Universidade de Lisboa

Faculdade de Farmácia

Generic Medicines Development: a FTIR based methodology to assess the quantitative formula

Milene Isabel Rosa Fialho

Trabalho de campo orientado pelo Professor Doutor João Almeida Lopes,
Categoria Professor Auxiliar

Mestrado Integrado em Ciências Farmacêuticas

2023

Universidade de Lisboa

Faculdade de Farmácia

Generic Medicines Development: a FTIR based methodology to assess the quantitative formula

Milene Isabel Rosa Fialho

Trabalho Final de Mestrado Integrado em Ciências Farmacêuticas apresentado à Universidade de Lisboa através da Faculdade de Farmácia

Trabalho de campo orientado pelo Professor Doutor João Almeida Lopes, Categoria Professor Auxiliar

2023

Acknowledgements

There are so many people I would like to thank for having supported me in the realization of this thesis, but mainly for having completed this master's degree. Although they never did it for me, they did it with me, supporting me every day, every step of the way and never letting me give in.

I would like to thank Professor João Lopes for all his patience with me in teaching, helping, listening to me and especially for challenging me to develop this final master's work.

I would like to thank my family who have always been with me, especially my parents who made this course possible, 130 kilometers far away from home. They were with me in every tear and celebrated with me every achievement. Never forgetting my sister, who is and will always be my shelter, my home, and my way forward.

Thanks to my boyfriend, Miguel Costa, who managed to be my support even when he needed it more than I did. He stood by me with all my insecurities, my fears and encouraged my dreams, my growth, and my achievements.

Finally, thank you to all my friends. Highlighting my best friend of a lifetime, Ana Santana Fernandes; My partners during all these 5 years of the course: Ana Catarina Santos, Mariana Tibério, Miriam Sousa, Mariana Bastos and Constança Boigues; at last, but not least, Beatriz Costa, who helped me a lot in the development of this project.

Preface

This dissertation is submitted for the Master Integrated in Pharmaceutical Sciences at the Faculty of Pharmacy, University of Lisbon. The work was developed at faculty's laboratories under the supervision and orientation of Professor João Almeida Lopes.

Abstract

The structures of organic compounds can be discovered through the use of numerous methods, with Infrared Spectroscopy being one of the most used nowadays. One of the great advantages of this technique is that virtually any sample in any state can be studied. More recently, this method has become common practice in the pharmaceutical industry, both in the production itself and in quality control.

This work presupposes the application of two simple methods: MCR – Alternating Least Squares and MCR – Weighted Linear Combination, both based on Infrared Spectroscopy, to estimate the concentration of components present in solid formulations. The methods used are based on the Lambert-Beer Law and can be extremely advantageous, as they allow a first quantitative estimate of the formulation, requiring only the knowledge of the number of components present and their pure spectra (it does not require a calibration procedure). The methods were tested in formulations containing Paracetamol and Caffeine, as active substances, and Microcrystalline Cellulose and Lactose, as excipients.

The analysis of the MCR – Alternating Least Squares method was very detailed, having been divided into three parts, thus increasing the level of information available. The results showed that, without providing the pure FTIR spectra of the components, the method became inefficient. On the contrary, when pure FTIR spectra were applied, the values approached the real concentrations, however the difficulty of the method in quantifying the active substances was always noted.

The following method was limited to four formulations, having then divided the respective analysis into two parts, namely: in the first part any type of restrictions were not applied and in the second part restrictions were imposed only on the values of the excipients. It should be noted that this method considered the different molar absorptivities of each component of the formulation. The technique used obtained better results than the commercial one, especially in the detection of one of the active substances - Paracetamol, making it an excellent tool for estimating the composition of a solid formulation, which can be completed later with other more complex techniques.

Keywords: FTIR; Formulations; MCR; Concentrations; Spectroscopy

Resumo

As estruturas dos compostos orgânicos podem ser descobertas através da utilização de inúmeros métodos, sendo a Espetroscopia de Infravermelho uma das mais utilizadas atualmente. Uma das grandes vantagens desta técnica é que, virtualmente, qualquer amostra em qualquer estado pode ser estudada. Mais recentemente, este método tornou-se uma prática comum na indústria farmacêutica, tanto na produção em si, como no controlo de qualidade.

Este trabalho pressupõe a aplicação de dois métodos simples: MCR – Alternating Least Squares e MCR – Weighted Linear Combination, ambos baseados na Espetroscopia de Infravermelho, para a estimativa da concentração dos componentes presentes em formulações sólidas. Os métodos utilizados baseiam-se na Lei de Lambert-Beer e podem ser extremamente vantajosos, dado que permitem uma primeira estimativa quantitativa da formulação, requerendo apenas o conhecimento do número de componentes presentes e os seus espetros puros. Os métodos foram testados em formulações contendo Paracetamol e Cafeína, como substâncias ativas, e Celulose Microcristalina e Lactose, como excipientes.

A análise do método MCR – Alternating Least Squares foi muito pormenorizada, tendo sido repartida em três partes, aumentando-se assim o nível de informação disponibilizada. Os resultados demonstraram que, sem o fornecimento dos espetros de FTIR puros dos componentes, o método tornava-se ineficiente. Contrariamente, aquando da aplicação dos espetros de FTIR puros, os valores aproximaram-se das concentrações reais, no entanto notou-se sempre a dificuldade do método em quantificar as concentrações das substâncias ativas.

O método seguinte foi limitado a quatro formulações, tendo-se então divido a respetiva análise em duas partes, a saber: na primeira parte quaisquer tipos de restrições não foram aplicados e na segunda parte foram impostas restrições apenas aos valores dos excipientes. Importa sublinhar que este método teve em consideração as diferentes absortividades molares de cada componente da formulação. A técnica utilizada obteve melhores resultados do que o método comercial, especialmente na deteção de uma das substâncias ativas - o Paracetamol, tornando-se uma ferramenta de excelência para uma estimativa da composição de uma formulação sólida, que pode ser completada posteriormente com outras técnicas mais complexas.

Palavras-chave: FTIR; Formulações; MCR; Concentrações; Espetroscopia

Table of Contents

Index of Figures

Index of Tables

1. List of Acronyms and Abbreviations

ALS	Alternating Least Squares
API	Active Pharmaceutical Ingredient
ATR	Attenuated Total Reflection
EMR	Electromagnetic Radiation
FTIR	Fourier-transform infrared
IR	Infrared
MCC	Microcrystalline cellulose
MCR	Multivariate Curve Resolution
MgS	Magnesium Stearate
NMR	Nuclear Magnetic Resonance
PCA	Principal Component Analysis
RLD	Reference Listed Drug
WLC	Weighted Linear Combination

2. Introduction

There are approximately 10 million known organic chemicals, and approximately 100,000 new organic compounds are produced annually. The structures of organic compounds can be found by using a variety of instruments – NMR, UV/visible, infrared, mass spectrometry, and X-ray crystallography are a few examples. Infrared spectroscopy is one of the most widely used of the listed instrumental techniques. (1)

One of the major benefits of infrared spectroscopy is that any sample in any state can be studied virtually. Provided the right sampling method, liquids, solutions, pastes, powders, films, fibers, gases, and surfaces, all can be examined. (2)
Infrared spectroscopy is a technique based on the vibrations of the atoms of a molecule. By passing infrared radiation through a sample and determining the proportion of the incident radiation that is absorbed at a particular energy, an infrared spectrum is typically obtained. The frequency of a vibration of a portion of a sample molecule corresponds to the energy at which any peak in an absorption spectrum appears. (3)

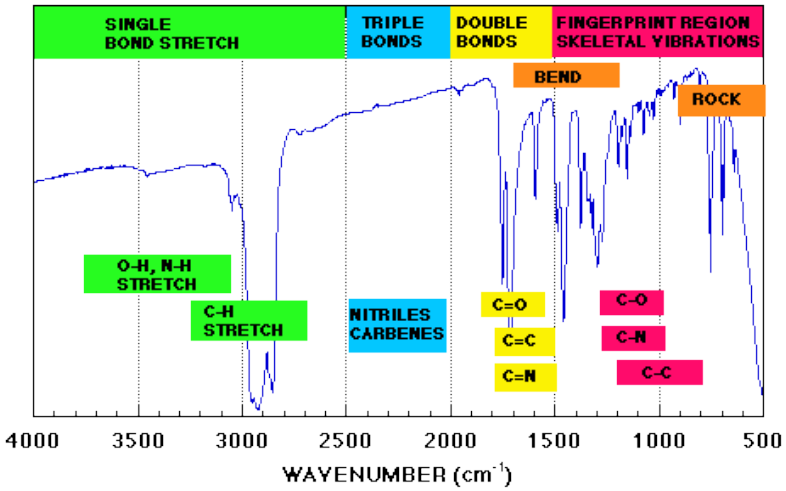

Figure 1: Scheme for interpreting the major regions of an FTIR spectrum (4)

Vibrational spectroscopy techniques based on infrared (IR) and Raman spectrophotometers are becoming common practice in the pharmaceutical industry in both manufacturing and quality control laboratories. Identification of raw materials and finished products as well as polymorphism characterization and quantification of active pharmaceutical ingredients (APIs) and excipients are the main applications found in scientific literature.

Vibrational spectroscopy has also proven to be one of the most powerful tools in reverse engineering. Each molecule has a unique infrared signature providing great specificity in the identification process. Although generics companies have been using reverse engineering for some time, the topic is little covered in the scientific literature. Reverse engineering is a useful tool for developing generic products to better ensure bioequivalence – a strategy that encompasses the decoding of the RLD's quantitative formula, the solid-state characterization of the API, and the manufacturing process. (5)

A new concept in the literature is the use of Attenuated Total Reflection Fourier-Transform infrared spectroscopy (ATR-FTIR) to detect counterfeit medicines. High amounts of counterfeit medicines enter the European market and detection of such pharmaceuticals at customs is not always simple due to misleading packaging. Therefore, easy to use equipment and techniques to perform an initial screening of these samples are required. For this purpose, Attenuated Total Reflection Fourier-Transform infrared spectroscopy (ATR-FTIR) is an interesting technique since it is user-friendly and little sample preparation is needed. (6)

3. Main Goals

The objective of this paper is to study the capability of using a methodology based on infrared spectroscopy to identify the components and obtain an estimation of the concentrations of each one, in a series of solid pharmaceutical formulations. In particular, formulations with a fixed value of each excipient and two active pharmaceutical ingredients (APIs) in very low concentrations and varying only slightly from each formulation to another. These formulations were purposely made for this work, which guarantee total knowledge of their composition.

To evaluate the spectra of the all the formulations developed, two distinct methods will be used: Multivariate Curve Resolution (MCR) – Alternating Least Squares and MCR – Weighted Linear Combination (created by Professor João Almeida Lopes).

4. Theoretical foundation

4.1. Vibrational Spectroscopy

Spectroscopy is the study of absorption, transmission, and/or emission of radiation by a molecule. Vibrational spectroscopy is used to identify and classify compounds based on the vibrations of their bonds. Depending on the type, vibrational spectroscopy either measures which frequencies of radiation are absorbed versus transmitted or which types of light scattering occur. The fingerprint region of a vibrational spectroscopy output refers to the set of vibrations produced by a molecule due to the vibrations of its bonds. Vibrational spectroscopy includes several techniques, the most important of which are mid-infrared (MIR), near-IR (NIR), and Raman spectroscopy. (7) (8)

4.2. Infrared Spectroscopy

Infrared light or infrared radiation, also known as thermal radiation, is electromagnetic radiation (EMR) with wavelengths longer than those of visible light. It is a type of radiant energy that is invisible to the human eye, however that can be felt as heat. The discovery of this region of the electromagnetic spectrum was made by English scientist William Herschel, in the last year of the 18^{th} century. (9)

In infrared spectroscopy, samples exposed to infrared radiation selectively absorb radiation of specific wavelengths, causing a change of dipole moment (separation of positive and negative charges) of sample molecules. Consequently, the vibrational energy levels of sample molecules transfer from ground state to excited state. The frequency of the absorption peak is determined by the vibrational energy gap. The number of absorption peaks is related to the number of vibrational modes of the molecule. The intensity of absorption peaks is related to the change of dipole moment and the possibility of the transition of energy levels. Therefore, by analyzing the infrared spectrum, one can readily obtain abundant structure information of a molecule. (9)

The commonly used region for Infrared spectroscopy is 4000 ~ 400 cm-1 because the absorption radiation of most organic compounds and inorganic ions is within this region. The region between 400 and 1500 cm^{-1} is known as the fingerprint region, so called because it's

difficult to assign all the absorption bands, and because of the unique patterns found there. Absorption bands in the 4000 to 1450 cm^{-1} region are usually due to stretching vibrations of diatomic units, and this is sometimes called the group frequency region. (9)

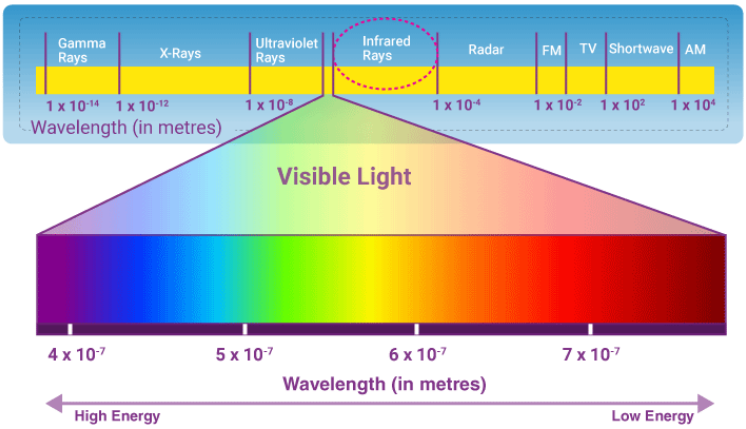

Figure 2: Electromagnetic Spectrum (10)

4.3. Fourier-transform infrared (FTIR) Spectroscopy

Infrared spectrometers have been commercially available since the 1940s. At that time, it consisted of three basic components: radiation source, monochromator, and detector; however, by the mid 1950s, diffraction gratings were introduced into dispersive machines. However, the most significant advances in infrared spectroscopy have come with the introduction of Fourier-transform spectrometers. This type of instrument employs an interferometer and exploits the well-established mathematical process of the Fourier transformation. Fourier-transform infrared (FTIR) spectroscopy has dramatically improved the quality of infrared spectra and minimized the time required to obtain data. In addition, with constant improvements to computers, infrared spectroscopy has made further great strides. Fourier-transform infrared (FTIR) spectroscopy is based on the idea of the interference of radiation between two beams to yield an interferogram. The latter is a signal produced as a function of the change in the path length between the two beams. The two domains of distance and frequency are interconvertible using the mathematical Fourier transformation method. (1)

The radiation emerging from the source is passed through an interferometer to the sample before reaching a detector. Upon amplification of the signal, in which high-frequency contributions have been eliminated by a filter, the data are converted to digital form by an

analog-to-digital converter and transferred to the computer for Fourier-transformation. The most common interferometer used in FTIR spectrometry is a Michelson interferometer, which consists of two perpendicularly plane mirrors, one of which can travel in a direction perpendicular to the plane. A semi-reflecting film, the beam splitter, bisects the planes of these two mirrors. (1)

In a perfect beam splitter, if a collimated beam of monochromatic light with wavelength (cm) is passed through, 50% of the incident radiation will be reflected to one mirror and 50% will be transmitted to the other mirror. These mirrors reflect the two beams, sending them back to the beam splitter where they combine and interfere. In the beam splitter, half of the beam that is reflected from the fixed mirror is transmitted while the other half is reflected in the direction of the source. The transmitted beam, which is the beam detected in FTIR spectrometry, is the beam that exits the interferometer at a 90- degree angle to the input beam. (1)

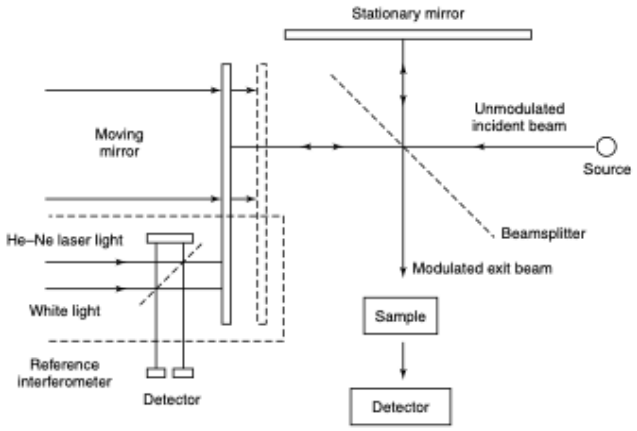

Figure 3: Schematic of a Michelson Interferometer (1)

4.4. Attenuated Total Reflection (ATR)

To overcome the disadvantages of KBr pellets and liquid cells, nowadays IR-measurements are mainly performed in ATR (Attenuated Total Reflection) mode as this technique is simpler to use than the conventional transmission mode. All types of samples are placed undiluted on the ATR crystal. The measurement is typically performed within a few seconds. Attenuated total reflectance (ATR) spectroscopy utilizes the phenomenon of total internal reflection. A beam of

radiation entering a crystal will undergo total internal reflection when the angle of incidence at the interface between the sample and crystal is greater than the critical angle, where the latter is a function of the refractive indices of the two surfaces. The beam penetrates a fraction of a wavelength beyond the reflecting surface and when a material that selectively absorbs radiation is in close contact with the reflecting surface, the beam loses energy at the wavelength where the material absorbs. The resultant attenuated radiation is measured and plotted as a function of wavelength by the spectrometer and gives rise to the absorption spectral characteristics of the sample. (1)

The basic principle is shown in Figure 4. The ATR crystal comprises an IR transparent material with a high refractive index and polished surfaces. (11)

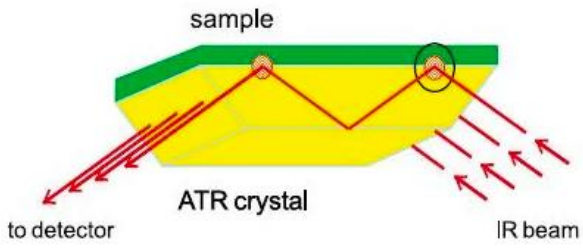

Figure 4: Schematic of an ATR crystal (11)

The fraction of light reaching into the sample is known as evanescent wave. Its penetration depth depends on the wavelength, the refractive indices of the ATR crystal and the sample and the angle of the entering light beam. In the spectral regions where the sample absorbs energy, the evanescent wave is attenuated. After one or several internal reflections, the IR beam exits the ATR crystal and is directed to the IR-detector. (11)

It has the advantages of faster sampling without preparation (as mentioned earlier), excellent sample-to-sample reproducibility, and minimal operator-induced variations. (11)

4.5. Solid Dosage Form – Powder

Powders are solid dosage forms containing finely divided drugs and other substances intended for ingestion or external applications. They may contain one or more active ingredients, various excipients, and, if necessary, flavoring and coloring. In addition to the pharmaceutical industry, powders have applications in many fundamental fields such as foods, cosmetics, and chemicals. Although tablets and capsules have largely replaced the use of powders, they still represent one of the oldest forms of dosage and continue to hold advantages that make them valuable:

- Powders are more stable, and the chance of incompatibility are less as compared to liquid dosage form;
- The onset of action is faster compared to other solid dosage forms (due to smaller particle size) and is more economical because it does not require any special machine or technique;
- This dosage form is especially important in small children and elderly patients, who have difficulties in swallowing tablets or capsules.

Powders can be classified in:

- Bulk Powders – in the oral form (antiacids, dietary supplements, laxatives, and pain relievers) and topical;
- Dusting Powders – intended to be dusted on the skin;
- Douche Powders – used to prepare vaginal cleansing solutions;
- Insufflations – injected to the body;
- Divided powders – powdered medicines in one or more individual packages, such as cellophane, metallic foil or paper). (12)

4.6. Chemometrics

Chemometrics is a discipline that manipulates data from chemical processes utilizing mathematics and statistic fundamentals. The advancement of the electronics and computer science have allowed a constant growth of chemometrics, expanding the applications of this discipline in practically all sub-areas of chemistry.

Chemometrics starts with measurements and data collection. Subsequently, mathematical, and statistical methods are applied to extract relevant information from the collected data. This

information is related to a chemical process that will give greater knowledge about a given system and thus a greater understanding of it. (13)

4.6.1. Multivariate Curve Resolution (MCR)

Multivariate data analysis is very important for the exploratory analysis of chemical data, especially spectroscopic and chromatographic data. In many different areas of application, Multivariate Curve Resolution (MCR) is a widely used and effective method for modeling and analyzing multi-way data. By putting in place sufficient constraints, this method provides a bilinear description of the observed data variation within the confines of the chemical realm.

The broad definition of multivariate resolution includes all methods that aim to decompose a data matrix into a linear model of dyads (the bilinear model). Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Multivariate Curve Resolution (MCR) all have the same overarching goal from this point of view: the conversion of intricate mixtures into pure-component contributions when little or no prior data is available.

According to the Beer-Lambert law, which states that the quantity of light absorbed by a substance dissolved in a fully transmitting solvent, is directly proportional to the concentration of the substance and the path length of the light trough the solution (which is typically 1).

$$\boldsymbol{A = \varepsilon l x C} \qquad \text{(Equation 1)}$$

A is the absorbance, ε is the molar attenuation coefficient or absorptivity, l is the path length, and the C is the concentration. (14)

When mixing the components into one mixture and taking a spectrum of that, the spectrum obtained will be a linear combination of the spectra of all the pure components. This can be written as follows:

$$\boldsymbol{D = C S^T + E} \qquad \text{(Equation 2)}$$

where D is the vector of spectral values representing the spectrum of the mixture, C is the concentration of the compounds, S is a matrix containing spectra of all the pure components and E is a matrix with residuals.

The purpose of the MCR methods is to get C and S by knowing D, so we sort of resolve the mixtures into individual components and their concentrations. This is not a trivial task as the expression above does not have a unique solution. (15)

MCR strategies generally don't need a ton of requirements or earlier data about the substance personality of the parts. However, it is convenient to have an assessment of the number of pure components in the system. To have an optimized model the implementation of this method, it must have adequate constraints. The constraints that are more important are the non-negative, equality and the closure. Non-negative means that negative values for C and/or S are not accepted. The equality is applied when there is knowledge (total or partial) about C or S. Closure limits the total concentration of the components, which means that it gives the information to the model that it is a formulation and so, the total mass fraction is one). (16)

5. Materials and Methods

5.5. Characterization of the Active Substances

5.1.2. Acetaminophen (Paracetamol)

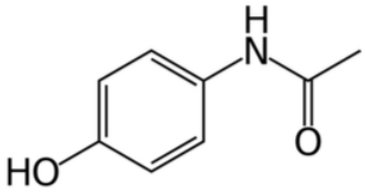

Figure 5: Paracetamol molecular structure ($C_8H_{10}NO_2$)

The World Health Organization (WHO) recommends Acetaminophen as the first- line treatment for pain disorders since it is the most widely used pain killer in the world. It is also employed for its antipyretic properties, which lower fever. This medication was first authorized by the U.S. FDA in 1951 and is offered in several dosage forms, including syrup, injection, suppository, normal tablets, effervescent tablets, and other forms. (17)

Brand used: FAGRON, Batch: 17H28-BO2-344184

5.1.2. Caffeine

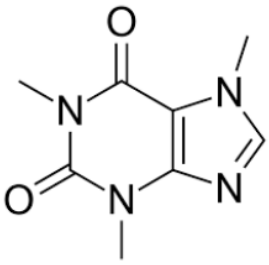

Figure 6: Caffeine molecular structure ($C_8H_{10}N_4O_2$)

A medication of the methylxanthine class known as caffeine is used for several problems including pain relief, the treatment of sleepiness, and several respiratory issues in premature

infants. Theophylline and Theobromine share a molecular structure with Caffeine. Although it can be obtained from coffee beans, it also naturally occurs in a few teas and cacao beans, which are not the same as coffee beans. Caffeine can be applied topically, taken orally, inhaled, injected, or used in cosmetics items. (18)

Brand used: CARLO ERBA, Batch: VSL048061E

5.2. Characterization of the excipients

5.2.1. Microcrystalline Cellulose

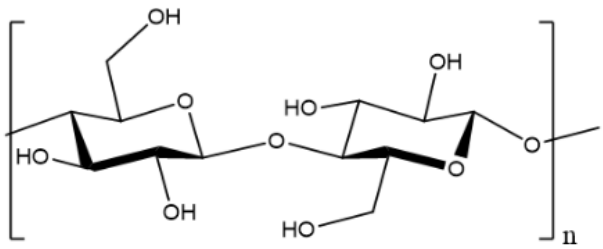

Figure 7: Microcrystalline cellulose molecular structure ($(C_6H_{10}O_5)n$)

Microcrystalline cellulose is a commonly used excipient in the pharmaceutical industry as a disintegrant. It has excellent compressibility properties and is used in solid dosage forms, such as tablets. Although tablets are hard, they dissolve quickly. (19)

- Appearance: white or almost white, fine, or granular powder.

- Solubility: solution insoluble in water, in acetone, in anhydrous ethanol and in toluene, in dilute acids and in the hydroxide solution at 50 g/l. (FPIX)

Brands used: Avicel PH-101 from FMC Corporation, Batch: 60842C (in set 1); Avicel PH-101 from Supelco – Sigma Aldrich, Batch: BCCG2024 (in set 2)

5.2.2. Lactose Monohydrate

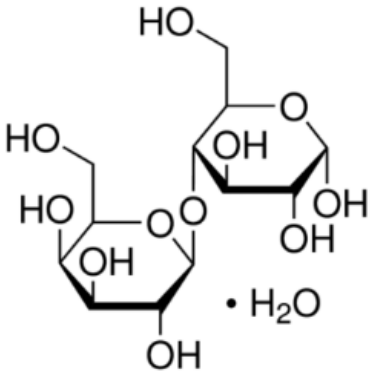

Figure 8: Lactose Monohydrate molecular structure ($C_{12}H_{22}O_{11}.H_2O$)

Lactose is one of the most used excipients within the pharmaceutical industry in the production of solid dosage forms. Lactose is commonly used as a diluent/binder to produce tablets of sufficient hardness whilst maintaining good disintegration properties. (20)

- Appearance: White or almost white crystalline powder.
- Solubility: easily but slowly soluble in water, practically insoluble in 96 percent ethanol. (FPIX)

Brand used: ERBA pharm, Batch: V6F689118N

5.3. Production of the Formulations

Two equal sets of 11 formulations (powders) with 30 grams each were planned. The placebo (composed by Microcrystalline Cellulose and Lactose) has a fixed value of 95% in each formulation (which equals to 28,5 grams in each formulation), i.e., each formulation will always have 57 grams of Lactose and 38 grams of MCC. The Active Pharmaceutical Ingredients, Paracetamol and Caffeine, are in very low concentration and are variable along the 11 formulations from each set. However, the two compounds together always make up 5% of the formulation.

The first step in the production of each set was to prepare the placebo. It was decided to make 20% more than necessary for all the 11 formulations. As described in table 1, the placebo was composed by 60% of Lactose and 40% of MCC, this mixture was made in a Cube Mixer for 20 minutes. This mixer was a Erweka® AR400 (ID number: EQ.0048).

Table 1: Composition of the placebo used in the formulations (in percentage mass by mass and in grams)

Placebo	% (m/m)	gram
MCC	40	150.4
Lactose	60	225.6
		Total= 376

After the preparation of the placebo, the weighing of the active substances followed. All weightings were carried out on an XS balance model BL-303, with a maximum capacity of 300 grams. The 4 components of each formulation were placed together and mixed in the same container (a plastic bottle of 150 mL). After that, the container was in an orbital mixer to mix for 15 minutes at 25 rpm. Finally, each formulation was manually mixed for 2 minutes. The orbital mixer was a 12-811 model from ThermoFisher Scientific®. (ID number: EQ.0057).

In the flowchart represented bellow it is possible to observe the production method better.

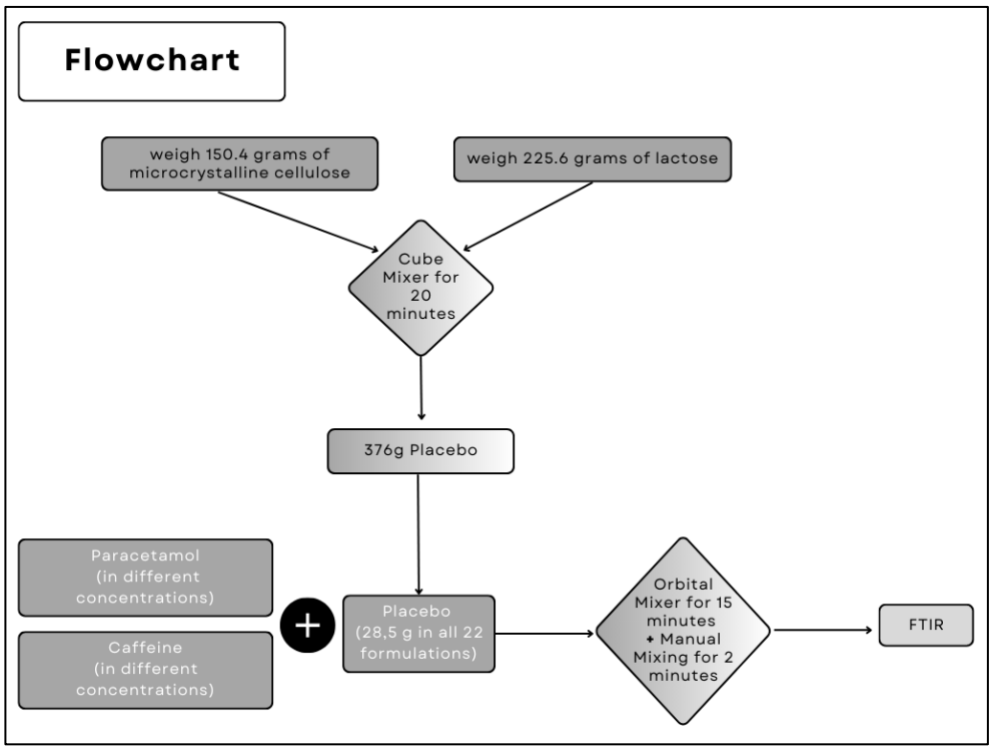

Figure 9: Formulations production flowchart

In the table 2, we can see which concentrations of paracetamol and caffeine were defined for each formulation, both in set 1 and set 2 (since they are replicas).

Table 2: Concentrations of Paracetamol and Caffeine defined for each formulation for both set 1 and 2

Number of Sets	Exp no	Caffeine (%)	Paracetamol (%)	Total (%)
2	1	0.0	5.0	5.0
	2	0.5	4.5	5.0
	3	1.0	4.0	5.0
	4	1.5	3.5	5.0
	5	2.0	3.0	5.0
	6	2.5	2.5	5.0
	7	3.0	2.0	5.0
	8	3.5	1.5	5.0
	9	4.0	1.0	5.0
	10	4.5	0.5	5.0
	11	5.0	0.0	5.0

In the tables 3 and 4, the weighted values carried out for each formulation of each set are recorded.

Table 3: Record of the weighted values of each component for each one of the 11 formulations of set 1

Sample no:	Caffeine (g)	Paracetamol (g)	Placebo (g)	Sample Mass (g)
1	0.000	1.499	28.505	30.004
2	0.150	1.350	28.567	29.917
3	0.300	1.198	28.529	30.027
4	0.444	1.051	28.501	29.996
5	0.596	0.898	28.601	30.095
6	0.755	0.755	28.534	30.044
7	0.915	0.602	28.526	30.043
8	1.066	0.447	28.519	30.032
9	1.202	0.295	28.600	30.097
10	1.348	0.147	28.510	30.005
11	1.497	0.000	28.509	30.006
Total:	8.273	8.242	313.901	330.266

Table 4: Record of the weighted values of each component for each one of the 11 formulations of set 2

Sample no	Caffeine (g)	Paracetamol (g)	Placebo (g)	Sample Mass (g)
1'	0.000	1.501	28.511	30.012
2'	0.153	1.356	28.523	30.032
3'	0.306	1.229	28.564	30.099
4'	0.449	1.051	28.511	30.011
5'	0.600	0.902	28.583	30.085
6'	0.749	0.748	28.537	30.034
7'	0.931	0.614	28.522	30.067
8'	1.053	0.457	28.505	30.015
9'	1.212	0.302	28.545	30.059
10'	1.353	0.153	28.503	30.009
11'	1.499	0.000	28.529	30.028
Total:	8.305	8.313	313.50	330.451

5.4. FTIR Spectral Measurements

The formulations and each pure compound were analyzed in a FTIR spectrometer, model Nicolet^{TM} iS 5 from ThermoFisher Scientific®. The measurement conditions are explained in table 5. Before each scan, the ATR crystal was cleaned with propranolol and a background was collected. As the formulations are solid, unlike liquids and semi-solids, it is necessary to use the punch to compress each sample against the crystal. The spectra of each formulation and each pure component were obtained using MATLAB version R2021b (MathWorks, Natick, US).

Table 5: Parameters applied to FTIR readings

Parameters	Condition applied
Mode	Attenuated Total Reflectance
Detector	Diamond ATR Crystal Plate
Resolution	2 cm^{-1}
Scans	32
Spectral Range	4000-500 cm^{-1}
Replicates	Triplicates

5.5. Chemometric Analysis

The MCR – Alternating Least Squares analysis was done using Solo - Stand Alone Chemometrics Software ((Eigenvector Research Incorporated, Manson, Washington, US). The MCR – Weighted Linear Combination Analysis was performed on MATLAB (R2021b version).

6. Results and Discussion

6.1. Pure Component Spectra

Figures 11 to 13 contain the pure FTIR - spectrum of each component present in the formulations. The figure 14 corresponds to the overlap of all the four pure component FTIR - spectra. All figures were obtained in MATLAB (R2021b version) software.

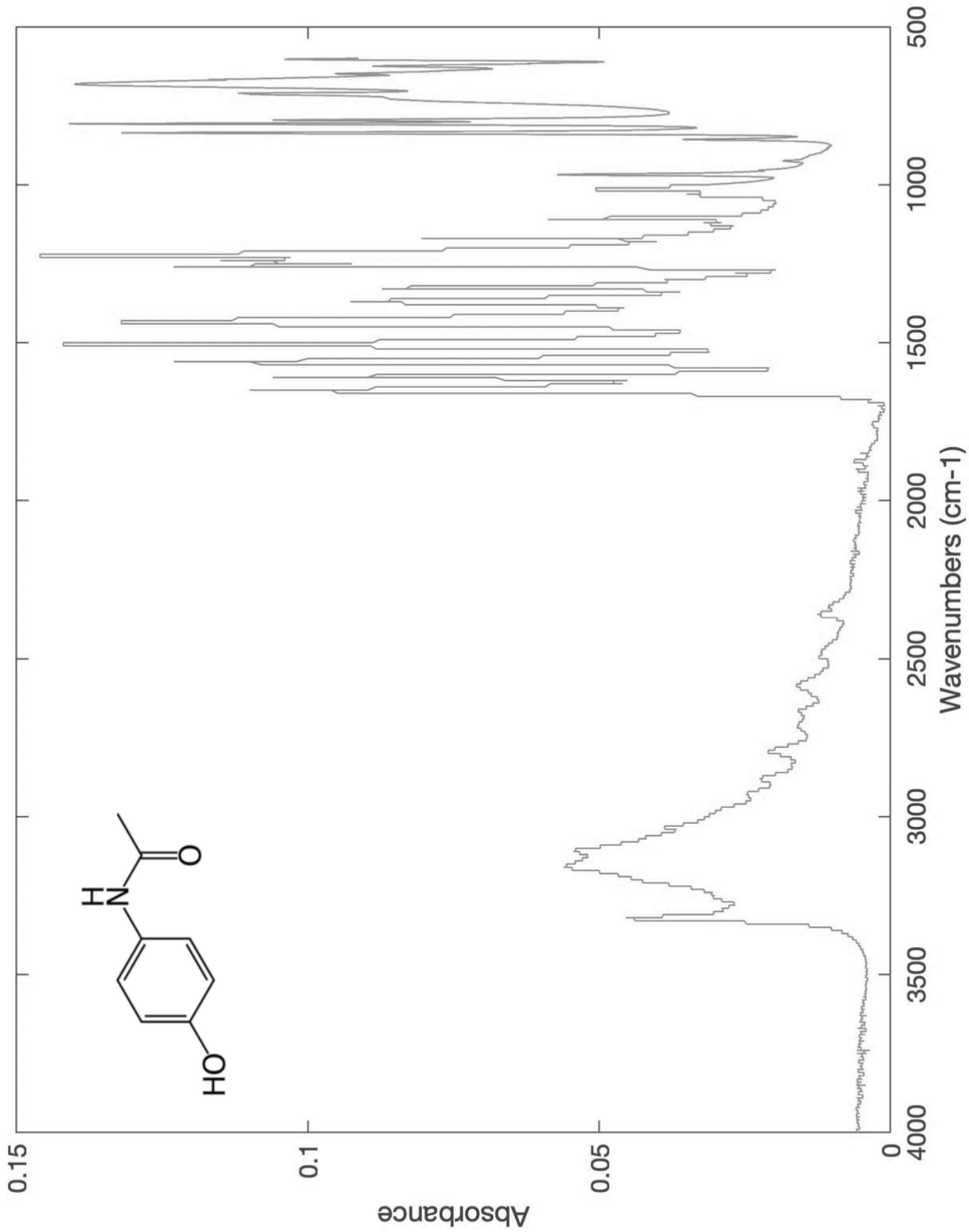

Figure 10: Pure Paracetamol FTIR-Spectrum

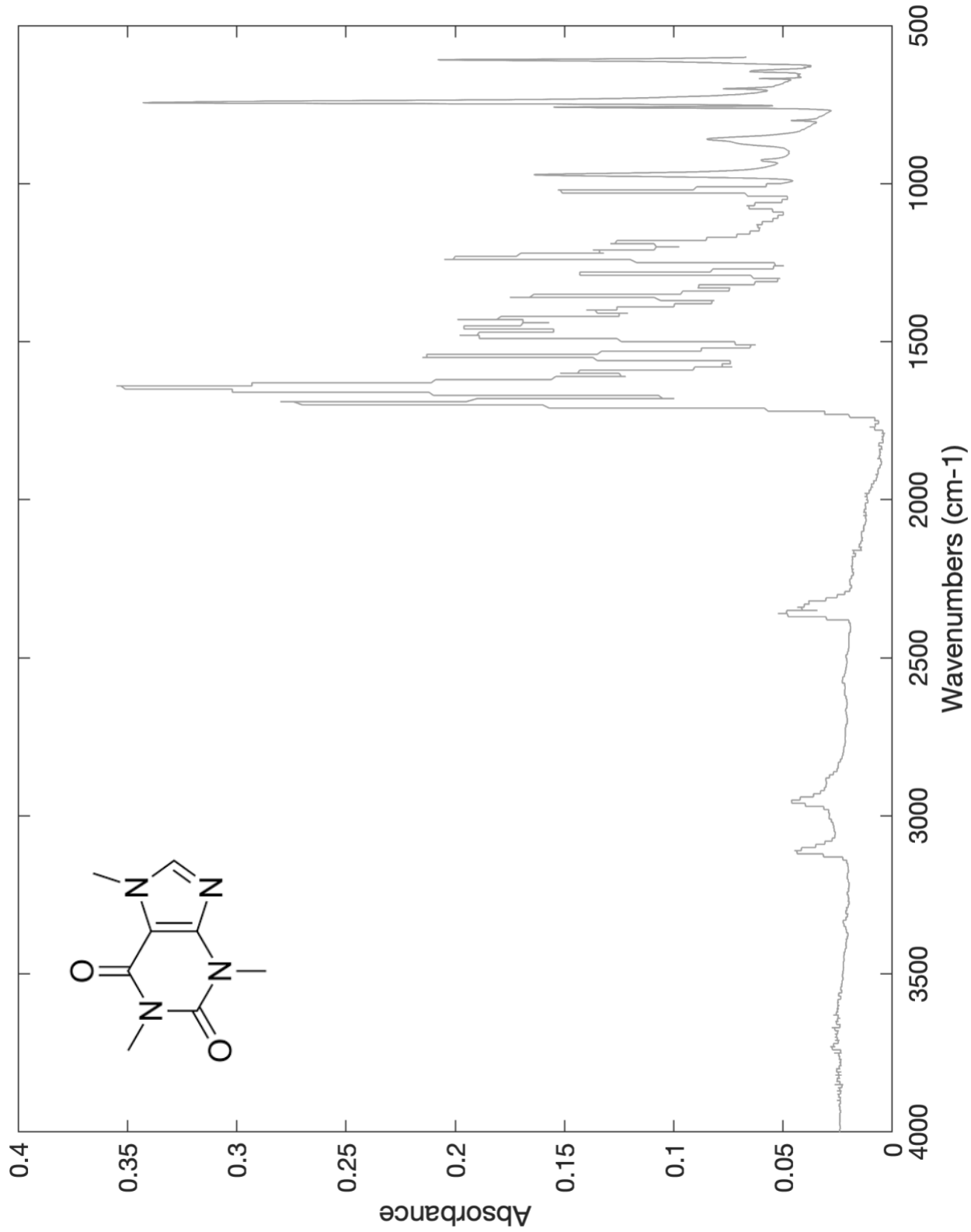

Figure 11: Pure Caffeine FTIR - Spectrum

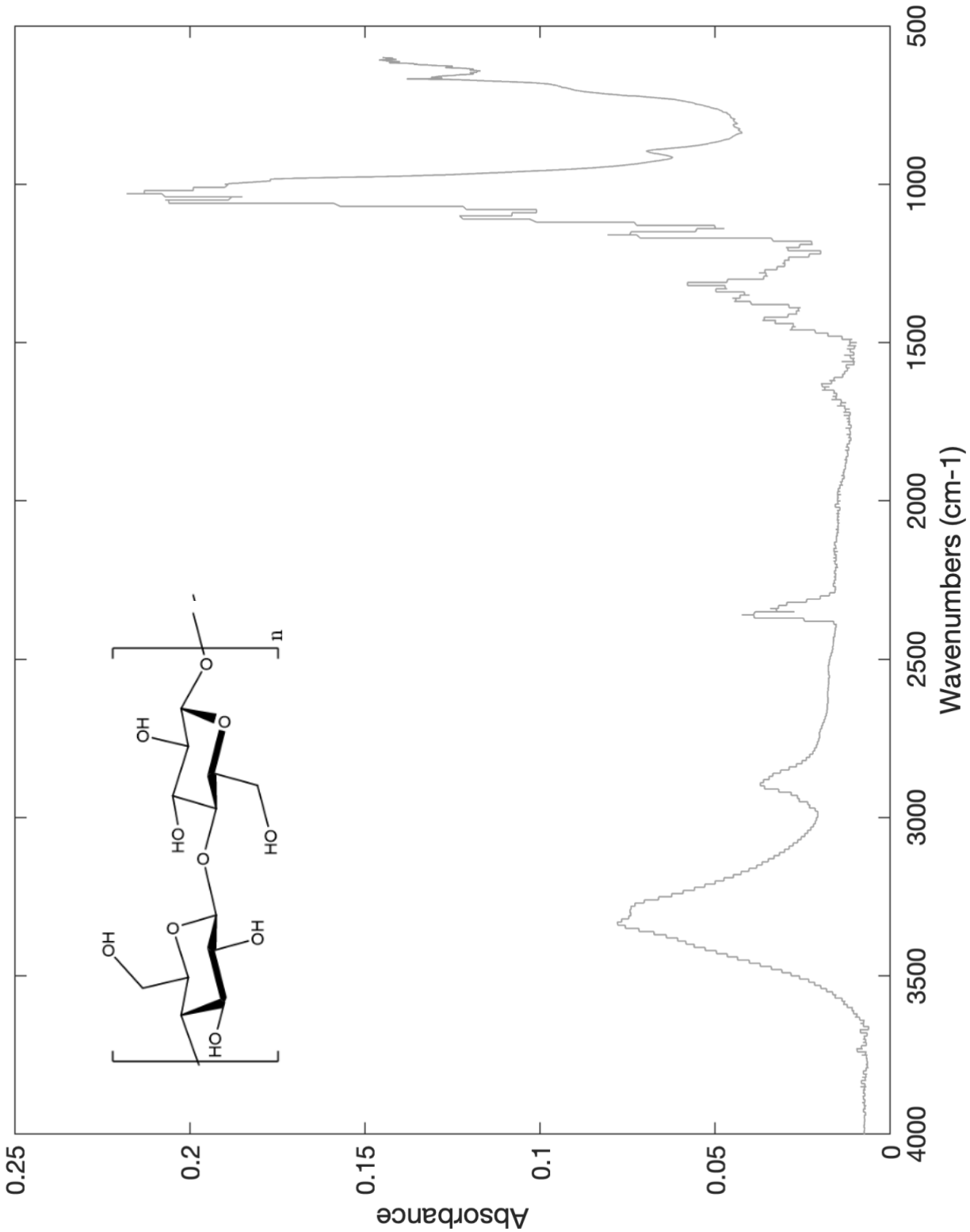

Figure 12: Pure Microcrystalline Cellulose FTIR-Spectrum

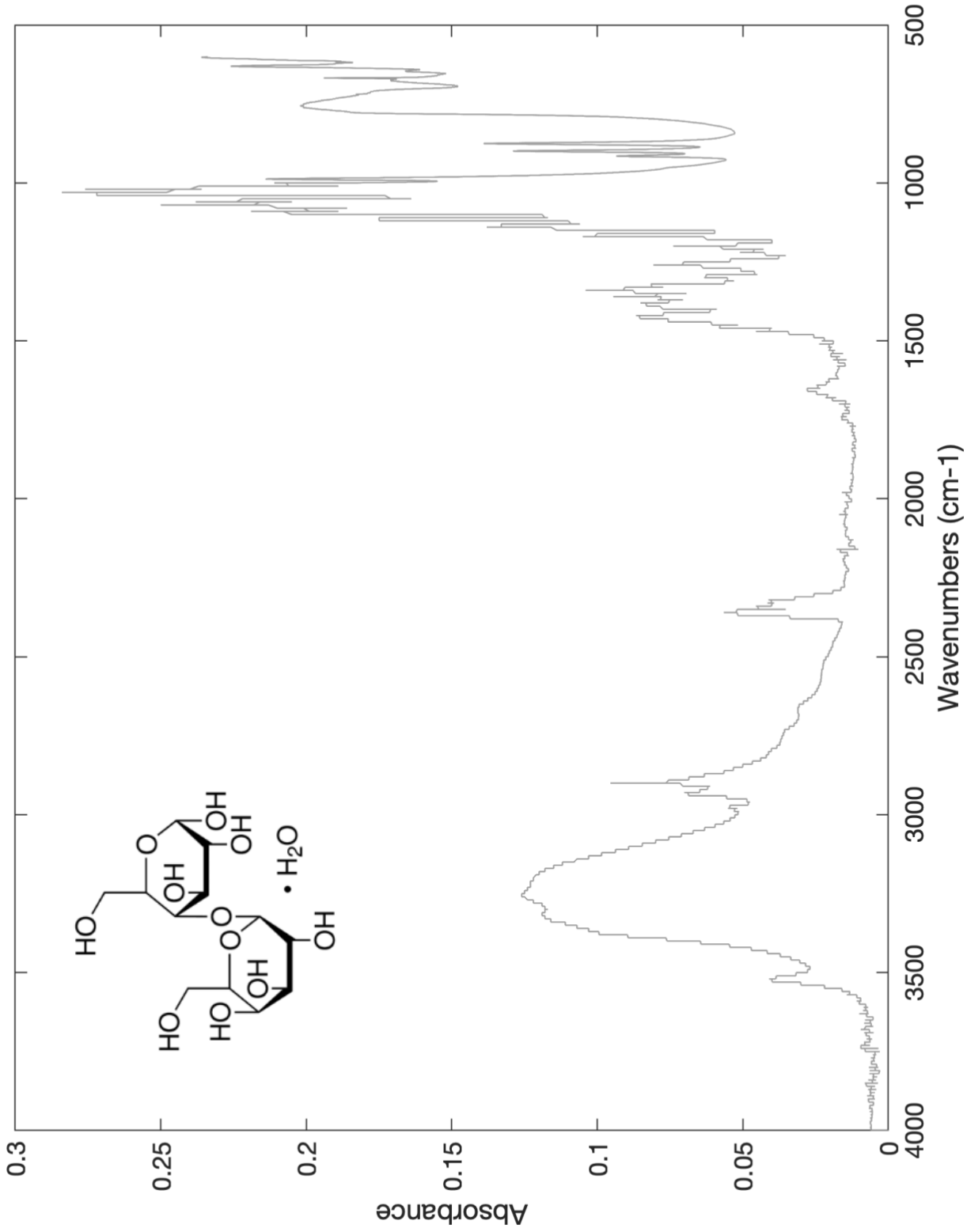

Figure 13: Pure Lactose FTIR-Spectrum

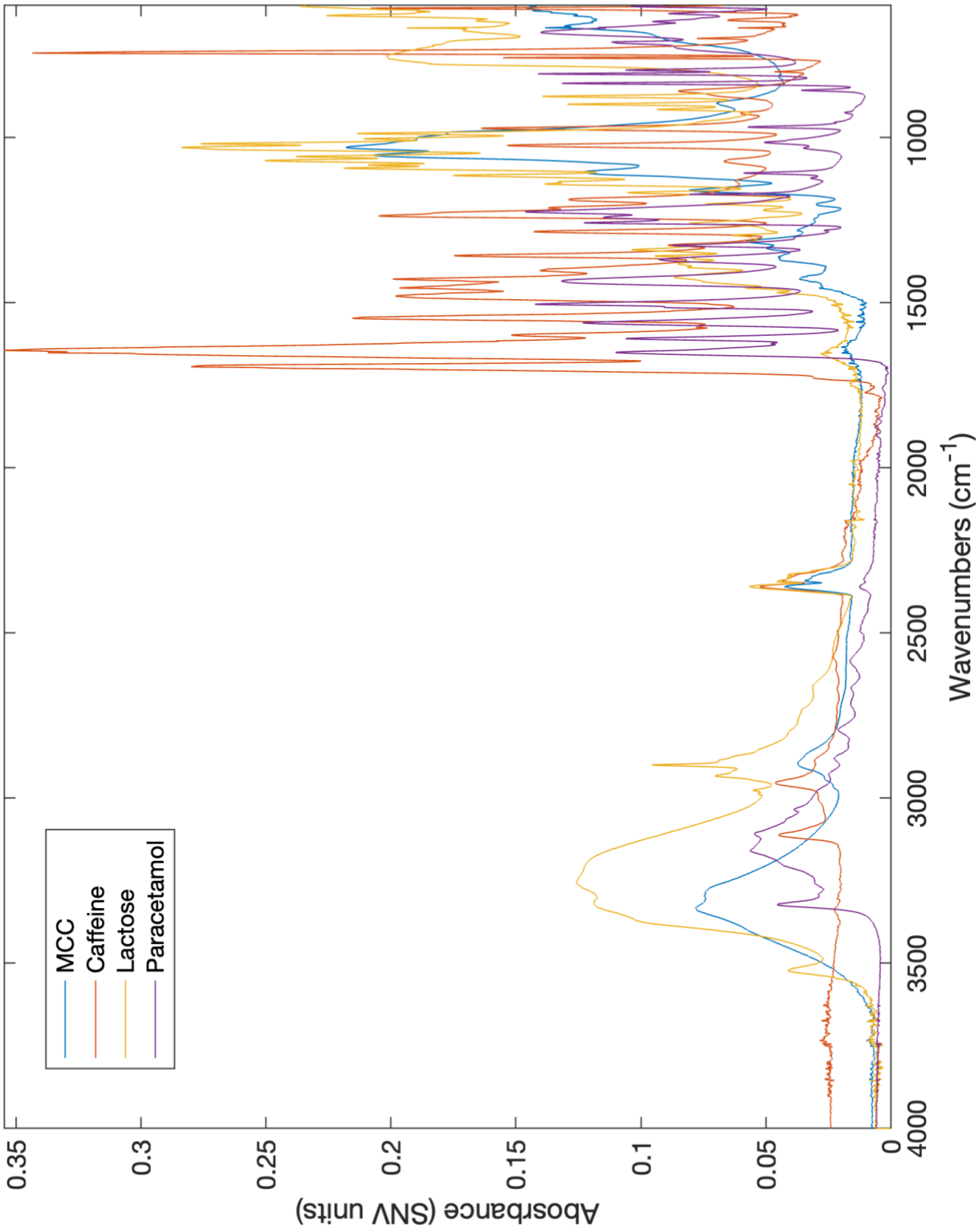

Figure 14: Overlap of all the four pure component FTIR - spectra

6.2. Spectra of the formulations

Figures 15 to 18 contain the FTIR - spectra of four formulations. Only four formulations were selected to be shown, two from each produced set.

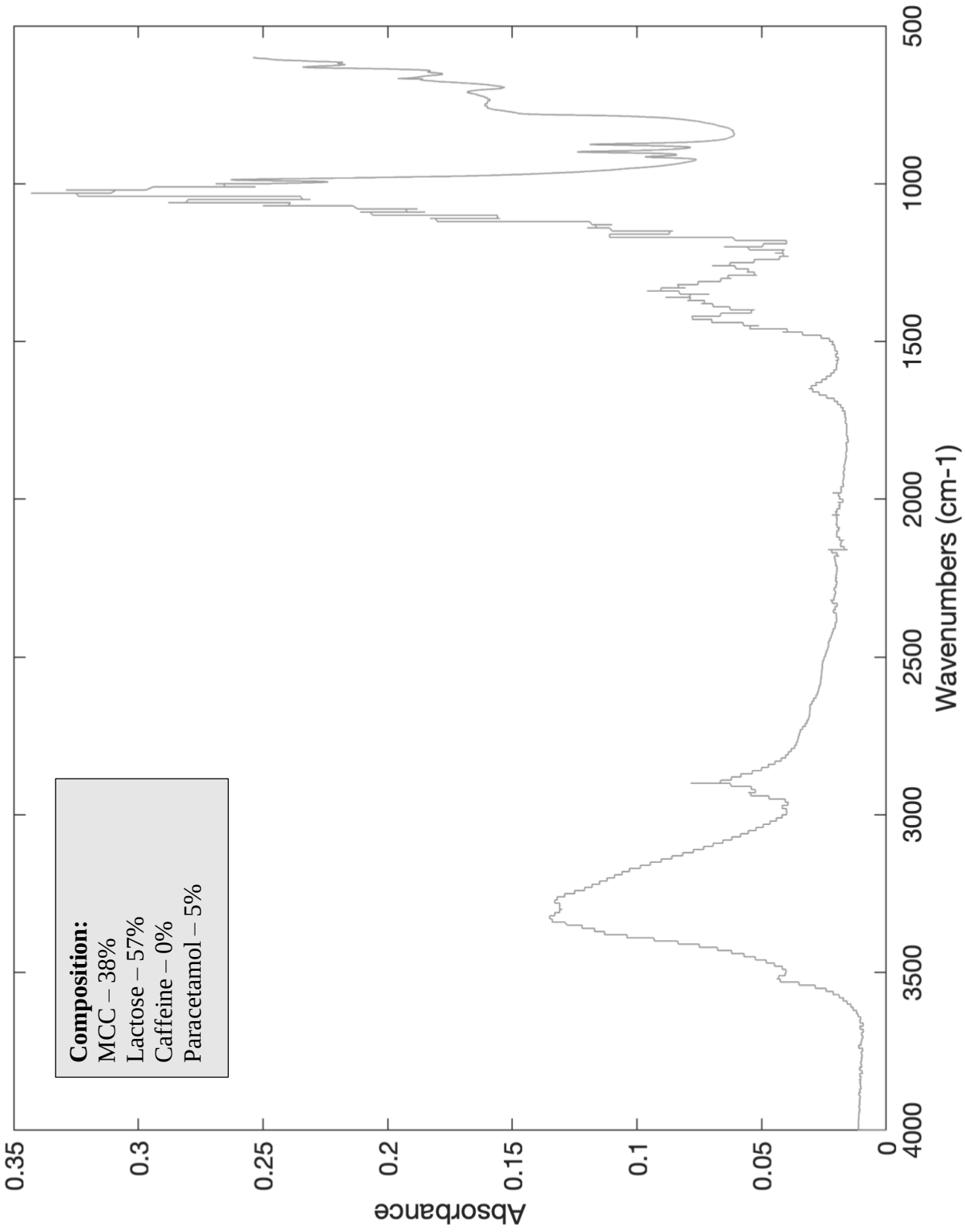

Figure 15: Formulation 1 from set 1 Spectrum

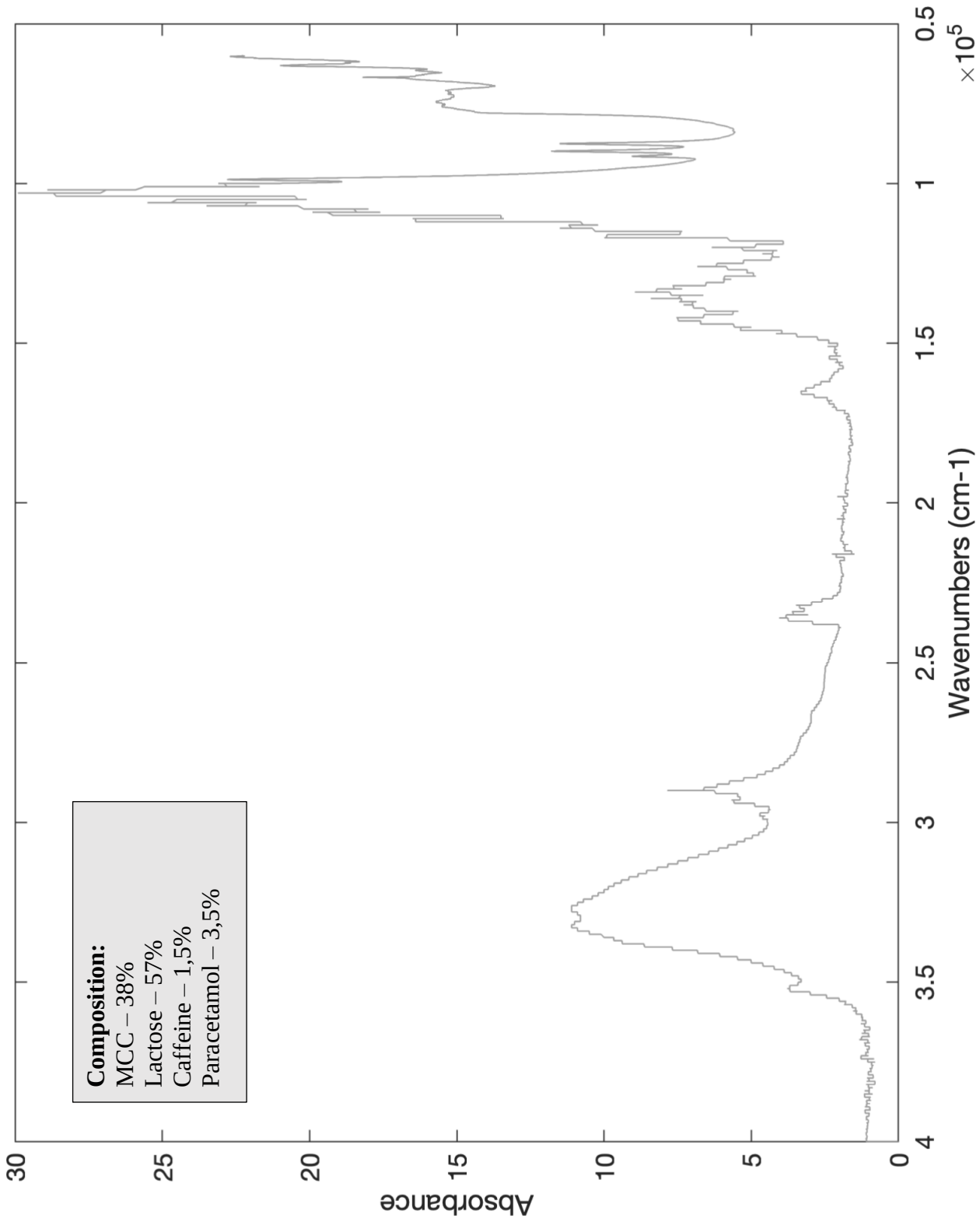

Figure 16: Formulation 4 from Set 2 Spectrum

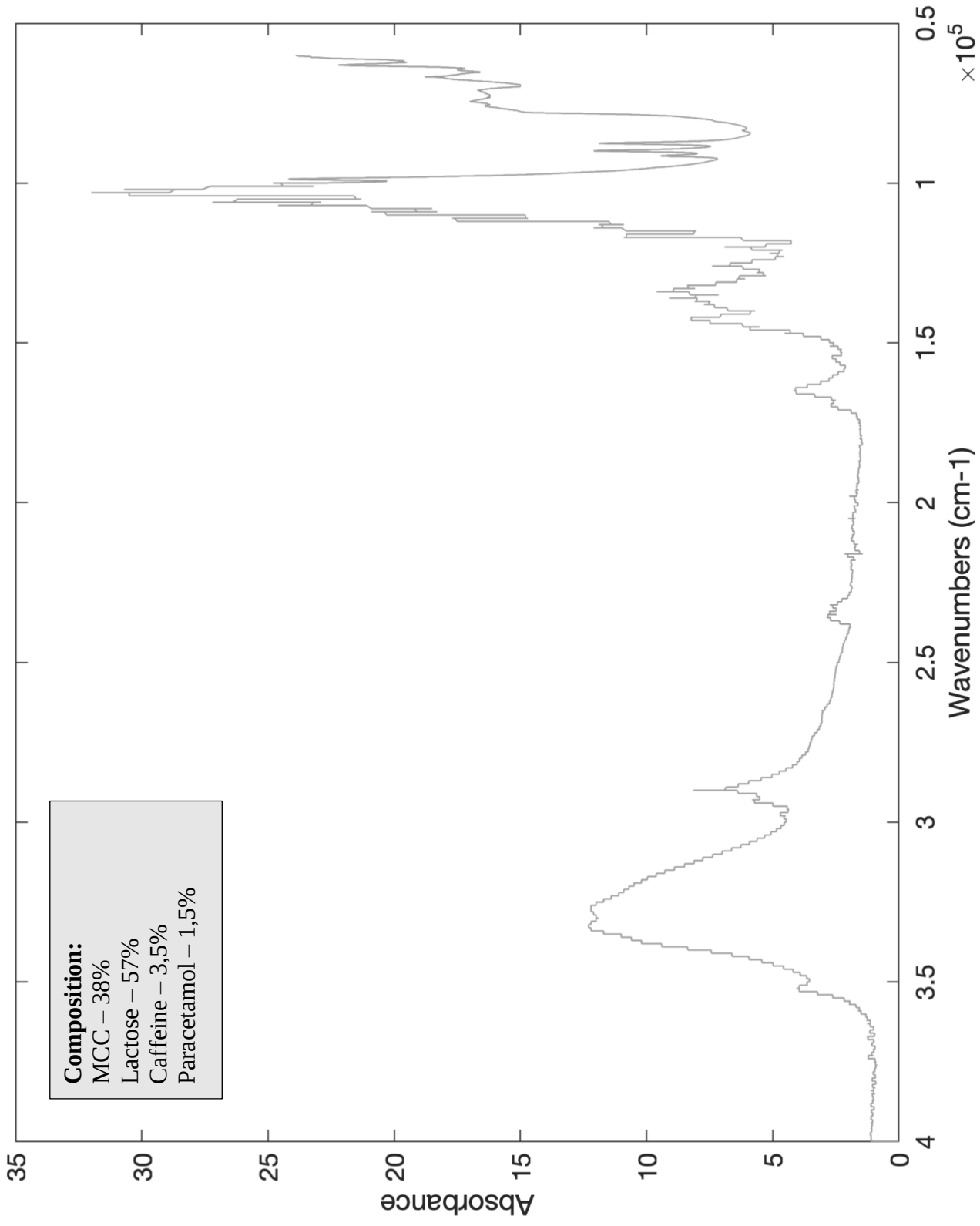

Figure 17: Formulation 8 from Set 1 Spectrum

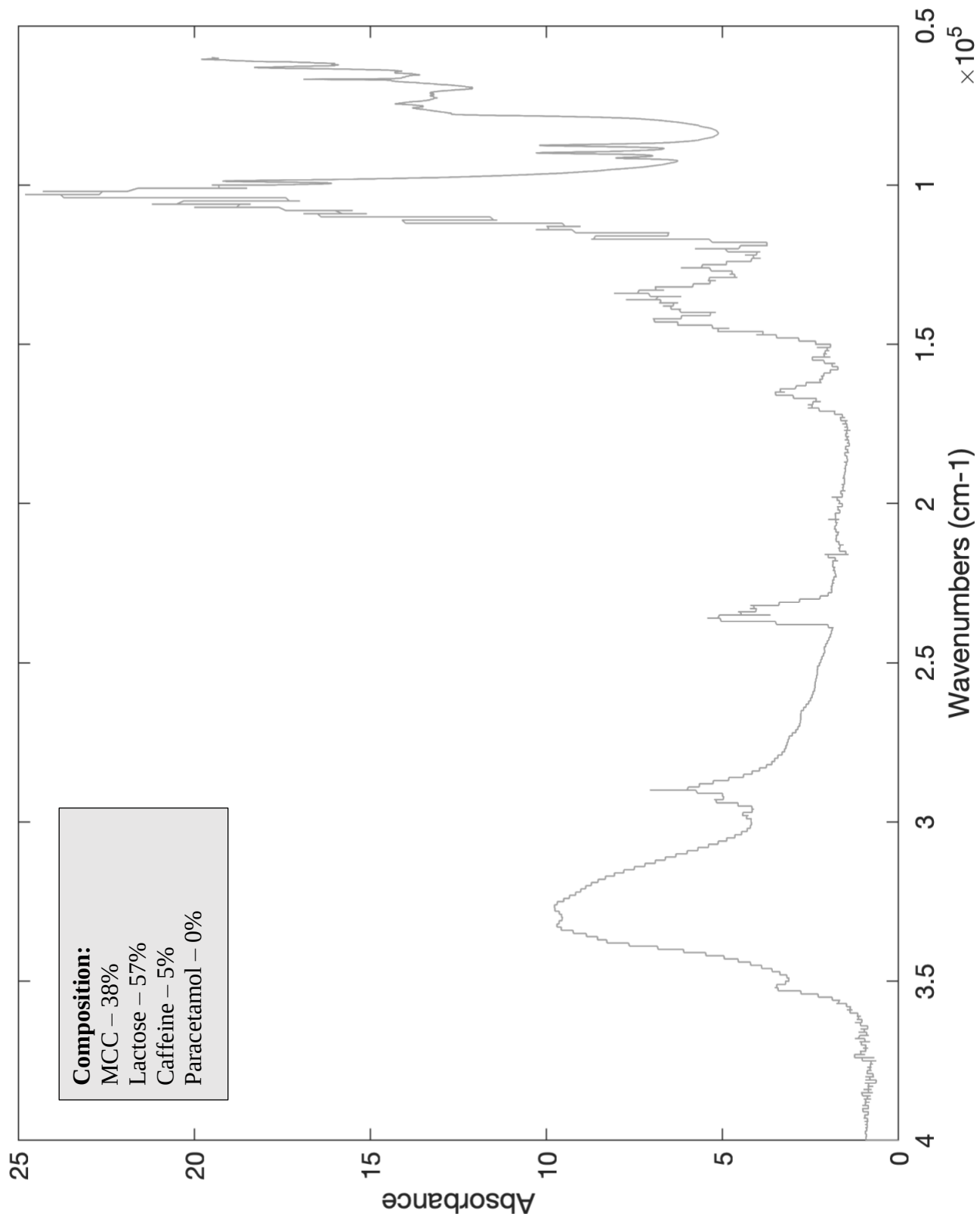

Figure 18: Formulation 11 from Set 2 Spectrum

6.3. PCA maps

Principal component analysis (PCA) is a popular technique for analyzing large datasets containing a high number of dimensions/features per observation, increasing the interpretability of data while preserving the maximum amount of information, and enabling the visualization of multidimensional data.

For the PCA maps, the zone 1805 – 2649 cm^{-1} has been excluded from all FTIR – spectra of the formulations.

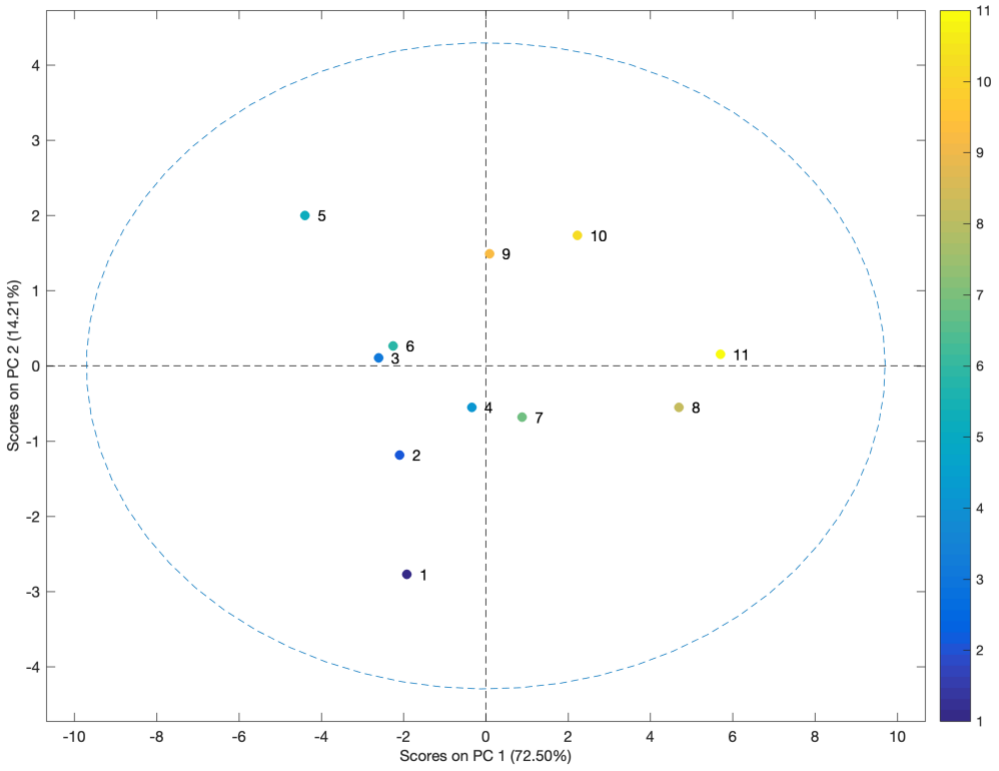

Figure 19: PCA map of the 11 formulations from set 1

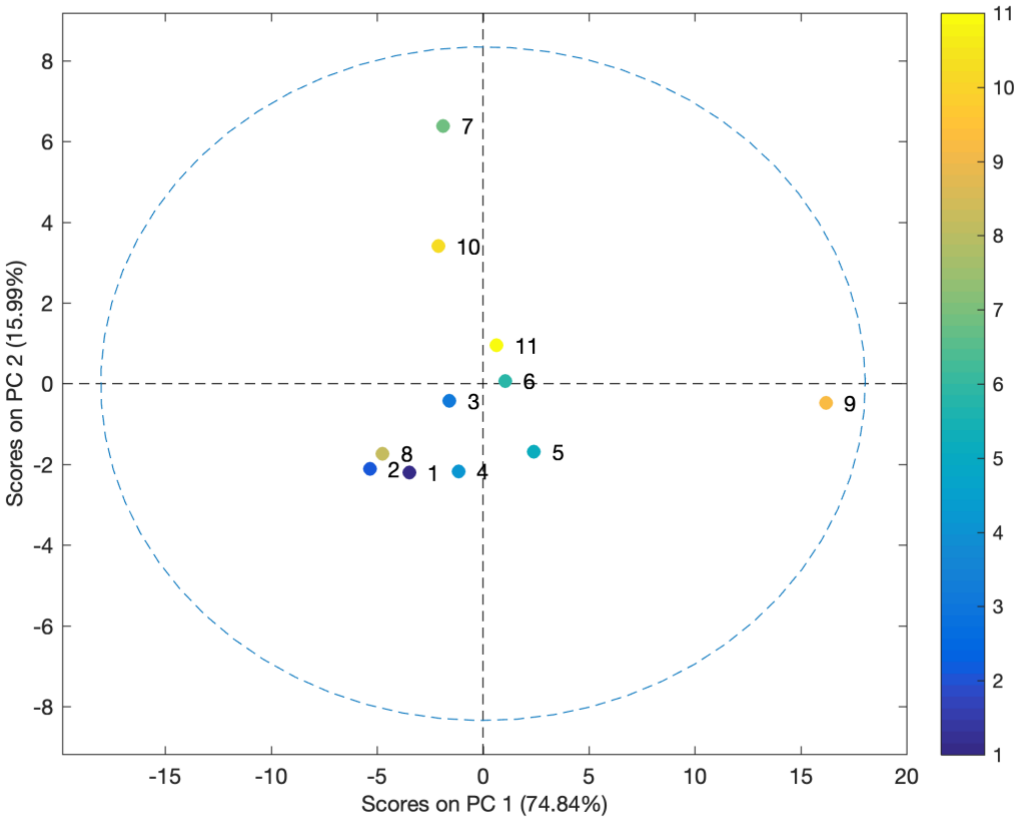

Figure 20: PCA map of the 11 formulations from set 2

By observing the figure 20, it can be seen that formulation 9 is an outlier, which is why it was excluded from the analyses.

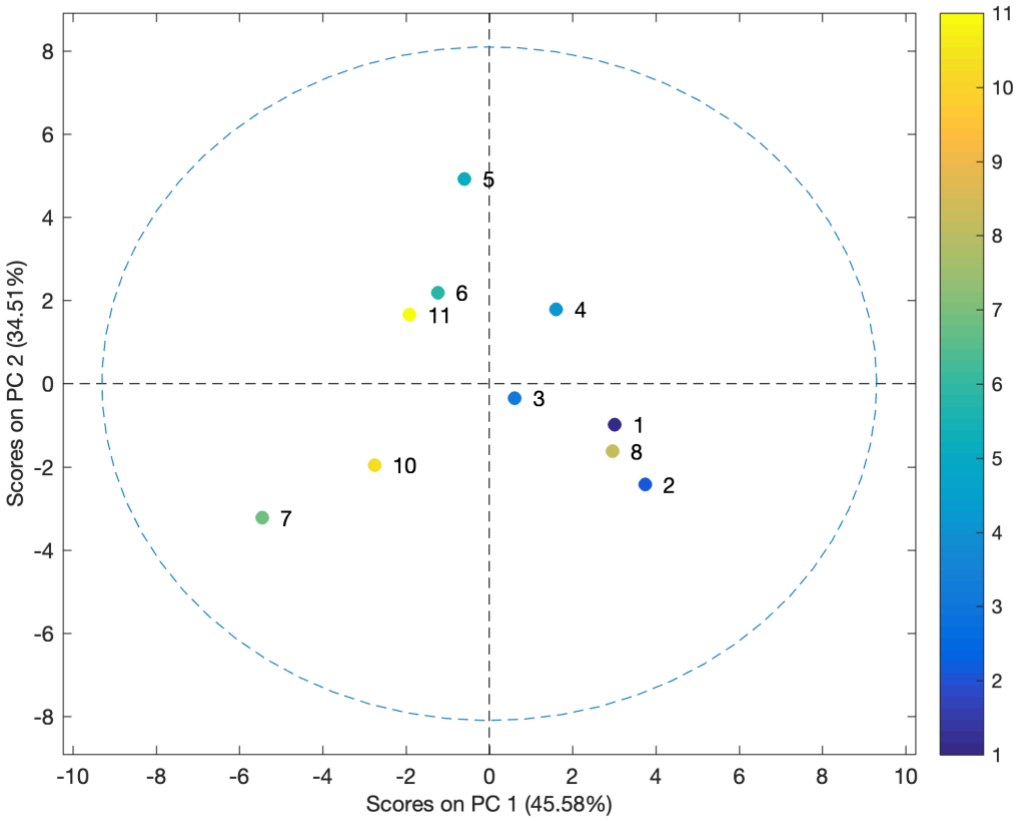

Figure 21: PCA map of 10 formulations from set 2 (formulation 9 has been removed)

In this figure, it is concluded that by removing the formulation 9, the 2 sets become very similar.

6.4. MCR - Alternating Least Squares Analysis

Methodology: The MCR analysis consisted in three parts. Since the amount of information about the formulations provided to the software was increasing.

- In a first approach, only information about the number of components present in each formulation was provided (in this case, four components).
- In a second phase, the pure spectra of the components were added and only the concentrations are going to be estimated.
- Finally, to reproduce what happens with a real drug marketed, where the concentration of the active substances is always known, we provide their concentrations to the software. First, only the concentrations of Paracetamol, then of Caffeine only and, at the end, of both.

The constraints applied in this analysis were the non-negative, equality (in the second and last part) and the closure - Previously explained in the theoretical foundation).

6.4.1. Part 1 of the MCR - Alternating Least Squares Analysis

Table 6: Results of the Set 1 – Part 1 of MCR – Alternating Least Squares Analysis (only provided the number of components present in each formulation)

Formulation no	Scores on Component 1 (28.78 %)	Scores on Component 2 (28.96 %)	Scores on Component 3 (17.15%)	Scores on Component 4 (25.10 %)	Q residuals (0.01 %)
1	0.000	0.766	0.234	0.000	0.7
2	0.238	0.407	0.000	0.355	0.9
3	0.0509	0.331	0.233	0.386	2.8
4	0.156	0.391	0.303	0.150	1.5
5	0.000	0.231	0.127	0.642	2.2
6	0.195	0.231	0.123	0.451	1.7
7	0.313	0.391	0.172	0.124	3.3
8	0.634	0.120	0.234	0.013	1
9	0.341	0.009	0.222	0.428	0.7
10	0.244	0.095	0.559	0.101	0.3
11	0.702	0.000	0.298	0.000	0.7

Table 7: Results of the Set 2 – Part 1 of MCR – Alternating Least Squares Analysis (only provided the number of components present in each formulation)

Formulation no	Scores on Component 1 (15.35 %)	Scores on Component 2 (12.34 %)	Scores on Component 3 (25.49 %)	Scores on Component 4 (21.60 %)	Q residuals (0.02 %)
1	0.096	0.086	0.000	0.818	2.2
2	0.000	0.073	0.087	0.840	1.7
3	0.149	0.199	0.054	0.598	5.9
4	0.201	0.019	0.085	0.696	2.8
5	0.358	0.002	0.127	0.512	10
6	0.251	0.129	0.256	0.364	3.4
7	0.007	0.922	0.071	0.000	0.8
8	0.002	0.000	0.331	0.666	4
9	1.000	0.000	0.000	0.000	1.3
10	0.000	0.202	0.798	0.000	2.9
11	0.189	0.076	0.579	0.155	1.6

Observing the 2 tables of results, both for set 1 and set 2, it is noticeable that the information provided was not enough for the model to work.

In set 1, the fit of the model was 99.99% and in set 2, 99.98%. The results obtained for both sets were very similar. In this part of the analysis, it is important to realize that few restrictions were applied to the method, so, mathematically, the model fit very well. However, it is impossible to match each score to a compound used in the formulations.

6.4.2. Part 2 of the MCR - Alternating Least Squares Analysis

Once the pure component spectra have been fixed, it is possible to associate each component (score) with a particular compound that depends on the order in which they were placed in the matrix of pure spectra (loadings).

Table 8: Results of the Set 1 – Part 2 of MCR – Alternating Least Squares Analysis Set 1 (Addition of pure components FTIR spectra)

Formulation no	Scores on MCC (34.11 %)	Scores on Caffeine (0.12 %)	Scores on Lactose (65.03 %)	Scores on Paracetamol (0.00 %)	Q residuals (0.74 %)
1	0.439	0.000	0.561	0.000	109
2	0.414	0.001	0.584	0.001	94
3	0.446	0.010	0.532	0.013	114
4	0.426	0.015	0.559	0.000	86
5	0.456	0.015	0.528	0.002	142
6	0.426	0.009	0.562	0.003	90
7	0.399	0.020	0.578	0.000	120
8	0.352	0.030	0.608	0.008	98
9	0.405	0.027	0.568	0.000	90
10	0.410	0.043	0.547	0.000	98
11	0.339	0.043	0.618	0.000	94

Table 9: Results of the Set 2 – Part 2 of MCR – Alternating Least Squares Analysis (Addition of pure components FTIR spectra)

Formulation no	Scores on MCC (22.01 %)	Scores on Caffeine (0.10 %)	Scores on Lactose (77.39 %)	Scores on Paracetamol (0.07 %)	Q residuals (0.43 %)
1	0.394	0.006	0.600	0.000	50
2	0.423	0.010	0.567	0.000	51
3	0.358	0.009	0.626	0.007	43
4	0.359	0.020	0.620	0.000	61
5	0.304	0.022	0.674	0.000	74
6	0.316	0.027	0.656	0.000	48
7	0.357	0.023	0.554	0.066	108
8	0.410	0.026	0.564	0.000	54
9	0.072	0.003	0.925	0.000	32
10	0.353	0.038	0.601	0.009	85
11	0.316	0.036	0.648	0.000	54

In the second part of the analysis, the method resulted much better with the addition of pure spectra, however the results were still far from the real concentrations.

For set 1 the fit was 99.26% and for set 2 99.57%. In general, MCC values can be seen to be around 30-40%, a value closer to the real (38%) and Lactose to be around 50-60% closer of the real as well (57%). However, the detection of Paracetamol was practically nil in every formulation.

The justification for these results has to do with the fact that in this model it was assumed that both components contribute equally to the spectrum of the mixture (have the same molar absorptivity); and it could also be related to the fact that Paracetamol is in low concentration and has absorption zones in the IR that overlap with other compounds.

Looking at concrete examples:

Table 10: Results of Part 2 of MCR – Alternating Least Squares Analysis: Experimental values vs. predicted values of formulation 1 from set 1

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	43,9	15,5	5.9
Caffeine	0	0	-	0
Lactose	57	56.1	1.6	0.9
Paracetamol	5	0	-	5

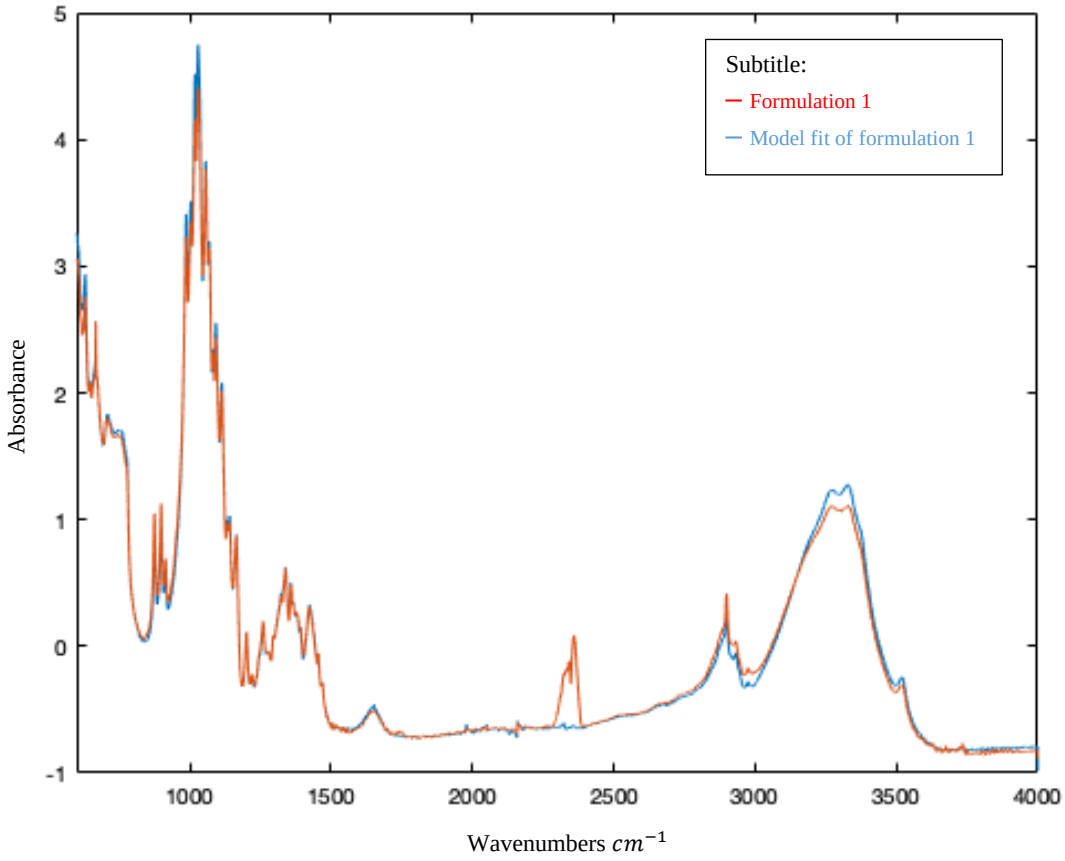

Figure 22: Overlap of the real spectrum of formulation 1 and the fit of the model

Table 11: Results of Part 2 of MCR – Alternating Least Squares Analysis: Experimental values vs. predicted values of formulation 4 from set 2

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	35.9	5.5	2.1
Caffeine	1.5	0	100	1.5
Lactose	57	62	8.7	5
Paracetamol	3.5	2	42.9	1.5

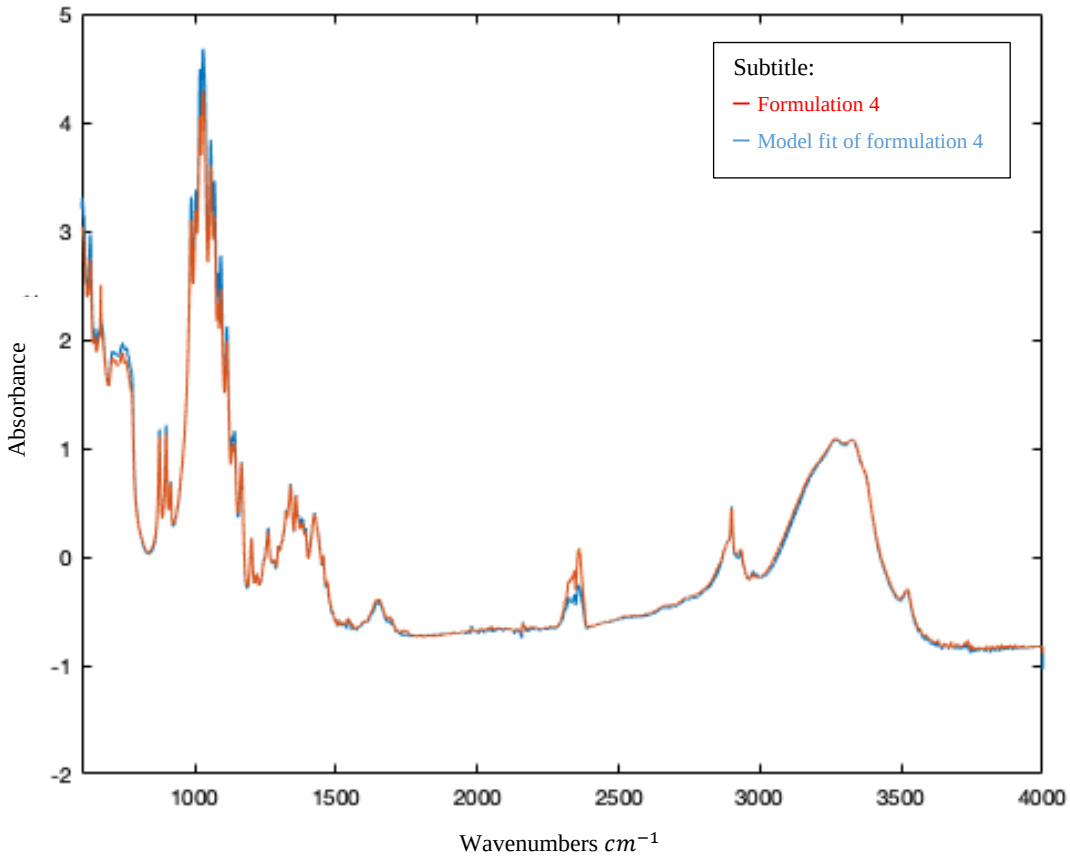

Figure 23: Overlap of the real spectrum of formulation 4 and the fit of the model

Table 12: Results of Part 2 of MCR – Alternating Least Squares Analysis: Experimental values vs. predicted values of formulation 8 from set 1

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	35.2	7.4	2.8
Caffeine	3.5	3	14.3	0.5
Lactose	57	60.8	6.7	3.8
Paracetamol	1.5	0.8	46.7	0.7

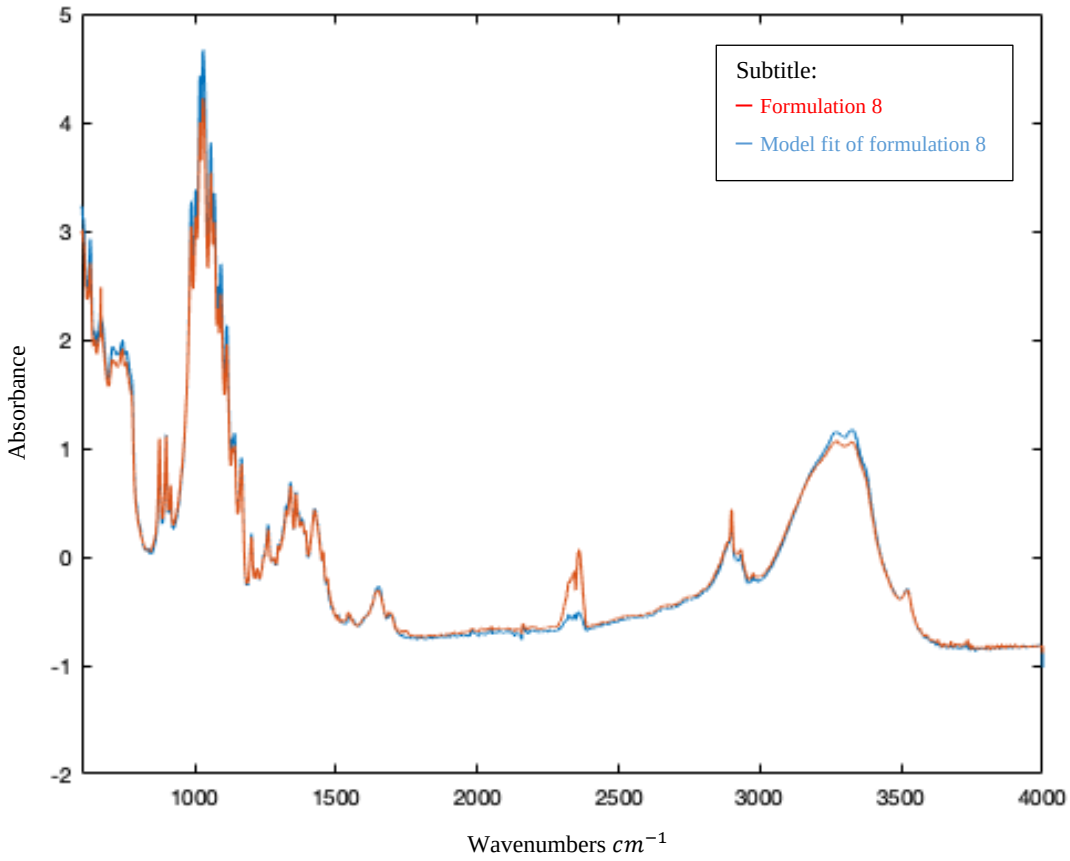

Figure 24: Overlap of the real spectrum of formulation 8 and the fit of the model

Table 13: Results of Part 2 of MCR – Alternating Least Squares Analysis: Experimental values vs. predicted values of formulation 11 from set 2

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	31.6	16.8	6.4
Caffeine	5	3.6	28	1.4
Lactose	57	64.8	13.7	7.8
Paracetamol	0	0	-	0

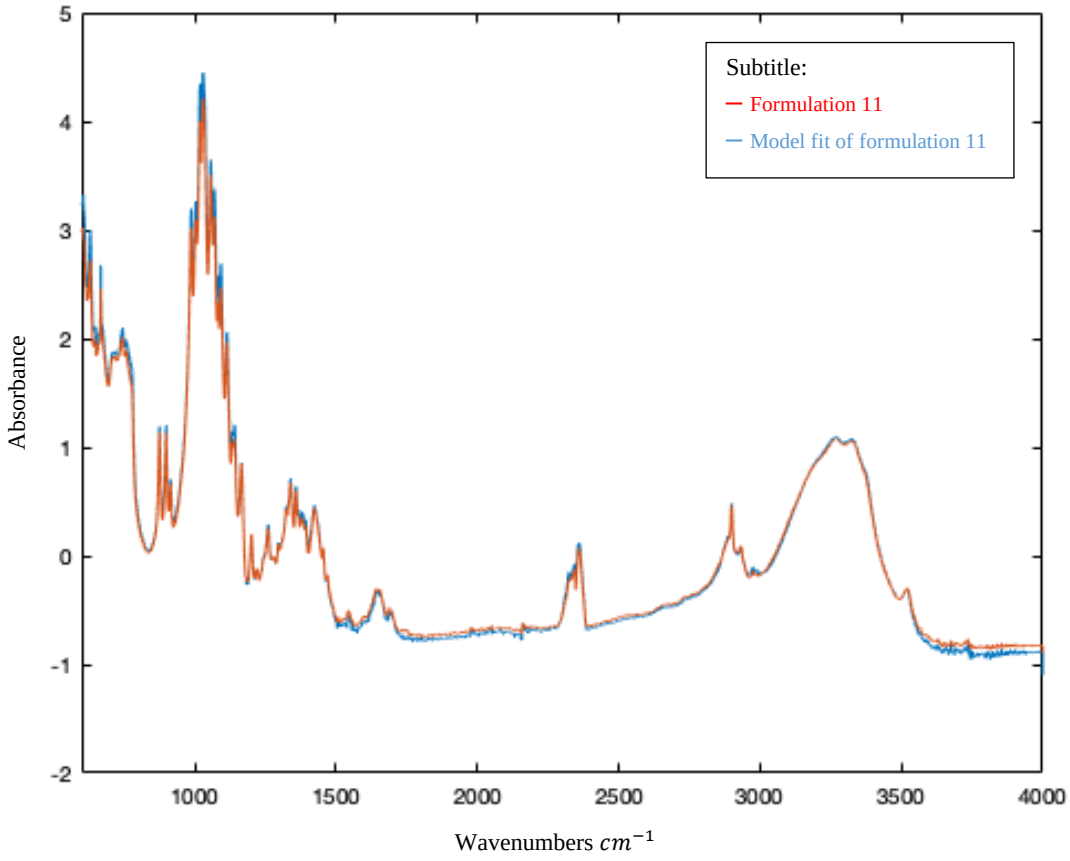

Figure 25: Overlap of the real spectrum of formulation 11 and the fit of the model

6.4.3. Part 3 of the MCR - Alternating Least Squares Analysis

Table 14: Results of the Set 1 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the Paracetamol 's concentrations)

Formulation no	Scores on MCC (36.42 %)	Scores on Caffeine (0.09 %)	Scores on Lactose (62.48 %)	Scores on Paracetamol (0.16 %)	Q residuals (0.85 %)
1	0.446	0.000	0.507	**0.047**	167
2	0.421	0.000	0.536	**0.043**	138
3	0.452	0.000	0.510	**0.038**	128
4	0.434	0.000	0.532	**0.033**	103
5	0.463	0.002	0.506	**0.029**	155
6	0.431	0.000	0.545	**0.024**	99
7	0.404	0.012	0.565	**0.019**	127
8	0.355	0.028	0.604	**0.014**	100
9	0.407	0.023	0.560	**0.010**	94
10	0.412	0.041	0.543	**0.005**	100
11	0.356	0.043	0.601	**0.000**	97

Table 15: Results of the Set 1 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the Caffeine's concentrations)

Formulation no	Scores on MCC (34.42 %)	Scores on Caffeine (0.16 %)	Scores on Lactose (64.65 %)	Scores on Paracetamol (0.00 %)	Q residuals (0.77 %)
1	0.439	**0.000**	0.561	0.000	109
2	0.413	**0.005**	0.582	0.000	96
3	0.446	**0.010**	0.532	0.013	114
4	0.426	**0.014**	0.560	0.000	85
5	0.454	**0.019**	0.526	0.000	145
6	0.422	**0.024**	0.554	0.000	104
7	0.397	**0.029**	0.574	0.000	128
8	0.352	**0.033**	0.610	0.007	100
9	0.402	**0.038**	0.560	0.000	104
10	0.412	**0.042**	0.547	0.004	97
11	0.355	**0.047**	0.598	0.000	102

Table 16: Results of the Set 1 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the concentrations of both APIs)

Formulation no	Scores on MCC (36.71 %)	Scores on Caffeine (0.16 %)	Scores on Lactose (62.02 %)	Scores on Paracetamol (0.17 %)	Q residuals (0.94 %)
1	0.446	**0.000**	0.507	**0.047**	167
2	0.420	**0.005**	0.533	**0.043**	146
3	0.449	**0.009**	0.503	**0.038**	140
4	0.431	**0.014**	0.522	**0.033**	122
5	0.458	**0.019**	0.495	**0.028**	177
6	0.426	**0.023**	0.528	**0.023**	133
7	0.400	**0.028**	0.553	**0.019**	149
8	0.354	**0.033**	0.599	**0.014**	107
9	0.404	**0.038**	0.549	**0.009**	115
10	0.411	**0.042**	0.542	**0.005**	102
11	0.355	**0.047**	0.598	**0.000**	102

Table 17: Results of the Set 2 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the Paracetamol 's concentrations)

Formulation no	Scores on MCC (22.77 %)	Scores on Caffeine (0.10 %)	Scores on Lactose (75.43 %)	Scores on Paracetamol (0.15 %)	Q residuals (0.55 %)
1	0.403	0.000	0.549	**0.047**	101
2	0.432	0.000	0.526	**0.043**	89
3	0.366	0.000	0.596	**0.038**	64
4	0.369	0.005	0.593	**0.033**	83
5	0.313	0.009	0.650	**0.029**	92
6	0.323	0.016	0.637	**0.024**	59
7	0.343	0.045	0.593	**0.019**	109
8	0.414	0.019	0.553	**0.014**	62
9	0.078	0.000	0.914	**0.010**	36
10	0.352	0.040	0.604	**0.005**	84
11	0.357	0.037	0.606	**0.000**	63

Table 18: Results of the Set 2 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the Caffeine's concentrations)

Formulation no	Scores on MCC (22.79 %)	Scores on Caffeine (0.15 %)	Scores on Lactose (76.52 %)	Scores on Paracetamol (0.07 %)	Q residuals (0.47 %)
1	0.396	**0.000**	0.602	0.002	48
2	0.424	**0.005**	0.570	0.001	48
3	0.358	**0.010**	0.626	0.006	43
4	0.360	**0.014**	0.625	0.000	56
5	0.305	**0.019**	0.676	0.000	72
6	0.317	**0.024**	0.657	0.002	46
7	0.355	**0.027**	0.554	0.063	112
8	0.409	**0.033**	0.558	0.000	63
9	0.074	**0.038**	0.887	0.000	69
10	0.352	**0.042**	0.600	0.006	89
11	0.354	**0.047**	0.600	0.000	78

Table 19: Results of the Set 2 – Part 3 of MCR – Alternating Least Squares Analysis (Addition of the concentrations of both APIs)

Formulation no	Scores on MCC (24.13 %)	Scores on Caffeine (0.15 %)	Scores on Lactose (74.94 %)	Scores on Paracetamol (0.15 %)	Q residuals (0.62 %)
1	0.403	**0.000**	0.549	**0.047**	101
2	0.431	**0.005**	0.522	**0.042**	96
3	0.364	**0.009**	0.589	**0.038**	77
4	0.367	**0.014**	0.586	**0.033**	95
5	0.311	**0.019**	0.642	**0.028**	105
6	0.322	**0.024**	0.631	**0.024**	69
7	0.346	**0.028**	0.607	**0.019**	89
8	0.411	**0.033**	0.542	**0.014**	82
9	0.078	**0.038**	0.874	**0.010**	82
10	0.351	**0.042**	0.602	**0.005**	88
11	0.354	**0.047**	0.599	**0.000**	78

In this last part of the analysis, it is possible to see that, as expected, when the concentrations of the active substances are forced, the results obtained are much closer to the real ones.

6.4.4. Summary of the MCR - Alternating Least Squares Analysis

Table 20: Sum of Squared errors for each component of the 11 formulations from set 1, in each part of the MCR - Alternating Least Squares Analysis

Set 1	E^2 Sum	E^2 Sum	E^2 Sum	E^2 Sum
Component:	**MCC**	**Lactose**	**Paracetamol**	**Caffeine**
Analysis Part 2 (addition of the pure FTIR-spectra)	233.5	80.2	82.2	5.5
Analysis Part 3 (Addition of the Paracetamol's concentrations)	271.5	178.6	**x**	20.3
Analysis Part 3 (Addition of the Caffeine's concentrations)	214.9	69.9	88.5	**x**
Analysis Part 3 (Addition of the concentrations of both APIs)	247.6	226.6	**x**	**x**

When observing this table, contrary to what would be expected, in part two of the analysis, where only the FTIR – spectra of the pure components were provided to the method, the E^2 Sum for each one was smaller. It is possible to perceive the difficulty of the method in detecting Paracetamol, because when Paracetamol's concentrations are added, the error for the other components greatly increases and when Caffeine's concentrations were added, the lowest E^2 Sum was obtained.

Table 21: Sum of Squared errors for each component of the 11 formulations from set 2, in each part of the MCR - Alternating Least Squares Analysis

Set 2	E^2 Sum	E^2 Sum	E^2 Sum	E^2 Sum
Component	**MCC**	**Lactose**	**Paracetamol**	**Caffeine**
Analysis Part 2 (addition of the pure FTIR-spectra)	187.4	320.9	77.5	4.5
Analysis Part 3 (Addition of the Paracetamol's concentrations)	151.3	171.41	**x**	11
Analysis Part 3 (Addition of the Caffeine's concentrations)	160.6	281.9	101.3	**x**
Analysis Part 3 (Addition of the concentrations of both APIs)	153.1	162.9	**x**	**x**

For set 2, the conclusions are the same as mentioned for set 1. Although less noticeable, since the E^2 Sum for MCC and Lactose is bigger in the part 2 of the analysis, i.e., this is what is expected given that in this part is where less information was provided to the method, we were also able to observe an increase of the E^2 Sum of Caffeine from 4.5 to 11, once the Paracetamol concentrations are added.

6.5. MCR – Weighted Linear Combination Analysis

This method is very similar to the MCR – Alternating Least Squares, this one only approximates the Lambert-Beer law, as it considers that the result of the spectrum of each final mixture is only the multiplication of the pure spectra of each component involved with the respective concentration in which they are present. That is, that all components contribute equally to the spectrum of the final mix.

This alternative algorithm considers that all compounds have different molar absorptivities. What this algorithm manages to do is randomly generate 10 000 different weights with values fluctuating the value 1, for each component. Then it is possible to compare with the control (which in this case is to assume that all compounds weigh the same in the mixture). By reconstructing the spectra with different weights, it is possible to select the one with the smallest error and therefore the best estimate.

For this analysis, only four formulations were selected. We opted for the same ones that had previously shown their spectra: Formulation 1 from Set 1, Formulation 4 from Set 2, Formulation 8 from Set 1 and Formulation 11 from Set 2.

This analysis was divided into two parts: the first part in which concentrations were not restricted and the second part where Microcrystalline cellulose and Lactose were restricted between 30-60% each. In both parts, the pure components spectra were known. In the first part, is also possible to see the difference in the results, when the weights defined for the components were applied and when it was considered that all the components contributed equally.

For this analysis of this method and in all spectra, the zone of 1805 – 2649 cm^{-1} was deleted.

Weights defined for the components:

- Microcrystalline cellulose – 1.1318
- Caffeine – 0.9148
- Lactose – 1.0425
- Paracetamol – 0.7041

6.5.1. MCR – Weighted Linear Combination Analysis - with no restrictions

Table 22: Results of MCR – Weighted Linear Combination Analysis: Formulation 1 from set 1 – with no restrictions and with the defined weights applied

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	41	7.9	3
Caffeine	0	1.50	-	1.50
Lactose	57	55.2	3.2	1.8
Paracetamol	5	2.29	54.2	2.71

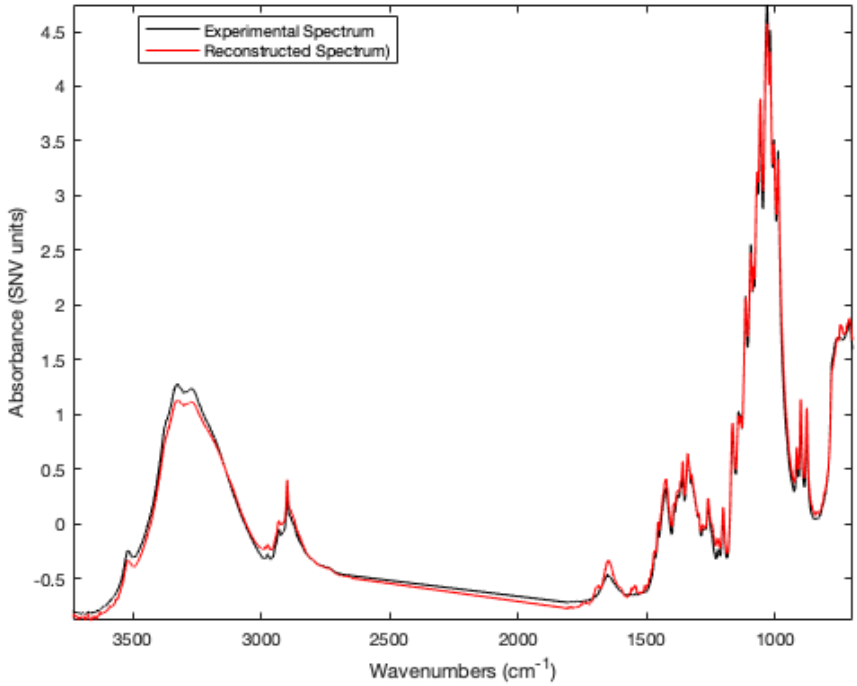

Figure 26: Overlap of the experimental spectrum of formulation 1 from set 1 and the reconstructed one (with the defined weights applied)

Table 23: Results of MCR – Weighted Linear Combination Analysis: Formulation 1 from set 1 – with no restrictions and without the defined weights

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	46.4	22.1	8.4
Caffeine	0	0.54	-	0.54
Lactose	57	51.9	8.95	5.1
Paracetamol	5	1.11	77.8	3.89

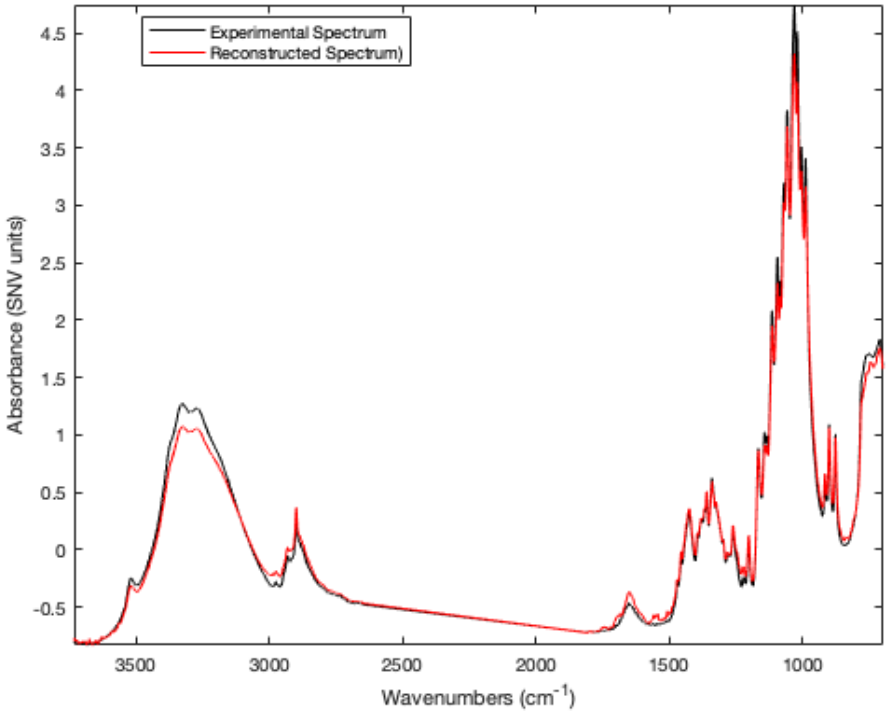

Figure 27: Overlap of the experimental spectrum of formulation 1 from set 1 and the reconstructed one (without the defined weights)

Table 24: Results of MCR – Weighted Linear Combination Analysis: Formulation 4 from set 2 – with no restrictions and with the defined weights applied

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	39.9	5	1.9
Caffeine	1.5	1.71	14	0.21
Lactose	57	56.6	0.7	1.6
Paracetamol	3.5	1.77	49.4	1.73

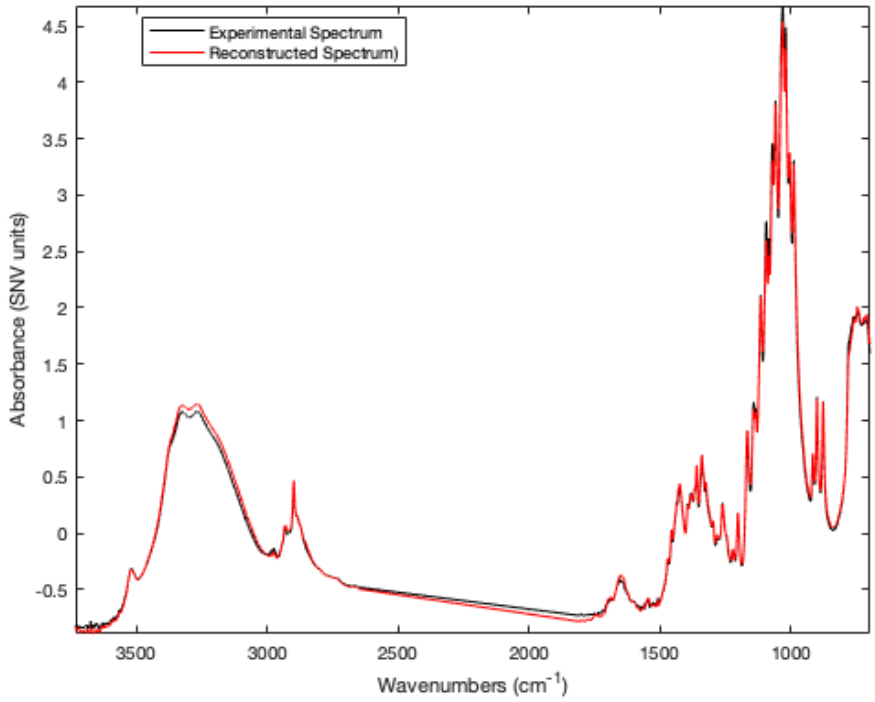

Figure 28*:* Overlap of the experimental spectrum of formulation 4 from set 2 and the reconstructed one (with the defined weights applied)

Table 25: Results of MCR – Weighted Linear Combination Analysis: Formulation 4 from set 2 – with no restrictions and without the defined weights

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	35.2	7.4	2.8
Caffeine	1.5	0.398	73.5	1.1
Lactose	57	63.2	10.9	6.2
Paracetamol	3.5	1.21	65.4	2.3

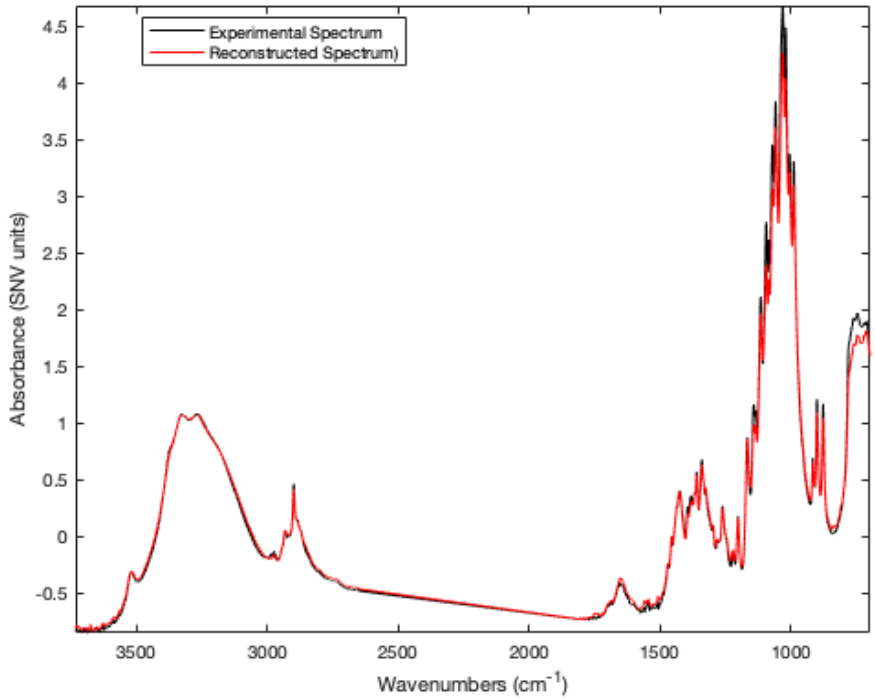

Figure 29: Overlap of the experimental spectrum of formulation 4 from set 2 and the reconstructed one (without the defined weights)

Table 26: Results of MCR – Weighted Linear Combination Analysis: Formulation 8 from set 1 – with no restrictions and with the defined weights applied

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	35.1	7.6	2.9
Caffeine	3.5	2.54	27.4	0.96
Lactose	57	60.7	6.5	3.7
Paracetamol	1.5	1.61	7.3	0.11

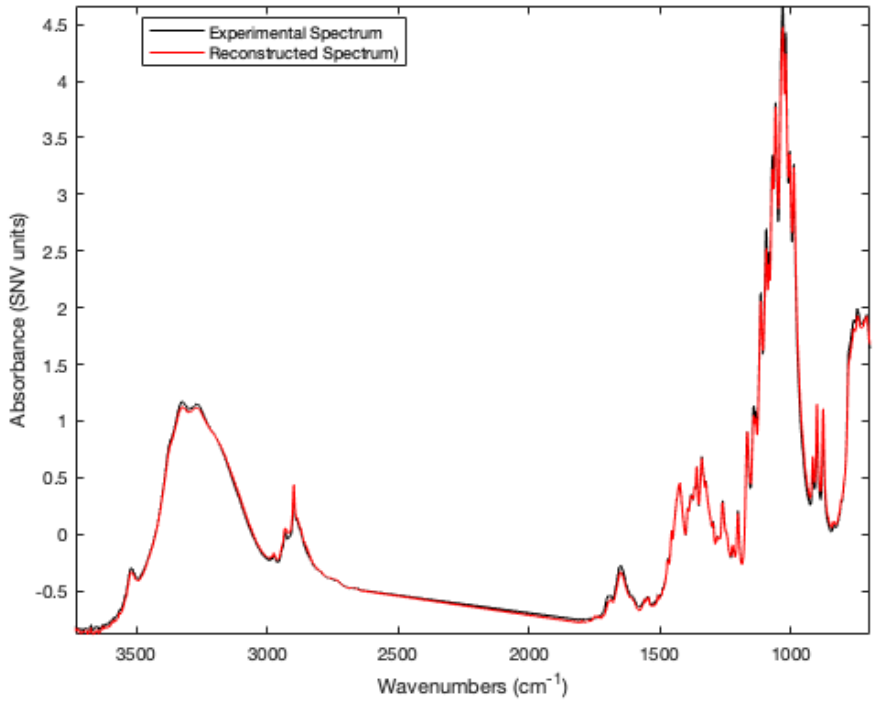

Figure 30: Overlap of the experimental spectrum of formulation 8 from set 1 and the reconstructed one (with the defined weights applied)

Table 27: Results of MCR – Weighted Linear Combination Analysis: Formulation 8 from set 1 – with no restrictions and without the defined weights

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	39.1	2.9	1.1
Caffeine	3.5	2.01	42.6	1.49
Lactose	57	58.1	1.9	1.1
Paracetamol	1.5	0.735	51.3	0.77

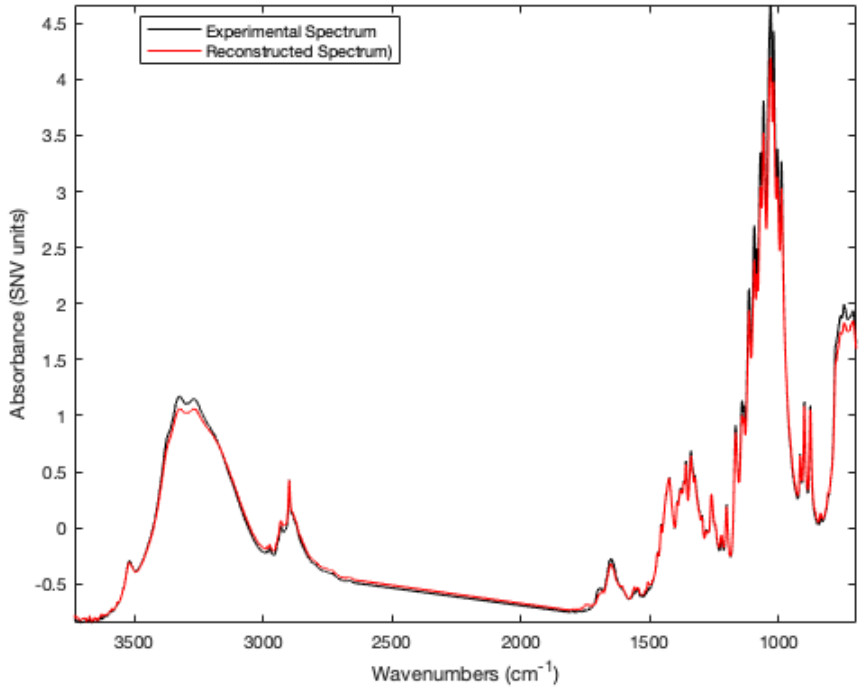

Figure 31: Overlap of the experimental spectrum of formulation 8 from set 1 and the reconstructed one (without the defined weights)

Table 28: Results of MCR – Weighted Linear Combination Analysis: Formulation 11 from set 2 – with no restrictions and with the defined weights applied

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	31.6	16.8	6.4
Caffeine	5	3.67	26.6	1.33
Lactose	57	63.7	11.8	6.7
Paracetamol	0	1.04	-	1.04

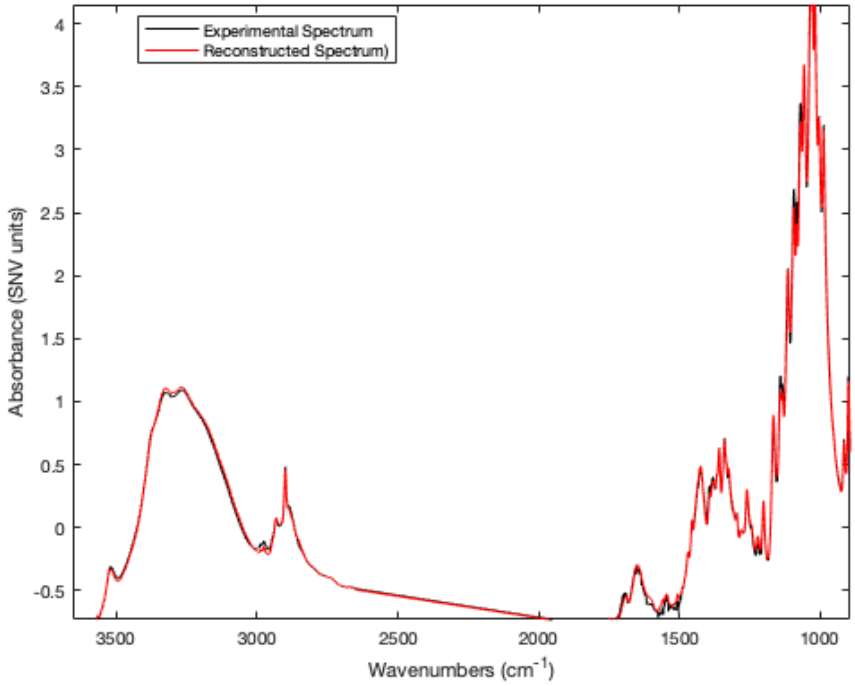

Figure 32: Overlap of the experimental spectrum of formulation 11 from set 2 and the reconstructed one (with the defined weights applied)

Table 29: Results of MCR – Weighted Linear Combination Analysis: Formulation 11 from set 2 – with no restrictions and without the defined weights

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	37.5	1.3	0.5
Caffeine	5	1.99	60.2	3.01
Lactose	57	58.5	2.6	1.5
Paracetamol	0	2.01	-	2.01

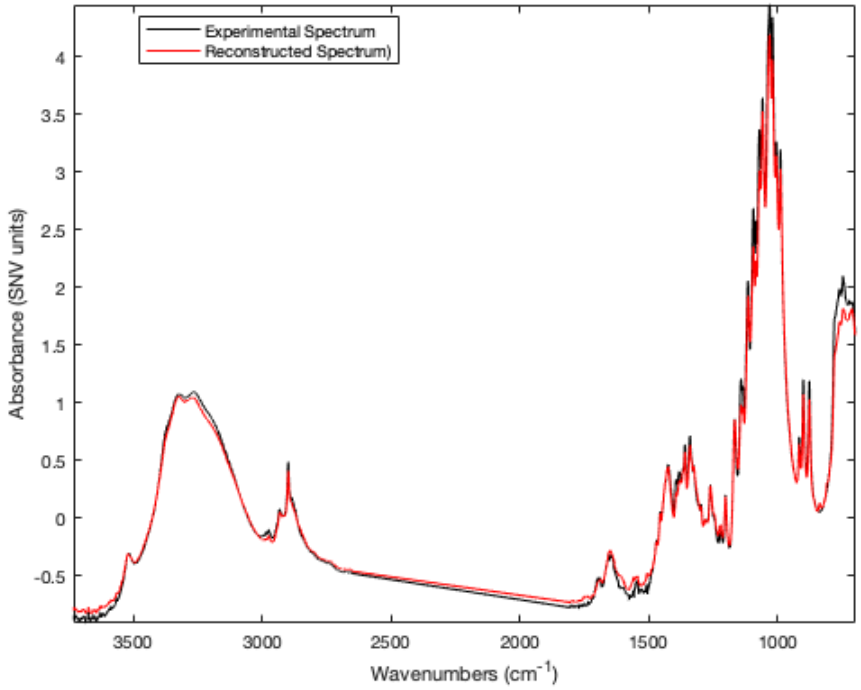

Figure 33: Overlap of the experimental spectrum of formulation 11 from set 2 and the reconstructed one (without the defined weights)

When observing both the figures with the overlap of the experimental graph with the one reconstructed by the method, and the absolute and relative errors, it is easily concluded that when the defined weights are not applied, the results obtained are much further away from the real concentrations. This difference is even more noticeable in the concentrations of the active substances – Paracetamol and Caffeine.

6.5.2. MCR – Weighted Linear Combination Analysis - with restrictions

Table 30: Results of MCR – Weighted Linear Combination Analysis: Formulation 1 from set 1 – with restrictions

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	39.7	4.5	1.7
Caffeine	0	0.402	-	0.402
Lactose	57	57	0	0
Paracetamol	5	2.91	41.8	2.09

Table 31: Results of MCR – Weighted Linear Combination Analysis: Formulation 4 from set 2 – with restrictions

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	34.8	8.4	3.2
Caffeine	1.5	2.73	82	1.23
Lactose	57	60.2	5.6	3.2
Paracetamol	3.5	2.29	34.6	1.21

Table 32: Results of MCR – Weighted Linear Combination Analysis: Formulation 8 from set 1 – with restrictions

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	36.7	3.42	1.3
Caffeine	3.5	2.95	15,7	0.55
Lactose	57	57.6	1,1	0.6
Paracetamol	1.5	2.67	78	1.17

Table 33: Results of MCR – Weighted Linear Combination Analysis: Formulation 11 from set 2 – with restrictions

Component	Experimental (%)	Predicted (%)	Relative Error (%)	Absolute Error (%)
MCC	38	39.9	5	1.9
Caffeine	5	1.71	14	3.29
Lactose	57	56.6	0.7	0.4
Paracetamol	0	1.77	49.4	1.77

When restrictions on excipient concentrations – Lactose and MCC are applied, both relative and absolute errors decrease. However, the difference is not very pronounced.

6.6 Comparison between the two methods – Summary Tables

Table 34: Comparison between the MCR - ALS method (Part 2 – addition of the pure components spectra) with the MCR – WLC (with the defined weights and without restrictions on concentrations) for formulation 1 from set 1

Component	Experimental (%)	MCR – ALS: Relative Error (%)	MCR – ALS: Absolute Error (%)	MCR – WLC: Relative Error (%)	MCR – WLC: Absolute Error (%)
MCC	38	15.5	5.9	7.9	3
Caffeine	0	-	0	-	1.50
Lactose	57	1.6	0.9	3.2	1.8
Paracetamol	5	-	5	54.2	2.71

Table 35: Comparison between the MCR - ALS method (Part 2 – addition of the pure components spectra) with the MCR – WLC (with the defined weights and without restrictions on concentrations) for formulation 4 from set 2

Component	Experimental (%)	MCR – ALS: Relative Error (%)	MCR – ALS: Absolute Error (%)	MCR – WLC: Relative Error (%)	MCR – WLC: Absolute Error (%)
MCC	38	5.5	2.1	5	1.9
Caffeine	1.5	100	1.5	14	0.21
Lactose	57	8.7	5	0.7	1.6
Paracetamol	3.5	42.9	1.5	49.4	1.73

Table 36: Comparison between the MCR - ALS method (Part 2 – addition of the pure components spectra) with the MCR – WLC (with the defined weights and without restrictions on concentrations) for formulation 8 from set 1

Component	Experimental (%)	MCR – ALS: Relative Error (%)	MCR – ALS: Absolute Error (%)	MCR – WLC: Relative Error (%)	MCR – WLC: Absolute Error (%)
MCC	38	7.4	2.8	7.6	2.9
Caffeine	3.5	14.3	0.5	27.4	0.96
Lactose	57	6.7	3.8	6.5	3.7
Paracetamol	1.5	46.7	0.7	7.3	0.11

Table 37: Comparison between the MCR - ALS method (Part 2 – addition of the pure components spectra) with the MCR – WLC (with the defined weights and without restrictions on concentrations) for formulation 11 from set 2

Component	Experimental (%)	MCR – ALS: Relative Error (%)	MCR – ALS: Absolute Error (%)	MCR – WLC: Relative Error (%)	MCR – WLC: Absolute Error (%)
MCC	38	16.8	6.4	16.8	6.4
Caffeine	5	28	1.4	26.6	1.33
Lactose	57	13.7	7.8	11.8	6.7
Paracetamol	0	-	0	-	1.04

7. Conclusion

The purpose of this thesis was to evaluate two different methods to estimate the concentration of solid formulations.

The first part of the MCR - Alternating Least Squares (commercial method) analysis revealed that the method is inefficient when the pure FTIR - spectra of the components are not applied. When the spectra were added, the results were much better, but the active substances continued with scores far away from the real concentrations. It is important to realize that this model only allow scores to be obtained, never exact concentrations.

The MCR – Weighted Linear Combination analysis showed, in geral, the importance of considering the different molar absorptivities of each component of the formulation. This was mainly noticeable in the detection of the active substances.

Through the calculation of the errors (relative and absolute), the second method had results closer to the real ones. Especially in the detection of concentrations of APIs, paracetamol was better detected, which was the main flaw of the commercial method.

It is also important to consider that there is always some important experimental error in both methods, which makes the use of FTIR a good complementary method for starting a generic development or even identifying a counterfeit medicine, undoubtedly. However, it may be necessary to refine this method with more complex techniques.

As a possible improvement of this work, the components of the formulations could have been dosed to certify that the planned formulations were, corresponded to those produced, and if the mixture was well executed.

As future perspectives, it would be interesting to carry out both analysis in semi-solid or liquid formulations, or even in solid formulations again, but increasing the number of components or the concentration of the active substances.

8. References

1. Thompson JM. Infrared Spectroscopy. 2018 Jan 19 [cited 2023 Jun 11]; Available from: https://www.taylorfrancis.com/books/mono/10.1201/9781351206037/infrared-spectroscopy-james-thompson
2. Stuart BH. Infrared Spectroscopy: Fundamentals and Applications. John Wiley & Sons, Ltd. 25 Jun 2004. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/0470011149?fbclid=IwAR3ql88SunyNIu5YzKqm8t5bBcHEnVaVzf2K8SfrgMu0Wa7ptrhjMWU9rcg
3. InfraRed spectoscropy lecture notes - 247 Chapter 15 Infrared Spectroscopy C.-P. Sherman Hsu, Ph. - Studocu [Internet]. [cited 2023 Jun 11]. Available from: https://www.studocu.com/row/document/mbarara-university-of-science-and-technology/pharmaceutical-sciences/infrared-spectoscropy-lecture-notes/21999820
4. Facilities: Optical Spectroscopy Instrumentation - [Internet]. [cited 2023 Jun 11]. Available from: https://www.marshall.edu/chemistry/facilities/optical-spectroscopy/
5. Bansal A. The Role of Reverse Engineering in the Development of Generic Formulations. Pharmaceutical Technology [Internet]. 2005 [cited 2023 Jun 11];29(8). Available from: https://www.pharmtech.com/view/role-reverse-engineering-development-generic-formulations
6. Custers D, Cauwenbergh T, Bothy JL, Courselle P, De Beer JO, Apers S, et al. ATR-FTIR spectroscopy and chemometrics: An interesting tool to discriminate and characterize counterfeit medicines. J Pharm Biomed Anal. 2015 Aug 10; 112:181–9.
7. Siesler HW. Vibrational Spectroscopy. Reference Module in Materials Science and Materials Engineering [Internet]. 2016 [cited 2023 Jun 11]; Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128035818013187
8. Vibrational Spectroscopy Overview & Types | What is Vibrational Spectroscopy? - Video & Lesson Transcript | Study.com [Internet]. [cited 2023 Jun 11]. Available from: https://study.com/learn/lesson/vibrational-spectroscopy-overview-types.html
9. Infrared Radiation - Infrared Spectroscopy - Advancing Materials [Internet]. [cited 2023 Jun 11]. Available from: https://www.thermofisher.com/blog/materials/all-about-infrared-radiation-and-spectroscopy/
10. Selina Spectrum by Byjus | PDF | Electromagnetic Radiation | Electromagnetic Spectrum [Internet]. [cited 2023 Jun 11]. Available from: https://pt.scribd.com/document/446342155/Selina-Spectrum-by-Byjus

11. ATR: Advantages for FT-IR Spectroscopy [Internet]. [cited 2023 Jun 11]. Available from: https://www.azom.com/article.aspx?ArticleID=5958
12. Definition, Classification, Advantages and Disadvantages of Powders: Pharmaguideline [Internet]. [cited 2023 Jun 11]. Available from: https://www.pharmaguideline.com/2021/10/definition-classification-advantages-disadvantages-of-powders.html?utm_content=cmp-true
13. Wulandari L, Idroes R, Noviandy TR, Indrayanto G. Application of chemometrics using direct spectroscopic methods as a QC tool in pharmaceutical industry and their validation. Profiles Drug Subst Excip Relat Methodol. 2022 Jan 1; 47:327–79.
14. Lambert-Beer's law - WikiLectures [Internet]. [cited 2023 Jun 11]. Available from: https://www.wikilectures.eu/w/Lambert-Beer%27s_law
15. Multivariate Curve Resolution | Getting started with mdatools for R [Internet]. [cited 2023 Jun 11]. Available from: https://mdatools.com/docs/mcr.html
16. Ruckebusch C, Blanchet L. Multivariate curve resolution: a review of advanced and tailored applications and challenges. Anal Chim Acta [Internet]. 2013 Feb 26 [cited 2023 Jun 11];765:28–36. Available from: https://pubmed.ncbi.nlm.nih.gov/23410623/
17. Ennis ZN, Dideriksen D, Vægter HB, Handberg G, Pottegård A. Acetaminophen for Chronic Pain: A Systematic Review on Efficacy. Basic Clin Pharmacol Toxicol [Internet]. 2016 Mar 1 [cited 2023 Jun 11];118(3):184–9. Available from: https://pubmed.ncbi.nlm.nih.gov/26572078/
18. Michael Z, Spyropoulos F, Ghanta S, Christou H. Bronchopulmonary Dysplasia: An Update of Current Pharmacologic Therapies and New Approaches. Clin Med Insights Pediatr [Internet]. 2018 Jan [cited 2023 Jun 11];12:117955651881732. Available from: https://pubmed.ncbi.nlm.nih.gov/30574005/
19. Microcrystalline Cellulose: What is it and where is it used? [Internet]. [cited 2023 Jun 11]. Available from: https://www.drugs.com/inactive/microcrystalline-cellulose-48.html
20. Hebbink GA, Dickhoff BHJ. Application of lactose in the pharmaceutical industry. Lactose. 2019 Jan 1;175–229.