorbance units. Pixel correction was applied in order to remove areas of non-uniform illumination and to remove the effect of unresponsive pixels on the detector. The spectra were normalized using Standard Normal Variate (SNV) and a Savitzky-Golay second derivative (filter order 3, filter width 7) in order to eliminate variations in slope. The spectral range was restricted to 2000-2300 nm. Pure component imaging data was collected from powder samples of the API and the excipients where spectral pretreatment was performed in the same manner.

A reference library with the pure components spectra was built. The pure component training spectra were used as predictors to build a PLS-DA in the classification model. The intensity of each pixel in the resultant scores image is determined by the degree of membership (scaled from 0 to 1). The variation in pixel intensity is related to the variation in GF abundance across the sample. The abundance values obtained were used to create the score images shown in this report. These score images were evaluated morphologically using particle statistics tools. The ISys software (Version 5.0.0.14) has a complete particle statistics package for determining particle/domain size and distribution of individual chemical species within chemical images.

7.5 Results

NIR chemical imaging was used to study the spatial distribution of the GF throughout the 34 x 42 mm films. The drug distribution within the film was evaluated with the abundance values determined with the PLS-DA. This algorithm was used to create all score images shown in Figure 7.1. This figure divides the films analyzed into three columns; each column representing one of the

preparation conditions evaluated. The preparation conditions evaluated were viscosity, air temperature, and air flow. The films were concatenated to observe the impact of each one of these preparation conditions. The criteria used for the concatenations are based in the similar conditions of their preparation. For example, the first score image of the column of air flow shows the concatenated films A1 and A3. These two films were prepared with the same viscosity and with the same air temperature, but with different air flow. This condition (air flow) is different in each concatenated score image, as shown in the first column. This same criterion was used to create the other two columns (one related to viscosity and the other to temperature).

Each column shows concatenated scores images with their respective score values which is an indicator of the abundance of GF. The top of the scale (red color in the colorbar) indicates the highest contributor of GF to any pixel. The maximum percentage of GF in each pixel ranges between: 0.25 (25%) -0.58 (58%). Pixels rich in GF are observed in each one of the analyzed areas.

Areas of pure GF have an abundance value of 1.0 or 100%. Therefore, pixels showing areas of pure GF are not observed. The mean value obtained indicates the abundance of GF in each polymeric film. The mean values shown in Table 7.4 are based on the abundance values of more than 80,000 pixels (encompassed within three standard deviations from the mean). This mean value is therefore comparable to the concentration value that could be obtained with a single standard NIR spectrum. Mean values, range, and standard deviation of different films are presented in Table 7.4. The mean values obtained were close to the nominal values. The standard deviation values also help to explain the variation in GF abundance in the films.

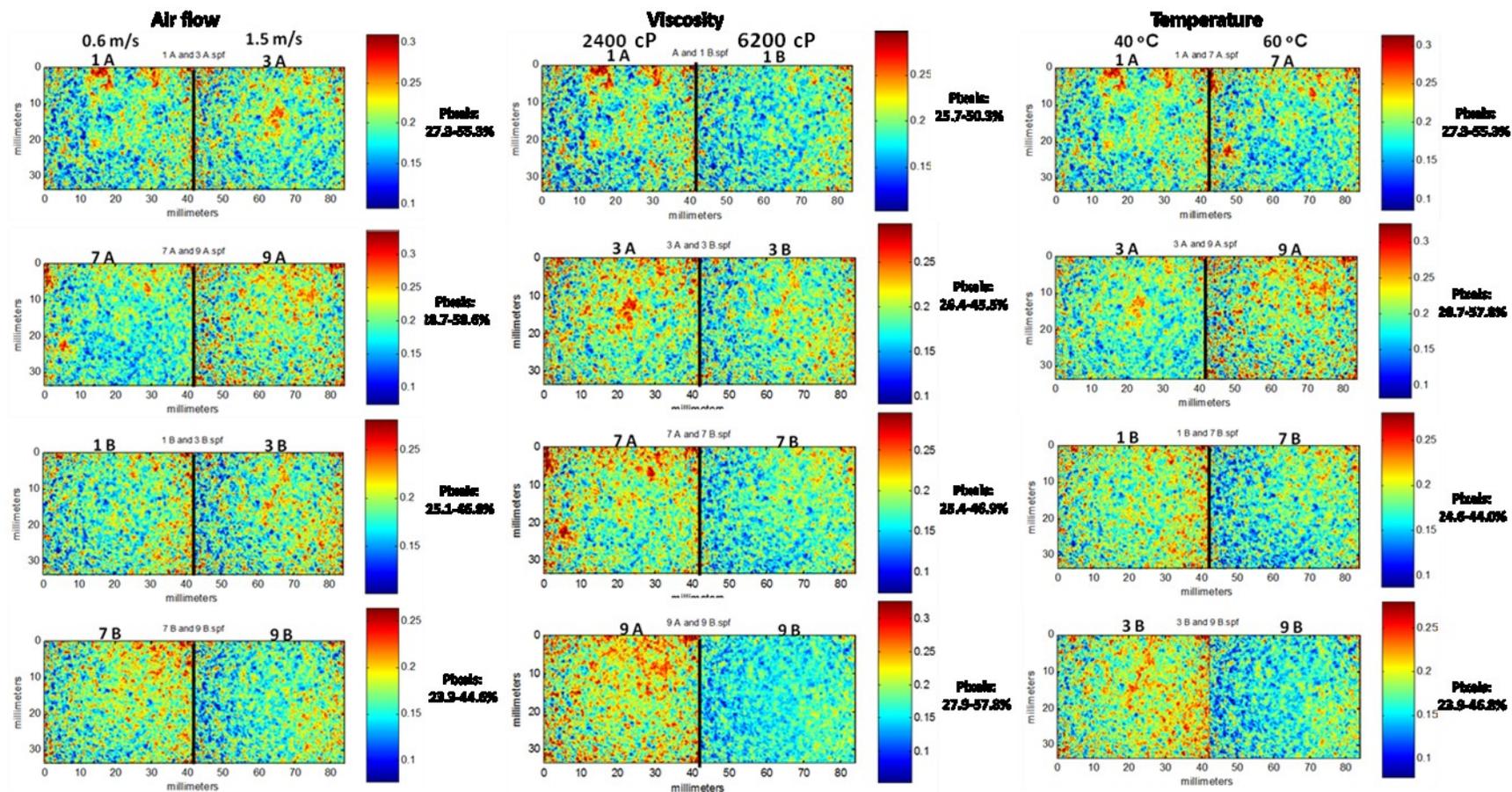


Figure 7.1 PLS predicted API images are displayed on the same scale color using the 131 μm /pixel objective. The images show how the variation in air flow (m/s), viscosity (cP), and temperature ($^{\circ}\text{C}$) affects the API distribution. Pixel values used to create binary images are shown on each colorbar. Binary images were created for pixels where the intensity is greater than the mean value + 2 standard deviations.

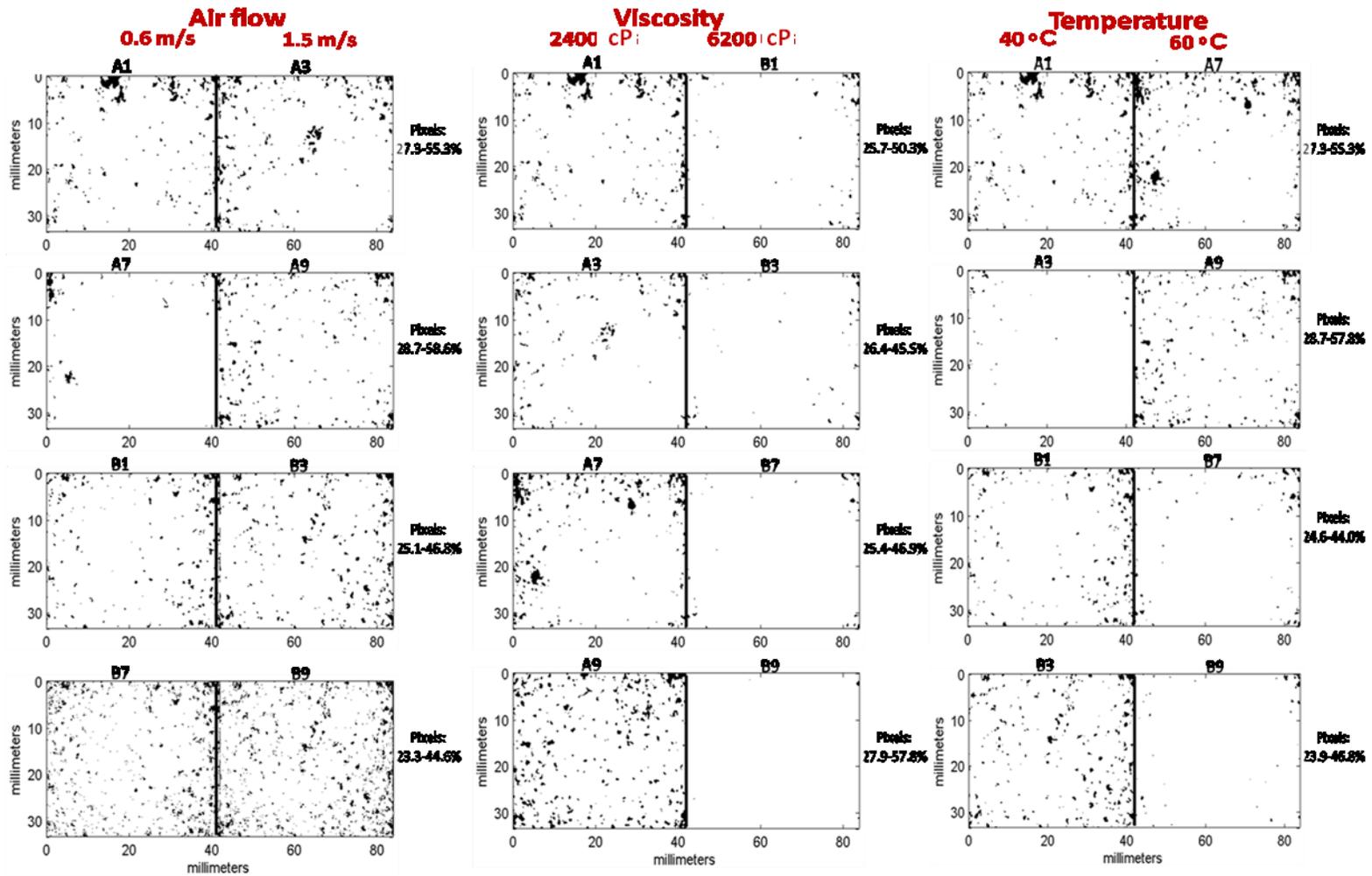


Figure 7.2 Binary image generated from the scores images using the 131 μm /pixel objective. A threshold was established such that only those pixels with values beyond + 2 standard deviations of the mean are shown.

Table 7.4 Nominal and predicted values of the different films analyzed.

Film	Nominal value	131µm/pixel			
		Predicted value	Range		STD
			Lowest	Highest	
1A	19	19.6	3.16	43.1	0.039
3A	19	19.8	4.28	39	0.036
7A	19	19.4	1.16	46.5	0.040
9A	19	21.2	1.41	54.1	0.047
1B	17	19	2.59	39.8	0.031
3B	17	19.1	1.53	47.9	0.034
7B	17	17.4	1.33	39.1	0.033
9B	17	16.4	1.59	40.4	0.031

Table 7.5 Domain Size Statistic Results of API based on binary images generated from scores images.

<u>Air flow (m/s)</u>	2400 cP - 40 °C		2400 cP - 60 °C		6200 cP - 40 °C		6200 cP - 60 °C	
	0.6	1.5	0.6	1.5	0.6	1.5	0.6	1.5
Films	1 A	3 A	7 A	9 A	1 B	3 B	7 B	9 B
Number of Drug Rich Particles	185	185	61	208	225	264	809	885
Mean Area (mm ²)	0.322	0.273	0.326	0.217	0.235	0.228	0.108	0.110
Area STD (mm ²)	0.905	0.405	0.674	0.427	0.489	0.352	0.375	0.287
Largest domain size (mm ²)	9.49	2.78	3.04	4.31	4.61	3.36	9.35	6.11
Lowest domain size (mm ²)	0.0172	0.0252	0.0257	0.0172	0.0172	0.0267	0.0172	0.0273

* Largest clusters are in bold.

— Hyperspectral Image Analysis to evaluate the effect of drying on the drug dispersion. —

	40 °C - 0.6 m/s		40 °C - 1.5 m/s		60 °C - 0.6 m/s		60 °C - 1.5 m/s	
Viscosity (cP)	2400	6200	2400	6200	2400	6200	2400	6200
Films	1 A	1 B	3 A	3 B	7 A	7 B	9 A	9 B
Number of Drug Rich Particles	186	71	147	66	198	48	317	20
Mean Area (mm ²)	0.317	0.191	0.228	0.138	0.358	0.159	0.314	0.083
Area STD (mm ²)	0.905	0.305	0.296	0.172	1.05	0.165	0.576	0.108
Largest domain size (mm ²)	9.49	1.7	2.04	0.841	11.9	0.755	6.57	0.515
Lowest domain size (mm ²)	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172

	2400 cP - 0.6 m/s		2400 cP - 1.5 m/s		6200 cP - 0.6 m/s		6200 cP - 1.5 m/s	
Temperature (° C)	40	60	40	60	40	60	40	60
Films	1 A	7 A	3 A	9 A	1 B	7 B	3 B	9 B
Number of Drug Rich Particles	180	181	60	207	191	87	259	55
Mean Area (mm ²)	0.333	0.345	0.125	0.216	0.234	0.19	0.236	0.137
Area STD (mm ²)	0.933	0.97	0.147	0.426	0.509	0.226	0.391	0.155
Largest domain size (mm ²)	9.49	10.5	0.704	4.31	5.32	1.2	3.67	0.738
Lowest domain size (mm ²)	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172	0.0172

* Largest clusters are in bold.

Relatively good particle distributions are observed for all film samples. Binary images were created to facilitate the visualization of GF rich areas. These binary images were created by applying a threshold value, where only pixels with intensity greater than the mean + 2 standard deviations are observed. In other words, the distribution of the pixels observed is the one that has the highest API content per pixel.

Figure 7.2 shows the binary images created by applying the threshold value. The dark areas are those with GF abundance higher than the mean + 2 standard deviations. The white areas in the binary images correspond to the GF abundances that did not exceed + 2 standard deviations from the mean. Hence, Figure 7.2 provides qualitative visual information about drug distribution within the film in addition to the variation between films. Qualitatively, images with predominantly white areas indicate absence of drug rich areas and a better overall drug distribution. In images with a higher proportion of white areas, the GF is more uniformly distributed by pixel. Figure 7.2 shows that the most prominent influence is from the difference in viscosity, because all higher viscosity cases exhibited more uniformly distributed white areas. Consequently, qualitative observation suggests that viscosity has the most dominating effect on the distribution of drug.

For particles suspended in a solution, mobility is dominated by Brownian motion. The Brownian motion is regulated by the number of collisions as well as the probability of effective collisions. The collision frequency is a function of diffusion coefficient and particle radius. The collision frequency is inversely proportional to suspension viscosity. At lower polymer concentrations, collisions between hardly or partially covered particles lead to higher collision frequencies and higher probabilities of sticking. However, as polymer loading increases, the agglomeration decreased due to a reduction in collision rates.¹¹

Table 7.5 shows the results of the statistical analysis of binary images from Figure 7.2. The distribution of the pixels rich in GF is different in each one of the score images. The largest agglomerates have a size of 6.11mm² - 11.9 mm². Most of the agglomerates (GF rich areas) were found in films with low viscosities, thus indicating

that the viscosity is the factor with the major effect on the distribution of GF rich areas. Some tendencies observed using the 131 μm /pixels objective are:

- With a viscosity of 6200 cP, the air flow and air temperature does not affect the distribution of these pixels. At this viscosity, the drug is distributed uniformly per pixel.
- The best distribution in films with high viscosity was observed when working at 60 °C.
- The best distribution in films with minor viscosity was observed when working with an air flow of 0.6 m/s and a temperature of 40 °C.

The drying times was reduced from about 12 h to 1 h or less using the convective drying unit, this unit allow controlled the evaporation process. The evaporation is a diffusion process and when it occurs rapidly causes the agglomeration of GF particles. The evaporation also generates a flow through the surface of polymeric films that involves changing from a liquid to solid state. This flow causes gradient (horizontally and vertically) across of the film surface. If the air surrounding the drying film is flowing then evaporation is enhanced by convection. The convective drying unit was designed such that the air flow pattern is laminar and the air temperature over the sample stage does not vary across the chamber cross-section.

In summary, higher viscosity films have fewer GF rich areas, and their average sizes are smaller. This suggests that although there are GF rich areas in all cases, the distribution is much better for the higher viscosity cases. The lowest average size generates a viscosity of 6200 cP, a temperature of 60°C, and an air flow of 1.5 m/s. The lower viscosity has the largest domain and the bigger number of GF rich areas for each preparation condition. Large drug rich areas have a low viscosity while small drug rich areas have a high viscosity, further suggesting the improvement of distribution by using higher viscosities.

7.6 Conclusion

Drying of polymer films is one of the most important parameters to produce polymeric films with reproducible drug content. The drug content depends on the drug distribution throughout the film surface. An analytical imaging method was developed with the objective to understand how the drying parameters affect the drug distribution. NIR-CI evaluation helped to identify the best conditions of the drying process to obtain films with a good GF rich areas distribution. NIR-CI helped overcome API irreversible agglomeration, one of the problems of pharmaceutical film formulations in the drying process. This agglomeration affects the API content throughout the film surface. The film with the best drug distribution is obtained when using a viscosity of 6200 cP, a temperature of 60 °C, and an air flow of 1.5 m/s. The polymeric films prepared with a viscosity of 2400 cP have the highest number of GF rich areas and present the largest agglomerates. In contrast, the film prepared with a viscosity of 6200 cP presents the smallest number of drug rich areas, the smaller domains, and the best drug rich area distribution.

Chapter 8. Development of Raman and Near Infrared (NIR) Methods in a pharmaceutical chemical plant.

8.1 Objective

The objective of the following chapter is to present the research performed during the internship in a pharmaceutical chemical plant.

8.2 Summary

NIR calibration models were developed to determine the amounts of toluene and ethyl acetate in recovered toluene and the amount of *alcoholic monomethylamine* (MMAA) in reaction completion of an arylation. A feasibility study was conducted using Raman Spectroscopy as an alternate technique for the ID analysis of raw material. NIR and Raman techniques have advantages over the current methods, such as: *High-performance liquid chromatography (HPLC) and gas chromatography (GC)*; because sample preparation is not required since the analysis is non-destructive.

This chapter is divided in three sections, according to the original objectives of the internship: Perform a preliminary feasibility study to use portable Raman technology for the ID testing of raw and starting material used at manufacturing site; Perform a feasibility study for use NIR to monitor process reaction completion and Perform a feasibility study for implementation of NIR for solvent recovery process.

8.3 Perform a preliminary feasibility study to use portable Raman technology for the ID testing of raw and starting material used at manufacturing site

8.3.1 Problem

The requirement to perform identification for all raw materials and starting materials received at the manufacturing site significantly increases the lead time and work load of the Quality Control (QC) laboratory. An alternate analytical technique needs to be implemented to simplify and streamline the ID testing process.

8.3.2 Background

8.3.2.1 Raman Spectroscopy

Raman spectroscopy is a spectroscopic technique used to observe vibrational, rotational, and other low-frequency modes in a system. It is predominantly applicable to the qualitative and quantitative analyses of covalently bonded molecules. Raman scattering occurs only when the *net polarizability* of the molecule is altered during the vibration. Qualitatively, Raman spectroscopy measures changes in the electron cloud of the molecules. Covalent bonds are more polarizable than ionic bonds and the intensity of the vibration increases with bond order. A Raman spectrum is a plot of the intensity of Raman scattered radiation as a function of its frequency difference from the incident radiation (usually in units of wavenumbers, cm^{-1}). This difference is called the *Raman shift*. Wavenumber $\Delta\nu$ shift is defined as the difference in wavenumbers (cm^{-1}) between the observed radiation and that of the source.

$$\Delta\nu \text{ Raman shift} = \nu_{\text{Laser}} - \nu_{\text{Scattered}} \quad (8.1)$$

Shifts in wavenumber depend on the chemical structure of the molecule responsible for the scattering. Raman spectroscopy is good for fingerprinting, probing molecular symmetry and is also a non-destructive method of analysis.⁷³

8.3.3 Methodology

Raman spectra were acquired for each reagent and compared with a similar structure reagent in order to perform an exploratory analysis and evaluate the spectral differences (selectivity).

8.3.3.1 Instrumentation

Raman spectra were acquired using a Raman RXN1-RA-785 system from Kaiser Optical Systems (Ann Arbor, MI) and the PhAT™ probe accessory that was installed in vertical position. The laser wavelength in the system is 785 nm, with spectral coverage from 150 to 1900 cm^{-1} with a resolution of 4 cm^{-1} , with integration time of 15 seconds per scan. The Raman spectrum was the result of the average of 3 scans. Intensity and wavelength calibration were performed immediately before obtaining the samples, using Kaiser Software version HOLOGRAMS 4.1.

8.3.4 Reagents used in this report

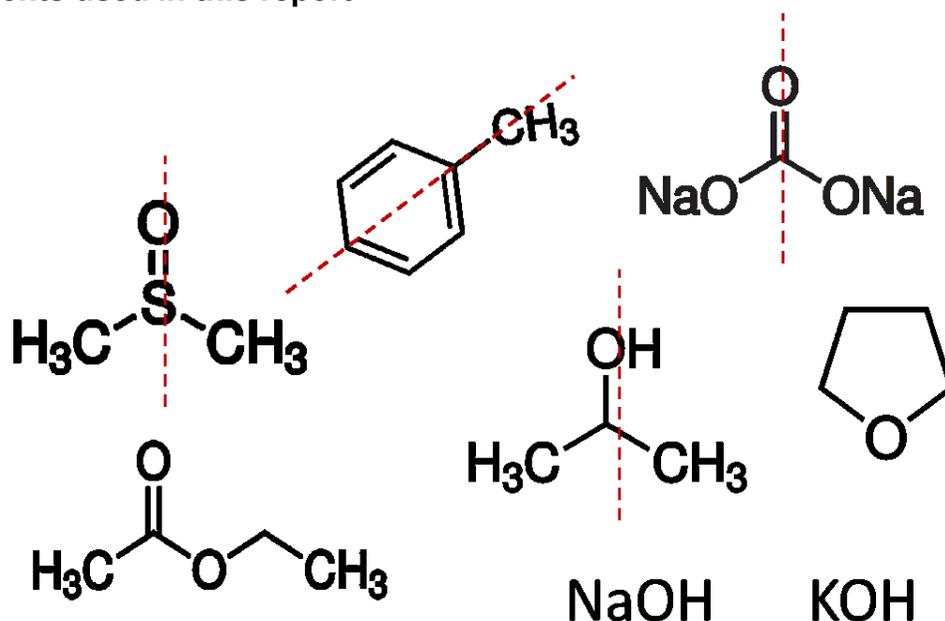


Figure 8.1 Structures of the eight reagents used in this report.

Figure 8.1 shows the chemical structures of the reagents used in this study. Dimethylsulfoxide, Isopropyl Alcohol, Sodium Carbonate, Tetrahydrofuran and Toluene have some level of symmetry which makes easier the study of specificity. Spectra will be different if changing the chemical structure of these reagents; causing the appearance of a new vibrational mode.

8.3.5 Results

Table 8.1 show the raw materials that are being used actually and the technique used for ID analysis. Based on chemical structure of the starting materials, it is very likely that Raman technique will provide the desired selectivity and performance. All the reagents are excellent candidates for analysis by Raman technique, except NaOH and KOH, whose signals are weak. This technique can be employed as additional method for identification in order to reduce the number of drums to be tested with registered analytical method. In order to change the registered ID method, regulatory change must be pursued.

Figures 8.2-8.8 show the Raman spectra of the raw materials. Spectral differences can be observed among raw material with similar chemical structure. Raman spectroscopy is very susceptible to changes in chemical composition and structure. The alteration of the atoms in a molecule causes changes in their vibrational frequencies; therefore, the Raman spectra also change.

Table 8.1 Raw materials of the pharmaceutical chemical plant.

Raw Materials	Current ID analysis
1-Fluoronaphthalene	Identification of 1- Fluoronaphthalene by Infrared Spectroscopy
Potassium Hydroxide	Identification of Potassium (USP/NF)
Sodium Hydroxide	Identification of Sodium Test – USP
Dimethylsulfoxide	Identification of Dimethyl Sulfoxide by GC
Ethyl Acetate	ID of Ethyl Acetate by GC
Isopropyl Alcohol	Identification of Isopropyl Alcohol by GC
Monomethyl amino alcohol	Identification of Monomethyl Amino Alcohol by IR Spectroscopy
Monomethylamine	Identification of monomethylamine by FTIR spectroscopy.
Sodium Carbonate	Carbonate Identification Test – USP
Tetrahydrofuran	Identification of Tetrahydrofuran by GC
Toluene	Determination of Toluene by GC
Chloroacetyl Carboline	Identification of QA500X using FT-RAMAN

8.3.6 Raman Spectra

8.3.6.1 Sodium carbonate

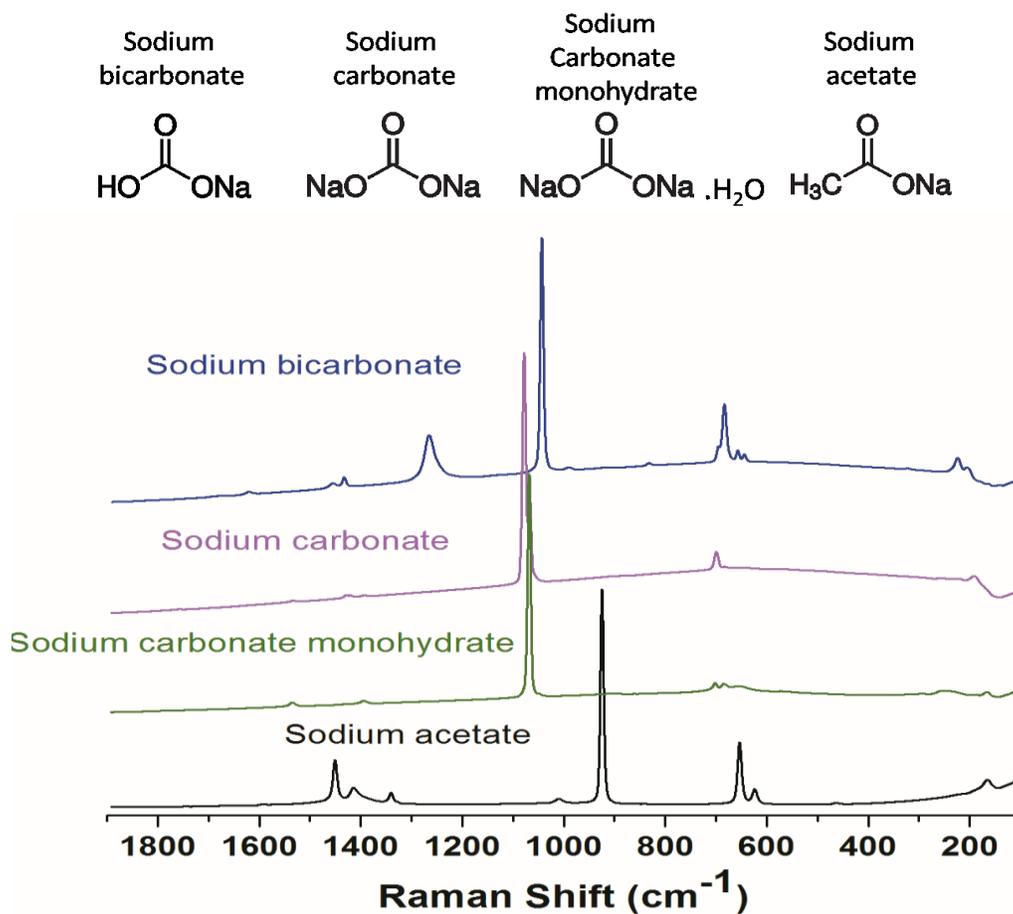
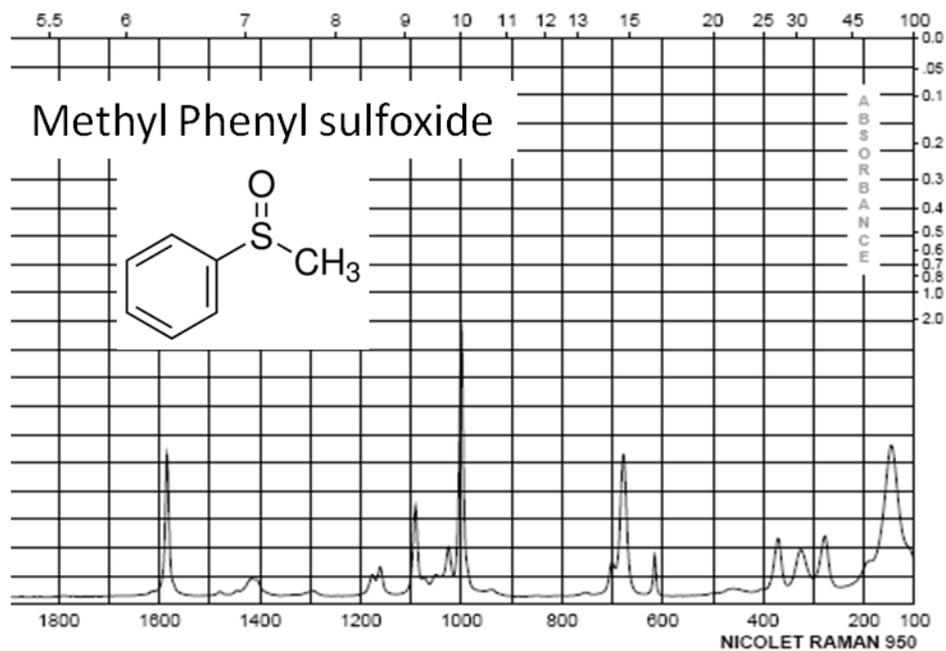


Figure 8.2 Structures and spectra of sodium bicarbonate, sodium carbonate, sodium carbonate monohydrate and sodium acetate.

Sodium carbonate shows an intense Raman vibrational mode at 1081 cm^{-1} , which can be assigned to the ν_1 symmetric stretch.⁷⁴ Additionally, the spectrum presents a vibrational mode at 701 cm^{-1} ; this vibrational mode is a reliable source for their identification.

8.3.6.2 DimethylSulfoxide



Source: Spectral Database for Organic Compounds, SDBS

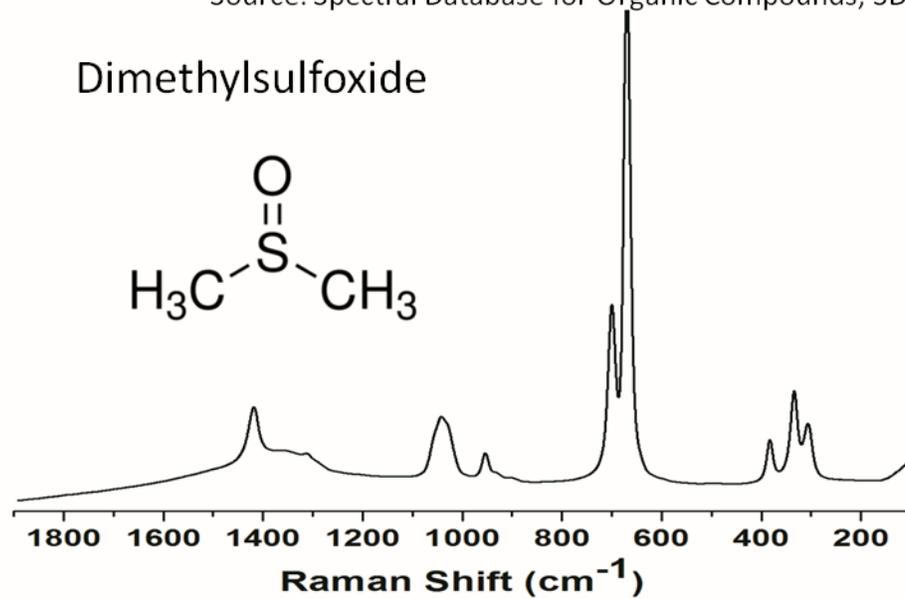


Figure 8.3 Structures and spectra of Methyl Phenyl sulfoxide and Dimethyl Sulfoxide

Dimethylsulfoxide shows three Raman vibrational modes at 305, 334 and 384 cm^{-1} , which can be assigned to C-S-C bend, C-S-O antisymmetric bend and C-S-O symmetric bend respectively. The 668 cm^{-1} vibrational mode is commonly assigned to the symmetric stretch of the C-S and the 698 cm^{-1} vibrational mode correspond to C-S anti-symmetric stretch. The other three vibrational mode correspond to the vibrations that involving the CH_3 functional group.⁷⁵

8.3.6.3 Ethyl acetate

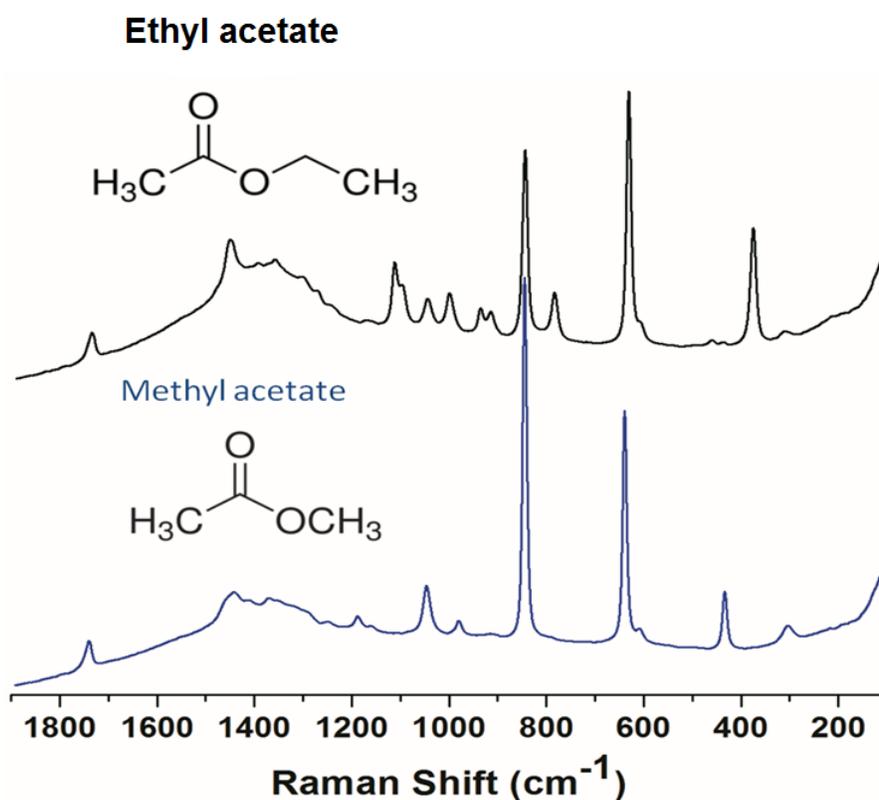


Figure 8.4 Structures and spectra of Ethyl acetate and Methyl acetate

Ethyl acetate has three characteristic vibrational mode at 376, 632 and 845 cm^{-1} ; the 376 cm^{-1} vibrational mode corresponding to C-C bending, the 632 cm^{-1} vibrational mode corresponding to C-C-O bending and the 845 cm^{-1} vibrational mode corresponding to CH_3 rocking.⁷⁶

8.3.6.4 Isopropanol

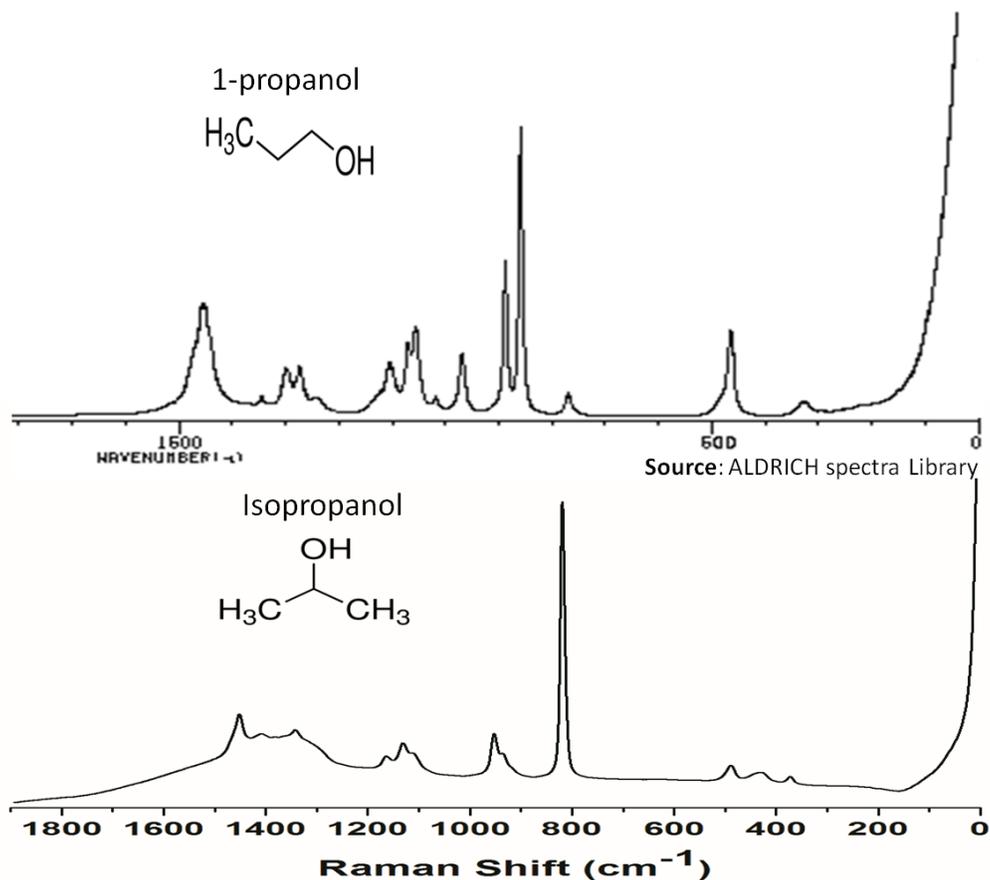


Figure 8.5 Structures and spectra of 1-propanol and Isopropanol

The intense Raman vibrational mode at 817 cm^{-1} corresponds to stretch in phase of C-C-O;⁷⁷ this is the most important mode vibrational in isopropanol. The CH_3 bend (deformation) is near 1465 cm^{-1} .

8.3.6.5 KOH

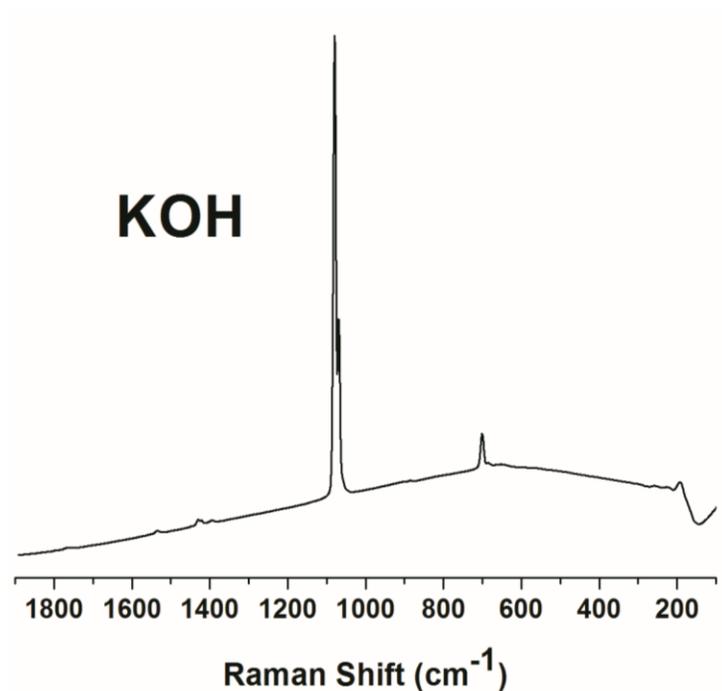


Figure 8.6 Spectrum of KOH

Each molecular unit has a preferred orientation due to interactions with the nearby molecules; additionally also has librational modes corresponding to small rotations about this preferred orientation. The librational mode for KOH occur in the region of 700-1200 cm⁻¹.⁷⁸

8.3.6.6 Tetrahydrofuran

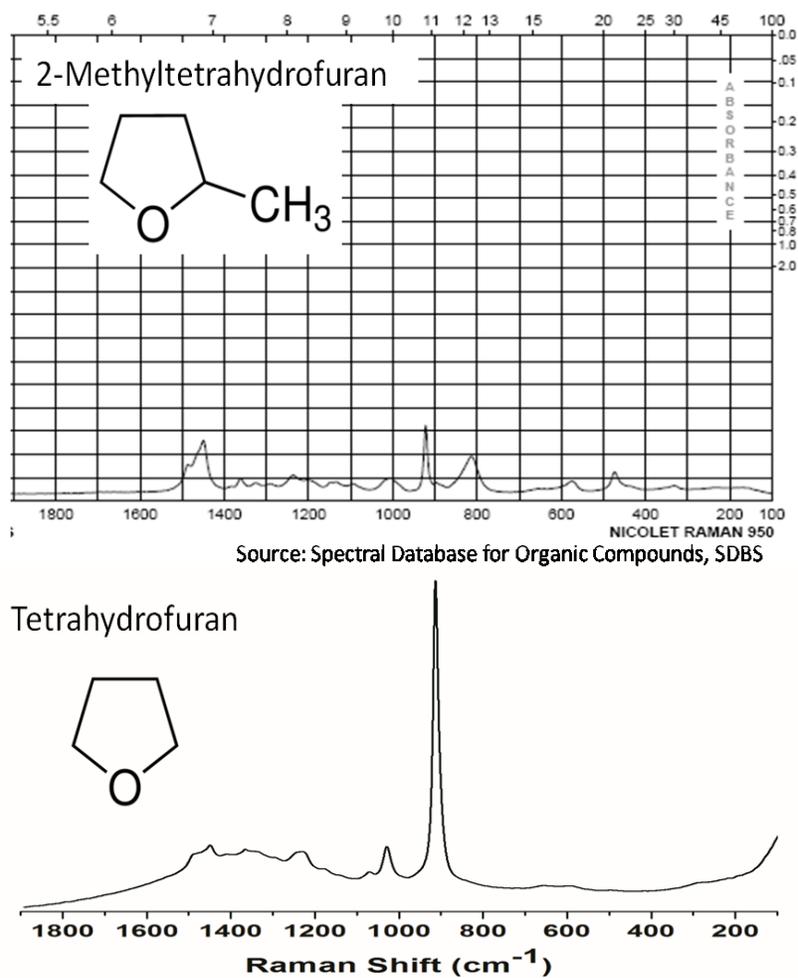


Figure 8.7 Structures and spectra of 2-methyltetrahydrofuran and Tetrahydrofuran

This spectra show an intense band at 913 cm^{-1} that corresponds to ring breathing and at 1027 cm^{-1} that corresponds to ring stretch.⁷⁹

8.3.6.7 Toluene

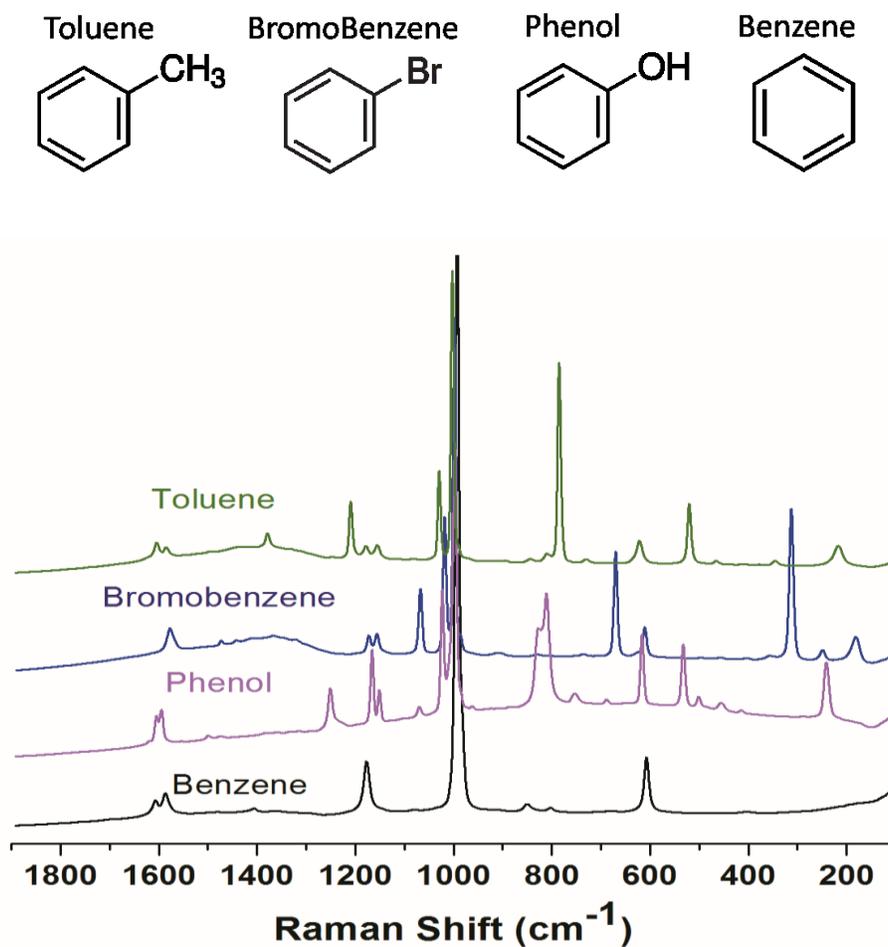


Figure 8.8 Structures and spectra of Toluene, Bromobenzene, Phenol and Benzene

Toluene spectrum shows the ring deformation vibrations at 1000-1040 cm^{-1} , also present the out-of-plane vibrational band at 786 cm^{-1} .⁸⁰

8.3.7 NIR Spectra

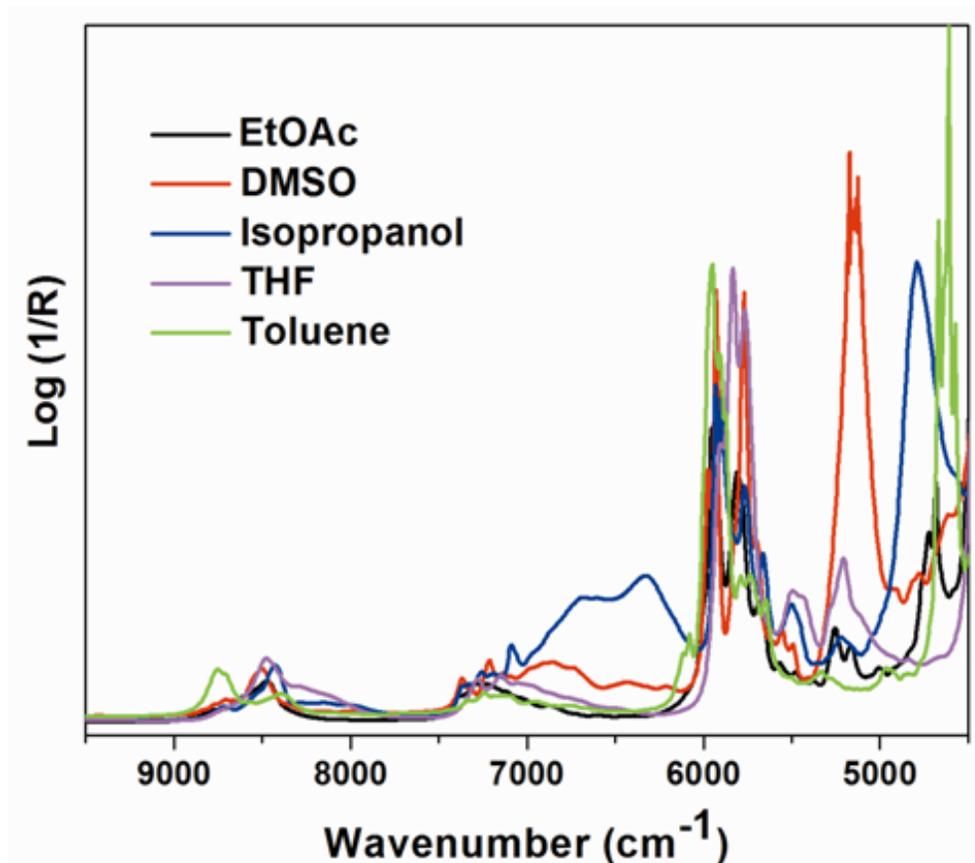


Figure 8.9 NIR Spectra of Ethyl acetate (EtOAc), Dimethylsulfoxide (DMSO), Isopropanol, Tetrahydrofuran (THF) and Toluene.

Figure 8.9 shows the NIR spectra of the different reagents. Spectra show overlapping bands; therefore it is difficult to differentiate between the spectra. Sodium Carbonate and KOH do not show NIR spectra. Table 8.2 and 8.3 summarizes the correlation between spectra using the two techniques. The correlation between the Raman spectra is low, this was the technique selected for implement as alternative technique for ID analysis of the pharmaceutical chemical plant.

Table 8.2 Correlation between Raman spectra.

	DMSO	EtOAc	Isopropanol	KOH	Na ₂ CO ₃	THF	Toluene
DMSO	1						
EtOAc	0.099	1					
Isopropanol	0.075	0.440	1				
KOH	0.061	0.065	-0.159	1			
Na ₂ CO ₃	0.196	0.071	0.046	0.339	1		
THF	0.050	0.346	0.291	0.026	0.056	1	
Toluene	0.017	0.251	0.148	-0.051	0.041	0.118	1

Table 8.3 Correlation between NIR spectra.

	<i>EtOAc</i>	<i>DMSO</i>	<i>Isopropanol</i>	<i>THF</i>	<i>Toluene</i>
<i>EtOAc</i>	1				
<i>DMSO</i>	0.651	1			
<i>Isopropanol</i>	0.686	0.454	1		
<i>THF</i>	0.751	0.610	0.419	1	
<i>Toluene</i>	0.778	0.452	0.529	0.453	1

Figure 8.10 show a scores plot of the Raman spectra, the main source of variation is the physical state of the reagents.

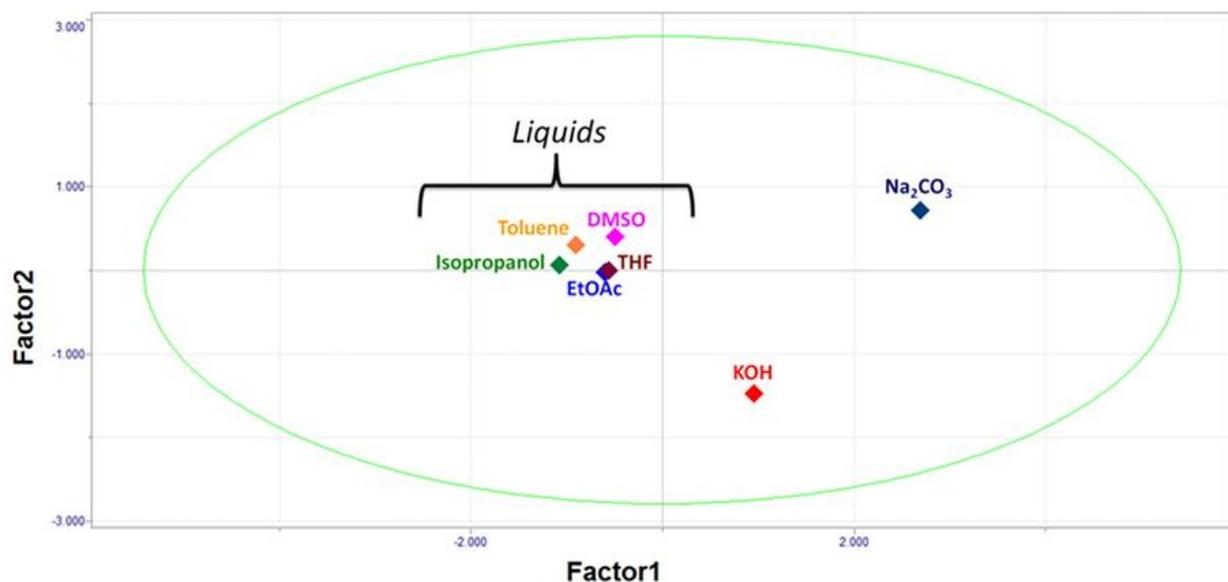


Figure 8.10 Score plot using Raman Spectra.

8.3.8 Future Studies and Tasks

The pharmaceutical chemical plant purchased two Thermo Scientific TruScan- Handheld Raman units for Pharmaceutical-Raw Material Identification. The Thermo Scientific TruScan analyzer is a lightweight, handheld instrument for rapid raw material identification. This equipment is easy to use; its non-destructive point-and-shoot operation enables material identification through sealed packaging to minimize risk of exposure and contamination. It is used for incoming raw material identity verification, process troubleshooting, counterfeit identification etc.

The resolution that offers the TruScan RM is sufficient to differentiate the spectra of the reagents used in PR06 as shown in Figures 8.2-8.8 and Tables 8.2 and 8.3; the main vibrational modes are well defined and separate. The spectral range of the equipment covers the range of the main vibrational modes and is very similar to used for the feasibility study. Currently, the units are being qualified and the next step is the development of the method.

8.4 Perform a feasibility study for use NIR to monitor process reaction completion.

8.4.1 Problem

The Arylation process consists of a single reaction step followed by isolation/purification steps. The Arylation is carried out under basic conditions using excess of potassium hydroxide (KOH) flakes at 85 °C over 10-13 hours. The target for reaction completion using HPLC method is NMT 1.0% MMAA remaining (area % relative to Duloxetine). The lead time of HPLC in-process testing for this analysis is approximately 2-3 hours.

8.4.2 Background

8.4.2.1 Near Infrared Spectroscopy

In recent years, the use of Near Infrared (NIR) spectroscopy has become widespread in process analysis and within pharmaceutical industry for raw material testing, product quality control and process monitoring. NIR absorption bands are typically broad, overlapping and 10–100 times weaker than their corresponding fundamental mid-IR absorption bands. NIR spectroscopy is a vibrational spectroscopic method belongs to the infrared light spectrum which is very close to the visible region (from about 750 to 2500 nm), where the most of organic and some inorganic compounds shows good reflectance or transmission properties, therefore exhibiting good absorption of light at the NIR region²⁵ as discussed in Chapter 2 of this dissertation.

8.4.3 Methodology

A NIR Partial Least Square (PLS) predictive model was developed by running arylation reactions at laboratory scale; using the Bruker FT-NIR spectrometer and HPLC to monitor the reaction completion. Samples were obtained at 5, 6, 7 and 10 hours of reaction. Plant samples also were used to create a NIR model. Figure 8.11 shows the set up of the arylation reaction of duloxetine.

8.4.4 Instrumentation

MPA (Multi Purpose Analyzer) FT-NIR spectrometer equipment was used for spectra acquisition. NIR spectra were obtained in transmittance mode using a PbS detector. Each spectrum was an average of 128 scans at a resolution of 8 cm^{-1} , over the range of $12,000 - 3500\text{ cm}^{-1}$. The NIR models were developed using OPUS Quant Software version 6.5.



Figure 8.11 Set up of the arylation reaction.

8.4.5 Results

The first step in developing the NIR model is to select the set up of the equipment. The most important parameters to select are the number of scans and the resolution (128 scans and 8 cm^{-1}); the noise in the spectra depends on these two parameter. Additionally, the spectral range selected should contain the main spectral variation.

For consistent results, it is necessary to analyze the samples at room temperature and to wait for that the excess KOH settle to the bottom of the vial (see Figure 8.12). NIR radiation source is obtained from a tungsten lamp, this lamp may cause an increase in temperature during the analysis, causing fluctuations within the sample. To prevent these

fluctuations affect the results, three spectra are acquired for each sample and the average value was reported.

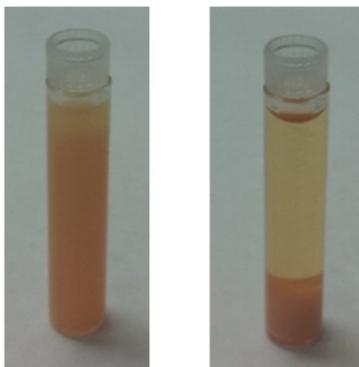


Figure 8.12 Interference of KOH in excess.

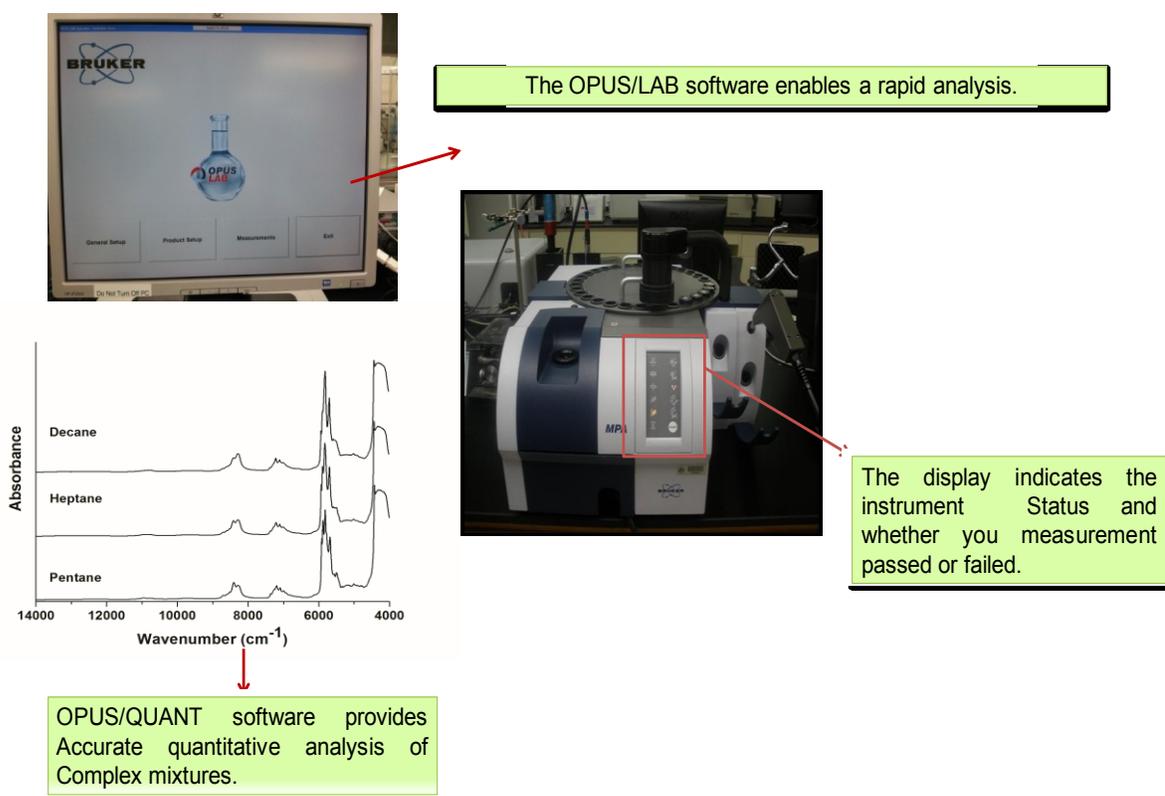


Figure 8.13 Spectra of an arylation reaction.

Figure 8.13 shows some of the advantages of using MPA FT-NIR. Bruker offers several software packages that make its use easier. Two packages used in this work are: OPUS/LAB: Dedicated QA/QC software and OPUS/QUANT: Self-optimizing chemometric quantification software.

The NIR model was created using the OPUS/QUANT software. OPUS/LAB provides the platform to implement the model. The advantage of using the computer software is that the measurement can be made by the operator (a person that is not a chemist and is not thoroughly familiar with the instrument). The operator introduces the date and the batch number. The vial is placed directly on the sample compartment and in few seconds, the results are obtained.

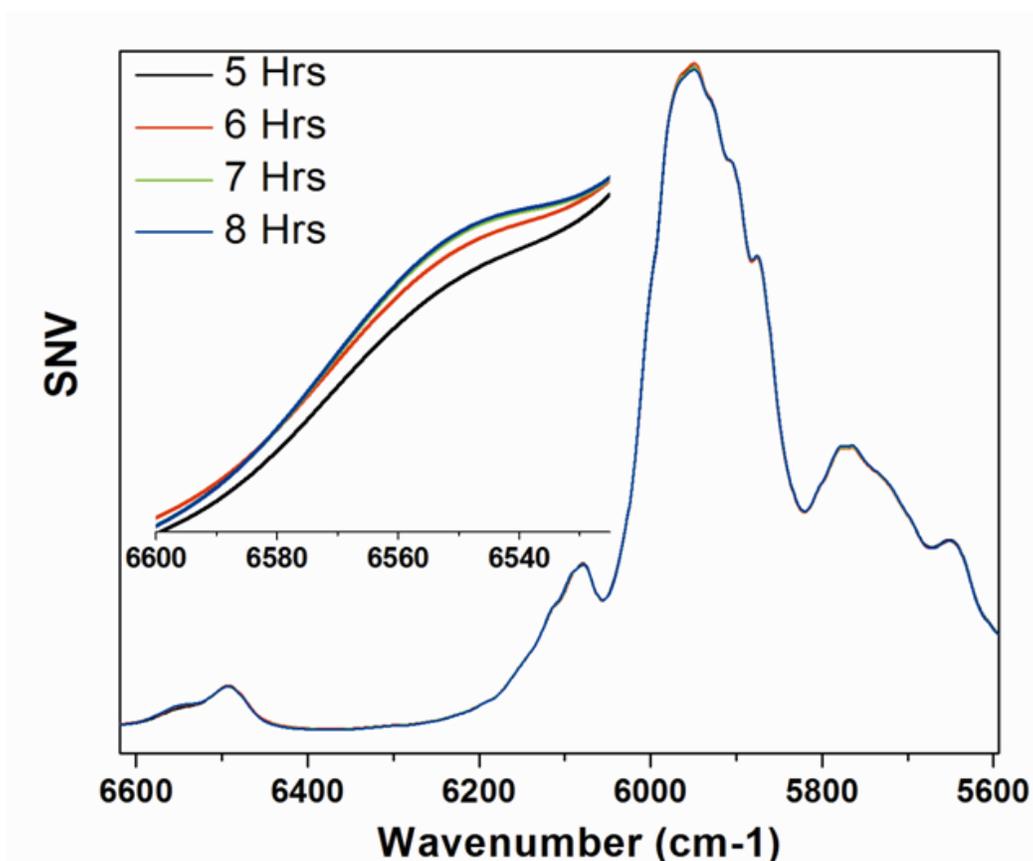


Figure 8.14 Spectra of an arylation reaction. Subtle differences are observed in this spectral range. These differences are increased with the pretreatment.

Figure 8.14 shows the spectra at four different times of the arylation reaction. These spectra show overlap between the main bands. Chemometrics tools are needed with this technique due the overlap between spectra. Chemometrics is the science of extracting information from chemical systems using tools of statistics and mathematics as discussed in Chapter 2 of this dissertation.

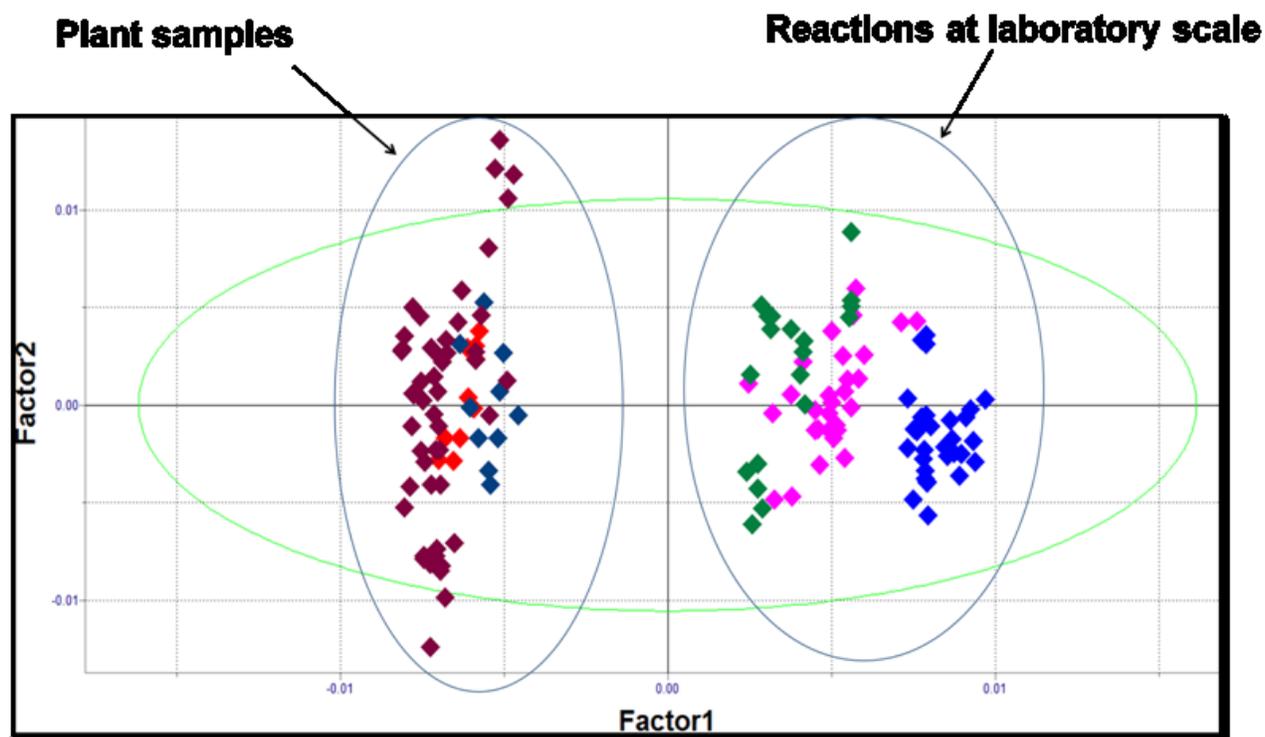


Figure 8.15 PCA of the laboratory and plant samples.

Principal Component Analysis (PCA) and pretreatments are examples of chemometrics tools used in this work. The function of the pretreatment is to increase the spectral differences between spectra. In this case, SNV and first derivative was applied. PCA defines the significant sources of variance within the sample set as a series of ranked components or factors and assigns each spectrum a score based on the relative contribution of each factor. PCA analysis is often a very effective method for highlighting subtle spectral variations between different objects or regions in a sample, also helps

visualize the principal sources of variation in data sets, removes irrelevant or random variation, by retaining only the principal components that capture relevant variation and helps to visualize relationships between samples in a data set.

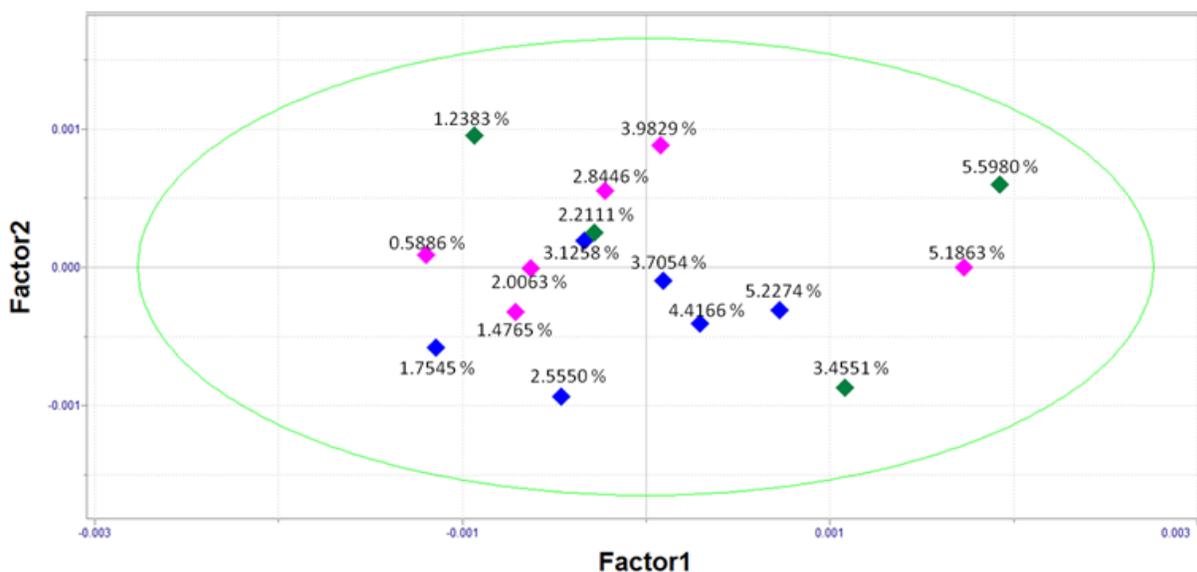


Figure 8.16 PCA of the laboratory samples.

The objective to create a model that includes the greatest possible variation, so the optimum is to create a model that includes plant samples and laboratory samples. Figure 8.15 shows a scores plot. This figure shows that laboratory and plant samples are separated; this indicates that the main source of variation is the origin of samples and not the changes in concentration of the target. Probably, these differences are because one of the components of the arylation was used with a purity different to that used in the plant process. Therefore the NIR model was created using only the plant samples.

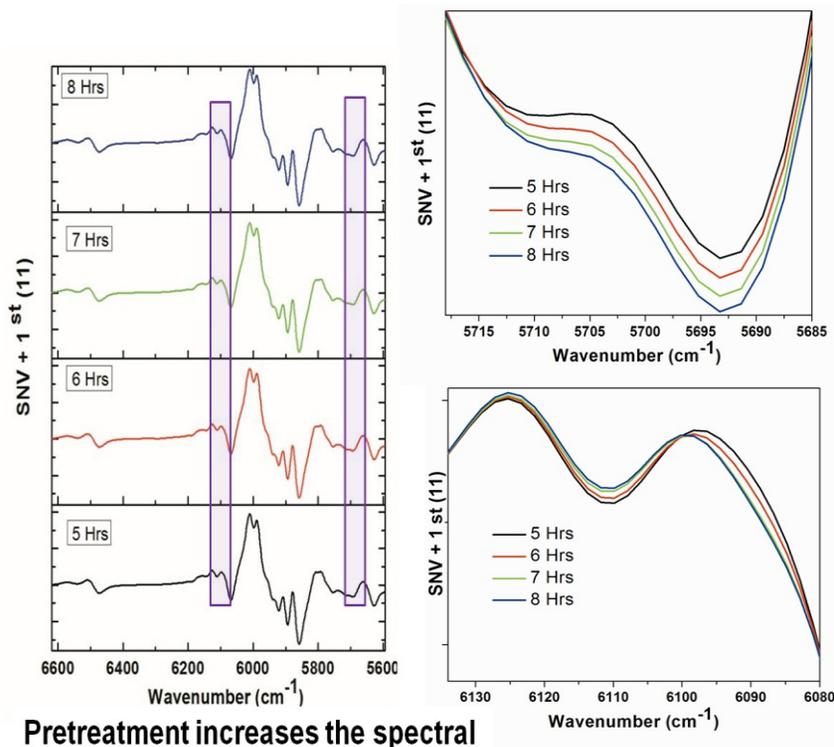
Figure 8.16 shows the score plot of the laboratory samples. Most of the scores are organized according to the changes in relative peak area of MMAA with respect to Duloxetine, except for samples 3.4551% and 3.1258%. This is a reaction very difficult of reproduce because the reaction must be under anhydrous conditions.

Figure 8.17 shows the PCA scores of the plant samples, each arylation is identified by colors. The samples are uniformly distributed; the samples are organized according to the changes in concentration, from high to lower concentration. Table 8.4 and 8.5 summarize the evaluation and results of the model obtained.

The true values are the result obtained using HPLC method and the fit and predicted values are the results obtained using the NIR model. The values obtained with the two methods are similar; the variability observed is of 0.1 %.

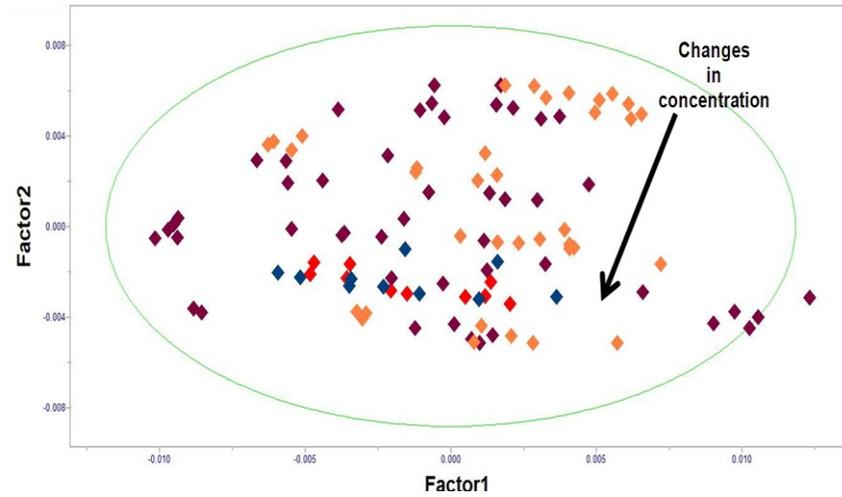
8.4.6 Next Steps

Validate the model with samples of other Arylations (samples that are not included in the model).



Pretreatment increases the spectral differences.

Principal Component Analysis (PCA)



Pretreatments: SNV + 1st (11)

Figure 8.17 Spectra and PCA of the plant samples.

Table 8.4 MMAA calibration model statistics. .

	MMAA
Range	4.9 - 0.2
Processing	First derivative and SNV
Smooth points	13
Mean centering	on
Region (cm⁻¹)	6619 -5593
R²-cal	99.78
RMSEE	0.0698
R²-val	99.63
RMSEP	0.0871
Rank	4

Table 8.5 Calibration and Validation values* obtained with the NIR model using plant samples.

Calibration		Validation	
TRUE (%)	Fit	TRUE (%)	Prediction
4.4	4.4	4.4	4.4
4.9	4.9	4.9	4.8
3.2	3.2	2.8	2.9
3.3	3.2	3.2	3.2
2.8	2.9	3.3	3.2
1.8	1.8	1.8	1.8
2.1	2.1	2.1	2.2
1.4	1.4	1.4	1.4
0.4	0.4	0.2	0.1
0.2	0.2	0.4	0.4

* This values correspond to relative peak area of MMAA with respect to Duloxetine

8.5 Perform a feasibility study for implementation of NIR for solvent recovery process.

8.5.1 Problem

The QC laboratory analyzes on a weekly basis X batches of recovered *ethyl acetate* (EtOAc) for testing (%) and from Y batches of recovered Toluene. Each analysis is performed with Gas Chromatography (GC) taking around 5 hours. Batch release has been delayed several times due to equipment upsets. Delay in ethyl acetate testing negatively impacts the Toluene release process. As a result, the Toluene batch has to be reprocessed. This generates a greater amount of waste and the need to buy more Toluene to continue with the normal process. It is necessary to implement an alternative technique of analysis faster.

8.5.2 Methodology

A new NIR predictive model was developed using data acquired previously. GC was used to determine the EtOAc content of the different samples.

8.5.3 Instrumentation

MPA (Multi Purpose Analyzer) FT-NIR spectrometer equipment was used for spectral acquisition. NIR spectra were obtained in transmittance mode using a PbS detector. Each spectrum was an average of 128 scans at a resolution of 8 cm^{-1} , over the range of $12,000 - 3500\text{ cm}^{-1}$. The NIR models were developed using OPUS Software version 6.5.

8.5.4 Results

NIR spectra were obtained using two different size vials. The vial size affects the optical path of the radiation. Figure 8.18 shows a greater spectral differentiation between samples in the spectral range selected. This spectral range corresponds to first and second overtones, the intensity of these bands increased according to the increase

sample absorption. The best predictions were obtained using a vial of 21 mm-od crew cap PTFE lined (Kimble, part no 60940A 16) or equivalent.

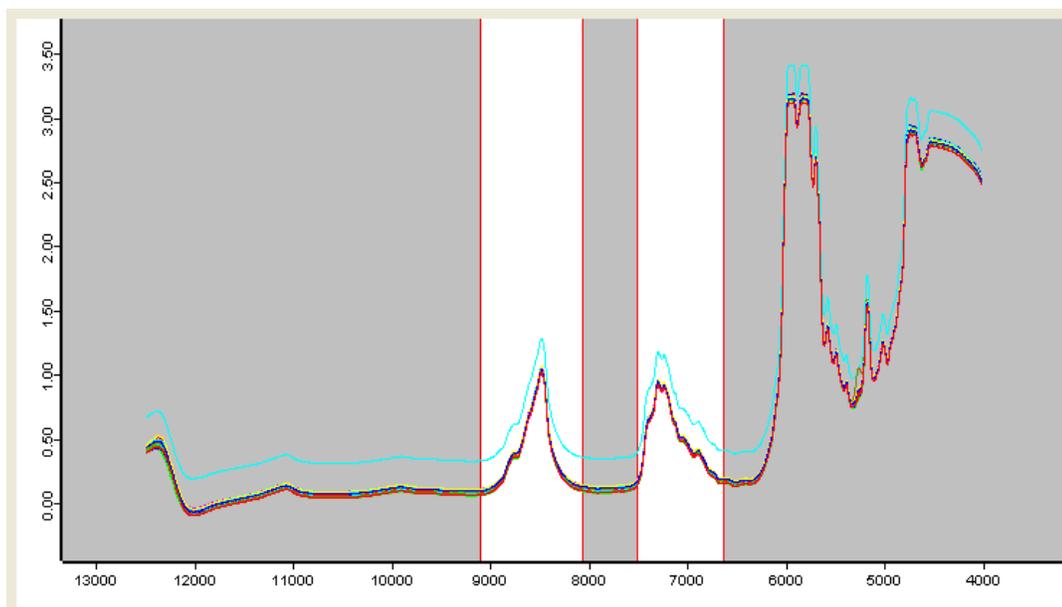


Figure 8.18 Selected spectral range.

Table 8.6 shows the parameters and evaluation of the NIR model. This model has a very narrow range due to that the target is 99.5% of EtOAc. The temperature is monitored to ensure that most of the distilled product is ethyl acetate.

Table 8.6 Specification of the EtOAc Model.

	EtOAc
Range	100-97.5
Processing	First derivative and SNV
Smooth points	13
Mean centering	on
Region (cm⁻¹)	9103.1-8061.6
	7510-6630.6
R²-cal	99.97
RMSEE	0.0173
R²-val	99.92
RMSEP	0.0165
Rank	3

Again, the true values are the result obtained using HPLC method and the fit and predicted values are the results obtained using the NIR model. Table 8.7 shows that the values obtained with the two methods are similar; the variability observed is of 0.1%.

Table 8.7 Calibration and Validation values* obtained with the NIR model using solvent recovery samples.

Calibration		Validation	
TRUE	Fit	TRUE	Prediction
100	100	99.73	99.73
100	100	99.73	99.73
100	100	99.72	99.72
100	100	99.72	99.72
100	100	99.74	99.74
99.5	99.5	99.73	99.73
99.5	99.5	98.96	98.97
99.5	99.5	98.99	99
99.5	99.5	98.98	98.99
99.5	99.5	98.98	98.98
98.7	98.7	98.99	99.01
98.7	98.7	98.97	98.98
98.7	98.7	98.43	98.43
98.7	98.8	98.43	98.44
97.6	97.7	98.44	98.44
97.6	97.6	98.43	98.44
97.6	97.6	98.42	98.42
97.6	97.6	98.41	98.42
97.6	97.7	98.26	98.27
97.5	97.5	98.25	98.25
97.5	97.5	98.26	98.26
97.5	97.5	98.23	98.22
97.5	97.5	98.27	98.28
97.5	97.5	98.21	98.21

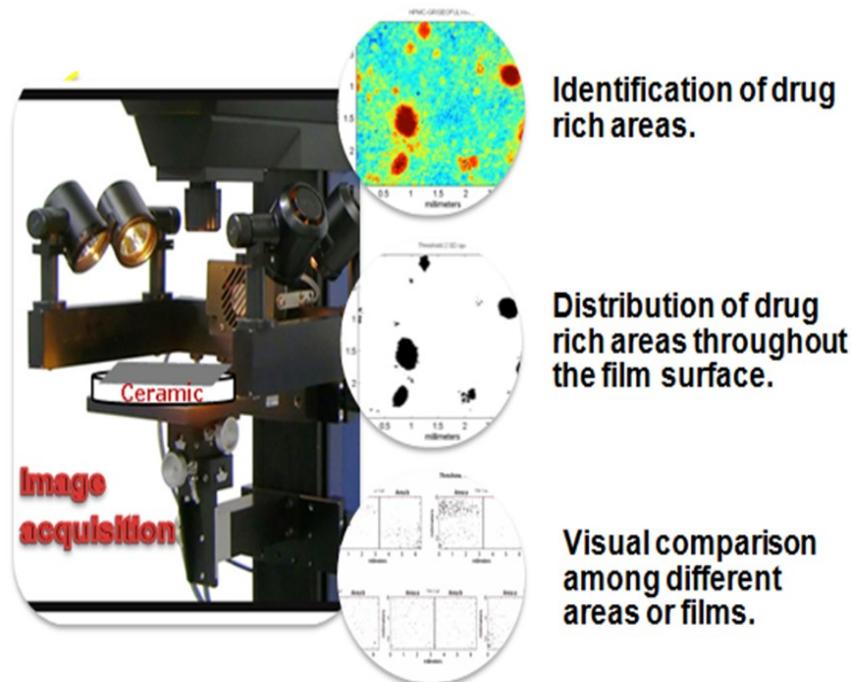
* These values correspond to relative peak area of EtOAc with respect to toluene

8.5.5 Next Steps

Continue the validation of the model analyzed more samples of solvent recovery process. Additionally, do a *repeatability study*.

9. Research Contribution

During the development of the polymeric films formulation a factor important to take in account is the distribution of API through the film surface. Drug distribution impacts the drug content of the polymeric films. The objective of this pharmaceutical formulation is to obtain a reproducible film with unagglomerated drug. Therefore, it is very important to develop a technique to evaluate the drug distributions in the polymeric films. This research was directed at developing an analytical method to analyze the API distribution in the polymeric films. This developed method is based on the analysis of hyperspectral images obtained in the near infrared range. Based in the literature review, this is the first methodology developed to evaluate drug distribution in this type of pharmaceutical formulation. Additionally, this research also introduced the term drug rich areas in the hyperspectral image analysis. Drug rich areas are those areas where pixels with greater contribution of drug per pixel are present. Pure drug pixels are observed only when the NIR image comes from a film with highly agglomerated drug. Drug rich areas are used when analyzing films with improved drug distribution.



The previous image show the approach developed during this dissertation.

The Identification of drug rich areas was based on analysis of score images or on analysis of images acquired at a specific wavelength. The identification of drug throughout the film surface was developed applying a threshold value to generate binary images. Binary images are a powerful in line tool that allow evaluating the image in a simple way. These binary images identify the drug rich areas, where the drug is agglomerated per pixel. The second evaluation consisted of determining the drug distribution through the entire film sample, thus representing drug distribution over an area of at least one potential dose. These drug distributions were evaluated in based of the number of drug rich areas, mean area, area STD, largest domain size and lowest domain size. Visual comparison among different areas of the same film or in films prepared with similar conditions was evaluated using concatenation.

9.1 Future Work

Develop a methodology to quantify the degree of distribution of the drug through the film surface. This is important because it would facilitate comparison of results between different batches. Develop a methodology to assess the stability and aging of the polymeric films.

10. REFERENCES

1. Gendrin, C.; Roggo, Y.; Spiegel, C.; Collet, C., Monitoring galenical process development by near infrared chemical imaging: one case study. *European journal of pharmaceutics and biopharmaceutics* **2008**, *68* (3), 828-37.
2. Gowen, A.; O'Donnell, C.; Cullen, P.; Bell, S., Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European journal of pharmaceutics and biopharmaceutics* : **2008**, *69* (1), 10-22.
3. Wu, Z.; Tao, O.; Cheng, W.; Yu, L.; Shi, X.; Qiao, Y., Visualizing excipient composition and homogeneity of Compound Liquorice Tablets by near-infrared chemical imaging. *Spectrochim Acta A Mol Biomol Spectrosc* **2012**, *86*, 631-6.
4. Clarke, F., Extracting process-related information from pharmaceutical dosage forms using near infrared microscopy. *Vibrational Spectroscopy* **2004**, *34* (1), 25-35.
5. Bakeev, K. A., *Process analytical technology: spectroscopic tools and implementation strategies for the chemical and pharmaceutical industries*. Wiley: 2010.
6. Bobiak, J. P. Raman and FTIR imaging of dynamic polymer systems. Case Western Reserve, 2006.
7. Sievens-Figueroa, L.; Bhakay, A.; Jerez-Rozo, J. I.; Pandya, N.; Romanach, R. J.; Michniak-Kohn, B.; Iqbal, Z.; Bilgili, E.; Dave, R. N., Preparation and characterization of hydroxypropyl methyl cellulose films containing stable BCS Class II drug nanoparticles for pharmaceutical applications. *Int J Pharm* **2012**, *423* (2), 496-508.
8. Garsuch, V.; Breitzkreutz, J., Novel analytical methods for the characterization of oral wafers. *European journal of pharmaceutics and biopharmaceutics*. **2009**, *73* (1), 195-201.
9. Prats-Montalbán, J. M.; Jerez-Rozo, J. I.; Romañach, R. J.; Ferrer, A., MIA and NIR Chemical Imaging for pharmaceutical product characterization. *Chemometrics and Intelligent Laboratory Systems* **2012**, *117*, 240-249.

10. Jerez, J. I.; Zarow, A.; Zhou, B.; Pinal, R.; Iqbal, Z.; Romanach, R. J., Complementary near-infrared and Raman chemical imaging of pharmaceutical thin films. *Journal of pharmaceutical sciences* **2011**, *100* (11), 4888-95.
11. Beck, C.; Sievens-Figueroa, L.; Gärtner, K.; Jerez-Rozo, J. I.; Romañach, R. J.; Bilgili, E.; Davé, R. N., Effects of stabilizers on particle redispersion and dissolution from polymer strip films containing liquid antisolvent precipitated griseofulvin particles. *Powder Technology*.**2012**.
12. Gosselin, R.; Rodriguez, D.; Gonzalez-Nunez, R.; Duchesne, C., Potential of hyperspectral imaging for quality control of polymer blend films. *Industrial & Engineering Chemistry Research* **2009**, *48* (6), 3033-3042.
13. Gendrin, C.; Roggo, Y.; Collet, C., Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review. *Journal of pharmaceutical and biomedical analysis* **2008**, *48* (3), 533-553.
14. Susarla, R.; Sievens-Figueroa, L.; Bhakay, A.; Shen, Y.; Jerez-Rozo, J. I.; Engen, W.; Khusid, B.; Bilgili, E.; Romañach, R. J.; Morris, K. R.; Michniak-Kohn, B.; Davé, R. N., Poorly water-soluble API nano-particle loaded HPMC films formed via convective drying. **2013**.
15. Franch-Lage, F.; Amigo Rubio, J. M.; Skibsted, E.; MasPOCH, S.; Coello, J., Fast assessment of the surface distribution of API and excipients in tablets using NIR-hyperspectral imaging. *International Journal of Pharmaceutics* **2011**, *411* (1-2), 27-35.
16. Coutts-Lendon, C. A.; Wright, N. A.; Mieso, E. V.; Koenig, J. L., The use of FT-IR imaging as an analytical tool for the characterization of drug delivery systems. *Journal of controlled release* **2003**, *93* (3), 223-248.
17. Banker, G. S.; Rhodes, C., *Modern pharmaceuticals*. Informa HealthCare: 2002; Vol. 121.
18. Chan, K.; Kazarian, S., Visualisation of the heterogeneous water sorption in a pharmaceutical formulation under controlled humidity via FT-IR imaging. *Vibrational spectroscopy* **2004**, *35* (1), 45-49.
19. Le Person, S.; Puiggali, J.; Baron, M.; Roques, M., Near infrared drying of pharmaceutical thin films: experimental analysis of internal mass transport. *Chemical Engineering and Processing: Process Intensification* **1998**, *37* (3), 257-263.

20. Blanco, M.; Coello, J.; Iturriaga, H.; Maspoch, S.; De La Pezuela, C., Near-infrared spectroscopy in the pharmaceutical industry. *Analyst* **1998**, *123*, 135-150.
21. Burns, D. A.; Ciurczak, E. W., Basic Principles of Near-Infrared Spectroscopy. In *Handbook of near-infrared analysis*, Third ed.; Siesler, H. W., Ed. Taylor & Francis Group, LLC: Boca Raton, FL, 2007; Vol. 35.
22. Ravn, C. Near-infrared Chemical Imaging in Formulation Development of Solid Dosage Forms. University of Copenhagen, Denmark, 2009.
23. Alcalà, M.; Blanco, M.; Menezes, J. C.; Felizardo, P. M.; Garrido, A.; Pérez, D.; Zamora, E.; Pasquini, C.; Romañach, R. J., Near-Infrared Spectroscopy in Laboratory and Process Analysis. In *Encyclopedia of Analytical Chemistry*, John Wiley & Sons, Ltd: 2012.
24. Amigo, J. M.; Ravn, C., Direct quantification and distribution assessment of major and minor components in pharmaceutical tablets by NIR-chemical imaging. *European Journal of Pharmaceutical Sciences* **2009**, *37* (2), 76-82.
25. Aenugu, H. P. R.; Kumar, D. S.; Parthiban, S.; Ghosh, S. S.; Banji, D., Near infrared spectroscopy-An overview. *International Journal of ChemTech Research* **2011**, *3* (2), 825-836.
26. Coddill, R. P.; Drennen, J. K., Near-Infrared Spectroscopy. In *Spectroscopy of Pharmaceutical Solids* Brittain, H. G., Ed. Taylor & Francis: New York., 2006.
27. Ely, D. R.; Carvajal, M. T., Determination of the scale of segregation of low dose tablets using hyperspectral imaging. *International journal of pharmaceutics* **2011**, *414* (1), 157-160.
28. Šašić, S.; Ozaki, Y., *Raman, Infrared, and Near-Infrared Chemical Imaging*. Wiley Online Library: 2010.
29. Buice, R. G.; Gold, T. B.; Lodder, R. A.; Digenis, G. A., Determination of Moisture in Intact Gelatin Capsules by Near-Infrared Spectrophotometry. *Pharmaceutical Research* **1995**, *12* (1), 161-162.
30. Svensson, O.; Josefson, M.; Langkilde, F. W., Classification of Chemically Modified Celluloses Using a Near-Infrared Spectrometer and Soft Independent Modeling of Class Analogies. *Applied Spectroscopy* **1997**, *51* (12), 1826-1835.

31. Gustafsson, C.; Nyström, C.; Lennholm, H.; Bonferoni, M. C.; Caramella, C. M., Characteristics of hydroxypropyl methylcellulose influencing compactibility and prediction of particle and tablet properties by infrared spectroscopy. *Journal of pharmaceutical sciences* **2003**, *92* (3), 494–504.
32. Furukawa, T.; Sato, H.; Shinzawa, H.; Noda, I.; Ochiai, S., Evaluation of homogeneity of binary blends of poly(3-hydroxybutyrate) and poly(L-lactic acid) studied by near infrared chemical imaging (NIRCI). *Analytical sciences : the international journal of the Japan Society for Analytical Chemistry* **2007**, *23* (7), 871-6.
33. Nerella, N. G.; Drennen, J. K., Depth-Resolved Near-Infrared Spectroscopy. *Applied Spectroscopy* **1996**, *50* (2), 285-291.
34. Blanco, M.; Romero, M. A., Near infrared transfectance spectroscopy: determination of dexketoprofen in a hydrogel. *Journal Pharmaceutical Biomedical Analysis* **2002**, *30* (3), 467-472.
35. Gendrin, C.; Roggo, Y.; Collet, C., Content uniformity of pharmaceutical solid dosage forms by near infrared hyperspectral imaging: a feasibility study. *Talanta* **2007**, *73* (4), 733-741.
36. Cruz, J.; Blanco, M., Content uniformity studies in tablets by NIR-CI. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *56* (2), 408-412.
37. Burger, J., *Hyperspectral NIR image analysis*. 2006; Vol. 2006.
38. Clark, D.; Henson, M.; LaPlant, F.; Šašić, S.; Zhang, L., Pharmaceutical applications of chemical mapping and imaging. *Handbook of Vibrational Spectroscopy* **2007**.
39. Clarke, F. C.; Jamieson, M. J.; Clark, D. A.; Hammond, S. V.; Jee, R. D.; Moffat, A. C., Chemical image fusion. The synergy of FT-NIR and Raman mapping microscopy to enable a more complete visualization of pharmaceutical formulations. *Analytical chemistry* **2001**, *73* (10), 2213-2220.
40. Cruz, J.; Bautista, M.; Amigo, J. M.; Blanco, M., NIR-Chemical Imaging study of acetylsalicylic acid in commercial tablets. *Talanta* **2009**, *80* (2), 473.
41. Amigo, J. M., Practical issues of hyperspectral imaging analysis of solid dosage forms. *Analytical and bioanalytical chemistry* **2010**, *398* (1), 93-109.

42. Bagchi, S.; Li, W.; Plakogiannis, F., Optimization and prediction of drug release from matrix tablets using response surface methodology and near infrared chemical imaging. *Pharmaceutical development and technology* **2012**, *17* (4), 398-406.
43. Cairós, C.; Amigo, J. M.; Watt, R.; Coello, J.; MasPOCH, S., Implementation of enhanced correlation maps in near infrared chemical images: Application in pharmaceutical research. *Talanta* **2009**, *79* (3), 657-664.
44. Chalmers, J. M.; Everall, N. J.; Schaeberle, M. D.; Levin, I. W.; Neil Lewis, E.; Kidder, L. H.; Wilson, J.; Crocombe, R., FT-IR imaging of polymers: an industrial appraisal. *Vibrational spectroscopy* **2002**, *30* (1), 43-52.
45. Reich, G., Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications. *Advanced drug delivery reviews* **2005**, *57* (8), 1109-1143.
46. Rinnan, A.; Berg, F.; Engelsen, S. B., Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry* **2009**, *28* (10), 1201-1222.
47. Tobias, R. D., Chemometrics: A Practical Guide. *Technometrics* **1999**, *41* (4), 375-376.
48. Smith, L. I., A tutorial on principal components analysis. *Cornell University, USA* **2002**, *51*, 52.
49. Roggo, Y.; Jent, N.; Edmond, A.; Chalus, P.; Ulmschneider, M., Characterizing process effects on pharmaceutical solid forms using near-infrared spectroscopy and infrared imaging. *European journal of pharmaceutics and biopharmaceutics*. **2005**, *61* (1), 100-110.
50. Salzer, R.; Siesler, H. W., *Infrared and Raman spectroscopic imaging*. Wiley Online Library: 2009.
51. Tauler, R., Multivariate curve resolution applied to second order data. *Chemometrics and Intelligent Laboratory Systems* **1995**, *30* (1), 133-146.
52. Prats-Montalbán, J. M.; de Juan, A.; Ferrer, A., Multivariate image analysis: A review with applications. *Chemometrics and Intelligent Laboratory Systems* **2011**, *107* (1), 1-23.

53. Prats-Montalbán, J. M.; Ferrer, A., Integration of colour and textural information in multivariate image analysis: defect detection and classification issues. *Journal of Chemometrics* **2007**, *21* (1-2), 10-23.
54. Barker, M.; Rayens, W., Partial least squares for discrimination. *Journal of chemometrics* **2003**, *17* (3), 166-173.
55. Navea, S.; Tauler, R.; Goormaghtigh, E.; de Juan, A., Chemometric Tools for Classification and Elucidation of Protein Secondary Structure from Infrared and Circular Dichroism Spectroscopic Measurements. *PROTEINS: Structure, Function, and Bioinformatics* **2006**, *63*, 527-541.
56. Musumarra, G.; Barrese, V.; Condorelli, D. F.; Fortuna, C.; Scire, S., Genome-based identification of diagnostic molecular markers for human lung carcinomas by PLS-DA. *Computational Biology and Chemistry* **2005**, *29*, 183-195.
57. Roggo, Y.; Edmond, A.; Chalus, P.; Ulmschneider, M., Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta* **2005**, *535* (1), 79-87.
58. Instrument, M. *ISys 5.0 Chemical Imaging Software. User's Manual* 2002-2008.
59. Shi, Z.; Anderson, C. A., Application of Monte Carlo simulation-based photon migration for enhanced understanding of near-infrared (NIR) diffuse reflectance. Part II: Photon Radial Diffusion in NIR Chemical Images. **2010**, *99* (5), 1-9.
60. Shi, Z.; Anderson, C. A., Application of Monte Carlo simulation-based photon migration for enhanced understanding of near-infrared (NIR) diffuse reflectance. Part I: Depth of penetration in pharmaceutical materials. *Journal of pharmaceutical sciences* **2010**, *99* (5), 2399-2412.
61. Hudak, S. J.; Haber, K.; Sando, G.; Kidder, L. H.; Lewis, E. N.; *Practical limits of spatial resolution in diffuse reflectance NIR chemical imaging*; Malvern Instruments: Columbia, MD 21046, USA.
62. Wu, Z.; Tao, O.; Dai, X.; Du, M.; Shi, X.; Qiao, Y., Monitoring of a pharmaceutical blending process using near infrared chemical imaging. *Vibrational Spectroscopy* **2012**, *63* (0), 371-379.

63. Blanco, M.; Cueva-Mestaza, R.; Cruz, J., Critical evaluation of methods for end-point determination in pharmaceutical blending processes. *Analytical Methods* **2012**, *4*, 2694-2703.
64. Piqueras, S.; Burger, J.; Tauler, R.; de Juan, A., Relevant aspects of quantification and sample heterogeneity in hyperspectral image resolution. *Chemometrics and Intelligent Laboratory Systems* **2012**, *117*, 169-182.
65. Offroy, M.; Roggo, Y.; Duponchel, L., Increasing the spatial resolution of near infrared chemical images (NIR-CI): The super-resolution paradigm applied to pharmaceutical products. *Chemometrics and Intelligent Laboratory Systems* **2012**, *117*, 183-188.
66. Rosas, J. G.; Blanco, M., A criterion for assessing homogeneity distribution in hyperspectral images. Part 2: Application of homogeneity indices to solid pharmaceutical dosage forms. *Journal of Pharmaceutical and Biomedical Analysis* **2012**, *70*, 691-699.
67. Post, G.; de Carvalho, F.; Poppi, R., Study of the similarity between distribution maps of concentration in near-infrared spectroscopy chemical imaging obtained by different multivariate calibration approaches. *Microchemical Journal* **2011**, *99*, 542-547.
68. de Juan, A.; Tauler, R., Spectroscopic imaging and chemometrics: a powerful combination for global and local sample analysis. *TrAC Trends in Analytical Chemistry* **2004**, *23* (1), 70-79.
69. Prats-Montalbán, J. M.; Ferrer, A.; Bro, R.; Hancewicz, T., Prediction of skin quality properties by different Multivariate Image Analysis methodologies. *Chemometrics and Intelligent Laboratory Systems* **2009**, *96* (1), 6-13.
70. Bilgili, E.; Afolabi, A., A combined microhydrodynamics–polymer adsorption analysis for elucidation of the roles of stabilizers in wet stirred media milling. *International Journal of Pharmaceutics* **2012**, *439*, 193-206.
71. Nair, R.; Nyamweya, N.; S. Gonen, S.; Martinez-Miranda, L. J.; Hoag, S. W., Influence of various drugs on the glass transition temperature of poly(vinylpyrrolidone): a thermodynamic and spectroscopic investigation. *International Journal of Pharmaceutics* **2001**, *225*, 83-96.

72. Hirasawa, N.; Danjo, K.; Haruna, M.; Otsuka, A., Physicochemical Characterization and Drug Release Studies of Naproxen Solid Dispersions Using Lactose as a Carrier. *Chemical and Pharmaceutical Bulletin* **1998**, *46* (6), 1027-1030.
73. Mansour, H. M.; Hickey, A. J., Raman characterization and chemical imaging of biocolloidal self-assemblies, drug delivery systems, and pulmonary inhalation aerosols: a review. *AAPS PharmSciTech* **2007**, *8* (4), 140-155.
74. de Veij, M.; Vandenabeele, P.; De Beer, T.; Remon, J. P.; Moens, L., Reference database of Raman spectra of pharmaceutical excipients. *Journal of Raman Spectroscopy* **2009**, *40* (3), 297-307.
75. Selvarajan, A., Raman spectrum of dimethyl sulfoxide (DMSO) and the influence of solvents. *Proceedings of the Indian Academy of Sciences - Section A* **1966**, *64* (1), 44-50.
76. Neelakantan, P., Raman spectrum of ethyl chloroacetate. *Proceedings of the Indian Academy of Sciences - Section A* **1964**, *59* (6), 385-389.
77. Schiel, D.; Richter, W., The OH-stretching vibrational raman profile and isomer assignment of 2-propanol. *Chemical Physics Letters* **1987**, *142* (5), 345-348.
78. Snyder, R. G.; Kumamoto, J.; Ibers, J. A., Vibrational Spectrum of Crystalline Potassium Hydroxide. *The Journal of Chemical Physics* **1960**, *33* (4), 1171-1177.
79. Shurvell, H. F.; Southby, M. C., Infrared and Raman spectra of tetrahydrofuran hydroperoxide. *Vibrational Spectroscopy* **1997**, *15* (1), 137-146.
80. Wilmshurst, J. K.; Bernstein, H. J., THE INFRARED AND RAMAN SPECTRA OF TOLUENE, TOLUENE- α -d₃, m-XYLENE, AND m-XYLENE- α '-d₆. *Canadian Journal of Chemistry* **1957**, *35* (8), 911-925.