CHEMOMETRIC ATTENUATED TOTAL REFLECTANCE FOURIER TRANSFORM INFRARED SPECTROSCOPY METHOD FOR ROUTINE SCREENING OF PARACETAMOL, IBUPROFEN, AND ASPIRIN ADULTERATION IN HERBAL MEDICINES

Mr. Mario Theodore

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Pharmaceutical Sciences and Technology
FACULTY OF PHARMACEUTICAL SCIENCES
Chulalongkorn University
Academic Year 2022

การใช้เคโมเมทริกซ์จาก attenuated total reflectance fourier transform infrared spectroscopy เพื่อตรวจคัดกรองยาสมุนไพรที่เจือปนด้วยยาพาราเซตามอล ไอบูโพรเฟน และ แอสไพริน

นายมาริโอ ธีโอดอร์

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเภสัชศาสตร์และเทคโนโลยี ไม่สังกัดภาควิชา/เทียบเท่า
คณะเภสัชศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2565

Thesis Title	CHEMOMETRIC ATTENUATED TOTAL REFLECTANCE FOURIER TRANSFORM INFRARED SPECTROSCOPY METHOD FOR ROUTINE SCREENING OF PARACETAMOL, IBUPROFEN, AND ASPIRIN ADULTERATION IN HERBAL MEDICINES
By	Mr. Mario Theodore
Field of Study	Pharmaceutical Sciences and Technology
Thesis Advisor	Associate Professor VORASIT VONGSUTILERS, Ph.D.

Accepted by the FACULTY OF PHARMACEUTICAL SCIENCES, Chulalongkorn University in Partial Fulfillment of the Requirement for the Master of Science

..... Dean of the FACULTY OF PHARMACEUTICAL SCIENCES
(Professor PORNANONG ARAMWIT, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Associate Professor BODIN TUESUWAN, Ph.D.)

..... Thesis Advisor
(Associate Professor VORASIT VONGSUTILERS, Ph.D.)

..... Examiner
(Assistant Professor CHAISAK CHANSRINIYOM, Ph.D.)

..... Examiner
(PATANACHAI LIMPIKIRATI, Ph.D.)

..... External Examiner
(Associate Professor Uthai Sotanaphun, Ph.D.)

มาริโอ ธีโอดอร์ : การใช้เคโมเมทริกซ์จาก attenuated total reflectance fourier transform infrared spectroscopy เพื่อตรวจคัดกรองยาสมุนไพรที่เจือปนด้วยยาพาราเซตามอล ไอบูโพรเฟน และแอสไพริน. (CHEMOMETRIC ATTENUATED TOTAL REFLECTANCE FOURIER TRANSFORM INFRARED SPECTROSCOPY METHOD FOR ROUTINE SCREENING OF PARACETAMOL, IBUPROFEN, AND ASPIRIN ADULTERATION IN HERBAL MEDICINES)
อ.ที่ปรึกษาหลัก : รศ. ภก. ดร.วรสิทธิ์ วงศ์สุทธิเลิศ

การวิเคราะห์หาสิ่งปลอมปนยาในผลิตภัณฑ์สมุนไพรจำเป็นต้องมีเครื่องมือวิเคราะห์ที่ง่าย และเชื่อถือได้ เช่น High performance liquid chromatography (HPLC) แต่อาจมีข้อจำกัดในเรื่องของโลจิสติกและกระบวนการทำงานที่ซับซ้อน โดยเฉพาะในพื้นที่ที่ไม่สามารถเข้าถึงวิธีการวิเคราะห์นี้ได้ สำหรับ Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) เป็นเครื่องมือควบคุมคุณภาพที่สามารถใช้พิสูจน์เอกลักษณ์ด้วยาและมีการใช้งานที่ง่าย การใช้ ATR-FTIR ร่วมกับ Linear Discriminant Analysis (LDA) chemometrics จะช่วยเพิ่มความจำเพาะของเครื่องมือในการวิเคราะห์ส่วนประกอบต่าง ๆ ของตัวอย่างยาสมุนไพร การสร้างวิธีวิเคราะห์ให้เป็นโมเดล สำหรับวิธีการทดสอบปกติ โดยการใช้ ATR-FTIR ร่วมกับ LDA เปรียบเทียบกับวิธีวิเคราะห์ด้วย HPLC พบว่าการวัดประสิทธิภาพของโมเดล มีค่า %correct classification เท่ากับ 99.2% และผลการสอบเทียบของชุดตัวอย่างแบบจำลอง มีค่า cross-validation เท่ากับ 95.8% การทดสอบเปรียบเทียบวิธีวิเคราะห์ ATR-FTIR ร่วมกับ LDA พบว่าชุดตัวอย่างแบบจำลอง จำนวน 16 ตัวอย่าง มีการแสดงผลที่ถูกต้อง 100% เมื่อเทียบกับวิธีวิเคราะห์ HPLC นอกจากนี้ ผลการทำนายของตัวอย่างทดสอบมีความสอดคล้องเมื่อเทียบกับผลของชุดตัวอย่างแบบจำลอง ซึ่งในการศึกษานี้พบว่ามีความสามารถในการจำแนกมากกว่า 90% และสามารถใช้เป็นเครื่องมือคัดกรองการปลอมปนของยา paracetamol ibuprofen และ aspirin ในผลิตภัณฑ์สมุนไพรได้

สาขาวิชา เภสัชศาสตร์และเทคโนโลยี ลายมือชื่อนิสิต .....
ปีการศึกษา 2565 ลายมือชื่อ อ.ที่ปรึกษาหลัก .....

6470038133 : MAJOR PHARMACEUTICAL SCIENCES AND TECHNOLOGY
KEYWOR D: chemometric, spectroscopy, ATR-FTIR, adulteration, Herbal medicines

Mario Theodore : CHEMOMETRIC ATTENUATED TOTAL REFLECTANCE FOURIER TRANSFORM INFRARED SPECTROSCOPY METHOD FOR ROUTINE SCREENING OF PARACETAMOL, IBUPROFEN, AND ASPIRIN ADULTERATION IN HERBAL MEDICINES. Advisor: Assoc. Prof. VORASIT VONGSUTILERS, Ph.D.

Simple and reliable screening tools for herbal medicine drug adulteration are required, especially in rural areas where complex instrumental methods like HPLC face challenges such as logistics and complex operational processes. Attenuated Total Reflectance Fourier Transform Infrared (ATR-FTIR) is a good identification tool in drug quality control, and its operation is much simpler compared to HPLC, as it requires less sample preparation and has a simple operation. To improve its selectivity in analyzing multicomponent samples like herbal medicine, ATR-FTIR could be combined with chemometric methods such as Linear Discriminant Analysis (LDA). In order to establish itself as a routine testing method, ATR-FTIR combined with LDA will be compared against the currently validated HPLC method. The model's overall % correct classification results are 99.2% with cross-validation at 95.8% in the training set samples. Comparative testing shows 100% correct results (16 out of 16) in classification compared to the HPLC method. The results of the testing set samples' predictions are consistent with the cross-validation of the training set samples. Against the HPLC method, the model shows over 90% correct classification and could become a good candidate for routine screening tools for detecting adulteration of herbal products with paracetamol, ibuprofen, and aspirin.

Field of Study:	Pharmaceutical Sciences and Technology	Student's Signature .....
Academic Year:	2022	Advisor's Signature .....

ACKNOWLEDGEMENTS

I greatly appreciate Chulalongkorn University's Faculty of Pharmaceutical Sciences and its Scholarship Funding Program for ASEAN and Non-ASEAN Countries. Additionally, I would like to express my gratitude to the members of the Food and Pharmaceutical Chemistry department, especially Associate Professor Vorasit Vongsutilers, Ph.D., for their invaluable assistance in completing the research and enhancing my understanding. I would also like to thank my lab partner, Chanbopha Tho, who has been a wonderful knowledge-sharing companion throughout the course and research. Lastly, I am deeply thankful to the Indonesia National Agency for Drug and Food Control for providing me with the opportunity to pursue a Master's degree and enhance my abilities

Mario Theodore

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATION

2D-COS	=	2-Dimensional Correlation spectroscopy
ASP	=	Aspirin
ATR	=	Attenuated Total Reflectance
AUC	=	Area Under Curve
ESI	=	Electron Spray Ionization
FPR	=	False Positive Rate
FTIR	=	Fourier Transform Infrared Spectroscopy
GC	=	Gas Chromatography
HPLC	=	High Performance Liquid Chromatography
IBU	=	Ibuprofen
LDA	=	Linear Discriminant Analysis
mg/ml	=	Miligram per mililiter
LDA	=	Linear Discriminant Analysis
MS	=	Mass Spectroscopy
NMR	=	Nuclear Magnetic Resonance spectroscopy

PCA	=	Principle Component Analysis
PCT	=	Paracetamol
PLS	=	Partial Least Square
QTOF	=	Quadrupole Time of Flight
RMS	=	Root Mean Square
ROC	=	Receiver Operating Characteristic
RSD	=	Relative Standard Deviation
S/N	=	Signal to Noise Ratio
SERS	=	Surface Enhanced Raman Spectroscopy
SPE	=	Solid Phase Extraction
TLC	=	Thin Layer Chromatography
TPR	=	True Positive Rate
WHO	=	World Health Organization
WT	=	Wooden Tip
w/w	=	Weight per weight

CHAPTER I INTRODUCTION

1.1 Introduction, Background Rationale and Significance

In recent days utilization of herbal medicine in therapy has been increasing since the discovery of medicinal plants. WHO reported in 2018, that 98 countries of 194 WHO member have established herbal medicine regulation, indicating the growth of herbal medicines usage globally (1). WHO even released selected medicinal plants monograph that some of the medicinal plants are presented in Table 1 (2):

Table 1 List of some of herbal medicinal plants listed in WHO selected medicinal plants monograph

No	Herbal Plants	Parts of the plants	Pharmacology activity	Major Constituent
1.	*Pimpinella anisum l.*	Dried fruit	Analgesic, antimicrobial, anticonvulsant	linalool (0.1–1.5%), methylchavicol (estragole, isoanethole; 0.5–6.0%), αterpineol (0.1–1.5%), cis-anethole (< 0.5%), trans-anethole (84–93%), panisaldehyde (0.1–3.5%)
2.	*Humulus lupus l.*	Dried strobile of female plants	Antimicrobial, anti-oedema, antioxidant	humulone and lupulone and their related derivatives, 2–10% and 2–6%, respectively
3.	*Passiflora incarnata l.*	Dried herbs	Analgesic-antipyretic, anti-inflammatory, antimicrobial	Flavonoids up to 2.5%
4.	*Rehmania glutinosa* var. *purpurea* makino	Dried roots	Antidiarrhea, antihepatotoxic, antimicrobial, antihyperglycemic,	Iridoid monoterpenes (2.6–4.8%)

No	Herbal Plants	Parts of the plants	Pharmacology activity	Major Constituent
			anti-inflammatory, antitumor, antiulcer	
5.	*Prunus armeniaca l.* var *ansu* maxim	Dried seeds	Analgesic-antipyretic, antitussive, antitumor	amygdalin (up to 4.9%), a cyanogenic glycoside (a plant compound that contains sugar and produces cyanide)
6.	*Crocus sativus l.*	Dried stigma	Anti-arteriosclerotic, anticoagulant, proliferation inhibition	essential oils (0.4–1.3%) with α- and β-pinene, 1,8-cineole (eucalyptol), a monoterpene glucoside, picrocrocin (4%), safranalm and carotenoid glucosides known as crocins (2%)
7.	*Foeniculum vulgare* mill	Dried fruit	Analgesic-antipyretic, antimicrobial, antispasmodic	essential oil (2–6%) containing *trans*-anethole (50–82%), (+)-fenchone (6–27%), estragole (methylchavicol) (3–20%), limonene (2–13%), p-anisaldehyde (6–27%), α-pinene (1–5%) and α-phellandrene (0.1–19.8%)

Herbal medicines are reported with some advantages compared to conventional pharmaceutical drugs such as having low adverse effect and relatively safe for use in the long term. Herbal medicines have been through modernization, although maintaining plant extract or fraction as their active ingredient. For example, Indonesia

classified herbal medicine into 3 categories based on its formulation and clinical trials. The lowest class is “jamu” which is formulated using plant extract or dried plant parts of common medicinal plants and safety assessment based is on empirical data. In the middle class of Indonesia herbal medicines is called “Obat Herbal Terstandar” which means standardized herbal medicines. The ingredients of herbal medicine product must be standardized according to regulation and the active ingredients is herbal extract or fraction. Standardized herbal drugs had through the pre-clinical testing for its safety. Top of Indonesia herbal medicine is “fitofarmaka” where the ingredient is standardized and efficacy of the product proven by clinical tests (2).

However, herbal medicines popularity is also followed by the crime practice that looms over it, through counterfeiting or the addition of prohibited substances. Herbal medicines counterfeiting conducted by substituted herbal crude raw material with other similar plants. The lowest risk of substitution could happen is the herbal medicine has no therapeutic effects. However, if the substitute herbal raw material contains a toxic substance that is not found in original ingredients it endangers consumers. The other way is adulterating the herbal medicine with a synthetic drug in order to enhance its therapeutic potency. Adding a synthetic drug could trigger various health issue.

In 2021, Indonesia’s National Drug and Food Control reported 53 herbal medicines product contain synthetic drug substance. Adulterated herbal medicines with a synthetic drug substance have a potential health issue in long-term condition such as kidney failure, hepatic damage etc. As explained in Calahan J. et. al in 2016 review article of Chemical Adulterants in Herbal Medicinal Products (3), the adulteration of herbal medicine product with synthetic potent drugs is one of the drug regulatory agencies main surveillance task, for they pose serious health risks. There are few reports of toxicity cases from adulterated herbal medicines with synthetic dug adulteration with common drugs for cases like sleep aids (clonazepam), weight loss (sibutramine and fenfluramine) diabetes (glibenclamide), and bodybuilding (steroids) products. Few reports of synthetic drug addition for example, Fenfluramine has now been banned by the FDA in the U.S. and in Hong Kong after reported induce pulmonary hypertension and valvular disease. The problems with sibutramine were

also reported in Japan and Taiwan. Sibutramine, is a drug that has been intentionally added to slimming products also reported banned in US and Europe since its potential cardiovascular risks or even strokes.

Analgesic herbal medicines are currently the most popular among Indonesia community. Generally, herbal plan used for analgesic indication or pain management is zingiber family plant. For example, *Zingiber officinale* has been used in India and China as muscle pain and swelling, arthritis, headaches, digestive and appetite problems, prevention of motion sickness, postoperative nausea and vomiting, hyperemesis gravidarum, and also cold and bacterial infections. Another herbal plant used for joint pain treatment is *Curcuma xantorrhiza* root with its germacrone. Inflammatory activity also show by ethyl acetate extract of *Elephantopus scaber* by inhibition of p38.activating Chan et. al, 2017 (4).

The addition of a synthetic drug such NSAID in analgesic herbal medicine commonly happens in first places. Malaysian National Pharmaceutical Regulatory reported in 2019 (5), that 3461 products of pain and fever indication herbal medicine were seized, it is 54% of all adulterated product seized from market (Fig.1). Paracetamol, Ibuprofen, and Aspirin are over-the-counter and cheap drug to buy in store or pharmacy, an easy access to use it as adulterant in herbal medicine. Excessive use of NSAID could trigger various health issues. As several report long-term adverse effect of paracetamol which could damage liver have potency to cirrhosis. While Ibuprofen also harm for gastro intestinal track and the other hand aspirin will enhance bleeding and had similar impact to gastro intestinal track like ibuprofen.

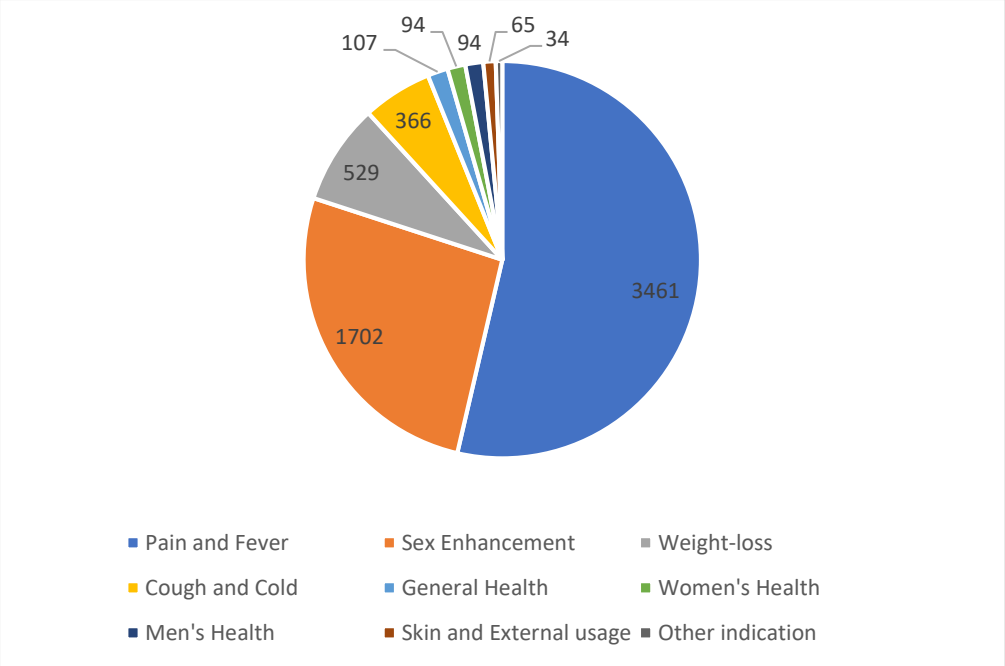


Figure 1 Adulterated Traditional Herbal Medicines Seized by Malaysian National Pharmaceutical Regulatory in 2018

Herbal medicines are containing various substances which requires special preparation to analyze the target substances. This situation also giving many difficulties in detection of adulterants or synthetic drug presence in herbal medicine samples.

Chromatography will give the best solution to separate the analyte with samples matrix or other substances contained. It is selective and accurate analytical method for multiple component analysis such herbal medicine. Despite its great capability, chromatography is time and reagent consuming analytical method.

Chromatography's main method is the separation of analyte by its interaction with the mobile phase and stationary phase. The mobile phase of chromatography could be a liquid or gas phase, while the stationary phase is commonly solid or liquid coated on a solid support. For example, simple chromatography techniques such as TLC (Thin Layer Chromatography) uses liquid solvent as the mobile phase and silica as the stationary phase. In order to get accurate and valid results, all reagents, including those used in the mobile phase of analysis, should be analytical grade reagents. The same way goes to HPLC, where HPLC grade reagent is more expensive.

Another consideration of chromatography technique is the complexity of the instruments. For advanced instruments like HPLC and GC, which require several accessories for their operational HPLC requires pump pressure to flow the mobile phase, whereas GC requires a high temperature oven. Chromatography's great accuracy is followed by a lot of preparation and logistics behind its execution. Pratiwi et al., 2017 summarize some analytical methods for undeclared synthetic drugs in herbal medicine in Table 2. below (6).

Table 2 Current instrumental based analytical method for synthetic drug presence in Traditional Herbal Medicine collected by Pratiwi et. Al (6)

.

No	Analyte	Samples	Method
1	Sildenafil	Herbal capsules	TLC-SERS
2.	Sibutramine	Herbal slimming tea	TLC-Densito
3.	Sildenafil, tadalafil, and vardenafil hydrochloride	Herbal sexual enhancer product	HPLC-MS-MS
4.	Caffeine, piroxicam, chlorpheniramine, betamethasone, oxethazaine	Herbal capsules	UPLC-QTOF-MS
5.	Fenfluramine, phenolphthalein, bumetanide, and sibutramine	Slimming supplement	SPE-UPLC-MS/MS
6.	Amitriptyline, acetaminophen, ibuprofen, chlorzoxazone, sulfamethoxazole, tadalafil, and sildenafil	Traditional Chinese medicines and food supplement	GC-MS
7.	Dexamethasone	Herbal joint pain	IR-Partial Least Square
8.	Sildenafil	Herbal sexual enhancer	FTIR-Stepwise Multiple Linear Regression

No	Analyte	Samples	Method
9.	Sibutramine, phenolphthalein	Herbal slimming product	Low-field H-NMR
10.	Rutin, quercetin, kaempferol	Ginkgo biloba capsules	FTIR-PLS-DA
11.	Fenfluramine, phenolphthalein, bumetanide, and sibutramine	Slimming supplement	Gtip SPE-UPLC-MS/MS
12.	Ephedrine, pseudo-ephedrine	Slimming herbal product	2D-COS
13.	Melatonin, doxepin, diazepam, chlorpheniramine, zopiclone, nitrazepam, zaleplon, alprazolam, clonazepam and chlordiazepoxide	Herbal dietary supplements (pills, Tablets, capsules, or soft-gel capsules)	WT-ESI-MS
14.	Paracetamol, naproxen, sulfamethoxazole, diclofenac, and phenylbutazone	Tablet and capsules (anti rheumatism health care products)	ESI-MS

In rural or remote islands, chromatography faces challenges of logistics and maintenance, where transportation could be one of the major obstacles. Several simple detection methods, such as the color test and dipstick, have already been developed in order to overcome this challenge. Philp et al. (2018) reviewed some spot tests used in the field by law enforcement, specifically their specificity and selectivity (7). The chemical spot test offers a fast, simple, and reliable test. Its portability makes it possible to conduct real-time presumptive tests (Fig 2).

The general concept of a chemical color spot test is the interaction between an analyte and reagents in a test that triggers color changes. An electron transfer is commonly

used to produce colored metal complexes and charged organic species. It also considers the pH of the reaction that gives an effect on the color and intensity of the product tested. The common reagent tests are shown in the Figure below.

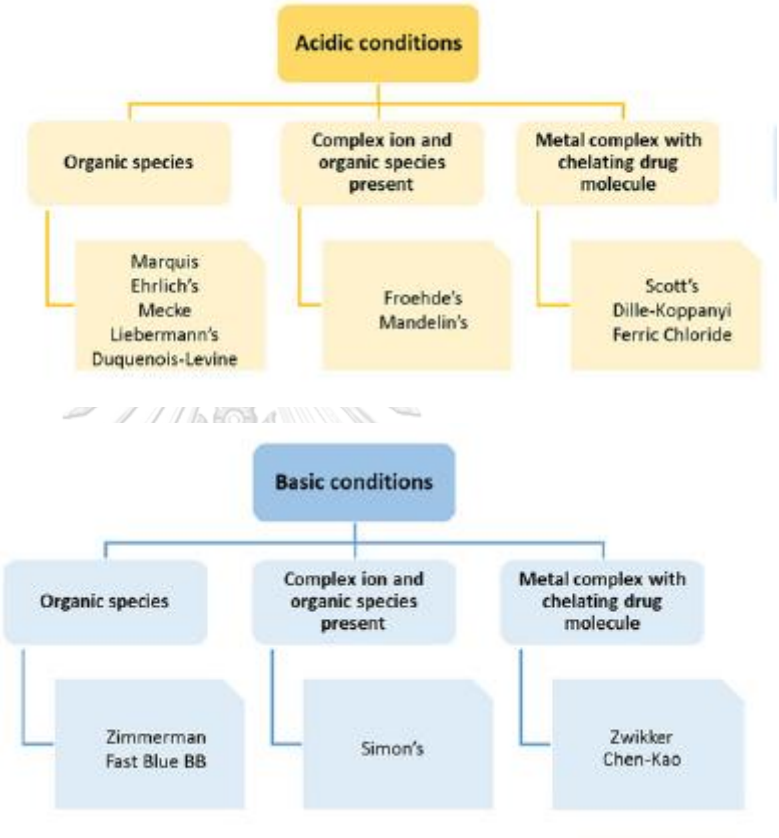


Figure 2 Common reagents on color spot test based on pH solution test

(8)

Even with its benefits for simple and reliable presumptive test there's still some drawback of chemical color spot test. False positive results could occur due to some different chemical substances analyte could give similar chemical reaction to reagent test. Chemical reagent test handling is also something that need to be concern to maintain in its reliability.

There are gaps between laboratory analysis test accuracy with on-field rapid testing technique. In order to overcome this gap, simple and rapid instrument based analytical method is required for adulteration detection. Instrument based offering better accuracy and selectivity compared to field chemical color spot test.

Spectroscopy compared to chromatography offers better portability despite no separation for mixture samples. However, several techniques have been developed to solve such issues, like derivative spectroscopy and chemometric combination. Pratiwi reviews that some spectroscopy techniques like FTIR combined with Partial Least Square-Discriminant Analysis could be the tools for herbal medicine authentication. Since chemometric could enhance spectroscopy data by processing it through a statistical function. developing validated spectroscopy method with chemometrics for rapid screening of adulteration could be the answer.

1.2 Literature Review

1.2.1. Fourier Transformed Infrared (FTIR) and Attenuated Total Reflectance (ATR)

Infrared Spectroscopy is an analytical technique that observes the chemical interaction of infrared light that is absorbed by chemical molecules. Interaction between infrared light and chemical substances occurs when the molecular structure of chemical substances has a dipole-moment covalent bond. The vibration of molecular structures bends or stretches bonds depending on the type of molecular bond, and it provides information about the functional groups contained in the molecule.

There are two most common types of infrared spectrometers used in drug quality control; they are the Dispersive Infrared Spectrometer and the Fourier Transform Infrared Spectrometer. The Dispersive Infrared Spectrometer works by scanning the frequency transmitting through a samples and a reference. Each frequency that passes through the sample is measured individually by the detector, which consequently slows the process of scanning the entire IR region (8). The Fourier Transform Infrared Spectrometer, on the other hand, introduced the Michelson interferometer, which can combine multiple beams of IR sources transmitting through a sample. Then, by the fourier transform function, the collected interferogram is converted into spectrum data. This gives the Fourier Transform Infrared

Spectrometer a faster measurement tool compared to the dispersed IR spectrophotometer.

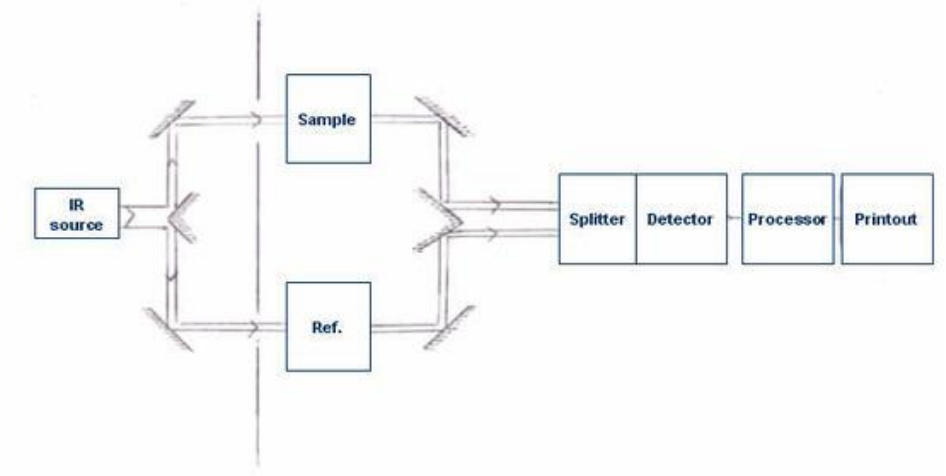
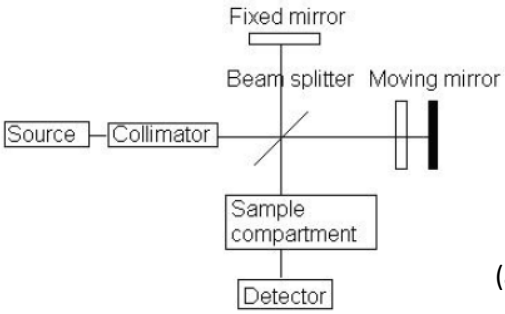


Figure 3 Schematic diagram of Dispersive-IR
(8)



(a)

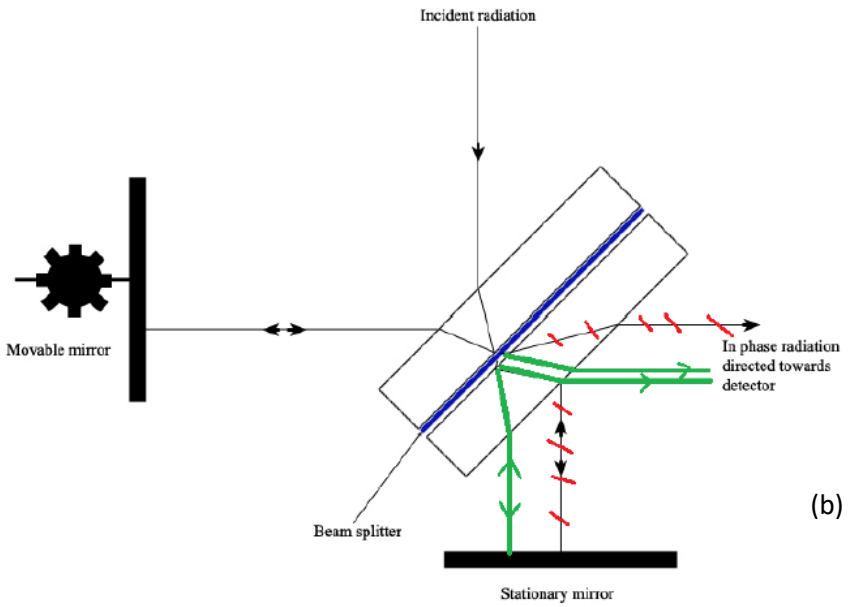


Figure 4 (a) FTIR schematic diagram; (b) Michelson interferometer (8)

Common FTIR technique probes used in analysis are Transmission and Reflectance. The transmission probe works by passing IR light through samples to the detector. On the other hand, the reflectance probe works with a reflectance IR (ATR) light source to sample through the reflectance medium to the detector.

Transmission IR requires KBr as a window, which makes it need a sample preparation for sample scanning. KBr requirement since it's not an IR active substance and can be mixed with samples into solid form without giving any interference. Another consideration of transmission is that the samples must be translucent enough in order for IR light to pass through them. On the other hand, ATR doesn’t need it since ATR could use air as a background and window (9).

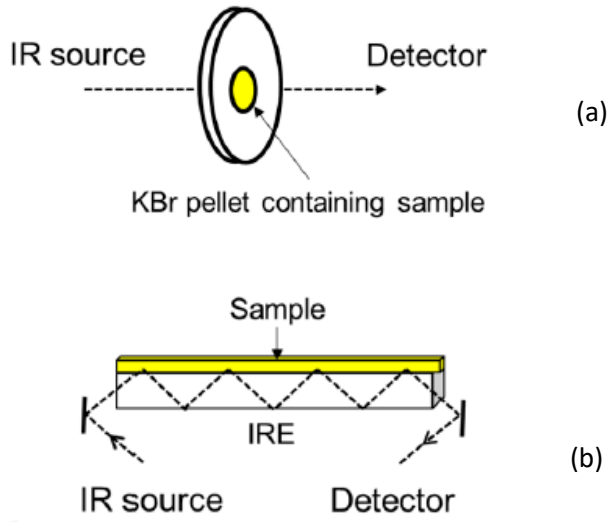


Figure 5 (a) Transmission FTIR; (b) Reflectance FTIR (ATR) (9)

1.2.2 FTIR Spectroscopy as Authentication Tools and Derivative Spectroscopy Technique

Spectroscopy offers a good technique in the identification of pharmaceutical products. It is fast and reliable compared to color test identification but lacks separation like chromatography. Despite its inability to separate, spectroscopy fingerprints can be used to identify an analyte in qualitative analysis.

Every organic substance has unique fingerprinting spectra shown in mid IR at 1500 to 400 cm^{-1}. This fingerprint spectra giving capability of IR in authentication or identification test. Even though in mixtures, we still can pick pointing some characteristic spectra of each substances in fingerprint region as marker of identification. For substances whose that chemical structures have been identified, functional group region could also play big role in differentiate each substance (Table 1). When using FTIR combined with chemometrics technique.

However peak intensity of analyte could be overshadowed by matrix samples peak. This condition could lead into miscalculation when we are going to use

intensity value to process the data. In order to overcome such issues, some researcher employing derivative spectroscopy.

Derivative spectroscopy is technique utilizing mathematic function to derived spectrum function. If spectrum is express in absorbance A as function of wavelength $\boldsymbol{\lambda}$, the derivative spectra will be calculated this way (10):

$$\text{Zero order} \quad : A = f(\boldsymbol{\lambda})$$

$$\text{First order} \quad : \frac{dA}{d\lambda} = f'(\lambda)$$

$$\text{Second order} \quad : \frac{d^2A}{d\lambda^2} = f''(\lambda)$$

A derivative spectrum could be obtained by such methods like optical, electronic, or mathematical. Commonly, mathematical methods are utilized since it's easier for mathematical methods to calculate or recalculate with different parameters or smoothing techniques. Derivative spectrum calculation is done by computer with the output of a new derivatized spectrum.

Derivative spectroscopy offering advantages of background elimination that makes the baseline shift, smoothing data point (utilizing savitzky-golay method), and increasing discrimination of mixtures samples. Increasing discrimination is obtained when derivatized spectrum give a different amplitude. Example of derivation in UV-Spectroscopy presented below.

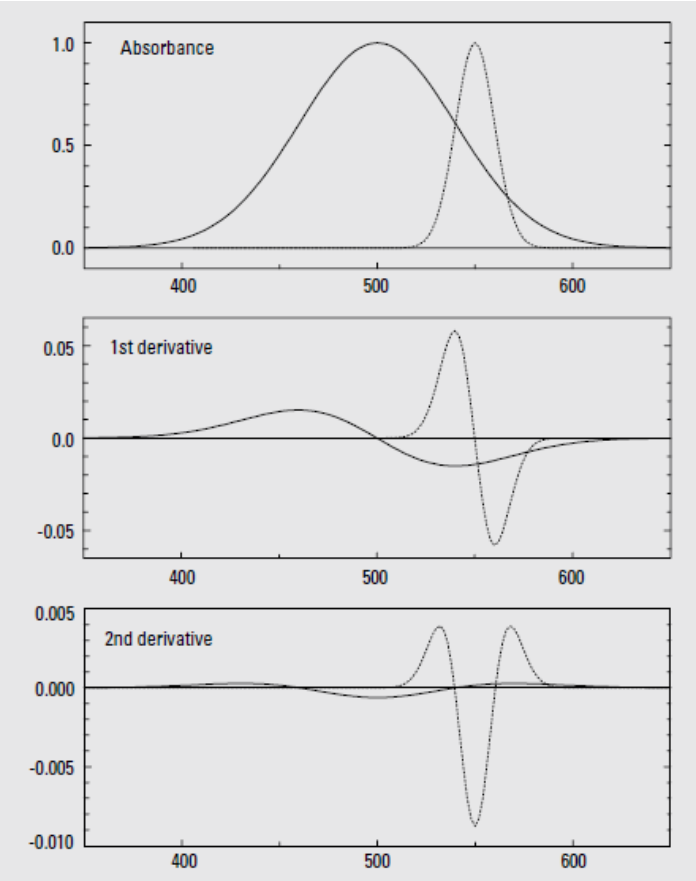


Figure 6 UV spectrum examples of derivative spectrum (10)

1.2.3 Chemometrics for Pattern Recognition in Spectroscopy

Chemometrics is multidisciplinary study where mathematics, statistic, and logic tools are used in chemical experiments (11). Chemometrics utilizes other field knowledge such as statistic, computer science, algebra, and recently chemometrics is also fused with machine learning. Main purpose of chemometrics is to extract meaningful information from data then process it into mathematical model in order to enhance quality of the data or assisting in decision making.

In this topic we are discussing chemometrics in pattern recognition and how it processes spectroscopy data (FTIR) for screening adulterated herbal medicines. The main concept of chemometrics for pattern recognition's is generating a mathematical model by a statistical algorithm that is able to

separate a group of samples. The mathematical model generated by chemometric for pattern recognition employs multivariate analysis of variance (MANOVA) since spectroscopy data contains a lot of variables. This mathematical model plays a big role in spectroscopy analysis, where the spectroscopy technique lacks separation ability compared to chromatography. Pattern recognition techniques consist of unsupervised and supervised techniques. The difference between the two is that prior classification is used in the supervised technique, which provides superior classification power over the unsupervised technique. As presented in Figure 7, according to USP chemometric modelling, qualitative analysis of authentication is included in the classification section (12).

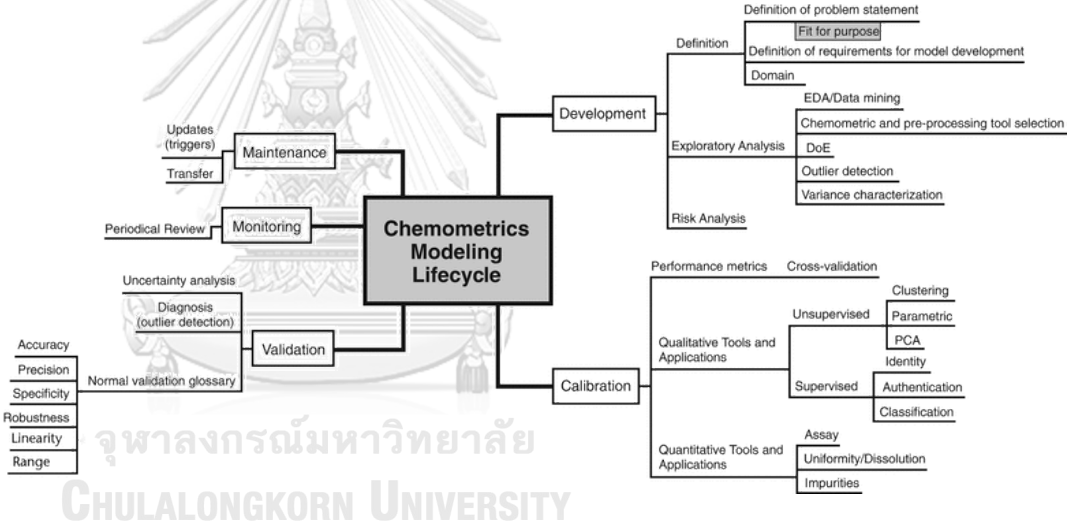


Figure 7 Chemometrics Modeling Lifecycle scheme
(12)

Spectroscopy data contains a large number of variables, where it is hard to point out which variable has a significant impact. For example, we have 10 samples of mixed herbal medicine capsules that come from different brands (D and E) with only IR spectrum data of two herbal medicines provided (see Figure 8). With eyes, both spectrums seem similar and it is difficult to distinguish which sample belongs to D or E. In order to give better classification between D and E, we could employ a statistical approach by treating the peak intensity of a certain wavenumber as an independent

variable and the dependent variable as a group of samples D and E, as shown in the Figure below.

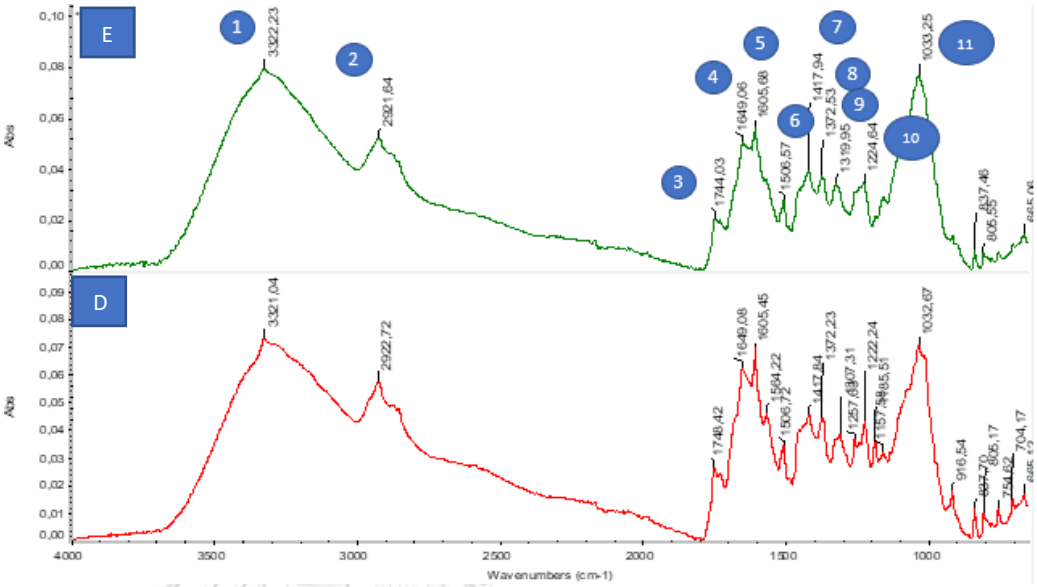


Figure 8 Ilustration of assigned FTIR spectrum as variable in PCA

We could point out some characteristic peaks of the sample (1 to 10), and each peak is an independent variable that will undergo a statistical classification test. The dependent variable in this statistical classification test is the group of samples D and E. However, a lot number of variables could lead to collinearity problem. Where the correlation between variables occurs, that could lead to biased classification results to distinguish sample D or E. An illustration of the transformation from spectrum measurement into a data matrix Table could be seen in Figure 9.

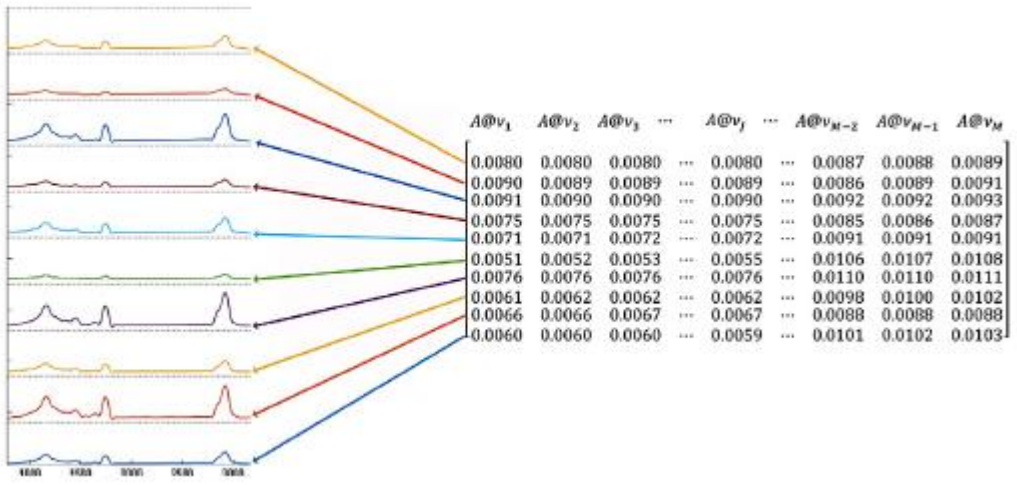


Figure 9 graphical illustration spectra measured in IR transformed into data matrix

Observation of collinearity between variables could be done by drawing a graph of each variable's correlation, but it is not efficient and consume a lot of times. We could draw a 10 dimensions variable at the same time to observe it, but this approach method is not possible since we are only able to draw a 3 dimensions graph.

Multivariate analysis in chemometrics could assist pointing a significant spectral data by variable compression method such PCA, PLS, LDA, etc. This traditional statistic technique is relatively easy to operate compared to advance machine learning or deep learning algorithm which requires computer science background. Using spectroscopy as a tool for classification it is important to determine what significant variable that differentiate samples in the group. A brief explanation of how PCA, PLS, and LDA reduces variables then classifying samples is presented in next few paragraphs.

Principal Component Analysis (PCA) is unsupervised data (sample in this case) classification technique utilizing variable compression by orthogonal projection of individual variables into a new dimensional subspace. The PCA algorithm of variable compression is used to flatten the data that comes from a lot of dimensions into smaller dimensional data that is meaningful to

differentiate groups of data. which could tackle collinearity issues and make it easier to observe data distribution when we present it in a graph.

The PCA mechanism of dimensional or variable reduction is by drawing a new line between all the data distributions that has the most variance, then projecting all the data into the line. It's not only drawing one line but several lines that cover most of all the data distribution. A line that is used to project the data becomes a new dimensional data that is called the principal component (PC). PC1 will always be the line that covers the most variation of data, followed by PC2, the second most variation of data, and so on. The schematic process of how PCA generates new dimensional data is presented in Figure 10.

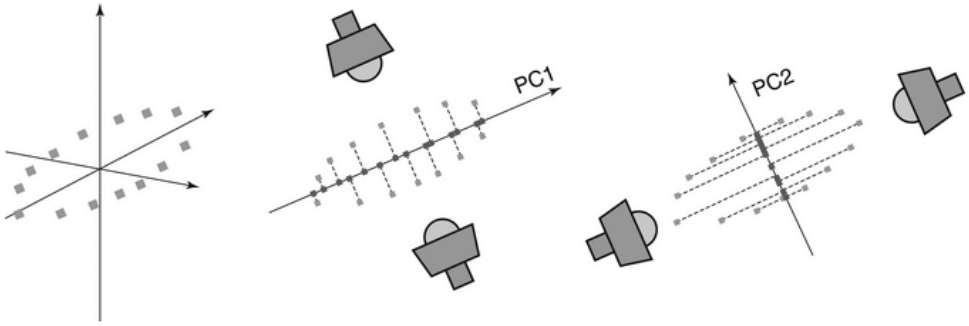


Figure 10 PCA projection scheme algorithm

(12)

After we form a new dimensional graph from principal components, we are going to plot our data (samples) by first calculating the principal component score (PC score) of each data point. The PC score will become the new coordinate of the data in the new PC graph. Data with similar or close PC score will be grouped together in the graph, and we will have visual data of the data distribution in this case, samples. The PC score is calculated by the formula below :

$$PCx\ score = \sum_{n=1}^{x} \pm nm$$

n is the data value and m is the loading value of the variable in the PC plot. Loading is the value of the coefficients, or weights, of the linear combination in each compositive variable. Loading signs and magnitude indicate the magnitude and direction of corresponding variables. Generally, PCA calculation of dimensional reduction follows this formula:

$$X = TP^T + E$$

Where T are PC score, P are loading values of variables, and E are residuals. This calculation is generated by computer

PCA output is a graph of data distribution and is followed by data classification results. An example of a PCA distribution graph of raw materials from 3 different suppliers is presented in Figure 11. Using PCA, we find a way to group samples better with spectroscopy measurement data.

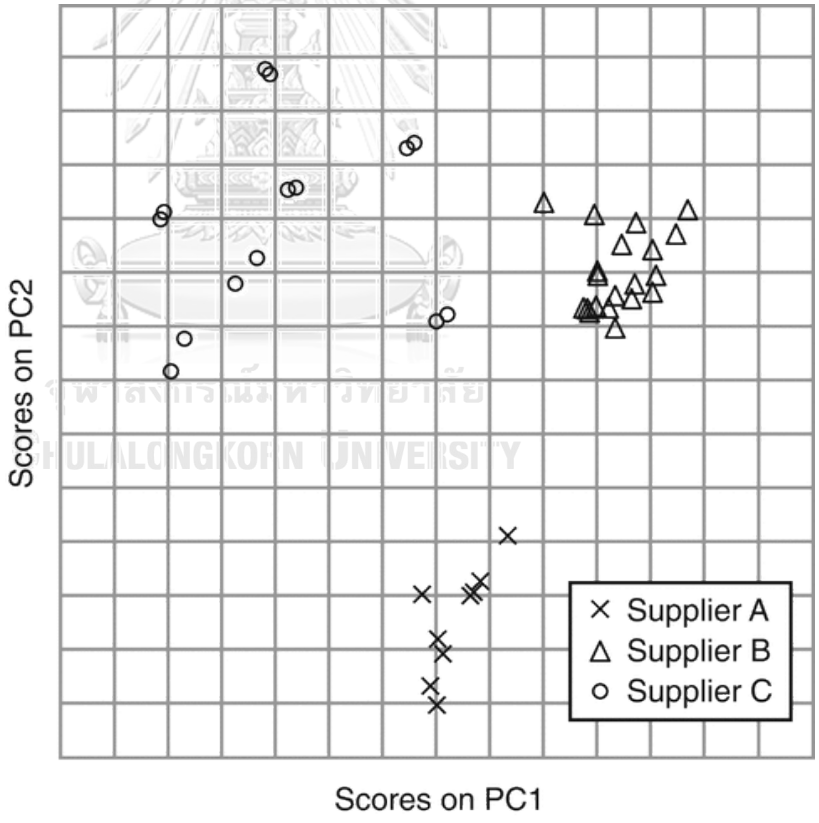


Figure 11 PCA graph examples of raw material authentication
(12)

Partial Least Squares (PLS) is another similar group classification that uses a variable compression method like PCA. PLS tries to find the maximum fit of the dependent variable (Y) by drawing a line of Y distribution in a graph and then projecting it into a new line that covers the most variation (Figure 12). The same is true for independent variable (X), except that it seeks an X distribution that best explains Y rather than the most variation of X. After projecting the data of Y and X, PLS creates a new dimensional graph where a line that covers the most variation of Y (Latent Vector 1 or LV1) and a line of X (Latent Vector 2 or LV2) that explains Y the best becomes the new axis. PLS is commonly used for a group of data that has several dependent variables (Y).

This is a different approach compared to PCA. By this way, PLS is trying to maximize the loading of LV. The maximum loading value earned also shows a strong correlation between the dependent variable and the independent variable. LV is similar to PC in PCA in that it generates a new dimensional graph by projecting data onto a new axis. Maximizing loading score will give a better prediction (classification) of the dependent variable and minimize the loss of data after variable compression. The PLS calculation formula is the same as PCA as shown in the formula below:

$$X = TP + E$$

However, T value on PLS is different with T in PCA calculation. In PLS T is matrix of empirical experiment data and P is regression coefficient of matrix. While E are also residuals.

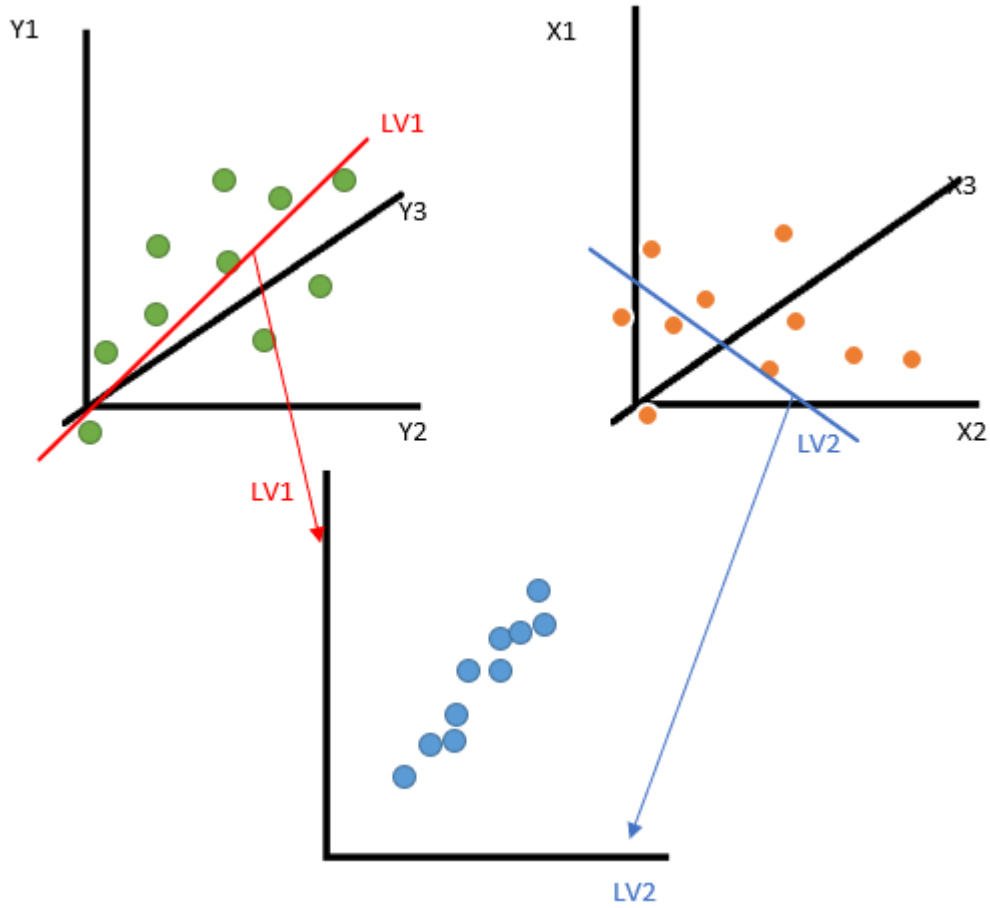


Figure 12 Scheme of PLS projection and variable compression

PLS commonly used in quantitative analysis for in spectroscopy analysis, for qualitative analysis commonly it combined with discriminant analysis (PLS-DA). PLS is also good for predicting dependent variable of more than one block set of data. Similar to PCA, PLS output also presenting distribution graph (Fig. 13) and classification results Table.

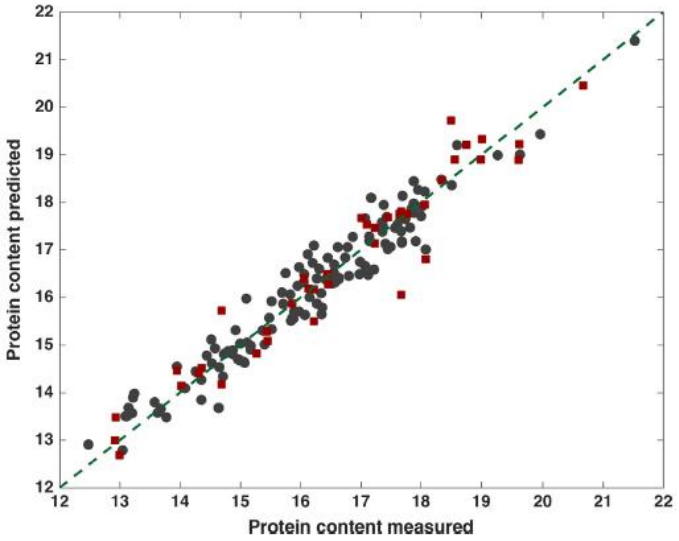
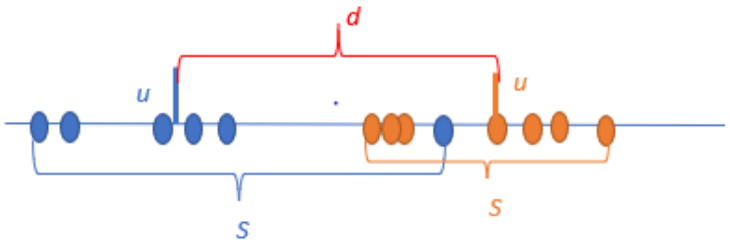


Figure 13 Examples of PLS regression analysis in protein content measurement (11)

Linear Discriminant Analysis (LDA) is a robust discrimination method that is commonly used in authentication tasks. However, LDA assigns data into certain groups by maximizing the ratio between-group variance and within-group variance. Maximum ratio reduces the possibility of samples being scattered and groups being separated from one another. It works in a similar way to PCA by projecting the data into the new axis, then calculating the maximum distance and separation of the data in order to get the best separation, an algorithm where PCA lacks. It measures the ratio of group means (u), distances (d) and scatter (s) of the data, which is depicted by the scheme below.



$$\frac{(u-u)^2}{s^2+s^2} = \frac{d^2}{s^2+s^2}$$

Figure 14 LDA separation scheme and calculation

The LDA involves prior discriminatory categories and the model allows users to determine significant differences between prior defined groups with variables that give significant means different across the group. Prior discriminatory also gives better selection for sample grouping compared to PCA. That's why LDA is included as a supervised multivariate analysis technique. LDA graph output examples are presented in the Figure below.

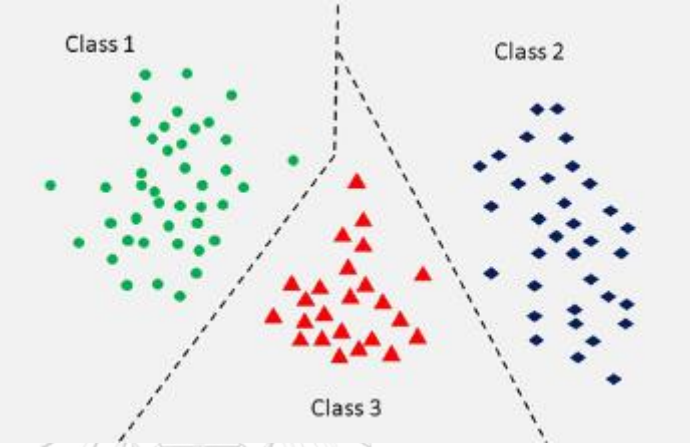


Figure 15 Discriminant Analysis Method for data grouping (13)

LDA operational relatively much simpler compared from others variable compression. Despite its only employ PCA algorithm for variable compression where PLS is more effective, however citing from Brereton in Chemometrics of Pattern Recognition 2009, LDA could have similar classification performance to PLS-DA if retaining all non-zero data (13).

In 2021, Dhamarastuti et al. able to classify the adulterated jamu (Indonesia Traditional Herbal Medicine) using multivariate analysis (14). Adulterated samples and unadulterated samples were classified in ternary mixtures utilizing PLS-DA with ATR-FTIR spectrum data. Adulterant analytes in the experiment were metamizole and prednisone adulteration in herbal pain reliever product. While Cebi et. al., 2017 also successfully discriminated sibutramine contaminated herbal slimming tea by ATR-FTIR technique coupled with hierarchical cluster analysis and PCA (15).

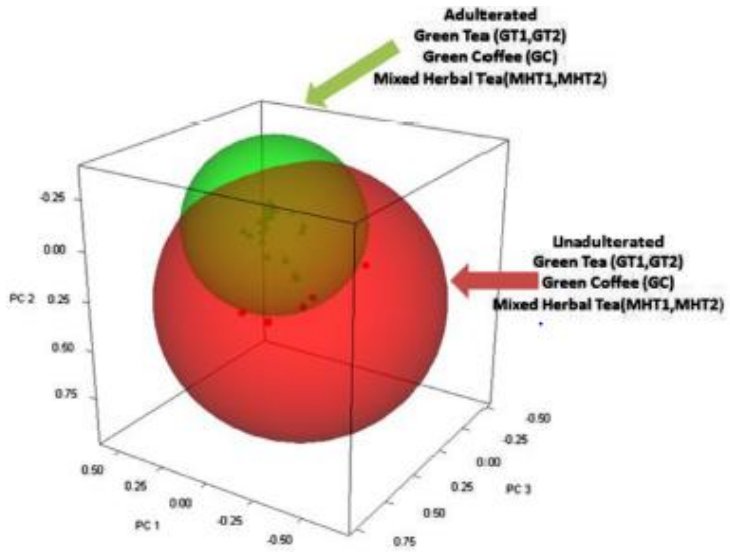


Figure 16 PCA Classification of adulterated Herbal Slimming Tea by Cebi et.al., 2017
(15)

Chemometrics data processing requires several tools, like statistical software. There is a lot of statistical software that is able to perform multivariate analysis in chemometrics. However, generally, scientific software uses coding to run its functions, making it only personnel with good computer coding able to operate it.

For example, Matlab is an open-source scientific software where users can add various tools that help them in scientific experiments. Despite its modularity and flexibility, Matlab requires coding operations in order to execute any programs that run on it. It requires well trained personnel to operate the software.

Routine analytical testing demands consistent results between personnel who run the analysis. Easiness of operation of analytical methods and other tools that support is required. IBM SPSS, compared to Matlab or a similar type of statistical software, is relatively easy to use. It is able to operate without coding operations and is almost similar to office software operations.

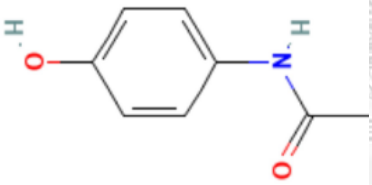
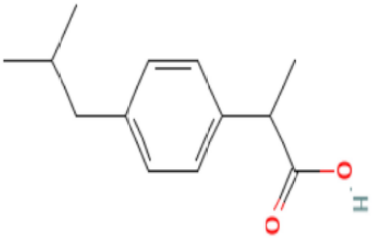
Multivariate analysis is also provided in SPSS, i.e., PCA, Cluster Analysis, and linear regression. However, SPSS is still severely limited in multivariate analysis applications, for example, it cannot executing PLS-DA simultaneously where Matlab is capable to. Even with its limitations, SPSS is still able to give reliable and simple multivariate analysis classification tests.

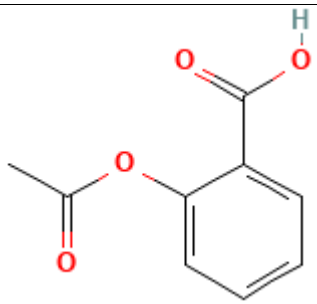
Utilizing ATR-FTIR coupled with chemometrics could be a good option for a reliable method of herbal medicine adulteration in remote areas. Since it is able to analyze samples with minimal preparation, where analysis logistic support challenges could be minimized.

1.3 Hypothesis

1.3.1 Adulterant analytes

Table 3 Spectral analysis box of Paracetamol, Ibuprofen, and Aspirin

Structures	Functional Group and Wavelength
 Paracetamol	• Amides N-H : 3300 cm^{-1} C=O : 1680 – 1630 cm^{-1} C-N : 1400 cm^{-1} • Phenol -OH : 3700 – 3584 cm^{-1} Free C-O : 1260 - 1000 cm^{-1} =C-H : 900 – 690 cm^{-1} C=C : 1600 – 1475 cm^{-1}
 Ibuprofen	• Carboxylic Acid C=O : 1730 – 1700 cm^{-1} C-O : 1320 – 1210 cm^{-1} -OH : 3400 – 2400 cm^{-1} • Aromatic Benzene Ring =C-H : 900 – 690 cm^{-1} C=C : 1600 – 1475 cm^{-1}

Structures	Functional Group and Wavelength
 Acetylsalicylic acid	• Ester & Carboxylic Acid C=O : 1730 – 1700 cm^{-1} C=O : 1750 – 1725 cm^{-1} (ester) C-O : 1320 – 1210 cm^{-1} -OH : 3400 – 2400 cm^{-1} • Aromatic Benzene Ring =C-H : 900 – 690 cm^{-1} C=C : 1600 – 1475 cm^{-1}

Three analytes have the similarity of containing benzene and carbonyl groups. However Paracetamol carbonyl group is amide which could gives a specific marker of peaks at 3300 cm^{-1} and 1400 cm^{-1}. While Aspirin and Ibuprofen have the same ester group, the only difference is that Aspirin's ester group has a peak absorption of C=O (ester) region close to carboxylic acid. The strong peaks of the Aspirin ester group will be the key point.

Based on the spectrum acquired in IR, we then selected the characteristic peak region of each analyte and measured its intensity. The wavenumber region is a variable we put in the LDA method as an independent variable, while the dependent variable we put in the method is a group of samples.

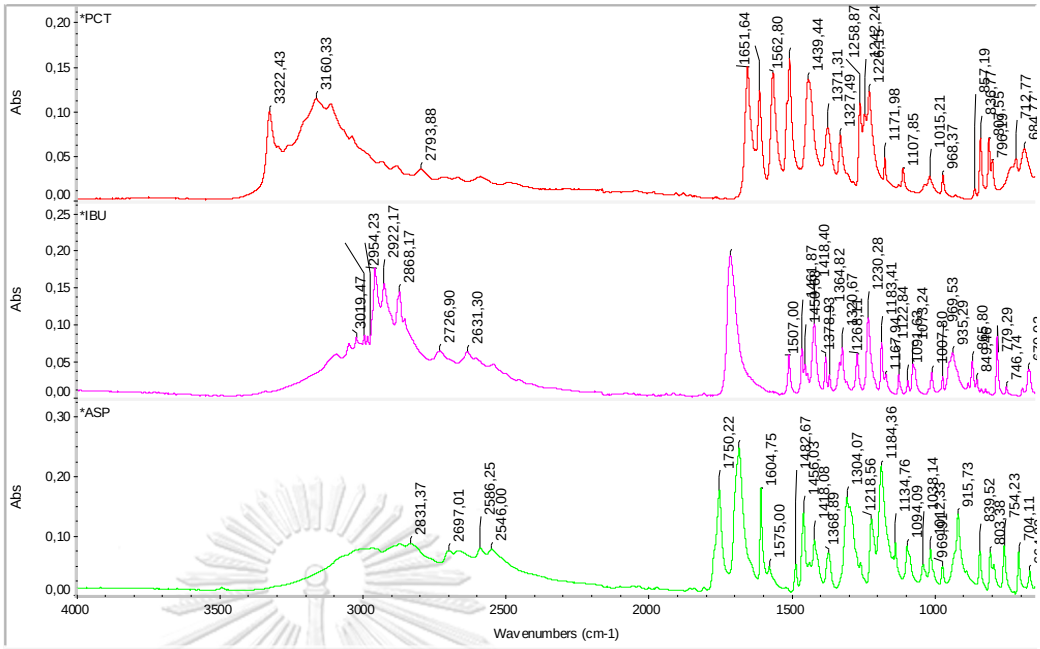


Figure 17 IR spectra of Paracetamol (PCT), Ibuprofen (IBU), and Aspirin (ASP) in absorbance mode

The characteristic peaks assigned before, as expected, show up when mixed with herbal medicine samples. Typical characteristic peaks are selected from functional group regions to fingerprint regions. The intensity of the characteristic peak of each analyte selected as a variable will be processed in multivariate analysis, in this case, LDA. The LDA model was then build using selected region characteristic peaks as predictors variable. The classification model will also present the distribution of data groups in a graph.

1.3.2 Expected Results

Collecting the wavenumber region of each analyte provides critical information for separation in chemometric statistical functions. The behavior of wavenumbers will then be checked in the mixture system of matrix samples. Regions that consistently appear and specifically refer to the adulterant will be marked and assigned as variable inputs. Hopefully, through statistical calculations, the model will be able to classify adulterated and unadulterated samples.

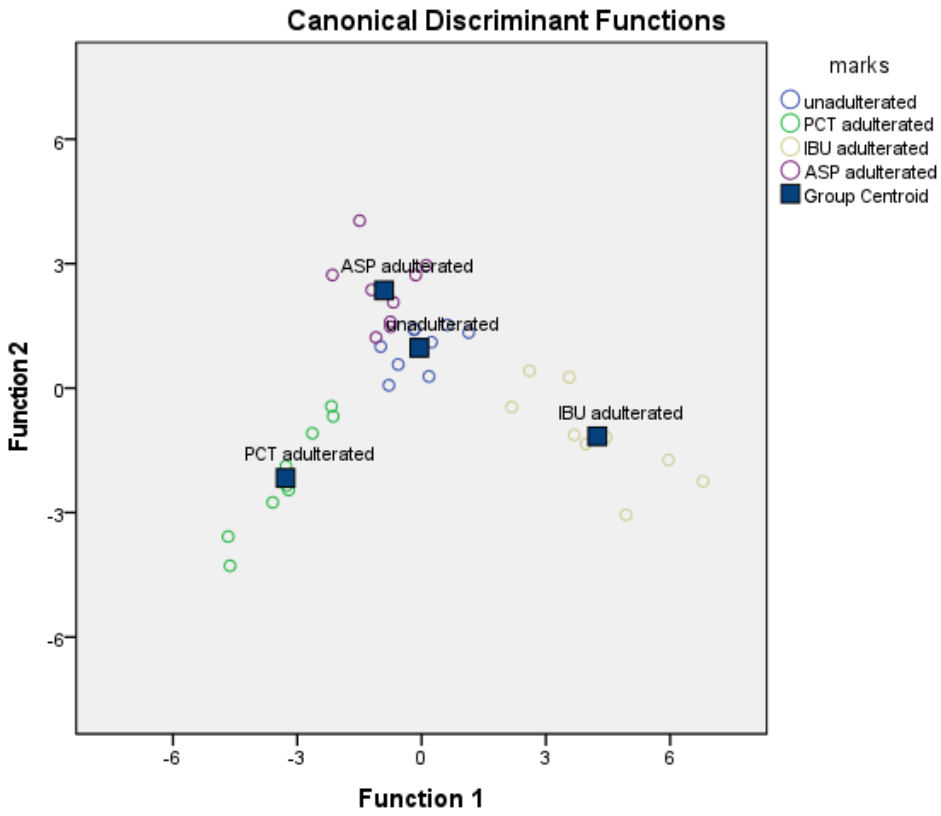


Figure 18 Expected results of ATR-FTIR Chemometrics Plot of Discriminant Analysis

Graph information :

- Unadulterated : Unadulterated samples
- PCT adulterated : Paracetamol adulterated samples
- ASP adulterated : Aspirin adulterated samples
- IBU adulterated : Ibuprofen adulterated samples

ATR-FTIR data acquired from experiment will be processed in SPSS discriminant analysis, where we expect that discriminant analysis will able to group each type of samples. The classification results shown in SPSS as Table below.

Table 4 Classification results of LDA classification model

Classification Results[a,c]

			Predicted Group Membership				
		marks	unadulterated	PCT adulterated	IBU adulterated	ASP adulterated	Total
Original	Count	unadulterated	9	0	0	0	9
		PCT adulterated	0	9	0	0	9
		IBU adulterated	0	0	9	0	9
		ASP adulterated	1	0	0	8	9
	%	unadulterated	100.0	.0	.0	.0	100.0
		PCT adulterated	.0	100.0	.0	.0	100.0
		IBU adulterated	.0	.0	100.0	.0	100.0
		ASP adulterated	11.1	.0	.0	88.9	100.0
Cross-validated[b]	Count	unadulterated	9	0	0	0	9
		PCT adulterated	0	9	0	0	9
		IBU adulterated	1	0	8	0	9
		ASP adulterated	1	0	0	8	9
	%	unadulterated	100.0	.0	.0	.0	100.0
		PCT adulterated	.0	100.0	.0	.0	100.0
		IBU adulterated	11.1	.0	88.9	.0	100.0
		ASP adulterated	11.1	.0	.0	88.9	100.0

a. 97,2% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 94,4% of cross-validated grouped cases correctly classified.

1.4 Research Objectives

Developing validated chemometric models with ATR-FTIR as screening tools for Paracetamol, Ibuprofen, and Aspirin adulteration in herbal medicines

1.5 Research Benefit

Cost effective, valid, and reliable screening tools for adulteration with comparable results to HPLC

CHAPTER II EXPERIMENTAL

2.1 Research Plan and Experiment Detail

2.1.2 Samples

1. 10 Thai herbal medicines
2. 10 Jamu
3. Paracetamol working standard (99.89% purity)
4. Ibuprofen working standard (99.60%)
5. Aspirin raw material (95.97% potency assay by DMSc working standard)

2.1.3 Equipment

1. Mortar and stamper
2. Pike Diamond Attenuated Total Reflectance (ATR) module.
3. FTIR Nicolet is50 by Thermo Scientific with OMNIC as software data acquisition
4. IBM SPSS 28 statistic software
5. Analytical balance

2.1.4 Sample preparation

A total of 24 samples will be spiked with mono-adulteration, binary adulteration, and ternary adulteration. The adulteration level for single spiking is 12%, for double combination spiking it is 12+12%, and for triple combination spiking it is 12+12+12% w/w. In total, there are 192 samples to process, with an additional 24 samples not spiked and considered as original or unadulterated samples.

The 12% adulteration level is based on the assumption that paracetamol, ibuprofen, and aspirin are present at a concentration of 50 mg per 400 mg in the average weight of herbal medicine dosage forms, with 48 mg expected to provide a pharmaceutical therapeutic effect. One common oral liquid dosage form for paracetamol is 48 mg/ml, and ibuprofen is available in a chewable dosage form of 50 mg. Aspirin is typically administered at a dose of 50 mg for ischemic stroke. The sample preparation design is shown in the table below..

Table 5 Sample treatment design

Samples	Adulteration spiking		
	Single (12%)	2 combination (12%+12%)	3 combination (12%+12%+12%)
A	PCT IBU ASP	PCT+IBU IBU+ASP ASP+PCT	PCT+IBU+ASP

2.1.5 Data Acquisition and signal pre-processing

ATR-FTIR instrumental settings

- Reflectance medium : Diamond
- Beam Splitter : KBr
- Correction method : ATR correction
- Number of scans : 32
- Resolution : 4 cm^{-1}
- Wavenumber scans : 4000-650 cm^{-1}
- Background : Air
- FTIR software : Thermo Scientific OMNIC
- Pressure clamp : 6 kg/m^2
- Spectral processing : auto-baseline; Savitzky-Golay Derivative

2.1.6 Variable selection

Characteristic peak of each analyte (adulterants) is marked and its intensity is recorded. The selected wave number will be processed on the classification model, then percent classification will be observed. Biggest percentage of classification from various combination is determined to select the best variable combination for classification model.

2.1.7 Classification model build

The multivariate analysis chosen in the experiment is Linear Discriminant Analysis (LDA), which is processed by IBM SPSS 28 (16), (Fig. 19). Variable selection used in SPSS is based on the selected wavenumber region in the spectral analysis box of the spectrum. The intensity of peak value acquired from the experiment will be processed in the LDA classification model. Overall, in much of the research, PLS-DA (Partial Least Squares-Discriminant Analysis) gives a better variable compression compared to PCA

where it is employed in LDA. According to Brereton, PLS-DA, which retains all non-zero data, produces classifiers that are similar to LDA. LDA operational in SPSS is relatively simple compared to PLS-DA, since it only has a one-step operational function.

Statistical assumption requirements in the LDA model are tested in SPSS by Univariate statistic and Box-m with Fisher as the coefficient function. The Univariate statistic test confirms equality of group means, while the Box-m test examines the homogeneity of data variance. While the Fisher coefficient function will determine the classification function value that is assigned to each case.

Features selection is also playing important role in tuning the model capability. Provided features in SPSS Discriminant Analysis is put all the variable together in other words selecting variable manually or utilizing stepwise method function.

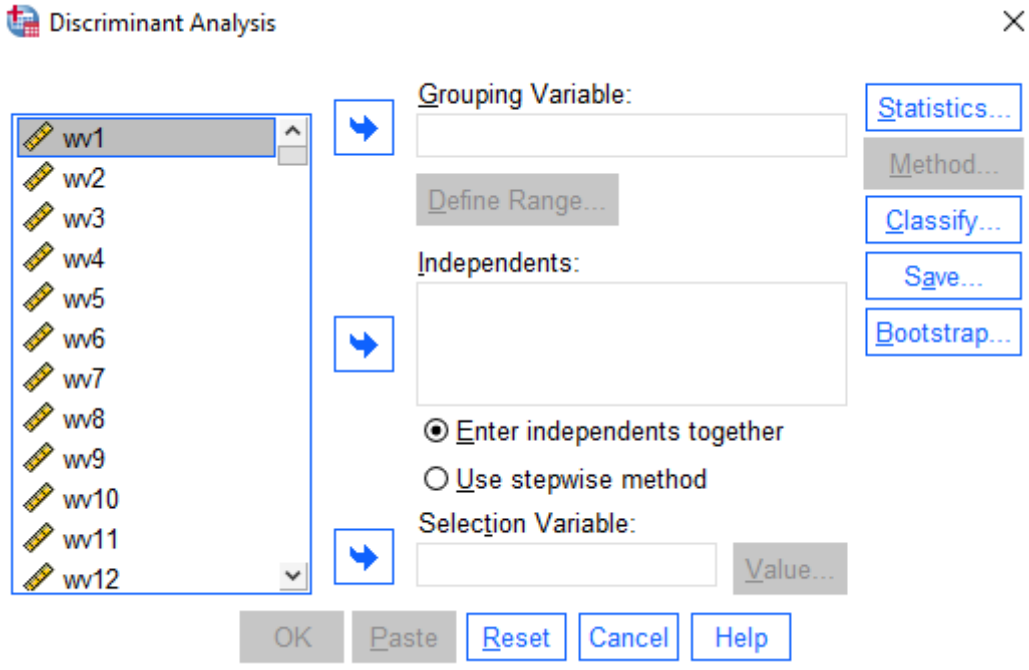


Figure 19 SPSS dialog box for Discriminant Analysis

The stepwise method is an automatic variable selection run by SPSS. A variable selection algorithm is employed in the stepwise method as listed in features selection. Utilizing features selection gives various results depending on how the features work to shape the model. Since it determines variables

entered into the classification model, different combinations of variables will give different classification results (Fig. 20). features selection in the SPSS Discriminant Analysis Stepwise Method, including:

- Method provided in features :
  1. Wilks Lambda

     A variable selection method for stepwise discriminant analysis that chooses variables for entry into the equation on the basis of how much they lower Wilks' lambda. At each step, the variable that minimizes the overall Wilks' lambda is entered.

  2. Unexplained Variance

     The variable that minimizes the sum of the unexplained variation between groups is entered.
  3. Mahalanobis Distance

     A measure of how much a case's values on the independent variables differ from the average of all cases. A large Mahalanobis distance identifies a case as having extreme values on one or more of the independent variables
  4. Smallest F-ratio

     Variable selection in stepwise analysis based on maximizing an F ratio computed from the Mahalanobis distance between groups.
  5. Rao-V

     A measure of the differences between group means. Also called the Lawley-Hotelling trace. At each step, the variable that maximizes the increase in Rao's V is entered.
- Criteria for variable selection:
  1. Use F Value

     A variable is entered into the model if its F value is greater than the Entry value and is removed if the F value is less than the Removal value. Entry must be greater than Removal, and both values must be positive. To enter more variables into the model, lower the Entry value. To remove more variables from the model, increase the Removal value.

2. Use Probability of F

   A variable is entered into the model if the significance level of its F value is less than the Entry value and is removed if the significance level is greater than the Removal value. Entry must be less than Removal, and both values must be positive. To enter more variables into the model, increase the Entry value. To remove more variables from the model, lower the Removal value.

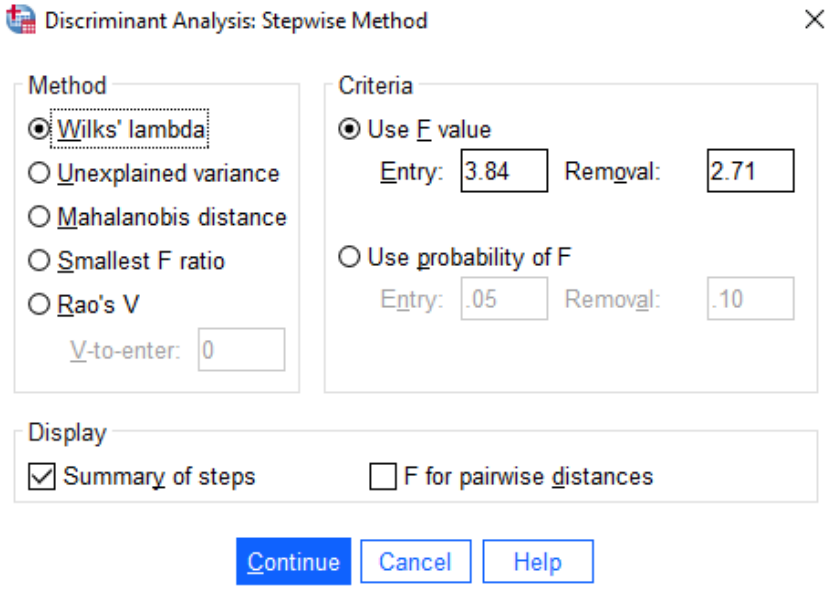


Figure 20 Stepwise selection option in SPSS Discriminant Analysis

A combined plot diagram is chosen in order to give visual information of classification distribution of samples. Leave-one-out cross-validation is also performed to simulate model prediction using training set data. It works where one of each data in the analysis taken and used as testing set while the remaining used to build the model. The correct classification percentage is observed with a target value of over 90% for overall classification and cross-validation.

2.1.8 Validation of Classification Model

In order to prevent overfitting of the classification model, certain validation steps are required. Overfitting is a condition where a model works well with a training set but has poor performance when tested with testing set data.

A validation test will be performed on an independent testing data set. The sample set will be divided into a training set and an independent testing set (13). Training set samples are a set of sample data used to construct the classification model. While the testing sample set is a set of samples to test the prediction capability of the classification model. Testing set samples are excluded from the training sample set in order to avoid bias prediction results. Following his recommendation, the training data set is 2/3 of total samples, while the rest is a testing data set. The sample set is split into the following schemes:

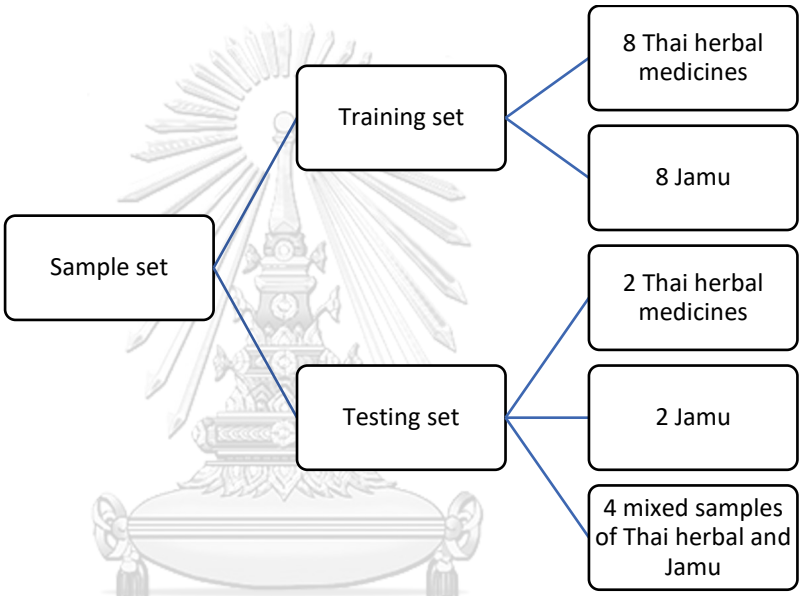


Figure 21 samples split set scheme

Acceptance criteria of prediction test is % classification of testing set is not less than %classification results of training set.

Prediction test is performed by SPSS scoring wizard. SPSS scoring wizard will showing prediction of sample class based on classification model that have been made before. Scoring wizard dialog box could be seen in Figure below.

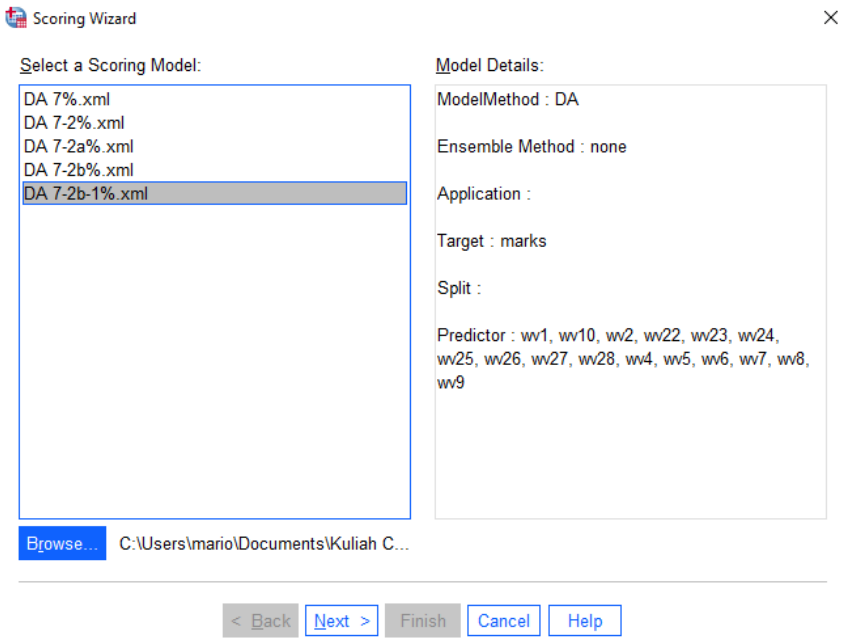


Figure 22 SPSS Scoring Wizard Function

Following USP general chapter 1225, "Validation of Compendial Procedure" (category IV), the specificity of the model is validated (17). Specificity and sensitivity tests by measuring model performance prediction true positive rate (sensitivity), and false positive rate (1-specificity). The calculation formula is presented below.

- True Positive Rate (TPR)

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

  Where TP = True Positive and FN = False Negative

- False Positive Rate (FPR)

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

  Where FP = False Positive and TN = True Negative

TPR values expected not less than 90% according to %classification results of model while FPR results expected not more than 10% following performance target of the model.

2.2 Comparative testing studies

Prediction results of the model will be challenge with HPLC (High Performance Liquid Chromatography) method in testing set samples. HPLC test conducted to detect or screening the presence of Paracetamol, Ibuprofen, and Aspirin in samples.

Proposed HPLC analytical method of Indonesia National Agency of Drug and Food Control in-house analytical method for Paracetamol and Aspirin detection, as shown below:

- Samples and Equipment:
  1. Paracetamol working standard
  2. Aspirin working standard
  3. Herbal medicines samples
  4. Sep-pak® Vac 3 cc (500mg) C18 Cartridges Solid Phase Extraction (SPE).
  5. Methanol and demineralized water
  6. O-phosphoric acid
  7. Potassium hydroxide
  8. Acetonitrile
  9. Vacuum pump
  10. Membrane filter 0.22 µm pore
- Solvent : Methanol 60% water
- Standard concentration (PCT) : 1.67 µg/ml
- Standard concentration (ASP) : 20 µg/ml
- Column : C18-250 x 4.6mm; 5µm
- Detector : Photo Diode Array 200-400 nm
- Flow rate : 1.0 ml/min
- Injection volume : 20 µl
- Solid Phase Extraction of samples and spiked samples procedure:
  1. 50 mg of fine powdered samples transferred into 5 ml volumetric flask. Then add 2 ml solvent (methanol 60%) and shake for 30 minutes, add solvent to volume.
  2. Prepared the SPE cartridges and conditioning the column using 1.5 ml methanol and water respectively. Note, do not let the SPE column dried.

3. Add 500 µl of samples solution on SPE cartridge then let it drip slowly (around 15 drops/minute) through SPE column. After all sample solution flows through SPE column, washed the sample in column using 1.5 ml 5% methanol solution. When all wash solution dripping out, eluate the column using 60% methanol and retain all solution that flows through the column

- Mobile Phase : Gradient elution. Phosphate buffer made

  From 6.64 o-phosphoric acid in 1 L water

  then adjust the pH to 3.74 ± 0.03 with

  Potassium Hydroxide 10%

Time (Min)	Composition %	
	Acetonitrile	Phosphate Buffer
0.01	15	85
10	25	75
11.25	60	40
12.50	40	60
32.50	50	50
35.00	15	85
39.99	15	85
40.00	Stop	

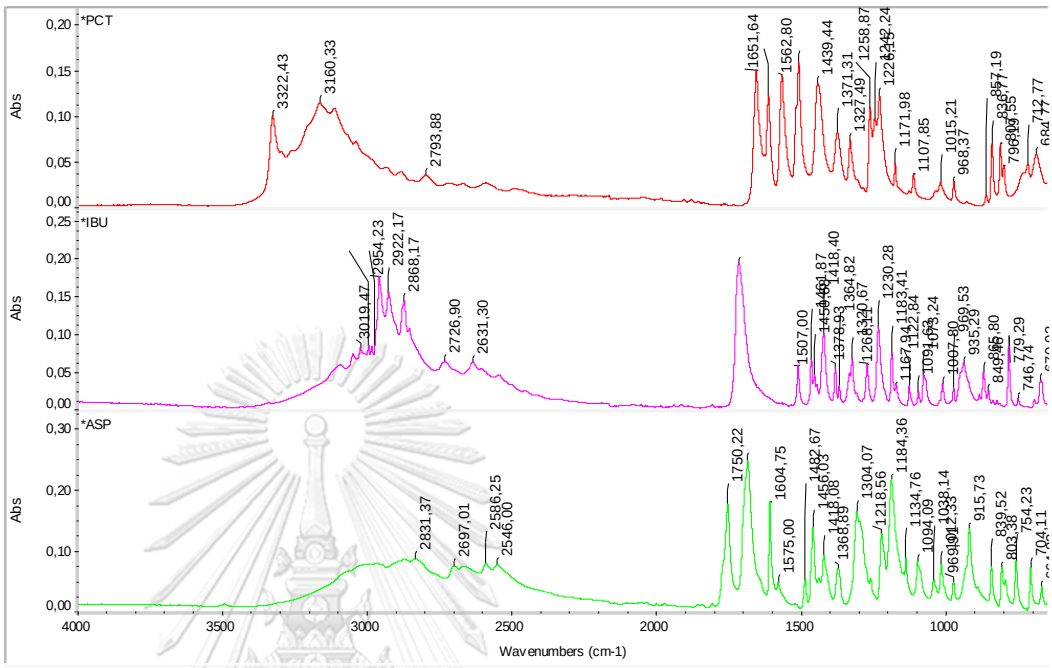
2.3 Preliminary studies results

- Spectrum analysis and variable selection

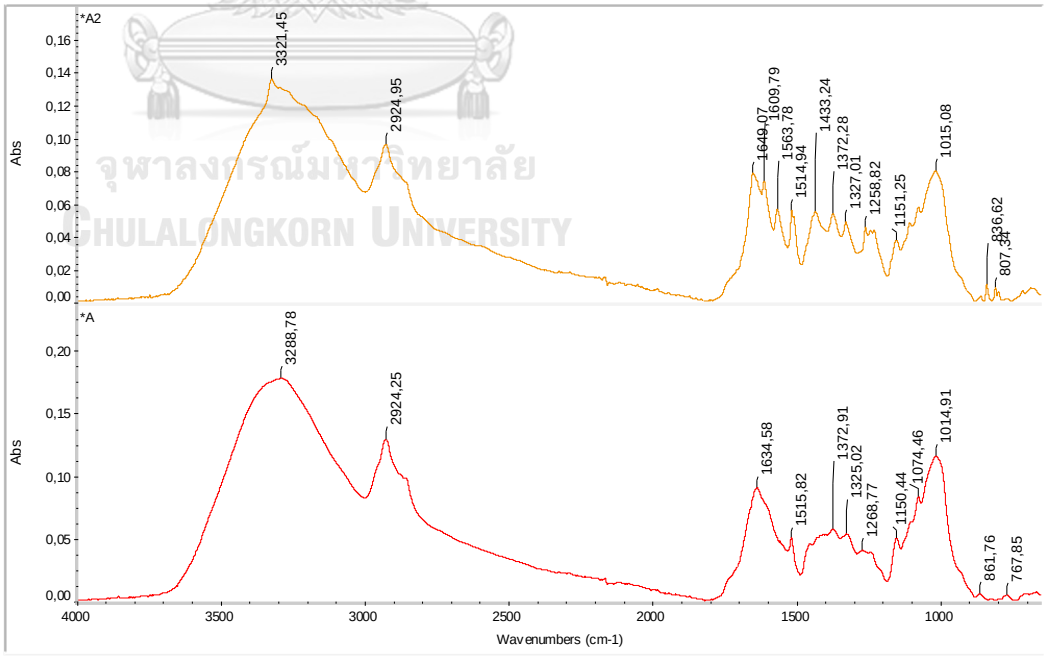
  Each analyte (adulterant) is scanned and observed; each characteristic peak is marked. The characteristic peaks of an analyte will be the marker of analyte presence after it has mixed with the sample matrices.

  Characteristic peak of analyte was also observed after being mixed with sample matrices. As shown in the Figure below, there are several characteristic peaks of paracetamol, ibuprofen, and aspirin that only each analyte had. Characteristic peaks of analytes that consistently appear in

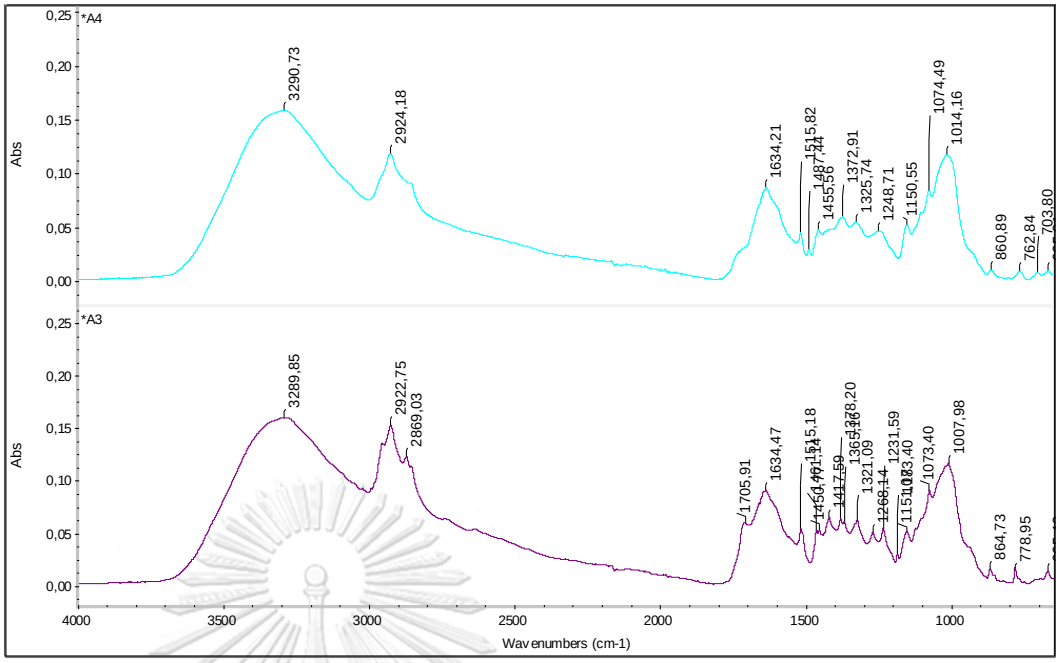
sample matrices are collected and shown in Table 6. Matrices of common peaks are also observed and recapped inside the Table.



(a)



(b)



(c)

Figure 23 (a) peak marker of analyte (adulterants); (b) original sample matrices A stacked with PCT adulterants A2; (c) sample matrices A adulterated with Ibuprofen A3 and Aspirin A4 stacked together

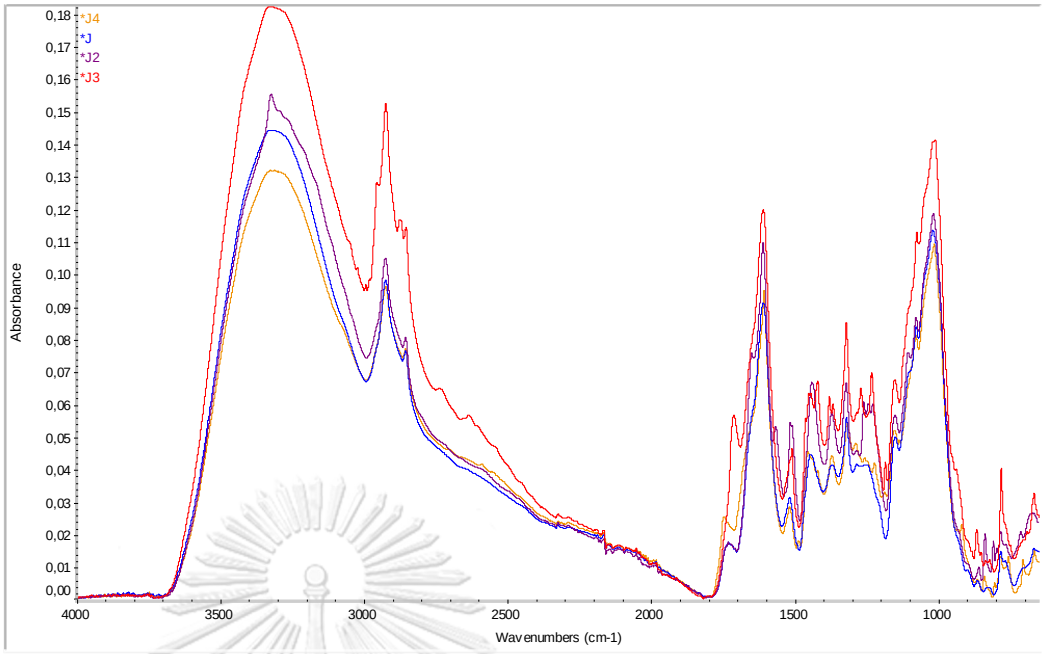
Table 6 Assigned peak region of samples in preliminary study

No	Analyte	Peaks Region Observed
1.	Paracetamol	3320 cm^{-1}
		1650 cm^{-1}
		1560 cm^{-1}
		830 cm^{-1}
		800 cm^{-1}
2.	Ibuprofen	2950 cm^{-1}
		2920 cm^{-1}
		1700-1710 cm^{-1}
		1500 cm^{-1}
		1450 cm^{-1}
		1440 cm^{-1}
		830 cm^{-1}

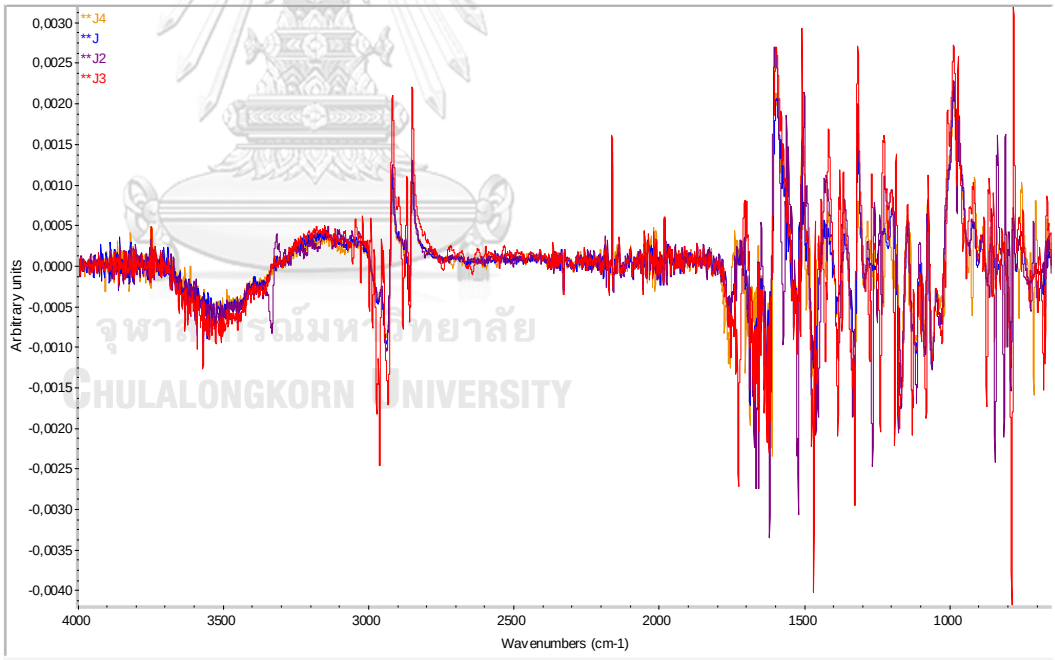
No	Analyte	Peaks Region Observed
		770 cm^{-1}
3.	Aspirin	1750 cm^{-1}
		1680 cm^{-1}
		1600 cm^{-1}
		1410 cm^{-1}
		916 cm^{-1}
		860 cm^{-1}
		750 cm^{-1}
		700 cm^{-1}
4.	Sample matrices common peak	2920 cm^{-1}
		1370 cm^{-1}
		1300 cm^{-1}
		1280 cm^{-1}
		1260 cm^{-1}
		1160 cm^{-1}
		1070 cm^{-1}
		1010 cm^{-1}

- Spectrum signal processing and classification model build trial

  Signal preprocessing performed in the experiment is auto-baseline and savitzky-golay 1st derivative spectrum in order to extract qualitative information from the spectrum using mathematical derivative. According to Rohman et al. (2014), derivative spectroscopy also helps eliminate baseline shift and baseline tilts (18). Signal preprocessing is done in OMNIC software for FTIR measurement and spectrum analysis. Signal processing of the original spectrum and derivative spectrum is presented in Figure 24.



(a)



(b)

Figure 24 (a) original sample spectrum; (b) Savitzky-Golay derivative spectrum of the samples in preliminary study

The classification model for LDA is built using derivative spectrum data of original samples and samples adulterated with paracetamol, ibuprofen, and

aspirin separately. The trial was conducted using eight Thai herbal medicine samples. The classification model build test purpose is to examine whether selected variables (assigned characteristic peaks) are good classifiers.

The original spectrum intensity value and the derivative one give different values in the model. Since the background interference is removed and the difference in intensity value of the variable between samples starts to give a contrast number. It affects the classification results of the model built and increases the classification results of the model. Derivative spectrum results give better classification results compared to the original spectrum, as shown in Figure 24b.

%classification results of model using original spectrum data is lower than derivative one due to background and baseline shift which affecting intensity value. If intensity value of assigned marker peak is low, the classification model couldn't recognize samples correctly. %classification results of original spectrum model is 94.4% with cross-validation results 50.0%, while derivative spectrum model gives 97.2% correct classification results and 88.9% cross-validation results.

Classification Results[a,c]

			Predicted Group Membership				
		marks	unadulterated	PCT adulterated	IBU adulterated	ASP adulterated	Total
Original	Count	unadulterated	8	1	0	0	9
		PCT adulterated	0	9	0	0	9
		IBU adulterated	0	0	9	0	9
		ASP adulterated	1	0	0	8	9
	%	unadulterated	88.9	11.1	.0	.0	100.0
		PCT adulterated	.0	100.0	.0	.0	100.0
		IBU adulterated	.0	.0	100.0	.0	100.0
		ASP adulterated	11.1	.0	.0	88.9	100.0
Cross-validated[b]	Count	unadulterated	2	3	2	2	9
		PCT adulterated	2	6	1	0	9
		IBU adulterated	3	0	5	1	9
		ASP adulterated	4	0	0	5	9
	%	unadulterated	22.2	33.3	22.2	22.2	100.0
		PCT adulterated	22.2	66.7	11.1	.0	100.0
		IBU adulterated	33.3	.0	55.6	11.1	100.0
		ASP adulterated	44.4	.0	.0	55.6	100.0

a. 94.4% of original grouped cases correctly classified.

b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 50.0% of cross-validated grouped cases correctly classified.

(a)

Classification Results[a,c]

			Predicted Group Membership				
		marks	unadulterated	PCT adulterated	IBU adulterated	ASP adulterated	Total
Original	Count	unadulterated	9	0	0	0	9
		PCT adulterated	0	9	0	0	9
		IBU adulterated	0	0	9	0	9
		ASP adulterated	1	0	0	8	9
	%	unadulterated	100.0	.0	.0	.0	100.0
		PCT adulterated	.0	100.0	.0	.0	100.0
		IBU adulterated	.0	.0	100.0	.0	100.0
		ASP adulterated	11.1	.0	.0	88.9	100.0
Cross-validated[b]	Count	unadulterated	8	0	0	1	9
		PCT adulterated	1	8	0	0	9
		IBU adulterated	1	0	8	0	9
		ASP adulterated	1	0	0	8	9
	%	unadulterated	88.9	.0	.0	11.1	100.0
		PCT adulterated	11.1	88.9	.0	.0	100.0
		IBU adulterated	11.1	.0	88.9	.0	100.0
		ASP adulterated	11.1	.0	.0	88.9	100.0

a. 97.2% of original grouped cases correctly classified.

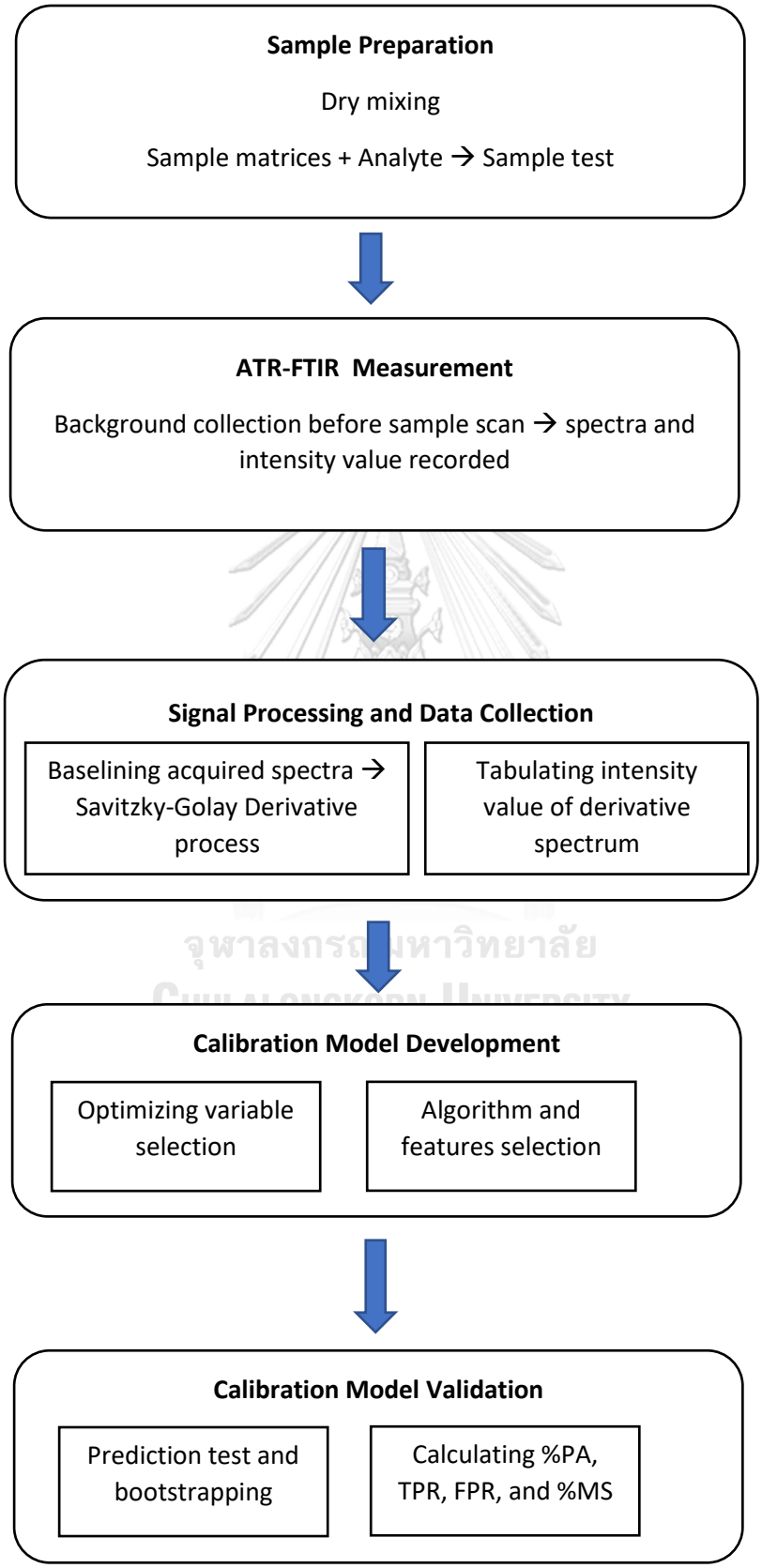
b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

c. 88.9% of cross-validated grouped cases correctly classified.

(b)

Figure 25 SPSS Discriminant Analysis classification results capture of 8 Thai herbal medicines in preliminary studies. (a) original spectrum data while (b) using Savitzky-Gollay derivative spectrum data

2.4 Research framework



Chapter III RESULTS AND DISCUSSION

3.1 Method Development

3.1.1 FTIR measurement and data matrix

FTIR spectra were measured using dry-mixed analyte with the samples matrix for ATR-FTIR absorbance intensity is interfered by the water presence (19). Samples code in the experiment presented as follows:

Table 7 Sample information for training and testing set

No	Name	Code	*Origins	Group	Indications
1	Thai Ginger Capsule	A	THAI	Training set	Ease abdominal pain
2	Thai Curcuma - Xanthorrhiza Capsule	B	THAI	Training set	Ease abdominal pain and inflammation
3	Thai Finger Root Capsule	C	THAI	Training set	Support inflammation treatment
4	Thai Andrographis - Paniculata Capsule	D	THAI	Training set	Fever reliever
5	Thai Andrographis - CAPSULE Mix	E	THAI	Training set	Fever reliever
6	Thai Embilica Extract Caps	F	THAI	Training set	Sore throat
7	Thai Turmeric Capsule	G	THAI	Training set	Support treatment for abdominal pain
8	Thai Cinnamon Capsule	H	THAI	Training set	Support inflammation treatment
9	Thai Herbal- Pain Pill	I	THAI	Testing set	Common pain
10	Thai Herbal - Antipyretic Tab	J	THAI	Testing set	Fever relieve
11	Indonesia Common Pain Herbal Tablet	K	INA	Training set	Common pain
12	Indonesia Light Fever Herbal Tablet	L	INA	Training set	Cold and fever
13	Indonesia HERBAL PAIN Tablet	M	INA	Training set	Common pain
14	Indonesia Oral Herbal Drink Powder For Flu	N	INA	Training set	Cold and fever
15	Indonesia Common Pain Herbal Drink	O	INA	Training set	Common pain

No	Name	Code	*Origins	Group	Indications
	Powder				
16	Indonesia Common Pain Herbal Drink Powder With Ginseng	P	INA	Training set	Common pain
17	Indonesia Common Pain And Fatigue Herbal Tablet	R	INA	Training set	Joint pain and common pain
18	Indonesia Common Pain Herbal Drink Powder Mix	Q	INA	Training set	Common pain
19	Indonesia Ginseng Pill	S	INA	Testing set	Common pain
20	Indonesia Joint Pain Herbal Pill	T	INA	Testing set	Joint pain
21	Thai Herbal Pain Pill – Indonesia Ginseng Pill	IS	THAI-INA	Testing set	Common pain
22	Thai Herbal Pain Pill – Indonesia Herbal Joint Pain Pill	IT	THAI-INA	Testing set	Common and joint pain
23	Thai Herbal Antipyretic Tab - Indonesia Ginseng Pill	JS	THAI-INA	Testing set	Fever relieve and common pain
24	Thai Herbal Antipyretic Tab - Indonesia Herbal Joint Pain Pill	JT	THAI-INA	Testing set	Fever relieve and common pain

Background spectra were collected before every sample measured. Air used as a background in measurement and the sample chamber is wiped with ethanol for cleaning.

Acquired spectra are processed with OMNIC software to get derivative IR spectra using the Savitzky-Golay function in the software. Spectra are saved as SP and CSV file types in order to observe the intensity value. The intensity value of each wavenumber region was compiled into one data matrix file and then copied to SPSS. The spectrum region included in the data matrix selected group regions at 3320 to 2950 cm^{-1} and 1750 to 750 cm^{-1} (Appendix A).

The FTIR detection limit was determined by the mixtures of matrix samples and analyte. The detection is observed by signal to noise ratio (S/N), S/N calculated by formula : $S/N = \frac{signal}{rms}$,where rms is root mean square of voltage noise ratio by FTIR machine (20). According to USP S/N ratio for detection limit should above 3.0.

Table 8 Detection limit test of ATR-FTIR Nicolet i50

Analyte	S/N				
	4%	6%	8%	10%	12%
PCT	4.19	6.97	4.21	4.03	4.16
IBU	4.80	5.10	5.24	5.15	5.33
ASP	4.89	5.23	5.14	5.15	5.29

As presented in Table 8, detection limit test results indicating that in 12% w/w adulteration level the instrument still gives S/N ratio above 3.0. However S/N ratio in experiment showing that intensity and noise linearly increasing at the same time with concentration increases. Therefore the S/N ratio of the analyte in IR spectra not significantly different.

3.1.2 Chemometric Model Development

The feature selection included in the test is a stepwise method for variable selection using mahalanobis distance with criteria selectio use of F value and probability of F. Sample treatment consist of unadulterated, single spiking and combination spiking (Figure 26). Model developed through various data input condition in trial and features selections. Model development built using derivative spectra data. since the intensity of native spectra is floating in each measurement that comes from background intensity effects, such unwanted signal variations(21). Data matrix of each variable selection method is tested in order to get the highest %correct classification that also meets descriptive statistical test requirement (ANOVA and covariance equality test) and also gives the highest value in cross-validation (Table 9). Statistic test of models provided in Appendix C.

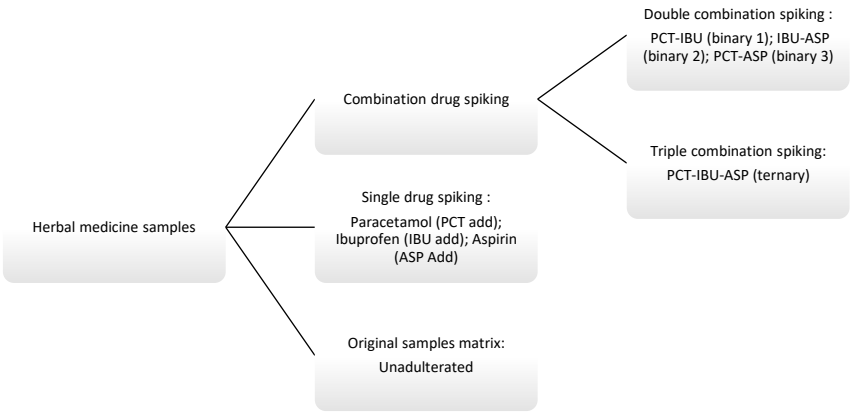


Figure 26 Sample naming scheme

The" use F value" as criteria in variable selection used to calculate the F ANOVA statistics value of each variable then compare it with the F value from the statistic Table. Variable included in function if the F value is above 3.40. The same way goes with Probability of F, however probability of F calculating significance of the F value. Probability of F includes the variable in the function if significance of variable F statistics value is below 0.05 and is removed when the significance value is greater. Region (variable) that is included in the separation function present in Table 10.

The stepwise method involves compressing 28 variables into 6 variables using the F-value criteria for selection, with 12 variables selected based on the Probability of F and 14 variables selected manually. The final step of the stepwise method involves calculating the tolerance value, where a higher value indicates that the variable has a significant impact on group separation.

Table 9 LDA model development with native and processed spectra trial

No	Data Input Test	Performance	
		%correct classification	cross-validation
	Native Spectra		
1	Use of F value	77.8%	69.4%
2	Probability of F	77.8%	63.9%
	Processed Spectra 1st derivative		
1	Use of F value	93.8%	93.8%
2	Probability of F	97.7%	93.8%
3	Manual selection	97.7%	88.9%
	Processed Spectra 2nd derivative		
1	Use of F value	96.1%	93.0%
2	Probability of F	96.1%	93.0%
3	Manual selection	99.2%	95.3%

The model was developed using 2nd derivative spectra data. The intensity of the original spectra varies in each measurement due to background intensity effects. These floating values result in inconsistent input data, as they come from signal variation. Consequently, the range of data to the group centroid (mean) can be wide(21), indicating poor data scattering and leading to numerous outliers that adversely affect the classification of the samples.

To address the background intensity effects, an auto-baseline function followed by Savitzky-Golay derivative spectroscopy is applied. This process eliminates the background intensity effects and provides more consistent peak intensities in the spectra measurements.

Table 10 Variable included in Discriminant Analysis SPSS

No	Assigned Variables	Use F Value		Probability of F		Manual selection	
		Entered	*Tolerance	Entered	*Tolerance	Entered	*Tolerance
1	3323	1650	0.444	2880	0.268	2959	
2	2959	1440	0.447	1650	0.348	1600	
3	2923	1370	0.736	1560	0.191	1560	
4	2880	1070	0.883	1500	0.426	1500	
5	1750	836	0.272	1440	0.227	1450	NA
6	1720	803	0.712	1300	0.226	1440	
7	1685			1280	0.493	1410	
8	1650			1260	0.201	1280	
9	1600			865	0.198	1260	
10	1560			836	0.327	916	
11	1500			803	0.067	836	
12	1450			704	0.059	803	
13	1440					775	
14	1410						
15	1370						
16	1300						
17	1280						
18	1260						
19	1160						
20	1070						
21	1010						
22	916						
23	865						
24	836						
25	803						
26	775						
27	756						
28	704						

Mahalanobis-distance stepwise method variable selection with use of F and Probability of F criteria selection gives similar %correct classification. While manual selection gives highest correct classification among other variable selection methods and chosen as prediction model (Table 11).

Table 11 Classification results of SPSS discriminant analysis model development

No	Samples Group	Criteria Selection					
		Use of F value		Probability of F		Manual selection	
		Overall %	CV %	Overall %	CV %	Overall %	CV %
1	Unadulterated	100	100	100	100	100	93.8
2	PCT Add	100	100	100	100	100	100
3	IBU Add	100	93.8	93.8	93.8	100	93.8
4	ASP Add	93.8	93.8	93.8	93.8	93.8	87.5
5	Binary 1	100	87.5	93.8	87.5	100	93.8
6	Binary 2	81.3	81.3	93.8	81.3	100	93.8
7	Binary 3	93.8	93.8	93.8	93.8	100	100
8	Ternary	100	93.8	100	93.8	100	100

The stepwise method tends to have lower correct classification rates because it only includes variables based on statistical calculation significance (probability of F) or their statistics value, potentially resulting in a loss of information during the variable selection process. However, the stepwise method can still be chosen when there is uncertainty about which variables have a significant impact on separation. On the other hand, the manual selection of variables involves building models through trial and error, along with visual inspection of the spectra. This approach allows us to select spectra regions that contain critical information for separation. However, it can be time-consuming to find the best variable combination compared to the stepwise method.

SPSS discriminant analysis also provides a distribution graph of the sample groups. A good classification should exhibit minimal scattering of sample data relative to the group centroid. However, Linear Discriminant Analysis (LDA), follows Gaussian distribution assumption, the results can be sensitive and

biased towards outlier data. Another drawback of LDA is that groups with similar mean (centroid) values can lead to biased classification grouping(22).

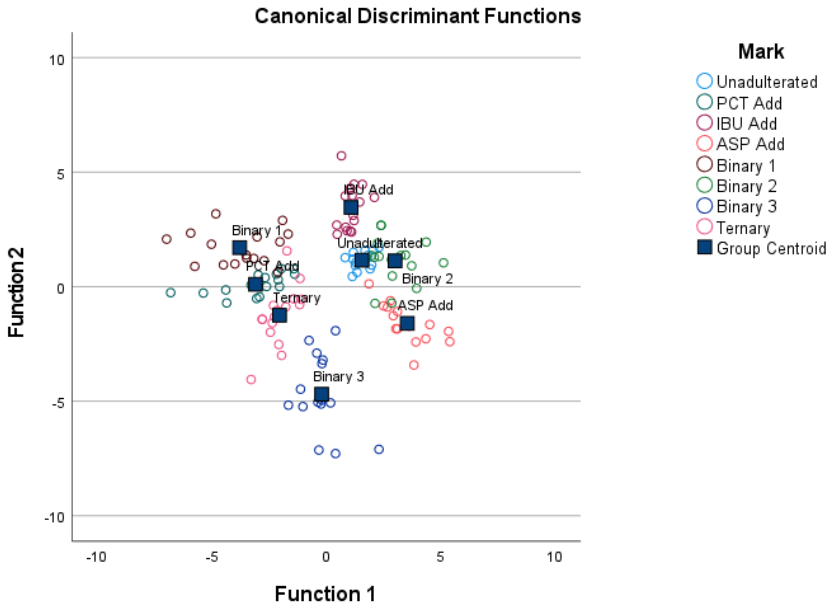


Figure 27 LDA graph of the manual selection model

3.2 Method Validation

3.2.1 Discriminant Analysis Model Evaluation

3.2.1.1 Auto-prediction test results

An auto-prediction test was conducted using 8 independent testing set samples, with 8 different treatments were used as training set samples, resulting in a total of 64 samples. The test was performed using the SPSS scoring wizard function. Overall, the auto-prediction test achieved 63 correct classifications out of 64, resulting in an accuracy of 98.44%. The classification results for each group are presented in the confusion matrix below:

Table 12 Confusion matrix of testing set samples auto-prediction results

		True Value								Total
	Count	Unadulterated	PCT Add	Ibu Add	ASP Add	Binary 1	Binary 2	Binary 3	Ternary	
Prediction	Unadulterated	8	0	0	0	0	0	0	0	8
	PCT Add	0	8	0	0	0	0	0	0	8
	Ibu Add	0	0	8	0	0	0	0	0	8
	ASP Add	0	0	0	8	0	0	0	0	8
	Binary 1	0	0	0	0	8	0	0	0	8
	Binary 2	0	0	0	0	0	8	0	0	8
	Binary 3	0	0	0	0	0	0	8	0	8
	Ternary	0	1	0	0	0	0	0	7	8
		True Value								Total
	Percent (%)	Unadulterated	PCT Add	Ibu Add	ASP Add	Binary 1	Binary 2	Binary 3	Ternary	
Prediction	Unadulterated	100	0	0	0	0	0	0	0	100
	PCT Add	0	100	0	0	0	0	0	0	100
	Ibu Add	0	0	100	0	0	0	0	0	100
	ASP Add	0	0	0	100	0	0	0	0	100
	Binary 1	0	0	0	0	100	0	0	0	100
	Binary 2	0	0	0	0	0	100	0	0	100
	Binary 3	0	0	0	0	0	0	100	0	100
	Ternary	0	12.5	0	0	0	0	0	87.5	100

Misclassification errors occurred in the ternary sample group. These errors may be attributed to the close proximity in distribution between the ternary group and the PCT add group.

3.2.1.2 ROC curve results

Evaluation of the model following USP Chapter 1039 was also conducted with a ROC (receiver operating characteristic) curve. The data used to run ROC analysis came from the probability of prediction model in the SPSS scoring wizard function.

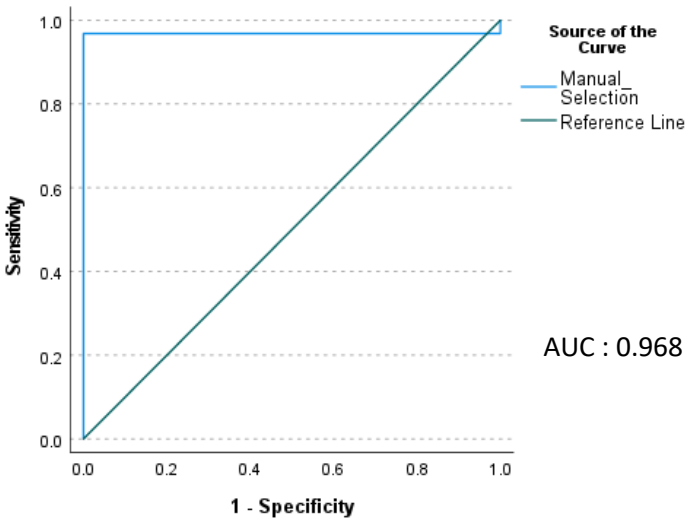


Figure 28 ROC Curve of LDA model from auto-prediction result

ROC curve is metrics for binary classification data (true/false), and the experiment data that are multiclass classification turned into binary

AUC: 0.979

prediction. ROC area under the curve (AUC) of the graph is parameter of how well the model make true or false prediction. The graph is made from 1-specificity in the horizontal axis and sensitivity in the vertical axis, with a large AUC indicating a high value of specificity and sensitivity (23). According to USP, the ideal AUC of the ROC curve is close to 1. TPR (sensitivity) and FPR (1-specificity) calculations are determined from the cut-off value of the ROC curve that gives the minimum value of probability prediction from the scoring wizard giving correct classification. TPR, FPR, PA, and MS calculations are presented in Table 13. The chosen cut-off value is 0.6537 probability prediction, which gives 96.8% TPR and 0.0% FPR.

Table 13 Cut-off value of TPR and FPR from ROC Curve

Prediction Probability	Cut-off	
	TPR	FPR
0.5982	98.4%	100.0%
0.6382	96.8%	100.0%
0.6537	96.8%	0.0%

3.2.2 Comparative testing with HPLC Method

Comparative testing was conducted on binary-3 samples from the testing set to evaluate the method's capability to simultaneously detect paracetamol, aspirin, salicylic acid, and caffeine. Before the test began, the system suitability was checked by injecting a combined standard solution of PCT at a concentration of 0.0012 mg/ml and aspirin at a concentration of 0.0064 mg/ml. Five consecutive injections showed that the relative standard deviation (RSD) values for retention time and peak area of both analytes were below 2.0%. Additionally, the tailing factor for each analyte was below 2.0, and the theoretical plate count was above 2000 (24), as presented in Table 14.

Table 14 System Suitability test results of PCT (a) and ASP (b)

PCT

No	Injection	Ret. Time	Peak Area	Tailing	TP
1	SST1	5.223	44715	1.603	4566
2	SST2	5.22	43512	1.687	4601
3	SST3	5.225	42998	1.681	4660
4	SST4	5.213	43096	1.672	4578
5	SST5	5.22	43135	1.688	4567
Average		5.220	43491.20		
STDev		0.00	711.41		
%RSD		0.09	1.64		

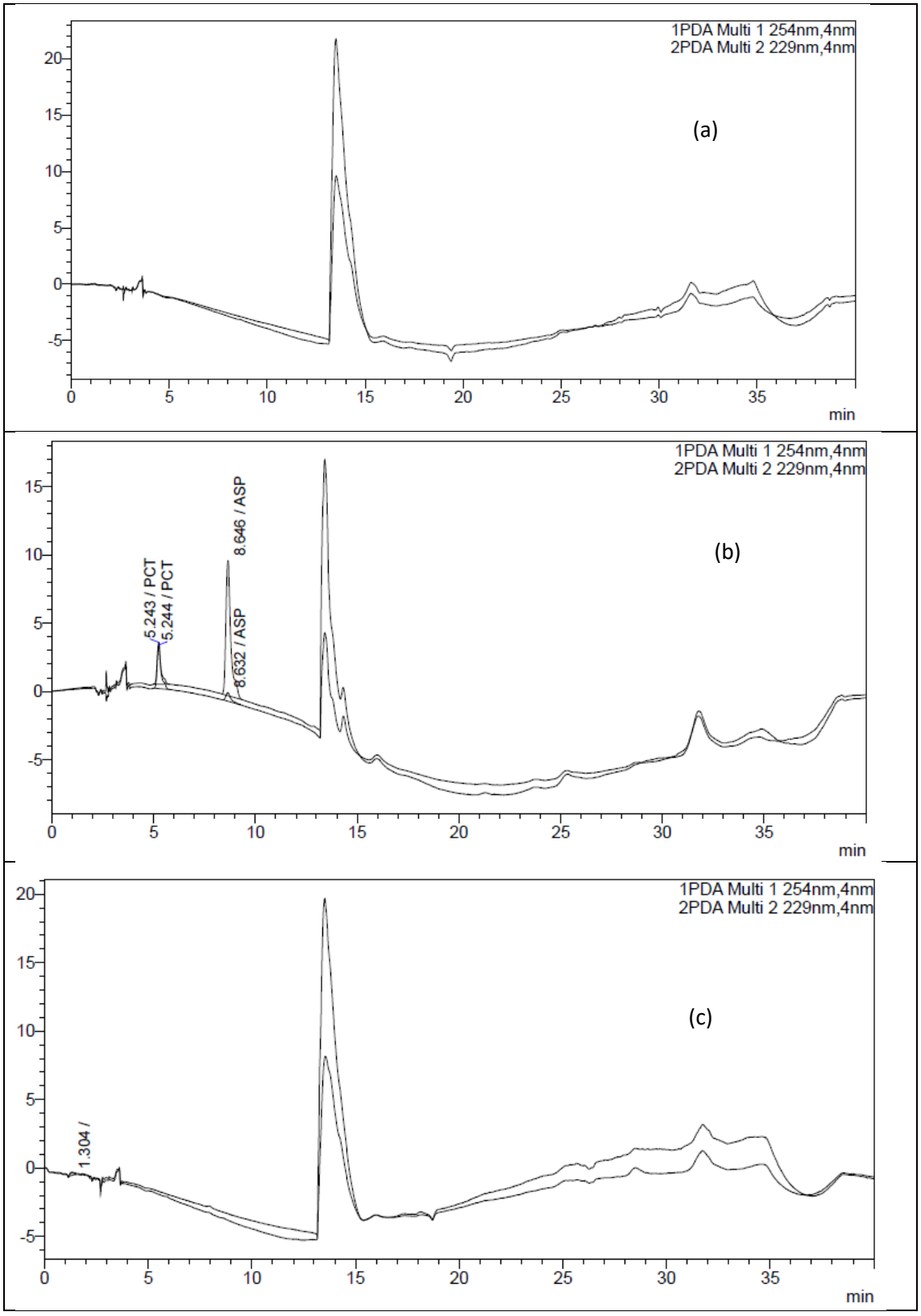
(a)

ASP

No	Injection	Ret. Time	Peak Area	Tailing	TP
1	SST1	8.595	167296	1.769	7652
2	SST2	8.596	167161	1.738	7645
3	SST3	8.600	166092	1.722	7761
4	SST4	8.579	165997	1.717	7733
5	SST5	8.604	166995	1.724	7616
Average		8.595	166708.20		
STDev		0.01	616.10		
%RSD		0.11	0.37		

(b)

The specificity test shows the PCT retention time around 5 minutes and the ASP around 8–9 minutes at 254 nm and 229 nm detection wavelengths. Spectra acquired by LC-PDA show PCT lambda max at 243 nm and ASP at 296 nm. The systems detection limit was determined at 1% w/w for PCT and 5% w/w for ASP. (Appendix B)


1PDA Multi 1 254nm,4nm
2PDA Multi 2 229nm,4nm
(a)
min
1PDA Multi 1 254nm,4nm
2PDA Multi 2 229nm,4nm
5.243 / PCT
5.244 / PCT
8.646 / ASP
8.632 / ASP
(b)
min
1PDA Multi 1 254nm,4nm
2PDA Multi 2 229nm,4nm
1.304 /
(c)
min

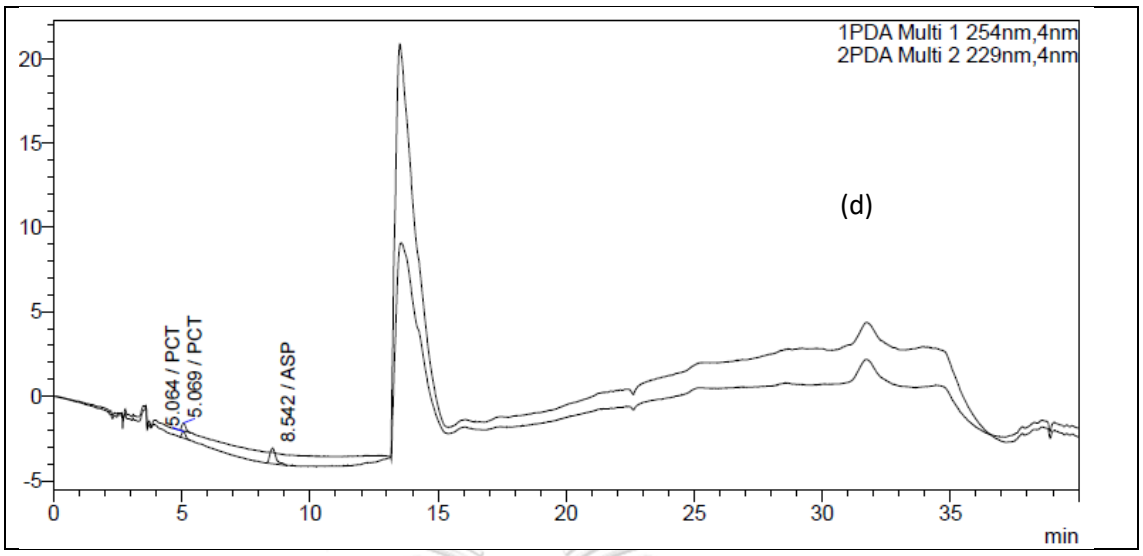


Figure 29 . HPLC chromatogram of (a) diluent (MeOH 60%), standard solution (b), one of the sample matrix (c) with the spiked one (d)

The HPLC detection yielded a total of 16 correct results out of 16, matching the predictions made by the prediction model. The sample used for comparative testing involved a double combination spiking of PCT-ASP. It should be noted that the HPLC method was only validated to detect the presence of PCT and ASP in herbal products.

While the HPLC method offers better sensitivity and specificity, it requires complex preparation procedures, which can increase the likelihood of analyte loss. In this experiment, the HPLC method exhibited low recovery after the solid-phase extraction (SPE) process. Based on the area per area ratio compared to standard spiked samples, the peak area of the analyte dropped by up to 90%.

The low recovery in the SPE process could be attributed to various factors, including analyte pass-through during sample loading, elution of analytes during the washing sequence, and the use of elution solvents that may not be sufficiently strong to fully elute the analyte of interest(25).

Table 15 Comparative testing results between HPLC method and predictive model

No	Samples	HPLC Results	Model Prediction
1	I	-	-
2	J	-	-
3	S	-	-
4	T	-	-
5	IS	-	-
6	IT	-	-
7	JS	-	-
8	JT	-	-
9	I6	+	+
10	J6	+	+
11	S6	+	+
12	T6	+	+
13	IS6	+	+
14	IT6	+	+
15	JS6	+	+
16	JT6	+	+
Total		16	16
Percentage (%)		100	100

*Predicted as ternary mixtures

CHAPTER IV CONCLUSION

4.1 Conclusion

ATR-FTIR spectra obtained require a derivative and smoothing process through the Savitzky-Golay derivative process in order to eliminate the background intensity effect. The intensity of the spectra will become data input in the statistical process.

28 wavenumber regions assigned as variables are reduced into 12 variables that give a significant separation function. Manual selection model picked for it gives highest correct classification and cross-validation results.

The selected models were tested using independent testing sets of samples to conduct an auto-prediction test. The results of the auto-prediction test demonstrated an overall correct classification rate exceeding 90% as the target. Additionally, the model was evaluated using ROC curve analysis, yielding an AUC value of 0.968 (close to 1, as per the USP standard). The ROC curve also indicated a probability prediction cut-off value of 0.6537, which achieved a high true positive rate (TPR) and a low false positive rate (FPR).

Comparative testing with HPLC with testing samples that contain PCT and ASP shows that the prediction model gives a correct classification on par with HPLC results. Indicating the model could be a good candidate for routine screening tools (Table 16).

Table 16 Comparison of the ATR-FTIR LDA model with established HPLC Method

Comparison	HPLC	ATR-FTIR and LDA Model
Sample type	All dosage forms	Only solid dosage form
Sample preparation	Complex	Simple (only grinding)
Operational time	12-24 hours	1-2 hours
Operational cost	High	Cheap
Accuracy	100%	98.55%
Sensitivity	1% (PCT); 5% (ASP)	12%
Special requirement	Analytical grade reagent	Statistical software

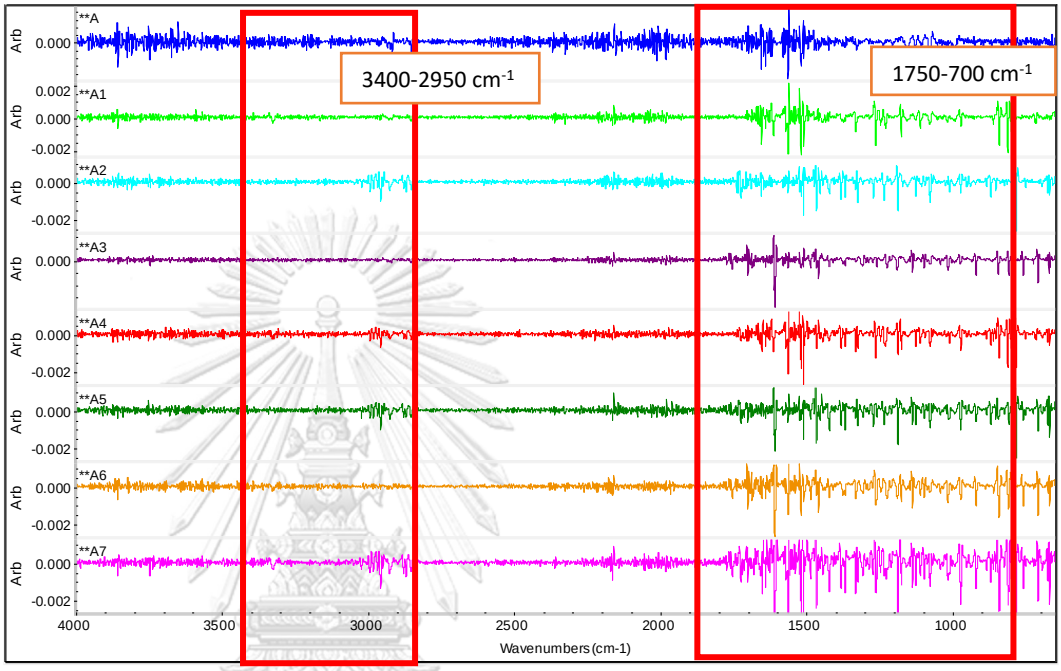
Future project suggestion for the model needed are confirmation of ibuprofen detection with established confirmatory method, testing the model against analyte that have similar functional group as adulterant or the degradation products of analyte, and pilot project of screening test in real sample.

Appendix A

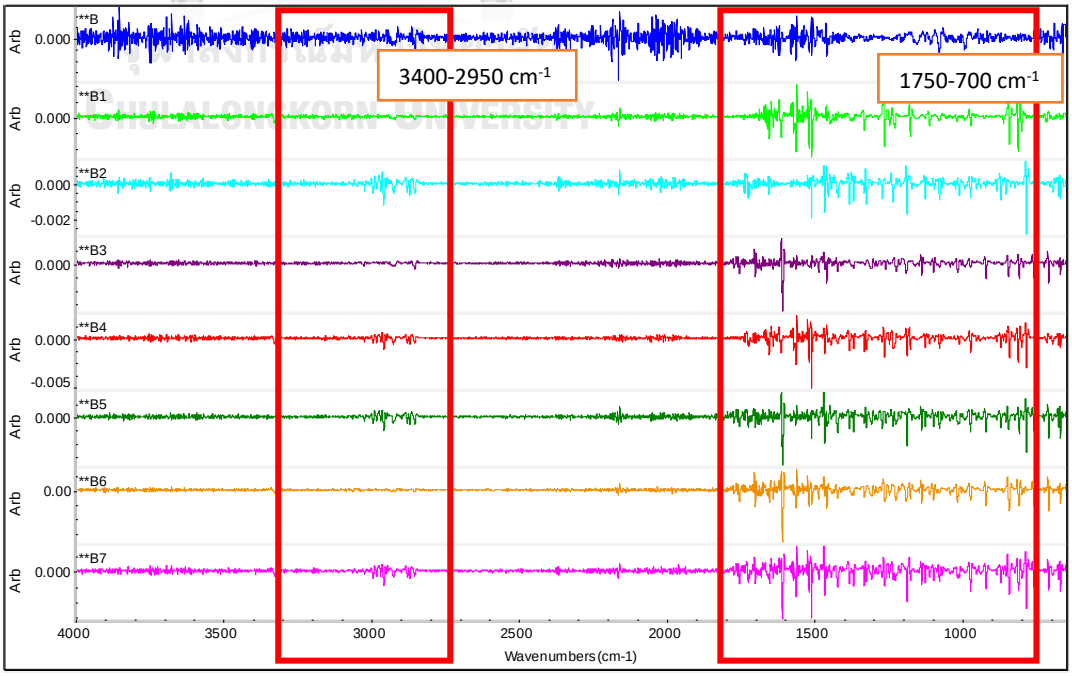
- 2nd- Derivative FTIR Spectra of Samples

  Selected region for data processing at 3400-2950 cm^{-1} and 1750-700 cm^{-1}.

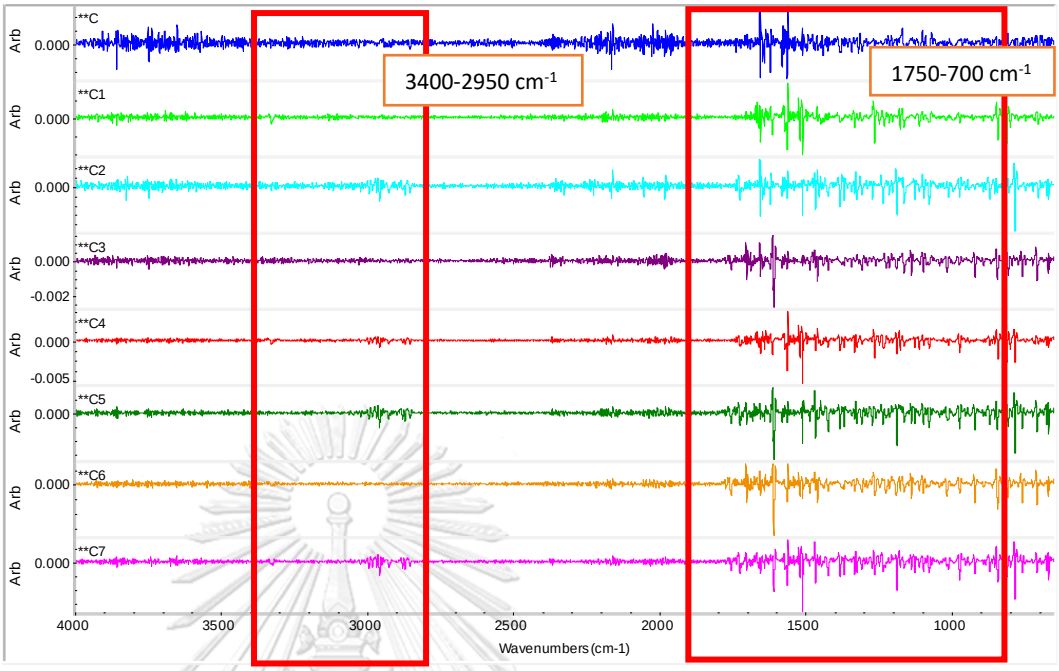
  a. Sample A



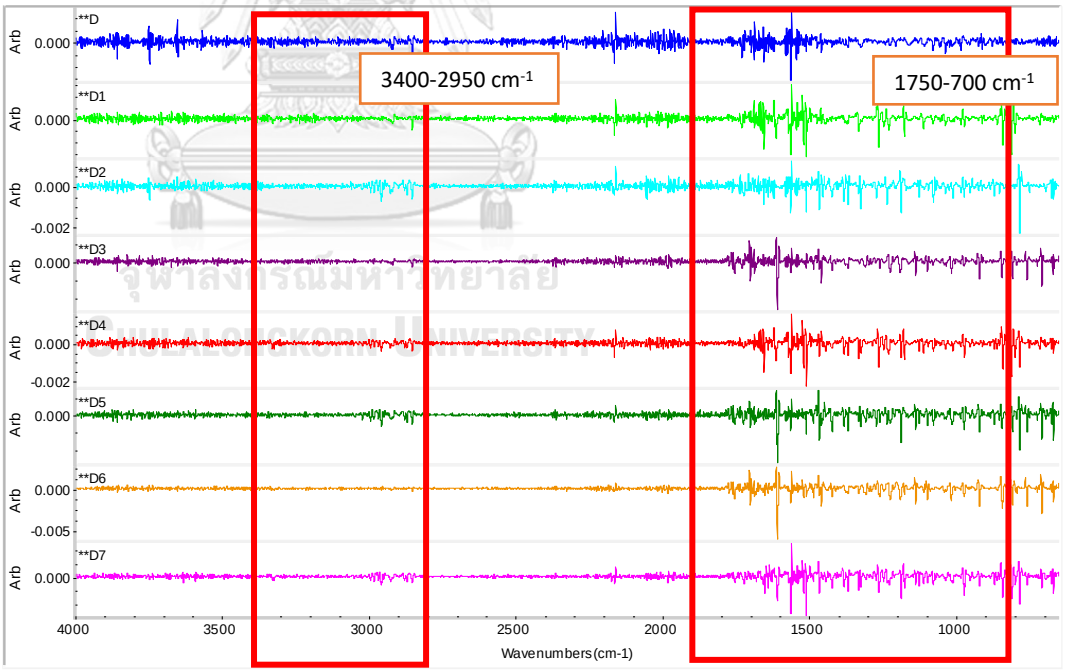
  b. Sample B



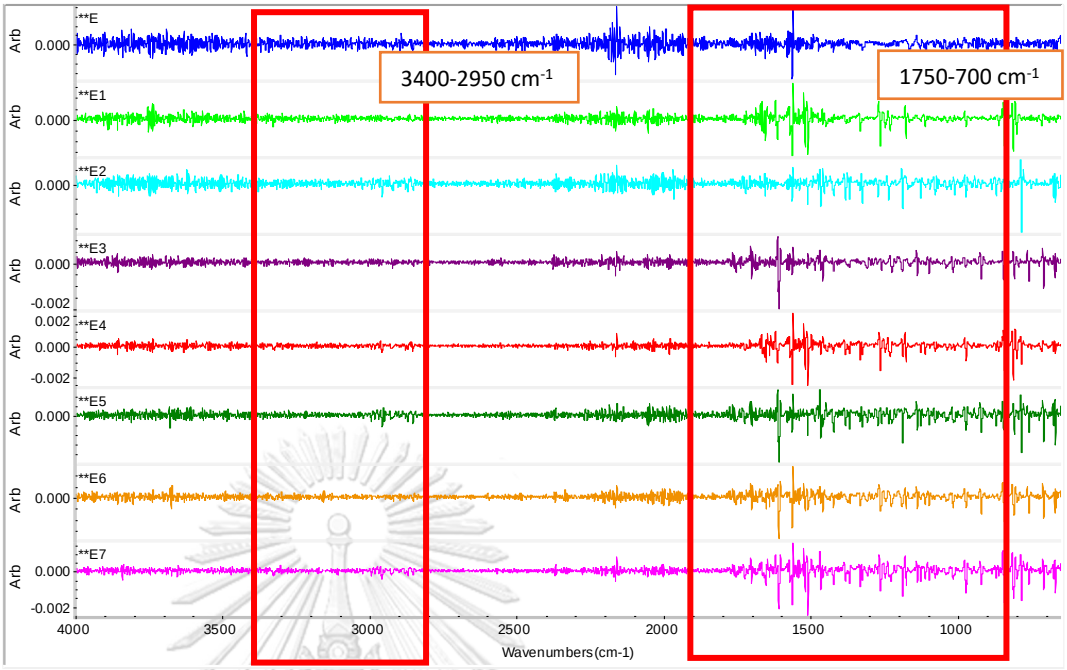
c. Sample C



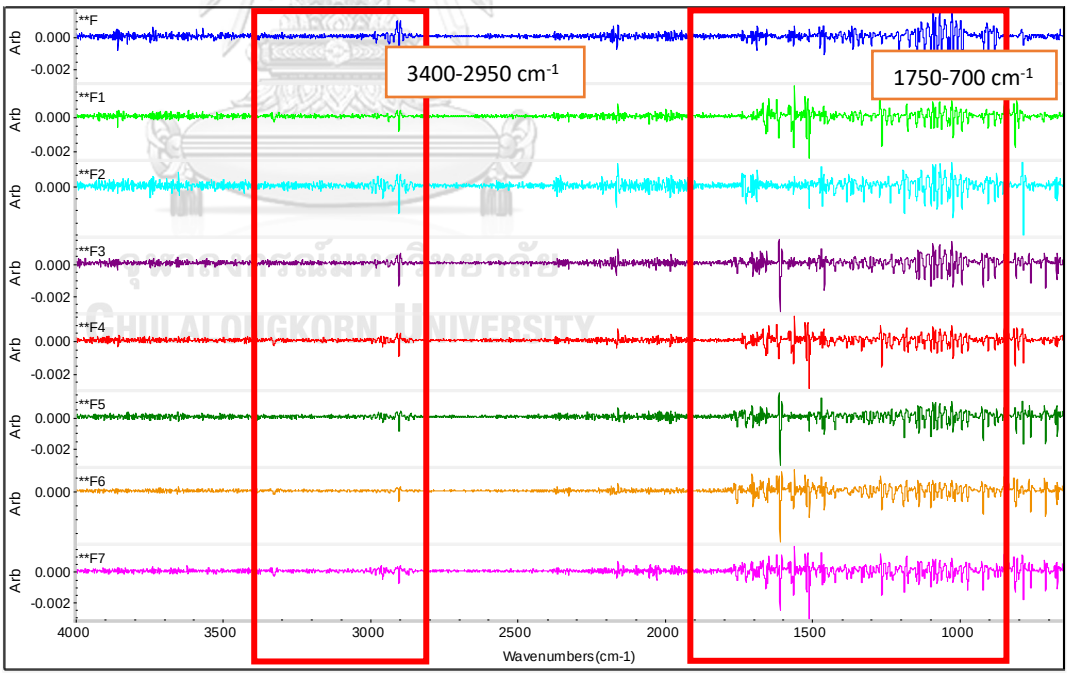
d. Sample D



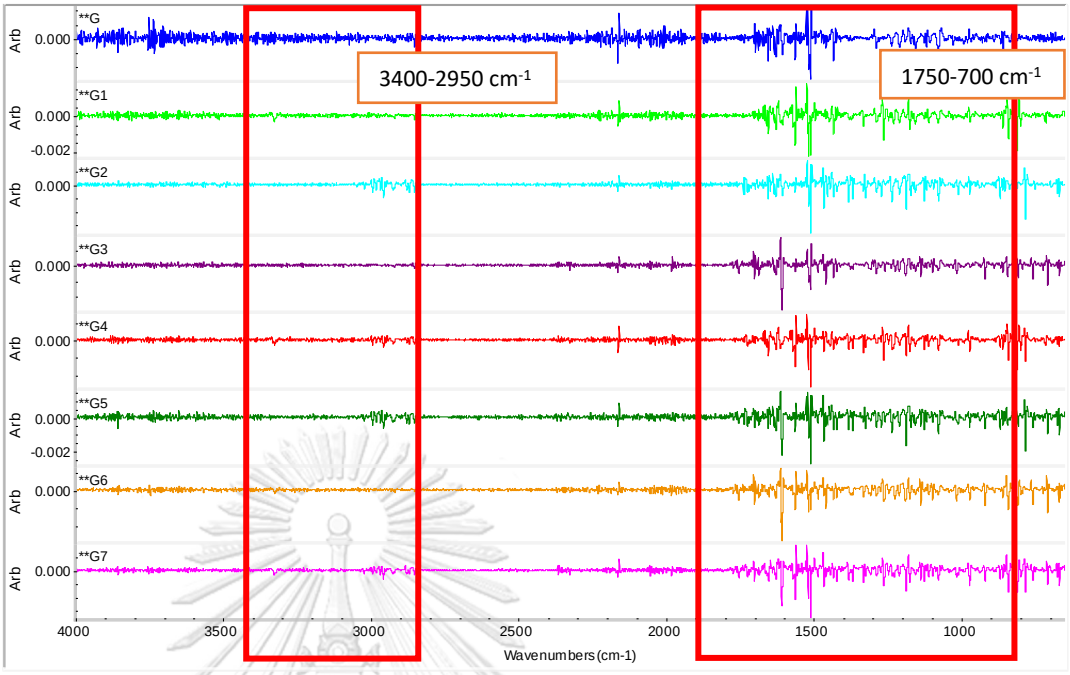
e. Sample E



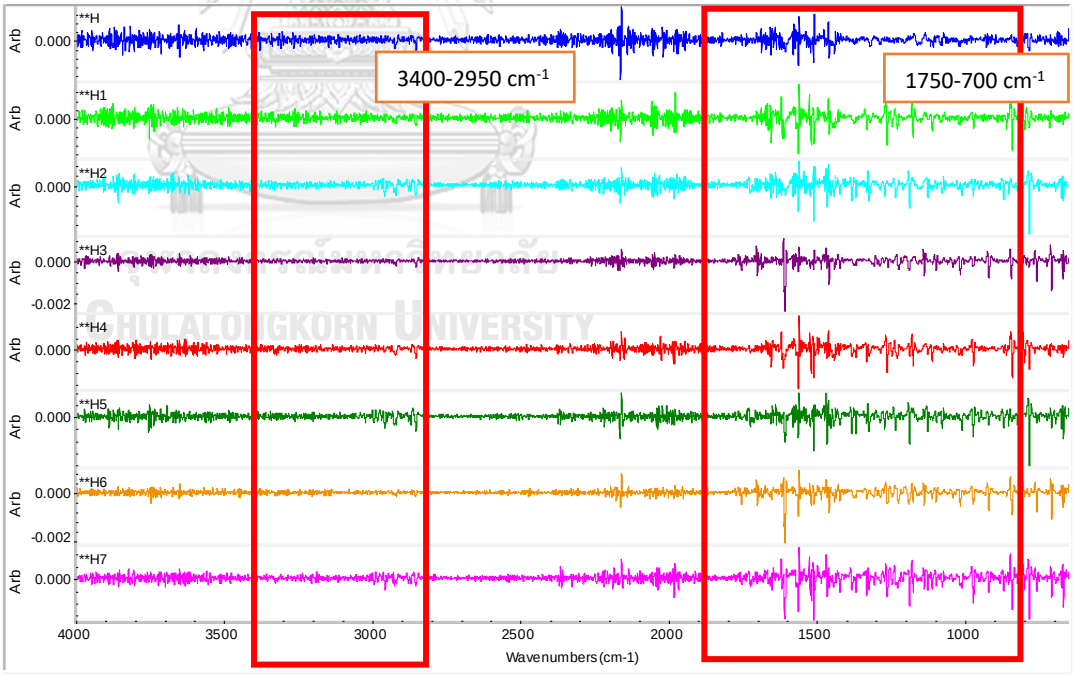
f. Sample F



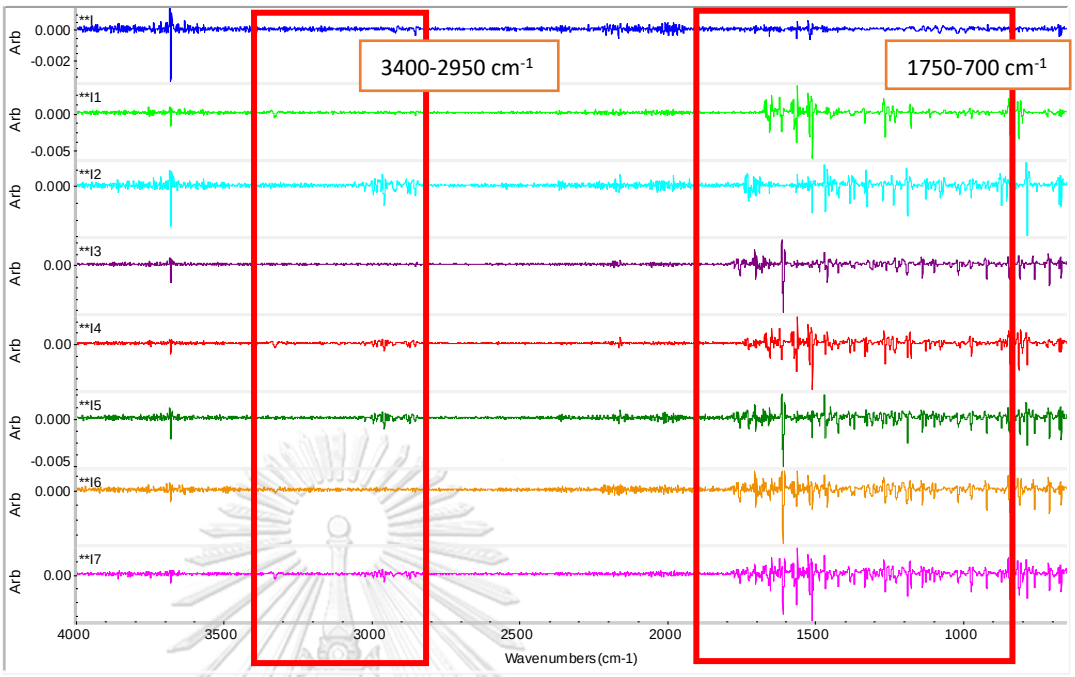
g. Sample G



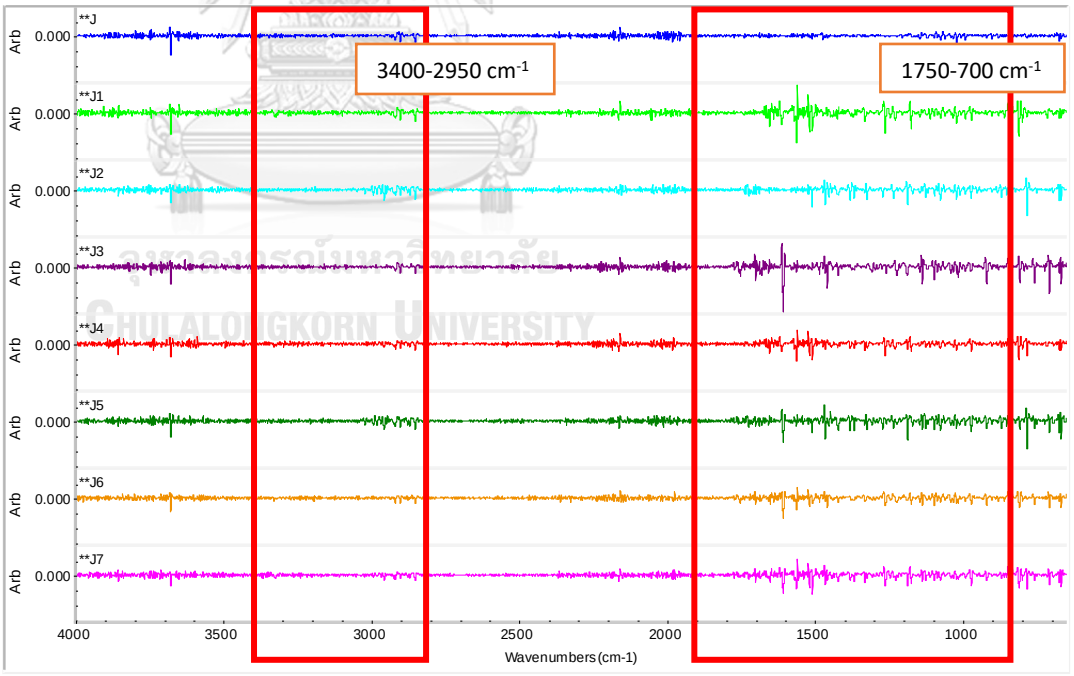
h. Sample H



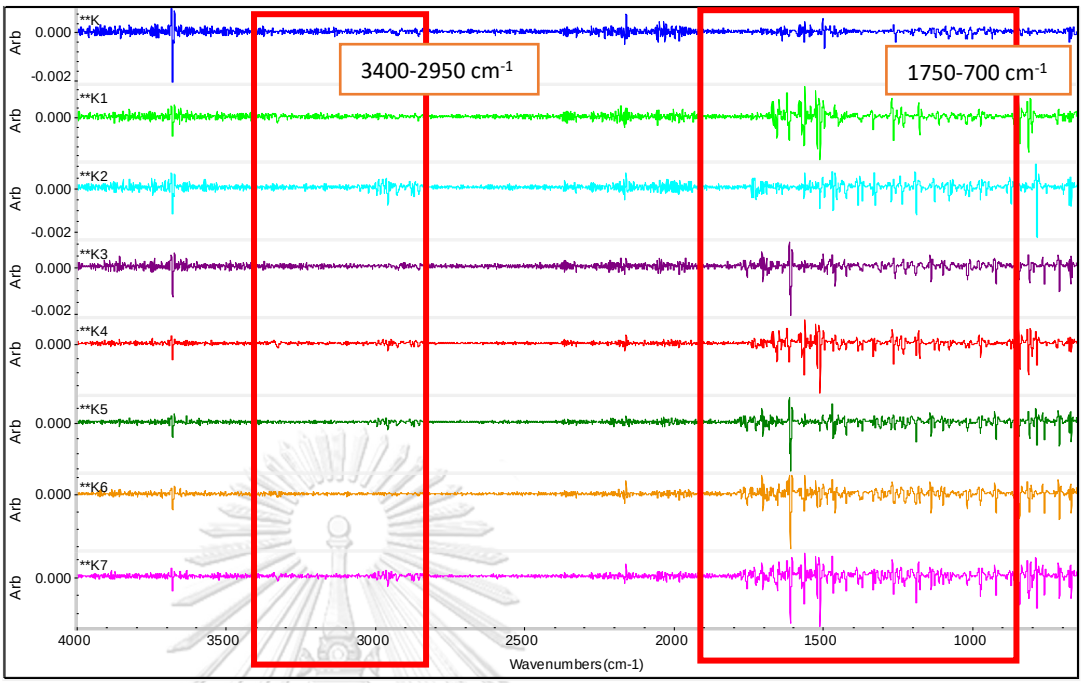
i. Sample I



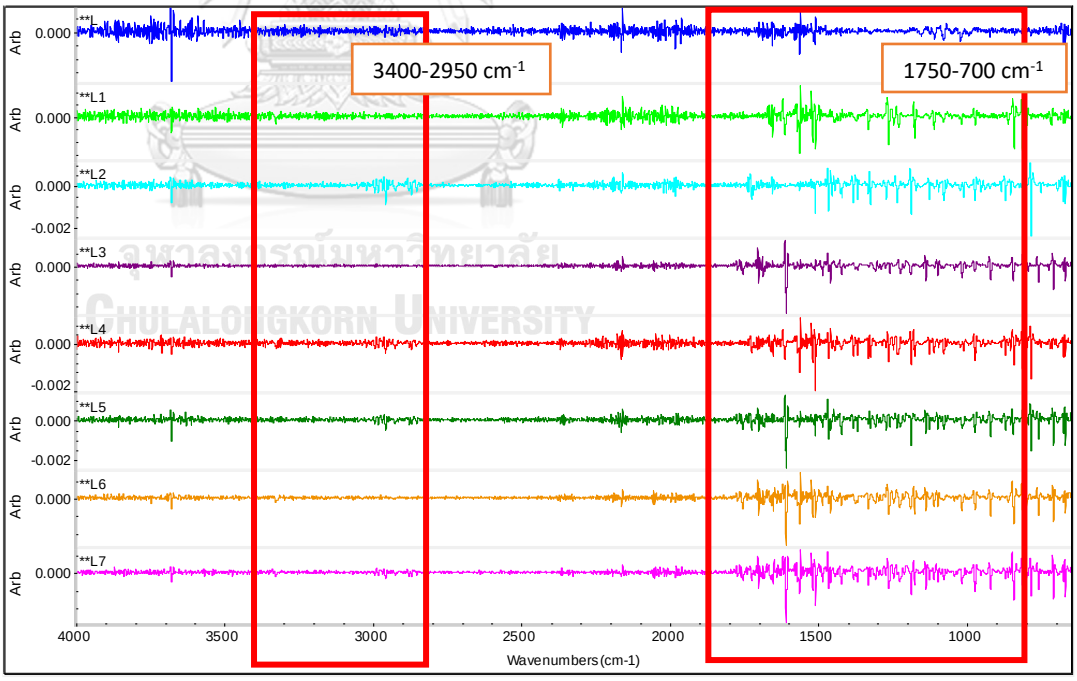
j. Sample J



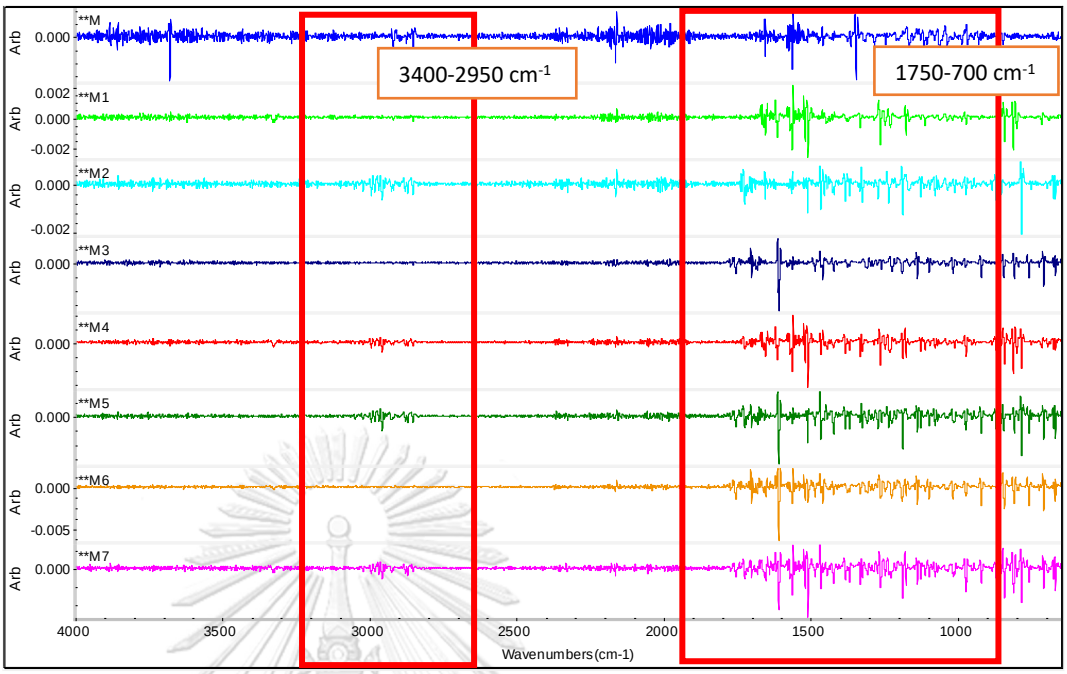
k. Sample K



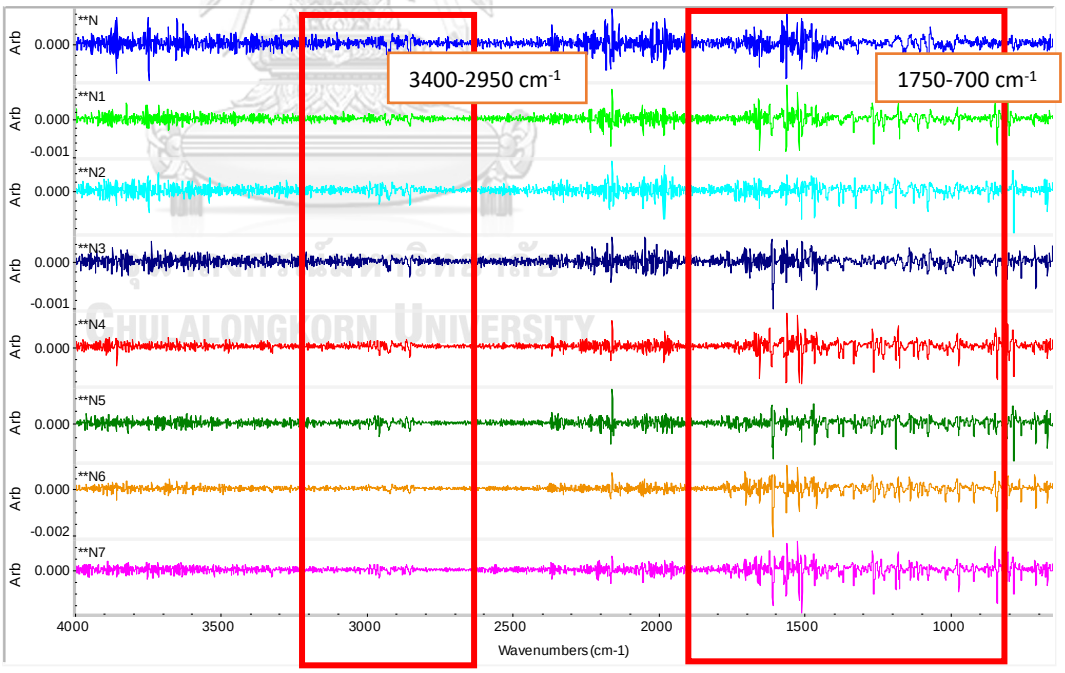
l. Sample L



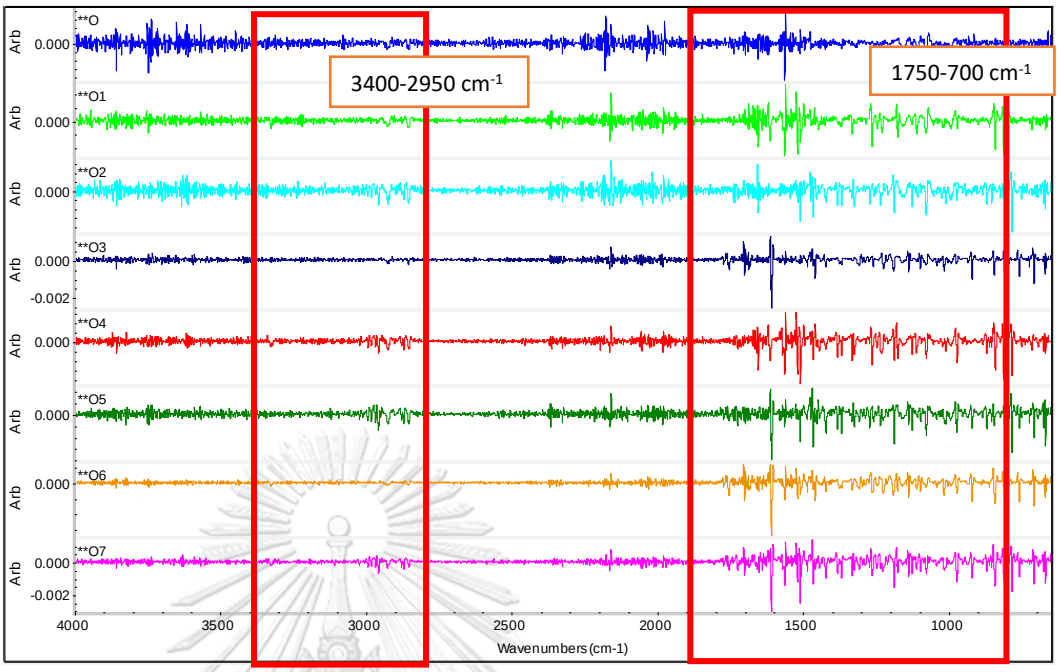
m. Sample M



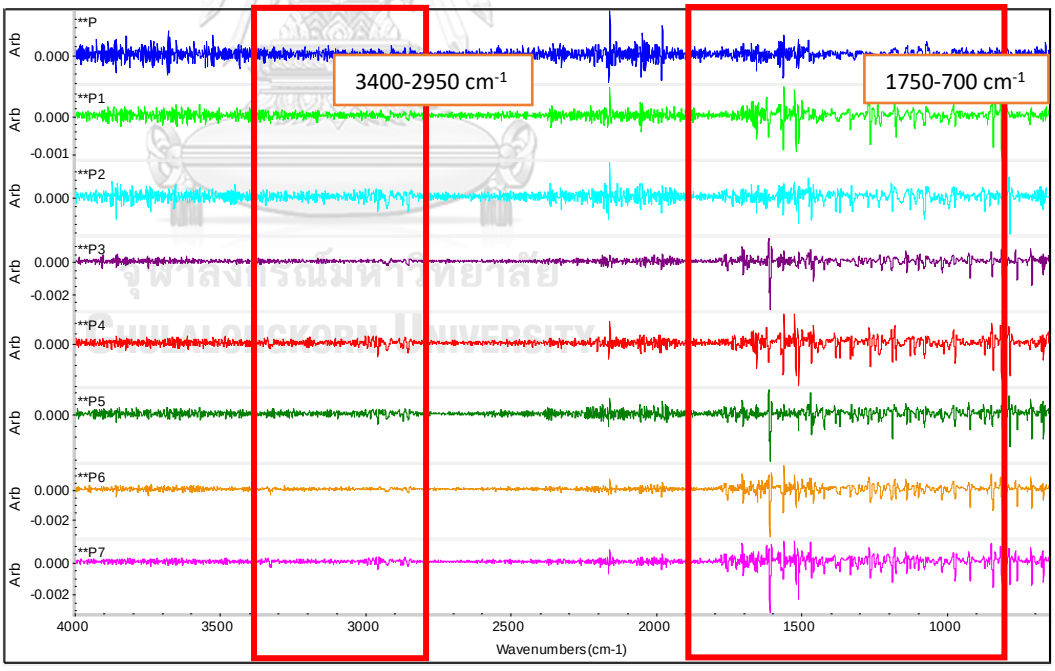
n. Sample N



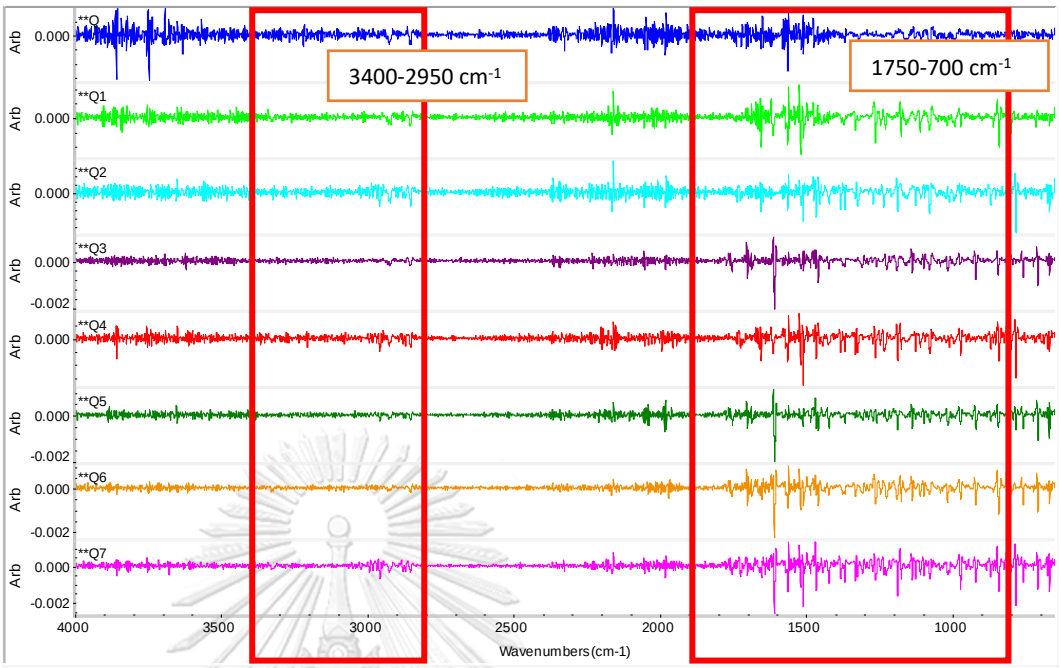
o. Sample O



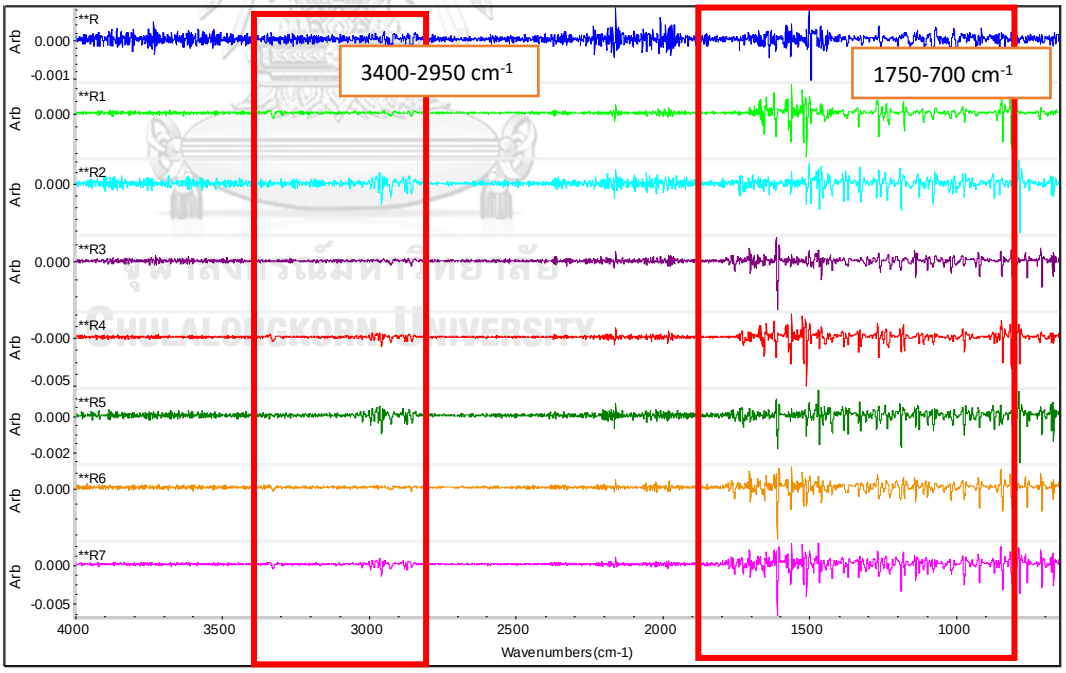
p. Sample P



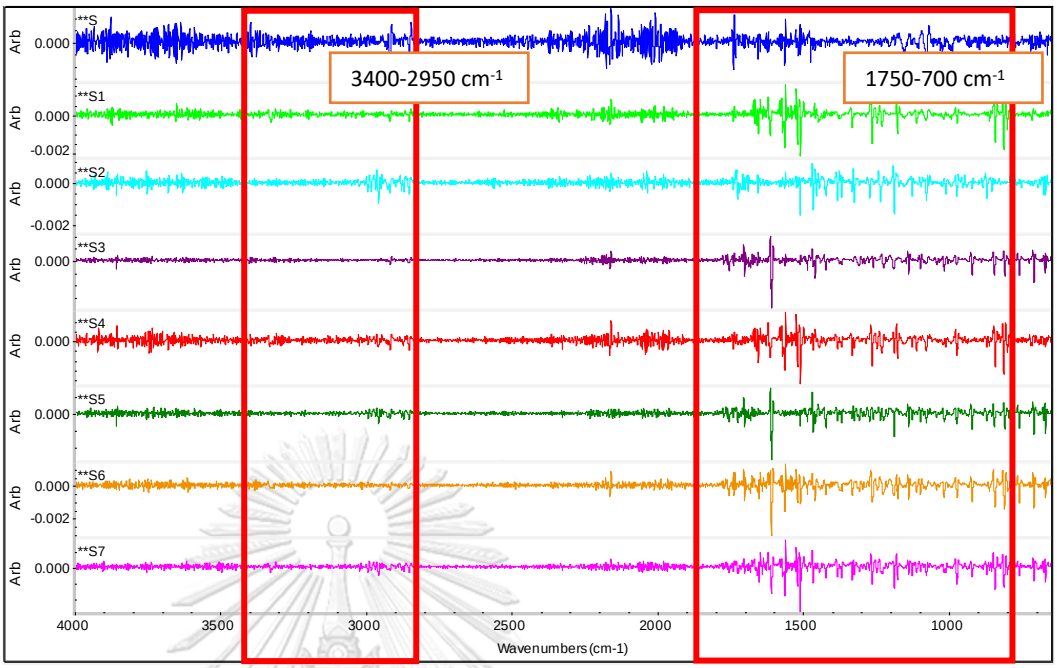
q. Sample Q



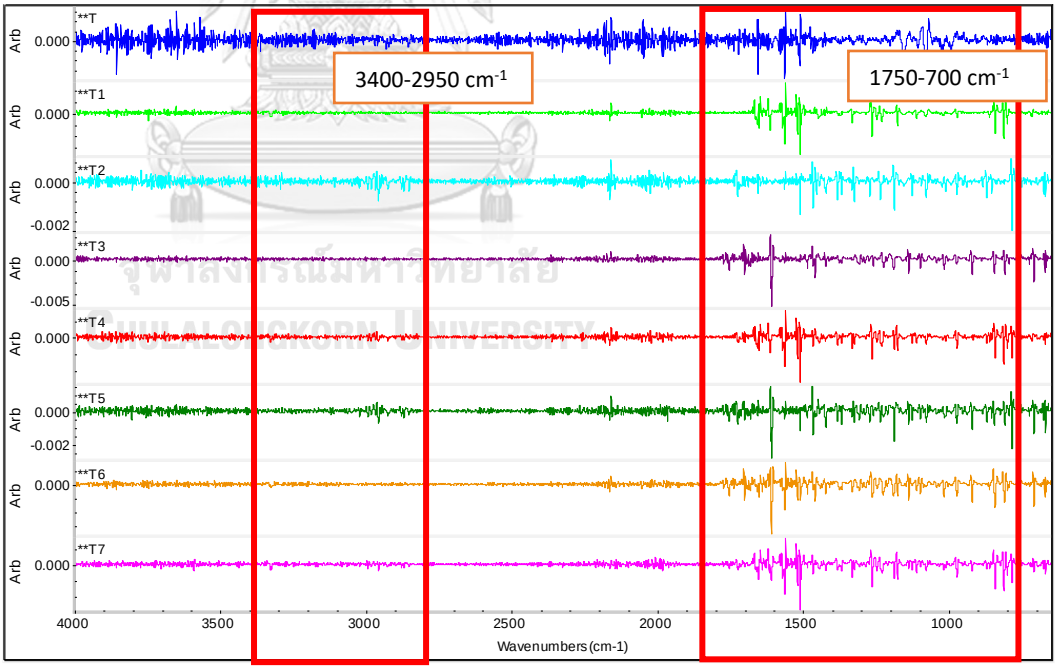
r. Sample R



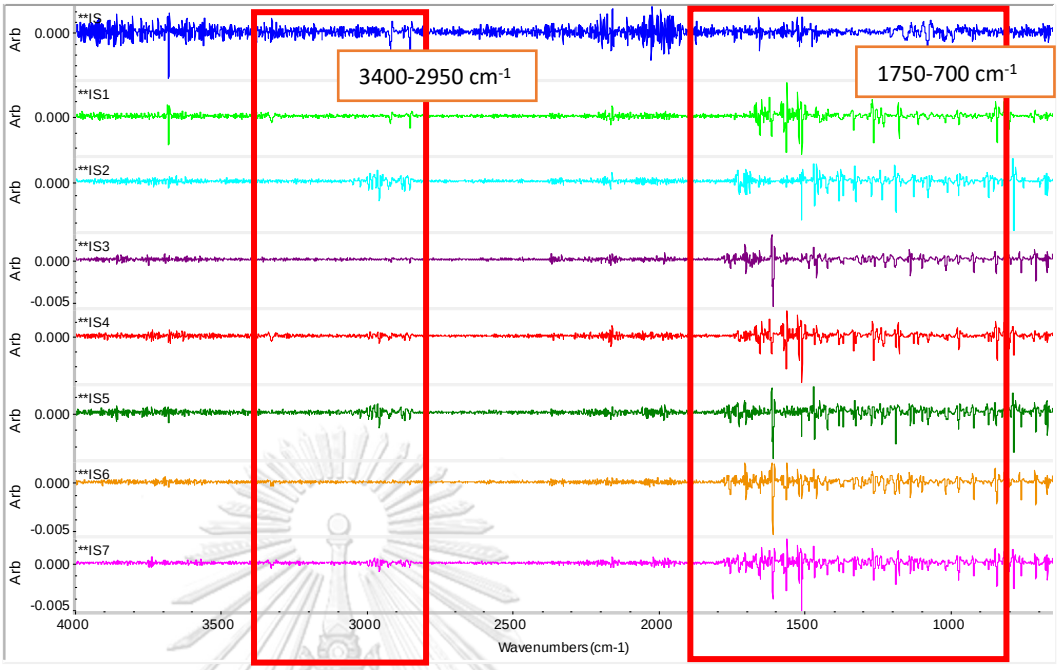
s. Sample S



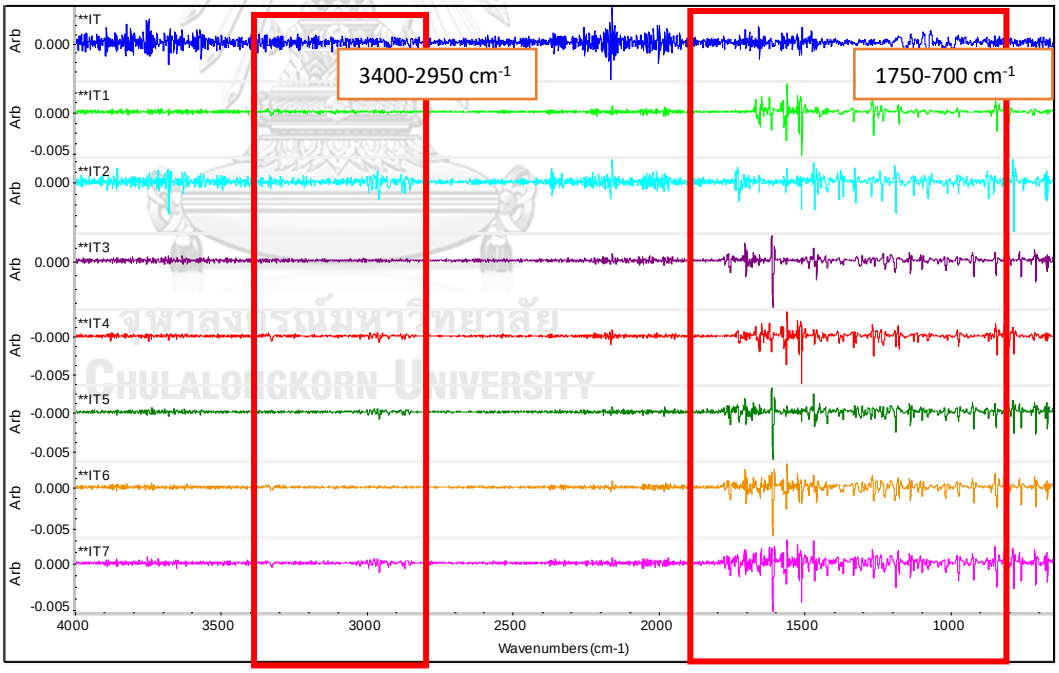
t. Sample T



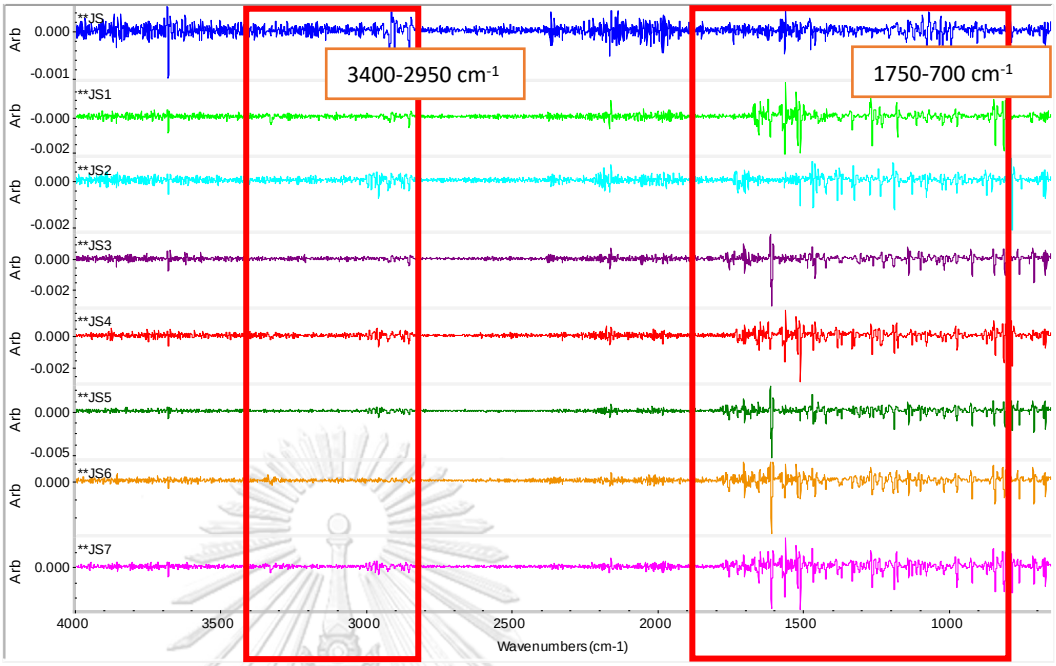
u. Sample IS



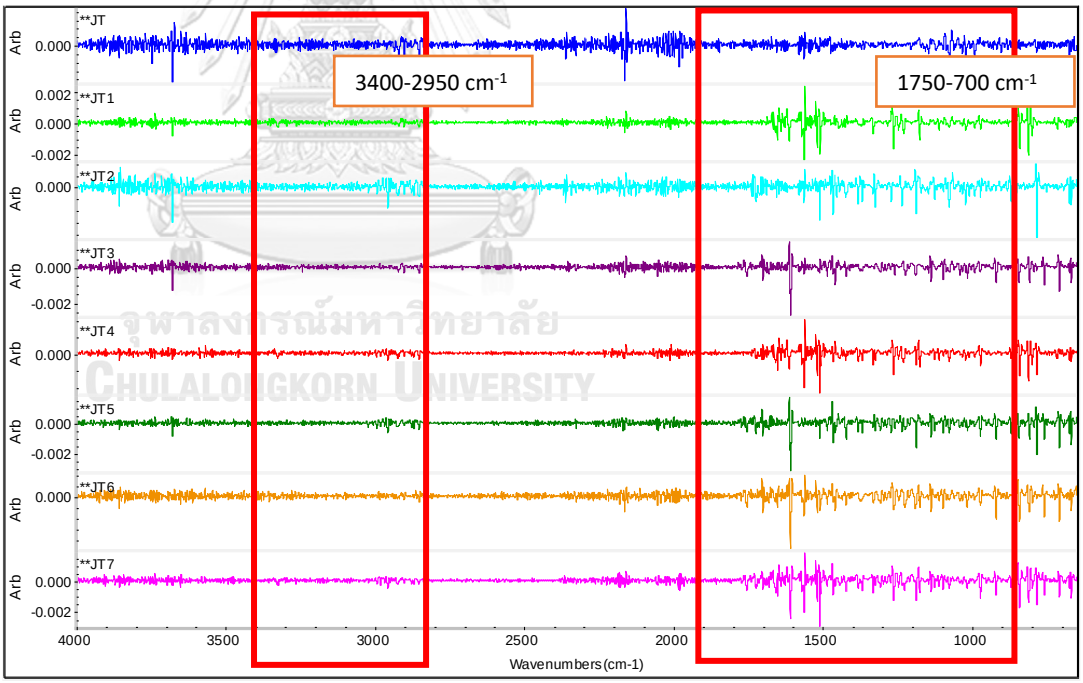
v. Sample IT



w. Sample JS



x. Sample JT



Legends:

-- : code 0 for unadulterated samples

-- : code 1 for PCT spiked samples

-- : code 2 for IBU spiked samples

-- : code 3 for ASP spiked samples

-- : code 4 for double combination spiking of PCT and IBU

-- : code 5 for double combination spiking of IBU and ASP

-- : code 6 for double combination spiking of ASP and PCT

-- : code 7 for triple combination spiking of all adulterants

Appendix B

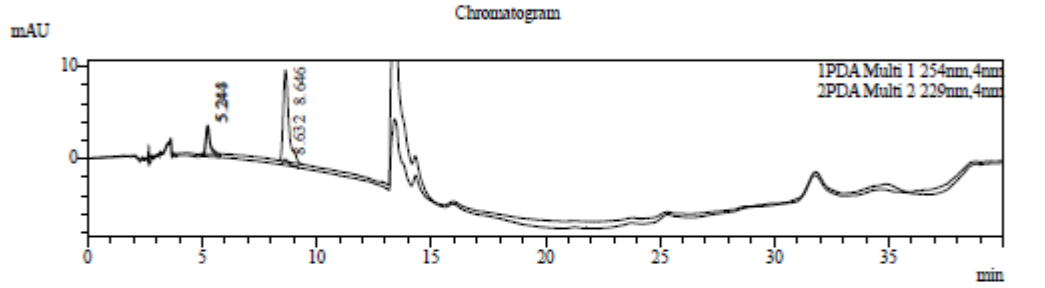
HPLC comparative test chromatogram

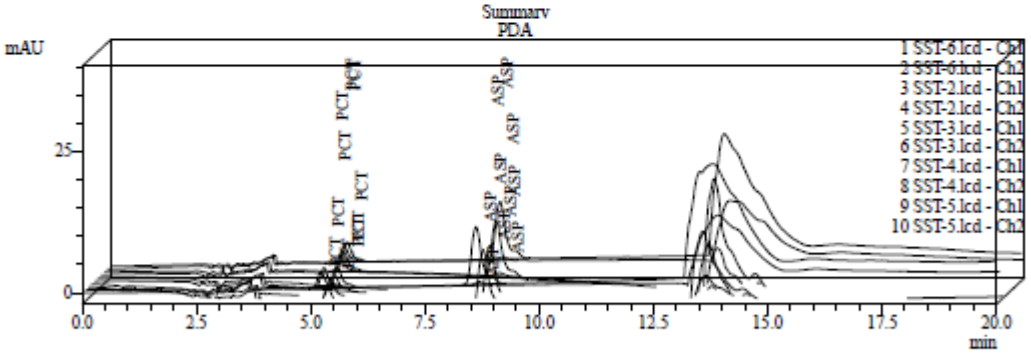
- System Suitability Test (SST)
  a. SST part-1

SHIMADZU LabSolutions **Analysis Report**

<Sample Information>

Sample Name	: SST-1		
Sample ID	: SST-1		
Data Filename	: SST-2.lcd		
Method Filename	: Trial MA OT2411 postrun.lcm		
Batch Filename	: MA OT 24-11 160523.lcb		
Vial #	: 1-1	Sample Type	: Unknown
Injection Volume	: 20 uL		
Date Acquired	: 17/5/2566 8:34:10	Acquired by	: System Administrator
Date Processed	: 17/5/2566 17:26:59	Processed by	: System Administrator





<< PDA >>

ID#1 Compound Name: PCT

Title	RT	Sample Name	Area	Height	Tailing	Plate	Rs	SF	S/N	Peak Purity Index
SST-6.lcd	5.185	SST-5	43446	3430	1.631	4524	--	--	16.95	Not calculated
SST-2.lcd	5.244	SST-1	37582	2869	1.684	4498	--	--	7.40	Not calculated
SST-3.lcd	5.246	SST-2	36743	2881	1.674	4592	--	--	13.75	Not calculated
SST-4.lcd	5.240	SST-3	36637	2899	1.637	4600	14.086	--	10.64	Not calculated
SST-5.lcd	5.225	SST-4	37058	2927	1.625	4582	--	--	10.14	Not calculated
Average	5.228		38293	3001	1.630	4559	14.086	--	11.78	Not calculated
%RSD	0.483		7.583	8.022	1.637	0.995	0.000	0.000	31.126	0.000

ID#4 Compound Name: ASP

Title	RT	Sample Name	Area	Height	Tailing	Plate	Rs	SF	S/N	Peak Purity Index
SST-6.lcd	8.597	SST-5	16696/	10183	1.659	7755	9.710	--	30.84	Not calculated
SST-2.lcd	8.632	SST-1	10464	623	1.853	7522	9.558	--	2.66	Not calculated
SST-3.lcd	8.647	SST-2	9876	619	1.921	7919	9.752	--	3.79	Not calculated
SST-4.lcd	8.631	SST-3	8024	592	1.189	8795	10.045	1.769	3.87	Not calculated
SST-5.lcd	8.619	SST-4	8205	597	1.129	8902	10.084	--	3.30	Not calculated
Average	8.625		40707	2523	1.550	8179	9.830	1.769	8.89	Not calculated
%RSD	0.215		173.406	169.751	23.895	7.688	2.303	0.000	138.099	0.000

b. SST part-2

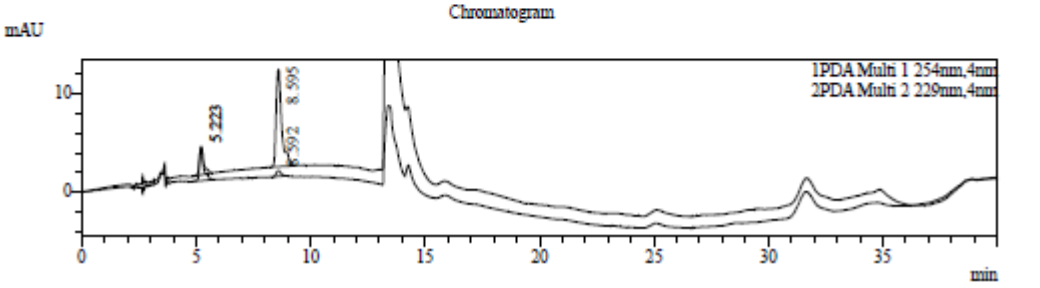
SHIMADZU LabSolutions

Analysis Report

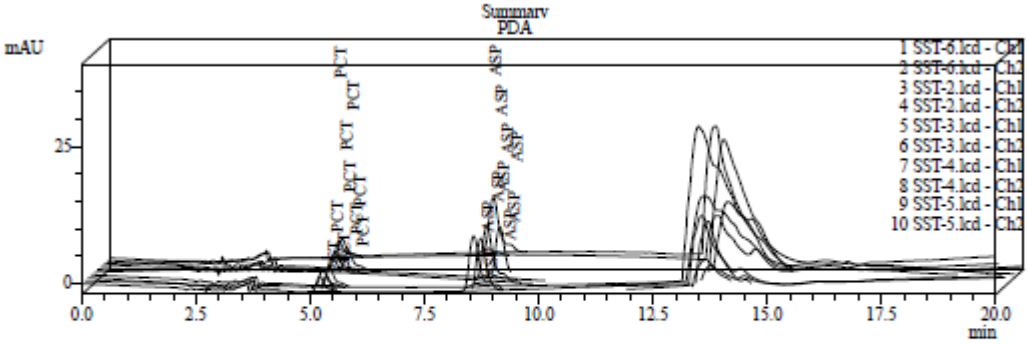
<Sample Information>

Sample Name	: SST-1		
Sample ID	: SST-1		
Data Filename	: SST-2.lcd		
Method Filename	: Trial MA OT2411 postrun.lcm		
Batch Filename	: MA OT 24-11 180523.lcb		
Vial #	: 1-1	Sample Type	: Unknown
Injection Volume	: 20 uL		
Date Acquired	: 18/5/2566 7:58:07	Acquired by	: System Administrator
Date Processed	: 18/5/2566 15:29:50	Processed by	: System Administrator

Chromatogram



Summary
PDA



<< PDA >>

ID#1 Compound Name: PCT

Title	RT	Sample Name	Area	Height	Tailing	Plate	Rs	SF	S/N	Peak Purity Index
SST-6.lcd	5.220	SST-5	36719	2835	1.680	4491	--	--	6.62	Not calculated
SST-2.lcd	5.223	SST-1	38113	2872	1.616	4510	--	--	6.25	Not calculated
SST-3.lcd	5.220	SST-2	37038	2865	1.706	4537	--	--	10.42	Not calculated
SST-4.lcd	5.225	SST-3	36653	2863	1.665	4601	--	--	10.02	Not calculated
SST-5.lcd	5.212	SST-4	36313	2836	1.660	4554	--	--	6.65	Not calculated
Average	5.220		37091	2854	1.665	4539	--	--	7.99	Not calculated
%RSD	0.089		2.045	0.599	1.971	0.931	0.000	0.000	25.594	0.000

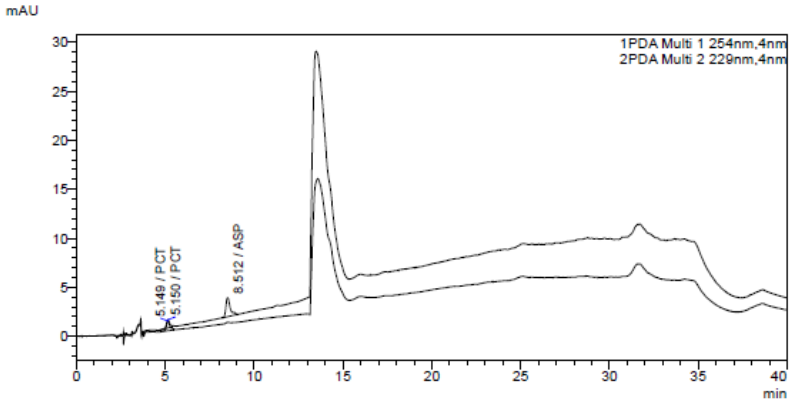
ID#4 Compound Name: ASP

Title	RT	Sample Name	Area	Height	Tailing	Plate	Rs	SF	S/N	Peak Purity Index
SST-6.lcd	8.598	SST-5	8274	595	1.183	8419	9.881	--	2.69	Not calculated
SST-2.lcd	8.592	SST-1	7845	579	1.141	8826	9.981	--	2.77	Not calculated
SST-3.lcd	8.586	SST-2	10671	629	1.928	7666	9.617	--	4.13	Not calculated
SST-4.lcd	8.596	SST-3	7837	578	1.149	9036	10.096	--	3.88	Not calculated
SST-5.lcd	8.582	SST-4	7826	572	1.084	9016	10.063	--	1.94	Not calculated
Average	8.591		8494	590	1.297	8593	9.928	--	3.08	Not calculated
%RSD	0.078		14.491	3.933	27.315	6.683	1.936	0.000	29.384	0.000

- Detection Limit

Sample Name : LOD1
Sample ID : LOD1
Data Filename : LOD1.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 180523.lcb
Vial # : 1-5
Injection Volume : 20 uL
Date Acquired : 18/5/2566 13:23:03
Date Processed : 18/5/2566 15:35:50
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

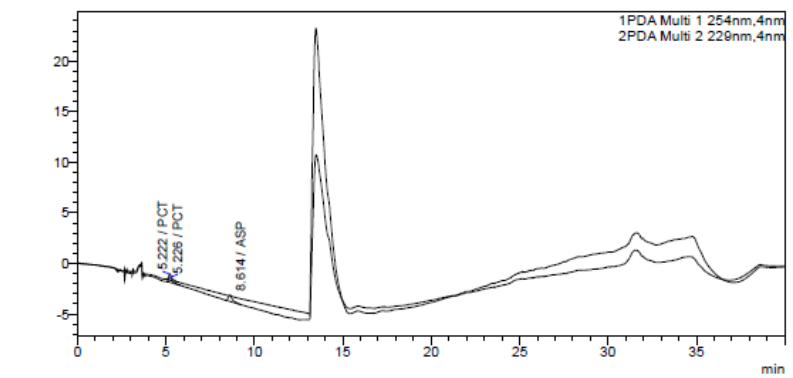
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.149	PCT	13207	--	3991	985	--	--	4.59
Total			13207			985			

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.150	PCT	11168	--	4020	847	--	--	3.32
2	8.512	ASP	32595	1.66	7442	1929	9.34	0.65	7.55
Total			43763			2776			

Sample Name : LOD3
Sample ID : LOD3
Data Filename : LOD3.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 180523.lcb
Vial # : 1-7
Injection Volume : 20 uL
Date Acquired : 18/5/2566 14:44:16
Date Processed : 18/5/2566 15:36:14
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.222	PCT	7880	1.74	4467	594	--	--	2.65
Total			7880			594			

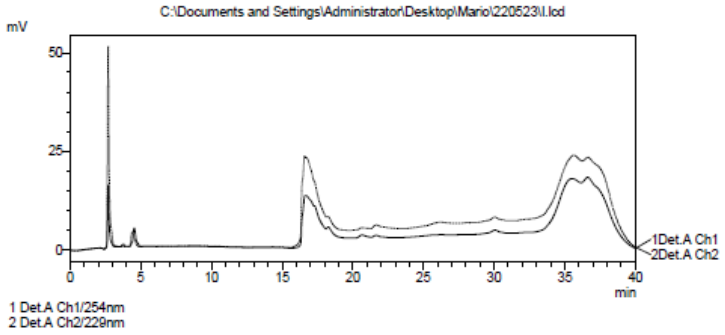
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.226	PCT	6964	1.77	4248	512	--	--	1.24
2	8.614	ASP	10219	1.79	7588	618	9.46	0.65	1.50
Total			17183			1130			

- Samples
  a. Sample I-I6

Acquired by : Mario
Sample Name : I
Sample ID : I
Vail # : 17
Injection Volume : 20 uL
Data File Name : I.lcd
Method File Name : MA24OT 220523.lcm
Batch File Name : MA24OT 220523.lcb
Report File Name : LC Peak Table(DetA-Ch1).lcr
Data Acquired : 23/5/2566 13:48:01
Data Processed : 23/5/2566 14:28:04

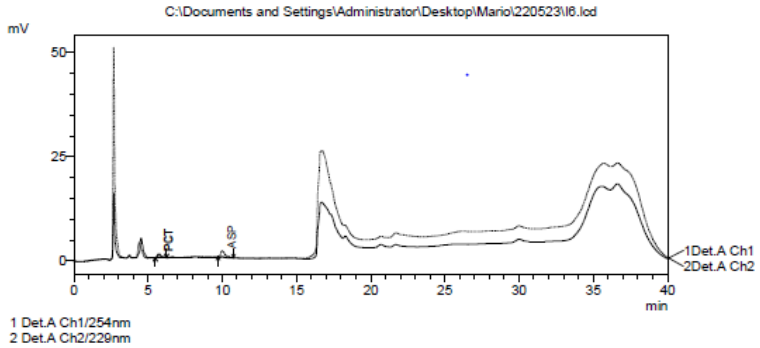
<Chromatogram>



No peaks detected at retention time of ASP and PCT

Acquired by : Mario
Sample Name : I6
Sample ID : I6
Vail # : 19
Injection Volume : 20 uL
Data File Name : I6.lcd
Method File Name : MA24OT 220523.lcm
Batch File Name : MA24OT 220523.lcb
Report File Name : LC Peak Table(DetA-Ch1).lcr
Data Acquired : 23/5/2566 15:08:59
Data Processed : 23/5/2566 15:49:02

<Chromatogram>



PeakTable

Detector A Ch1 254nm

Peak#	Ret. Time	Name	Area	Resolution	Tailing Factor	Theoretical Plates	Height
1	5.717	PCT	15453	0.000	1.688	3036	911
Total			15453				911

PeakTable

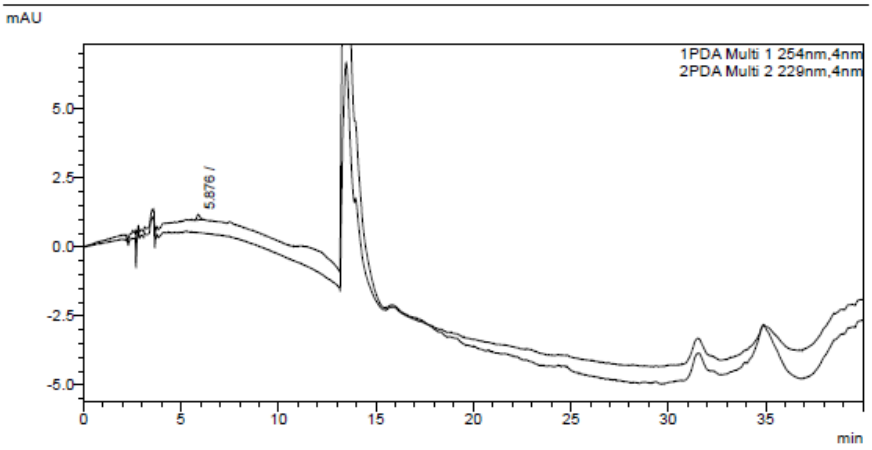
Detector A Ch2 229nm

Peak#	Ret. Time	Name	Area	Resolution	Tailing Factor	Theoretical Plate	Height
1	5.718	PCT	11442	0.000	1.732	3244	678
2	9.989	ASP	31114	9.561	1.825	6598	1468
Total			42556				2147

b. Sample J-J6

Sample Name : J
Sample ID : J
Data Filename : J.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-2
Injection Volume : 20 uL
Date Acquired : 19/5/2566 12:40:21
Date Processed : 19/5/2566 18:07:49
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

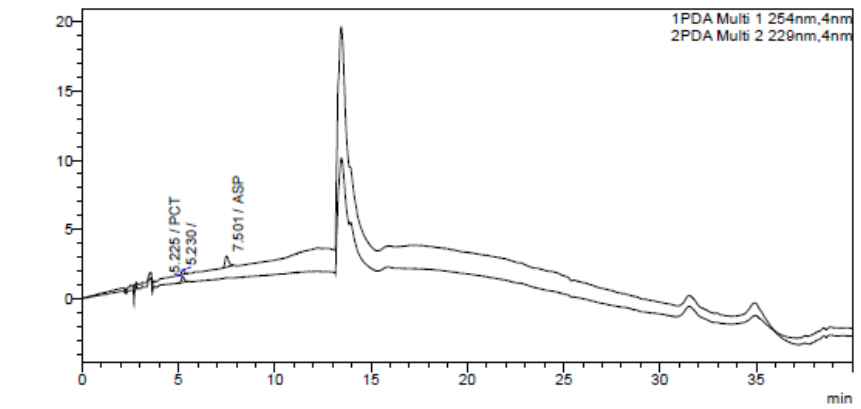
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.876		2218	1.30	6246	200	--	--	0.84
Total			2218			200			

Sample Name : J6
Sample ID : J6
Data Filename : J6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-3
Injection Volume : 20 uL
Date Acquired : 19/5/2566 13:20:58
Date Processed : 19/5/2566 18:07:50
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.225	PCT	4147	1.20	5396	394	--	--	2.62
Total			4147			394			

PDA Ch2 229nm

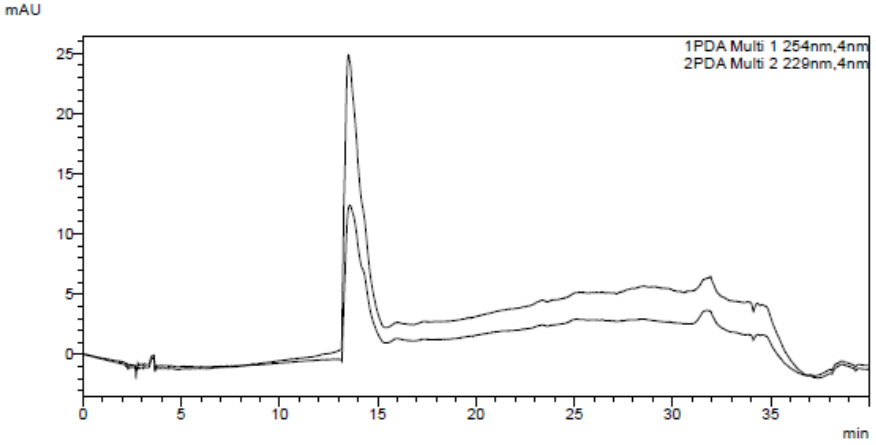
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.230		3496	1.16	5497	333	--	--	1.15
2	7.501	ASP	10378	1.24	7681	794	7.27	0.43	2.74
Total			13873			1127			

c. Sample IS-IS6

Sample Name : IS
Sample ID : IS
Data Filename : IS.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-6
Injection Volume : 20 uL
Date Acquired : 17/5/2566 14:39:42
Date Processed : 17/5/2566 17:27:20

Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

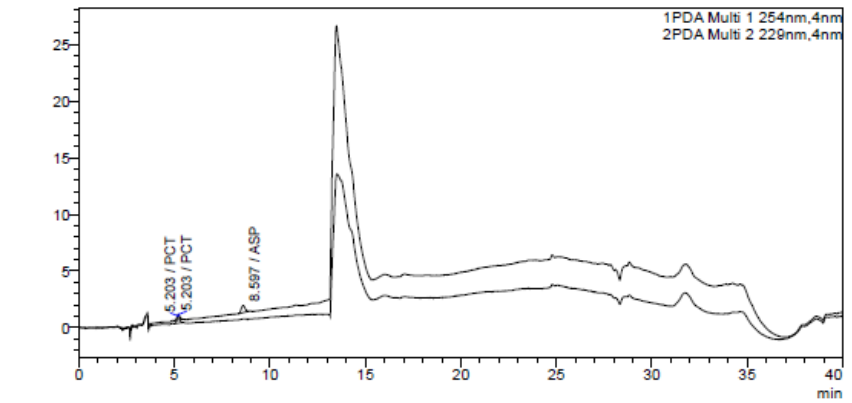
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

Sample Name : IS6
Sample ID : IS6
Data Filename : IS6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-7
Injection Volume : 20 uL
Date Acquired : 17/5/2566 15:20:18
Date Processed : 17/5/2566 17:27:21

Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.203	PCT	6238	1.16	5554	615	--	--	2.89
Total			6238			615			

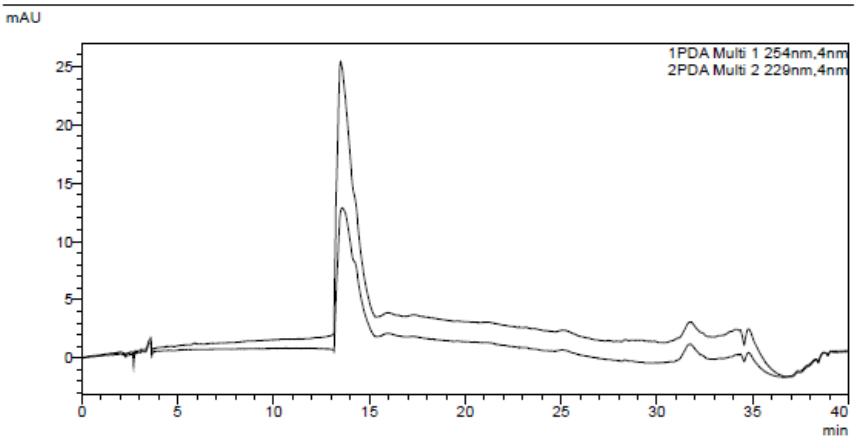
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.203	PCT	4978	1.08	5766	495	--	--	1.41
2	8.597	ASP	8521	1.11	8870	625	10.62	0.65	1.78
Total			13499			1121			

d. Sample IT-IT6

Sample Name : IT
Sample ID : IT
Data Filename : IT.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-8
Injection Volume : 20 uL
Date Acquired : 17/5/2566 16:00:54
Date Processed : 17/5/2566 17:27:22
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

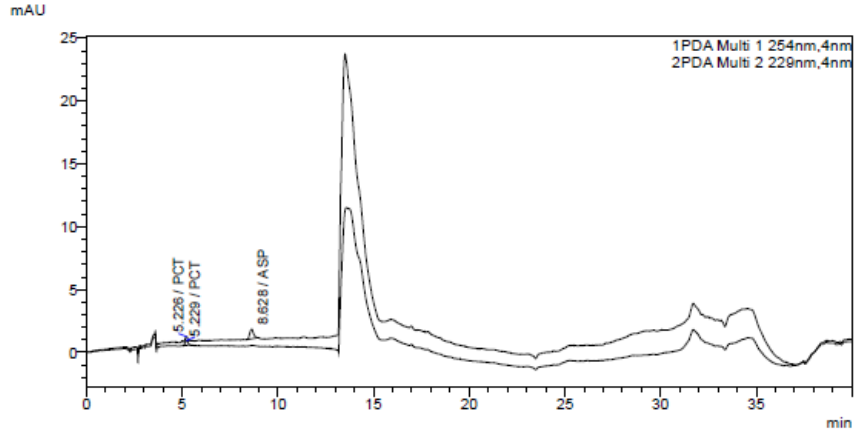
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

Sample Name : IT6
Sample ID : IT6
Data Filename : IT6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-9
Injection Volume : 20 uL
Date Acquired : 17/5/2566 16:41:30
Date Processed : 17/5/2566 17:27:23
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.229	PCT	3993	1.07	5939	409	--	--	1.96
Total			3993			409			

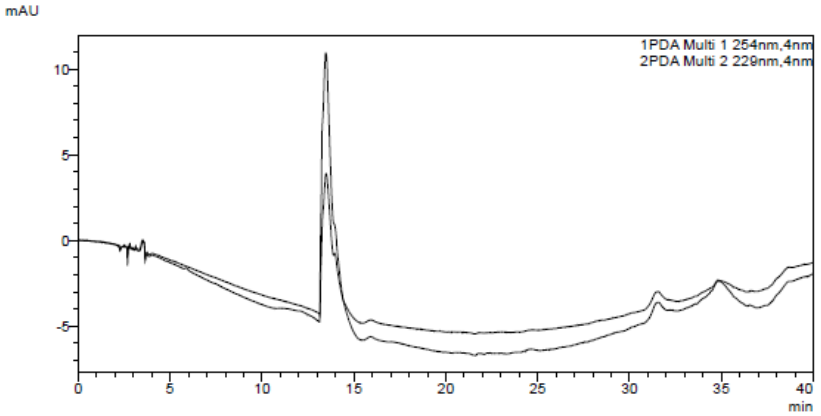
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.226	PCT	3496	1.10	5769	346	--	--	1.11
2	8.628	ASP	10546	1.14	8799	762	10.58	0.65	2.45
Total			14041			1108			

e. Sample JS-JS6

Sample Name : JS
Sample ID : JS
Data Filename : JS.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-6
Injection Volume : 20 uL
Date Acquired : 19/5/2566 15:22:46
Date Processed : 19/5/2566 18:07:53
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

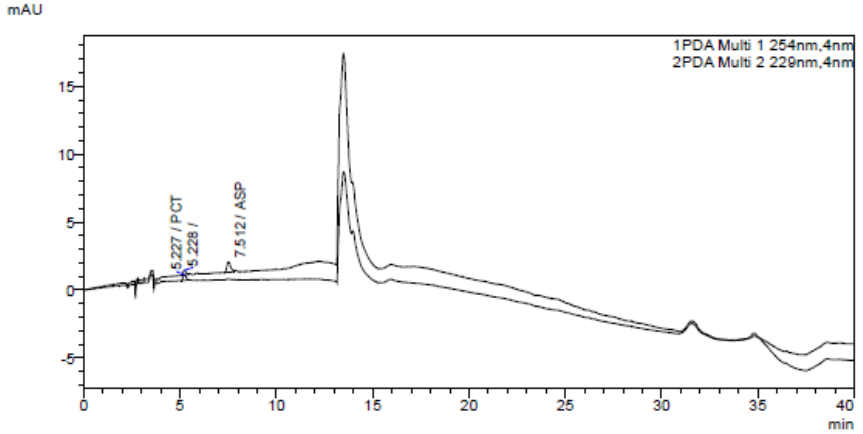
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

Sample Name : JS6
Sample ID : JS6
Data Filename : JS6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-7
Injection Volume : 20 uL
Date Acquired : 19/5/2566 16:03:21
Date Processed : 19/5/2566 18:07:54
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.227	PCT	4510	1.22	5348	423	--	--	3.03
Total			4510			423			

PDA Ch2 229nm

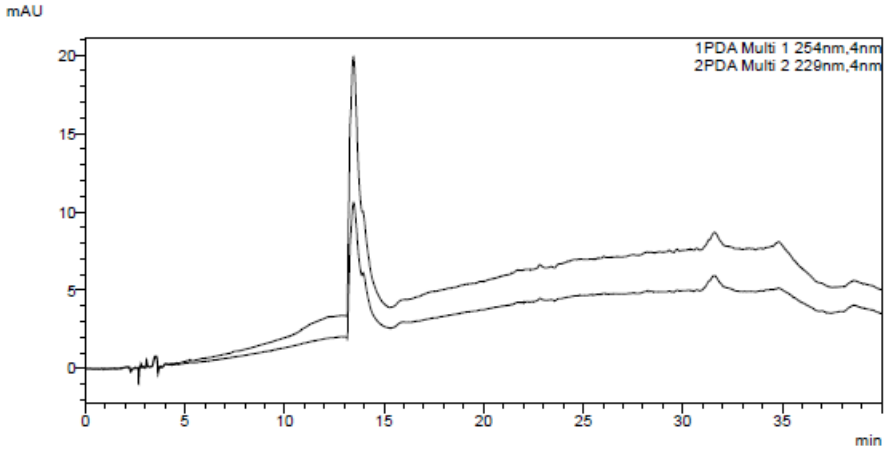
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.228		3893	1.19	5282	357	--	--	1.24
2	7.512	ASP	10724	--	6934	766	7.04	0.44	2.67
Total			14617			1123			

f. Sample JT-JT6

Sample Name : JT
Sample ID : JT
Data Filename : JT.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-8
Injection Volume : 20 uL
Date Acquired : 19/5/2566 16:43:58
Date Processed : 19/5/2566 18:07:55

Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

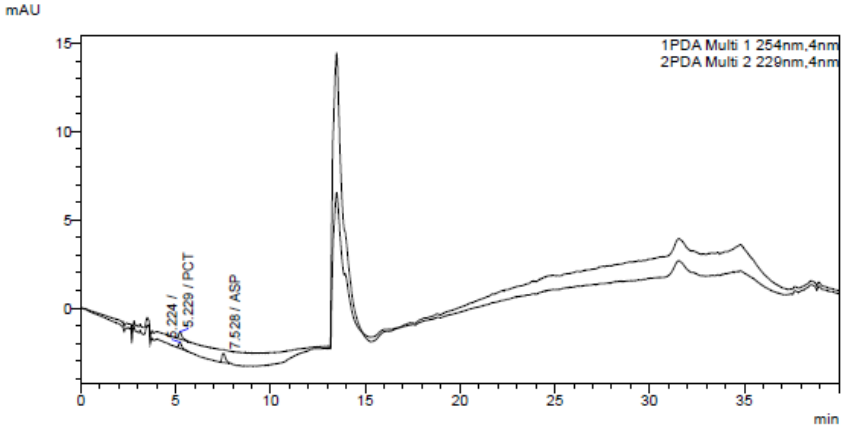
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

Sample Name : JT6
Sample ID : JT6
Data Filename : JT6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-9
Injection Volume : 20 uL
Date Acquired : 19/5/2566 17:24:34
Date Processed : 19/5/2566 18:07:56

Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.229	PCT	5106	1.83	4595	416	--	--	2.58
Total			5106			416			

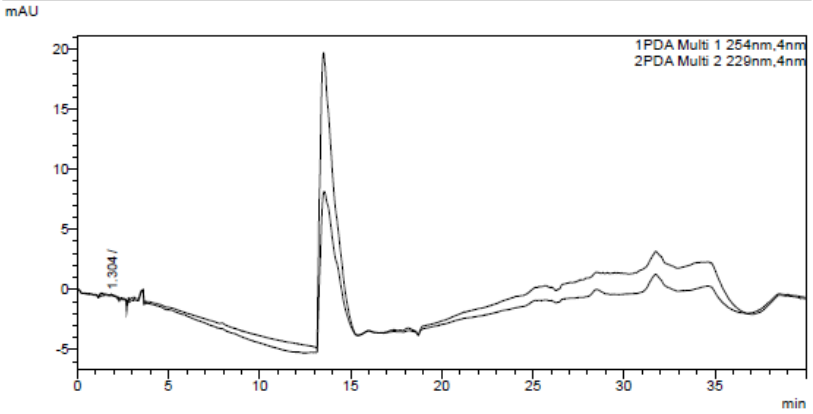
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.224		4021	1.97	4810	345	--	--	1.19
2	7.528	ASP	6708	1.25	7756	506	7.16	0.44	1.75
Total			10729			851			

g. Sample S-S6

Sample Name : S
Sample ID : S
Data Filename : S.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-4
Injection Volume : 20 uL
Date Acquired : 17/5/2566 13:18:31
Date Processed : 17/5/2566 17:27:18
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

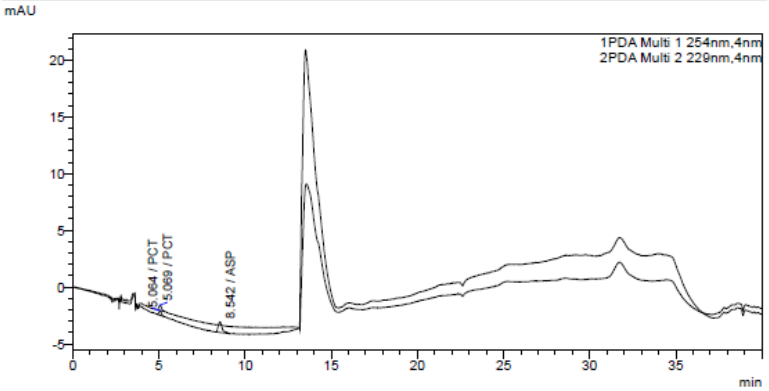
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	1.304		2355	1.08	241	183	--	--	0.67
Total			2355			183			

Sample Name : S6
Sample ID : S6
Data Filename : S6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 160523.lcb
Vial # : 1-5
Injection Volume : 20 uL
Date Acquired : 17/5/2566 13:59:06
Date Processed : 17/5/2566 17:27:19
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.089	PCT	4852	1.21	5305	479	--	--	2.01
Total			4852			479			

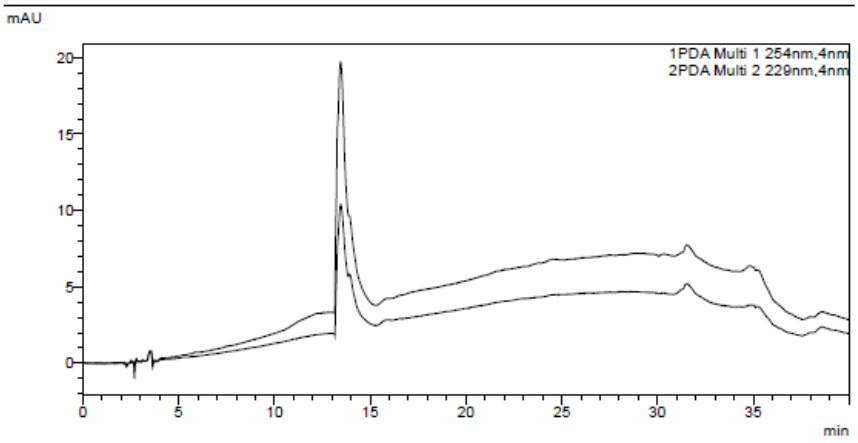
PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.064	PCT	3790	1.17	5367	387	--	--	1.64
2	8.542	ASP	14747	1.53	7696	921	10.45	0.69	3.90
Total			18537			1308			

h. Sample T-T6

Sample Name : T
Sample ID : T
Data Filename : T.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-4
Injection Volume : 20 uL
Date Acquired : 19/5/2566 14:01:35
Date Processed : 19/5/2566 18:07:51
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

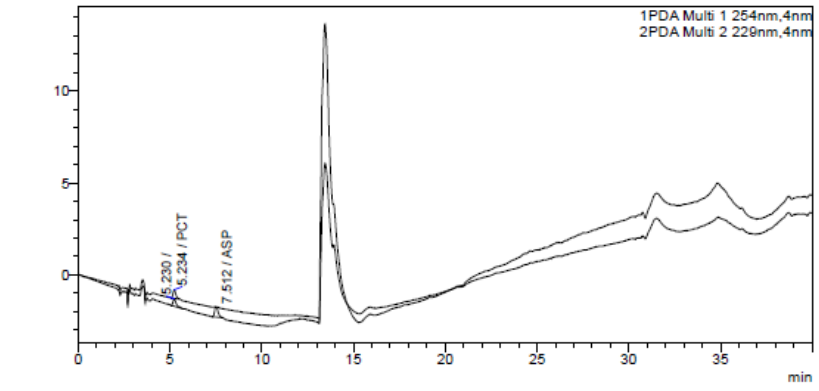
Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
Total									

Sample Name : T6
Sample ID : T6
Data Filename : T6.lcd
Method Filename : Trial MA OT2411 postrun.lcm
Batch Filename : MA OT 24-11 190523.lcb
Vial # : 1-5
Injection Volume : 20 uL
Date Acquired : 19/5/2566 14:42:11
Date Processed : 19/5/2566 18:07:52
Sample Type : Unknown
Acquired by : System Administrator
Processed by : System Administrator

<Chromatogram>



<Peak Table>

PDA Ch1 254nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.234	PCT	5045	1.29	5262	476	--	--	3.24
Total			5045			476			

PDA Ch2 229nm

Peak#	Ret. Time	Name	Area	Tailing	Plate	Height	Resolution	k'	S/N
1	5.230		5160	1.85	4868	415	--	--	1.51
2	7.512	ASP	7109	1.24	7587	550	7.08	0.44	1.99
Total			12269			965			

Appendix C

- Box-m test

Use of F Value	Probability of F	Manual Selection
Test Results	**Test Results**	**Test Results**
Box's M: 590.072	Box's M: 1524.604	Box's M: 1453.267
F — Approx.: 3.342	F — Approx.: 1.831	F — Approx.: 1.756
df1: 147	df1: 546	df1: 462
df2: 17451.383	df2: 16281.557	df2: 8793.880
Sig.: <.001	Sig.: <.001	Sig.: <.001
Tests null hypothesis of equal population covariance matrices.	Tests null hypothesis of equal population covariance matrices.	Tests null hypothesis of equal population covariance matrices.

- Eigenvalue of the model

Use of F Value

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	6.092[a]	44.1	44.1	.927
2	4.667[a]	33.8	77.8	.907
3	2.848[a]	20.6	98.4	.860
4	.182[a]	1.3	99.8	.392
5	.034[a]	.2	100.0	.181
6	.000[a]	.0	100.0	.010

a. First 6 canonical discriminant functions were used in the analysis.

Probability of F

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	6.092[a]	44.1	44.1	.927
2	4.667[a]	33.8	77.8	.907
3	2.848[a]	20.6	98.4	.860
4	.182[a]	1.3	99.8	.392
5	.034[a]	.2	100.0	.181
6	.000[a]	.0	100.0	.010

a. First 6 canonical discriminant functions were used in the analysis.

Manual Selection

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	7.095[a]	41.7	41.7	.936
2	5.822[a]	34.2	75.9	.924
3	3.560[a]	20.9	96.8	.884
4	.328[a]	1.9	98.7	.497
5	.121[a]	.7	99.4	.329
6	.057[a]	.3	99.7	.233
7	.044[a]	.3	100.0	.206

a. First 7 canonical discriminant functions were used in the analysis.

- Linear Function of each model

Use of F Value

Standardized Canonical Discriminant Function Coefficients

	Function 1	2	3	4	5	6
Var8	-.221	-.271	-.320	1.328	-.512	.023
Var13	-.025	-.420	.519	.322	1.283	.205
Var15	-.017	.765	-.617	.156	.516	-.321
Var20	-.039	.481	-.333	.052	.052	.885
Var24	-.156	.889	.835	-1.146	-.915	.107
Var25	.881	.325	.688	-.142	-.070	.153

Probability of F

Standardized Canonical Discriminant Function Coefficients

	Function 1	2	3	4	5	6	7
Var4	-.081	-.257	.297	.709	.390	1.257	-.447
Var8	-.215	-.303	-.665	.072	1.206	-.159	.303
Var10	-.935	-.332	-.778	-.117	-.205	1.210	-.159
Var11	.522	.030	.422	-.645	.096	.176	.422
Var13	-.100	-.709	.685	1.143	-.139	.508	-.666
Var16	-.065	-.222	.566	1.431	-.746	-.125	.792
Var17	.292	.448	.139	-.426	-.021	-.238	-.220
Var18	1.172	.299	.881	-.490	-.223	.131	.205
Var23	.651	-.531	1.436	-.282	.217	.681	.047
Var24	.958	1.243	-.127	-1.309	.371	-.192	.037
Var25	.530	.682	.872	-1.873	.952	1.427	-.212
Var28	-.741	.231	.325	2.645	-1.515	-1.371	.560

Manual Selection

Standardized Canonical Discriminant Function Coefficients

	Function						
	1	2	3	4	5	6	7
Var2	-.168	.035	.239	-1.487	.111	1.127	-.508
Var9	-.236	.713	.201	1.201	.513	.435	1.249
Var10	-.806	-.065	-1.043	.791	-1.290	-.163	.453
Var11	.313	.066	.525	.217	-.418	.234	-.477
Var12	.260	.108	.808	.356	.697	1.711	-.306
Var13	-.102	-.520	.408	-.618	-.013	-.266	.729
Var14	-.046	.136	-.614	1.185	.389	.776	1.182
Var17	.149	.400	.072	.500	-.101	.021	.043
Var18	.871	.175	1.348	-.997	.473	1.483	-.905
Var22	-.478	.251	.106	.054	-1.789	.660	1.980
Var24	.605	1.276	-.002	1.058	1.216	.294	.414
Var25	-.263	.964	1.295	.053	.694	1.270	.051
Var26	.017	.381	-.339	.374	.392	-.388	-.074
Var27	.225	.545	-.307	1.334	2.030	-.315	-.966

REFERENCES

1. Organization WH. WHO GLOBAL REPORT ON TRADITIONAL AND COMPLEMENTARY MEDICINE 2019. Organization WH, editor. Luxembourg: World Health Organization; 2019.
2. KRITERIA DAN TATA LAKSANA PENDAFTARAN OBAT TRADISIONAL,

OBAT HERBAL TERSTANDAR DAN FITOFARMAKA HK.00.05.41.1384 (2005).
3. Calahan J, Howard D, Almalki A, Gupta M, Calderon A. Chemical Adulterants in Herbal Medicinal Products: A Review. Planta medica. 2016;82.
4. Chan CK, Tan LT, Andy SN, Kamarudin MNA, Goh BH, Kadir HA. Anti-neuroinflammatory Activity of Elephantopus scaber L. via Activation of Nrf2/HO-1 Signaling and Inhibition of p38 MAPK Pathway in LPS-Induced Microglia BV-2 Cells. Front Pharmacol. 2017;8:397.
5. Ariffin SH, I AW, Hassan Y, Abd Wahab MS. Adulterated Traditional-Herbal Medicinal Products and Its Safety Signals in Malaysia. Drug Healthc Patient Saf. 2021;13:133-40.
6. Pratiwi R, Dipadharma RHF, Prayugo IJ, Layandro OA. Recent Analytical Method for Detection of Chemical Adulterants in Herbal Medicine. Molecules. 2021;26(21).
7. Philp M, Fu S. A review of chemical 'spot' tests: A presumptive illicit drug identification technique. Drug Test Anal. 2018;10(1):95-108.
8. Birkner N. How an FTIR Spectrometer Operates. In: Qian W, editor.: Chem.libretexts.org; 2022. p. 1844.
9. Chen Y, Zou C, Mastalerz M, Hu S, Gasaway C, Tao X. Applications of Micro-Fourier Transform Infrared Spectroscopy (FTIR) in the Geological Sciences--A Review. Int J Mol Sci. 2015;16(12):30223-50.
10. Owen AJ. Uses of Derivative Spectroscopy. Agilent Technology Inc.; 1995.
11. Biancolillo A, Marini F, Ruckebusch C, Vitale R. Chemometric Strategies for Spectroscopy-Based Food Authentication. Applied Sciences. 2020;10(18):6544.
12. Pharmacopoeia US. USP-NF. 1039 <Chemometrics>. United States: United States Pharmacopeia; 2022.
13. Brereton RG. Chemometric of Pattern Recognition: A John Wiley and Sons, Ltd., Publication; 2009.
14. Dharmastuti Cahya F, Ratna Asmah S, Respati Tri S, Abdul R. Application of Ftir-Atr Spectroscopy in Combination With Multivariate Analysis to Analyse Synthetic Drugs Adulterant in Ternary Mixtures of Herbal Medicine Products. Indonesian Journal of Pharmacy. 2022;33(1).
15. Cebi N, Yilmaz MT, Sagdic O. A rapid ATR-FTIR spectroscopic method for detection of sibutramine adulteration in tea and coffee based on hierarchical cluster and principal component analyses. Food Chemistry. 2017;229:517-26.
16. IBM. IBM SPSS Statistics Base 28. 2021.
17. Pharmacopoeia US. USP-NF. 1225<Validation of Compendial Procedures>. United States: United States Pharmacopoeia; 2022.
18. Rohman A, Setyaningrum D, Riyanto S. FTIR Spectroscopy Combined with Partial Least Square for Analysis of Red Fruit Oil in Ternary Mixture System. International Journal of Spectroscopy. 2014;2014:1-5.
19. Pavia DL, Lampman GM, Vyvyan JR. Introduction To Spectroscopy Fourth Edition. Australia: Brooks/Cole Cengage Learning; 2009.

20. Griffiths P, Haseth J. Fourier Transform Infrared Spectrometry, Second Edition. 2006.
21. Oliveri P, Forina M. Data Analysis and Chemometrics. Chemical Analysis of Food: Techniques and Applications. 2012:25.
22. Tharwat A, Gaber T, Ibrahim A, Hassanien AE. Linear discriminant analysis: A detailed tutorial. Ai Communications. 2017;30:169-90.
23. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. CJEM. 2006;8:19-20.
24. Pharmacopoeia US. <621>Chromatography. 2023.
25. Watson D, Raynle D, Majors R. Understanding and Improving Solid-Phase Extraction. LC GC Europe. 2014;27:645-52.

VITA

NAME	Mario Theodore
DATE OF BIRTH	09 November 1992
PLACE OF BIRTH	Surabaya, Indonesia
INSTITUTIONS ATTENDED	Airlangga University
HOME ADDRESS	Griya Kebraon Selatan FA-10, Surabaya, Indonesia