CHAPTER II

LITERATURE REVIEW

1. THE MODEL DRUG

Azithromycin dihydrate is the parent compound in a class of antibacterial obtained from acid-stable erythromycin A. The active form is a crystalline dihydrate ($C_{38}H_{72}N_2O_{12}.2H_2O$). However, its chemical structure differs from erythromycin in that a CH_3-substituted nitrogen atom is substituted into the 15-membered ring (as shown in Figure 1). Azithromycin is the most important drug of a subclass of antibacterial, known as azalides. It is an effective therapeutic agent in oral administration for infections, showing very high antibacterial effect (bacteriostatic) against both Gram-positive and Gram-negative bacteria. Its action of mechanism is inhibiting protein synthesis on the ribosome with preventing growth of bacteria. (Timoumi 2014, Robaina 2013)

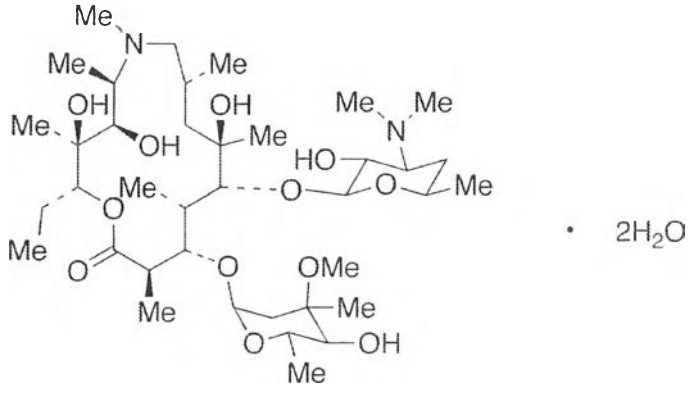


Figure 1 Chemical structure of azithromycin dihydrate.

Azithromycin dihydrate (AZD) is a white crystalline powder with a 785 molecular weight. It has intensively bitterness taste and slow rate of dissolution which decreases patient compliance and leads to the poor bioavailability respectively (Gandhi 2002).

N-atom is substituted in 15-membered lactone ring or azalide structure. The carboxyl ester group and N-dimethyl group are the major functional group in macrolide antibacterial such as clarithromycin, roxithromycin, erythromycin and telithromycin (Tsuji 2006). The taste in macrolide antibacterial is much bitter, due to these functional groups that affinitive and selective to the bitter taste receptor. Their bitter taste is carried by couple of G-protein receptors as subtype TR, especially N-atom which is the major component of protein substance. (Matsunami, 2000)

The bitter taste is the last and one of the four taste receptors in taste buds of tongue. The bitter taste receptor can be found in the back area of the tongue. It is activated by a variety of substances, especially organic compounds. The inhibition and reduction of Bitterness are necessary characteristics of a good oral pharmaceutical dosage forms (Mennella 2013). Determination for amount of progress has been succeeded in the development of taste-masked formulations in recent years. Various techniques are available for masking drug bitter taste, including taste masked with ingredients such as sweeteners, flavors, and amino acids; taste masked by conventional granulation; taste masked by polymer coating; taste masked with ion-exchange resin etc. (Douroumis 2007, Tawakkul 2009)

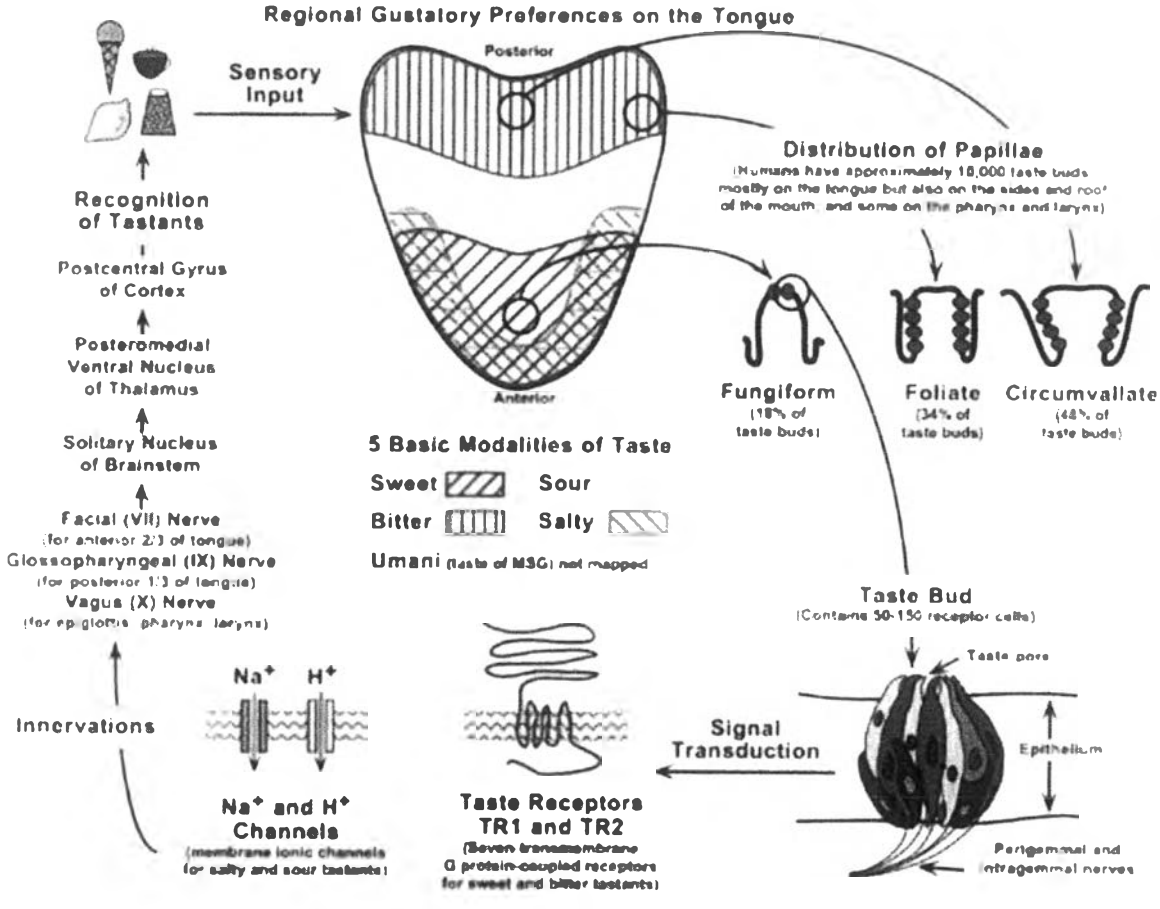


Figure 2 General diagram of taste receptor function.

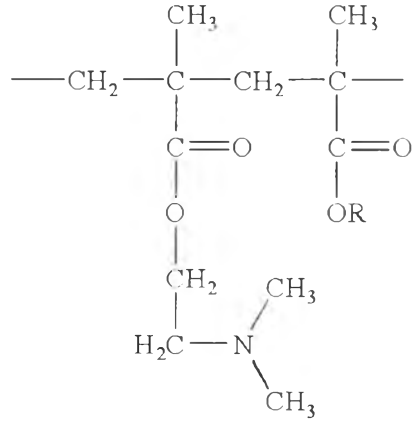
(http://www.bio.davidson.edu/courses/molbio/molstudents/spring2000/kazama/restricted/tastegram.jpg [cited 2013 May 5])

The taste reception is a mostly important part of an animal's life. Mammals are thought to distinguish between only 5 basic taste modalities: bitter, salty, sour, sweetness and umami. The taste cells of mammals are located in taste buds on the tongue and mediated by signal transduction pathways and expressed into the receptor cells. It shows that bitter taste receptor cells act as using specialized membrane channels. The function of these membrane channels using G protein-

coupled receptors, which originate signaling order, ending in neurotransmission by cranial nerves transduction (Gupta A.K. 2010, Mennella, Spector 2013).

2. TASTE MASKING AGENT

Recently, the well-known effective taste masking polymer of bitter drugs is Eudragit® E PO. The polymer can be formulated as a Eudragit® E PO solution (Yan 2010). It acts as an effective humidity barrier coating method (Maniruzzaman 2012). Aqueous Eudragit® E PO dispersions exhibit efficient taste masking by coating on tablets and particles (Guhmann 2014, Pawar and Joshi 2014). Above pH5, It is permeable, swellable and soluble below pH5. The glass transition temperature (T_g) of Eudragit® E PO is about 50°C.



R = CH_3, C_4H_9

Figure 3 Chemical structure of Eudragit® E PO.

(http://patentimages.storage.googleapis.com/US20090061048A1/US20090061048A1-20090305-C00001.png [cited 2013 July 27])

Polymers may be sprayed onto various types of cores from solutions or dispersions in solvents (Joshi and Petereit 2013). Taste masking is critical importance for active pharmaceutical ingredients with a decreasing of bitterness, due to the need for increasing patient adherence. In the pharmaceutical industry, taste masking technologies involve the development of a system that produces the barrier of active substance interacting with the taste buds, thereby inactivating or reducing the negative response sensation.

3. TASTE MASKING TECHNOLOGIES

Taste masking is described as a perceived minimization of an unpleasant taste that would otherwise exist. Taste masking technologies offer a great scope for invention and patents. The ideal solution to taste masking of bitter substances is the finding of a universal inhibitor of all bitter tasting drugs that can not affect the other taste reception such as sweetness (Gittings 2014). Reliable methods for bitterness minimization and prevention have resulted in modified palatability of oral dosage forms (Sohi 2004, Dhakane 2012). Taste masking technology is an important factor for improving patient compliance. This technology includes mainly two aspects:-

1. Selection of compatable taste masking agents such as sweeteners, polymers, flavors, amino acids,
2. The taste masking method was selected by suitability.

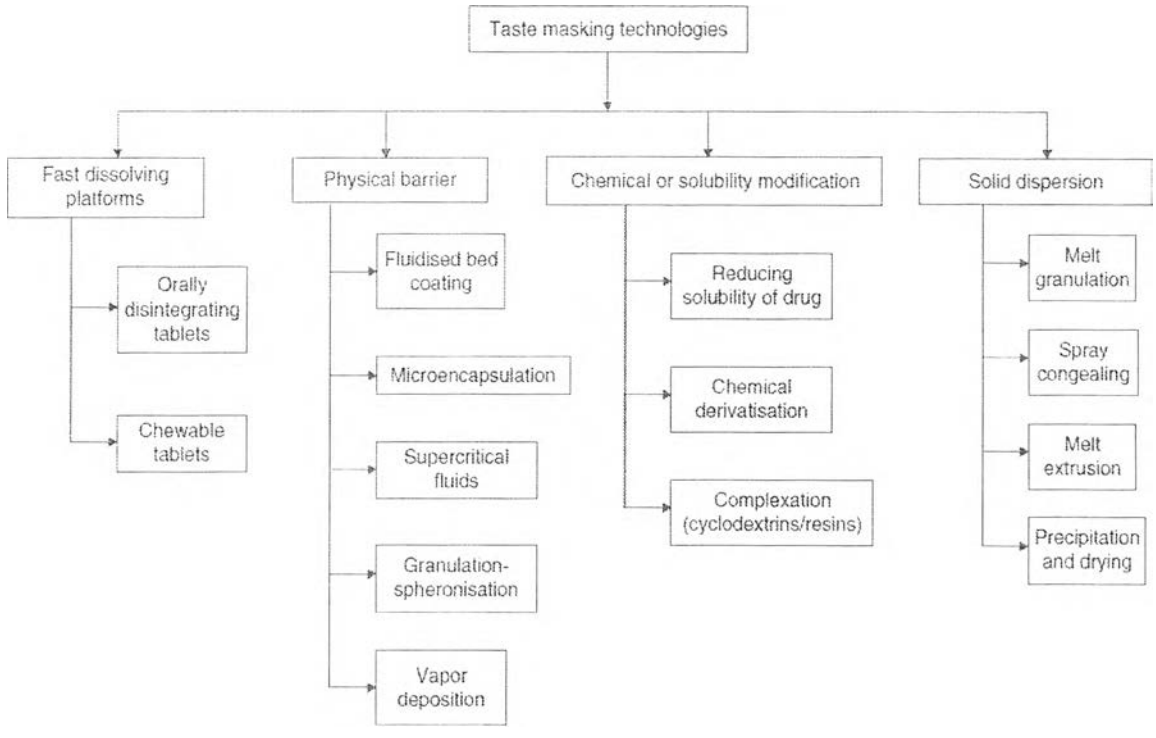


Figure 4 General taste masking principles (Douroumis 2007).

Recent developments in taste masking technology have presented obvious and robust dosage platforms with excellent taste masking properties. Most of these technologies enable excellent taste and odor masking, reduced drug dosage, enhanced therapeutic effect, improved drug absorption rates and patient compliance.

From the Figure 4, Physical barrier and chemical modification are used for many taste masking technologies in nowadays. Polymer coating is an excellent method in taste masking of bitter drugs. It avoids the direct contact of bitter drugs on the taste bud cells. The selection of polymer is based on its physicochemical property and compatability with drug (Sikandar 2011). Granule preparation is inexpensive, fast method and it is suitable for taste masking of slightly bitter tasting substances. Granulation decreases the effective surface area of the bitterness that

comes in contact with the tongue upon oral administration and thus provides effective taste masking. The techniques of dry, wet, and melt granule preparation are used to process the mixture of bitter substances, hydrophobic polymers, sweeteners and lipids or waxes to prepare taste masked oral solid or liquid formulations (Douroumis 2007). Among various procedures two are commonly used to eliminate the bitterness of drug.

1. By minimizing the drug solubility in the saliva pH (5.6-6.8).
2. By adjusting the affinity and nature of drug which will activate with the taste receptor.

Taste masking process is a pharmaceutical procedure with variability. Due to its various technologies of taste masking that can put into use. From literature review of PAT application for taste masking, the studies of Rahman (2012) is the only one that used NIRs-CI as process analyzer to determine the taste masked chlorpheniramine complexes preparation by solvent evaporation method. In this study, we focused on application of PAT to monitor the end-point of taste-masked azithromycin dry powder preparation by granulation and polymer coating method.

4. PROCESS ANALYTICAL TECHNOLOGY (PAT)

The aims of PAT are to alter the present concept and production within the pharmaceutical industry move to real-time process control or monitoring instead of the intermediate or final product off-line analysis. The PAT procedure utilized to the pharmaceutical manufacturing is trended to increase the understanding of process, stimulate analysis of in-process and end product and initiate the suitable process control.

PAT is the other word for PAC (process analytical chemistry) that transforms the processing variances into scientific information. These well-known methods have been applied in many industries fields. (Knop and Kleinebudde 2013, Chen 2008). The process understanding and PAT key concepts, which recommended by the PAT guidance according to USFDA, are

(1) Analysis tools and the acquisition of multivariate data,

(2) Process analyzers and,

(3) The end-point of process monitoring and control, and

(4) The continual improvement of information management.

Chemometric acts as a PAT concept critical function. PAT means to design and analyze methods that obtain timely data information about the process control and develop the process understanding (Mark 2010). Chemometric procedure can compress concerning information from the variances of process that can be used for timely multivariate analysis model construction. The quality of patterns between starting materials, in-process and final product properties can be built up (Skibsted and Engelsen 2010). Moreover, the reproducible manufacturing with minimal errors could check and monitor critical process parameters. The process for risk-based analysis is important for PAT processing. The goal of the risk-based analysis is to reduce and inhibit the high risks such as an error constituents amounts in formulations, the distributions of particle size error or incorrected data being analyzed from the process (Menezes 2009, Moes 2008, Haleem 2013). The real time release (RTR) process for the final product without traditional pharmaceutical analysis is the main objective of PAT. (Puchert 2011, Wu 2011, Aksu 2012). This refers to the prediction of process and final product quality. (Momose 2011, Rosas 2012, Kumar

2013). The quality of off-line tests for final and in-process products may not be totally instead by real-time PAT process. PAT helps with decreasing the release for sale process consuming time. RTR process also improves products quality, including reduces the sampling errors. (Wang 2011, Karande 2010).

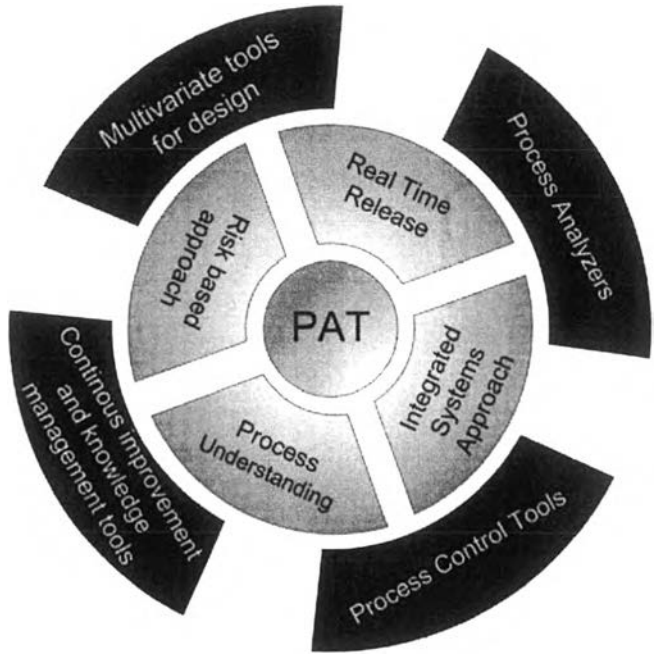


Figure 5 Scheme of a process analytical technology (PAT).
(http://www.intechopen.com/source/html/44781/media/image1.jpeg [cited 2013 October 13])

In the conceptual of PAT process is the measurement from the production line, which consist of 3 real-time measurements; in-line, on-line and at-line. In-line refers to the measurement fuses in the production line, which could inform the data real-timely. (Chablani 2011) . On-line is quite similar to the in-line (the measurement is fuse in), but the sample has to take out of the production line and return back after the measurement. (Schaefer 2013). Both of these measuring processes are not cause the loss of products. Lastly, at-line, the sample is randomly takes out of the

production line and measure with the machine, which not fuses in the production line, the sample is not return back to the production line (Corredor 2011). On the other hand, off-line is traditional time consuming pharmaceutical analysis. The PAT key concepts consist of; (Rajalahti and Kvalheim 2011):

(1) non-destructive sampling and monitoring methods that timely process data and,

(2) multivariate analysis in chemometric that transform and analyze data variances into detailed information.

The benefit of the process control and optimization for quality of end or in-process product could be described by using chemometric models. The outcomes of in-process can be demonstrated with the critical alert limit of process (Momose and Imai 2011, Puchert, Holzhauer 2011, Troup and Georgakis 2013).

5. NEAR-INFRARED SPECTROSCOPY (NIRs)

5.1 Spectral region of near infrared radiation

The associated wavelength region was discovered and described by Sir William Herschel as early as in 1800, when he diffracted sunlight with the aid of a prism. But for a long time, nearly no use was made of this discovery. It was in the 1960s, when Karl Norris published his work using near infrared spectroscopy for analyzing agricultural products, reprinted in 1996 (Pasquini 2003).

The exploit of NIR spectroscopy began with the progress of the computational power in the 1980s. Computers are of great importance for NIRs data acquisition and evaluation. Nowadays, NIRs spectra are recorded and evaluated extremely fast and uncomplicated software allows the interpretation of complex

spectra even without profound expertise. As a result, NIRs is now used in plenty of different industries. Besides the agricultural and food industry, the technique is also applied in most manufacturing industries as in semiconductor industry and in refining industries for classification of motor oils. Its use is so widespread that tool is also applied in medicine e.g. for non-invasive blood glucose tests (Jamrogiewicz 2012).

5.2 Spectral region of near infrared radiation

The NIRs spectral region is defined as the region between the visible and the mid-infrared (MIR) wavelength range. The spectral region extends from 780 nm to 2500 nm according to the European Pharmacopoeia. As it is more common to refer to the wavenumbers as in the mid-infrared region, this means about 12500 cm^{-1} to about 4000 cm^{-1} (Pasquini 2003). The location of the NIRs radiation in the electromagnetic spectrum is depicted in Figure 6.

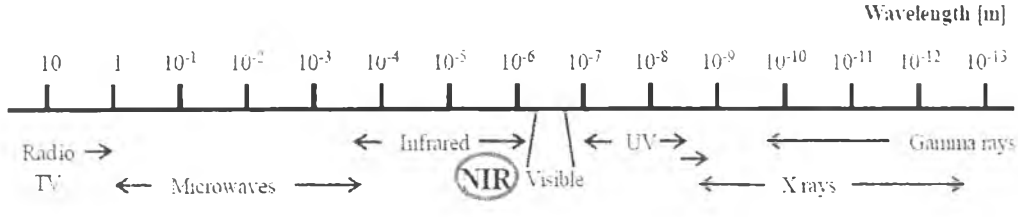


Figure 6 Electromagnetic wavelength.

The wavenumber $\tilde{\nu}$ is the reciprocal of the wavelength λ and is typically in spectroscopy expressed in the unit cm^{-1}.

$$\tilde{\nu} = 1/\lambda$$

Hence, the wavenumber is directly proportional to the frequency of the absorbed radiation and also to its energy.

5.3 Vibrations

NIR and MIR spectroscopy are both vibrational spectroscopic techniques which detect absorption of radiation due to vibrations of atomic bonds. The uptake of specific energy quanta induces stretching and deformation vibrations of atomic bonds and a change in the respective diploe moment. In contrast, changes in the dipole moment due to molecule rotations are induced and detected in far infrared spectroscopy (Skibsted and Engelsen 2010, Sacré 2014).

5.3.1 Stretching vibrations

The change of the distance of inter-atomic along the bond-axis is called stretching vibration. Stretching vibrations of several atomic bonds can be either symmetric or asymmetric. A schematical illustration of these two stretching vibrations is presented in Figure 7.

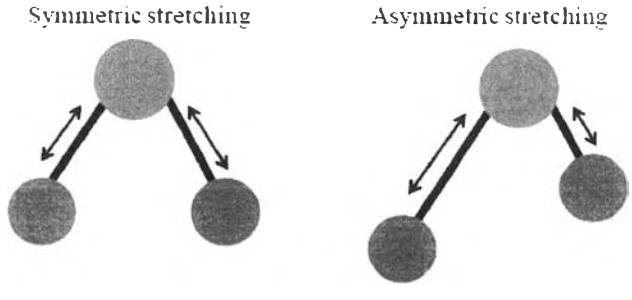


Figure 7 Symmetric and asymmetric stretching vibrations.

5.3.2 Bending vibrations

Bending or deformation vibrations change the angle between two atomic bonds. These changes occur as in-plane or out-of plane vibrations. The differences are illustrated in Figure 8.

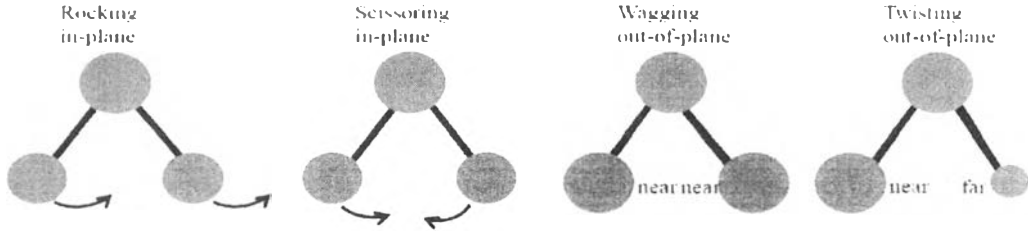


Figure 8 Types of bending vibrations.

5.3.3 Fundamental, overtone and combination vibrations

At room temperature, molecules are in their ground state of energy. By irradiation of material with infrared light, specific energy quanta can be absorbed and higher levels of energy can be attained. Figure 9 illustrates schematically the potential energy and the related energy levels according to the harmonic and anharmonic oscillator model. Contrary to the harmonic oscillator model, which is symmetric and parabolic in shape, the anharmonic oscillator model is characterized by a dissociation level, which means that the interatomic distance cannot be infinite but approaches a dissociation level. Moreover, experiments showed that deviating from the harmonic oscillator model, the distances between succeeding vibrational energy levels (v) are not equal (Jamrógiewicz 2012, Zhang 2010).

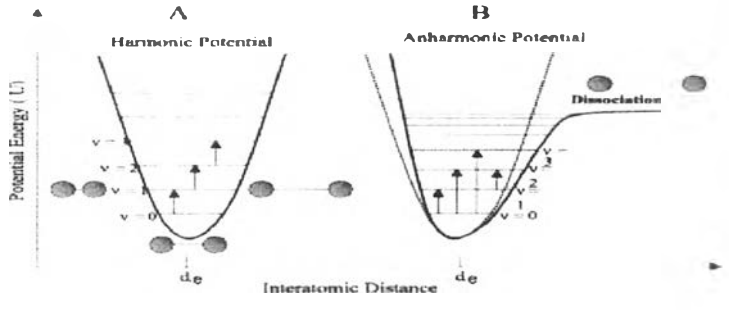


Figure 9 Schematic representation of harmonic (A) and anharmonic (B) oscillator model. (Pasquini 2003)

The transition from the ground state (v=0) to excited state at the first level (v=1) is called fundamental transition of which the occurrence is most probable and the required excitation energy is lowest. The resulting absorption peaks, which can be found in the mid-infrared wavelength range, are of high intensity.

Selection rules according to classical quantum mechanics allow transitions to the next energy level, whereas transitions over multiple energy levels are "forbidden transitions". These so-called overtone vibrations are mainly detected in the NIRs range.

The first overtone vibration is thus the transition from the ground stage of energy v=0 to the level v=2, second overtone vibrations overcome the energy difference from v=0 to v=3. The probability for the transitions of 1^{st} and 2^{nd} overtone vibrations is low, thus the intensity of their absorption peak decreases compared to IR spectroscopy which focuses on fundamental oscillations. NIRs radiation is applicable to excite these oscillations since its energy is higher than of mid-infrared light. Beside of overtone vibrations are also induced by NIRs radiation. The frequency of combination vibrations corresponds approximately to the sum of frequencies of multiple fundamental vibrations (Sacré, De Bleye 2014, Martínez 2013).

Due to the fact that overtone vibrations require a high degree of mechanical anharmonicity of the vibrating atoms in the molecule which means a profound difference in the mass of the vibrating atoms, mainly O-H, S-H, N-H and C-H bonds are NIRs-active. Detection of overtone and combination vibrations results in an absorption spectrum in which the peaks are in a high degree overlapping and not as distinct as in an IR spectrum. Hence, chemometric techniques are required to interpret the complex NIRs spectra (Skibsted and Engelsen 2010, Pasquini 2003).

Near infrared spectrophotometer consists of a light source, a filter, grating or interferometer system with a range of wavelength or wavenumber in the NIRs region and an suitable detector typically of indium gallium arsenide or lead sulphide to measure and collect the light transmitted or reflected by the sample. Several types of NIRs are widespread used: interference-filter spectrophotometers, grating based dispersive, acousto-optic tunable filter units or multichannel Fourier-transform- (FT) spectrophotometer (Jamrógiewicz 2012, Schönbichler 2013).

5.4 Fourier-transform spectrophotometer (Michelson interferometer)

FT-spectrophotometer collects all wavelengths simultaneously. FT-NIR spectrophotometer equipped interferometers recover the intensities of individual wavelengths in the NIRs region. This technique is advantageous concerning the achieved wavelength precision and accuracy as well as the resulting high signal-to-noise ratio and scan speed. FT-spectrophotometer is usually based on the Michelson interferometer (see Figure 10). This interferometer consists of a beam splitter (e.g. a semi-transparent mirror) and two mirrors, a fixed one and a precisely movable one. Light from the source is divided into two beams by the beam splitter. There the beams are superimposed, but again separated into two beams: one illuminating the sample and reaching the detector, the other one being reflected back to the source (Räsänen and Sandler 2007, Ozaki 2012). For measurements, distance of the movable reflection mirror to the beam splitter is varied and the intensity of interference is measured at the detector. This intensity of interference is called interferogram.

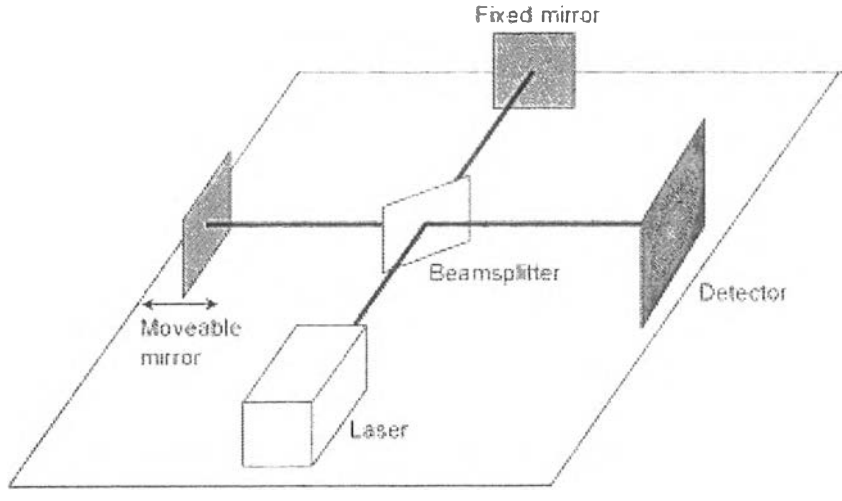


Figure 10 Schematic of the Michelson interferometer.

(http://upload.wikimedia.org/wikipedia/en/d/d0/Michelson-interferometer.png [cited 2013 July 27])

The collected interferogram is transformed to a single-beam frequency-domain spectrum. For calculating the reflectance spectrum, a reference spectrum is compressed from the single-beam spectrum. Usually, reflectance spectra are converted into pseudo-absorbance spectra which are calculated as the negative logarithm of the reflectance (Krämer and Ebel 2000, Freitas 2005).

5.5 Acquisition of spectral data

In general, NIRs spectra are collected using one of the following three measuring modes: transmission, diffuse reflection or transflection.

5.5.1 Transmission

Measurements in transmission mode detect the radiation which has passed through the sample. This mode of data acquisition is typically known from UV-VIS spectroscopy of liquid samples (Bodson 2006). Transmission is defined as the fraction

of incident radiation I at defined wavelengths which can be detected after radiation with the intensity I_0 has been passed through a sample.

$$T=I/I_0$$

For measurements in transmittance mode, the sample is settled between the optical light source and the detector. The result can be either expressed in transmission or absorbance values.

$$A= -\log_{10}T = \log_{10}(1/T)$$

The transmission mode is often applied when tablets or liquids are to be analyzed. The relatively high penetration depth in the sample is advantageous, while ate the same time, the sample thickness which can be measured is limited (Bodson, Dewé 2006, Gendrin 2008).

5.5.2 Diffuse Reflection

The cumulative effect of the phenomena absorption, refraction and reflection on solid surfaces is defined as diffuse reflection. The intensity of light ratio reflected by the sample to the one reflected by a reference surface is measured in reflectance spectroscopy. Most frequently applied reference surfaces are gold and ceramic standards. Considerations should be given to physical attributes of the sample such as packing density, sample depth and variations in measuring e.g. coverage of detector window, sample cell variation and probe pressure (Kramer and Ebel 2000).

5.5.3 Transflection

Transflection measuring mode is a hybrid of transmission and reflection. A reflection plate behind the sample reflects the transmitted radiation back through the sample to the detector which is located in front of the sample and thus on the

same side as the light source (Blanco and Peguero 2010, Blanco and Romero 2002). Using this technique, the incident radiation passes twice the sample.

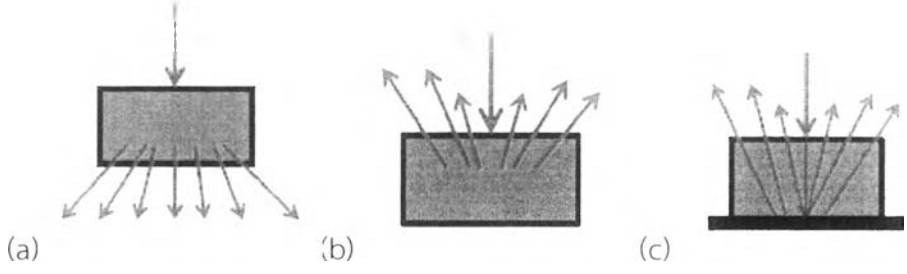


Figure 11 Schematic of measuring mode for (a) Transmission (b) Diffuse reflection on solid samples and (c) Transflection. (Pasquini 2003)

The development of a NIRs calibration model for quantitation is challenging and time-consuming. Plenty of samples have to be manufactured especially for calibration purpose, which is particularly difficult in a manufacturing environment. Hence, calibration samples are usually manufactured on laboratory equipment using respective material and processes. Nevertheless, even deviations such as using another press type for compression into tablets can lead to calibration models which are not applicable for real manufacturing processes. Furthermore, NIRs models for quantitation rely on reference methods, which themselves contain a certain degree of variation. This test method can be the major source of variation in NIRs prediction results of calibration models (Vanarase 2013, Wahl 2014). Moreover, it is necessary to correct or update existing NIRs calibration methods when a drift whether slope or bias error is obvious. This can be due to process or measurement equipment age, conditions change of environment or the application of new lots of starting materials. The effective sample size in near infrared analysis is often significantly smaller than for off-line methods because the illuminated area of the sample is usually relatively small due to the area of the beam. Therefore, heterogeneity on a micro-scale can be

detected by NIRs, but may not be consistent with results of reference methods (Blanco and Peguero 2010, Bu 2013).

Both quantitative and qualitative multivariate analysis modelling aims to have been used in taste masking blending process monitoring. NIRs instruments are well-known used PAT application to detect the success of the blending process (Févotte 2004, De Beer 2011, Jamrógiewicz 2012, Mantanus 2012).

6. CHEMOMETRIC

Chemometric is the important part of PAT that use in chemical data analysis. Chemometric is the statistical and mathemathical methods application, especially multivariate analysis, to manage process or chemical information. The need of chemometric methods is massive amounts of data by modern measuring devices (Rajalahti and Kvalheim 2011). Chemometric trends to deal with data tables or matrices consisting of several variables and measurement targets as a whole, rather than as the variations of single variables. The multivariate analysis is able to find out the latent variables or correlated variables information in the extracted original data from matrix. (Gendrin and Roggo 2008, El-Gindy and Hadad 2012).

There are many definitions about the meaning of chemometric. Its application is the methods for solving the problems regarding to statistical and mathematical methods. (Menezes and Ferreira 2009, Zhang and Chen 2010, Jamrógiewicz 2012,).

Chemometricians have developed methods from other study fields, for example linear partial least squares (PLS) and multi-way methods, respectively. Nowadays, the pharmaceutical industries have generally applied the chemometric methods. (Menezes and Ferreira 2009, Samad 2013).

The use of chemometric was met to three specific needs. First, during the acquisition of NIRs spectra, uncontrolled spectral variations can occur. Many factors could be responsible for these variations: change of optical path length between the samples, moving particles effect during in-line process monitoring, probe displacement. The impact of such uncontrolled parameters on the NIRs spectra can be significant as they can hide the relevant spectral features. Therefore, signal pretreatment is used to reduce their impact. Regarding the investigated property of interest, attention must be paid on the selection of the signal pretreatment (Sacré and De Bleye 2014). One would not want to select a signal pretreatment that reduces the effect of the property of interest in the NIRs spectra. Second, classification methods were developed to compress and extract the information contained in the spectra to find out the relation between the samples of the data set. Third, for quantification purpose, algorithms for multivariate regression were programmed to link the spectral data to a numerical value of the property of interest (Rajalahti and Kvalheim 2011).

6.1 Signal pretreatment examples

6.1.1 Multiplicative Scatter Correction (MSC)

Multiplicative Scatter Correction (MSC) is a well-known signal pretreatment that decreases the result of scattered light on diffuse reflectance spectra. On a calibration set for example, MSC performs as follows: first a reference spectrum is selected. Most of the time of the reference spectrum consists in the calibration set of the mean spectrum. All the calibration set spectra are then regressed linearly against the reference spectrum. Further, the slope and intercept are calculated. Finally, the spectra are corrected according to the slope and intercept. The extracted

information then represents the information that cannot be modeled: the relevant chemical information (Bakeev 2010).

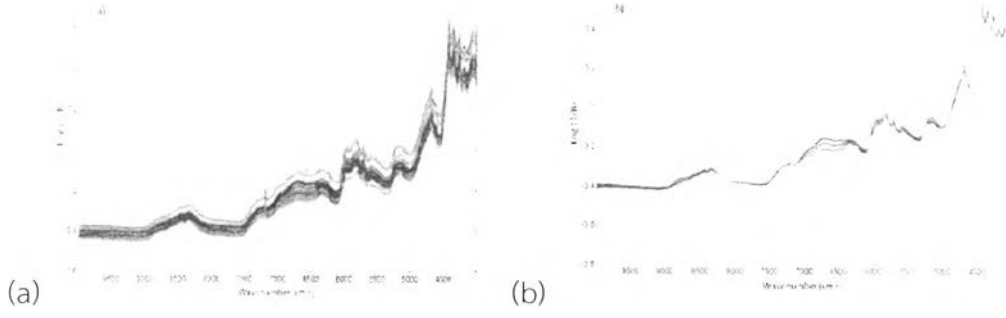


Figure 12 Spectra of moving pharmaceutical pellets with different API contents: comparison between NIRs spectra (a) before and (b) after MSC signal pretreatment (Mantanus 2012).

Figure 12 shows the spectra of moving pharmaceutical pellets with three different API contents before and after MSC signal pretreatment. One can see that MSC highlights the chemical information discriminating the samples: the three API levels of the spectral data set.

6.1.2 Standard Normale Variate (SNV)

Standard Normale Variate (SNV) corrects uncontrolled spectral variations. On a sample spectrum, it performs as follows: each variable absorbance is first corrected by the mean absorbance of the spectrum. The previous result is then divided by the standard deviation of the variables under investigation. Figure 13 shows the spectra of moving pharmaceutical pellets with three different API contents before and after SNV signal pretreatment. In the same way as MSC did, SNV allows to highlight the chemical information related to the three API levels (Bakeev 2010, El-Gindy and Hadad 2012).

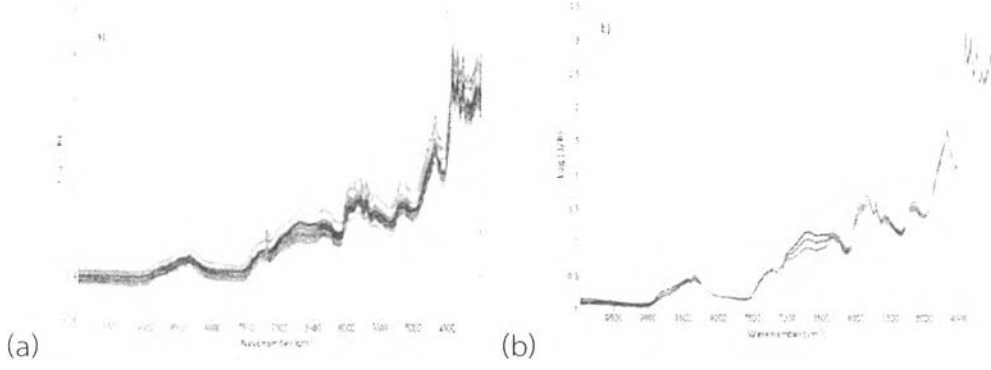


Figure 13 Spectra of moving pharmaceutical pellets with different API contents: comparison between NIRs spectra (a) before and (b) after SNV signal pretreatment (Mantanus 2012).

6.1.3 Smoothing and derivative signal pretreatment

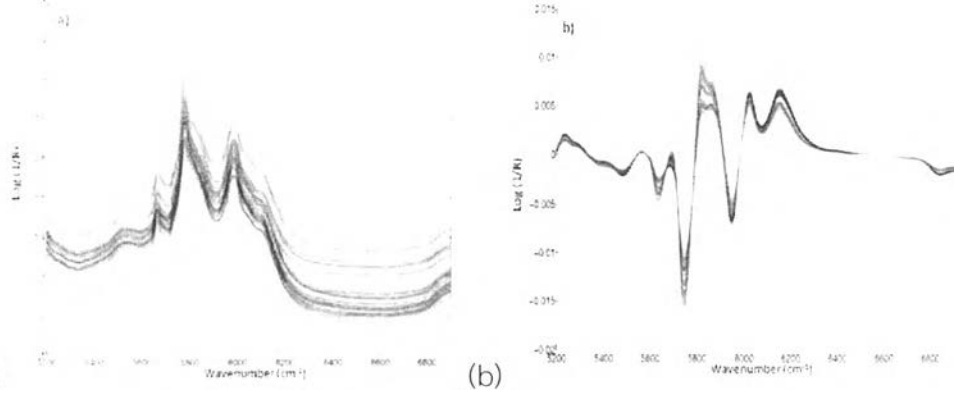


Figure 14 Comparison between NIRs spectra: Caffeine granules in polyethylene bag (red) and caffeine powder in polyethylene bag (green) (a) before and (b) after first derivative signal pretreatment. The Savitsky-Golay smoothing filter is applied before the first derivative signal pretreatment. The smoothing filter polynomial order is 2 and the filter width is 25 corresponding to 96.25 cm^{-1} (Mantanus 2012).

Smoothing aims to reduce the spectral noise, taking into account that spectral noise can hide the relevant physical or chemical information (El-Gindy and Hadad 2012). The most used smoothing technique is the Savitzky-Golay smoothing filter which basically performs as follows: the spectrum is divided into different spectral windows which all have the same width size (Savitzky and Golay 1964). A

polynomial regression is then performed on each window allowing to smooth the data. Parameters such as the window size and the polynomial degree can be set in the algorithm knowing that the larger the window size and the lower the polynomial degree will be, the stronger the smoothing effect will be (Rajalahti and Kvalheim 2011, Jamrógiewicz 2012).

Derivative signal pretreatment can be used to reduce baseline shifts between the spectra in order to extract physical spectral information. Besides, it can also improve the resolution between overlapping absorption bands. However, derivative transformation is highly sensitive to spectral noise which could hide the relevant information. Consequently, the Savitzky-Golay smoothing filter can be performed prior to derivative signal pretreatment (Bakeev 2010, Sacré and De Bleye 2014). Figure 14 displays the comparison between the NIRs spectra of caffeine granules and caffeine powder before and after the signal pretreatment (Savitzky-Golay [polynomial order: 2 and filter width 25] followed by 1st derivative). From the Figure 14 (a) and 14 (b), it can be seen that the signal pretreatment highlighted the physical difference between the granules and powder at 6180 cm^{-1}.

6.2 Classification method example

Principal Component Analysis (PCA)

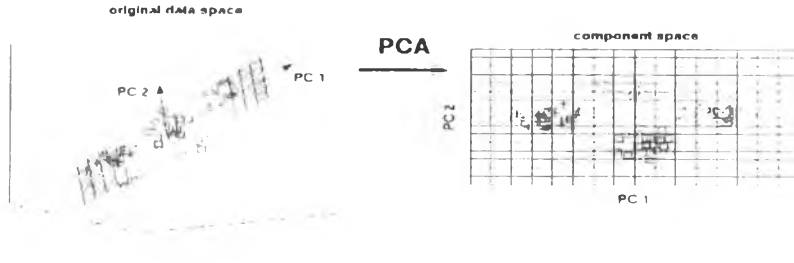


Figure 15 Geometric principle of PCA transformation.

(http://www.nlpca.org/fig_pca_principal_component_analysis.png [cited 2013 June 10])

Figure 15 geometrically explains the principle of PCA which can be easily understood as follows: given a series of spectra recorded with absorbance values corresponding to three wavelengths: λ_1, λ_2 and λ_3, PCA will reduce this 3-dimensional space into the component space, a 2-dimensional space that maximizes the variance between the samples, easing the understanding of the major trends in the data. Even though these major trends in the data are not so difficult to understand when considering data point in a 3-dimensional space, it becomes obvious that PCA is truly necessary when dealing with data sets having absorbance values for more than 2000 different wavelengths, which is usually the case with NIRs spectra (El-Gindy and Hadad 2012, Rahman 2012).

PCA aims to reduce plenty of observed variables to a small number of latent variables. The latent variables are called factors or principal components (PC). These factors are linear combinations of the original ones meaning that they are constituted of the linear sum of the weighted original variables (Märk, Andre 2010).

The statistical and mathematical method uses an orthogonal transformation: the PC_1 accounts for most of data sets with the variability and each succeeding one account itself for most of the remaining variability. Moreover, each PC is supposed to be orthogonal to the previous one. The loading describes the proportion of the factor with regard to the total variance of the matrix. Each object of the original data set is transformed using the principal components as factor space. The scores can be considered as the coordinates in the multidimensional factor space (Gendrin, Roggo 2008, Li 2014).

PCA is one of the mostly used classification method for data exploration. As it describes only the major trends within X without maximizing the covariance between

X and reference measurements, e.g. the active content of each sample row, this classification method is unsupervised. Other Classification methods (supervised or unsupervised) include: Linear Discriminant Analysis (LDA), Soft Independent Modeling of Class Analogy (SIMCA), Partial Least Squares Discriminant Analysis (PLS-DA), K-Nearest Neighbors (KNN) (Wu, White 2011, Bakeev 2010, Chen 2011).

6.3 Multivariate regression method example

Partial Least Squares (PLS) regression

The objective of Partial Least Squares (PLS) regression is to combine some reference measurements with their corresponding spectral data in order to create a model that will further only need spectral data to predict some reference measurements. Consequently, two data blocks are needed to perform PLS regression: the y block which contains the reference method results and the x block containing the corresponding data set (Zhang and Chen 2010, Martínez and Peinado 2013). Accordingly, contrary to PCA, PLS is a supervised method. Reference measurements represent the analytical results obtained with the conventional analytical method, in the case of an API content determination method (Castellanos Gil 2012), the reference method could be a HPLC method for instance (Blanco and Alcala 2006, Vanarase 2010, Martínez 2013).

PLS regression is related to PCA. The principal components are not chosen in the way that they account for as much as possible of the variance in the data sets, but to explain as much as possible of the observed variables (De Beer, Burggraeve 2011). Thus the structure of the y-data is not only used for the regression, but also for determination of the principal components. A major limitation of this technique is the probability of correlations. Spectral variations are presented in the calibration

spectra that may be correlated to the targeted analyte without being caused by features of the analyte (Bakeev 2010).

Facing the fact that many variables, many wavelengths, of the NIRs spectra are highly collinear, PLS regression selects a subset of a new variables orthogonal to each other, PLS factors which capture the greatest amount of variance in x while at the same time maximize the covariance between x and y. (El-Gindy and Hadad 2012).

PLS is the one of the mostly used multivariate regression method, especially in the PAT field. Other multivariate regression methods include Multi-Linear Regression (MLR), Principal Component Regression (PCR), Artificial Neural Networks (ANN) (Cogdill 2005, Gendrin and Roggo 2008, Bakeev 2010).

Test set validation is an external validation approach. Two separate sample sets are required for calibration and validation. A calibration model is computed based on one sample set which is then used to predict the values of the other sample set. The Root Mean Square Error of Prediction (RMSEP) is appropriate to assess the validity of the calibration model (Cogdill 2005, Ferreira 2013). Some caution is required when performing a test set validation: the test set has to be representative for the collective data set composed of calibration and test data. Hence, test set validation is not suitable for assessment of trends if data is classified by time. The calculation speed for test set validation is higher than for cross validation especially if high numbers of spectra are covered (Karande and Heng 2010, Vanarase and Alcalà 2010, Koller 2011, Momose and Imai 2011).