Assessment of Low Dose Tablets by Near-Infrared and Raman Spectroscopy

Erik Jansson

Master's Thesis

Lund Reports on Atomic Physics, LRAP-342
Lund, May 2005

Abstract

Near-infrared (NIR) spectroscopy has been used in the analytical world since the early 1960s, and it is today a valuable technique for qualitative and quantitative analysis in the pharmaceutical industry. A major drawback of this technique when compared to traditional analytical techniques is its relatively low sensitivity, due to broad spectral overlap in spectra. Raman spectroscopy has in recent years, due to instrumental developments as compact and stable lasers in addition to more efficient sampling techniques, become a most promising tool for chemical analysis.

In this thesis, a comparison between NIR and Raman spectroscopy for quantitative analysis has been carried out. Measurements were made on pharmaceutical tablets designed for this project, with a weight percent concentration of active pharmaceutical ingredient (API) ranging from 0.1% up to 10%. To allow comparison of the spectroscopic techniques, both analyses used the same set of tablets. The data was evaluated using multivariate techniques and UV absorbance spectroscopy was used as a reference method.

The multivariate calibrations showed that both techniques are able to accurately predict the amount of active substance in tablets with concentrations as low as 1%. The models based on Raman spectra indicated that accurate predictions may be obtained below concentrations of 0.5%. Models based on Raman were also more robust than those based on NIR spectra.

Preface

This Master Thesis has been carried out at the division of Analytical Development, Astrazeneca, Mölndal, Sweden, during the period October 2004 to April 2005. The thesis has been made in collaboration with the division of Atomic Physics at Lund Institute of Technology (LTH), Lund University, Sweden.

I would like to thank Jonas Johansson and Jonas Eriksson for the opportunity to write this thesis and for their support and supervision throughout the course of this project. I would also like to thank Anders Sparén for taking time to discuss the multivariate data analysis and for proofreading the report. Last but not least, I would like to thank Magnus Swenson for his help during the tablets manufacturing process.

Thank you!

Erik Jansson

Table of Contents

1. Introduction

Vibrational spectroscopy has been used in the industry to identify raw materials and to make quantitative determinations for decades. Infrared (IR), Near-Infrared (NIR) and Raman spectroscopy are the most common methods for vibrational analysis. Instrumental development and refined data analysis, merely by chemometric procedures, have improved the techniques and increased their area of use [1]. Spectroscopic techniques are both fast and easy to use, but the most important advantage is the possibility of non-destructive analysis. Wet chemical analyzing techniques, such as HPLC (High-Performance Liquid Chromatography) and UV (Ultra-Violet) absorbance spectrophotometry, demand that the solid form is dissolved into solution. Therefore, important solid-state information is lost and the measurements give average chemical content only. These traditional analytical methods are also very time-consuming compared to spectroscopic techniques. Another instrumental feature, making spectroscopic techniques very attractive for process control and monitoring, is the possibility of coupling the instruments with suitable fibre optics. Infrared and Raman instrument can be applied "at-line" in the laboratory control, "on-line" with the use of measurement through optically clear windows and "in-line" with the use of specific probes for the measurement inside a pipe or a tank [1].

Up to this point, NIR spectroscopy has been the most utilized spectroscopic technique for quantitative analysis in the pharmaceutical industry [2, 3]. The technique is both simple and fast, but it has some limitations. In NIR spectroscopy broad spectral features overlap, which may hide relevant chemical information, and in the analysis of solid samples light scattering will distort the spectra even further. These drawbacks make quantitative analysis of samples containing low concentrations very difficult.

During the last few years there has been a growing interest for applications of Raman spectroscopy in pharmaceutical analysis [4, 5], but only a limited number of quantitative Raman studies have been reported. Szostak et al. [6] have reported on a quantitative determination by FT (Fourier Transform) Raman spectroscopy. The study was made on powders of ground tablets containing acetyl salcylic acid at a high level of concentration. Their study showed a high potential for FT-Raman spectroscopy when combined with chemometric methods and prediction errors of less than 2% were achieved. In the study by Dyrby et al. [7], transmittance NIR spectroscopy and FT Raman spectroscopy were compared for chemometric

quantitation of the active substance in low dose pharmaceutical tablets. The prediction ability of tablets containing 5% active substance, calculated as the relative prediction error, was 5.4% for the NIR transmittance calibration and 10.1% for the FT-Raman calibration. The FT-Raman calibration gave larger prediction errors, mainly due to a smaller sampled volume [7].

Analysis of compact powders such as pharmaceutical tablets require some considerations other than the more fundamental aspects of spectroscopy of solid materials. Because of spatial variations and inhomogeneities in the tablet matrix different measurements may give different results. This subsampling problem is more serious in Raman spectroscopy compared to NIR spectroscpy, since the irradiated sampled area is very small. In most FT-Raman systems the irradiated area is represented by a laser spot diameter < 0.5 mm. To achieve a representative quantitative analysis of a sample the sampled area has to be increased when using FT-Raman instruments. By translating the sample to multiple positions during the acqusition, for example via a rotating sample holder, subsampling errors can be reduced. Another way to lower these subsampling errors is to use dispersive systems fitted with fiber optic probes manufactured for large-spot sampling.

1.1 Aim

The aim of this project was to make a comparison between the analytical techniques NIR spectroscopy and Raman spectroscopy regarding quantitative analysis of pharmaceutical tablets with a low concentration of active substance. The underlying hypothesis was based on differences in spectral features in the two techniques. A major drawback of NIR spectroscopy when compared to traditional analytical techniques is its relatively low selectivity, due to broad spectral overlap, making analysis of low dose pharmaceutical tablets difficult. Since the Raman spectrum shows much sharper features compared to the NIR spectrum, it may be better suited for low dose quantitative analysis. The outcome of these experiments may help to set up where the respective method is the most capable. Since the results are highly dependent on the properties of the manufactured tablets, results should be discussed in view of the chemical homogeneity of the formulation.

1.2 Chapter summary

This summary aims to give the reader a brief outline of the different chapters in this Master Thesis. Chapter 2 contains the theory and background concerning vibrational spectroscopy, the spectroscopic techniques and chemometrics. Chapter 3 describes the methods used and the tablet manufacturing process. Results are displayed and discussed in Chapter 4. Chapters 5 and 6 contain conclusions made from the study and suggestions of how to improve future work. Chapter 7 contains references. Specifications of the spectroscopic systems used can be found in Chapter 8.

2. Theory

This Chapter offers an introduction to vibrational spectroscopy, the historic background to the spectroscopic techniques and the physical theory behind absorption and molecule scatter.

2.1 Vibrational Spectroscopy

Vibrational spectroscopy is one component of spectroscopic analysis, where NIR Spectroscopy and Raman Spectroscopy are two different techniques. Vibrational spectroscopy implies discrete vibrational transitions in the electronic ground state of a molecule. The vibrational energy levels are characteristic for every type of molecule and may therefore be used as fingerprints in analysis.

Spectroscopic techniques have many analytical advantages beside the ability of non-destructive analysis. They require no or limited sample preparation, sample size and form is not a restriction and they may be used to analyse samples in all three physical states - solids, liquids and gases [8].

2.1.1 Electromagnetic Radiation

By classical theory, electromagnetic radiation can be described by two orthogonal oscillating fields, one electric and one magnetic [1]. The two fields are sinusoidal oscillating, orthogonal to one another and also to the direction of propagation. Their movement can be described as

$$Y = A\sin\omega t \tag{2.1}$$

where Y is the amplitude with a maximum A, ω is the angular velocity and t is time. Equation (2.1) can be rewritten as

$$Y = A\sin 2\pi\nu t \tag{2.2}$$

where the frequency ν [s^{-1} or Hz] expresses the number of cycles per second [1]. An additional property of electromagnetic radiation is the wavelength λ [m], which is defined as

$$\lambda = \frac{c}{\nu} \tag{2.3}$$

where c is the velocity of electromagnetic radiation in vacuum. To explain phenomena like absorption and emission, electromagnetic radiation is described as a stream of discrete particles, light quantum or photons. In 1905 Albert Einstein created the quantum theory of light, based on Max Planck's theories, which states that the photon energy, E, is directly proportional to the frequency and inversely proportional to the wavelength. The relations read

$$E = h\nu = \frac{hc}{\lambda} \tag{2.4}$$

where h is Planck's constant [1].

In Raman spectroscopy, it is customary to describe the energy of the electromagnetic radiation by the frequency term *wavenumber*, $\bar{\nu}$ [cm^{-1}]. The wavenumber is the number of waves in a 1 cm long wave train [1].

2.1.2 Molecule Structure

A molecule is formed by the binding of two or more atoms in such a way that the total energy is lower than the sum of the constituents energy. The bonds in a molecule can be of ionic or covalent nature. Also week van der Waals bonds appear in molecules [9].

The molecular energy consists of translation energy and potential energy, where the potential energy is conserved as electronic, vibrational and rotational energy. The electronic energy depends on the structure of the electrons in the molecule, while the vibrational and rotational energy depends on the vibration of the atoms according to their equilibrium position and the molecule axis-perpendicular rotations respectively.

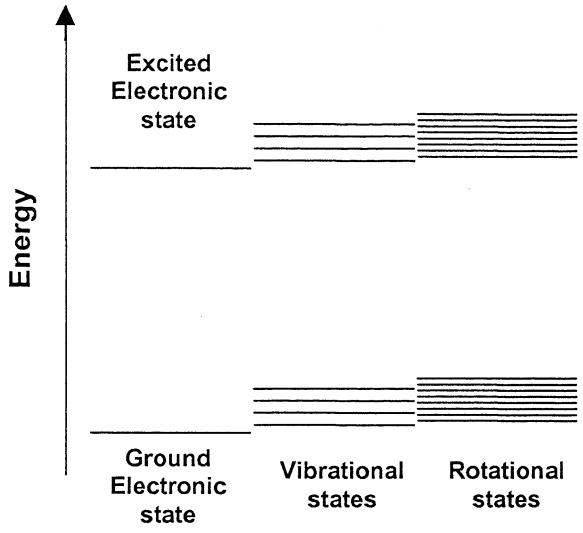


Figure 2.1.

Schematic view of the potential energy in a molecule showing the electronic, vibrational and rotational levels.

2.1.3 Molecular Vibrations

A molecule can be described by point masses representing the atomic nuclei, connected by massless springs representing the internuclear bonds. The springs tend to restore bond lengths and angles to specific values, the equilibrium values. To define the position in space of each individual point mass, three coordinates are needed, i.e. each point mass has three independent degrees of freedom. A molecule consisting of N atomic nuclei possesses 3N degrees of freedom [10].

If the centre of mass position instead defines the molecule position, three coordinates are needed. Another three degrees of freedom consists of molecular rotation (two if the molecule is linear, since rotation around the symmetry axis of a linear molecule does not affect the position of the atoms or the molecular energy). The number of vibrational degrees of freedom is therefore either 3N-6 or 3N-5, depending on the shape of the molecule [8, 10]. Each individual vibrational degree of freedom corresponds to a normal mode of vibration. A normal mode implies that all atoms vibrate at the same frequency and pass through equilibrium position simultaneously [10, 11].

As an example, water (H_2O), which is a non-linear molecule consisting of three atoms, has three independent normal modes (See Figure 2.2).

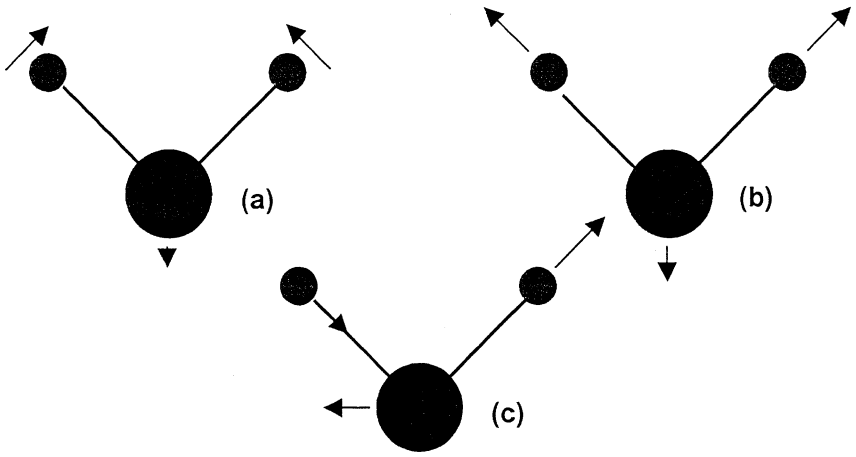


Figure 2.2.

The three independent normal modes of water: (a) bending vibration (b) symmetric stretch (c) asymmetric stretch.

2.1.4 Vibration Frequencies

According to the simple model of springs and point masses, any movement away from the equilibrium position creates a restoring force that is proportional to the displacement. The basic concept of molecular vibration can be described in terms of a simple diatomic molecule, discussing the vibration of a single isolated bond. The potential energy of a diatomic molecule depends on the inter-nuclear distance r. The molecule potential around the equilibrium position, r_0, can be approximated by a parabola which is a harmonic oscillator approximation of the molecule potential [11, 12]. Figure 2.3a shows a typical molecule potential and an approximated harmonic oscillator potential. The model of a simple harmonic oscillator is based on Hooke's law, which defines the vibration frequency $\bar{\nu}$ in wavenumbers as

$$\bar{\nu} = \frac{1}{2\pi c}\sqrt{\frac{k}{\mu}}, \qquad \mu = \frac{m_1 m_2}{(m_1 + m_2)} \tag{2.5}$$

where c is the speed of light, k is the bonding force constant (defined by the force expression $F = -k(r - r_0)$) and μ is the reduced mass of the atoms forming the molecular bond [8, 11].

The vibration energy of a harmonic oscillator, expressed in cm^{-1} (E/hc), is according to quantum mechanic theory

$$\frac{E_{vib}}{hc} = (\upsilon + \tfrac{1}{2})\bar{\nu}, \quad \upsilon = 0,1,2,... \tag{2.6}$$

where υ is the vibrational quantum number [8, 9]. A harmonic oscillator may change the quantum number only by ± 1, which imply evenly spaced vibration energy levels (See Figure 2.3b). The energy difference between two consecutive vibration levels is $((E_{(\upsilon+1)} - E_{\upsilon})/hc) = \bar{\nu}$, the so-called *fundamental frequency* [10].

From Figure 2.3a it is obvious that the model of molecular vibration potential, approximated by a harmonic oscillator, is not valid for vibrational states with higher energy. In reality, molecular vibrations are more or less anharmonic and deviations about the equilibrium positions are non-symmetric [8]. The anharmonic behaviour lowers the energy of each vibrational level below that of the harmonic behaviour, and the discrepancy gets larger as the displacement increases (See Figure 2.3c) [11]. A correction to the harmonic oscillator model, that improves consistency between theoretical and experimental data, is to include an additional higher order term into Hooke's law. Thus the energy for each level will be given by

$$\frac{E_{vib}}{hc} = \bar{\nu}_e\left((\upsilon + \tfrac{1}{2}) - \chi_e(\upsilon + \tfrac{1}{2})^2\right) \tag{2.7}$$

where χ_e is the non-harmonicity constant (a measure of the approximated parabola deviation from the actual potential function) and $\bar{\nu}_e$ the equilibrium vibrational frequency expressed in wavenumbers [8, 10]. Adding a quadratic term to Hooke's law changes the selection rule to $\Delta\upsilon = \pm 1, \pm 2, ...$ As a consequence, other, higher frequency transitions known as *overtones* or *harmonics* appear in addition to the fundamental transitions. Simultaneous changes in two or more vibrational modes are also possible. These transitions result in frequency bands called *combination* and *subtraction* bands. The combination and subtraction bands are weaker than the overtone bands and their transition frequencies are the sum or the difference between the individual fundamental frequencies [13].

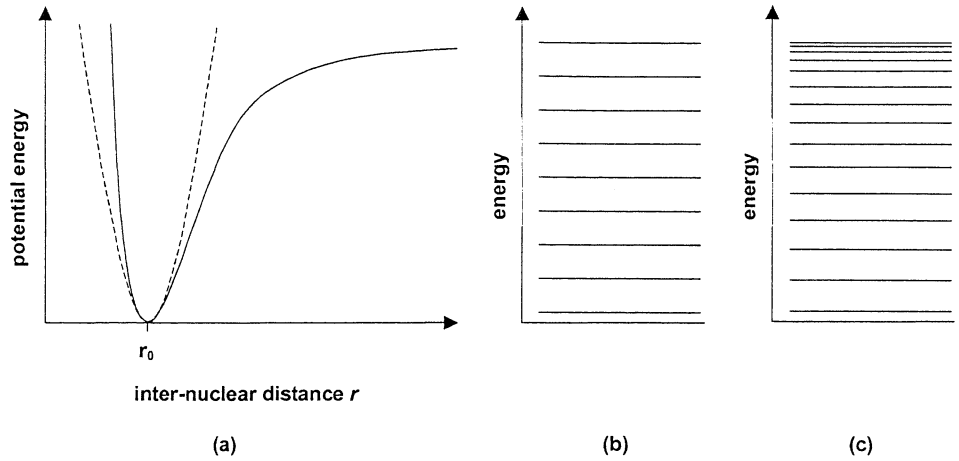


Figure 2.3.

(a) Potential energy versus inter-atomic distance for a diatomic molecule. The dashed line represents an approximated harmonic oscillator potential function. (b) Vibration energy levels of a harmonic oscillator. (c) Vibration energy levels of an anharmonic oscillator.

2.1.5 Vibration in Polyatomic Molecules

Even if all necessary information about a certain molecule is present, calculating the vibrational frequencies becomes very difficult if the molecule is complex. It has been found that different molecules with similar sub-molecular groups of atoms have similar vibrational patterns with frequencies in a characteristic region of the vibrational spectrum. These bands are the characteristic *group frequencies* [11]. As early as 1890 scientists noticed that all hydroxyl-containing species absorbed infrared radiation at roughly the same wavelength, near 3500 cm^{-1} and again at approximately 1650 cm^{-1}. These bands have been assigned to hydroxyl (—OH) stretching and bending vibrations. Another example is that all molecules containing the carbonyl functionality (—C=O) have a characteristic frequency in the range 1600-1800 cm^{-1} [11]. Beside these group frequencies each individual molecule has a number of bands that are specific of the individual chemical compound. These so-called *fingerprint bands* can be used to distinguish similar compounds from each other [11].

2.2 NIR Spectroscopy

The German astronomer Fredrick William Herschel uncovered the first evidence of solar energy output outside of the visible spectrum, in what is now known as the infrared, as early as 1800 [1]. The NIR spectral region of the electromagnetic

spectrum, which borders to the visible region, is generally from 750-2500 nm (13 300 to 4000 cm^{-1}) [13]. His discovery was largely ignored and even spectroscopists of the first half of the last century believed that it lacked analytical interest. The earliest applications of NIR spectroscopy was reported in the 1950s [14]. Over the last 25 years, NIR spectroscopy has been a widely used analytical tool, particularly by the food and agricultural industries, but also in the pharmaceutical industry [2]. The interest in NIR spectroscopy, as well as in Raman spectroscopy, lies in its advantages of recording spectra from solid and liquid samples using simple sample presentation, and predicting physical and chemical parameters from a single spectrum. Instrumental breakthroughs and the incorporation into equipment-bundled software of mathematical procedures for processing NIR spectra have increased its usefulness [15]. Absorption in the NIR region arises from overtones or combinations of fundamental vibrational motions, in addition to combinations of overtones. There are a great number of overtones and combination bands in a normal NIR spectrum. Since many of these bands overlap, NIR spectra lack sharp spectral features. The majority of overtone bands in the near IR region arise from X-H stretching modes, where X is O, N, C or S [2, 13].

A molecule absorbs infrared radiation by increasing the vibrational energy. The absorption process is analogous to a mechanical oscillator that has a few resonance frequencies where it may absorb energy from an external oscillating disturbing force [10]. A similar process occurs for an absorbing molecule. The disturbing external oscillating force consists of the photon electrical field. Molecular vibrational absorption occurs at certain specific frequencies that match the natural vibration frequencies of the molecule. At non-resonant frequencies no absorption occurs [10].

2.2.1 NIR Analysis techniques

Analysis of solid samples is often done by diffuse reflectance NIR. Diffuse reflectance measures the intensity ratio of scattered light from the sample compared to the light reflected from a reference surface. Since the absorptivity in the NIR region is quite low, the sampling depth is typically 1-3 mm in a solid sample. Penetration depth depends on particle size, their packing and the radiation wavelength [16]. NIR analysis can also be performed in transmission mode, where instruments measure the decrease in radiation intensity when radiation passes through the sample, as a function of the wavelength. The transmission mode has an advantage in a larger sampled volume, but the wavelength range is often narrower since the detector

signal becomes noisy when too little light penetrates the sample [17]. According to Gottfries et.al [17], multivariate models based on transmission spectra of solid pharmaceuticals have better prediction ability compared to models based on diffuse reflectance spectra even if the avaliable wavelength range is narrower.

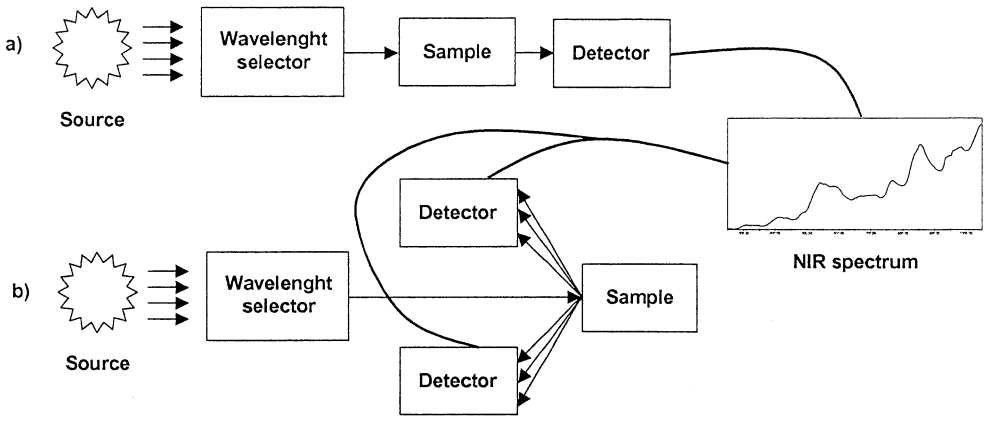


Figure 2.4.

Schematic configuration of the two different types of NIR instrument described above; (a) transmission mode (b) diffuse reflectance mode.

2.3 Raman Spectroscopy

During the 1920s, there was a great interest in light scattering processes. The Compton effect, which describes changes in frequency of electron-scattered X-ray light quantum, was documented in 1923. In the next few years inelastic molecular scattering of electromagnetic radiation was theoretically predicted. The Indian scientist Chandrasekhar Venkat Raman and the Russian physicists Grigory Landsberg and Leonid Mandelstam made the first experimental observations, of what is today known as the Raman effect, virtually simultaneously in 1928. Raman alone was credited for the discovery and he was awarded with the Nobel Prize in physics in year 1930 [9], a decision that has been debated thereafter.

The Raman effect is a scattering process that changes the frequency for a tiny part of an incoming monochromatic light beam. Before the first commercial and reliable lasers, that were available in the 1960s, the technique was very time-consuming and an experts-only method [18]. The transition to a standard analytical technique has been propelled by several technological advances in optics and electronics [18]. Raman scattering can occur due to a change in vibrational, rotational

or electronic energy of a molecule. The term Raman effect will hereafter refer to the vibrational Raman effect only.

2.3.1 Raman Instruments

There are two fundamentally different designs of the Raman spectrometer: dispersive Raman and FT-Raman [1]. The dispersive instruments use gratings to disperse the scattered light in order to detect the whole spectrum simultaneously. Dispersive spectra may be obtained with photodiode arrays or CCDs [19]. Charge transfer devices have drawbacks with performance deterioration in the near-infrared spectral region [11]. Therefore, excitation sources with shorter wavelengths that often induce fluorescence, has to be used in dispersive systems. The induced fluorescence may hide the Raman signal. In FT-Raman spectrometers, detection is often achieved by solid-state detectors that perform well in the near-infrared. Solid-state detectors require cooling to liquid nitrogen temperatures to perform well and to reduce the noise level. A commonly used radiation source is the Nd-YAG laser radiating at a wavelength of 1064 nm, which makes spectra virtually free from fluorescence. A drawback with the FT-Raman technique is the fact that water absorbs in the 1000 nm region, limiting the ability to study aqueous samples [19].

2.3.2 The Raman Effect

An incident photon interacting with a molecule can be absorbed, transmitted or scattered. Figure 2.5 shows a schematic view over the absorption and scattering processes. Molecule scatter is either elastic or inelastic. In quantum mechanics scatter processes are described as excitations to a virtual state with lower energy than an excited electronic state [9]. When photons are scattered elastically no change in energy occur and the scattering process can be described as an elastic collision between photons and molecules. Elastic scattering is also referred to as *Rayleigh* scattering and it is by far the strongest component of scattered radiation, generally ~10^3 times stronger than the Raman lines [9]. However, a small fraction (approximately 1 in 10^7 photons) is scattered inelastically at frequencies different from the frequency of the incident photon [1, 20].

The process leading to inelastic scatter, the Raman effect, is basically an inelastic collision between an incident photon and a molecule where, as a result of the collision, an amount of energy, ΔE, has either been taken from or given to the

molecule. The collision results in a change in vibrational energy of the molecule. The physic law of energy conservation states that the energy difference between the scattered photon, $h\nu_s$, and the incident photon, $h\nu_i$, must be equal to ΔE.

$$\Delta E = h\nu_i - h\nu_s \tag{2.9}$$

The type of scatter where photon energy has decreased is known as *Stokes Raman* scatter. Photons that have gained energy are called *anti-Stokes Raman* scattered photons. The scattered photons give an exact picture of energy differences between vibrational states in the molecule. The intensity ratio between the different types of Raman scattered radiation depends on the population ratio between the vibrational states, which is determined from the Boltzmann distribution function

$$\frac{n_{\upsilon=1}}{n_{\upsilon=0}} = \frac{g_1}{g_2}\exp\left(\frac{\Delta E_{vib}}{kT}\right) \tag{2.10}$$

where g_1 and g_2 are the degeneration of the individual vibrational states, k is the Boltzmann constant and T is the temperature expressed in degree Kelvin [10]. Since Stokes Raman scatter bands are several orders of magnitude more intense than the Anti-Stokes Raman scatter bands at room temperature these are the bands detected in most spectrographs [10].

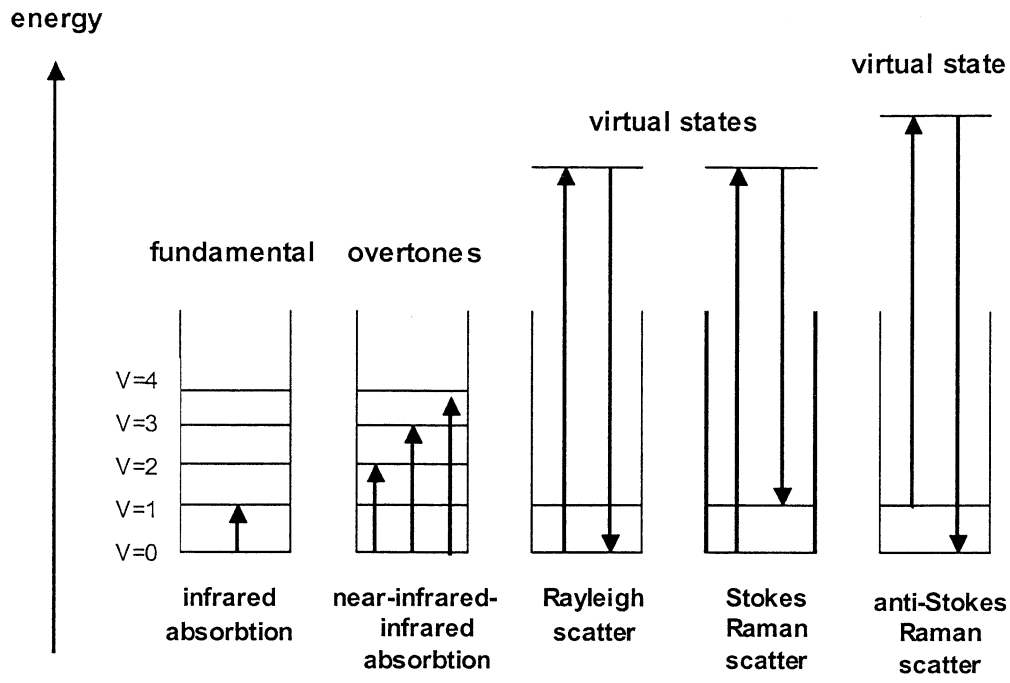


Figure 2.5.

Summary of the processes observed in vibrational spectroscopy.

2.3.3 Raman Scatter Frequencies

The oscillating photon electric field induces a dipole moment, P, when interacting with molecules, which is directly proportional to the amplitude of the field via

$$P = \alpha \cdot E \tag{2.11}$$

where the proportional constant, α, is the polarizability of the molecule. The polarizability can be described as the deformability of the electron cloud [10]. As the electric field of the electromagnetic radiation oscillates, Equation 2.11 can be rewritten as

$$P = \alpha E_o \cos(2\pi \nu_0 t) \tag{2.12}$$

According to the classical theory of an oscillating dipole moment it will emit electromagnetic radiation with the frequency ν_0 in all directions. However, as vibrating molecules are alternately compressed and extended, the shape of the electron cloud changes and so does the polarizability [10].

A molecular vibration can be expressed as

$$Q = Q_0 \cos(2\pi \nu_{vib} t) \tag{2.13}$$

where the constant Q_0 is the maximum displacement of a normal coordinate common for the molecule, and ν_{vib} is the frequency of the normal coordinate [10]. The polarizability can be expanded in Taylor series (if the displacements are small) as

$$\alpha = \alpha_0 + \left(\frac{\delta\alpha}{\delta Q}\right)_0 Q_0 + \text{higher order terms} \tag{2.14}$$

where α_0 is the equilibrium polarizability. Higher order terms are neglected in the harmonic approximation [10]. By combining equations 2.12-2.14 the expression for the dipole moment of a vibrating molecule reads

$$P = \alpha_0 E_0 \cos(2\pi \nu_0 t) + \left(\frac{\delta\alpha}{\delta Q}\right)_0 Q_0 \cos(2\pi \nu_{vib} t)\cos(2\pi \nu_0 t) \tag{2.15}$$

By using the trigonometric identity

$$\cos A \cos B = \tfrac{1}{2}(\cos(A+B) + \cos(A-B)) \tag{2.16}$$

Equation 2.15 can be rewritten as

$$P = \alpha_0 E_0 \cos(2\pi \nu_0 t) + \frac{1}{2}\left(\frac{\delta\alpha}{\delta Q}\right)_0 Q_0 [\cos(2\pi(\nu_0 - \nu_{vib})t) + \cos(2\pi(\nu_0 + \nu_{vib})t)] \tag{2.17}$$

The expression above expresses an induced dipole that varies with three different frequencies, ν_0, $\nu_0 - \nu_{vib}$, and $\nu_0 + \nu_{vib}$ (*Rayleigh*, *Stokes* and *anti-Stokes*). The value of ν_{vib} is known as the *Raman shift*. A more rigorous description of the classical and quantum mechanical theory of Rayleigh and Raman scattering can be found in [21].

2.3.4 Raman Sampling

The most important aspect of Raman sampling is the fact that the obtained Raman spectrum is only representative of the material irradiated by the light source. It is therefore of great importance that the sampled portion is truly representative of the material. In most FT-Raman systems the irradiated area is represented by a laser spot diameter < 0.5 mm. Different laser irradiation methods with rotating sample holders have been characterized to enlarge the irradiated area and improve accuracy in quantitative analysis [22]. The laser can be defocused to a larger area by using fibre optic probes. The laser spot size in the Raman system used in this project was approximately 4 mm in diameter at sample position. If the sample consists of a tablet with a diameter of 9 mm the radiated area covers approximately 20% of the sample surface.

Sample heating is another important issue when using lasers as excitation sources. Intense radiation on a small area may heat the sample and change its properties. Sample heating is a considerable problem in FT-Raman systems where the laser output power is high [23]. In systems using large spot sampling and dye lasers sample heating effects can be neglected.

2.3.5 Fluorescence

Fluorescence is, as mentioned before, a major issue in dispersive Raman spectroscopy. This is because most molecules have electronic band gaps corresponding to the visible region of the electromagnetic spectrum. Fluorescence is emitted when electronically excited molecules are de-excitated into lower lying electronic states. The frequency of the fluorescence depends not only on the energy gap between the electronic states but also on those vibrational and rotational states involved in the process. The frequency of the fluorescence may therefore possess many different values.

The reason why fluorescence is a major problem in Raman spectroscopy is that it is much stronger than the Raman scattered fraction, and may therefore hide the Raman signal.

2.4 Vibrational Spectroscopic Activity and Intensity

Even if NIR and Raman spectra both involve vibrational energy levels they are not duplicates of each other. This is because the intensity of the spectral bands depends on how effective the photon energy is transferred to the molecule, and the mechanism for photon energy transfer differs in the two techniques [10]. To get the full picture of the vibrational modes and their energies a complete investigation with both NIR spectroscopy and Raman spectroscopy is needed.

While the infrared absorption frequency depends on the molecular vibrational frequencies, the infrared absorption activity and intensity depends on the *change in the dipole moment* that occurs when a molecule vibrates [10]. Since the wavelength of infrared radiation is much larger than the inter-nuclear distance in a molecule, the electric field of a photon can be considered as uniform over the whole molecule. An applied electric field will force the molecular charges in perpendicular directions, and as the electric field oscillates it induces a dipole oscillating with the same frequency. At certain frequencies a forced dipole moment tends to stimulate a molecular vibration. This is when the molecular vibration itself causes a change in dipole moment. The more the dipole moment changes during a molecular vibration, the more easily the photon electric field can activate that vibration. A selection rule for infrared activity can be summarized as: "In order to absorb infrared radiation, a molecular vibration must cause a change in the dipole moment of the molecule" [10]. The selection rule implies that infrared spectroscopy only observes vibrations that have nonzero values of $(\delta P / \delta Q)_0$ [11]. Symmetric molecules are therefore inactive to infrared absorption. The intensity of an infrared absorption band is proportional to the square of the change in dipole moment $(\delta P / \delta Q)^2$ of the molecular vibration, giving rise to the absorption band [1, 10].

Raman spectroscopy only observes vibrations that have nonzero values of $(\delta \alpha / \delta Q)_0$ [11]. The selection rule reads: "In order for a molecular vibration to be Raman active, the vibration must be accompanied by a change in the polarizability of the molecule" [10]. From classical theory it is known that the intensity of scattered radiation is proportional to the square of the maximum value of the excitation field (E_0^2). The intensity of Raman scattered radiation is therefore proportional to $(\delta \alpha / \delta Q)^2$ [10]. The intensity of the Raman signal is also linearly proportional to the intensity of the incident laser light and inversely proportional to the forth power of the

laser wavelength [18]. Raman intensity from an active pharmaceutical ingredient in a pharmaceutical tablet can be expressed as

$$I_{Raman} \sim \frac{I_{laser} \cdot a \cdot C}{\lambda_{laser}^4} \tag{2.18}$$

where a is a constant that depends on sample presentation and the instrument used and C is the active substance concentration.

2.5 UV Absorbance Spectrophotometry

UV absorbance spectrophotometry is a chemical analysis method that measures absorbtion in liquid samples. This method, which is an old, but still useful instrumental method, is typically used to provide quantitative information about a substance. It was used to determine reference values regarding active substance in the tablets. Absorption in the ultra-violet region is associated with electron excitation in molecules.

The amount of light, I, transmitted through a solution of absorbing chemical in a transparent solution can be related to the concentration via Beer's Law:

$$-\log\frac{I}{I_0} A = \varepsilon_\lambda bc \tag{2.19}$$

where I_0 is the incident light intensity, A is the absorbance, b is the cell path in cm, c is the concentration of a chemical in moles/litre, and ε_λ is the molar absorptivity. ε_λ is the quantity that represents the spectrum of the solution and is therefore wavelength dependent: $\varepsilon = \varepsilon(\lambda)$.

2.6 Chemometrics

Chemometrics was founded when the need to analyse data containing many variables emerged in analytical and physical organic chemistry during the 1960s [24]. Chemometrics are mathematical and statistical methods, used to analyse and excavate maximal information from spectroscopic and other multivariate data. Multivariate models often carry out the chemometric analysis. Chemometrics is still

evolving and the definition may have to be modified in order to include new developments in the area [24].

2.6.1 Multivariate Data Analysis (MVDA)

Spectroscopic data consists of many variables collected from different samples and perhaps at different times. The raw data is difficult to analyse by just looking at the digitized spectral data or study plotted graphs. With MVDA any table of data can be converted into a few meaningful plots that display the information in a more accessible way [15].

Spectroscopic data is often displayed as a spectrum where the intensity of light is plotted as a function of the wavelength. The detectors used today present digital spectra, which facilitates the data analysis. The spectroscopic data can be described as $x = \eta + \varepsilon$, where x is the raw data, η is a deterministic component modelling the data and ε is stochastic component including the noise. The noise is descended from measurement errors, sampling errors and other experimental errors [15].

The multivariate data analysis is not founded upon a theory, but uses regularities in the data to build models. Its simplicity makes the analysis very useful in handling large amounts of noisy, non-linear and incomplete data [15]. Projection methods were used for multivariate analysis in this project. All chemometric analysis were performed by a chemometric data program called SIMCA (Soft Independent Modelling of Class Analogy version P 10.0.4.0, Umetrics, Umeå, Sweden)

2.6.2 Principal Component Analysis (PCA)

PCA is a multivariate projection method used to extract and display systematic variations in a data matrix **X**. A spectroscopic data matrix normally consists of *N* rows and *K* columns, where the rows represent observations in the form of spectra. The *K* columns are variables containing information from certain wavelengths or wavenumbers in the spectra. The idea of PCA is to describe the multivariate data and its numerous variables by a few new variables, *principal components,* also known as *latent variables*, that transform the multidimensional data table into an understandable lower dimensional world [15, 23].

The method of PCA is best understood by a geometrical presentation where data matrix **X** is transferred to a *K*-dimensional variable space, where the *K* variables

represent *K*, individual orthogonal, coordinate axes. Points in this variable space represent the N observations. The first principal component, PC_1, is the straight line in *K*-space that best approximates the data in least square sense and it is therefore oriented in such a way that it reflects the largest variation in the data. The orthogonal projection of an observation onto a PC gives the score value. The score values define the properties of each observation (Figure 2.6). By graphically illustrating the score values, dataset properties such as clustering of observations, trends and outliers can be observed. The influence and correlation of variables can be visualized in a loading plot [24]. The loadings express the orientation of the principal components in *K*-space.

If one PC does not model the whole variation in the data set, one has to enlarge the model by adding one or more PCs. A new line in *K*-space, orthogonal to the earlier PCs, represents every new PC. PC_2 represents the second largest variation in the data. Every new PC is introduced the same way. Variation in data that is not explained and covered by the PCs is included in a residual matrix **E**.

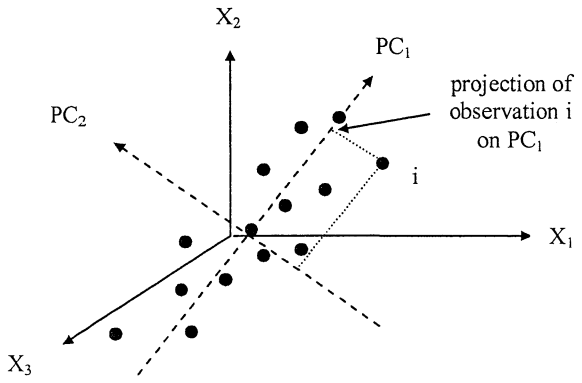


Figure 2.6.

*The first and second principal component fitted to a data matrix **X** containing 14 observations and 3 variables.*

PCA-modelling can be described mathematically as

$$\left.\begin{aligned}\mathbf{X} &= \mathbf{t}_1\mathbf{p}_1 + \mathbf{E}_1 \\ \mathbf{E}_1 &= \mathbf{t}_2\mathbf{p}_2 + \mathbf{E}_2\end{aligned}\right\} \quad \mathbf{X} = \mathbf{t}_1\mathbf{p}_1 + \mathbf{t}_2\mathbf{p}_2 + \ldots + \mathbf{t}_A\mathbf{p}_A + \mathbf{E}_A \qquad (2.20)$$

where $\mathbf{t}_x$ and $\mathbf{p}_x$ consists of the score and loading values belonging to PC_x.

2.6.3 Multivariate Calibration

Multivariate calibration is a very useful application of chemometrics on spectroscopic data, and it is widely used in analytical chemistry [24]. Multivariate calibration builds relationships between the chemical properties of samples and the corresponding spectra. The regression relationship is

$$\mathbf{y} = f(\mathbf{X}) \tag{2.21}$$

where vector $\mathbf{y}$ is a vector containing measured concentrations and $\mathbf{X}$ is a matrix where each row contains a digitized spectrum. When the function f is found, concentrations of new samples can be calculated via their spectra. Since solving equation 2.21 becomes impossible when $\mathbf{X}$ consist of many variables it is simplified as

$$\mathbf{y} = \mathbf{Xb} + \mathbf{f} \tag{2.22}$$

where $\mathbf{b}$ is a vector of regression coefficients and $\mathbf{f}$ is a vector containing the residuals. Equation 2.22 is a soft equation, where the regression coefficients can be found without much prior knowledge of the system.

2.6.4 Partial Least Square Regression to Latent Structures (PLS)

PLS is a method used to relate two matrices and to find the best solution to the matrix equation 2.22. PLS models use principal components to relate a data matrix $\mathbf{X}$ to a reference vector $\mathbf{y}$. The PLS principal components are oriented in the way that they reflect the maximal covariance in the data.

In analytical chemistry PLS is most often used to create multivariate calibration with which the active substance concentration of new samples can be predicted. To estimate the prediction accuracy and ability, models are created using a selected part of the matrix $\mathbf{X}$ and the corresponding part of the vector $\mathbf{y}$, a *calibration set*. The created model is then used to predict another part of the matrix $\mathbf{X}$, the *validation set* or the *prediction set*, measuring the prediction ability.

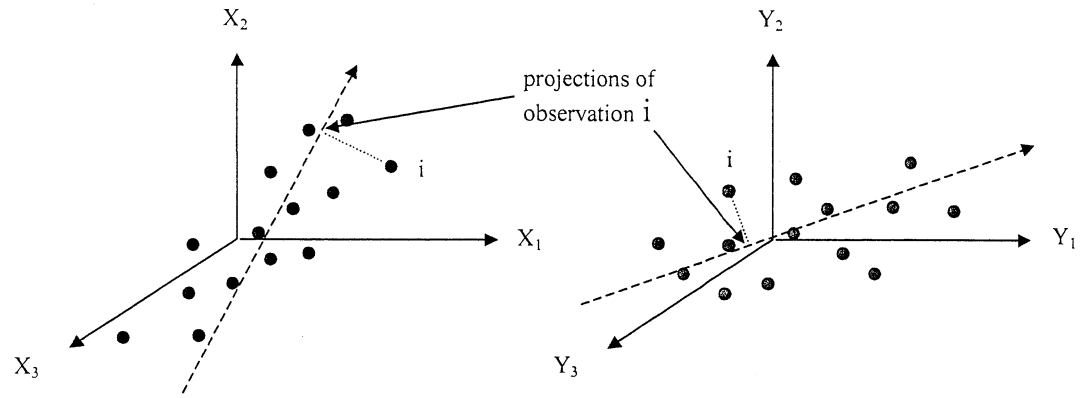


Figure 2.7.

A geometrical presentation of the first principal components in a PLS model.

PLS modeling can be described mathematically as

$$\begin{aligned}\mathbf{X} &= \mathbf{t}_1\mathbf{p}_1 + \mathbf{t}_2\mathbf{p}_2 + ... + \mathbf{t}_A\mathbf{p}_A + \mathbf{E}_A \\ \mathbf{y} &= \mathbf{u}_1\mathbf{c}_1 + \mathbf{u}_2\mathbf{c}_2 + ... + \mathbf{u}_A\mathbf{c}_A + \mathbf{F}_A\end{aligned} \tag{2.23}$$

Prediction accuracy can be estimated by the Root Mean Square Error of Prediction (RMSEP) value. The RMSEP is defined as

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \tag{2.24}$$

where $\hat{y}_i$ are the predicted variables, y_i are the values from the reference vector $\mathbf{y}$ and n is the number of predicted samples. The RMSEP value can be divided by the mean value of the y_i:s to obtain the percentage of RMSEP [15].

2.6.5 Data Pre-Treatment

In order to enhance the predictive power of multivariate calibration models, spectral data is often pre-processed prior to data analysis. The reason is that NIR and Raman spectra contain systematic variations that are unrelated to the responses $\mathbf{y}$. Systematic variations may interfere and disturb the multivariate modeling. For solid

samples, this variation can be due to light scattering and differences in spectroscopic path length. Other reasons for systematic and unwanted variations are different detector sensitivity in various regions of the spectrum, baseline drift and regions of wavelength that lack important information [25].

- **Scaling and centering of mean-value**

Scaling data can be performed in many different ways dependent on the prior knowledge of the variables and their effect on the models. The two basic scaling techniques are scaling to unit-variance and mean-centering. Unit-variance scaling implies that every individual variable is divided by its standard deviation and result in a spectrum where all variables get the variance one and thereby have the same influence on the analytical model. Mean-centering implies that each variable is subtracted by their average values, which gather all the variables around the zero baseline [15].

- **Orthogonal signal correction (OSC)**

OSC is a signal correction method that removes bilinear components, which are orthogonal to **y,** from **X**. OSC components are calculated via an algorithm called the NIPALS algorithm [25]. The NIPALS algorithm implies that the first principal component score vector, **t**, of **X** is orthogonalized with respect to **y** to give **t***, the actual correction vector. Subsequently, the PLS weights, **w**, are computed such that **Xw** becomes as close as possible to **t***. The correction vector **t*** is then processed in the NIPALS algorithm through **X** to give an updated score vector **t**, which is again orthogonalized with respect to **y**. This is iterated until convergence. After convergence, a loading vector **p** is computed and then **t*p** is subtracted from **X** [25].

- **Derivatives**

Differences between spectra are often better visualized by a first or second derivative, which only includes curve-shape information. By using a derivative method called Savitsky-Golay, signal to noise ratio is preserved. This method implies that a polynomial is fitted to a selected part of the spectrum after which the derivative is calculated. The derivative process is repeated over the whole curve [26, 27].

- **Fluorescence subtraction by polynomial curve fitting**

Polynomial curve fitting has been found to be a simple but effective method to lift off the fluorescence background in dispersive Raman spectra. The method has distinct advantages over other fluorescence reduction techniques in its ability to retain the spectral contours and the band intensities of the input spectra. To efficiently reproduce the fluorescence background, a polynomial fit can not merely be applied in a least-square manner since subtraction of such a fit results in a spectrum that varies about the zero baseline. A modified polynomial fitting method described in [28] was used. The basis for this method is a least-square (4:th or 5:th degree) polynomial curve-fitting function. However, to eliminate the Raman bands from the fit, the function is modified so that all data points in the generated curve that have an intensity value higher than their respective pixel value in the input spectrum are automatically reassigned to the original intensity [28]. This process (curve fitting and subsequent reassignment) is repeated, gradually eliminating the weakest Raman peaks from the underlying fluorescence. Fluorescence subtraction by polynomial curve fitting was used to subtract fluorescence from the Raman spectra used in the homogeneity analysis of the tablets.

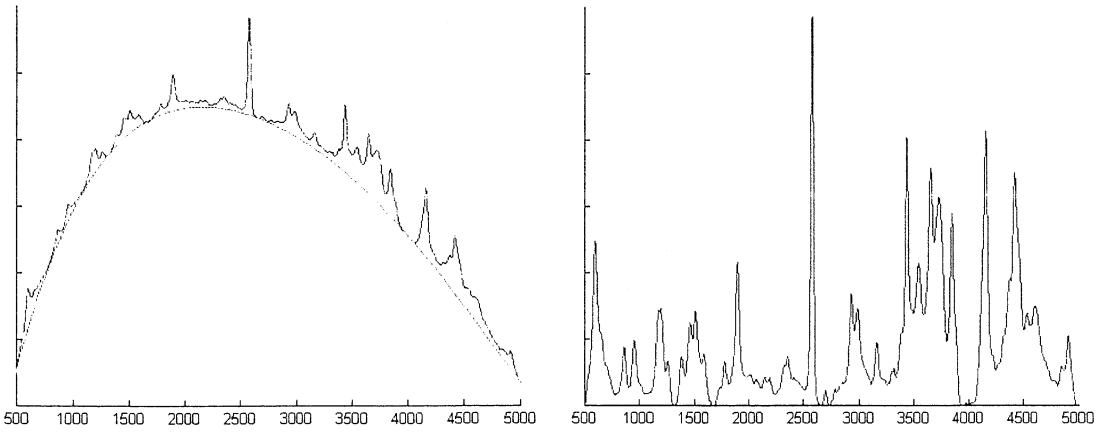


Figure 2.8.

Fluorescence subtraction applied on a raw Raman spectrum acquired from a 10% salicylic acid tablet. The fit is made with a 5:th degree polynomial and 1000 iterations. (a) Raman spectrum and fitted curve (b) fluorescence subtracted Raman spectrum.

3. Material, Instruments and Methods

This Chapter offers a description of the instruments and the measurement procedures. It also includes a description of the material used and a brief introduction to the data analysis.

3.1 Diffuse Reflectance NIR System

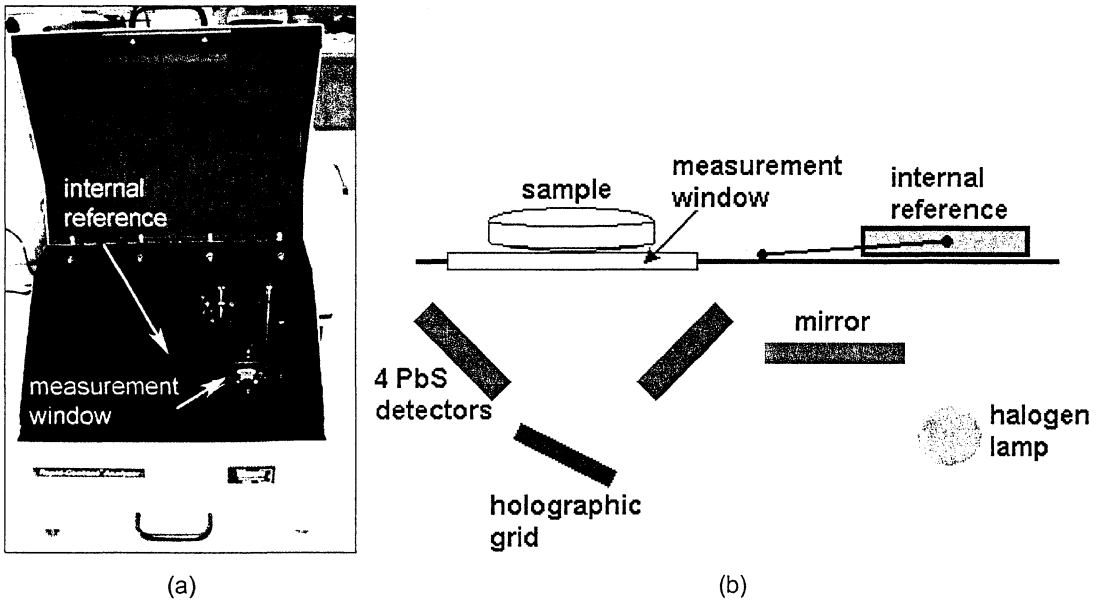


Figure. 3.1.

(a) Foss NIR Systems Model 6500 spectrometer fitted with a Rapid Content Analyzer® (RCA). (b) Schematic drawing of the optics in a Rapid Content Analyzer®.

Diffuse reflectance NIR spectra of single tablets were obtained using a Foss NIR Systems Model 6500 spectrometer fitted with a Rapid Content Analyzer® (RCA). Measurements were controlled and acquired by Vision® 2.51 Spectral analysis software for windows, Foss NIR Systems. Specifications of the NIR system can be found in Appendix 1.

3.1.1 NIR Measurement Procedure

Diffuse reflectance NIR measurements where performed according to the intrumental guidelines. Spectra were collected over the full spectral range 400 to 2500 nm. An iris was used for centering the tablets on the measurement window. To control

instrument performance a polystyrene standard was used and the built-in reference was used for instrument calibration. The acquisition time for a full range spectrum was approximately 40 seconds. The recorded spectra were exported to SIMCA for evaluation.

3.2 Dispersive Raman System

(a) (b)

Figure 3.2.

(a) RamanRxn1™ Analyzer and (b) measurement setup.

Dispersive Raman spectra of single tablets were obtained using a RamanRxn1™ Analyzer, manufactured by Kaiser Optical Systems, Inc (KOSI). The RamanRxn1™ Analyzer is a dispersive Raman system that consists of a base unit connected to a fiber optic probe and a computer system. The base unit contains the laser, a spectrograph, a 2D-array detector (CCD) and associated power supplies. The system uses an Invictus™ laser as light source. The Invictus™ laser is an external-cavity-stabilized diode laser emitting light at 785 nm. Sampling is performed by a fiber optic probe designed for large-spot sampling. The probe focuses the excitation beam light onto the sample via an optic fiber bundle and samples the Raman scattered light via another optic fiber bundle. Excitation light is focused to minimize interference from areas outside the sampling area. A HoloPlex™ grating is used for a simultaneous full spectral collection of Raman data. The HoloPlex™ is a holographic transmissive grating that separates the incoming photons by wavelength and passes them onto the CCD-detector. Before the scattered light enters the detector it passes through a holographic SuperNotch Plus™ filter, which has a U-shaped transmission curve that transmits Stokes scattered radiation, but prevents the reflected and

elasticallly scattered light from reaching the detector. The RamanRxn1™ Analyzer uses a CCD for detection of the Raman scattered light. The CCD is a rectangular two-dimensional array of independent, integrating detectors (pixels) on a chip of silicon. Photo-generated charge from each individual pixel is collected and stored in a potential well during the measurements. When an exposure is finished, charge from the potential wells is serially read out and via an on-chip amplifier processed and transferred to the computer to produce the Raman spectrum [20]. Specifications of the Raman system can be found in Appendix 1.

3.2.1 Raman Measurement Procedure

The RamanRxn1™ Analyzer needed approximately half an hour to let the laser warm up to, and the CCD detector to cool down to, operating temperature. The probe was aligned in a vertical position above the stationary sample holder (see Figure 3.2b) and the vertical positioning of the probe was adjusted to maximize Raman signal intensity. The laser beam diameter was approximately 4 mm at the sample position. The laser output power was set to 400 mW and exposure time to 10 s. The spectral resolution was 5 cm^{-1} in average according to the manufacturer [29]. All spectra were corrected for dark signal. Measurements were acquired and controlled by the HoloGRAMS™ software package, and the recorded spectra were exported to SIMCA for evaluation.

3.3 Reference Analysis

The content of active substance in the tablets was determined by UV absorbance spectrophotometry. The spectrophotometer used, a Hewlett Packard model HP 8453, detected unabsorbed light with a linear array detector consisting of 1024 light sensitive diodes. The linear array detector allows instantaneous measurement of the complete spectrum ranging from 190-1000 nm with a maximum spectral resolution of 1 nm. Absorption was measured at 303 nm and background, measured at 500nm, was subtracted. A quarts cuvette, transmitting UV radiation, with a path length of 1 cm were used as sample holder.

The weighted tablets were put in volumetric flasks and dissolved in 0.1 molar hydrochloric acid. The solutions were mechanically shaken for approximately one hour until the tablets were completely disintegrated, after which the solutions were diluted to a concentration of 0.04 mg/ml and left to let the insoluble microcrystalline

cellulose sediment for three hours. At this level of concentration the detector response was linear. Another solution consisting of pure active substance, dissolved and diluted to the same concentration, was prepared and used as a standard. The standard solution absorbance at 303 nm was used to calculate the amount of active substance in the dissolved tablets. To simplify the measurements an automatic sampler, connected to a pump, was used to fill the quarts cuvette.

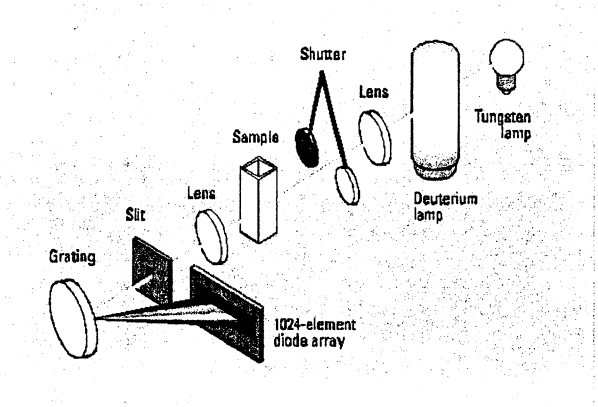


Figure. 3.3.

Schematic view of the Hewlett Packard HP 8453 UV Diode Array Spectrophotometer used for reference analysis.

3.4 Matrix NIR Imaging System

To get a picture of the particle size and the distribution of constituents in the tablets, a MatrixNIR™ NIR Spectral Imaging System (Spectral Dimensions Inc., Olney, USA) was used. A focal-plane array based imaging spectrometer provides the individual spectrum in 320x240 pixels. Every pixel contributes to an approximately 1600 μm^2 (40 µm x 40 µm) area of the sample surface. The vibrational spectra collected for each pixel on the array detector create a third dimension. The X- and Y-axes represent spatial information, and the Z-axis represents reflectance at the selected NIR wavelengths. By comparing the spectra from the sample with pure constituents spectra, it was possible to locate the constituents and get information about their spatial distribution within the sample. This instrument has been described in detail in [27].

3.5 Pharmaceutical Tablets

Pharmaceutical tablets contain one or more active substances, and a few excipients added to aid formulation and manufacturing. The properties of the final dosage form (i.e. its bioavailability and stability) are, for the most part, highly dependent on the excipients chosen, their concentration and interaction with both the active compound and each other [30].

The two set of tablets used in this project were produced by direct compression. Direct compression implies that the dry powder substances are mixed together and compressed into tablets. Tablet formulation by direct compression is an easy and fast method but it causes problems from a spectroscopic analytical point of view, since the homogeneity of direct compressed tablets often is quite poor. Most pharmaceutical tablets today are produced by wet granulation. Wet granulating is a procedure, where a solution is added to the blend of powder while mixing them. The formed granules are then dried and milled before they are compressed into tablets. Tablets prepared by the wet granulation method have better content uniformity and suitable tablet dissolution and stability properties, but the method is quite laborious.

3.5.1 Substances and their Applications in Pharmaceutical Formulation or Technology

Acetyl salicylic acid and salicylic acid were used as active substances in the tablets. Acetyl salicylic acid is a well-known substance among pharmacists, used in painkillers for its ability to reduce pain and inflammation and lower fever. Salicylic acid is often used in skin care pharmaceuticals. The excipients used were microcrystalline cellulose, lactose (anhydrous) and sodium stearyl fumarate. Microcrystalline cellulose is widely used in pharmaceuticals, primarily as a binder/diluent in tablets and capsule formulations, where it is used in both wet-granulation and direct-compression processes. Lactose is often used as a filler or binder in tablets and capsules. Sodium stearyl fumarate is a lubricant used in capsule and tablet formulations, normally at a concentration of 0.5–2.0% [31].

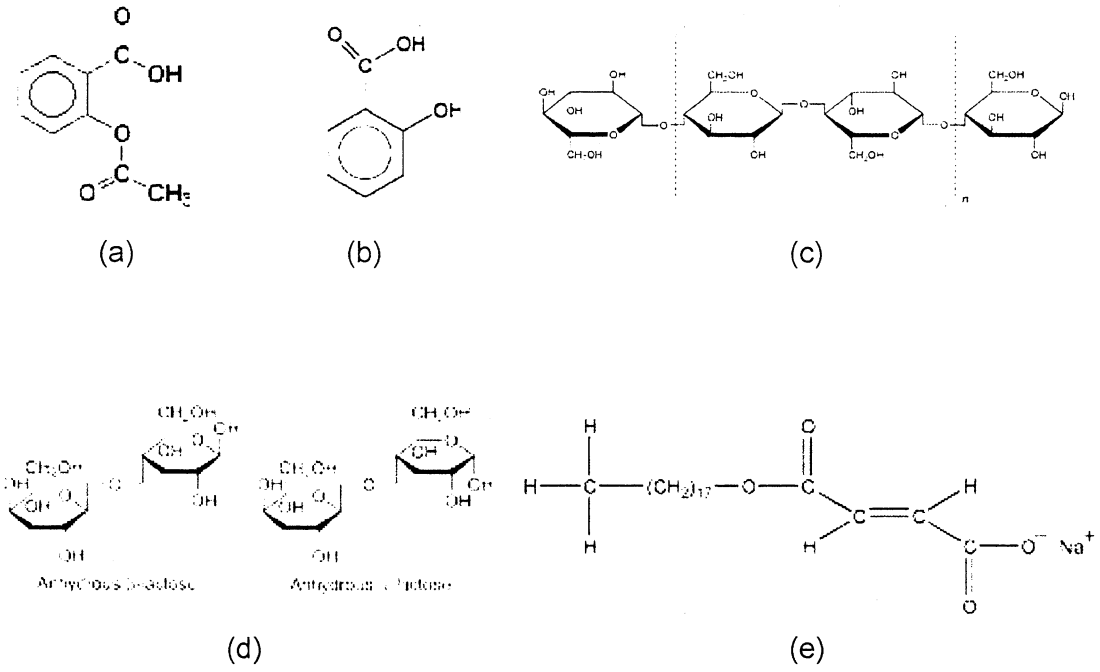


Figure 3.4.

Molecule structure of the substances used in the tablets; (a) acetyl salicylic acid (ASA), ($C_8H_9O_4$) (b) salicylic acid (SA), ($C_7H_6O_3$) (c) microcrystalline cellulose (MCC), (($C_6H_{10}O_5$)·n), where n ≈ 220 (d) anhydrous lactose (LA), ($C_{12}H_{22}O_{11}$) (The anhydrous lactose typically contains 70–80% anhydrous β-lactose and 20–30% anhydrous α-lactose) and (e) sodium monostearyl fumarate (SMF), ($C_{22}H_{39}NaO_4$)

3.5.2 Manufactoring of the Tablets

Tablets were manufactured according to an experimental design shown in Table 3.2, where the amounts are stated as weight percentage. Dry powder mixing was carried out in two steps by a blender apparatus (Turbula Blender T2C). In the first step the active pharmaceutical ingredient, microcrystalline cellulose and lactose were weighed and dry mixed for ten minutes. The excipients were delivered with a particle size ensuring good mixing, but the active substances were delivered with a coarser particle size. The acetyl salicylic acid substance was grinded and sieved to improve tablet homogeneity, and the salicylic acid, which contained both smaller and bigger particles, was sieved to a particle size < 180 μm before mixing. In step two, sodium stearate fumarate was added to the mixture, which was finally mixed for five more minutes. Manufacturing of the tablets was made with a single punch tablet machine (Diaf, model TM20, Copenhagen, Denmark).

Table 3.1. *Tablet specifications*

diameter	9 mm
geometric shape	arched, cylindrical
weight	~250 mg
height	~3 mm (at center)
press force	~10 kN
coil	uncoiled
quantity in each concentration group	~100

Table 3.2. *Experimental design of the tablets.*

concentration group	active pharmaceutical ingredient (%)	microcrystalline cellulose (%)	anhydrous lactose (%)	sodium monostearyl fumarate (%)
1	0.10	52.90	45.00	2.00
2	0.25	52.75	45.00	2.00
3	0.50	52.50	45.00	2.00
4	0.75	52.25	45.00	2.00
5	1.00	52.00	45.00	2.00
6	2.50	50.50	45.00	2.00
7	5.00	48.00	45.00	2.00
8	7.50	45.50	45.00	2.00
9	10.00	43.00	45.00	2.00

3.6 Method of Data Analysis

Multivariate PLS models were developed by SIMCA in order to predict the content of active substance, measured by the reference method (y-variable), from the dispersive Raman and diffuse reflectance NIR spectra (X-variables). Reference values were expressed as the weight percent concentration of active substance in each tablet. The X variables were mean-centered prior to all model building. Models were built using PLS, and several different pre-processing methods were tested in order to obtain the best prediction ability. Since the prediction ability of a PLS model is highly dependent on the number of relevant X-variables [15], many different variable intervals were tested. Four observations were removed from all models because of unrealistic reference values caused by measurement errors. 10 observations, marked with the letter c, from each group of concentration level were used in the calibration set. The remaining 10 observations in each group were included in the prediction set, and marked with the letter p. The predictive ability of the calibrations was calculated by the RMSEP value. The RMSEP value divided by the nominal active substance content, y_{nom}, was used to compare prediction errors from prediction sets with different nominal active substance concentrations. The number of PLS principal components used in the models was selected by studying

loading plots. Additional PLS principal components may bring irrelevant information into the model and cause loss in the prediction ability.

4. Results and Discussion

In this Chapter, results from the data analysis are described and discussed. Pictures taken by the NIR imaging system were included to illustrate the problem with inhomogeneous tablet matrices.

4.1 Dispersive Raman Spectra

Raman wavenumber shifts were collected in the range 0-1800 cm^{-1}, with an interval of 0.3 cm^{-1}. Raman spectra of pure powdered substances, acquired with an exposure time of 16 seconds, are shown in Figure 4.1-4.3. The acetyl salicylic acid Raman spectrum (Figure 4.1) has a spectral feature that is promising for quantitative analysis. The pre-dominant peak at 1606 cm^{-1} lies in a wavenumber range where the excipients show no or little Raman activity. The salicylic acid Raman spectrum shown in Figure 4.2 has a dominant peak at 772 cm^{-1}, which is in a region with intense fluorescence and where the excipients show some Raman activity. These factors make accurate calibrations, based on salicylic acid spectra, harder to obtain compared to acetyl salicylic acid spectra. Raman spectra of the excipients are shown in Figure 4.3. The notch filter, which is used to filter the Rayleigh line, affected all Raman spectra in the spectral region 0-150 cm^{-1}.

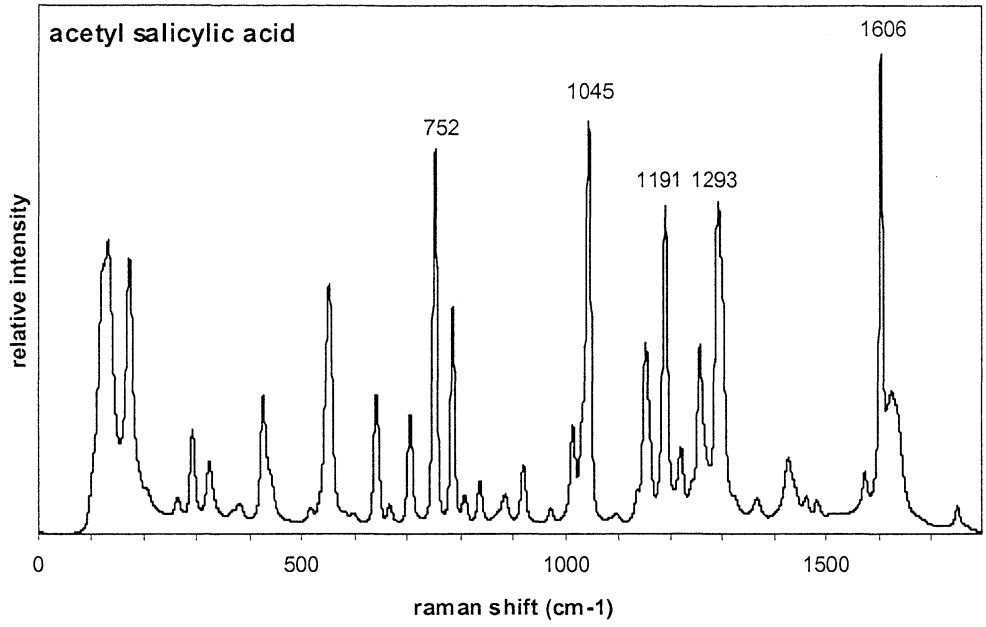


Figure 4.1.

The Raman spectrum of pure powdered acetyl salicylic acid substance. The five most intense Raman peaks are identified by their Raman shift expressed in cm^{-1}.

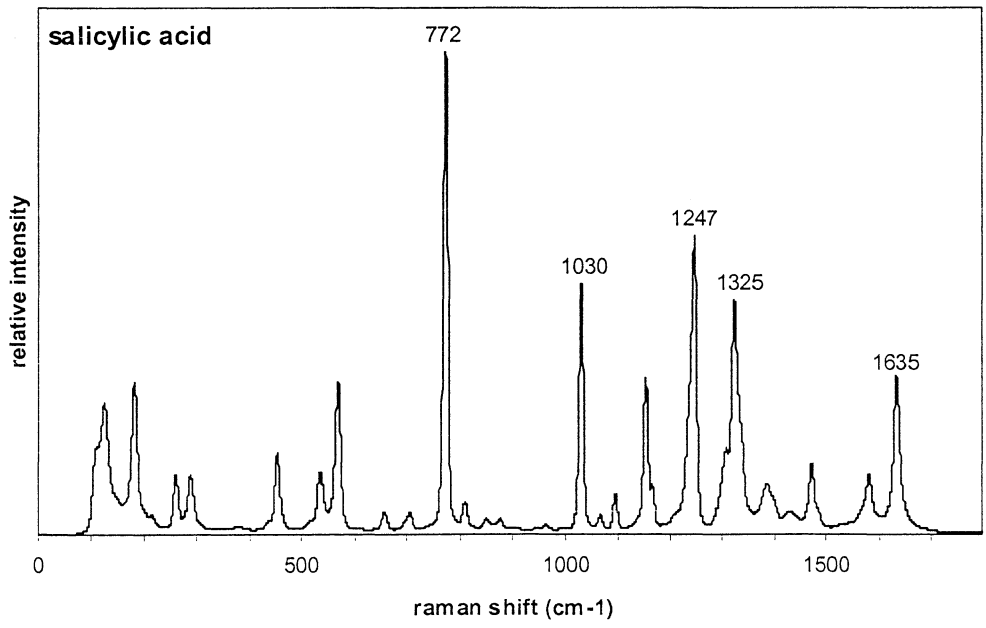


Figure 4.2.

The Raman spectrum of pure powdered salicylic acid substance. The five most intense Raman peaks are identified by their Raman shift expressed in cm^{-1}.

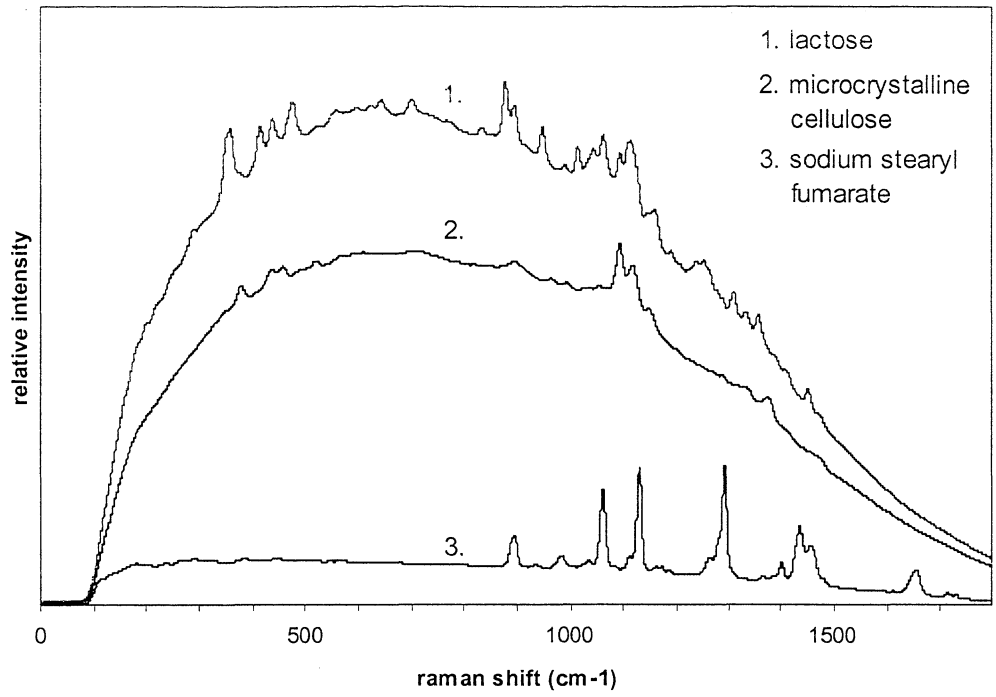


Figure 4.3.

The Raman spectra of the excipients included in the tablets. As can be seen fluorescence is mainly due to the lactose and the microcrystalline cellulose.

4.1.1 Exposure Time

Selecting exposure time is a crucial step in Raman measurements. While an increased exposure time gives higher signal to noise ratio, as demonstrated in Figure 4.4, it may also involve non-linear bleaching effects. In the bleaching process, which is yet not fully understood, molecules in a tablet matrix are destroyed due to the intense radiation [11]. The bleaching process causes the flourescence to reduce by time. Since the bleaching process is a non-linear effect, it affects the multivariate calibration in a negative way. The exposure time was adjusted to 10 seconds to achieve a good signal to noise ratio and to avoid bleaching effects.

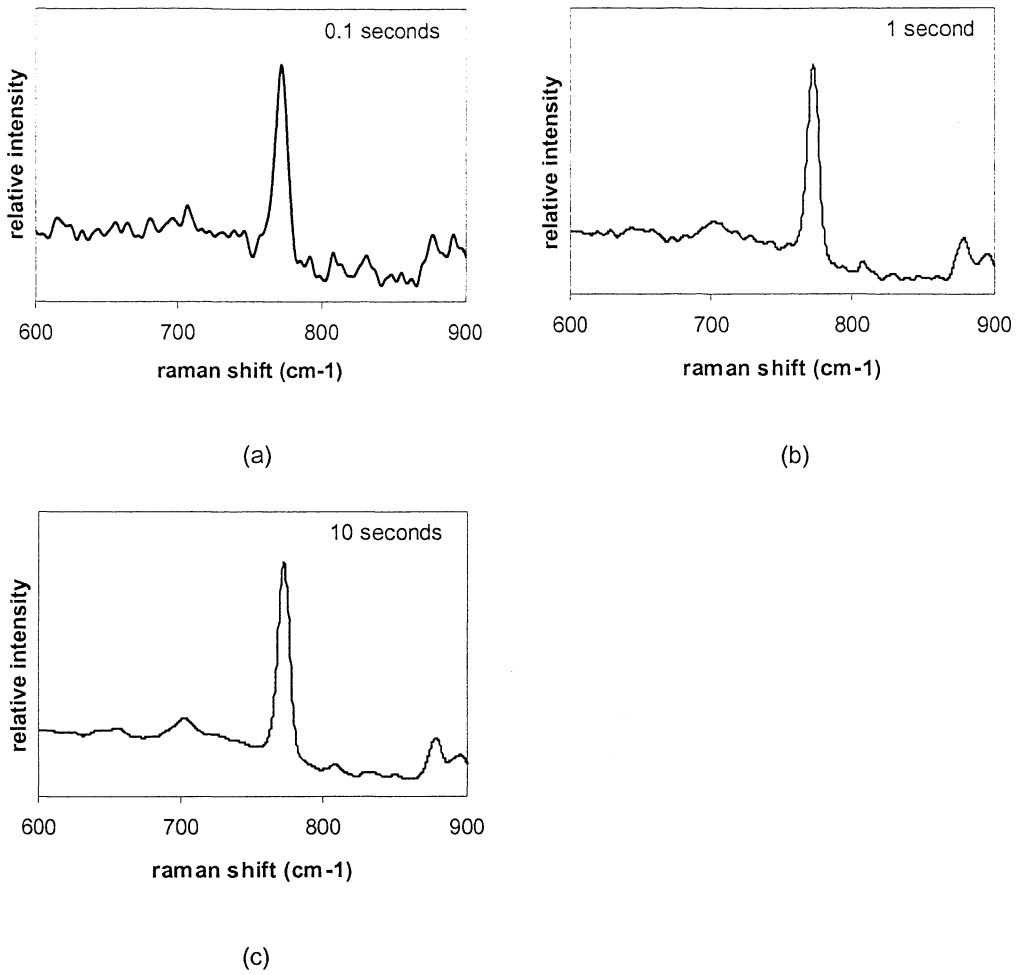


Figure 4.4.

The signal to noise ratio after 0.1, 1 and 10 seconds of exposure time in dispersive Raman spectra of a tablet containing salicylic acid at a concentration of 10%. The Raman shift interval is 600-900 cm^{-1} and the distinct Raman peak is the dominant peak at 772 cm^{-1}.

4.1.2 Intensity Calibration

Intensity calibration of the Raman instrument corrects for uneven instrument response across the CCD-camera [29]. Unfortunately, intensity calibration of the RamanRxn1^{TM} Analyzer could not be carried out in a reliable way by the manufacturer instructions. Intensity calibrated spectra were therefore only used for illustration, and uncalibrated spectra were used in the data analysis. The intensity calibrated spectra did also show a slightly higher noise level, which would deteriorate the data analysis. Differences between an intensity calibrated spectrum and an uncalibrated spectrum of a tablet containing salicylic acid at a concentration of 10% is shown in Figure 4.5.

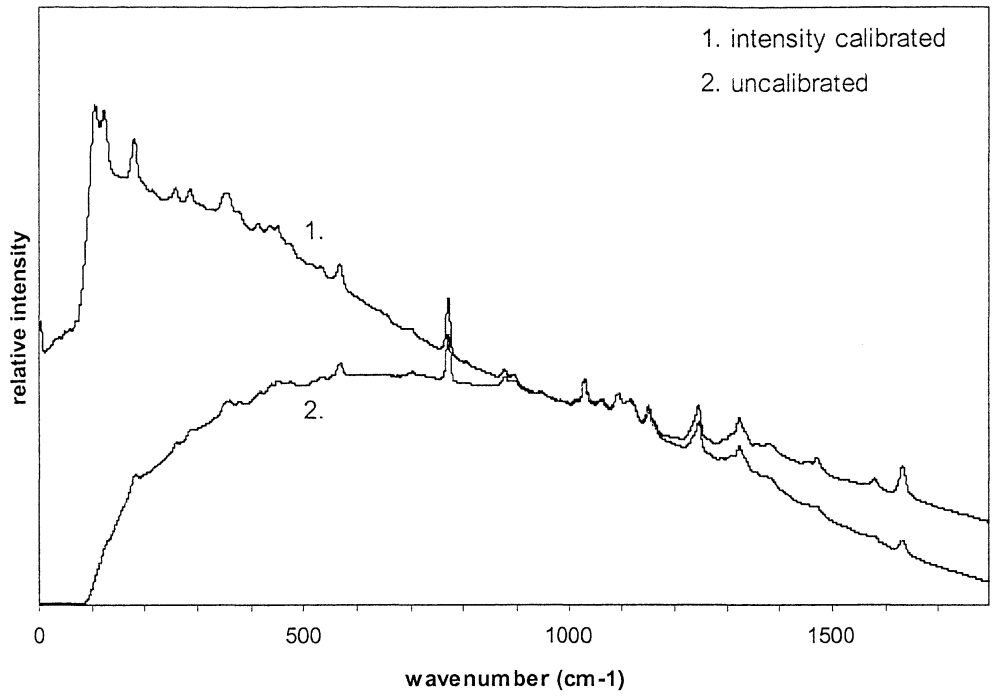


Figure 4.5.

Intensity calibrated and uncalibrated dispersive Raman spectra of a tablet containing salicylic acid at a concentration of 10%.

4.2 Diffuse Reflectance NIR Spectra

Figure 4.6 shows one NIR spectrum from a tablet containing salicylic acid at a concentration of 10% and one spectrum aquired from pure powdered salicylic acid substance. The full range NIR spectra cover the wavelength range 400-2500 nm, and were collected with an interval of 2 nm. In the spectrum from the pure salicylic acid, two distinct absorption peaks are observed at wavelengths 1132 and 1660 nm, originating from overtone absorption in the –OH vibration. The absorption bands in the spectral region above 2000 nm are due to combinations of fundamental vibrations. A NIR spectrum acquired from pure powdered acetyl salicylic acid is shown in Figure 4.7. Since the molecule structure of the active substances are quite similar, the absorption spectra of the two different active substances resemblance each other.

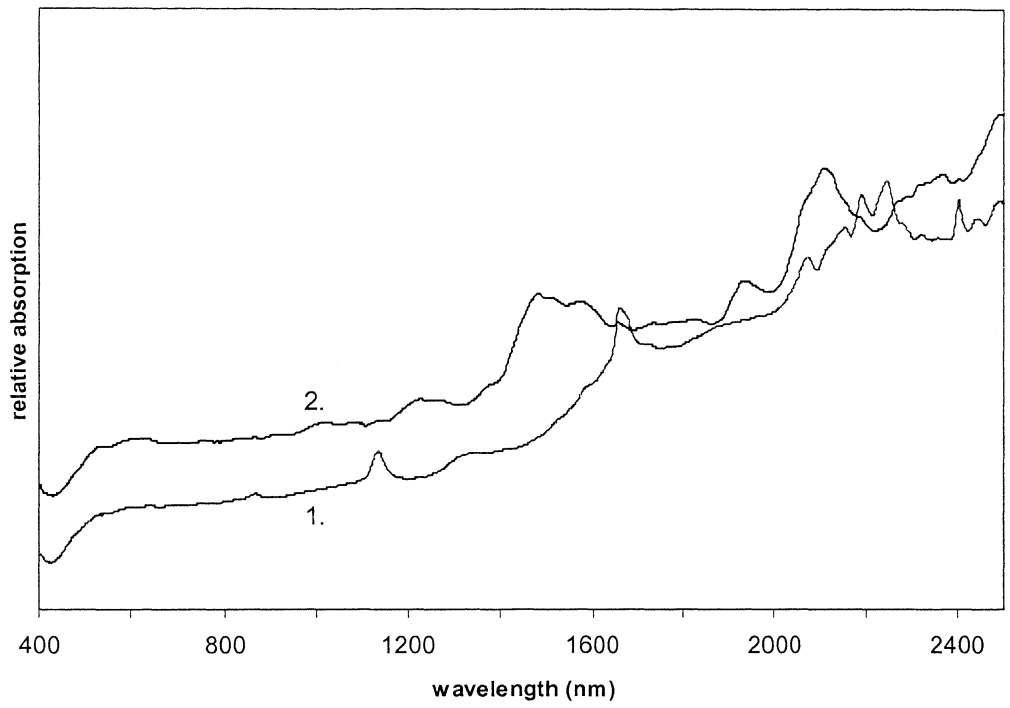


Figure 4.6.

Diffuse reflectance NIR spectra from pure powdered salicylic acid substance (1) and a tablets containing salicylic acid at a concentration of 10% (2).

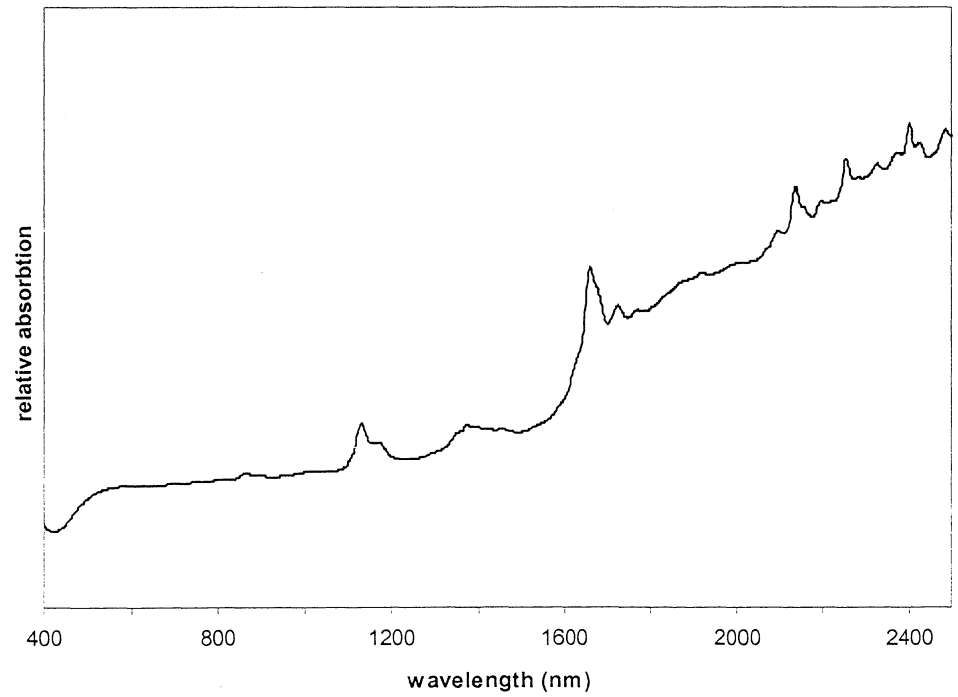


Figure 4.7.

NIR spectra from pure powdered acetyl salicylic acid.

4.3 Properties of the Manufactured Tablets

The tablets used for analysis were manufactured with two different active substances, acetyl salicylic acid and salicylic acid. The acetyl salicylic acid has several advantages compared to the salicylic acid. It is relatively harmless to handle, and its spectral properties makes it suitable for quantitative analysis by vibrational spectroscopy. However, it is transformed into salicylic acid when dissolved in both acidic and basic environment. Reference analysis of tablets containing acetyl salicylic acid is therefore a very complicated process. No multivariate calibrations were made on the tablets containing acetyl salicylic acid.

The final result of the calibrations is highly dependent on the tablet properties, especially the tablet homogeneity regarding the active pharmaceutical ingredient. Results from the reference analysis made on tablets containing salicylic acid can be seen in Table 4.1. When using NIR spectroscopy in the diffuse reflectance mode and dispersive Raman, only a limited volume of the tablets is sampled. It is therefore of great importance that the sampled volume is representative for the tablet as a whole. Images of the active substance distribution in bisected tablets were acquired by the Matrix NIR imaging system. The images where collected at 5 nm interval from 1100 to 1750 nm using a 1x objective. The spectral images were pre-processed with first derivatives and compared with images of pure substances. In Figure 4.8, first derivative images at 1675 nm are shown for two bisected salicylic acid tablets. The images illustrate the distribution of the active component (blue areas) with both smaller and larger regions within the tablets. The largest regions were estimated to 3 mm in diameter. The excipients appeared to be evenly distributed within the tablets.

Table 4.1. *Properties of the tablets containing salicylic acid measured by the reference method.*

concentration group	nominal concentration (%)	average concentration ± relative standard deviation (%)
1	0.100	0.105 ± 0.036
2	0.250	0.250 ± 0.032
3	0.500	0.504 ± 0.021
4	0.750	0.748 ± 0.013
5	1.000	0.994 ± 0.020
6	2.500	2.483 ± 0.011
7	5.000	5.053 ± 0.020
8	7.500	7.593 ± 0.035
9	10.000	9.853 ± 0.020

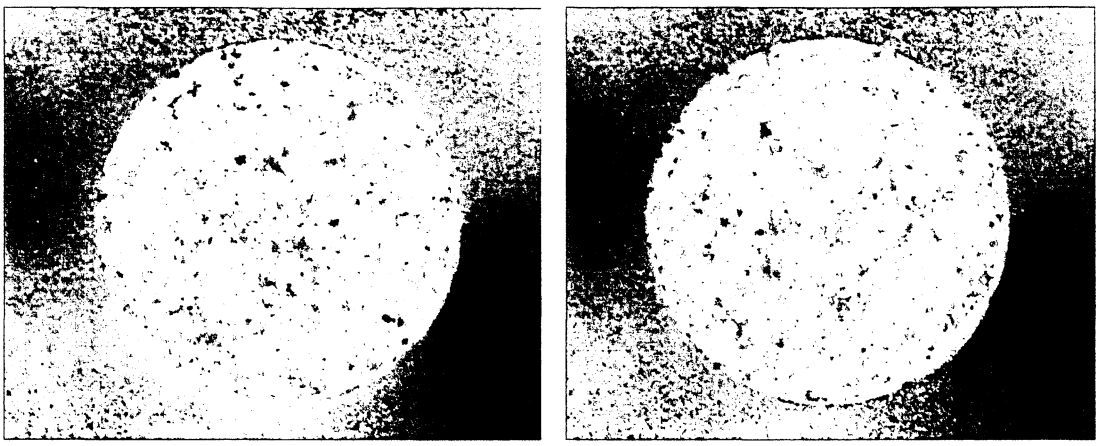


Figure 4.8.

Distribution of salicylic acid in two different tablets originating from concentration group 8 (7,5%). Regions with a high concentration of salicylic acid are coloured in blue.

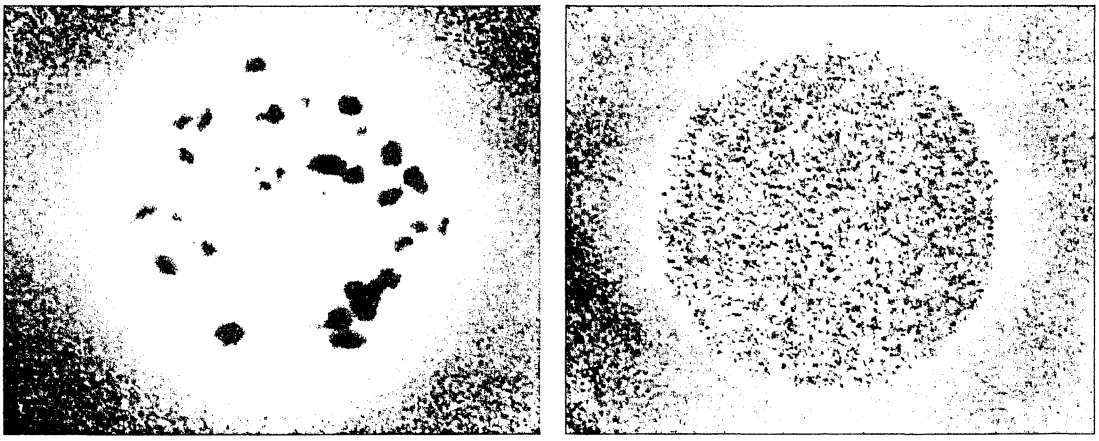


Figure 4.9.

Distribution of acetyl salicylic acid in two different tablets originating from concentration group 8 (7,5%) before (left) and after (right) the substance was grinded and sieved. Regions with a high concentration of acetyl salicylic acid are coloured in red.

To estimate the inhomogeneity regarding active substance, five tablets from each concentration level were measured on both sides by both spectroscopic techniques. The Raman spectra were corrected and flourescence reduced via 4th order polynomial fitting in Matlab. The estimations were based on the peak height ratio of the most dominant peak of the active substance and of the one excipient choosen (microcrystalline cellulose). The distribution of the normalization excipient was supposed to be uniform throughout the tablets. In spectra from tablets containing salicylic acid, the pre-dominant peak at 772 cm^{-1} was used for evaluation, and in the tablets containing acetyl salicylic acid, the peak at 1606 cm^{-1} was used. In both cases, the excipient peak at 1094 cm^{-1} was used for normalization. Homogeneity was calculated as the difference in the peak/peak ratio divided by the average peak/peak

ratio from both sides of the tablet. Figure 4.10 shows the average percental value for the five tablets in each group as a function of concentration. The two lowest concentration levels (groups 1 and 2) were left out, since the active substance Raman peaks could not be distinguished from noise. Differences in homogeneity between tablets containing acetyl salicylic acid and tablets containing salicylic acid were in good agreement with the NIR images. Tablets containing acetyl salicylic acid (see Figure 4.9) were better mixed compared to the salicylic acid tablets (see Figure 4.8) and contained no obvious lumps. In the tablets containing salicylic acid lumps as big as 0.3 mm in diameter can be observed.

The same investigation was made with peak areas instead of peak heights. It gave approximately the same results and therefore only data based on peak height ratios is displayed here. Since the NIR spectra lack sharp features, these spectra could not be used to calculate peak height ratios.

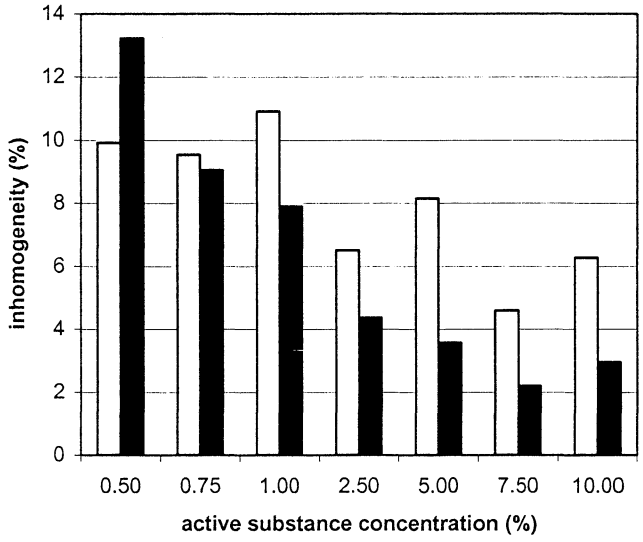


Figure 4.10.

Inhomogeneity calculated from peak height ratios in raw Raman spectra; acetyl salicylic acid tablets (blue) and Salicylic acid tablets (yellow).

Another estimation of the salicylic acid tablet inhomogeneity was made by multivariate models. In model 1, a calibration set consisting of spectra from one side of five tablets in each group of concentration was used. Spectra from the other side of the same tablets were used for prediction. In model 2, the former prediction set was used as calibration set and vice versa. This method can only be valid if the loadings for model 1 and 2 are similar, i.e. the active substance concentrations are calculated in the same 'basis'. The loadings plots belonging to these two models can

be seen in Figure 4.12a and b. Spectra used in the NIR calibration were pre-treated by 2nd derivative and the models used one PLS component. The Raman model was based on raw Raman spectra and used three PLS components. The difference between predicted amounts of salicylic acid from both sides divided by the average predicted amount from both sides were calculated to estimate the inhomogeneity. The calculated inhomogeneity of the five tablets in each group of concentration was averaged. The averaged percental values as a function of concentration level are plotted in Figure 4.11. The results from the NIR calibration are quite surprising and indicate an inhomogeneity much higher than expected. Since the sampled volume is much larger in NIR compared to Raman, lower values of inhomogeneity were expected. One reason for the surprising results may be that the models based on NIR spectra are more sensitive to calibration sets that include few observations. The calculated inhomogeneity according to the Raman model agrees well with the results from the peak height ratio estimation.

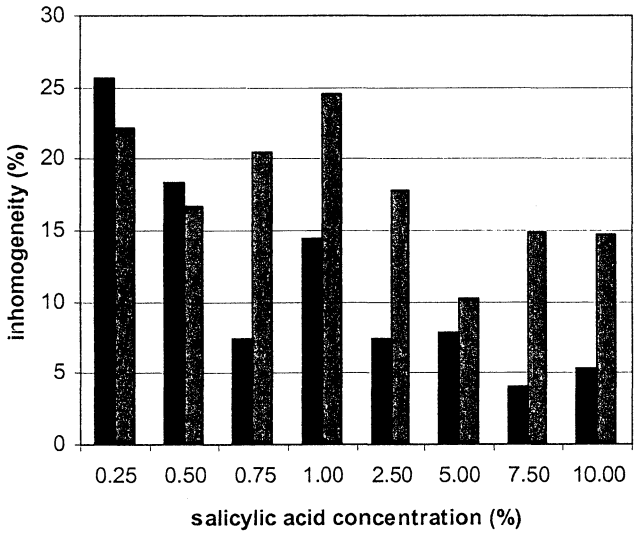
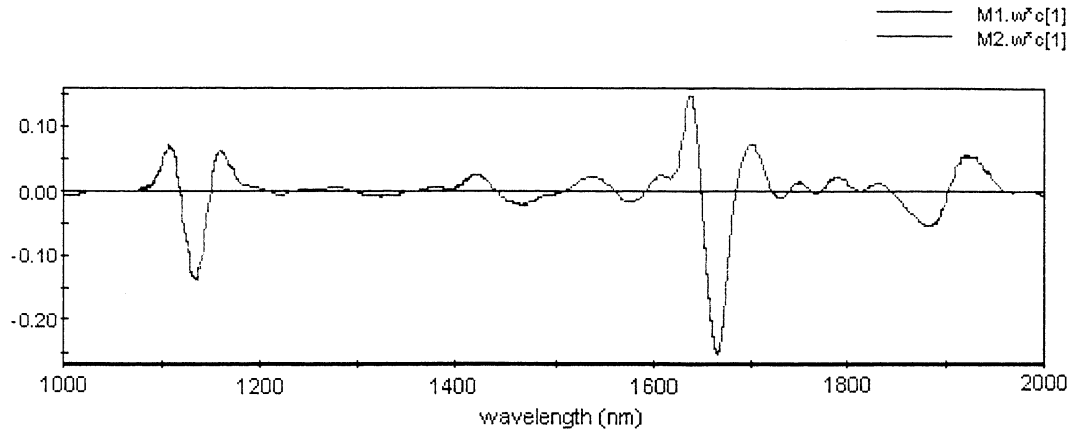
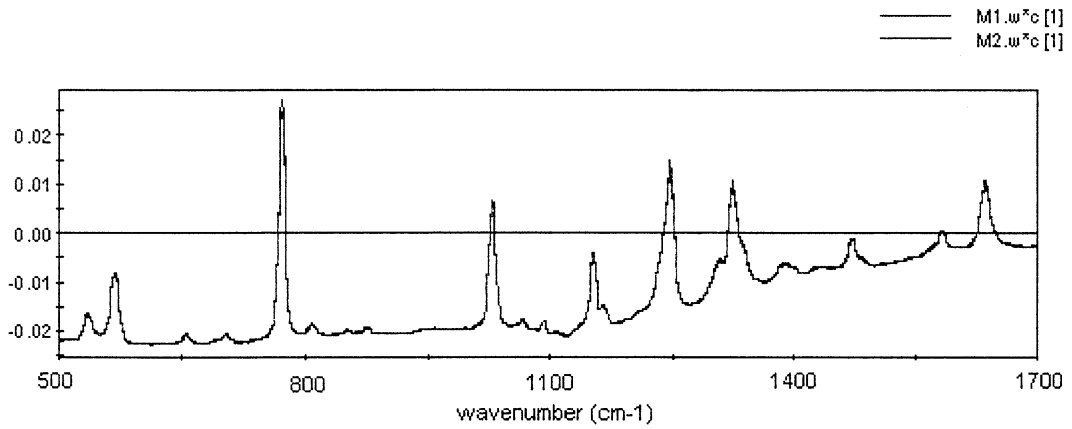


Figure 4.11.

Salicylic acid tablets inhomogeneity according to multivariate models built from NIR spectra (red) and Raman spectra (black).



(a)



(b)

Figure 4.12.

(a) *Loadings for model 1 and model 2 in the NIR calibration (b) Weighted loadings for model 1 and 2 in the Raman calibration.*

4.4 Calibration Based on Dispersive Raman Spectra

Raman spectra of tablets containing salicylic acid were collected in the range 0-1800 cm^{-1}, with an interval of 0.3 cm^{-1}. Spectra of 20 tablets from each group of concentration level formed an **X** matrix used for PLS modelling, with each spectrum as a row. Different pre-treatment methods were utilized. It was found that models built on raw Raman spectra and Raman spectra pre-treated by 2nd derivative resulted in the lowest prediction errors. The selected variable interval 500-1700 cm^{-1} (4000 variables) was chosen since it gave the best prediction ability. A calibration set containing 10 observations from each concentration group was used in the models, giving a calibration set consisting of 90 observations in total. The PLS models built

was used to predict the amount of salicylic acid in the remaining 10 tablets of each concentration group. In the model based on raw Raman spectra, three PLS components were used. The model based on spectra pre-treated by 2nd derivative used one PLS component. Results from the calibrations are displayed in Table 4.2 where the prediction ability is presented both as RMSEP and RMSEP/y_{nom}. The predictive ability RMSEP/y_{nom} versus concentration level is plotted in Figure 4.13. The prediction error of concentration group 2 in the model based on raw Raman spectra is much higher than the error for concentration group 1. To investigate this strange behaviour, a new model was built using the former prediction set as calibration set and vice versa. The result from this calibration can be seen as a dashed line in Figure 4.13. As can be seen in Table 4.2, the prediction errors are quite high for all prediction sets, which to some extent can be explained by the tablet inhomogeneity estimated in Section 4.3. Another conclusion is that no accurate predictions can be made on those tablets contained in concentration groups 1 and 2.

Table 4.2. *Prediction ability of the Raman spectra calibrations.*

		2nd derivative		Raw Raman	
cal. set	**pred. set**	**RMSEP (%)**	**RMSEP (%)/ Ynom**	**RMSEP (%)**	**RMSEP (%)/ Ynom**
1-9c	1p	0.13	130.90	0.04	40.85
1-9c	2p	0.06	23.91	0.18	73.16
1-9c	3p	0.06	11.34	0.05	9.69
1-9c	4p	0.09	11.77	0.09	11.38
1-9c	5p	0.08	8.37	0.06	6.42
1-9c	6p	0.14	5.44	0.16	6.56
1-9c	7p	0.24	4.84	0.26	5.14
1-9c	8p	0.40	5.29	0.36	4.75
1-9c	9p	0.46	4.55	0.40	3.98
1-9c	1-9p	0.21	6.99*		

*RMSEP(%)/Yaverage

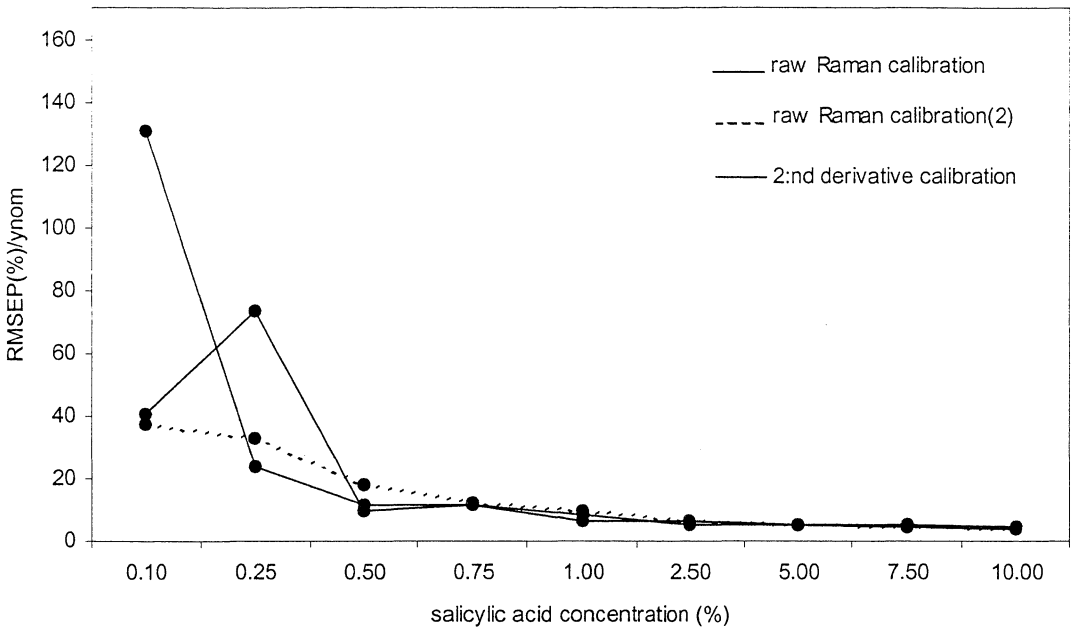


Figure 4.13.

Predictive ability vs. concentration level for the three different Raman calibrations

Figure 4.14 shows the loadings for the first three PLS components in the model based on raw Raman spectra. As can be seen loading $\mathbf{p}_1$ handles most of the fluorescence background and the second loading, $\mathbf{p}_2$, is in good agreement with the pure substance Raman signal in Figure 4.2. The third loading contains some additional information about the active substance. Figure 4.15 shows the loading for the PLS component in the model built on Raman spectra pre-treated with 2nd derivative.

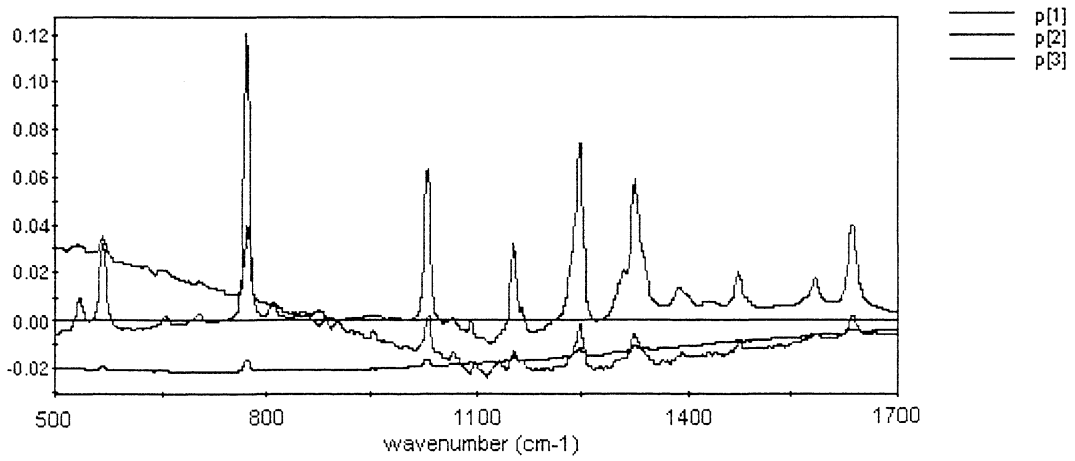


Figure 4.14.

The loadings for PLS components 1-3 in the model based on raw Raman spectra. The wavenumber range 500-1700 cm^{-1} was used for the model.

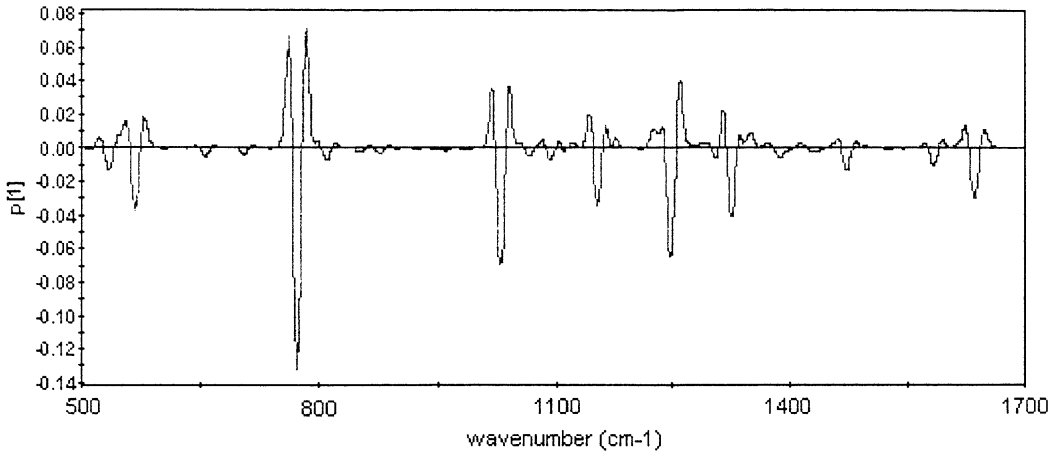


Figure 4.15.

Loading plot of the principal component in the PLS model based on Raman spectra pre-treated by 2nd derivative. The wavenumber range 500-1700 cm^{-1} was used for the model.

Figure 4.16 shows the scores (orthogonal projections) for the first two PLS components, where the concentration level groups 3-9 can be distinguished as separate groups, whereas concentration level groups 1-2 (black and blue coloured diamonds) are more or less bundled together as one. This means that spectra originating from the two lowest concentration groups are very similar, and therefore difficult to predict accurately. Values measured by the reference method versus values predicted by the PLS-model based on raw Raman spectra are plotted in Figure 4.17. The variations in predicted values indicate that there are differences in the spectra that are not explained by the reference values. Figure 4.18 shows measured versus predicted values of the tablets in the four lowest concentration levels. As can be seen, predictions are not very accurate and predictions of tablets from concentration group 2 are particularly bad.

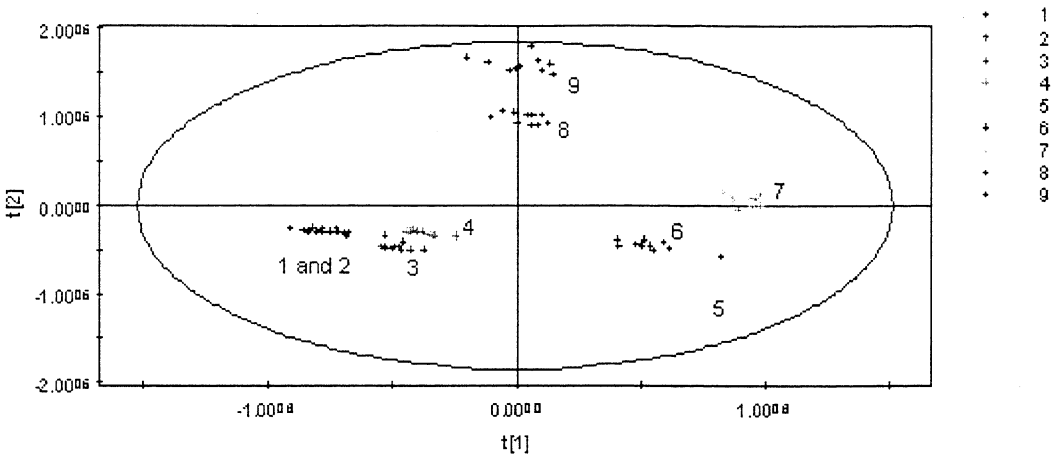


Figure 4.16.

Score values for the first two PLS components in the raw Raman spectra model.

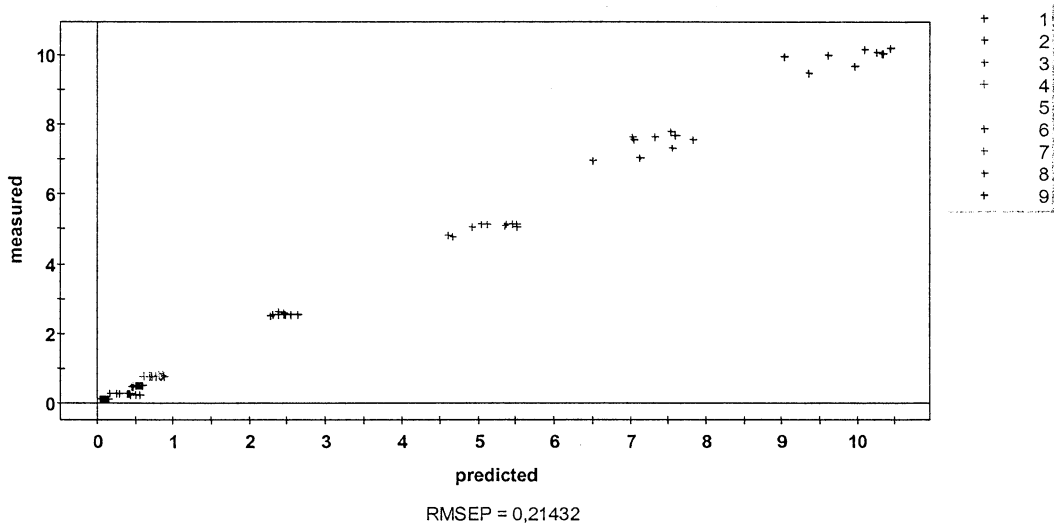


Figure 4.17.

Measured vs. predicted plot of the tablets includes in the prediction set, calculated from the PLS-model based on raw Raman spectra.

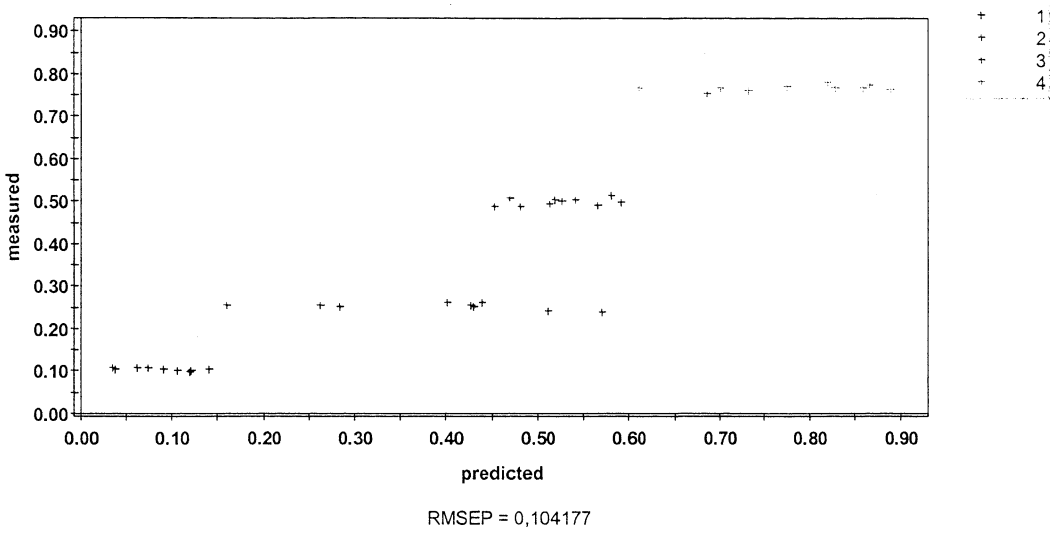


Figure 4.18.

Measured vs. predicted plot of tablets originating from the lowest four groups of concentration level. The model is based on raw Raman spectra.

4.5 Calibration Based on Diffuse Reflectance NIR Spectra

NIR spectra of tablets containing salicylic acid were collected in the wavelength range 400-2500 nm, with an interval of 2 nm. Spectra of 20 tablets from each group of active substance concentration formed an **X** matrix for PLS calibration, with each spectrum as a row. Different pre-treatment methods were tested in order to obtain the best prediction ability. The best prediction ability was found using the pre-treatment methods 2nd derivative and OSC. The wavelength range 1000-2000 nm (500 variables) was used in the models, since it resulted in the best prediction ability. This region includes two sharp salicylic acid absorption peaks (see Figure 4.6). A calibration set containing 10 observations from each concentration group was used. The PLS model built was used to predict the remaining 10 tablets in each concentration group. The prediction ability is presented as RMSEP and as $RMSEP/y_{nom}$ in Table 4.3, and plotted as a function of concentration level in Figure 4.19. To investigate eventual differences between calibration set and prediction set, an additional model, based on NIR spectra pre-treated by 2nd derivative, was built. In this model, the former calibration set was used as calibration set and vice versa. The results from this calibration can be seen as a dashed line in Figure 4.19. The prediction errors of the NIR models are comparable with the level of inhomogeneity for concentration groups 5-9. Another conclusion is that no predictions can be made on those tablets contained in concentration group 1 and 2 independent of pre-treatment method.

Table 4.3. *Prediction ability of the NIR spectra calibrations.*

		2nd derivative		OSC	
cal. Set	pred. Set	RMSEP (%)	RMSEP (%)/ Ynom	RMSEP (%)	RMSEP (%)/ Ynom
1-9c	1p	0.16	164.15	0.11	110.43
1-9c	2p	0.18	70.32	0.16	65.27
1-9c	3p	0.11	21.50	0.19	37.92
1-9c	4p	0.13	16.94	0.28	37.63
1-9c	5p	0.14	13.94	0.31	31.41
1-9c	6p	0.17	6.75	0.24	9.61
1-9c	7p	0.37	7.46	0.14	2.80
1-9c	8p	0.55	7.33	0.29	3.87
1-9c	9p	0.67	6.70	0.24	2.44
1-9c	1-9p	0.33	10.8*		

*RMSEP(%)/Yaverage

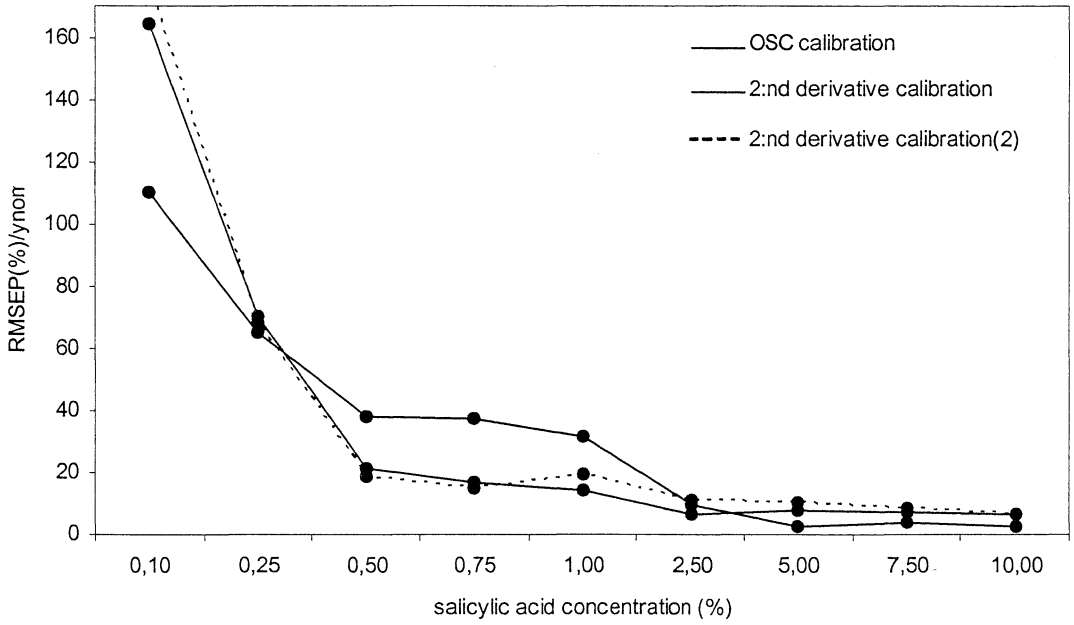


Figure 4.19.

Predictive ability vs. concentration level for the three different NIR calibrations.

Loadings for the PLS components in the two different models based on NIR spectra are plotted in Figure 4.20-4.21. The two peaks in Figure 4.20, which are also visible in Figure 4.21, are the salicylic acid absorption peaks at 1132 and 1660 nm respectively. In the spectral region 1800-2000 nm (variable number 400-500) influence from the excipients can be noticed.

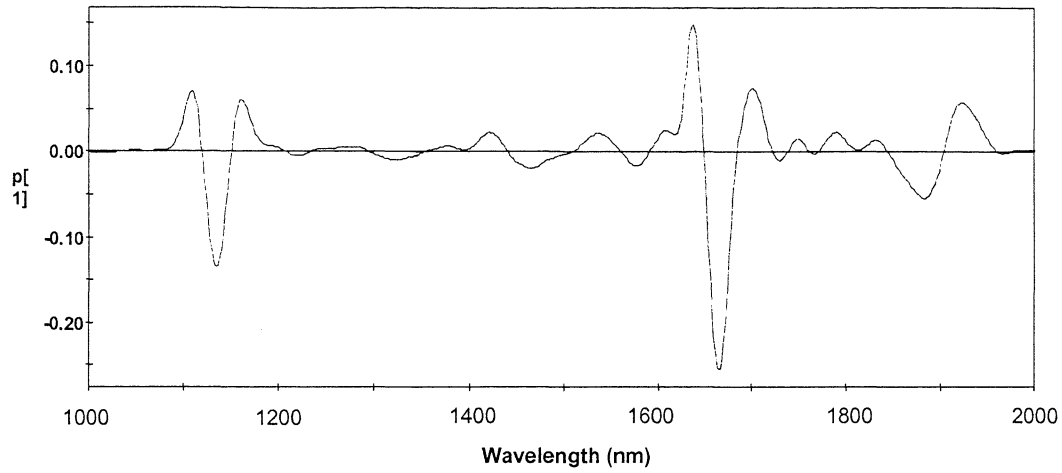


Figure 4.20.

Loading plot of the principal component in the PLS model based on NIR spectra pre-treated by 2nd derivative. The wavelength range 1000-2000 nm was used for the model.

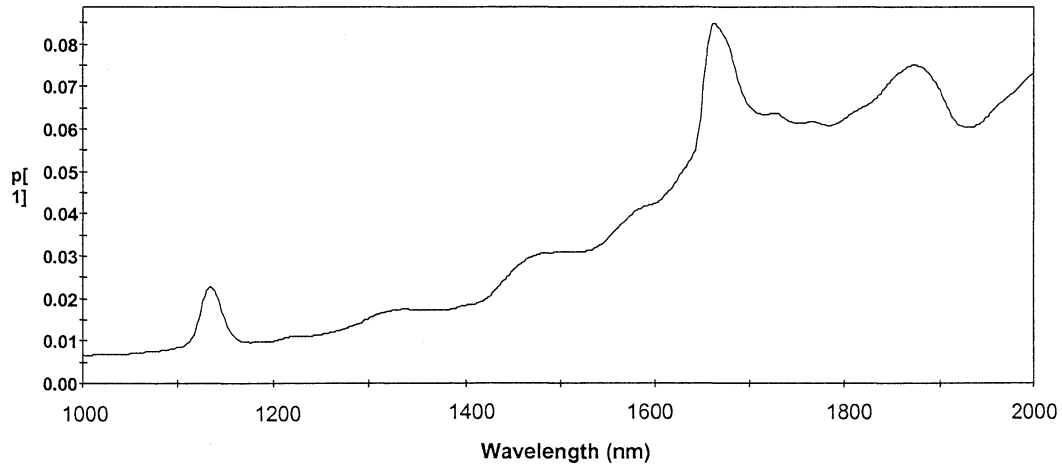


Figure 4.21.

Loading plot of the principal component in the PLS model based on NIR spectra pre-treated by OSC. The wavelength range 1000-2000 nm was used for the model.

Values measured by the reference method versus values predicted by the calibration based on NIR spectra pre-treated by 2nd derivative are plotted in Figure 4.22-4.23. The variations in predicted values indicate that there are differences in the spectra that are not explained by the reference values. Figure 4.23 shows measured versus predicted values of the tablets in the four lowest concentration levels. As can be seen, one of the tablets originating from concentration level group 1 is predicted negatively.

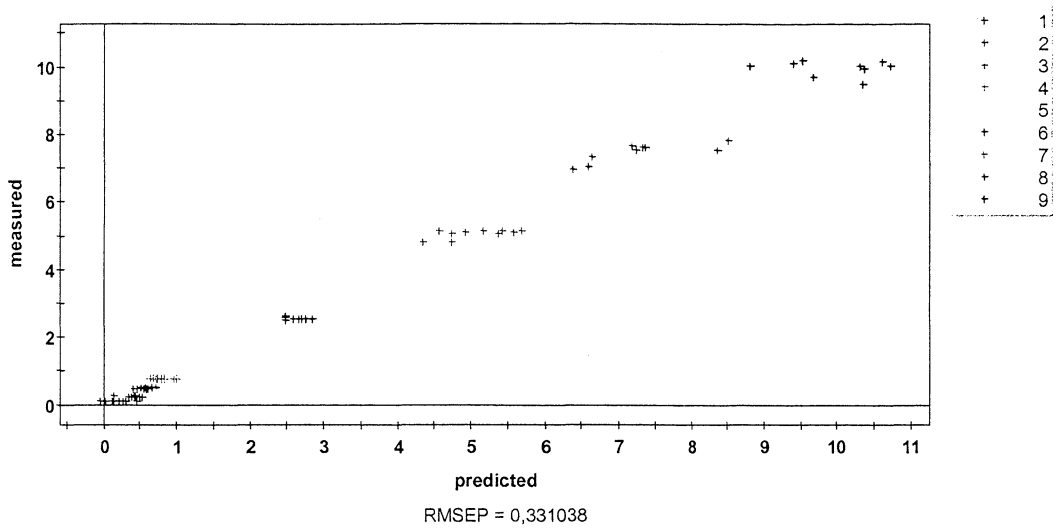


Figure 4.22.

Measured vs. predicted plot of the tablets includes in the prediction set, calculated from the calibration based on NIR spectra pre-treated by 2nd derivative.

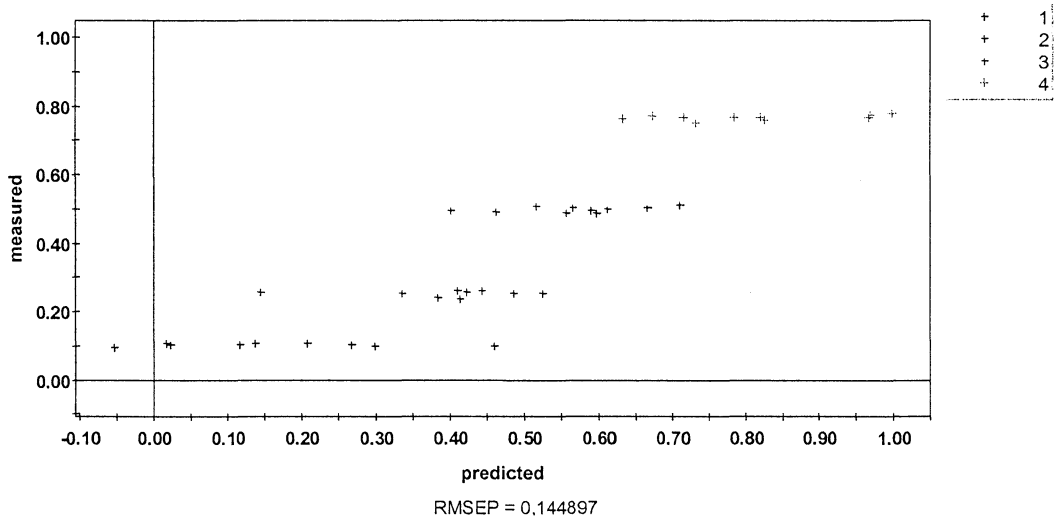


Figure 4.23.

Measured vs. predicted plot of tablets originating from the lowest four groups of concentration level. The calibration is based on NIR spectra pre-treated by 2nd derivative.

The averaged prediction errors from the two models based on raw Raman spectra, and the two models based on NIR spectra pre-treated by 2nd derivative, versus concentration level are plotted in Figure 4.24. As can be seen, the Raman models have an overall better prediction ability compared to the models based on NIR spectra. Another interesting observation is that the models based on Raman spectra indicate prediction ability of observations in concentration groups 1 and 2, even if the predictions are not very accurate.

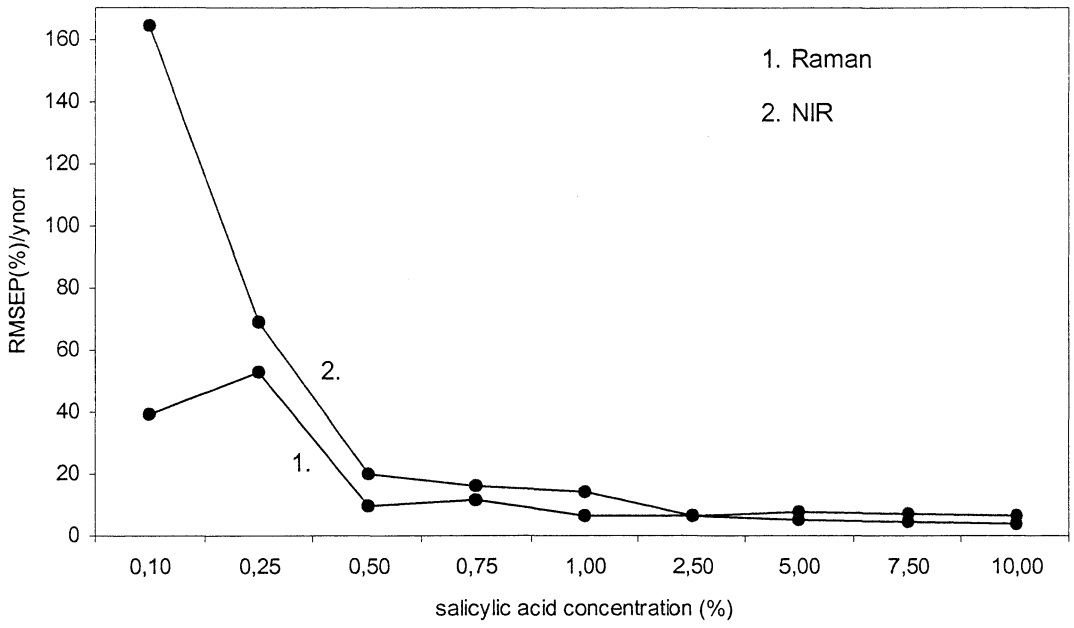


Figure 4.24.

Averaged prediction errors vs. concentration level for models based on the two spectroscopic techniques.

4.6 Model Robustness

Robust multivariate models are able to predict samples (observations) with spectral properties deviating from the samples used in the model. To estimate the robustness of the PLS models, a calibration set containing 10 observations from concentration groups 6-9 were used to predict 10 spectra in concentration group 3 (0.5% salicylic acid). Predicting observations outside the calibration set concentration interval is unavoidable troublesome and demands a linear model. The amount of microcrystalline cellulose, which was varied to account for the reduced proportion of active substance, induces a non-linearity and affects the predictions. Two different approaches regarding the spectral region used in the models were tested. Both model based on NIR and on Raman spectra gave, however, high prediction errors

when applied to a concentration level not included in the model. The models based on NIR spectra showed poor robustness and the prediction ability was highly dependent on the selected spectral interval. The models based on Raman spectra pre-treated by second derivative, showed better robustness and the selected spectral interval was not as important for the prediction ability (see Table 4.4). The lowest relative prediction error was as high as 30%. In the Raman calibrations where all concentration groups were included, the corresponding prediction error was ~10 %.

Table 4.4. *Model robustness estimation.*

Dispersive Raman

cal. set	cal. interval	pre-treatment	prediction set	RMSEP(%)	RMSEP(%)/y_{nom}	PLS-components
6-9c	500-1700 cm-1	2nd derivative	3p	0.15	30	1
6-9c	740-800 & 1000-1060 cm-1	2nd derivative	3p	0.16	32	1

Diffuse reflectance NIR

cal. set	cal. interval	pre-treatment	prediction set	RMSEP(%)	RMSEP(%)/y_{nom}	PLS-components
6-9c	1000-2000 nm	2nd derivative	3p	0.55	110	1
6-9c	1070-1200 & 1600-1720 nm	2nd derivative	3p	0.20	40	1

4.7 Local Models

A narrower concentration interval should reduce the problem with excipient influence in the model. Local models, based on observations from two consecutive concentration levels, were built in order to achieve better prediction ability. Unfortunately, local models could only be built based on spectra from the highest concentration levels and not on spectra from the concentration interval of interest (0.1-1%). There are many reasons why local models based on low dosage tablets are hard to obtain. Examples are the limited number of observations and the limited spectral features.

5. Conclusion

Both spectroscopic techniques are able to predict accurately, according to the multivariate models, down to concentrations of 1% active substance. The relative prediction error at this level of concentration is of the same order as the estimated tablet inhomogeneity. The models based on Raman spectra indicated that accurate predictions may be obtained below concentrations of 0.5%, but further investigations have to be made to confirm this indication. The NIR calibrations show a distinct deterioration in prediction ability of samples with a concentration $\leq$ 0.5 %. The observed relationship between the spectroscopic techniques and their prediction ability give support the hypothesis behind this project. The sharper features in Raman spectra make Raman Spectroscopy well suited for quantitative analysis of low dose tablets.

Since the manufactured tablets did not reach the desired properties of an inhomogeneity less than 5%, conclusions regarding detection limits from the calibrations are difficult.

Robust models are known to be an advantage of Raman spectroscopy compared to NIR spectroscopy. This was confirmed by the models robustness calibrations, even if the prediction ability was far from satisfying.

The tablets containing salicylic acid contained lumps and were less homogeneous than the tablets containing acetyl salicylic acid. According to the estimation based on peak height, homogeneity increased with the active substance concentration. The estimated inhomogeneity of the salicylic acid tablets, calculated from peak height ratios, ranged from 5-11% in the concentration interval 0.5-10%. Estimations based on multivariate models indicated an even larger inhomogeneity.

6. Future Work

As the study has shown, tablet properties have a big influence on the measurements and the calibrations. It is of great importance to make sure that the sampled volumes are representative. By using wet-granulated tablets, a higher level of homogeneity is assured and a more fair comparison can be made.

To get a better understanding of the reliability, limits and accuracy of each technique, another experimental design may be used. By using a design with concentration levels closer spaced in the lowest concentration interval (0-1%), a more trustworthy analysis may be achieved.

Better prediction ability may be achieved using transmission NIR instead of diffuse reflectance NIR, since a larger volume is sampled with the former.

Future work is also required to determine the optimal Raman spectrometer and model conditions. In large spot sampling systems, where the radiation intensity is lower, longer exposure times may be used without inducing bleaching. A longer exposure would ensure a higher signal to noise ratio.

7. References

1. V. Baeten, *Spectroscopy: Developments in instrumentation and analysis.* Grasas y aceites, 2002. **53**(1): p. 45-63.

2. M. Blanco, J. Coello, H. Iturriaga, S. Maspoch, and C. de la Pezuela, *Near-Infrared spectroscopy in the pharmaceutical industry.* The Analyst, 1998. **123**: p. 135-150.

3. M. Blanco and I. Villarroya, *NIR spectroscopy: a rapid-response analytical tool.* Trends in analytical chemistry, 2002. **21**(4): p. 240-250.

4. T. Vankeirsbilck, A. Vercauteren, W. Baeyens, G.V.d. Weken, F. Verpoort, G. Vergote, and J.P. Remon, *Applications of Raman spectroscopy in pharmaceutical analysis.* Trends in Analytical Chemistry, 2002. **11**(12): p. 869-877.

5. G. Fini, *Applications of Raman spectroscopy to pharmacy.* Journal of Raman Spectroscopy, 2004. **35**: p. 335-337.

6. R. Szostak and S. Mazurek, *Quantitative determination of acetylsalicylic acid and acetaminophen in tablets by FT-Raman spectroscopy.* The Analyst, 2002. **127**: p. 144-148.

7. M. Dyrby, S.B. Engelsen, L. Nørgaard, M. Bruhn, and L. Lundsberg-Nielsen, *Chemometric Quantitation of the Active Substance (Containing C≡N) in a Pharmaceutical Tablet Using Near-Infrared (NIR) Transmittance and NIR FT-Raman Spectra.* Applied Spectroscopy, 2002. **56**(5): p. 579-585.

8. J. Coates, *Vibrational Spectroscopy: Instrumentation for Infrared and Raman Spectroscopy.* Applied Spectroscopy Reviews, 1998. **33**(4): p. 267-425.

9. S. Svanberg, *Atomic and molecular spectroscopy: basic aspects and practical applications.* 3rd ed. 2001.

10. N.B. Colthup, L.H. Daly, and S.E. Wiberley, *Introduction to infrared and Raman spectroscopy.* 3rd ed. 1990: Academic Press.

11. P.J. Hendra and J.K. Agbenyega, *The Raman Spectra of Polymers.* 1993: John Wiley & Sons Ltd.

12. P. Hendra, C. Jones, and G. Warnes, *Fourier transform Raman spectroscopy.* 1991: Ellis Horwood Limited.

13. D.E. Bugay, *Characterization of the solid state: spectroscopic techniques.* Advanced Drug Delivery Reviews, 2001. **48**(1): p. 43-65.

14. I.V. M. Blanco, *NIR spectroscopy: a rapid-response analytical tool.* Trends in analytical chemistry, 2002. **21**(4): p. 240-250.

15. L. Eriksson, E. Johansson, N. Kettaneh-Wold, and S. Wold, *Multi- and Megavariate Data Analysis - Principles and Applications*. 2001, Umeå: Umetrics AB.

16. *"Raw Material Identification with FT-NIR"*, http://library.abb.com/GLOBAL/SCOT/SCOT205.nsf/VerityDisplay/C0E75C94E1718F2D85256FA5005E8DD9/$File/4218appnotePH30RawMID.pdf, Accessed in Dec. 2004.

17. J. Gottfries, H. Depui, M. Fransson, M. Jongeneelen, M. Josefsson, F.W. Langkilde, and D.T. Witte, *Vibrational spectrometry for the assessment of active substance in metoprolol tablets: a comparison between transmission and diffuse reflectance near-infrared spectrometry.* Journal of Pharmaceutical and Biomedical analysis, 1996. **14**(11): p. 1495-1503.

18. A. Deninger and A. Zach, *"An XTRA Source for Raman Spectroscopy: Towards Process Control"*, http://www.toptica.com/products/itemlayer/44/AN-XT-001v02.pdf, Accessed in Dec 2004.

19. D.P. Strommen, "Raman Spectroscopy" In *Handbook of Instrumental Techniques for Analytical Chemistry.* F. A. Settle, Ch 16, 1997.

20. *RamanRxn1™ Operations Manual.* 2004, Kaiser Optical Systems, Inc.

21. D.A. Long, *The Raman Effect: A Unified Treatment of Raman Scattering by Molecules*. 2002, John Wiley & Sons Ltd. p. 31-48, 49-84.

22. J. Johansson, S. Pettersson, and S. Folestad, *Characterization of Different Laser Irradiation Methods for Quantitative Raman Tablet Assessment.* Journal of Pharmaceutical and Biomedical Analysis. *(in press).*

23. J. Johansson, S. Pettersson, L.S. Taylor, *Infrared Imaging of Laser-Induced Heating during Raman Spectroscopy of Pharmaceuticals.* J. Pharm. Biomed. Anal, 2002. **3**: p. 1223-1231.

24. P. Gelaldi, *Chemometrics in spectroscopy. Part 1. Classical chemometrics.* Spectrochimica Acta Part B, 2003. **58**: p. 767-782.

25. S. Wold, H. Antti, F. Lindgren, and J. Öhman, *Orthogonal signal correction of near-infrared spectra.* Chemometrics and Intelligent Laboratory Systems, 1998. **44**: p. 175-185.

26. A. Savitzky, M.J.E. Golay, *Smoothing and Differentiation of Data by Simplified Least Squares Procedures.* Analytical Chemistry, 1964. **36**(8): p. 1627-1639.

27. K. Wiinikka, *"Utvärdering av det avbildande NIR-systemet MatrixNIR. Instrumentella egenskaper och tillämpningar".* Examensarbete 20 poäng, Analytisk och marin kemi, Göteborgs Universitet, 2003.

28. C.A. Lieber and A. Mahadevan-Jansen, *Automated Method for Subtraction of Fluorescence from Biological Raman Spectra.* Applied Spectroscopy, 2003. **57**(11): p. 1363-1367.

29. *P^{h}AT System^{TM} Operations Manual.* 2004, Kaiser Optical Systems, Inc.

30. http://www.medicinescomplete.com/mc/excipients/current/noframes/introduction.htm. Pharmaceutical Press and American Pharmacists Association 2004. Accessed in Dec 2004.

31. http://www.infomedica.com. Infomedica AB. Accessed in Dec 2004.

8. Appendix

Specifications of the NIR system

Scan Rate	1.8 scans/second
Wavelength Range	1100-2500nm (Standard Mode)400-2500nm (Wide-Range Mode)
Wavelength Accuracy	±0.3nm (Basis: Polystyrene Standard peaks 1100-2500nm)
Wavelength Repeatability	±0.010nm (Standard Deviation of 10 consecutive scans 1100-2500nm
Noise (peak-to-peak)	<3.0 (400-700nm) <.4 (700-2500nm)
Noise (RMS Avg)	<.2 (400-700nm) <.04 (700-2500nm)
Spectral Bandwidth	10 ± 1nm
Photometric Range	3.0 Au (1100-2500nm)
Linearity (2% to 99% reflectivity or transmissivity)	±1.0%
Stray Light	Less than 0.1% at 2300nm
Operating Temperature Range	15°C to 33°C (60°F to 92°F)
Power Consumption	150 watts (maximum

Specifications of the Raman system

General

Operating temperature range	5 to 35 °C

Base Unit

Warm-up time	20 minutes
Operating voltage	100–240 V, 50–60 Hz
Power consumption(maximum)	200 W
(typical startup)	175 W
(typical running)	135 W
Dimensions (width x length x height)	23" x 17.63" x 8"
Weight	61 lbs

Camera

Chip	1024 × 256 EEV MPP type
Pixel size	26 µm square
Full well capacity (single pixel)	400 000 electrons
Full well capacity (register)	800 000 electrons
Gain	typically 4 electrons/count
AD converter	16 bit
Minimum exposure time	25 ms
Minimum detector temperature	–50 °C
Detector linearity	Better than 1%
Detector readout noise	1–1.2 counts rms

Spectrograph

Type	Proprietary axial transmissive
Aperture ratio	*f*/1.8
Focal length	85 mm
Grating	holographic transmissive
Spectral coverage	150 to 1890 cm–1
Slit	50 µm fixed, removable
Spectral resolution	5 cm–1 average

Laser-785-nm Invictus™

Excitation wavelength	785 nm
Warranty	unlimited hours for 1 year

Probe

Laser power at sample	> 20 mW
Nominal focal length	
3-mm lens	120 mm
4.5-mm lens	175 mm
6-mm lens	250 mm
Nominal beam diameter at focal position	6 mm
Weight	2 lbs (with cable)
Length	12 in
Diameter	1.8 in
Cable length	3 m