การพิสูจน์เอกลักษณ์และลำดับกรดอะมิโนของโปรตีนจากพืชวงศ์ขิง

นาย โสภณ แก้วทิพย์

IDENTIFICATION AND AMINO ACID SEQUENCES
OF PROTEINS FROM ZINGIBERACEOUS PLANTS

Mr. Sophon Kaeothip

Thesis Title	IDENTIFICATION AND AMINO ACID SEQUENCES OF PROTEINS FROM ZINGIBERACEOUS PLANTS
By	Mr. Sophon Kaeothip
Field of Study	Chemistry
Thesis Advisor	Assistant Professor Polkit Sangvanich, Ph.D.
Thesis Co-advisor	Chantragan Phiphobmongkol, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial Fulfillment of the Requirements for the Master's Degree

…………………………………..... Dean of the Faculty of Science

(Professor Piamsak Menasveta, Ph.D.)

THESIS COMMITEE

………………………………………………….. Chairman

(Professor Sophon Roengsumran, Ph.D.)

………………………………………….……. Thesis Advisor

(Assistant Professor Polkit Sangvanich, Ph.D.)

………………………………………….. Thesis Co-advisor

(Chantragan Phiphobmongkol, Ph.D.)

…………………………………………….. Member

(Associate Professor Amorn Petsom, Ph.D.)

…………………………………………….. Member

(Associate Professor Chaiyo Chaichantipyuth, Ph.D.)

โสภณ   แก้วทิพย์ : การพิสูจน์เอกลักษณ์และลำดับกรดอะมิโนของโปรตีนจาก
พืชวงศ์ขิง (IDENTIFICATION AND AMINO ACID SEQUENCES
OF PROTEINS FROM ZINGIBERACEOUS PLANTS) อ. ที่ปรึกษา : ผศ.ดร.
พลกฤษณ์ แสงวณิช อ. ที่ปรึกษาร่วม : ดร. จันทรกานต์ พิภพมงคล, 128 หน้า
ISBN 974-53-1499-4

        พืชวงศ์ขิงเป็นพืชสมุนไพรที่สำคัญชนิดหนึ่งในประเทศไทย ในสมัยก่อนนั้นการจำแนกหรือพิสูจน์
เอกลักษณ์พืชแต่ละชนิดนั้นโดยส่วนใหญ่จะใช้ลักษณะทางพฤกษศาสตร์ แต่อย่างไรก็ตามพบว่าวิธีการดังกล่าว
นั้นค่อนข้างที่จะยุ่งยากและซับซ้อน ในการศึกษาครั้งนี้มีวัตถุประสงค์ที่ทำการพิสูจน์เอกลักษณ์ของพืชโดยใช้
โปรตีนที่มีอยู่ในพืชนั้นๆ ด้วยเทคนิคเจลอิเล็กโตรโฟเรซิสแบบ 2 มิติ   แผ่นเจลของพืชแต่ละชนิดนั้นจะให้
จุดโปรตีนที่แตกต่างกันซึ่งก็เพียงพอที่จะจำแนกพืชแต่ละชนิดในวงศ์ขิงได้ จุดของโปรตีนที่แตกต่างกันนั้นจะใช้
เป็นโปรตีนเครื่องหมายในการจำแนกพืชแต่ละชนิด โปรตีนส่วนใหญ่ของพืชวงศ์ขิงจะแสดงจุดอยู่ในช่วงมวล
13-20 กิโลดาลตัน และบริเวณที่เป็นกรดเล็กน้อยบนแผ่นเจล จากการหาลำดับกรดอะมิโนของโปรตีนด้วย
เทคนิคทางแมสสเปคโทรเมทรีและสืบค้นจากข้อมูลพื้นฐานโปรตีนพบว่า โปรตีนพื้นฐานจุด   A23 มีลำดับ
กรดอะมิโนของเปปไทด์ 4 เส้น ตรงกับลำดับกรดอะมิโนของโปรตีน Glutelin type-B2 [precursor] แต่พบว่า
น้ำหนักโมเลกุลและค่า pI ของที่คำนวณได้ของโปรตีนจุด A23 นี้มีค่าน้อยกว่าโปรตีน Glutelin type-B2
[precursor] ซึ่งเป็นไปได้ที่ว่าโปรตีนจุด A23 นั้นอาจจะเป็นหน่วยหนึ่งของโปรตีน Glutelin type-B2
[precursor]เนื่องจากโปรตีนตัวนี้มีลักษณะเป็นหน่วยย่อย 6 หน่วยด้วยกัน โปรตีนพื้นฐานจุด D12 มีลำดับ
กรดอะมิโนของเปปไทด์ 2 เส้นใกล้เคียงกับลำดับกรดอะมิโนของโปรตีน   แมนโนสสเปซิฟิก เล็กติน
(Mannose-specific lectin) แต่พบว่ามวลและค่า pI ของโปรตีนที่ได้นั้นไม่ตรงกัน ซึ่งอาจจะเป็นไปได้ที่โปรตีน
จุด D12 เป็นโปรตีนตัวใหม่ที่อยู่ในกลุ่มของโปรตีน lectin โปรตีนจุด Z1 และ Z2 มีลำดับกรดอะมิโนตรงกับ
ลำดับกรดอะมิโนของโปรตีนในกลุ่ม heat shock protein ของพืชชนิดอื่นจำนวน 1 และ 2 เส้น ตามลำดับ ซึ่ง
อาจเป็นไปได้ที่ว่าโปรตีนทั้งสองจุดนั้นเป็นโปรตีนตัวใหม่ในกลุ่มของ heat shock protein   จากการทดสอบ
ฤทธิ์ทางชีวภาพของสารสกัดโปรตีนที่ได้จากพืชวงศ์ขิงด้วยวิธี Mg(NP-40)A พบว่ามีฤทธิ์ในการทำให้เลือด
แข็งตัว (Hemagglutination) ซึ่งสารสกัดโปรตีนที่มีฤทธิ์แรงที่สุดนั้นได้มาจากเหง้าของว่านสาลิกาลิ้นทอง

ภาควิชา...........เคมี....................   ลายมือชื่อนิสิต..
สาขาวิชา.........เคมีอินทรีย์.........   ลายมือชื่ออาจารย์ที่ปรึกษา...........................
ปีการศึกษา.......2547...................   ลายมือชื่ออาจารย์ทีปรึกษาร่วม.....................

4672478923 : MAJOR CHEMISTRY

KEY WORD : ZINGIBERACEAE/ 2D-GEL ELECTROPHORESIS/ PROTEIN/ MASS SPECTROMETRY/ TANDEM MASS SPECTROMETRY

Zingiberaceous plants are important natural herb in Thailand. Previously, the identification of Zingiberaceous plants was mainly performed by their morphological origin. However, these methods were difficult to identify. The present study aimed to identify different species of Zingiberaceous plants using a proteome approach by two-dimensional gel electrophoresis (2-DE) and amino acid sequence by mass spectrometry. 2-DE gels of different Zingiberaceous plant samples contain sufficient different protein spots to permit easy discrimination. These "marker protein" will be used to identify of each species. Many proteins in Zingiberaceous plants are displayed in mass range 13.00-20.00 kDa and small acidic region on 2-DE gels. From amino acid sequence by tandem MS and amino acid sequence database searching, the amino acid sequence of common protein spot A23 is shown four-peptide sequence similar to partial amino acid residue of Glutelin type-B2 [precursor]. However, the experimental values of mass and pI were lower than theoretical values. Since Glutelin type-B2 [precursor] is a six subunit (hexamer), protein spot A23 could be subunit of Glutelin type-B2 [precursor]. Some parts of the amino acid sequence of two-peptide sequences from common protein spot D12 were matched to Mannose-specific lectin but mass and pI value was not matching. This protein could be a new protein in the same class of lectin. The amino acid sequences from protein spots Z1 and Z2 are similar to Class I heat shock protein. Crude proteins from Zingiberaceous plants with Mg(NP-40)A exhibited potent hemagglutinating activity. Crude protein from *Curcuma* sp. (Saligalinthong) exhibited strongest hemagglutinating activity.

Department.......Chemistry..................   Student's signature.......................................

Field of study....Organic Chemistry...   Advisor's signature......................................

Academic year.....2004......................   Co-advisor's signature.................................

ACKNOWLEDGEMENTS

CONTENTS

LIST OF FIGURES

Figure   Page

Figure   Page

Figure   Page

LIST OF TABLES

Figure   Page

Figure   Page

LIST OF ABBREVIATIONS

2-DE	Two-dimensional gel electrophoresis
µg	microgram
µl	microliter
ACN	Acetronitrile
APS	Ammonium persulfate
Ar	Argon
Bis	*N,N'*-methylenebisacrylamide
°C	degree Celsius
C	Crosslinking factor [%]
CAD	Collision activated dissociation
CCA	α-Cyano-4-hydroxycinnamic acid
CHAPS	3-(3-cholamidopropyl)dimethylammonio-1-propane sulfonate
CID	Collision induced dissociation
Da	Dalton
DC	Direct current
DHB	Dihydroxybenzoic acid
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI	Electrospray ionization
ESI-Q/TOF	Electrospray ionization Quadrupole time-of-flight
eV	electron Volt
IEF	Isoelectric focusing
IPG	Immobilized pH gradients
He	Helium
HPLC	High-performance liquid chromatography
kDa	Kilodalton
kVh	kilovolt-hour
LC	Liquid chromatography
LMW	Low molecular weight

nM	Nanomolar
mA	Milliampere
MALDI	Matrix Assisted Laser Desorption Ionization
MALDI-TOF	Matrix Assisted Laser Desorption Ionization/Time of flight
mg	Milligram
mg/ml	Milligram per milliliter
ml	Milliliter
min	Minute
mm	Millimeter
mM	Millimolar
MS	Mass spectrometry
MS/MS	Tandem Mass spectrometry
m/z	Mass per charge
PAGE	polyacrylamide-gel electrophoresis
pI	Isoelectric point
PMM	Peptide mass mapping
ppm	parts per million
PSD	Post source decay
Q/TOF	Quadrupole Time of flight
RF	Radiofrequency
rpm	Revolutions per minute
SDS	Sodium-dodecyl sulfate
SDS-PAGE	Sodium-dodecyl sulfate polyacrylamide-gel electrophoresis
T	Total acrylamide concertration [%]
TCA	Trichloroacetic acid
TEMED	$N,N,,N',N'$-tetramethylethylenediamine
TFA	Trifluoroacetic acid
TOF	Time of flight
Tris	Tris(hydroxymethyl)-aminoethane
UV	Ultraviolet spectroscopy
V	Volt

CHAPTER I

INTRODUCTION

Zingiberaceae family is one of the largest families of the plant kingdom. It is important natural resources that provide many useful products for food, spices, medicines, dyes, perfume and aesthetics to man. *Zingiber officinale* (1), for example, has been used for many years as spices and in traditional forms of medicine to treat a variety of diseases.

In Thailand Zingiberaceous plants are more than 200 species (2) in this family and each has very similar botanical characteristics such as flowers, leaves and rhizomes. It is very difficult to distinguish one from another or clearly identify them from their physical appearances expect by botanists and herbalists (3). For example, *Curcuma aeruginosa* has similar leaves with *Curcuma aromatica*, the only difference is texture on leaves and the color of their rhizomes. Rhizomes in *Curcuma* group also have similar color and physical appearance (2). In addition to botanical characteristics (taxonomy), genetic (DNA) sequences can also be used to distinguish this type of plants. The latter is rather complicated, time consuming and expensive although it seems to give good results. From investigated, proteomics has previously been used to identify different morphologically similar organisms, including bacteria, yeast and fungi. Proteomic is defined as the study protein expressed by genome and tissue[4]. Then, we interested whether proteomic approach could be used to identify different species in Zingiberaceae family. From assume, in general plants in the same family should have similar protein characteristics. However, some of their proteins could be different which is unique for each plant. Therefore, it is interesting to study protein pattern performed from this family by identifying their pI and molecular weights. Separation of proteins will be by 2-dimensional gel electrophoresis. Amino acid sequence will be used to verify the extracted proteins, which will be stored in database for reference.

Objectives of this research

The objectives of this research are to identify protein characteristic of plants in Zingiberaceae by two-dimensional gel electrophoresis and determine amino acid sequences by Mass spectrometry

CHAPTER II

THEORETICAL AND LITERATURE REVIEW

Theoretical background

2.1 ZINGIBERACEAE

Zingiberaceae have been known as natural herbs because of their medical application. In the rural areas of Thailand, these plants have been widely used for treatments for sickness (3). From the results of studies by various organizations, Zingiberaceae can be used to treat for more than 50 symptoms. In food, Zingiberaceae are also used heavily because they add colors, exotic tastes and valuable nutrients into food. Other than that, they are main ingredients in dyes, perfumes and aesthetics. Recently, scientific study has sought to reveal the bioactive compounds of the rhizome. It has been found to be effective in the treatment of thrombosis, seasickness, migraine and rheumatism (5).

2.1.1 General Characteristics of the Family ZINGIBERACEAE (2, 6)

Perennial rhizomatous herbs. Leaves simple, distichously. Inflorescence terminal on the leafy shoot or on the lateral shoot. Flower delicate, ephemeral and highly modified. All parts of the plant aromatic. Fruit a capsule.

2.1.2 Habitats

Species of the Zingiberaceae are the ground plants of the tropical forest. They mostly grow in damp and humid shady places. They are also found infrequently in secondary forest. Some species can fully expose to the sun, and grow on high elevation.

2.1.3 Distribution

Zingiberaceae are distributed mostly in tropical and subtropical areas. The center of distribution is in South East Asia. The greatest concentration of genera and species is in the Malaysian region (Thailand, Indonesia, Malaysia, Singapore, Brunei, the Philippines and Papua New Guinea).

2.1.4 Chemical Constituents

Curcumin (yellow pigment), essential oil (artumerone, zingberene, borneol), valepotriates, alkaloids, terpenoids, flavanoids, steroids, sesquiterpenes, monoterpene, phenolic compounds, protein etc.

2.1.5 Representation of Zingiberaceae

For example of some species in Zingiberaceae

1) *Curcuma longa* **Linn.**

Common names: Turmeric, Curcuma

Thai names: Khamin Chan

Distinguished characters: Rhizome branched, bright yellow. Flower bracts greenish yellow. Coma bracts pale yellow.

Distribution in Thailand: Cultivated throughout the country.

Uses: Treament of pepticulcer, dyspepsia, indigestion, and flatulence

2) *Curcuma zedoaria* **Rosc.**

Common name: Zedoary

Thai names: Khamin Oi

Distinguished characters: The rhizome pale yellow or white inside. The coma bracts dark red to purple.

Distribution in Thailand: Cultivated throughout the country

Uses: Young rhizome is eaten as vegetable soup. Rhizomes are used in relief of stomachache and as a carminative. It is also used in perfumery. Leaves are used for flavoring fish and other food.

3) *Zingiber officinale* **Rosc.**

Common name: Ginger

Thai name: Khing

Distinguished characters: Rhizome pale yellow or white, flower geenish yellow, bracteoles pale green.

Distribution in Thailand: Cultivated throughout the country

Uses: Carminative, antiemetic, expectorant, and diaphoretic.

4) *Zingiber purpureum* Rosc. (*Z. cassumunar* Roxb.)

Common name: Cassumunar

Thai name: Phlai

Distinguished characters: Inflorescence ellipsoid or spindle shaped. Bracts dark purple. Rhizome bright greenish-yellow, bitter with a strong smell.

Distribution in Thailand: Cultivated.

Uses: Rhizome is used to treat fevers and intestinal disorder. Various lotions and decoctions applied to swellings, rheumatism, bruise, numb feet, and painful parts.

Curcuma longa

Curcuma aeruginosa

Zingiber ottensii

Curcuma sp. (Khamindum)

Curcuma sp. (Khantamala)

Cucuma parviflora

Figure 2.1 Plants in Zingiberaceae family

Curcuma sp. (Enleung)

Curcuma sparganifolia

Curcuma amarissima

Curcuma sp. (Payawan)

Curcuma sp. (Maleung)

Curcuma comosa

Curcuma comosa

Curcuma sp. (Haroynang)

Figure 2.1 Plants in Zingiberaceae family (continued)

Curcuma sp. (Saligalinthong)

Zingiber cassumunar

Alpinia niga.

Zingiber officinale

Curcuma zedoaria

Curcuma mangga

Figure 2.1 Plants in Zingiberaceae family (continued)

2.2 Sample Preparation (7, 8)

2.2.1 Cell Disintegration and Protein Extraction

The purpose of the present section is to describe methods of disrupting the cells and releasing the enzyme into an aqueous extract. There are many methods of cellular disintegration, for there are many type of cell. Choice of disruption depends on whether the sample is from cells, solid tissue, or other biological material and whether the analysis is targeting all proteins or just a particular subcellular fraction.

Plant tissues are quite different as mentioned and only a small fraction of the volume of plant tissue is truly intracellular, large vacuoles (being regarded here as extracellular) and intercellular spaces mean that on disruption much liquid is released, making additional extractant liquid almost unnecessary. The residue after centrifugation may occupy only 20-40% of the volume of the original plant tissue. Nevertheless, it may be important to use some added extractant liquid to control undesirable process during homogenization. These include acidification and oxidation of susceptible compounds. A particular problem with many plants is their content of phenolic compounds, which oxidize-mainly under the influence of endogenous phenol oxidase to from dark pigments. These pigments attach themselves to proteins and react covalently to inactivate many proteins. Two approaches to this problem are useful. Firstly, the inclusion of a thiol compound such as β-mercaptoethanol minimizes the action of phenol oxidase. Secondly, addition of powdered polyvinylpyrrolidine is often beneficial as it adsorbs the phenolic compounds.

2.2.2 Precipitation with Organic Solvent

The method of protein precipitation by water-miscible organic solvents has been employed since the early days of protein purification. Addition of a miscible solvent such as ethanol or acetone to an aqueous extract containing proteins has a variety of effects that, combined, lead to protein precipitation. The principal effect is the reduction in water activity. The solvating power of water for a charged, hydrophilic protein molecule is decreased as the concentration of organic solvent increases. This can be described in terms of reduction of the dielectric constant of the

solvent, or simply in terms of a bulk displacement of water, plus the partial immobilization of water molecules through hydration of the organic solvent. A two-dimensional representation of proteins in water-organic solvent mixture is shown in Figure 2.2 Here aggregation is occurring by interactions between opposite-charged areas on the protein's surfaces.
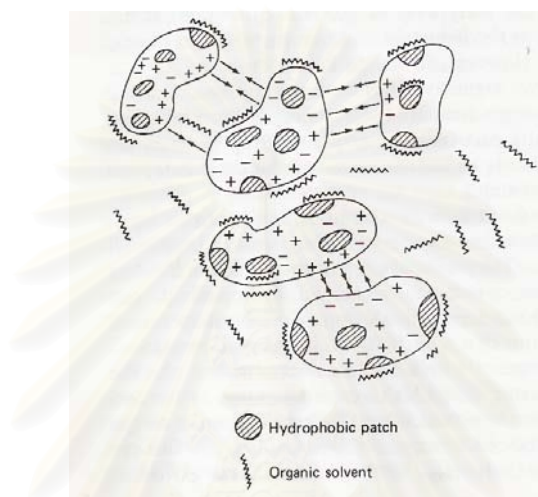
Figure 2.2 Aggregation of proteins by interactions in an aqueous-organic solvent mixture

2.3 Electrophoresis (9)

Electrophoretic separation techniques are at least as widely distributed as chromatographic methods. The main fields of application are biological and biochemical research, protein chemistry, pharmacology, forensic medicine, clinical investigations, veterinary science, food control as well as molecular biology. It will become increasingly important to be able to choose and carry out appropriate electrophoresis technique for specific separation problem.

Electrophoresis is the process of moving charged molecules in solution by applying an electric field across the mixture. Because molecules in an electric field move with a speed dependent on their charge, shape and size, electrophoresis has been extensively developed for molecular separations. Electrophoretic separations are carried out in free solution such as in capillary and free flow system, or in stabling media such as thin-layer plates, films or gels.

2.3.1 Gel electrophoresis (9, 10)

Gel electrophoresis is the most widely used method for proteomic research because it provides a powerful separation, quantization and characterization method for protein in complex mixture, including whole homogenates of cultured cells and tissues. Electrophoresis separations are carried out in gels such as agarose gel or polyacrylamide gel. In gel electrophoresis system: continuous and discontinuous. A continuous system uses only one buffer for the tanks and the gel. In a discontinuous system, first develop by Ornstein and Davis (1964), a non-restrictive large-pore gel called a stacking gel is layered on top of a separating (resolve) gel. The two gel layers are each mode with a different buffer, and the tank buffer from the gel buffer.

In a discontinuous system, when electrophoresis is started the ions and the protein begin migrating into the stacking gel. The proteins concentrate in a very thin zone, called the stack between the leading ion and the trailing ion. The proteins continue to migrate in the stack until they reach the separating gel. At that point, due to a pH or an ion charge the proteins become the trailing ion and unstack as they separate on the gel (see figure 2.3)

The discontinuous Laemmli system (Laemmli, 1970) (11), are a denaturing modification of Ornstein and Cavis (1964), in the most widely used system for research protein electrophoresis. The resolution of a Laemmli gel is excellent because treated peptides or proteins are concentrated in a stacking zone before entering the separating gel (Figure 2.3).
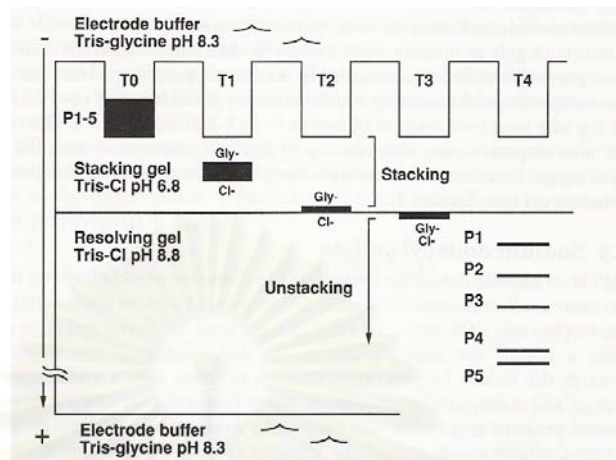
Figure 2.3 Explanation of Kohlrausch boundary migration under Laemmli multiphase
buffer conditions, Electrophoresis through the stacking (pH 6.8) and
resolving (pH 8.8) gels of a hypothetical sample containing five proteins
(P1-5) is shown as a function of time (T) T0: sample loading into well;
T1: partial stacking of sample between loading (Cl-) and trailing edge
(glycine) ions; T2: complete stacking of P1-5; T3: unstacking takes place
with increased ionization of glycine in the pH 8.8 resolving gel; T4:
resolution P1-5

For protein, most commonly used gel is acrylamide. Sodium dodecyl sulfate
polyacrylamide gel electrophoresis (SDS-PAGE) is a widely used analytical
technique for the separation and characterization of complex mixture of proteins and
peptide and for estimation of the relative molecular weight (M.W.). SDS-PAGE is
two major components are as fellows:

1) Sodium dodecyl sulfate (9)

Sodium dodecyl sulfate (SDS) is an anionic detergent that is used to denature
proteins, giving them all the same conformation properties and to prevent protein
interactions during electrophoresis. SDS solution in water, forms globular micelles
composed of 70-80 molecules with dodecyl hydrocarbon moiety in the core and the
sulfate head groups in the hydrophilic shell. SDS masks the charge of the proteins
themselves and the formed anionic complexes have a roughly constant net negative
charge per unit mass.

2) Polyacrylamide (12)

Polyacrylamide gels are chemically inert and mechanically stable. By chemical copolymerization of acrylamide monomers with a cross-linking reagent. It's usually *N,N'*-methylenebisacrylamide (Figure 2.4) and free radical-induced polymerization of acrylamide is catalyzed by ammonium persulfate and an accelerator (*N,N,N',N'*–tetramethylenediamine, TEMED). Polymerization should take place under an inert atmosphere since oxygen can act as a free radical trap. The polymerization is temperature dependent: to prevent incomplete polymerization the temperatures should be maintained above 20°C
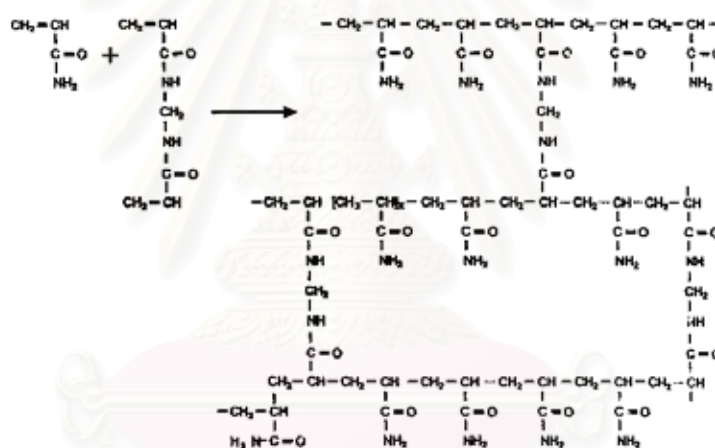
Figure 2.4 The polymerization reaction of acrylamide and methylenebisacrylamide

The total acrylamide concentration T and the degree of cross-linking C can be exactly and reproducibly control the pore size

$$\%T = \frac{\text{acrylamide (g) + bis-acrylamide (g)}}{100 \text{ ml}} \times 100\%$$

$$\%C = \frac{\text{bis-acrylamide (g)}}{\text{acrylamide (g) + bis-acrylamide (g)}} \times 100\%$$

In general, best resolution is achieved when the pore size approximates the protein molecular size (12). The %T of the gel determines the range over which the

protein M.W. is proportional to the rate of migration (Table 1). Gels may also be composed of a gradient of polyacrylamide (typically 5-20 %).

Table 2.1 Protein M.W. range resolved by SDS-PAGE

Polyacrylamide gel (%T)	M.W.range
6	30,000-200,000
8	20,000-175,000
10	15,000-150,000
12	10,000-100,000
15	6,000-50,000
5-20 gradient	6,000-250,000

2.3.2 Two-Dimensional Gel Electrophoresis (10, 13)

Two-dimensional gel electrophoresis (2-D electrophoresis) is nowadays frequently used in the laboratory. It is the most powerful protein separation technique and is applicable to soluble as well as membrane proteins. By combining isofocusing with sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), a high resolution and a high separation capacity is obtained. This technique sort's protein according to two independent properties in two discrete steps: the first dimension step, isoelectric focusing (IEF), separates proteins according to their isoelectric points (pI); the second dimension step, separates proteins according to their molecular weighs by SDS-PAGE. These properties make two-dimension polyacrylamide gel electrophoresis (2D-PAGE) the methods of choice for the analysis of very complex protein mixtures. More than thousand, radioactively labeled proteins can be resolved and detected in a single two-dimension gel (14).

Two-dimensional electrophoresis was first introduced by P.H.O'Farrell and J.Klose in 1975. In original technique, the first-dimension separation was performed in carrier ampholyte-containing polyacrylamide gels cast in narrow tube.

1) First-dimension isoelectric focusing (10, 15)

Isoelectric focusing (IEF) is an electrophoretic method that separates proteins according to their isoelectric points (pI). Proteins, enzymes and peptides are such amphoteric molecules. They carry either positive, negative, or zero net charge,

depending on the pH of their surroundings. The net charge of a protein is the sum of all negative and positive charges of the amino acid side chains and amino- and carboxyl-terminal but the three-dimensional configuration of the protein also plays a role. The isoelectrtic point (pI) is the specific pH at which the net charges of the protein are zero. Proteins are positively charged at pH values below their pI and negatively charged at pH above their pI. If the net charge of a protein is plotted versus the pH of its environment (Figure 2.5), a continuous curve that intersects the x-axis at the isoelectric point pI will result.
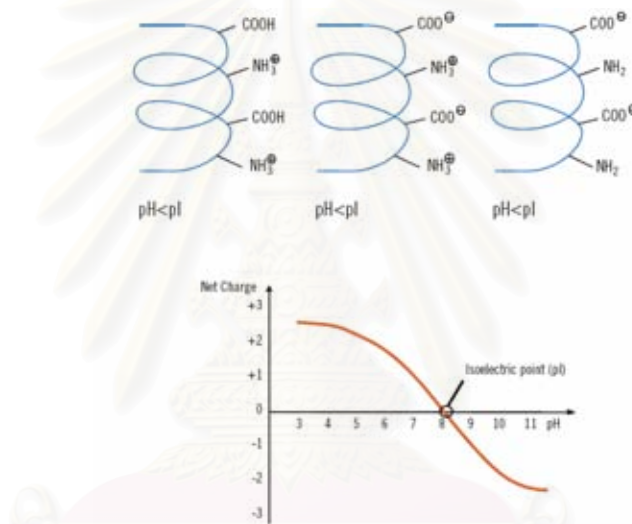
Figure 2.5 Plot of the net charge of a protein versus the pH of its environment. The point of intersection of the curve at the x-axis represents the isoelectric point of the protein

In a pH gradient and under the electric field, a protein will move towards the anode or the cathode until its reach a position on the pH gradient where its net charge is zero. If a protein should diffuse away from its pI, it immediately gains charge and migrate backs. This is the focusing effect of IEF, which concentrates proteins at their pI and allows proteins to be separated based on very small charge differences. In an isoelectric focusing gel, the pH conditions are established in polyacrylamide gel by two technique, carrier ampholytes (low molecular weight amphoretic species) and immobilines (acrylamide derivative). In the present, because of some limitations of the carrier ampholytes method, an alternative technique was developed: immobilized

pH gradients or IPG. An immobilized pH gradient (IPG) is created with acrylamide derivatives with buffering groups (acidic or basic buffering group), the Immobilines, by co-polymerization of the acrylamide monomers in a polyacrylamide gel.

The general structure of immobiline reagent is:

$CH_2=CH-CO-NH-R$,

Where, R = weakly acidic or basic buffering groups

To be able to buffer at a precise pH values, at least two different immobilines are necessary, an acid and base. Figure 2.6 is a graphic representation of the polyacrylamide matrix with attached buffering groups. The advantages of using IPG gel are allowing along focusing time to ensure the focusing of the analyst proteins, loading relatively large amount of proteins, available to purchasing IPG in variety of pH range and simplify the physical handing.
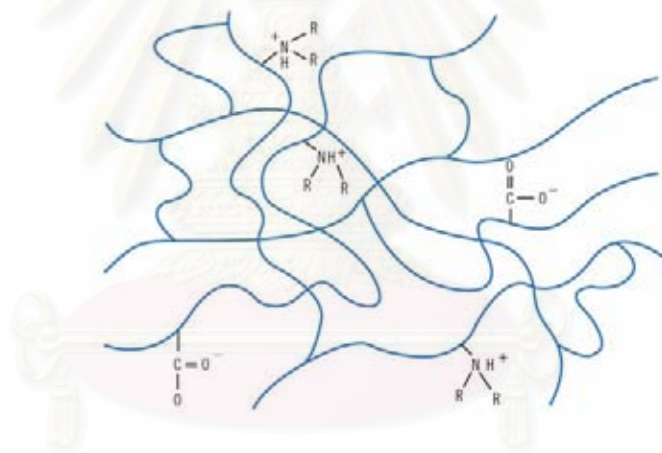
Figure 2.6 Immobilized pH gradient polyacrylamide gel matrix showing attached
buffering group

2) Second-dimension SDS-PAGE (9)

SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis) is an electrophoretic method for separating proteins, enzymes and polypeptides according to their molecular weights (M.W.). In this technique is performed in polyacrylamide gels containing sodium dodecyl sulfate (SDS). The intrinsic electrical charge of the sample proteins is not a factor in the separation due to the presence of SDS in to the sample and the gel. SDS is an anionic detergent, when loading in to the protein sample, the charge of the proteins is so well masked that anionic micelles with a

constant net charge per mass unit result: 1.4 g SDS per g protein. For separation, proteins have small molecular weight move more rapidly through the gel than larger proteins such that mixtures proteins can be separated.

2.4 Protein Detection (16)

After the electrophoresis run is complete, the gel must be analyzed qualitatively or quantitatively to answer analytical or experimental questions. Because most proteins and all nucleic acids are not directly visible, the gel must be processed to determine the location and amount of the separated molecules.

The mostly applied methods in the detection of separated proteins from gel electrophoresis are Coomassie blue staining and Silver-staining. Silver staining is the most sensitive non-radioactive method (below 1 ng). Silver staining is a complex, multistep process utilizing numerous reagents for which quality is critical. Coomassie blue staining, although 50- to 100- fold less sensitive than silver staining, is simple method and more quantitative than silver staining. Coomassie blue staining is preferable when relative amounts of protein are to be determined by densitometry. For another method used to detect protein in general laboratory such as Autoradiography and Fluorography, Negative Zinc, Fluorescent staining etc.

2.5 Data Analysis (10, 17)

In 2D gels, the molecular weight of the protein is determined based on its position the SDS-PAGE dimension, the pI is determined based on its position in the isoelectric focusing dimension, and it relative amount of protein is determined based on the density of the staining. In SDS-PAGE, the molecular weight scale can be calibrated with standard run at one edge of gel. The pI scale can be calibrated based on how the immobilized pH gradient strip was formed. One should be aware that these strips are available in linear and non-linear formats. In either case, plots of pH in the strip versus position relative to the end of the strip. The amount of protein can be estimated by the degree of staining of the gel band. This value is most accurately measured as a volume measurement that accounts for both the size of the band or spot and the density of staining. Scanning densitometers equipped with computer-controlled data recording system can make these measurements.

2.6 Mass Spectrometry (11, 18)

Mass spectrometry is a powerful technique for identifying unknown compounds, studying molecular structure and probing the fundamental principles of chemistry. In protein study, mass spectrometry is typically employed for protein primary structure analysis in the last stages of protein purification. Mass spectrometry not only allows the precise determination of the molecular weight of peptides and of proteins but also the determination of their sequences, especially when operate in tandem mode.

Principle of mass spectrometry (10, 18)

Mass spectrometer can be divided into three fundamental parts, namely the ion source, mass analyzer, and detector. The sample under investigation has to be introduced into the ion source of the instrument. Once inside the ion source the sample molecule are ionized to gas phase ions. These ions are extracted into the analyzer region of the mass spectrometer where they are separated according to their mass-to-charge ratios. The separated ions are detected and this signal sent to a data system where the m/z ratios are stored together with their relative abundance for presentation in the format of a mass spectrum.

The analyzer and detector of the mass spectrometer, and often the ionization source too, are maintained under high vacuum (10^{-4}-10^{-6} torr) to give the ions a reasonable chance of traveling from on end of the instrument to the other without any hindrance from air molecule, reduce the collision and loss of ions (Figure 2.7).

Sample introduction		Ion source	Mass analyzer	Detector		Data
			Vacuum system			System
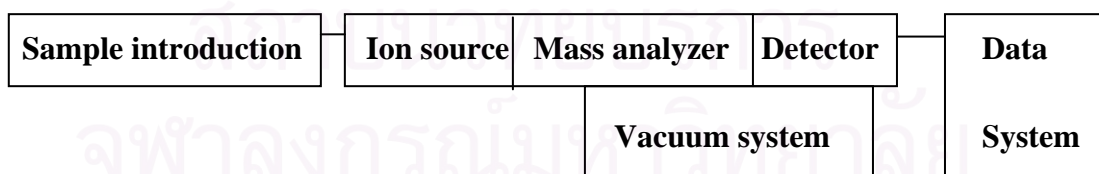
Figure 2.7 Mass spectrometer block diagram

There are several ionization techniques which suitable for biological analystes; matrix assisted laser desorption/ionization and electrospray ionization. Mass analyzers are widely used time-of-flight mass analyzer and quadrupole mass analyzer.

In this work, electrospray ionization and matrix assisted laser desorption/ionization was used as ionization technique quadrupole and time-of-flight mass analyzer is used to mass analyzers.

2.6.1 Matrix Assisted Laser Desorption/ Ionization (MALDI) (10, 19-20)

Karas and Hillenkamp first described Matrix Assisted Laser Desorption/Ionization in 1988 (19). This method is soft ionization technique of analyst sample. The analyst (biomolecule) is embedded in a crystal of matrix molecules absorbing at the laser wavelength (UV laser) of the instrument. By the energy, absorption the matrix is volatilized entraining the enclosed biomolecules in the gas phase, where they became ionized (Figure 2.8). The ions are than accelerated into the MS analyzer. As MALDI is a pulsed ionization technique, which is ideally coupled with a TOF analyzer.
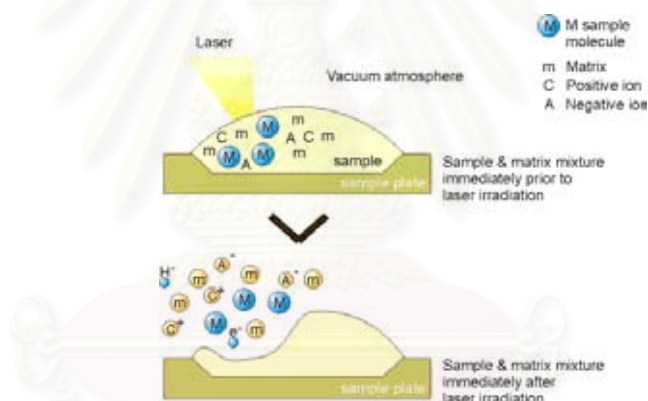
Figure 2.8 Matrix Assisted Laser Desorption/Ionization (MALDI) source.

MALDI can be generated positive- and negative-ion. The mass spectra of protein and peptides typically contain signals due to the soft ionization of MALDI mainly single charged molecules are formed, allowing determination of the molecular mass. The typical wavelength of UV laser utilized is 337 nm. Wide range matrixes for biological mass spectrometry applications have been adopted for use with UV lasers (Table 2.2).

Table 2.2 Some common MALDI matrices used in biological applications.

Matrix	Application
α-cyano-4-hydroxycinnamic acid	UV laser Peptide analysis & protein digests Analytes < 10 kDa
Sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid)	Analysis of large polypeptides& protein Analysts > 10 kDa
2,5-dihydroxybenzoic acid (2,5-DHB)	UV laser Protein digests & protein Oligasaccharides
2,4,6-trihydroxyacetophenone (THAP)	UV laser Oligonucleotide < 3 kDa

For peptides and small-molecular-mass proteins (<10,000 Da), good results are obtained with α-cyano-4-hydroxycinnamic acid (CCA), whereas high-mass proteins are analyzed with sinapinic acid. The use of 3-amino-4-hydroxybenzoic acid and 2,5-dihydroxybenzoic acid has been recommended for analysis of oligosaccharides.

2.6.2 Electrospray Ionization (ESI) (21-24)

Primarily Fenn and co-workers in the mid 1980s (20-21), an achievement for which Fenn won the 2002 Nobel Prize for Chemistry developed Electrospray Ionization (ESI). Initially exploited for the analysis of thermally fragile and high-molecular-weight materials including proteins and polymers, recently there has been an exponential increase in application involving coordination complexes. It a very gentle ionization process (involving transfer of solution ions to the gas phase), and therefore typically yields molecular ions with little or no fragmentation.

ESI requires very little sample for analysis, picomolar concentrations being easily accessible under ideal conditions, and the ability to couple with liquid chromatographic (LC or HPLC) separation techniques, giving LC-MS.

The ionization process

In the electrospray process, a dilute solution is produced applying a strong electric field, into a chamber, which is at atmospheric pressure, to a liquid passing through a capillary tube with a weak flux (normally 1-10 µl min^{-1}). The electric field is obtained by applying a potential difference of 3-6 kV between this capillary and the counter-electrode, separated by 0.3-2 cm, producing electric fields of the order of 10^6 Vm^{-1} (Figure 2.9).

A fine spray of charged solution droplets is produced, from which the solvent is evaporated by a stream of warm gas (usually nitrogen) to give gas-phase ions. The nature of the desolvation process has been the subject of much debate, with the two main theories proposed involving either ion evaporation from the surface of the charged droplet, or a more catastrophic process involving a Coulombic explosion.
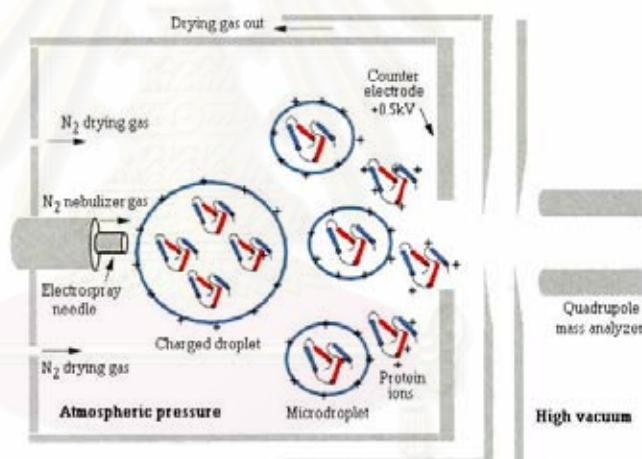
Figure 2.9 The Electrospray Ionization Source

The ESI process is the formation of a series of multiply charged ions for large biopolymers. The positive ions of the general nature [M+nH]$^{n+}$ are formed by the protonation of basic sites in biopolymers. In negative-ion mode, the [M-nH]$^{n-}$ type ions are formed. Because a mass spectrometer analyzers ions on the basis of their m/z ratios rather than their masses, the effect of multiple charging is to reduce significantly the m/z of the intact macromolecule, a process that intact macromolecule, a process that brings high-mass compounds within the usable mass range of an ordinary mass spectrometer. The charge state of a particular peak and hence the molecular mass of an ordinary mass spectrometer.

The charge state of a

$$n = \frac{m'' \, H}{m' - m''}$$

Particular peak and hence, the molecular mass of a biopolymer can be calculated by solving the following simultaneous equation:

Then $M = n \, (m'-H)$

Here, m' is the mass of the $[M+nH]^{n+}$ and m"(m'> m") of the adjacent $[M+(n+1)H]^{n+1}$ ion. Algorithms have been developed that can computer the value of M from successive pairs of adjacent ions, and provide the average value. In favorable cases, the molecular mass of macromolecules can be calculated with a precision of at least ±0.005 percentage.

2.6.3 Quadrupole Mass Filter (17-18, 23)

Quadrupole mass analyzers have been used in conjunction with electron ionization sources since the 1950s and are the most common mass spectrometers in existence today. Paul and Steinwegen described the principle of the quadrupole in 1953.

Principle

The Quadrupole mass filter is four precisely parallel rods (usually gold coated stainless steel) in hyperbolic or round shape, connected to a direct current (DC) voltage and a superimposed radio-frequency (RF) potential (Figure 2.9). Combined DC and RF potentials on the quadrupole rods can be set to pass only a selected mass-to-charge ratio. All other ions do not have a stable trajectory through the quadrupole mass analyzer and will collide with the quadrupole rods, never reaching the detector.
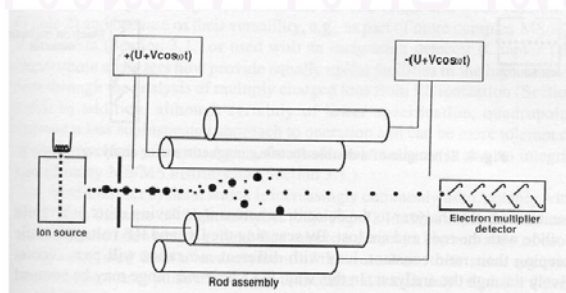
Figure 2.10 The Quadrupole Instrument

The ion trajectory (oscillation) governed by Mathieu equations, leading to definition of two parameters. The two opposite rods in the quadrupole have a potential of $+(U+V\cos(\omega t))$ and the other two $-(U+V\cos(\omega t))$ where 'U' is the fixed potential and $V\cos(\omega t)$ is the applied RF of amplitude 'V' and frequency 'ω'. The applied potential on the pairs of rods varies sinusoidally as $\cos(\omega t)$ cycles with time 't'. This results in ions being able to traverse the field free region along the central axis of the rods but with oscillations amongst the poles themselves. These oscillations result in complex ion trajectories dependent on the m/z of the ions. Specific combinations of the potentials 'U' and 'V' and frequency 'ω' will result in specific ions being in resonance creating a stable trajectory through the quadrupole to the detector. All other m/z values will be non- resonant and will hit the quadrupole and not detected. Length and diameter of rod determine the mass range and resolution of the instrument. Quadrupole mass analyzers are often called mass filters because of the similarity between m/z selection by a quadrupole and wavelength selection by an optical filter or frequency selection by an electric filter.

2.6.4 Time-of-Flight Mass Analyzer (10, 17-18, 24)

Time-of-Flight (TOF) analyzers are one of the simplest MS analyzer in use today. Stephens developed TOF in the late 1940's, but until the 1990's its popularity was limited. Recent improvements in TOF technology, including orthogonal acceleration, ion mirrors (reflectrons), and high-speed electronics, have significantly improved TOF resolution. This improved resolution, combined with powerful and easy-to-use electrospray (ESI) and matrix-assisted laser desorption/ionization (MALDI) ion source, have made ToF MS a core technology for the analysis of both small and large molecules.

Principle

A time of flight mass spectrometer measures the mass-dependent time it takes ions of different masses to move from the ion source to the detector. This requires that the starting time (the time at which the ions leave the ion source) be well-defined. Therefore, ions are either formed by a pulsed ionization method (usually matrix-assisted laser desorption/ionization, or MALDI), or various kinds of rapid electric field switching are used as a 'gate' to release the ions from the ion source in a very short time.

Recall that the kinetic energy of an ion leaving the ion source is:

$$\frac{mv^2}{z} = qV_s = zeV_s = E_k$$

The ion velocity, v, is the length of the flight path, L, divided by the flight time, t:

$$v = \frac{L}{t}$$

Substituting this expression for v into the kinetic energy relation, we can derive the

$$t^2 = \frac{m}{z}\left(\frac{d^2}{2V_s e}\right)$$

The mass-to-charge ratio can be calculated from a measurement of t^2, the terms in parentheses being constant. This equation also shoes that, all others factors being equal, the lower of an ion, the faster it will reach the detector. The simplest example is a linear TOF analyzer, where each ion is accelerated into the field-free region (the flight tube) and maintains the velocity it acquired by the acceleration until it hits the detector (Figure 2.11).

In principle, the upper mass range of a TOF instrument has no limit, which suitable for soft ionization techniques. For advantages of TOF mass analyzers are high sensitivity and very fast scan speed but it is poor mass resolution. Mass resolution is affected by a distribution in flight times among ions with the same m/z ratio .Such as the ions leaving the ion source of a time-of-flight mass spectrometer have neither exactly the same starting times nor exactly the same kinetic energies (recall the "chromatic aberrations" discussed for magnetic sector mass spectrometers). Various time-of-flight mass spectrometer designs have been developed to compensate for these differences. A reflectron is an ion optic device in which ions in a time-of-flight mass spectrometer pass through a "mirror" or "reflectron" and their flight is reversed (Figure 2.11). The reflectron is a series of rings or grids that act as an ion mirror. This mirror compensates for the spread in kinetic energies of the ions as they enter the drift region and improves the resolution of the instrument. The output of an ion detector is displayed on an oscilloscope as a function of time to produce the mass spectrum
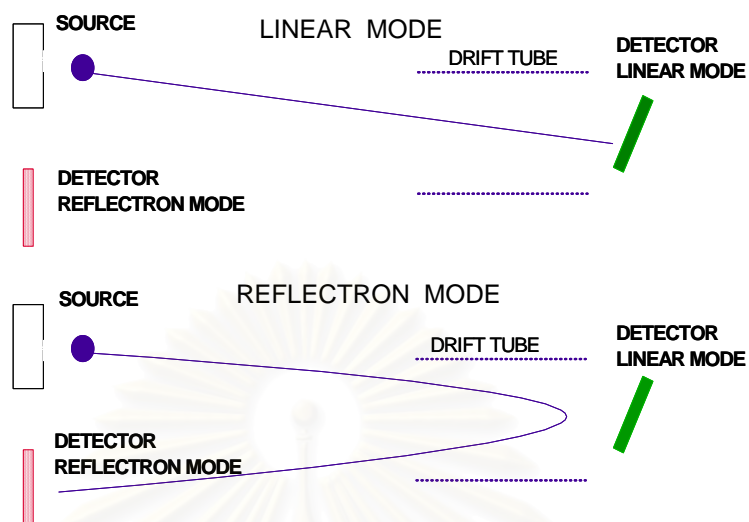
Figure 2.11 Linear Time-of-Fight and Reflectron Time-of-Fight Mass Spectrometer

2.7 Tandem Mass Spectrometry (MS/MS) (17, 24-25)

Tandem mass spectrometry (MS/MS) was first used in the late 1960s. There is used to produce structural information about a compound by fragmenting specific sample ions inside the mass spectrometer and identifying the resulting fragment ions. This information can then be pieced together to generate structural information regarding the intact molecule. Tandem mass spectrometry also enables specific compounds to be detected in complex mixtures because of their specific and characteristic fragmentation patterns. Tandem mass spectrometer is a mass spectrometer that has more than one analyser, in practice usually two. A collision cell into which an inert gas (e.g. argon, xenon) is admitted to collide with the selected sample ions and bring about their fragmentation separates the two analysers. The analysers can be of the same or of different types, the most common combinations being:

Quadrupole - quadrupole

Magnetic sector - quadrupole

Magnetic sector - magnetic sector

Quadrupole - time-of-flight

The four mains scan modes available using MS/MS are represented in Figure 2.12

First mode is product or daughter ion scanning: the first analyser is used to select user-specified sample ions arising from a particular component; usually the molecular-related (i.e. $(M+H)^+$ or $(M-H)^-$) ions. These chosen ions pass into the collision cell, are bombarded by the gas molecules which cause fragment ions to be formed, and these fragment ions are analysed i.e. separated according to their mass to charge ratios, by the second analyser. All the fragment ions arise directly from the precursor ions specified in the experiment, and thus produce a fingerprint pattern specific to the compound under investigation. This type of experiment is particularly useful for providing structural information concerning small organic molecules and for generating peptide sequence information.

Product ion scans

| Selected m/z | CID | Scanned |

Precursor ion scan

| Scanned | CID | Selected m/z |

Neutral loss scan

| Scanned m/z = x | CID | Scanned m/z = x–a |

Selected reaction monitoring

| Selected precursor m/z = a | CID | Selected fragment m/z = b |
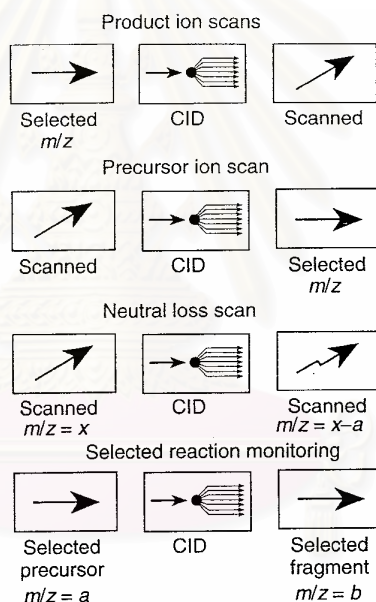
Figure 2.12 A Schematic Representation of The Tandem Mass Spectrometry Scan Modes

Precursor or parent ion scanning: the first analyser allows the transmission of all sample ions, whilst the second analyser is set to monitor specific fragment ions, which are generated by bombardment of the sample ions with the collision gas in the collision cell. Constant neutral loss scanning: this involves both analysers scanning, or collecting data, across the whole m/z range, but the two are offset so that the second analyser allows only those ions that differ by a certain number of mass units (equivalent to a neutral fragment) from the ions transmitted through the first analyzer.

The last mode is selected/multiple reaction monitoring: both of the analysers are static in this case as user-selected specific ions are transmitted through the first analyser and user-selected specific fragments arising from these ions are measured by the second analyser. The compound under scrutiny must be known and have been well characterised previously before this type of experiment is undertaken.

2.7.1 Collision-induced Dissociation (CID) (17, 24)

If an ion collides with a neutral atom or molecule, some of the ion's kinetic energy can be converted into internal energy. This is called collisional activation. If there is enough excess internal energy to break chemical bonds, the ion will decompose. This is called collision-induced dissociation (CID) or collisionally activated dissociation (CAD). Both terms mean the same thing. The CID processes a sequence of two steps. The first step is very fast and corresponds to the collision between the ion and the target molecule when a fraction of the ion translational energy is converted into internal energy, bringing the ion into an excited state. The second step is the unimolecular decomposition of activated ion. CID is accomplished by selecting an ion of interest with a mass analyzer and introducing that ion into the collision cell, where the selected ion collides with the collision gas (Ar or He) atoms, resulting in fragmentation. For collision energy conversion to internal energy, the collisions of mobile species (the ion) and a static target (the collision gas) control the kinetic energy for internal energy transfers. In practice, there are two groups of collision energy: high-energy collisions, these refer to collisions where the precursor ion is accelerated to kinetic energies of approximately one kilovolt or higher. Low-energy collisions are referring to collisions where the precursor ions have kinetic energies in the range of a few eV to a few hundred eV. Fragmentation within triple quadrupole, qaudrupole ion trap and hybrid quadrupole ToF analyzers occur at low collisional energy, whilst fragmentations within a magnetic sector or ToF/ToF analyzer occur at high collisional energy. The collision gas is more important than it is for the high-energy collisions. Heavier gases such as argon, xenon or kryton are preferred because they allow the transfer of more energy. In comparison of two-collision energy, the different fragmentation patterns are observed. The high-energy CID spectra give simpler, more clear-cut fragmentation, whereas low-energy CID spectra lead to more diverse fragmentation pathways, often including more rearrangements.

2.7.2   Electrospray   Ionization   Quadrupole/Time   of   Flight   Mass spectrometer (ESI-Q/TOF MS) (10, 24)

ESI-Q/TOF mass spectrometry has electrospray ionization for ion source, a quadrupole mass filter for the mass analyzer and a time-of-flight mass analyzer for the second mass analyzer.   ESI-Q-TOF MS used for molecular weight measurements, reaction monitoring, protein structural studies, peptide sequencing, nucleotide sequencing, macromolecule structure determination due to extended m/z range.

Quadrupole time-of-flight or hybrid analyzer is described in 1996 by Morris for oligosaccharide analysis and more recently Lobada et al., 2000; these instrument have rapidly become the instrument standard for MS/MS applications with in the theatre of proteomics.   The user is able to acquire MS and most notably MS/MS data with high mass accuracy, resolution and sensitivity.   The instrument is generally with interfaced with HPLC.
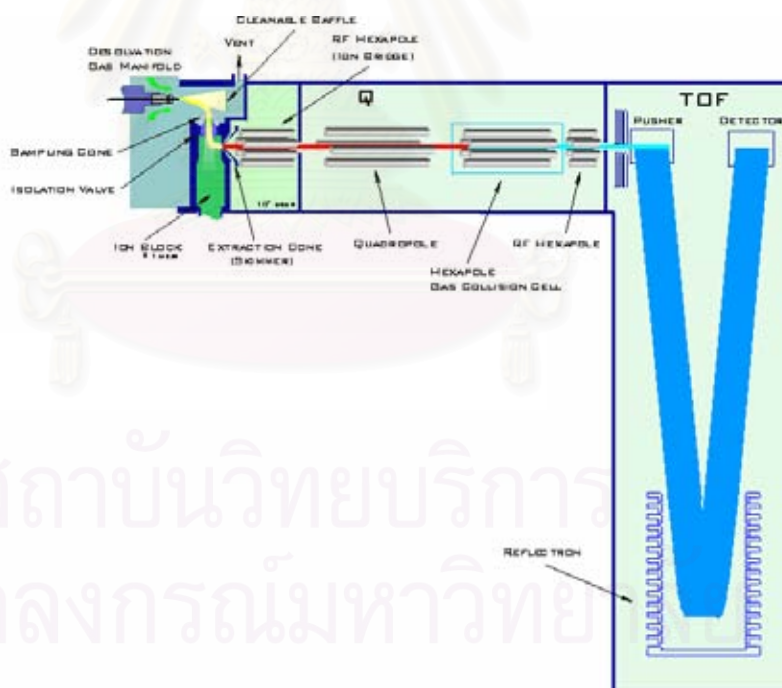
Figure 2.13 Schematic diagram of Electrospray Ionization Quadrupole/Time of Flight
Mass spectrometer

2.8 Peptide Sequencing by Tandem Mass spectrometry (17, 24)

The most common usage of MS/MS in biochemical areas is the product or daughter ion scanning experiment, which is particularly successful for peptide and nucleotide sequencing. Roepstorff and Fohhlman, 1984, described this fragmentation nomenclature.

Peptides are three different types of bonds that can fragment along the amino acid backbone: the NH-CH, CH-CO, and CO-NH bonds. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence, there are six possible fragment ions for each amino acid residue and these are labelled as in the figure 2.13, with a, b, and c ions having the charge retained on the N-terminal fragment. The x, y", and z ions having the charge retained on the C-terminal fragment (Figure 2.14). The most common cleavage sites are at the CO-NH bonds which give rise to the b and/or the y" ions.
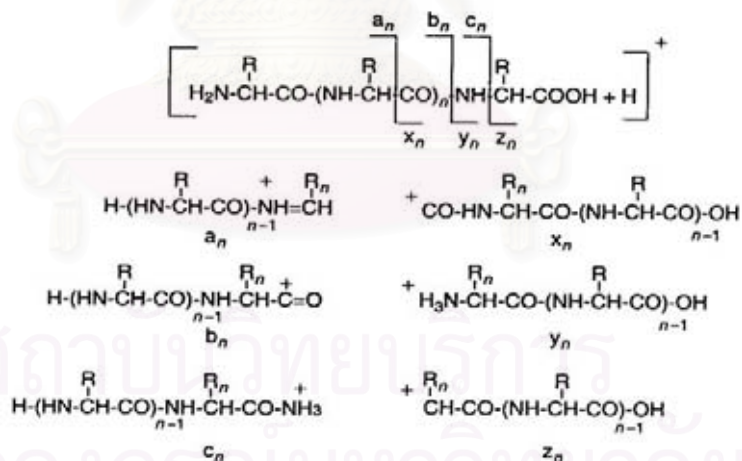
Figure 2.14 Structure of the peptide fragmentation

The extent of side-chain fragmentation detected depends on the type of analysers used in the mass spectrometer. A magnetic sector–magnetic sector instrument will give rise to high-energy collisions resulting in many different types of side-chain cleavages but also increasing the complexity and difficulty of interpretation. Quadrupole–quadrupole and quadrupole–time-of-flight mass

spectrometers generate low energy fragmentations with fewer types of side-chain fragmentations.

In addition, there are two other types of fragments, which appear among the low masses in the spectrum (Figure 2.15). The first type is called an internal fragment because these fragments have lost the initial N- and C- terminal side. The second type of fragment is immonium ions of amino acids; appear in the very low m/z range of the MS-MS spectrum. Each amino acid residue leads to a diagnostic immonium ion, with the exception of the two pair's leucine (L) and iso-leucine (I), and lysine (K) and glutamine (Q), which produce immonium ions with the same m/z ratio, i.e. m/z 86 for I and L, m/z 101 for K and Q. The immonium ions are useful for detecting and confirming many of the amino acid residues in a peptide, although no information regarding the position of these amino acid residues in the peptide sequence can be ascertained from the immonium ions.
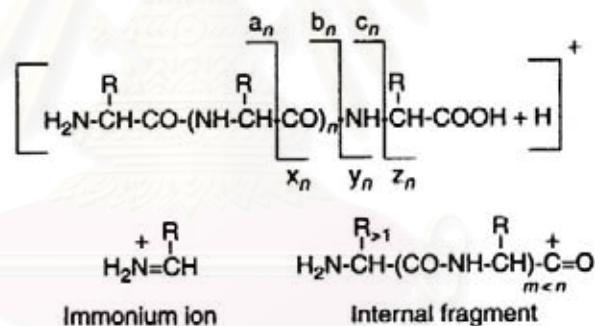
Figure 2.15 Immonium Ion and Internal Fragment

Protein and peptide structure differences relative to the residue masses in Table 2.3, which allow recognition of the N- and C- terminal of the peptide sequence.

Table 2.3 The residue masses of common 20 amino acids

Amino acid	Code (3 letters)	Code (1 letter)	Residue mass (Da)	Immonium ion (m/z)
Glycine	Gly	G	57.02	30
Alanine	Ala	A	71.04	44
Serine	Ser	S	87.03	60
Proline	Pro	P	97.05	70
Valine	Val	V	99.07	72
Threonine	Thr	T	101.05	74
Cysteine	Cys	C	103.01	76
Leucine	Leu	L	113.08	86
Isoleucine	Ile	I	113.08	86
Asparagine	Asn	A	114.04	87
Aspartate	Asp	D	115.03	88
Glutamine	Gln	Q	128.06	101
Lysine	Lys	K	128.09	101
Glutamate	Glu	E	129.04	102
Methionine	Met	M	131.04	104
Histidine	His	H	137.06	110
Phenylalanine	Phe	F	147.07	120
Arginine	Arg	R	156.10	129
Tyrosine	Tyr	Y	163.06	136
Tryptophan	Try	W	186.08	159

2.9 Protein Identification

Protein can be rapidly identified using mass spectrometry. It is especially the short time analysis, the high-sensitivity and high information content in protein sequencing. For general method to MS data can be used in four approaches for protein identification (26).

1) Peptide mass fingerprint (PMF) – MS mode. The mass measurement of each peptide derived from the enzyme digestion or chemical cleavage of the protein.

2) Peptide mass fingerprint and composition information. The molecular weights of each of the peptides derived from the enzyme digestion or chemical cleavage of the protein can be used alongside some composition information relating to one or more of the peptide.

3) Peptide mass fingerprint and sequence information. The molecular weight of each of the peptides derived from the enzyme digestion or chemical cleavage of the protein can be used alongside some direct sequence information realign to one or more the peptides.

4) Product ion MS/MS sequence data from one or more peptide-MS/MS mode.

As a rule, the mass spectrometric data use freely available databases containing amino acid sequences of protein. The example of the protein sequence databases were shown below (27):

- SWISS-PROT (www..expasy.ch/sprot-top.html) is a database of annotated protein sequence; it also contains addition information on function of the protein, its domain structure, posttranslational modification, etc.

- TrEMBL (www.expasy.ch/srs5/) is a supplement to SWISS-PROT, which contains all protein sequences, translated from nucleotide sequences of the EMBL database.

- NCBInr (www.ncbi.nLm.nih,gov/dbEST/) is a database containing sequences translated from DNA sequences of GenBank and also sequences from PDB, SWISS-PROT and PI database.

These databases are constantly updated and are usually characterized by the standardized data format.

For peptide, mass fingerprinting has several programs available to use. There is MASCOT at www.matrixscience.com, profound at www.prowl.com and MS-FIT at www.prospector.ucsf.edu/. The four important parameters for data search are peptides mass list, the cleavage agent, error tolerance (mass accuracy) and knowledge of peptide modification.

These existing algorithms and corresponding programs may be subdivided into three main groups (17) shown in Table 2.4.

Table 2.4 Examples of the databases search programs and their Internet addresses that
can be used for protein identification

Program Name	Internet Address
Programs that use amino acid sequences for the search query	
FASTA	fasta.bioch.virginia.edu/fasta
BLAST	www.ncbi.nlm.nih.gov/blast/blast.cgi
MS-Edman	prospector.ucsf.edu/mshome3.2.html
Programs that use peptide molecular weights for the search query	
MS-Fit	prospector.ucsf.edu/mshome3.2.htm
MOWSE	srs.hgmp.mrc.ac.uk/cgi-bin/mowse
Peptide	www.mann.-heidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html
Programs that use the data from uninterpreted product ion spectra for the search query	
SEQUEST	thompson.mbt.washington.edu/sequest/
MS-Tag	prospector.ucsf.edu/mshome3.2.htm
Peptide	www.mann.embl-heidelberg.de/Services/PeptideSearch/PeptideSearchIntro.html

For protein identification by 2-DE, in gel digestion has important
methodology to identify. There steps are excision of the relevant spots, in-gel
proteolytic cleavage and extraction of the resulting peptides, the mass spectrometric
(MALDI) analysis of this peptide mixture (peptide mass fingerprinting or PMF) can
be performed. A subsequent protein database search can result in the identification of
the organism is available. If this method does not result in unequivocal protein
identification, individual peptides (parent) of the proteolytic map can be subject to
further fragmentation analysis using MS/MS techniques (peptide fragmentation
fingerprint or PFF). Hence, the identity of the original protein can be determined
from the resulting fragment ion pattern (daughter ion spectrum) of a single proteolytic
peptide by a further database search. This method even allows the identification of
different components of an unseparated protein mixture. In addition to these database
dependent techniques, ESI-MS/MS also allows direct sequencing of individual
peptides from a proteolytic mixture without the availability of genetic information.

2.10 Literature Reviews

Form literature survey, there are a number of reports on the studies of plants in this Zingiberaceae family. Most of these studies involve their secondary metabolites and bioactivities. For example, in 1995 Apisariyakul A. *et al.* (28) found curcumin in volatile oil from *Curcuma long* Linn which it have antifungal activity. Jeenapongsa R. *et al.* (1995) reported (*E*)-1-(3, 4-dimethoxyphenyl) butadiene from *Zingiber cassumuna* (29)   These compound posse anti-inflammatory. Singh *et al.* (2002) (30) reported essential oils of some Indian *Curcuma* species have antifungal activity. Takahashi M. *et al.* (2004) (31) found parviflorene A which is a novel compound from Curcuma parviflora and have cytotoxic with B16 melanoma cell.

On the other hand, there have been only a few studies on primary metabolites especially on protein. From database searching, there are 2 and about 10 sequence proteins from Zingiberaceae found in SWISS-PROT and TrEMBL database (www.expasy.ch) (32), respectively. For example, In 1999 Hahn W.J. reported ATP synthase beta subunit (Fragment) from chloroplasts of *Riedelia aff* Wrayii. This protein have molecular masses 53,352 Da, 496 amino acid residues and produce ATP from ADP in the presence of a proton gradient across the membranes. In 2000, Cao H. *et al.* reported maturase-like polypeptide (62,102 Da, 515 amino acid residues) from leave of Curcuma xanthorrhiza. This protein was found in other plants in leaves of Curcuma genous such as *Curcuma longa*, *Curcuma zedoaria* and *Curcuma comosa* etc. Prince L.M. *et al.* reported NADH dehydrogenase subunit F (fragment) which it have 77,561 Da and 690 amino acid residues. This protein has catalytic activity NADH dehydrogenase. Maturase K, Ribosomal protein S4 , Ribulose 1,5-bisphosphate carboxylase and Cysteine protease GP I and GP II are another example protein in Zingiberaceae.

In previously, the mainly of identification of Zingiberaceae plants were used morphological. In 2001, Cao H. *et al.* (33) have developed a new method to identify *Curcuma* plants by using nucleotide differences at two sites and existence of 4-base idle on *trnK* gene. Recently, Samuel C. *et al.* (2002) (34) used proteome in ginseng to identified different species between Oriental ginseng (*Panax ginseng C.A.*) and American ginseng (*Panax quinquefolius).*

CHAPTER III

EXPERIMENTAL

3.1 Material

3.1.1 Plant materials

There were 20 species of Zingiberaceae in this study including *Curcuma longa, Curcuma zedoaria, Curcuma mangga, Curcuma parviflora, Curcuma amarissima, Curcuma aeruginosa, Curcuma aromatica, Curcuma comosa, Curcuma sparganifolia, Curcuma sp.* (Khamindum, Haroynang, Maleung, Saligalinthong, Payanwan, Enleung and Khantamala), *Zingiber cassumunar, Zingiber ottensii, Zingiber officinale* and *Alpinia niga.* Fresh rhizomes of *Curcuma longa* was purchased from three markets in Bangkok, Thailand (Pak Khlong Talat, Bang Kae and Chatuchak). Then, another species were purchased in one market (Chatuchak part). Zingiberceae plants were stored at room temperature until use when taking samples from rhizome, the most superficial layer and less than 0.2 cm in depth were pilled.

3.1.2 Instruments

Autopipette: Pipetman, Gilson, France

DUEL Gel Caster: Mighty Small SE245, Hoefer, U.S.A

IEF electrophoresis unit: Multiphor II, Amersham Pharmacia Biotech, Sweden

MALDI-TOF Mass spectrometer: Biflex, Bruker, Germany

ESI-Q-TOF Mass spectrometer: Micromass, UK.

Microcentrifuge: Biofuge pico Heraeus, Kendro, Germany

Microtrap desalting cartridge: Michrom Bioresource

Orbital Shaker: Kika-Werke GMBH&Co., Germany

Peptide Trap: Michrom BioResources, Germany

pH meter: Denver Instrument, U.S.A.

Power Supply: EPS 3500 XL, Pharmacia, England

Refrigarated centrifuge: Himac CR20B2, HITACHI, Japan

Sonicate: DHA-1000: Branson, U.S.A

Sonicator: BHA-1000, Branson, U.S.A

Spectrophotometer: TECAN, Austria

Speed vacuum centrifuge: Heto-Holten, Denmark

Vortex mixer: Vortex-Genie2, Scientific Industries, U.S.A

Water Bath Shaking: Memmert, Germany

3.1.3 Chemicals

Acetic acid: Merck Ag Darmstadt, Germany

Acetone: Merck Ag Darmstadt, Germany

Acrylamide PAGE: Plusone Pharmacia Biotech, Sweden

Agaraose: Plusone Pharmacia Biotech, Sweden

Ammonium persulfate: Plusone Pharmacia Biotech, Sweden

β-mercaptoethanol: Merck Ag Darmstadt, Germany

Bromophenol Blue: USB, U.S.A

CHAPS (3-[(3-cholamidopropyl)-dimthylammonio]-1-propane sulfonate): USB, USA

Coomassie Brilliant Blue G-250: USB, U.S.A

Dowex (2x8-400, Cal⁻form): Sigma, U.S.A

DTT (Dithiothrretiol): Plusone Pharmacia Biotech, Sweden

EDTA (Ethylenediaminetetraacetic acid): Merck Ag Darmstadt, Germany

Ethanol: Merck Ag Darmstadt, Germany

Glycerol: Plusone Pharmacia Biotech, Sweden

Glycine: Plusone Pharmacia Biotech, Sweden

Iodoacetamide: Sigma, U.S.A

Methanol: Merck Ag Darmstadt, Germany

N,N'-methylene-bis-acrylamide: Plusone Pharmacia Biotech, Sweden

Nonidet NP-40: Amresco, U.S.A

PMSF: Sigma, U.S.A

PVPP: USB, U.S.A

SDS (Sodium Dodecyl Sulfate): Plusone Pharmacia Biotech, Sweden

Tris: USB, U.S.A

Urea: Plusone Pharmacia Biotech, Sweden

3.1.4 Supplies

BIO-PROFIL Bio-2D: VILBER LOURMAT, France

Glass water Sills: GFL Gesellschaft für Labotechik mbh, Germany.

3.1.5 Kit

Immobilized pH gradient strip pH 4-7: Amersham pharmacia biotech, Sweden

Protein assay: BioRad, USA

Standard Low Molecular Weight Marker Protein: Amersham pharmacia biotech, Sweden

3.2 Methods

3.2.1 Extraction and Precipitation

1) Mg(NP-40)A and Mg(NP-40)B Extraction (35)

Two grams of plant tissues were placed in liquid nitrogen. The plant tissues were transferred to a prechilled mortar and ground with a pestle in liquid nitrogen to a fine powder. The powder was homogenized in 10 ml of ice-cold Mg/NP-40 extraction buffer containing 0.5 M Tris-HCl. pH 8.3, 2% v/v NP-40, 20 mM $MgCl_2$, 2% v/v β-mercaptoethanol, 1 mM phenylmethylsulfonylfluoride (PMSF), 1 % w/v polyvinylpolypyrrolidone (PVPP) and 5 % Dowex. After centrifugation at 12,000 x g for 15 min at 4 °C, proteins in the supernatant were precipitated by adding four volumes of cold acetone (Mg(NP-40)A or 10 % TCA in acetone (Mg(NP-40)B) at -20 °C for overnight for analysis of total protein by 2-DE. After centrifugation at 3,000xg for 10 min, the pellets were air-dried.

2) Phenol extraction (36)

Weight tissues (0.5 g to 1g) and place it in a chilled mortar. Freeze tissue in the mortar with liquid nitrogen and grind it immediately with a pestle to a fine powder. Add 3 ml extraction buffer containing 0.7 M sucrose, 0.5 M Tris, 30 mM HCl, 0.1 M KCl and 1% β-mercaptoethanol to the powder. Homogenize further to a fine suspension and transfer suspension to a Corex tube. Wash the mortar twice with 1 ml of extraction buffer and transfer this solution to the same tube. Vortex the extract solution and keep it at 4 °C for at least 30 min Centrifuge the suspension at 4,000xg for 10 min, transfer the supernatant to a new Corex tube, and keep at 4 °C. Resuspend the pellet with 3 to 5 ml of extraction buffer and repeat the extraction for three times. To combined supernatant, add 5 ml of water-saturated phenol (Appendix A). Mix phenol, extract thoroughly, and keep on ice for 1 hour. Separate phases by centrifugation at 8,000xg for 10 min and transfer the upper phenol phase to a new Corex tube. To water phase, interphase, add 3 ml of fresh water-saturated phenol and repeat the extraction. Combine both phenol phase to precipitate the proteins from the phenol by adding of 20 ml 0.1 M ammonium acetate dissolved in the methanol at -20 °C for overnight. Centrifuge at 4,000xg for 10 min to obtain a protein pellet. Pour off the supernatant, dissolve the pellet immediately in 1 ml of water, sonicate for 3 min, add 9 volumes of cold acetone and put the solution at -20 °C for at least 4 hours. Centrifuge the precipitated protein again at maximum 4,000xg, pour off the supernatant and air-dry the pellet briefly the smell of acetone has disappeared.

3.2.2 Protein Quantification

Protein content was determined by Bradford's method (37). Prepare five standards solution (1mL each) containing 0, 10, 20, 30, 40 and 50 μg/ml BSA. To a 1.4 ml plastic cuvelte, add;

0.2 ml Protein assay solution (BioRad).

0.8 ml of one of the protein standard solution or sample to be assayed (containing<100 μg of protein for <50 μg/ml standards)

Cover with Parafilm and gently invert several times to mix. Record the absorbance spectrum of sample at 595 nm and repeat the steps above for each of the protein standards and for samples to be assayed. Plot standard curve of Absorbance at

595 nm vs. [protein] for the protein standard. To determine the protein concentration of sample from it absorbance, use the standard curve to find the concentration of standard that would have the sample absorbance as the sample.

3.2.3 Two-Dimensional Gel Electrophoresis

1) Isoelectric focusing and SDS-PAGE

The 2-D electrophoresis procedure was described in the manual of 2-D electrophoresis using immobilized pH gradients; Principle and Method (15). For the dimension IEF, protein pellet dissolved in lysis buffer (8 M urea, 4% CHAPS and 2% IPG buffer pH 4-7 (Appendix B)). The total protein concentration was quantified by using Bradford's procedure (37). Thence, added and mixed rehydration solution (8 M urea, 2% CHAPS and 2% IPG buffer pH 4-7, 0.2% DTT and bromophenol blue (see Appendix B)) in the sample solution to total volume 125 μl contain an estimated 250 μg protein. Delivered the solution slowly at a central point in the slot of the Immobiline DryStrip Reswelling Tray and removed any lager bubbles. Placed the pH 4-7 IPG strips (length 7 cm) on the solution and overlaid each IPG strip with 3 ml of DryStrip Cover Fluid to minimize evaporation and urea crystallization. A minimum of 10 hours was required for rehydration at room temperature (Figure 3.1).
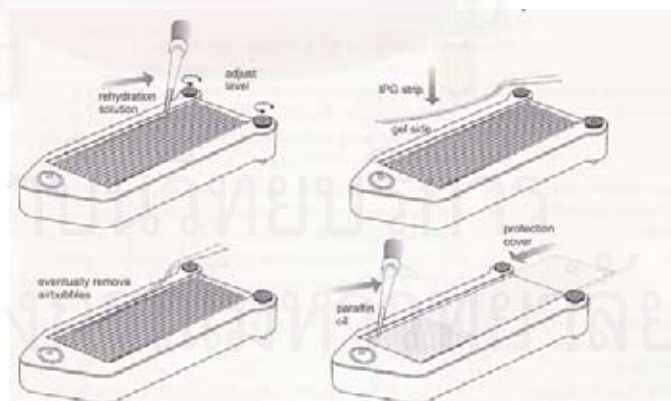
Figure 3.1 Rehydration of IPG strips in individual grooves in the reswelling tray

After rehydrated strip, removed a strip from the reswelling tray and transferred to the Immobiline Drystrip aligner. Placed the moistened electrode strips across the cathodic and anodic ends of the aligned IPG strips. The electrodes strips must be least partially contact the gel surface of each IPG strip. The IEF separation was carried out on the Multiphor II unit and connects the leads on the lid to the power supply (EPS 3500 XL). The separation conditions set the temperature at 20°C, current with 0.5 mA/strip and power with 5 W totals. The running condition according to the following gradient mode: 200 V for 1 min, 3,500 V for 1.5 h and 3,500 V for 1.5 h. Upon completion of IEF, each of the focused gel was put into a screw tube, the proteins were reduces by incubating with 1% w/v DTT for 10 min and alkylated with 2.5 % w/v iodoacetamide in 10 ml of equilibration buffer (30% v/v glycerol, 2% w/v SDS, 50 mM Tris-HCl pH 8.8, 6 M urea and 0.002% bromophenol blue) for 10 min. SDS-PAGE in the second dimension was carried out as described by Laemmli (1970) (11). The second dimensional, assemble glass plates for casting slab gel and label on plate with indelible marker at the resolving gel height. SDS gel was 1 mm thick and consisted of a 6 cm separation gel of 15% w/v acrylamide overlaid with a 0.5 cm stacking gel of 4% w/v acrylamide. The resolving and stacking gel consisted of solutions in the Table 3.1.

Table 3.1 Composition of the gel solutions for one gel

Stock solution*	15% *T* Resolving gel	5% *T* Stacking gel
30% T, 2.6 % C monomer (ml)	3.896	0.130
Resolving buffer (ml)	2	-
Stacking gel buffer (ml)	-	0.250
Double distilled H$_2$O (ml)	1.98	0.610
10% APS (µl)	40	5
10% SDS (µl)	80	10
TEMED (µl)	4	1
Total volume (ml)	8	1

* Preparation stock solution was described in Appendix B

The strips were then transferred to SDS-PAGE gel and sealed to SDS-PAGE gel with 0.5% agarose in SDS electrophoresis buffer. The SDS electrophoresis buffer contained 25 mM Tris-base, 192 mM glycine and 0.1% w/v SDS. For second dimension electrophoresis is used Hoefer miniVE vertical electrophoresis system and connects the leads on the lid to the power supply (EPS 3501XL). Electrophoresis was carried out at 10 mA/gel until the tracking dye had traversed the stacking gel and then at a 20 mA/gel until the dye had reached the bottom of the gel. The molecular weight markers were phosphorylase B (97,000 Da), bovine serum albumin (66,000 Da), ovalbumin (45,000 Da), carbonic anhydrase (31,000 Da), soybean trypsin inhibitor (21,000 Da) and lysozyme (14,400 Da).

2) Protein Detection by Coomassie blue Staining

After electrophoresis, proteins in the gel were stained by coomassie blue. Remove the polyacrylamide gel from between the glass plates and immersed in 50 mL of coomassie blue staining solution (filter through Whatman No. 1 filter paper before used). Placed on a rotary shaker and gently shake the gel overnight. Removed the staining solution and rinse the gel extensively with H_2O. Placed the in the destaining solution I on a rotary shaker for 1-2 hr, followed by changed to destaining solution II (Appendix B) until background of the gel was cleared.

3) 2-DE Data Analysis

The coomassie blue stained gels were scanned with at 300 dots per inch by Image densitometer (BIORAD, USA.). Spots were detected and quantified with Bio-2D software (38) based on their relative molecular weight, pI. Isoelectric points of protein spots identified were estimated by portion as the IPG strips used were of linear gradient. Low molecular weight marker estimated molecular weights of protein spots identified.

3.2.4 In-gel Digestion

In-gel digestion procedure has been modification from the method described in Chapter 6: the preparation of protein digests for mass spectrometric sequencing of the book of Protein Sequencing and Identification Using Tandem Mass Spectrometry (17). Rinse the gel with water. Excise bands of interest with clean scalpel cutting as close to the edge of the spot or band as possible. It was important to reduce the volume of "background" gel. Chop the excised bands into cubes (*ca*. 1x1mm). Transfer gel particles into a microcentrifuge tube (0.5 ml or 1.5 ml eppendorf). Wash the gel particle with 100-500 µl of water (5 min, *ca*. 2-3 times). Spin down and remove the liquid. Add acetonitrile/0.1M NH_4HCO_3 (1:1, *ca*. 3-4 times equal the volume of gel pieces) and wait for 10-15 min until the gel pieces shrunk-they become white and stick together. Spin the gel particles down and remove all liquid. Dry down gel particle in a vacuum centrifuge.

Swell the gel pieces in 10 mM dithiotreiol/0.1M NH_4HCO_3/1mM EDTA (add the liquid enough to cover gel) and incubate for 45 min at 60°C to reduce the protein. In-gel reduction was recommended even if proteins were reduced prior to an electrophoresis. Removed excess dithiotreitol solution and add 100 mM iodoacetamide/0.1 M NH_4HCO_3. Incubate for 30 min at room temperature in the dark. Remove iodoacetamide solution. Wash the gel particles with 150-200 µl of 0.05 M Tris-HCl pH 8.5/50% acetonitrile (*ca*. 3-4 times). Rehydrate gel particles in the 180 µl digestion buffer (containing 100 µl of 0.1 M Tris-HCl, pH 8.5, 2 µl of 100 mM $CaCl_2$, 20 µl of ACN and 78µl of distilled water) and 20 µl of trypsin solution. Incubate for overnight at 37 °C. After overnight incubate keep 100 µl supernatant in a microcentrifuge tube (0.5 ml or 1.5 ml Eppendorf) and add 100 µl of 2% TFA incubate for 30 min at 60 °C. After incubate, combine supernatant in before microcentrifuge tube and dry in a vacuum centrifuge. Add 30 µl of digestion buffer to the tube containing the gel pieces incubate for 10 min at 30 °C and sonicate 5 min. After then add ACN 30 µl in extraction buffer, incubate for 10 min at 30 °C and sonicate 5 min. Remove and combine supernatant in before tube containing supernatant. Add 5%formic acid/ACN in gel, incubate for 10 min at 30 °C and sonicate for 5 min. Combine all supernatant in tube and dry in a vacuum centrifuge.

3.2.5 Protein Identification Methods

1) Sample preparation for Mass spectroscopy

Peptides and proteolytic digests from the In-gel digestion was desalted and purified by a microscale reversed-phase cartridge (green stripe). The first step cleaned the cartridge with acetonitrile 500 µl and then equilibrated with 0.1 % formic acid 300 µl. Each sample, dissolved in an aqueous acidic solution (0.1 %(v/v) formic acid) and was loaded on the column. The salts were eluted with 100 to 300 µl of acetonitrile: water (20: 80). The peptide was eluted with 0.1% (v/v) formic acid in acetonitrile: water (80: 20, 75 µl).

2) MALDI-TOF Mass spectrometry

Mass measurement were carried out on a Biflex (Bruker, Germany) matrix assisted laser desorption/ionization time of flight mass spectrometer (MALDI-TOF). The dried samples were dissolved in 0.1% TFA and vortexes. The dried droplet method was used for MALDI-MS sample preparation (18). A saturated solution of α-cyano-4-hydroxycinnamic acid in 50% ACN/0.1%TFA was used as a matrix. For peptide samples, used 1 µl of sample solution was mixed with 5 µl of a matrix solution. Then, 1µl of the mixture solution was spotted on target and allowed to dry at room temperature. MALDI peptide spectra were performed in reflectron mode and internally calibrated with CCA ([2M+H]$^+$, 379.09) and Ubiquitin ([M+2H]$^{2+}$, 4283.45) oxidized form. The resulting peptide mass fingerprints (PMFs) were searched against protein database via the MASCOT program (www.matrixscience.com). The parameters used in the search were as follows: peptide masses tolerance 100 ppm and 1 missed cleavage trypsin cleavage.

3) Tandem Mass spectrometry

A Tandem mass spectrometer analysis was performed on the ESI-Q-TOF mass spectrometer (Q-TOF-Micro, Micromass, Manchester, UK). Sample were injected into a CapLC (Waters, Milfold, MA, USA) System equipped with an auto sampler, gradient and auxiliary pump. 6 µl was injected via "microliter pickup" mode and desalted on-line through a 300 µm × 5 mm C$_{18}$ Trapping cartridge (LC Packing, San Francisco, CA, USA). The sample was desalted at high flow (30µl/min) for 3 min.

The peptides was separated on a 75 μm × 15 cm × 3 μm Å C18 100 A Symmetry column (Waters, Milfold, MA, USA) prior to introduction into the mass spectrometer. A typical reversed-phase was used from low to high organic over about 60 min. Mobile phase A was 3% acetronitrile, 0.1% formic acid and B was 97% acetronitrile, 0.1% formic acid. The flow rate was 5 μl/min. The system utilized a split flow resulting in a column flow rate of approximately 200-300 nL/min.

MS/MS data was obtained using a Q-TOF MICRO (Micromass, Manchester, UK) filtered with Z-spray nanoflow electrospray ion source. The mass spectrometer was operated in positive ion mode with a potential of 3500 V applied to the nanoflow probe body. The collision energy was determined on the fly based on the mass and charge state of the peptides. Charge state recognition was used to switch only doubly and triply charged ions in MS-MS mode. Several trypsin autolysis ions were excluded. ProteinLynx Version (Micromass, Manchester, UK) to generate searchable peak lists processed the data. Initial protein identification was made by correlation of tandem mass spectra to entries in SWISS-PROT using Global Server (Version 2.0, Micromass). In addition, if clearly separated proteins were assigned to unreasonable such as different organism (animal, human, fungi) that could be traced to identical protein function an additional BLAST search (39) program of the sequence data against the clustered NCBI database (40).

3.2.6 Assay for Hemagglutinating activity (41)

A serial two-fold dilution of the lectin solution in micro titer U-plates (50 μL) was mixed with 50 μL of a 2% suspension of rabbit erythrocytes in phosphate buffered saline (pH 7.2) at room temperature. The results were read after about 1 h when the blank had fully sedimented. The hemagglutination titer, defined as the reciprocal of the highest dilution exhibiting hemagglutatination, was reckoned as one hemagglutination unit.

.

CHAPTER IV

RESULTS AND DISCUSSION

The proteins of Zingiberaceous plants were analyzed using all methods as described in the previous chapter. The results were shown and discussed in each part of this chapter, respectively.

4.1 Identification of Zingiberaceous plants by Two-dimensional Gel Electrophoresis

There analytical two-dimensional gel electrophoresis was established by using a pH 4-7 IPG strips in the first dimension and a 15% SDS-PAGE in the second dimension. Each spot in gel represents individual protein separated by pI in the first dimension and by the molecular weight in the second dimension. The appearance of the protein fingerprints was highly consistent with a variation of less than 5% in the number of spots detected by Bio-2D software (38).

4.1.1 Selection of Extraction and Precipitation procedure

For method selection, *Curcuma longa* was used as study model. Extraction and protein precipitate from *Curcuma longa* can be performed by three procedures described in the previous chapter. Pellet proteins were then resolubilized in lysis buffer and quantified. The prepared pellet proteins from Mg(NP-40)B were difficult to resolubilize. The comparison among Mg(NP-40)A , Mg(NP-40)B and Phenol preparations for 2-DE analysis has been done. Figure 4.1 shows the 2-DE results from Mg(NP-40)A (Figure 4.1A), phenol (Figure 4.1B) and Mg(NP-40)B (Figure 4.1C). From the image analysis using Bio-2D software, Mg(NP-40)A extraction gives more spots on 2-DE gels than phenol and Mg(NP-40)B procedure. From this result, Mg(NP-40)A acetone procedure was selected to used for extraction and precipitation. Although, the 2–DE gel resolution of proteins Mg(NP-40)B extracted was slightly better than others but the resolubilization of the pellet still gave problem and the phenol method is rather complicated and time consuming.
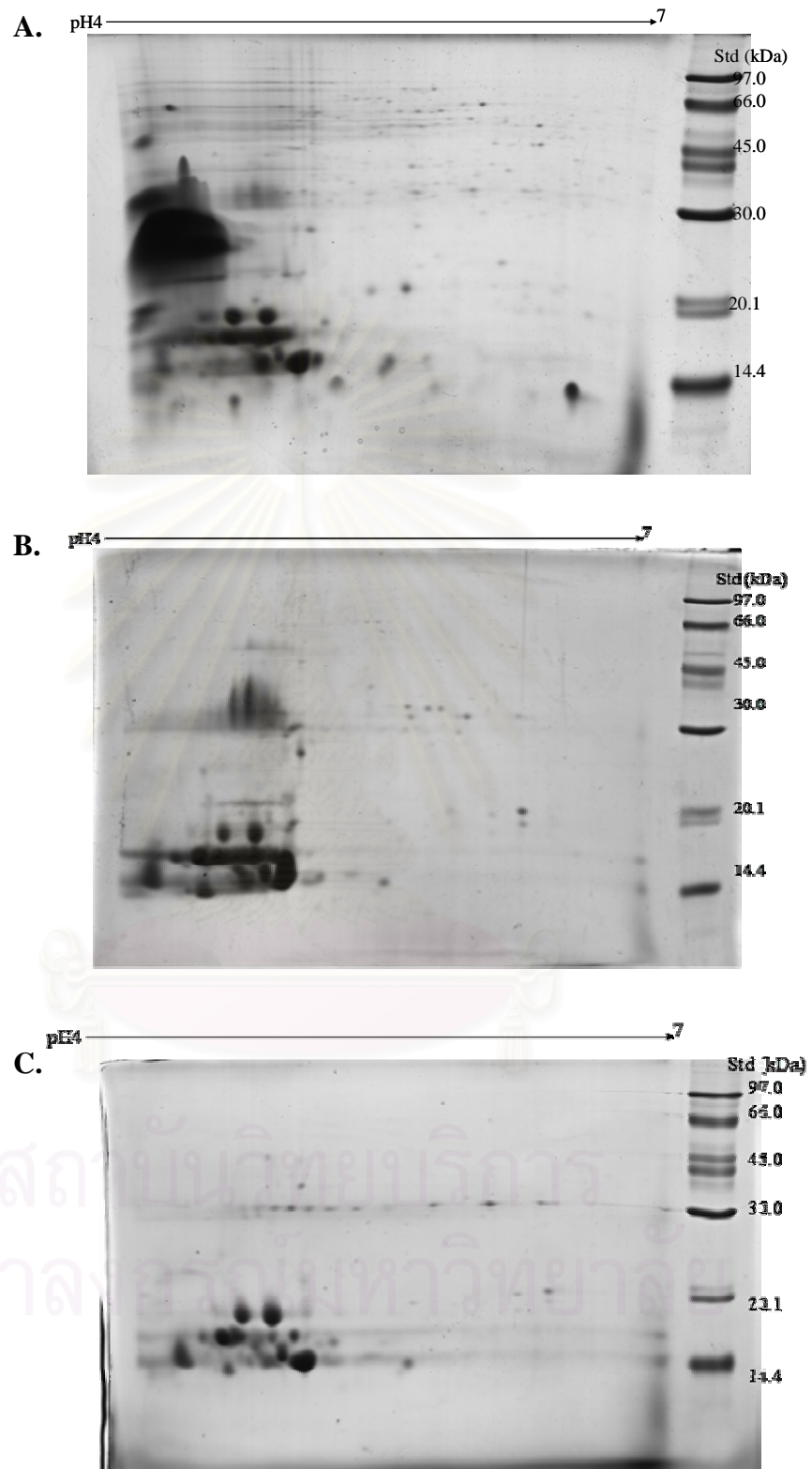
Figure 4.1 2-DE gels of *Curcuma longa* from three extraction and precipitation
procedures: A) Mg(NP-40)A, B) Phenol and C) Mg(NP-40)B extraction

4.1.2 2-DE of *Curcuma longa* from different sources

The 2-DE gels of *Curcuma longa* from three sources (Pak Khlong Talat, Bang Kae and Chatuchak) were shown in Figure 4.2. All 2-DE gels showed the similarity of protein pattern in range between 14 and 20 kDa.   2-DE results indicate that plants from difference sources had the same major proteins. Moreover, almost of all protein spots were shown in a small acidic region with pI range between 4 and 6.5.
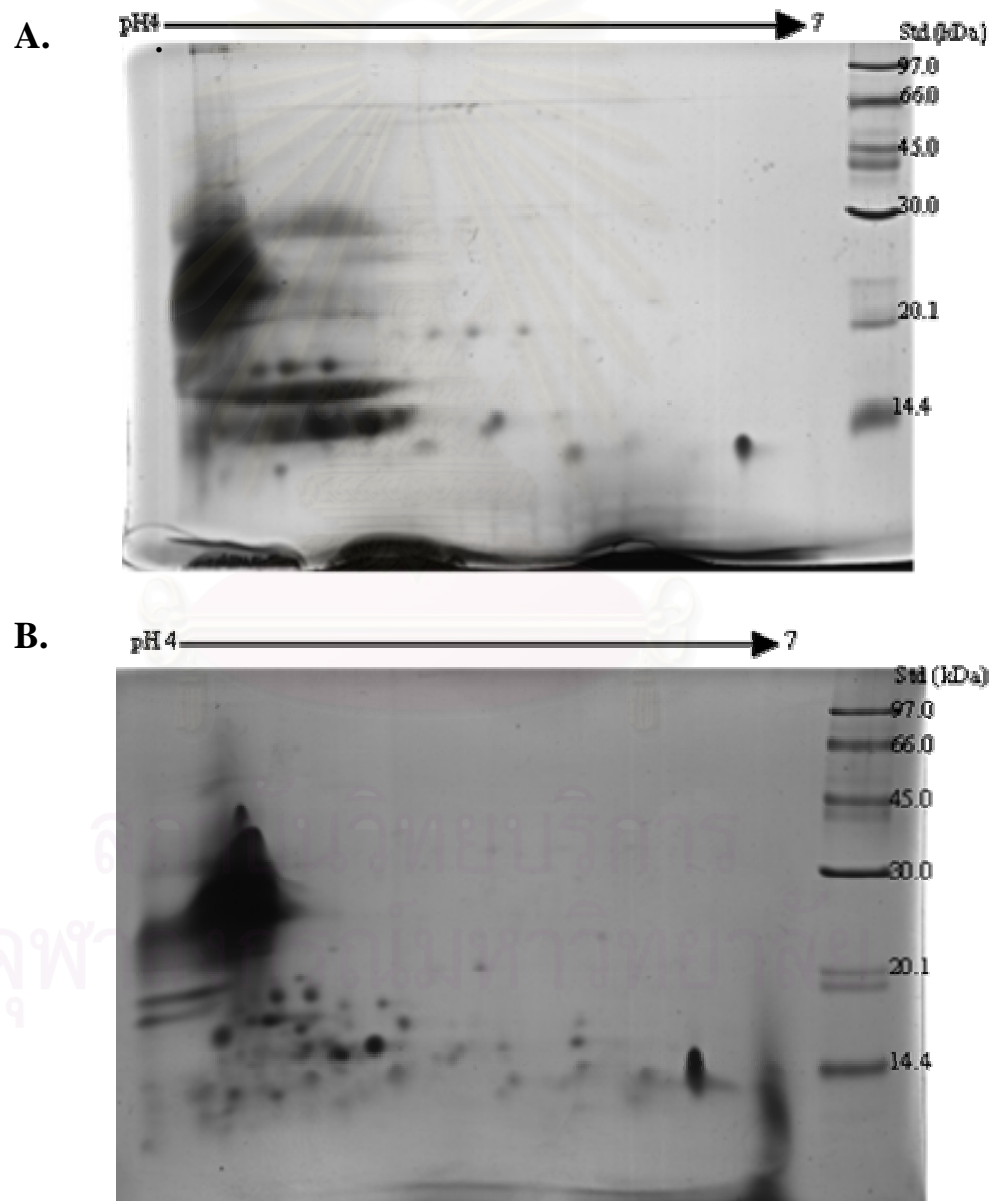
Figure 4.2 2-DE gels of *Curcuma longa* from different sources: A) From Pak Khlong Talat, B) Bang Kae and C) Chatuchak

Figure 4.2 2-DE gels of *Curcuma longa* from different sources: A) From Pak Khlong Talat, B) Bang Kae and C) Chatuchak (continued)

4.1.3 Precision of 2-DE method

Crude protein of *Curcuma longa* from Mg(NP-40)A extraction was analyzed by 2-DE three times. The protein patterns were the same for all extractions as shown in Figure 4.3. Molecular weight and pI of the main spots were determined by Bio-2D software (38) and demonstrated in Table 4.1. The results indicate good repeatability. The calculated values of molecular weight and pI of each spots were collected from each gel shown the similarity value with small standard deviation.
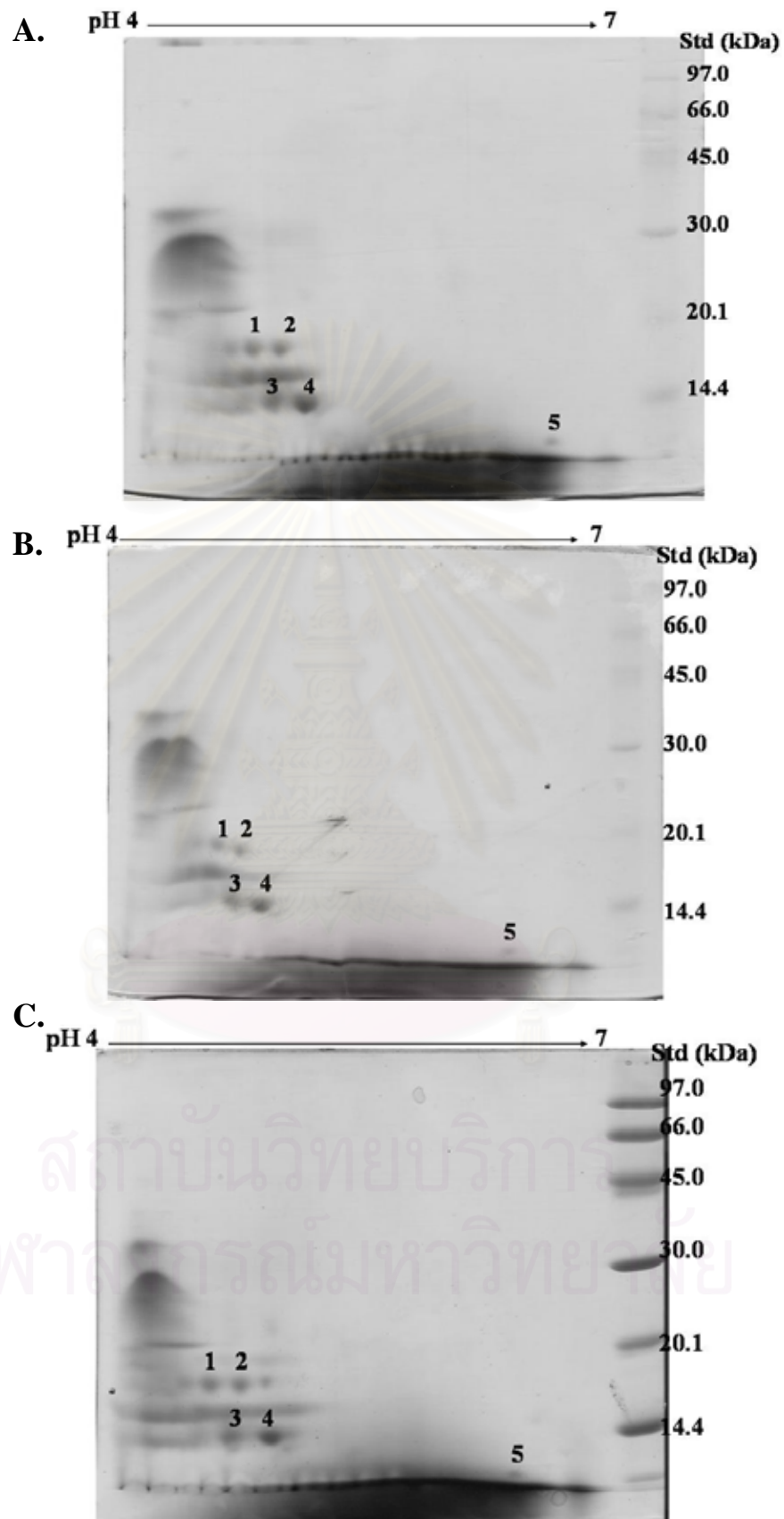
Figure 4.3 2-DE gels of *Curcuma longa* with three trials: A) first, B) second and C) third trial

Table 4.1 M.W. and pI of main spots from 2-DE gels of *Curcuma longa*

Trial No.	Molecular Weight					pI				
	Spot 1	Spot 2	Spot 3	Spot 4	Spot 5	Spot 1	Spot 2	Spot 3	Spot 4	Spot 5
1	17.49	17.49	14.36	14.4	11.62	4.46	4.63	4.58	4.79	6.11
2	17.86	17.82	14.2	13.92	10.97	4.48	4.62	4.6	4.75	6.08
3	18.84	18.39	14.69	14.2	10.81	4.48	4.61	4.64	4.71	6.05
Mean	18.06	17.90	14.42	14.17	11.13	4.47	4.62	4.61	4.75	6.08
Stdv.	0.70	0.46	0.25	0.24	0.43	0.01	0.01	0.03	0.04	0.03

4.1.4 Characterization of Zingiberaceae plants

The extracted proteins from 20 species of Zingiberaceous plants were analyzed by two-dimensional gel electrophoresis. The 2-DE gels of 20 species from Zingiberaceae were shown in Figure 1C-20C (Appendix C). Calculated molecular weight, pI of protein spots in 2-DE gel by Bio-2D software were presented in Table 1C-19C (Appendix C). Proteins that were specific to each species were also identified. These "marker proteins" may useful for characterize the species of Zingiberaceous plants. The M.W. and pI of marker proteins and number of protein spots from 2-DE gels of Zingiberaceous plant were represented in Table 4.2. Consequently, the results indicated that the 2-DE gels could be able to use for species identification in Zingiberaceous plants.

Table 4.2 Number of protein spots and marker proteins of each plant

Plants	Name of spot[a]	Number of protein spots (spot)	Marker proteins (pI, M.W.(kDa))	Common proteins
Curcuma longa	A_ _	24	A8 (4.56/19.92), A10 (4.70/19.92)	A13[b], A23[c]
Curcuma zedoaria	B_ _	29	B21 (5.16/15), B23 (5.22/15)	B16[b], B29[c]
Curcuma parviflora	C_ _	13	C3 (4.29/17.44), C5 (4.34/17.44)	
Curcuma sp. (Khamindum)	D_ _	32	D4 (4.54/12.45), D6 (4.62/12.45)	D12[b], D31[c]
Curcuma amarissima	E_ _	25	E1 (4.60/14.50), E12 (4.96,13.92)	
Curcuma sp. (Khantamala)	F_ _	27	F12 (4.44/14.65), F15 (4.54, 14.84)	F22[b], F33[c]
Curcuma aeruginosa	G_ _	32	G29 (5.96/10.51), G30 (5.98, 12.48), G31 (6.05, 13.08)	G19[b], G32[c]
Curcuma aromatica	H_ _	17	H14 (5.79, 13.89), H15 (5.87, 13.59)	
Curcuma comosa	I_ _	13	I8 (4.76, 11.19), I10 (4.85, 11.26)	
Curcuma sparganifolia	J_ _	22	J14 (5.07, 17.98), J22 (5.92, 14.40)	J25[c]
Curcuma sp. (Maleung)	K_ _	24	K1 (4.18, 12.82), K16 (4.81, 11.23)	K17[b], K24[d]
Curcuma sp. (Saligalinthong)	L_ _	15	L5 (4.33, 10.92), L8 (4.44, 11.65)	
Curcuma sp. (Payawan)	M_ _	11	M8 (5.03, 13.72), M9 (5.06, 14.16), M10 (5.67, 14.05)	
Curcuma sp. (Enleung)	N_ _	15	N14 (6.15, 11.63), N15 (6.15, 12.95)	N15[d]
Zingiber cassumunar	O_ _	35	O1 (4.09, 32.57), O3 (4.16, 35.83), O5 (4.25, 37.00)	

Table 4.2 Number of protein spots and marker proteins of each plant

Plants	Name of spot[a]	Number of protein spots (spot)	Marker proteins (pI, M.W.(kDa))	Common proteins
Zingiber ottensii	P_ _	21	P17 (5.40, 17.55), P18 (5.68, 20.00)	
Zingiber ofiicinale	Q_ _	11	Q7 (4.59, 30.85), Q9 (4.70, 30.85)	
Curcuma sp. (Haroynang)	R_ _	11	R4 (4.29, 11.11), R16 (5.93, 11.76)	R15[d]
Curcuma mannga	S_ _	2	S1 (5.23, 19.29), S2 (5.72, 19.75)	
Alpinia niga	T_ _	-	-	

[a] Name of spot in 2-DE gel such as A1, A2,_ _ , A23

[b] The same molecular weight and pI were 15.00 kDa and 4.83

[c] The same molecular weight and pI were 13.75 kDa and 6.12

[d] The sampe molecular weight and pI were 12.00 kDa and 6.12

For example, the 2-DE gel of *Curcuma longa* was used as a study model for Zingiberaceous plants. 2-DE gel of *Curcuma longa* showed twenty-three individual protein spots. Molecular weight and pI of each spots were shown in Table 1C. The spot in this gel was called "A" series. Several small protein spots in the same vicinity were group and called "A24". Many spots in 2-DE gel were represented with pI range between 4.00 to 5.00 and masses of 13.00-20.00 kDa. Specific protein spots for marker protein as A8 and A10, which had pI values of 4.56 and 4.70, respectively. These spots had a similar mass about 20 kDa. There are not all of samples showed the protein spot. The 2-DE gel of *Alpinia niga* has not shown any protein spot by Coomassie blue staining. It could be concentration of protein is lower than limit of detection the technique.

The protein patterns from 2–DE gels of different species from Zingiberaceous plant were compared. More or less protein spots were found in similar region; pI 4.00-5.00 and at M.W. 13.00-19.00 kDa for all species. The protein spots A13, B16, D12, F22, G19 and K17 in 2-DE gels of *Curcuma longa*, *Curcuma zedoaria*, *Curcuma* sp. (Khamindum), *Curcuma* sp. (Khantamala), *Curcuma aeruginosa* and *Curcuma* sp. (Maleung), respectively, which show the same molecular weight and pI, were use for the example of common protein spots from different species.   These spots were found as common protein in *Curcuma* genus. The mass spectra of tryptic fragment from protein spot A13 and D12 were shown similar peptide peaks at m/z of 2,166, 2,188, 2,203, 2,275, 2,298, 2,314 and 3,214 (Figure 5D and 11D in Appendix D). Then, the 2-DE gels of many species were compared and found spot with pI 6.12 were belong to A23, B29, D31, G32, F33, J25, K24, N15 and R15 (previous section). However, these spots are not show the same molecular weight. These spots can separate to two groups. The M.W. of first group were 13.75 kDa; A23, B29, D31, G32, F33 and J25. The M.W. of the second group is 12.00 kDa; K24, N15 and R15. The mass spectra of tryptic fragment were not similar to significant peptide peaks.

The protein patterns of *Curcuma longa* from phenol extraction, Mg(NP-40)A and Mg(NP-40)B were similar. The protein patterns from phenol extract of *Curcuma longa*, *Curcuma* sp. (Khantamala), *Curcuma zedoaria*, *Curcuma aeruginosa*, *Curcuma amarissima* and *Zingiber cassumunar* were shown two spots with the same

molecular weight (Figure 4.4) but they were not presented by other extraction procedures. These two spots were called Z1 and Z2, which had pI values 5.64 and 5.99, respectively and molecular weight of 22 kDa. The tryptic fragment mass spectra are shown in Figure 33D and 34D. The results suggested the similar peptide peaks at m/z of 1,626 (Figure 33D and 34D).
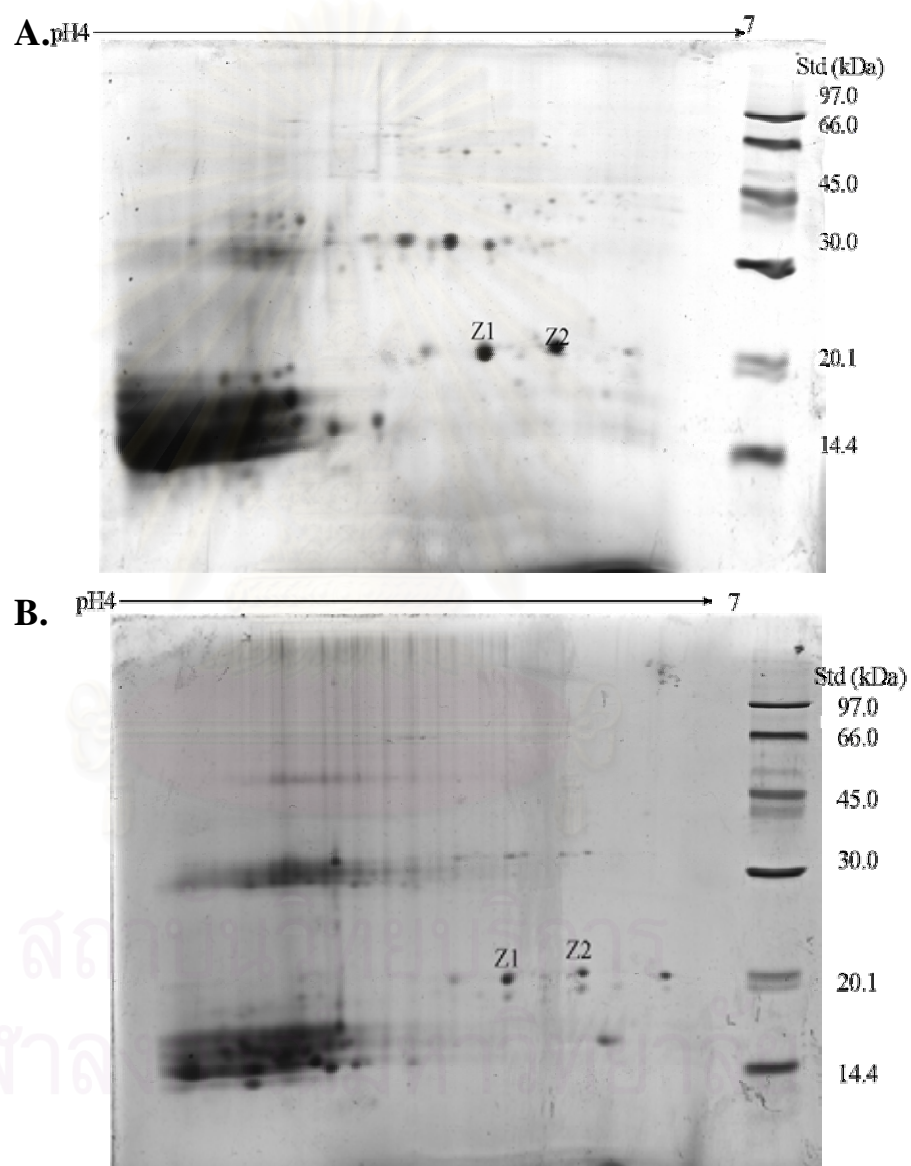
Figure 4.4 2DE-gels of Phenol extraction from A) *Curcuma* sp. (Khantamala),
B) *Curcuma zedoaria*, C) *Curcuma aeruginosa*, D) *Curcuma longa*,
E) *Curcuma amarissima* and F) *Zingiber cassumunar*
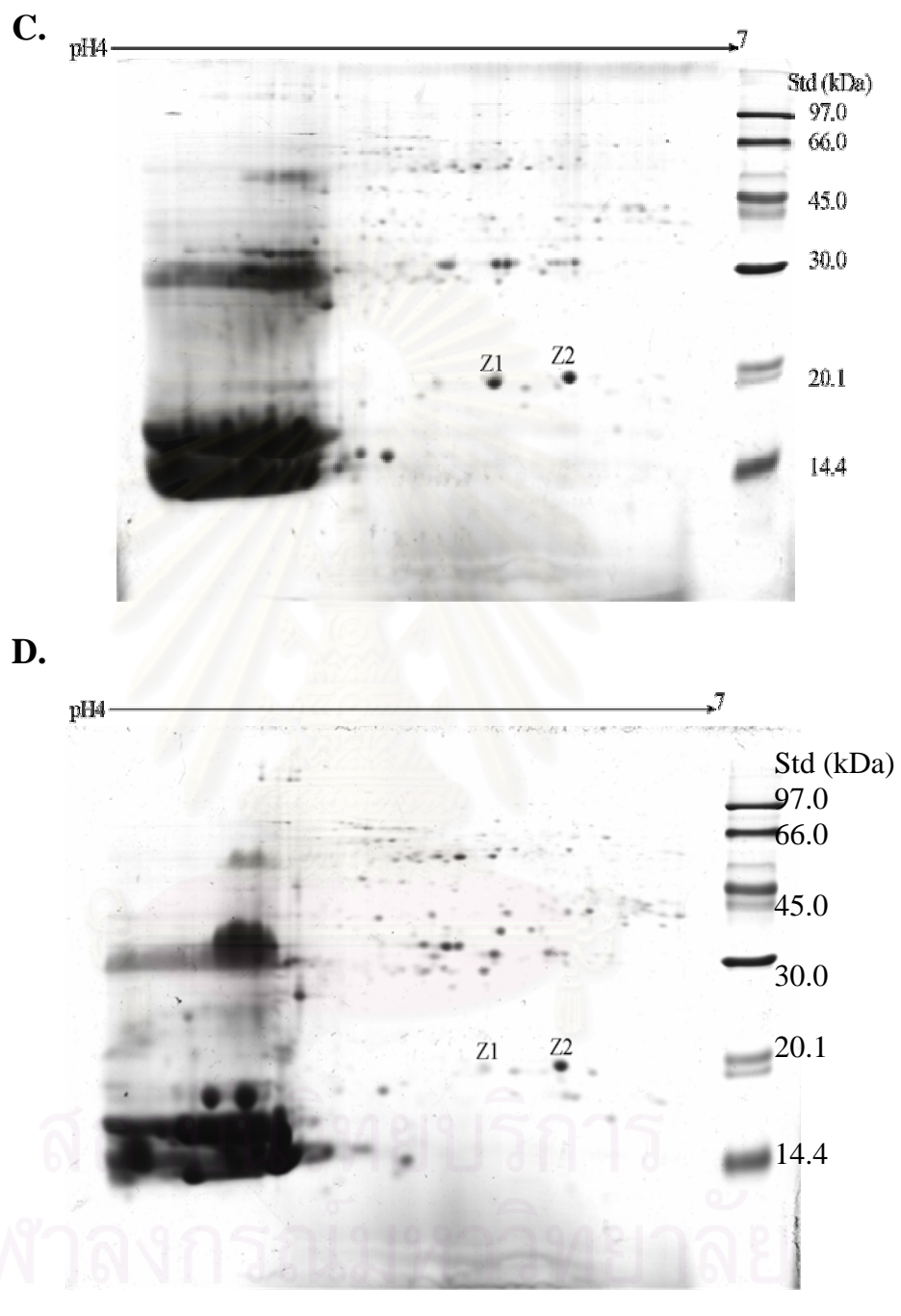
Figure 4.4 2DE-gels of Phenol extraction from A) *Curcuma* sp. (Khantamala),

B) *Curcuma zedoaria*, C) *Curcuma aeruginosa*, D) *Curcuma longa*,

E) *Curcuma amarissima* and F) *Zingiber cassumunar* (continued)
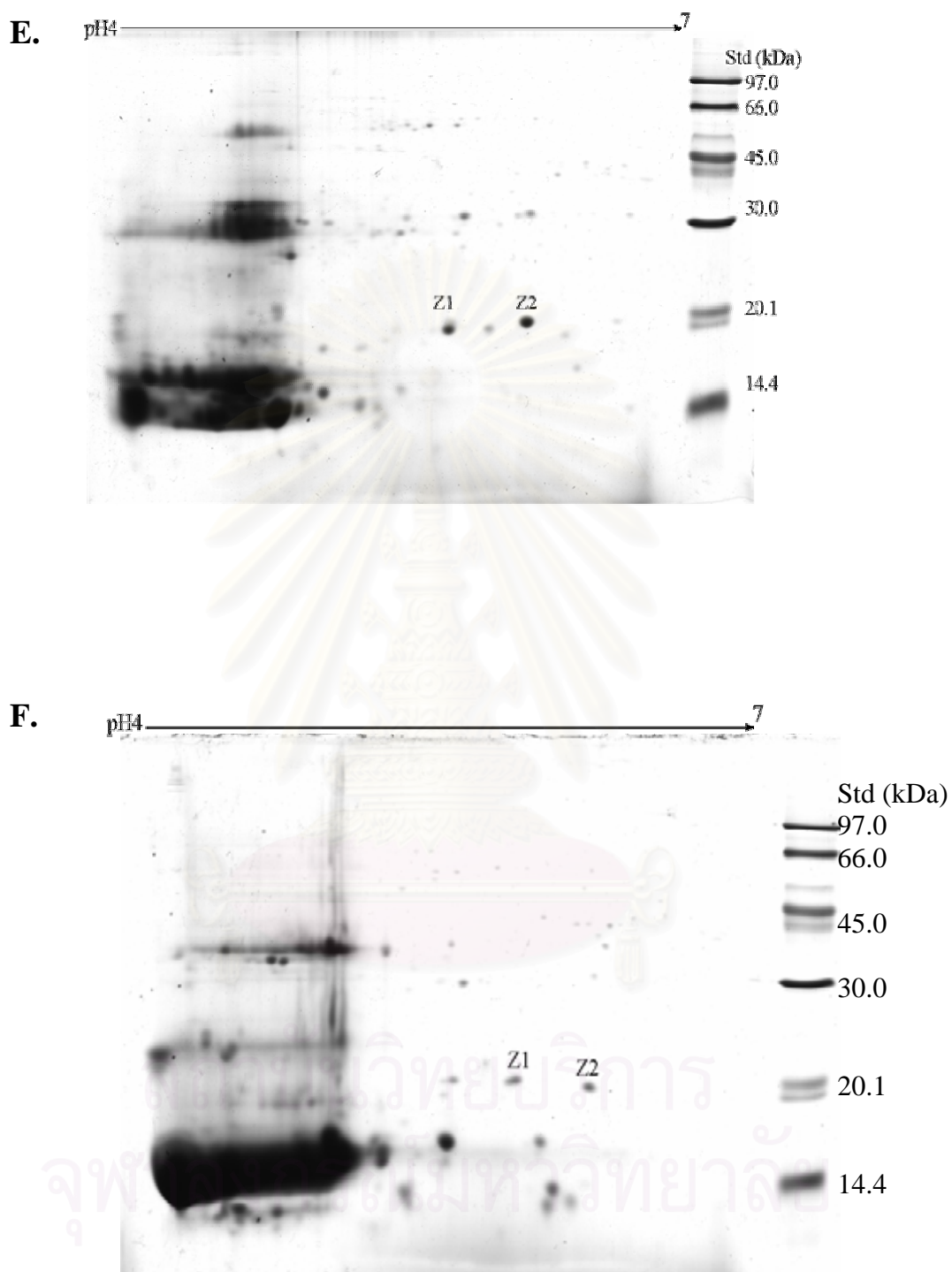
Figure 4.4 2DE-gels of Phenol extraction from A) *Curcuma* sp. (Khantamala),
B) *Curcuma zedoaria*, C) *Curcuma aeruginosa*, D) *Curcuma longa*,
E) *Curcuma amarissima* and F) *Zingiber cassumunar* (continued)

4.2 MALDI-TOF Mass Spectrometry

After the proteins from Zingiberaceous plants were analyzed by two-dimensional gel electrophoresis and detected with Coomassie blue stained, some visible protein spots were manually excised and subjected to digestion with trypsin as described in Section 3.2.4. The molecular weight of the tryptic fragments was determined using MALDI-TOF MS. The mass spectra of tryptic fragments are shown in Figure1D-34D (Appendix D). The significant peaks in mass spectra were shown in Table 4.3.

Table 4.3 Mass per charge of in-gel tryptic fragments of protein spots (spot number were shown in Section 4.1.4)

Spot[a]	Observed Mass[b] (m/z of)
A6	1113.09, 1435.35, 2229.06, 2308.71, 2308.71, 2344.02, 2615.95, 4392.65, 4913.94
A8	1071.20, 1086.42, 1458.38, 2189.45, 2277.92, 2300.46, 2306.25, 3181.82, 4932.49
A9	1112.18, 1416.64, 1434.44, 2274.63, 3138.87, 3155.69, 3213.26, 4913.53
A10	1112.51, 1430.56, 2274.67, 3139.63, 3154.72, 3212.89
A13	1108.34, 1141.64, 1243.49, 2188.12, 2203.97, 2276.85, 2314.17, 2364.10, 3181.22
A23	1084.56, 1999.22, 2190.44, 2206.45, 2279.06, 2300.76, 2316.93, 3184.07, 4922.53
B2	916.08, 930.19, 1231.68, 1909.79, 2230.86, 2784.86, 3183.96
B15	2179.67, 2288.16, 3379.58, 3630.85, 3733.35, 4907.24
D1	1275.85, 1320.05, 1364.55, 1408.10, 1452.44, 1496.48, 1584.69, 1628.48, 1672.84, 1716.64, 1761.57, 1907.96
D5	1070.39, 1086.44, 1292.47, 1382.52, 1690.77, 2861.45
D12	1083.64, 1113.16, 2187.57, 2203.29, 2275.71, 2313.60, 3139.71, 3178.77, 4912.91
E1	1248.50, 1946.26, 2227.25, 2274.65, 3165.44
E12	1280.82, 1950.90, 2279.92, 5100.76
F5	1453.45, 2749.74, 2843.98, 3364,38
F15	1084.46, 1393.94, 1817.68, 2231.57
F16	1084.88, 2231.17, 2279.09, 2295.25, 3182.99
F18	1099.76, 2289.98, 2629.32, 3670.19, 4906.65
G31	1069.10, 1100.86, 2191.75, 2279.12, 2302.09, 2318.38, 2862.42
I7	1220.14, 1236.00, 1280.23, 1324.60, 1368.35, 1456.90, 1500.68, 1572.93, 1633.46, 1676.70
K3	1068.08, 1083.99, 1100.78, 1244.22, 2188.35, 2195.51, 2275.84, 2856.65, 3156.82

Table 4.3 Mass per charge of in-gel tryptic fragments of protein spot (spot number
were shown in Section 4.1.4 (continued)

Spot[a]	Observed Mass[b] (m/z of)
K5	1067.43, 1083.30, 2233.45, 2275.84, 3173.25
L10	1068.21, 1113.38, 1511.04, 1554.58, 1616.53, 1643.53, 1686.29, 1731.45, 2276.61, 2308.95, 2616.74, 3203.95
L15	1493.53, 2616.19, 3152.16, 4902.77
M8	1214.32, 3063.25, 3073.48, 3084.06
M11	913.87, 930.96, 1083.49, 1098.92, 1143.02, 1159.07, 1373.03, 1606.61, 2187.55, 2274.34, 3139.17, 3177.12, 3214.31
N3	1113.85, 1437.62, 2170.32, 2616.00, 2804.28, 3215.22, 4914.63, 5106.10, 6585.04
O18	1171.59, 1760.82, 1802.38, 1818.52, 1998.92, 2124.48, 2198.49, 2231.24, 2278.93, 2295.52, 2619.61, 3199.19, 3220.63, 3535.32, 3548.71
P3	1100.79, 1863.84, 2230.99, 2686.10, 2860.55
P4	944.71, 1068.19, 1084.18, 1268.43, 1678.78, 2189.18
R5	1000.12, 1062.32, 1068.00, 1083.78, 1236.39, 1280.37, 1324.56, 1368.57, 1412.55, 1457.00, 1500.61, 1588.74, 1661.52, 2856.85
R7	1199.07, 1791.71, 2185.48, 2200.66, 2274.26, 3136.78
Z1	1062.72, 1103.67, 1119.37, 1449.55, 1625.39, 2857.01, 3527.57
Z2	1084.83, 1161.75, 1495.51, 1626.39, 2278.54, 2832.54

[a] Spot names correspond for protein found in Section 4.1.4

[b] Observed masses omit masses of keratin contamination and trypsin autolysis products

The molecular weight results were used for peptide mass mapping *via* the MASCOT program (www.matrixscience.com). The search parameters were included;

-100 ppm mass tolerance   - 1 missed cleavage, trypsin enzyme

- Source: green plant   - Database: MSDB

The search results were shown in Table 4.4. The scores expect and % protein coverage was too low to match any proteins. From the database, these were 2 match peptides from 8 peptides from spot A8. The search results suggested that these might be mixture of proteins or a novel protein, which have not yet been sequenced. Therefore, the sequencing techniques (tandem mass spectrometry) were required for further protein identification.

Table 4.4 Summary of proteins identified from Zingiberaceous plants by peptide mass fingerprinting

Spot[a]	Protein Name	Score	Expect	NCBI Accession No.c	Organism	Macthing peptides	% protein covered
A6	Putative DegPz protease	42	12	Q6Z806	*Oryza sativa* (japonica cultvar-group).	7(2)	11%
A8	Maturase	33	84	Q68LR1	*Phyllanthus epiphyllanthus*	8(2)	7%
A9	Maturase	38	31	Q9TK17	*Dalbergia sissoo*	8(2)	9%
A10	Maturase	36	42	Q6VFM1	*Globba leucantha*	7(3)	9%
A13	Halicase-like protein	34	79	Q84QRO_ORYSA	*Oryza sativa* (japonica cultvar-group).	9(2)	3%
A23	AF166114 NID	40	19	AAF43868	*Mesostigma viride*	9(2)	42%
B2	Nitrate transporter	41	14	Q76CO5	*Physcomitrella paten* (Moss).	7(2)	10%
B15	GblAAF07370.1.	51	15	Q9FN62	*Arabidopsis thaliana* (Mouse ear cress)	5(2)	23%
D1	Subtilisin-like serine protease	47	3.7	Q9FIM6_ARATH	*Arabidopsis thaliana* (Mouse ear cress)	12(3)	5%
D5	Maturase	36	40	Q8MFF5	*Duchesnea chrysantha*	6(2)	5%
D12	Maturase (Fragment)	39	26	O78239_9MAGN	*Chrysosplenium pilosum vas.*	9(2)	13%
E1	Maturase K	39	24	Q6DT98	*Primula capillaris*	5(2)	7%
E12	Ribosomal protein S7	40	18	Q67IE6	*Orchis rotundifolia*	5(2)	26%
F5	Maturase (Fragment)	38	28	Q8MDT9	*Cedrus deodara* (Deodar)	7(2)	8%
F15	Cytochrom P450	33	81	Q8HOI3	*Petumia hybrida* (Petamia)	4(2)	4%
F16	Putative subtilisin-like serine proteinase	44	6.7	Q6EPI5	*Oryza sativa* (japonica cultvar-group).	5(2)	7%
F18	sucrose-phosphate synthase (EC 2.4.1.14)	31	1.60E+02	T04103	Rice	5(1)	5%

Table 4.4 Summary of proteins identified from Zingiberaceous plants by peptide mass fingerprinting (continued)

Spot[a]	Protein Name	Score	Expect	NCBI Accession No.c	Organism	Matching peptides	% protein covered
G31	4-Coumarate-CoA ligase-like protein	38	27	Q84P21	*Arabidopsis thaliana* (Mouse ear cress)	8(2)	8%
I7	OSINBa0028M15.4 protein	32	1.20E+02	Q7FA45	*Oryza sativa* (japonica cultvar-group).	10(2)	25%
K3	Putative 4-Coumarate Co A ligase (EC 6.2.19.12) (Fragment)	32	1.00E+02	Q70CS9	*Festuca arumdimucea* (Tall fescue) (Schedonoris arudinaceus)	10(2)	23%
K5	Geranylgeranyl hydrogenase (E 1.3.1)	38	27	T12299	Common ice plant	5(2)	8%
L10	Putative receptor-type protein kinase LRK1.	34	70	Q6ZFB7	*Oryza sativa* (japonica cultvar-group).	12(2)	7%
L15	4-Coumarate-CoA ligase (EC 6.2.1.12)	43	9.5	Q42879	*Lithospermum erythrorhiza*	4(2)	6%
M8	Maturase (Fragment)	43	9.7	Q9MSU3	*Fitzroya curessoider*	4(2)	4%
M11	S-locus-specific receptor kinase (EC 2.7.1.-_)	29	-2.40E+02	Q719H4	*Welwitschia mirabilis* (Tree tumbo)		
N3	Maturase (Fragment)	34	66	JQ2387	*Oryza meyeriana*	11(2)	15%
O18	Putative protein	52	66	Q5W757	*Oryza sativa* (japonica cultvar-group).	15(3)	5%
P3	Maturase K	28	2.90E+02	Q71ML2	*Chimaphila maulata*	5(2)	7%
P4	ATP Synthase beta subunit	36	45	Q8MBF8	*Schizanthus pinnatus* (butterfly flower)	7(2)	7%

Table 4.4 Summary of proteins identified from Zingiberaceous plants by peptide mass fingerprinting (continued)

Spot[a]	Protein Name	Score	Expect	NCBI Accession No.c	Organism	Macthing peptides	% protein covered
R5	Putative isocitrate lyase	42	12	Q6Z6M4	*Oryza sativa* (japonica cultivar-group)	14(3)	6%
R7	Rust-resistance protein Lezi	41	1.40E+01	Q7XZE4	*Aegilops tauchii*	6(2)	2%
R17	Maturase K	36	44	Q8WJU5	*Laplacea portoricensis*	5(2)	5%
Z1	Integrase (Fragment)	38	28	Q9LKZ9_90RYZ	*Oryza rhizomatis*	7(2)	35%
Z2	AE016959NID	27	4.00E+02	AAP54086	*Oryza sativa* (japonica cultivar-group)	6(2)	3%

[a] Spot names correspond for protein found in Section 4.1.4

4.3 Tandem Mass spectrometry

The tryptic fragments of interest protein spots A8, A10, A23, D12, O3, O5, Z1 and Z2 were sequenced by Q-TOF mass spectrometer (Micromass, UK). The protein spot A23 and D12 are commonly proteins, which found in *Curcuma* species. Protein spot A8 and A10 are protein markers for *Curcuma longa* while O3 and O5 are protein markers for *Zingiber cassumunar*. Protein spot Z1 and Z2 are commonly found in phenol extraction of Zingiberaceous plant.

ProteinLynx Version (Micromass, Manchester, UK) to generate searchable peak lists processed the data. Initial protein identification was made by correlation of uniterpreted tandem mass spectra to entries in SWISS-PROT using Global server (Version 2.0, Micromass). In addition to, the amino acid sequence determination were obtained by de novo sequence interpretation. The searches were performed on the nr protein database and the EST others database at the BLAST interface of the US National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/BLAST). The search results were shown in Table 4.5.

Table 4.5 Putative amino acid sequences searches for Zingiberaceous plant proteins spots

Spot Number	Observed pI/MW[a]	Observed masses[b]	Putative amino-acid sequence[c]	Mass Diffrences[d]	Identification[e] (GenBank Accession No.)	Theoretical pI/MW[f]	Organism
A23	6.1/13.7	972.54	LQAFEPLR	-0.007	Glutelin typeB-2	9.4/56.2	*Oryza sativa*
		1100.63	LQAFEPLRK	-0.014	[precursor](Q02897)		(rice)
		1218.74	VIQPQGLLVPR	-0.011			
		1402.73	QKEFLLAGNNNR	-0.008			
D12	4.8/14.9	1113.62	LQTNG**D**LVV**R**	-0.01	Mannose-specific	4.8/11.9	*Aloe arborescens*
		1708.90	**GCNVAH**LQT**NG**D**LVVGL**	-0.046	Lectin (P49329)		(Kidachi aloe)
Z1	5.6/21.7	1316.72	**EF**NGVLTVT**L**PK	0.008	17.6 kDa class I	5.9/17.6	*Glycine max*
		1474.79	ACMENGVLTVTIPK	-0.038	heat shock protein		(soybean)
		1621.78	**MPSPVD**ETSAFANTR	-0.041	(P04795)		
		1659.77	**QPGFVD**ETSAFANTR	-0.006			
Z2	5.9/21.7	816.42	VEVEDGR	-0.029	17.6 kDa class I	5.3/17.6	*Helianthus annus*
		900.52	VLQISGER	-0.023	heat shock protein (P30693)		(common sunflower)
A8	4.5/19.9	1301.74	LSDLDSLLVVTK	-0.076	Nd	-	-
		1461.85	FLTETLTHELMK	-0.1			
A10	4.7/19.9	953.44	THNAELGVN	0.013	Nd	-	-
		2023.95	HVPCFNSNTLNNDVSHVK	-0.004			
O3	4.1/36.0	903.58	IYIVLRK	0.011	Nd	-	-
		2272.27	SIVHPSYNSNTLNNDIMLIK	-0.124			

Table 4.5 Putative amino acid sequences searches for Zingiberaceous plant proteins spots (continued)

Spot Number	Observed pI/MW[a]	Observed masses[b]	Putative amino-acid sequence[c]	Mass Diffrences[d]	Identification[e] (GenBank Accession No.)	Theoretical pI/MW[f]	Organism
		1134.59	SSGTSYPDVLK	-0.041			
O5	4.2/37.0	1065.58	YEDEINKR	-0.07	Nd	-	-
		1474.83	WELLQQVDTSTR	-0.094			
		1773.92	TPVTEAIATLKETLK	-00.68			

[a]Experimental pIs were estimated by portion as the linear gradient IPG strips and MWs were calculated accoding to those standard molecular weight (Amersham Bioscience) by using Bio2-D soft ware

[b]Masses are monoisotopic.

[c]In bold letters are indicated the amino acid which differ from retrieved sequence

[d]The mass in dalton correspond to difference between the observed peptide mass and predicted peptide mass deduced From de novo sequencing.

[e]The number in brackets corresponded to the GeneBank Accession numbers.

[f]Theoretical pIs and MWs calculated with computer pI/MW from Expasy (http://us.expasy.org/tools/pi_tool.html).

สถาบันวิทยบริการ

จุฬาลงกรณ์มหาวิทยาลัย

From the searching results, the peptide sequences in spot A23 were matched with Glutelin type B-2 [precursor] from *Oryza sativa* (rice). The product ion spectra of peptide sequence were shown in Figure 1E - 4E (Appendix E). The comparison of amino acid sequences of sample and Glutelin type B-2 [precursor] are displayed in Figure 4.5. Glutelin type–B 2 [precursor] is protein which found in Oryza sativa (rice). It has a molecular weight of 56,225 Da which consisted of 496 amino acids and 6 subunits (hexamer). The M.W. and pI of A23 protein was not matched with Glutelin type–B 2 [precursor]. Due to, protein spot A23 could be subunit or monomer of Glutelin type–B 2 [precursor].

```
Query: 11 RLQAFEPLRKVR------------------------VIQPQGLLVPRY 34
             LQAFEPLRKV                          VIQPQGLLVPRY
Sbjct: 49 RLQAFEPLRKVRSEAGVTEYFDEKNELFQCTGTFVIRRVIQPQGLLVPRY 98


Query: 35  RQKEFLLAGNNNRV 48
             RQKEFLLAGNNNRV
Sbjct: 195 RQKEFLLAGNNNRV 208
```

Figure 4.5 The amino acid sequence cover similar with amino acid residue of
           Glutelin type-B2 [precursor] from *Oryza sativa*

Two peptide sequences were obtained for spot D12 that matched with mannose-specific lectin from *Aloe arborescens* (Kidachi aloe). This mannose-specific lectin was found to have an agglutinating activity toward rabbit erythrocyte and mitogenic activity towards mouse lymphocytes. A comparison of amino acid sequences between peptides from D12 and mannose-specific lectin was shown in Figure 4.6.

```
Query:1 LQTNGDLVVRGCNVAH------------------LQTNGDLV-VG 26
          LQT+G+LVV    N A                  LQ NG+ V VG
Sbjct:56LQTDGNLVVQ--NSANRIIWQSNTGTGTNGDYLLVLQKNGNVVIVG 99
```

Figure 4.6 The amino acid sequence cover similar with amino acid residue of
           Mannose-specific lectin from *Aloe arborescens* (Kidachi aloe)

Note + are indicated the amino acid which differ from retrieved sequence

The product ion spectra of two peptide sequence were shown in Figure 5E-6E. However, M.W. and pI of protein spot D12 did not matching with Mannose-specific lectin. This could be due to protein spot D12 was a novel protein in a class of lectin.

Four sequences were obtained for spot Z1 that matched with 17.6 kDa class I heat shock protein from *Glycine max* (soybean). Finally, concerning Z2, matches were obtained with sequences from 17.6 kDa Class I heat shock protein from *Helianthus annus* (common sunflower). Protein spot Z1 and Z2 corresponded to a heat-shock protein. These proteins are high conserved, ubiquitous and abundant proteins essential for cellular viability. They are thought to protect cells from the stress of acute hyperthermia by binding and protecting partially denatured proteins during and after a heat-stress condition. The peptide sequences from protein spot Z1 and Z2 were covered on amino acid residue of heat shock protein that displayed in Figure 4.7-4.8, respectively.

```
                QPGFVDETSAFANTR              EFNGVLTVTLPK
   Query   1    MPSDVDETSAFANTR          ACMENGVLTVTIPK    29
                     E+SAF+NTR           ACMENGVLTVTIPK
   Sbjct  31   HVPTSSVSAENSAFVNTR -----ACMENGVLTVTIPK 139
```

Figure 4.7 The amino acid sequence cover similar with amino acid residue of
17.6 kDa class I heat shock protein from soybean

Note + are indicated the amino acid which differ from retrieved sequence

```
        Query: 1   VEVEDGRVLQISGER 15
                   VEVEDGRVLQISGER
        Sbjct: 72  VEVEDGRVLQISGER 86
```

Figure 4.8 The amino acid sequence cover similar with amino acid residue of
17.6 kDa class I heat shock protein from common sunflower

The product ion spectra of peptide sequence from protein spot Z1 and Z2 were shown Figure 7E-10E and 11E-12E, respectively. Moreover, molecular weight and pI values of two spots protein were not matched with heat shock protein. Protein spot Z1 and Z2 could be a novel protein in class of heat shock protein or had a partial amino acid sequence the same of heat shock proteins.

The MS/MS spectra of tryptic fragment from protein spot A8, A10, O3 and O5 were identified by using BLAST search. The results did not show any matched proteins because no more one-peptide sequence matched with one type of protein. From BLAST search were represent de novo sequencing of precursor ion of their peptides that shown in Table 1E-4E (Appendix E).

4.4 Hemagglutination activity

The biological activity of crude protein from *Curcuma longa, Curcuma zedoaria, Curcuma parviflora, Curcuma amarissima, Curcuma aeruginosa, Curcuma aromatica, Curcuma xanthorhiza, Curcuma sparganifolia, Curcuma sp.* (Khamindum, Haroynang, Maleung, Saligalinthong, Payanwan, Enleung and Khantamala), *Zingiber cassumunar* and *Zingiber ottensii* extraction with Mg(NP-40)A are shown hemagglutination activity (Table 4.6). A crude protein extract of *Curcuma sp* (Saligalinthong) exhibited strongest hemagglutinating activity, only at a concentration of 2×10^{-5} mg/ml.

Table 4.6 Hemagglutination activity of crude protein from Zingiberaceae with Mg(NP-40)A

Plants	Concentration of Heamagglution Activity ($\times 10^{-3}$ mg/ml)
Curcuma longa	2.00
Curcuma zedoaria	1.75
Curcuma amarissima	7.25
Curcuma aeruginosa	2.58
Curcuma aromatica	0.81
Curcuma xanthorhiza	10.00
Curcuma parviflora	10.00
Curcuma sparganifolia	5.00
Curcuma sp. (Khantamala)	0.059
Curcuma sp. (Maleung)	0.63
Curcuma sp. (Khamindum)	5.00
Curcuma sp. (Saligalinthong)	0.02
Curcuma sp. (Payawan)	10.00
Curcuma sp. (Enleung)	0.04
Curcuma sp. (Haroynang)	2.50
Zingiber cassumunar	0.039
Zingiber ottensii	1.25

CHAPTER V

CONCLUSION

In this research, the proteome approach was used to characterize different species in Zingiberaceous plants. Zingiberaceous plants were extracted and precipitated protein by Mg/MP-40 acetone procedure. Advantages of this method are this uncomplicated, short time, easy resolubilized and highly detection. The protein patterns of *Curcuma longa* from different sources were very similar. There was confirmed plant in different sources have the same protein. Moreover, from precision, method was shown a small standard deviation. After extraction and precipitation, crude proteins were analyzed by two-dimension gel electrophoresis. 2-DE gels of each plant are shown individual separate protein spots. The pI and molecular weight of each protein spots in 2-DE gels of 20 species were shown in Table 1D-19D. The marker protein spots of each species are described in section 4.1.4 such as *Curcuma longa* is shown specific or marker protein spots, which pI values 4.56 and 4.70. These proteins have similar mass about 20 kDa. *Curcuma* sp. (Khamindum) is shown specific protein spots, which pI values, 4.5 and 4.62. These proteins have similar mass at 12.5 kDa. From 2-DE gels, Zingiberaceous plants were displayed many protein spots in a similar region at pI range 4.0-5.0 and molecular weight 13-20 kDa. The common proteins were found which pI values about 4.83 and 6.12. Their proteins have masses about 17 and 13 kDa, respectively.

The in-gel tryptic digestion of protein spots were analyzed by MALDI-TOF MS and searched by MASCOT program. The results were not yield reasonable match with any proteins. Consequently, the tryptic fragment could not be identified by peptide mass fingerprint, which mean that these protein could be a novel protein and their sequences have not yet been reported. After that, the interest tryptic fragments of protein spots were selected to sequence by ESI-Q-TOF MS. Common protein spots A23 and D12 were found in *Curcuma* specie. Protein spot A8 and A10 were marker proteins in *Curcuma longa* while O3 and O5 were marker proteins in *Zingiber cassumunar*. Spots Z1 and Z2 were specific proteins found in 2-DE gel of phenol extraction of Zingiberaceous plants.

From amino acid sequence database search, the amino acid sequences of four peptides from tryptic fragment of spot A23 were RLQAFEPLR, RLQAFEPLRKV, RVIQPQGLLVPRY and RQKEFELLAGNNRV that found in the amino acid sequence of Glutelin type-B2 [precursor]. However, the result of pI and M.W. was not matched with that protein. Due to Glutelin type B2 [precursor] is a hexamer subunits, it may be protein spot A23 is a subunit of this protein. The amino acid sequence of two peptides from tryptic fragment of protein spot D12 were partial similar of Mannose-specific lectin from Kidachi aloe that are (-)LQTNGDLVVR(-) and (-)GCNVAHLQTNGDLVVGL(-). This protein may be a novel lectin. Protein spot Z1 has been sequenced. The results suggested that only one peptide, ACMENGVLTVTIPK is the same as part of amino acid residue of 17.6 kDa class I heat shock protein from soybean. The two amino acid sequences of tryptic fragment of protein spot Z2 are shown the same as amino acid residue of protein in 17.6 kDa class of heat shock protein from common sunflower which are (-)VEVEDGR(-) and (-)VLQISGER(-). Moreover, many peptide sequences of protein spot Z1 and Z2 have a partial similar amino acid residue of heat shock protein in many plants. Z1 and Z2 may be a novel heat shock protein. Proteins spot A8, A10, O3 and O5 are shown not significant match any protein in database. Finally, crude protein from seventeen species in Zingiberaceous plants have been tested the hemagglutinating activity. *Curcuma* sp. (Saligalinthong) exhibited strongest hemagglutinating activity.

REFERENCES

1. Mustafa, T. and Srivastava, K.C. Ginger (*Zingiber officinale*) in Migraine Headache, *Journal of Ethnopharmacology* **1990**, 29, 267-273.

2. Sirirugsa, P. Thai Zingiberaceae: Species diversity and their uses, *Pure Appl. Chem.* **1998**, 70, 2111-2118.

3. สมุนไพรเพื่อสุขภาพ, ปีที่ 1, ฉบับที่ 9, กรกฎาคม 2544, กรุงเทพฯ: บริษัท สำนักพิมพ์ ยูทิไลซ์ จำกัด, 31-42.

4. López, J.L., Role of proteomics in taxonomy: the *Mytilus* complex as a model of study, *J. Chromatography B* **2005**, 815, 261-274.

5. นันทวัน บุณยะประภัทร และ อรนุช โชคชัยเจริญพร, สมุนไพรไม้พื้นบ้าน, ครั้งที่ 1, กรุงเทพฯ :สำนักงานข้อมูลสมุนไพร คณะเภสัชศาสตร์ ม.มหิดล, **2539**.

6. Farnsworth, N.R., and Bunyapraphatsar, N. *Thai Medicinal plants: recommended for primary health care system*, Bangkok: Mahidol U., **1992**.

7. Scopes, K. and Robert, K. *Protein purification: A Laboratory Manual of Proteome Analysis*, 2nd ed. New York: Springer-Verlage, **1994**.

8. Harris, E.L.V and Angal, S. *Protein purification applications : a practical approach*, Oxford : IRL Press, **1990.**

9. Westermeier, R. *Electrophoresis in Practice*, 3rd ed. Weinheim: Wiley-VCH, **2001**.

10. Westermeier, R., and Naven, T. Proteomics Technology, *Proteomics in Practice*, Weinheim: Wiley-VCH, **2002**.

11. Laemmli, U.K. Cleavage of structure proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, 227, 680-685.

12. Cretghton, T.E. *Protein structure*, 2nd ed. New York: Oxford University Press, **1998**.

13. O' Farrell, P.H. High resolution two-dimensional electrophoresis analyses. *J.Biol.Chem.* **1975**, 250, 4007-4021.

14. Garrels, J.I. Two-dimensional gel electrophoresis and computer analysis of protein synthesized by clonal cell lines. *J.Biol.Chem.* **1979**, 254, 7961-7977.

15. Berkelman, T., and Stenstedt, T. *2-D Electrophoresis Using Immobilizes pH Gradients: Principle and Methods*, Amersham Biosciences AB, Sweden, **1998**.

16. Manufacture's instructor. *Protein Electrophoresis: technical manual*, Amersham Bioscience, USA, **1999**.

17. Kinter, M., and Sherman, N. E. *Protein Sequencing and Identification Using Tandem Mass Spectrometry*, New York: Wiley Interscience, **2000**.

18. Hoffmann, E.De, and Stroobant, V. *Mass Spectrometry: Principles and Applications*, 2nd ed. London: John Wiley& Sons, **2001**.

19. Karas, M., and Hillenkamp, F., Laser Desorption Ionization of Proteins with Molecular Masses Exceeding 10,000 Daltons, *Anal.Chem.* **1988**, 60, 2299-3201.

20. Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T., Matrix Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers, *Anal.Chem.* **1991**, 63, 1193A-1203A.

21. Whitehouse, C. M., Dreyer, R. N., Yamashita, M., and Fenn, J. B., Electrospray Interface for Liquid Chromatographs and Mass Spectrometers, *Anal. Chem.* **1985**, 57, 675-679.

22. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M., Electrospray Ionization for Mass Spectrometry of Large Biomolecules, *Science* **1989**, 246, 64-71.

23. Gaskell, S., Electrospray: Principles and Practice, *J. Mass Spectrom.*, **1997**, 32, 677-688.

24. Dass, C. *Principle and Practice of Biological Mass spectrometry*, United State: John Wiley& Sons, **2001**.

25. Hoffmann, E.De, Tandem Mass Spectrometry, *J. Mass Spectrom.*, **1996**, 31, 129-137.

26. Westermeier, R., and Naver, T. *Proteomic in Practice A Laboratory Manual of Proteome Analyses*, German: Wiley-VCH, **2000**.

27. Govorun, V.M., and Archakov, A.I., Proteomic technologies in Modern Biomedical Science, *Biochemistry (Moscow)* **2002**, 67, 1109-1113.

28. Apisariyakul, A., Vanittanakom, N., and Buddhasukh, D. Antifungal activity of turmeric oil extracted from *Curcuma longa* (Zingiberaceae). *Journal of Ethnopharmacology*. **1995**, 49, 163-169.

29. Jeenapongsa, R., Yoovathaworn, K., and Sriwatanakul, K. Anti-inflammatory activity of DMPBD, a phenylbutanoid from *Zingiber cassumuna .Annual research abstracts and bilography of non-formal publication, Mahidol univ*. **1995**, 22, 327.

30. Singh, G., Singh, O.P., and Maurya, S. Chemical and biocidal investigations on essential oils of some Indian Curcuma species. *Prog. Crystal Growth and Charact*. **2002**, 45, 75-81.

31. Takahashi, M., Kayano, T., Kowithayakorn, T., Hayashi, M., Komiyama, K., Ishibushi, M. Parviflorene A, a novel cytotoxic unsymmetrical sesquiterpene-dimer constituent from *Curcuma parviflora*. *Tetrahedron Letter*. **2003**, 44, 2327-2329.

32. Swiss-Prot and TrEMBL database (online). Available from: http://www.expasy.ch

33. Cao, H., Sasaki, Y., Fushimi, H., Komatsu, K. Molecular analysis of medicinally-used Chinese and Japanese *Curcuma* based on 18S rRNA gene and *trn*K gene sequence. *Biol.Pharm.Bull*. **2001**, 24, 1389-1394.

34. Lum, J.H.K., Fung, K.L., Cheung, M.S., Wong, M.S., Kwok, F.S.L., Hui, P.K., Lo, S.C.L., 2002. Proteome of oriental ginseng Panax ginseng C.A. Meyer and the potential to use it as an identification tool. *Proteomic*. **2002**, 2, 1123-113.

35. Kim, S.T., Cho, K.S., Jang, Y.S., Kang, K.Y. Two-dimensional electrophoresis analysis of rice protein by polyethylene glycol fractionation for protein arrays. *Electrophoresis*. **2001**, 22, 2103-2109.

36. Hurkman, W.J., and Tanaka, C.K. Solubilization of plant membrance proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol*. **1986**, 81, 802-806.

37. Bradford, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anla Biochem*. **1976**, 72, 248-254.

38. Bio-2D software handbook, VILBER LOURMAT, France.

39. www.ncbi.nlm.nih.gov/blast/blast.cgi

40. www.ncbi.nlm.nih.gov/Genbank/index.html

41. Wang, H.X and Ng, T.B. Alocasin, an anti-fungal protein from rhizome of the giant taro *Alocasia macrorhiza*. *Protein Expression and Purification*. **2003**, 28, 9-14

Appendices

Appendix A

Preparation of Water Saturated Phenol

Per each 500 g bottle of phenol: 500 ml of distilled H_2O

<u>Day 1:</u>

Place Phenol 500g and distilled H_2O in 50 °C water bath. Check on phenol every 15 min until completely melted. When melted, add 250 ml of distilled H_2O and shake/mix well. Be sure to warm water to 50 °C. Place in hood, store at room temperature overnight.

<u>Day 2:</u>

Add another 250 ml of room temperature distilled H_2O shake/mix well. Be sure to add room temperature water. Place in hood, store at room temperature overnight.

<u>Day 3:</u>

Do not shake or agitate. The phenol should be separated into 2 phases. The top phase being water the bottom being phenol. Remove the water phase using a 100 ml pipette, place waste in phenol collection bottle. Continue removing water and leave ~1/2 inch of water phase on top of phenol.

Appendix B

Solution in 1D and 2D Electrophoresis

A. Lysis buffer

Urea	4.8 g
CHAPS	0.4 g
Pharmalyte 3-10	200 μl
Double distilled H_2O	to 10 ml

Store at -20 °C

B. Rehydrate buffer

Urea	240 mg
CHAPS	10 mg
IPG Buffer or Pharmalyte	10 μl
DTT	1.4 mg
Double distilled H_2O	to 0.5 ml

Prepare just prior to use.

C. SDS equilibration buffer

	1st solution	2nd solution
1.5 M Tris-HCl, pH 8.8 (ml)	0.9	0.9
Urea (g)	3.24	3.24
Glycerol (ml)	2.7	2.7
SDS (mg)	90	90
DTT (mg)	90	-
Iodoacetamide (mg)	-	225
Double distilled H_2O (ml)	**3.015**	**3.015**

This solution prepared per 6 strips. Prepare just prior to use.

D. 30% T, 2.6 C monomer stock solution

Acrylamide	30 mg
N,N'-methylenebisacrylamide	0.8 g
Double distilled H_2O	to 100 ml

Filter solution through a 0.45 μm filter. Store at 4 °C in the dark

E. Resolving buffer

Tris base	4.5 g
Double distilled H_2O	70 ml
HCl	adjust to pH 8.8
Double distilled H_2O	to 100 ml

Filter solution through a 0.45 μm filter. Store at 4 °C

F. 10% SDS

SDS	5.0 g
Double distilled H_2O	to 50 ml

Filter solution through a 0.45 μm filter. Store at room temperature

G. 10% ammonium persulfate

Ammonium persulfate	0.1 g
Double distilled H_2O	1 ml

Prepare just prior to use.

H. SDS electrophoresis buffer

Tris-base	30.3 g
Glycine	144.0 g
SDS	10.0 g
Double distilled H_2O	to 10 l

Store at room temperature

I. Agarose sealing solution

SDS Electrophoresis buffer	100 ml
Agarose	0.5 g
Bromophenol blue	200 μl

Add all ingredients into a 500 ml Erlenmeyer flask. Swirl to disperse. Heat in a microwave oven low or heating until the agarose is completely dissolved. Do not allow the solution to boil over. Dispense 1.5 ml apendoft and store at room temperature.

J. Coomassie Blue staining solution

Coomassie Brilliant Blue G-250	0.1 g
Methanol A.R. grade	40 ml
Acetic acid	10 ml
Double distilled H_2O	50 ml

Dissolved Coomassie Brilliant Blue G-250 in methanol until it is completely

dissolved. Add Double distilled H_2O and acetic acid. Prepare just prior to use.

K. Destaining solution

	Destaining solution I	Destaining solution II
Methanol A.R. grade	40 ml	10 ml
Acetic acid	10 ml	5 ml
Distilled H_2O	to 100 ml	to 100 ml

Store at room temperature

Appendix C
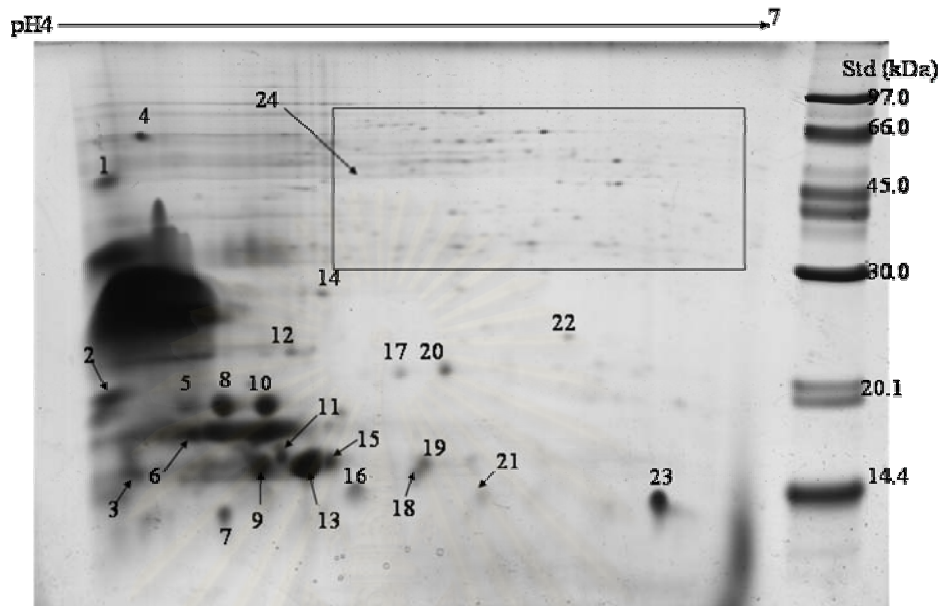
Figure 1C The 2-DE gel of *Curcuma longa* from Mg(NP-40)A extraction and
number of each spots was called A series (A1,A2,_ _ _, A24)
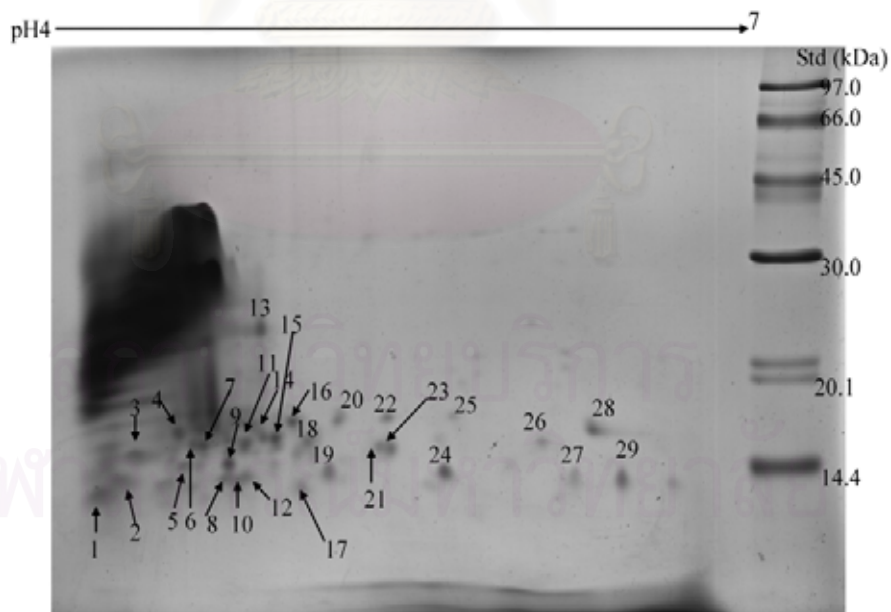
Figure 2C The 2-DE gel of *Curcuma zedoaria* from Mg(NP-40)A extraction and
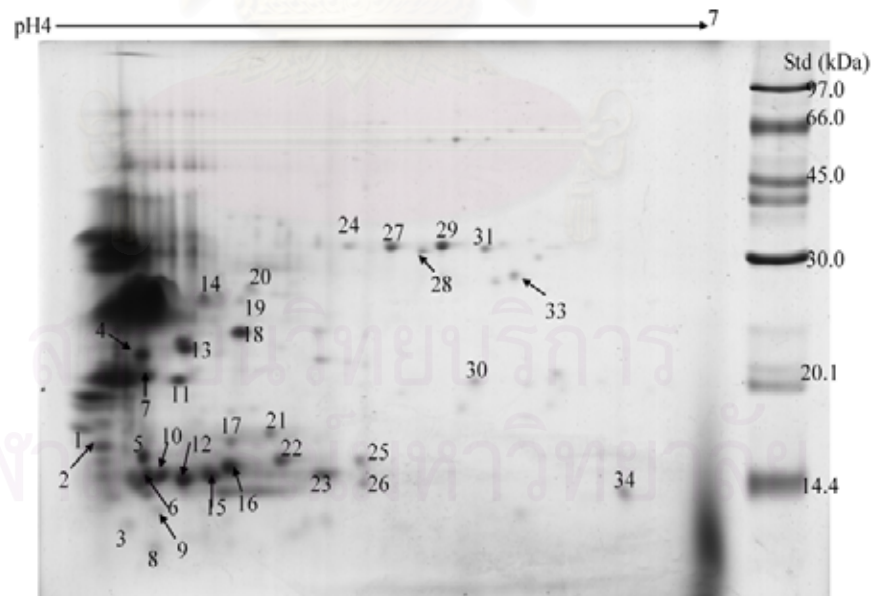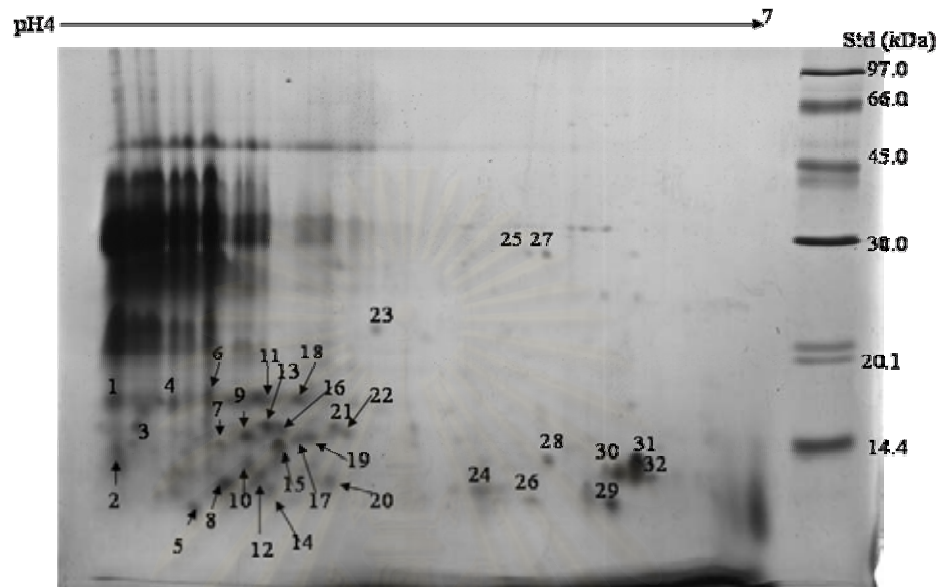number of each spot was named B series (B1,B2,_ _ _, B29)
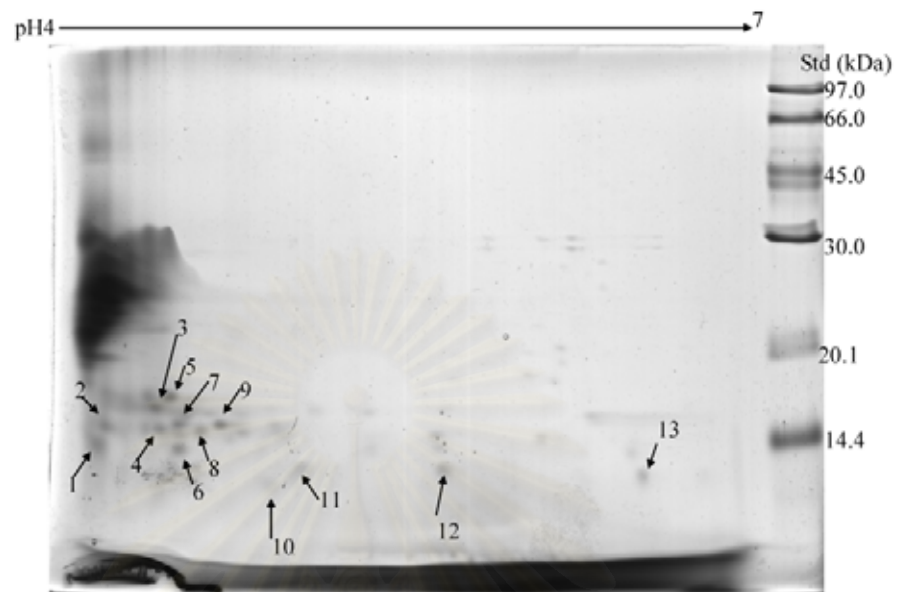
Figure 3C The 2-DE gel of *Curcuma parviflora* from Mg(NP-40)A extraction and
number of each spots was named C series (C1,C2,_ _ _, C13)
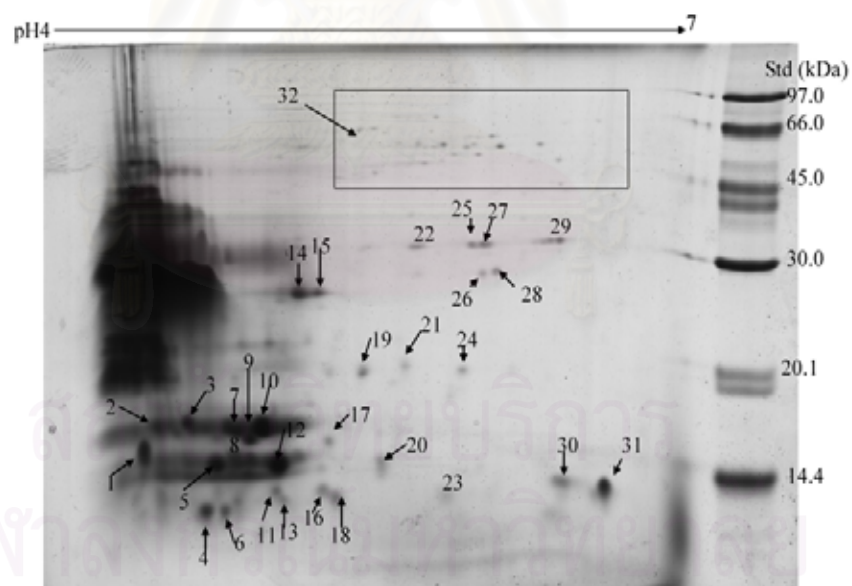
Figure 4C The 2-DE gel of *Curcuma* sp. (Khamindum) from Mg(NP-40)A extraction
and number of each spots was named D series (D1,D2,_ _ _, D32)
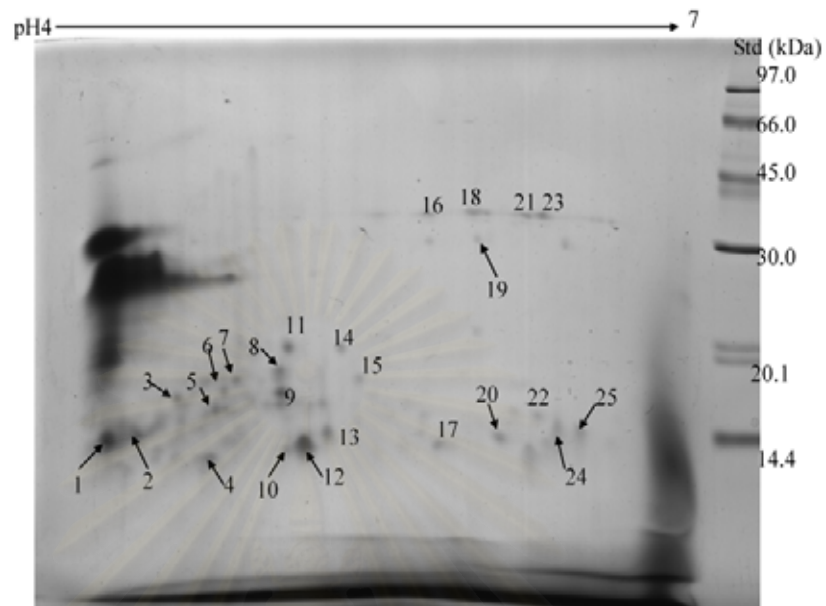
Figure 5C The 2-DE gel of *Curcuma amarissima* from Mg(NP-40) extraction
and number of each spots was named E series (E1,E2,_ _ _, E25)

Figure 6C The 2-DE gel of *Curcuma* sp. (Khuntamala) from Mg(NP-40)A extraction
and number of each spots was named F series (F1,F2,_ _ _, F34)

Figure 7C The 2-DE gel of *Curcuma aeruginaosa* from Mg(NP-40)A extraction
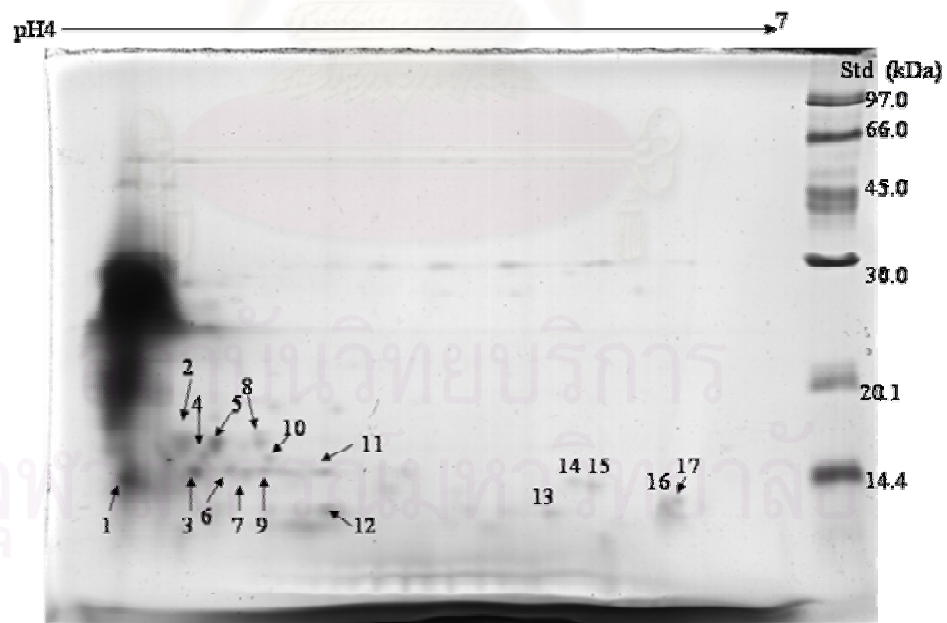and number of each spots was named G series (G1,G2,_ _ _, G32)

Figure 8C The 2-DE gel of *Curcuma aromatica* from Mg(NP-40)A extraction
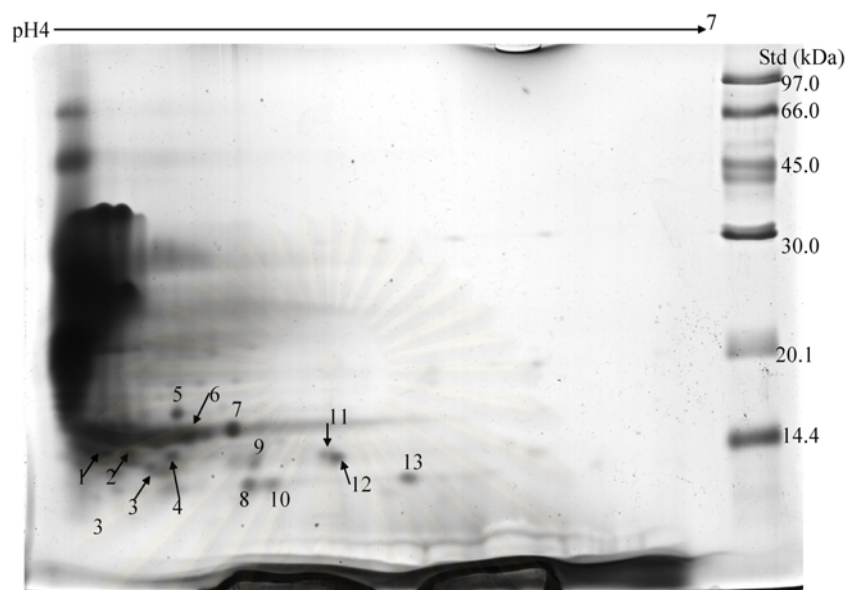and number of each spots was named H series (H1,H2,_ _ _, H17)

Figure 9C The 2-DE gel of *Curcuma comosa* from Mg(NP-40)A extraction
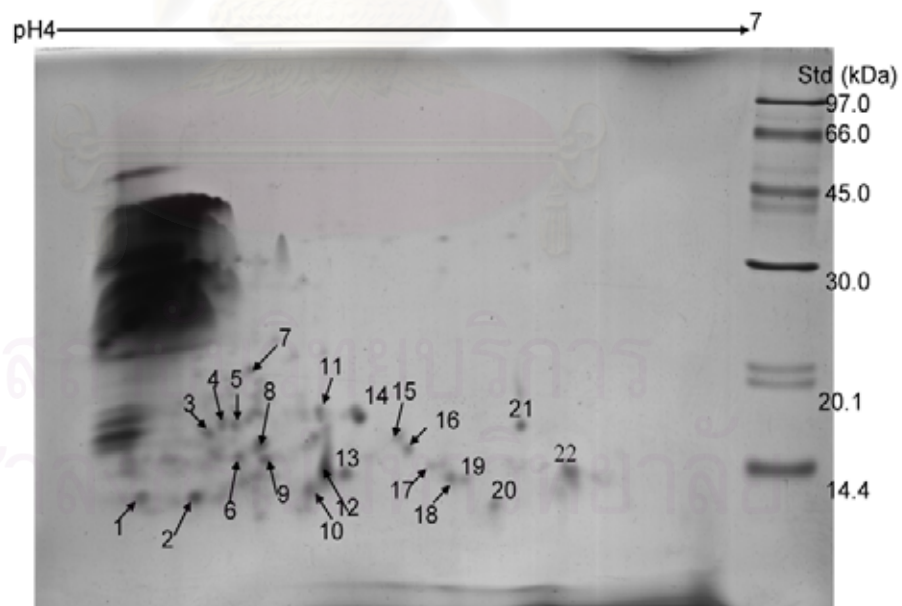and number of each spots was named I series (I1,I2,_ _ _, I13)

Figure 10C The 2-DE gel of *Curcuma sparganifolia* from Mg(NP-40)A extraction
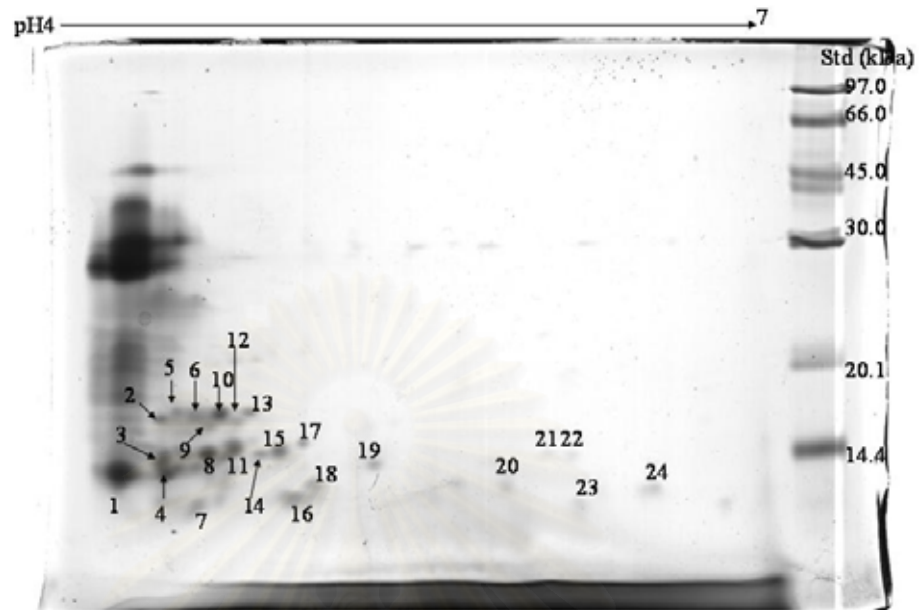and number of each spots was named J series (J1,J2,_ _ _, J22)

Figure 11C The 2-DE gel of *Curcuma* sp. (Maleung) from Mg(NP-40)A extraction
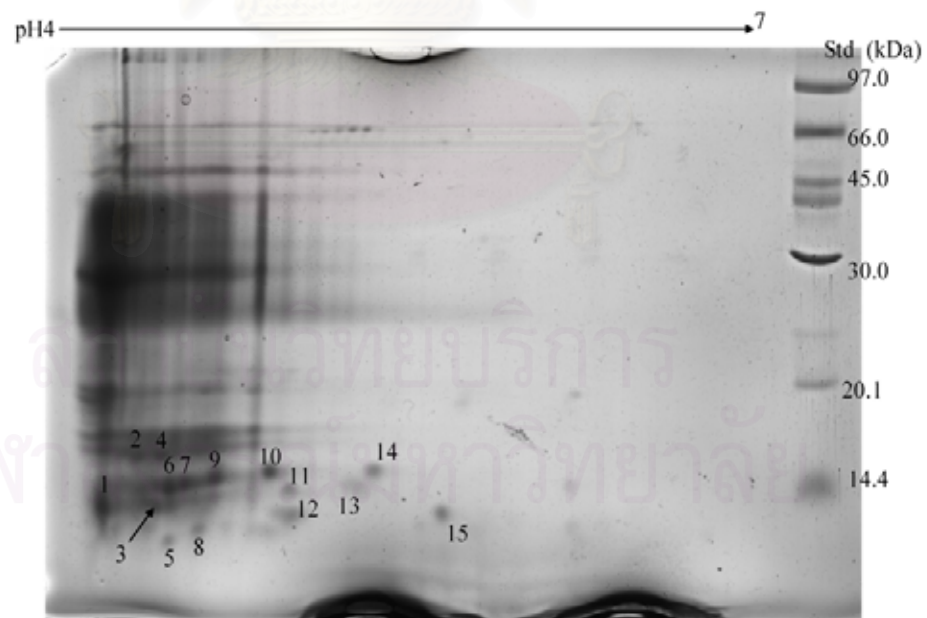and number of each spots was named K series (K1,K2,_ _ _, K24)

Figure 12C The 2-DE gel of *Curcuma* sp. (Saligalinthong) from Mg(NP-40)A
extraction and number of each spots was named L series (L1,_ _ , L15)
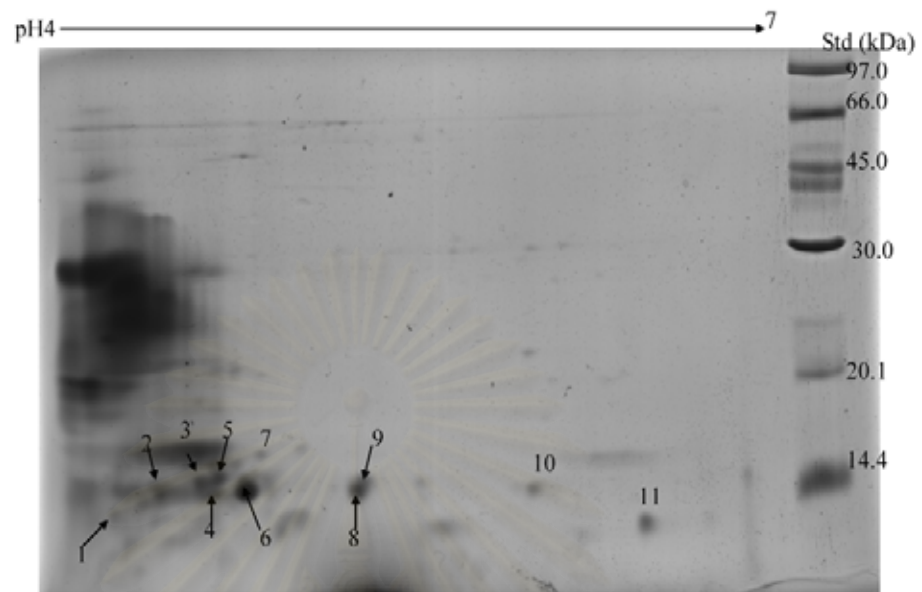
Figure 13C The 2-DE gel of *Curcuma* sp. (Payawan) from Mg(NP-40)A extraction
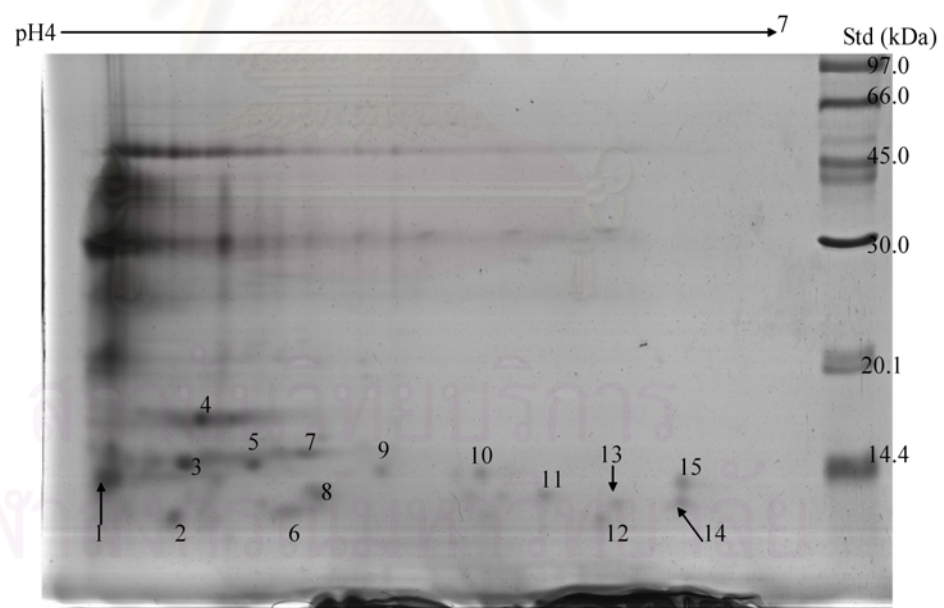and number of each spots was named M series (M1,M2_ _ _, M11)

Figure 14C The 2-DE gel of *Curcuma* sp. (Enlueng) from Mg(NP-40)A extraction
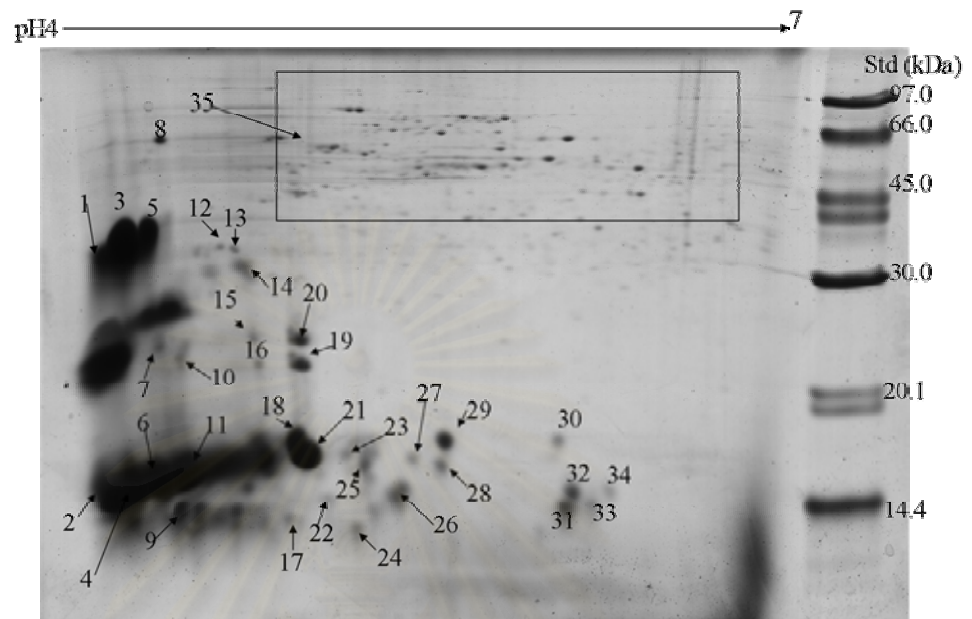and number of each spots was named N series (N1,N2,_ _ _, N14)

Figure 15C The 2-DE gel of *Zingiber cassumunar* from Mg(NP-40)A extraction and
number of each spots was named O series (O1,O2,_ _ _, O35)

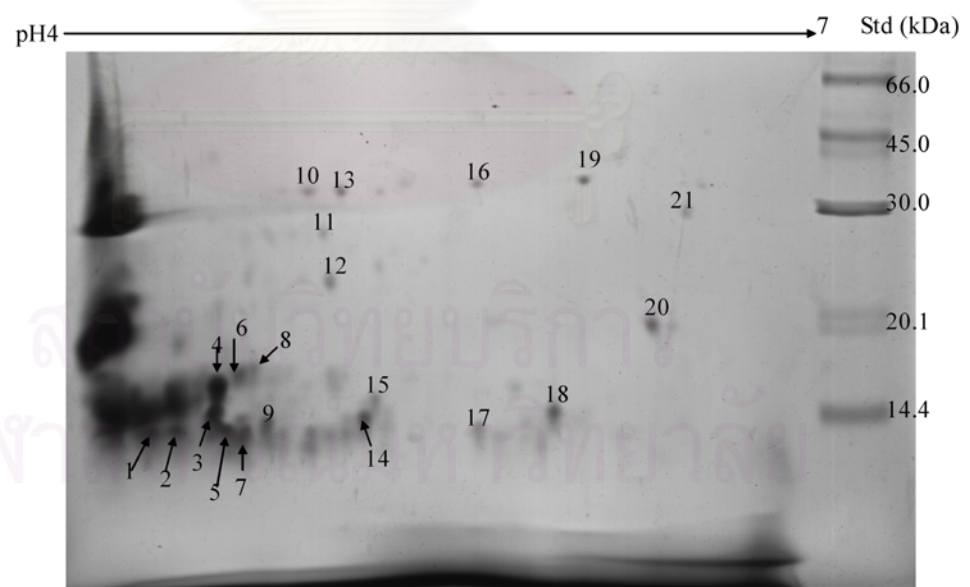**Figure 16C** The 2-DE gel of *Zingiber ottensii* from Mg(NP-40)A extraction and
number of each spots was named P series (P1,P2,_ _ _, P21)
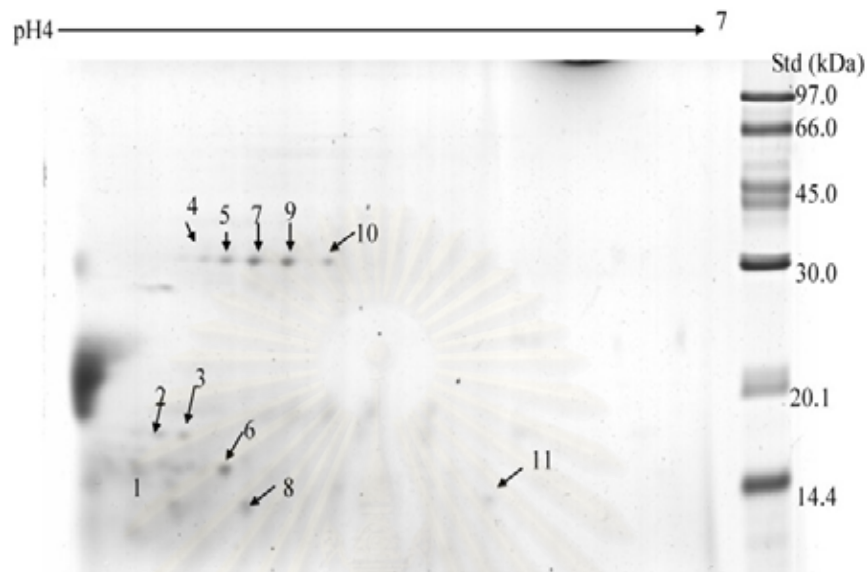
Figure 17C The 2-DE gel of *Zingiber officinale* from Mg(NP-40) extraction and
number of each spots was named Q series (Q1,Q2,_ _ _, Q11)
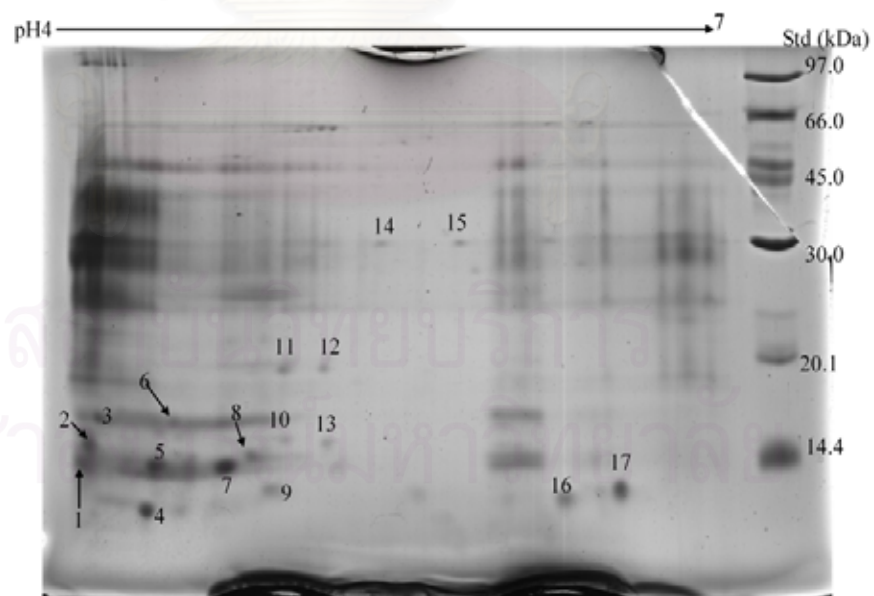

Figure 18C The 2-DE gel of *Curcuma* sp. (Haroynang)_ from Mg(NP-40)A
extraction and number of each spots was named R series (R1,R2,_ , R17)

Figure 19C The 2-DE gel of *Curcuma mangga* from Mg(NP-40)A extraction and
number of each spots was named S series
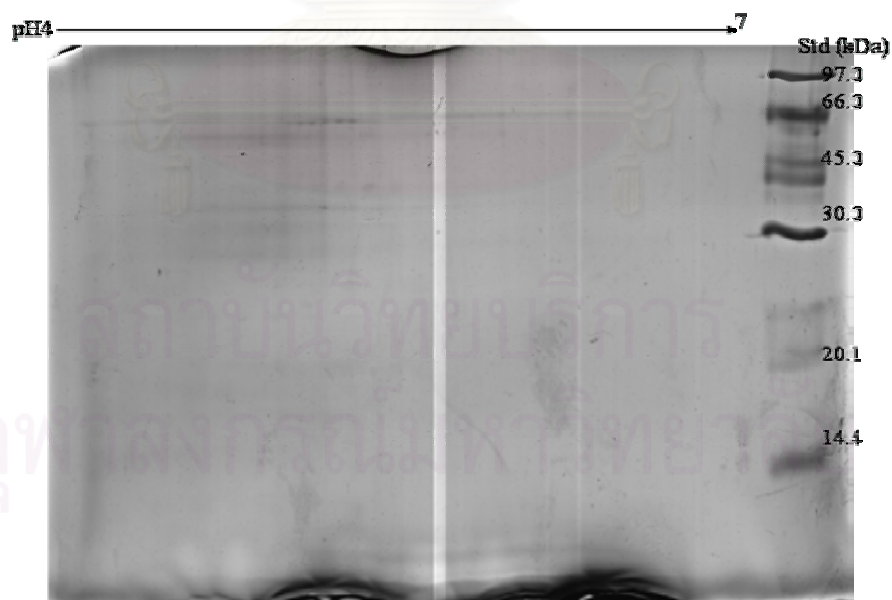
Figure 20C The 2-DE gel of *Alpinia niga* from Mg(NP-40)A extraction

Table 1C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma longa*

Spot No. (A_ _)	pI	M.W. (kDa)
1	4.14	20.33
2	4.14	49.53
3	4.23	15.72
4	4.26	66.00
5	4.43	19.92
6	4.45	18.35
7	4.56	13.19
8	4.56	19.92
9	4.69	16.14
10	4.70	19.92
11	4.75	18.37
12	4.80	22.98
13	4.86	16.29
14	4.91	28.11
15	4.94	16.61
16	5.03	14.61
17	5.19	21.72
18	5.24	15.67
19	5.28	16.35
20	5.35	21.90
21	5.48	14.30
22	5.79	24.11
23	6.11	13.97
24	5.12-6.13	35.12-88.53

Table 2C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma zedoaria*

Spot No. (B_ _)	pI	M.W. (kDa)
1	4.07	12.76
2	4.15	13.77
3	4.21	15.28
4	4.38	16.75
5	4.40	14.57
6	4.44	16.05
7	4.48	15.99
8	4.56	13.77
9	4.58	14.72
10	4.60	13.50
11	4.64	15.99
12	4.65	13.87
13	4.70	23.10
14	4.71	16.53
15	4.76	16.43
16	4.82	17.50
17	4.86	13.31
18	4.88	16.21
19	4.97	14.24
20	5.00	17.61
21	5.16	15.88
22	5.20	17.77
23	5.22	15.72
24	5.43	14.24
25	5.46	17.92
26	5.81	16.15
27	5.93	13.82
28	6.00	17.07
29	6.12	13.71

Table 3C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma parviflora*

Spot No. (C_ _)	pI	M.W. (kDa)
1	4.07	14.01
2	4.10	15.35
3	4.29	16.69
4	4.30	15.23
5	4.30	15.18
6	4.34	17.44
7	4.37	13.80
8	4.45	15.07
9	4.53	15.51
10	4.71	11.37
11	4.82	12.45
12	5.36	12.53
13	6.01	12.04

Table 4C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Khamindum)

Spot No. (D_ _)	pI	M.W. (kDa)
1	4.29	16.30
2	4.35	17.96
3	4.47	18.15
4	4.54	12.45
5	4.58	15.59
6	4.62	12.45
7	4.63	18.12
8	4.66	17.68
9	4.72	17.24
10	4.78	17.69
11	4.83	13.82
12	4.83	14.97
13	4.86	13.18
14	4.92	26.85
15	5.00	26.83
16	5.02	13.82
17	5.05	17.12
18	5.07	13.28
19	5.18	21.12
20	5.26	15.41
21	5.36	21.52
22	5.39	33.60
23	5.51	13.53
24	5.59	21.28
25	5.62	33.38
26	5.67	29.25
27	5.68	34.01
28	5.73	29.46
29	5.95	34.57
30	5.99	14.40
31	6.15	14.00
32	5.17-5.91	50.52-77.20

Table 5C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma amarissima*

Spot No. (E_ _)	pI	M.W. (kDa)
1	4.16	14.50
2	4.29	14.71
3	4.45	17.62
4	4.58	13.03
5	4.61	16.65
6	4.64	18.15
7	4.69	18.90
8	4.87	18.03
9	4.87	19.57
10	4.89	13.57
11	4.90	21.07
12	4.96	13.92
13	5.04	17.14
14	5.12	21.07
15	5.19	18.90
16	5.47	36.75
17	5.51	13.92
18	5.66	31.88
19	5.66	22.08
20	5.75	14.59
21	5.87	36.58
22	5.88	13.15
23	5.90	16.38
24	5.99	14.90
25	6.08	14.90

Table 6C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Khuntamala)

Spot No. (F_ _)	pI	M.W. (kDa)
1	4.01	17.46
2	4.13	16.59
3	4.23	11.97
4	4.28	21.73
5	4.29	16.01
6	4.3	14.69
7	4.3	20.60
8	4.33	10.43
9	4.34	12.72
10	4.36	14.89
11	4.42	20.35
12	4.44	14.65
13	4.45	22.08
14	4.52	25.66
15	4.54	14.84
16	4.62	15.33
17	4.63	16.83
18	4.66	23.08
19	4.67	25.51
20	4.70	26.67
21	4.77	17.36
22	4.82	15.82
23	4.98	14.89
24	5.08	32.21
25	5.12	15.72
26	5.14	14.31
27	5.24	32.21

Table 7C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma aeruginaosa*

Spot No. (G_ _)	Pl	M.W. (kDa)
1	4.09	17.38
2	4.11	13.19
3	4.18	15.56
4	4.31	17.21
5	4.39	10.12
6	4.44	17.54
7	4.49	14.50
8	4.50	11.82
9	4.58	15.10
10	4.59	12.97
11	4.63	17.65
12	4.64	12.48
13	4.66	15.79
14	4.67	11.27
15	4.71	14.24
16	4.72	15.62
17	4.78	17.60
18	4.78	14.86
19	4.83	14.86
20	4.90	11.88
21	4.92	15.62
22	4.96	15.27
23	5.07	21.89
24	5.45	11.33
25	5.64	28.27
26	5.65	13.36
27	5.71	28.37
28	5.72	13.36
29	5.96	10.51
30	5.98	12.48
31	6.05	13.08
32	6.10	12.21

Table 8C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel

of *Curcuma aromatica*

Spot No. (H_ _)	pI	M.W.
1	4.17	14.09
2	4.38	16.55
3	4.42	14.57
4	4.45	15.73
5	4.50	16.49
6	4.55	14.73
7	4.61	14.46
8	4.66	16.94
9	4.68	14.62
10	4.69	15.51
11	4.89	14.57
12	4.90	12.26
13	5.71	12.11
14	5.79	13.84
15	5.87	13.59
16	6.12	12.16
17	6.17	12.82

Table 9C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE

gel of *Curcuma comosa*

spot No. (I_ _)	pI	M.W. (kDa)
1	4.20	13.40
2	4.29	13.33
3	4.38	12.43
4	4.46	14.46
5	4.48	16.21
6	4.52	14.66
7	4.69	14.07
8	4.76	11.19
9	4.78	12.60
10	4.85	11.26
11	5.05	13.33
12	5.10	13.09
13	5.38	11.67

Table 10C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE
gel of *Curcuma sparganifolia*

Spot No. (J_ _)	pI	M.W. (kDa)
1	4.29	12.5
2	4.49	12.61
3	4.54	16.69
4	4.60	17.60
5	4.65	17.44
6	4.66	15.31
7	4.71	20.92
8	4.75	16.27
9	4.77	15.41
10	4.94	13.04
11	4.98	18.41
12	4.99	14.29
13	5.07	14.18
14	5.13	17.98
15	5.28	16.77
16	5.32	15.88
17	5.41	14.63
18	5.48	13.91
19	5.53	13.80
20	5.65	12.12
21	5.75	17.49

Table 11C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Maleung)

Spot No. (K_ _)	pI	M.W. (kDa)
1	4.19	12.82
2	4.34	16.57
3	4.35	13.71
4	4.36	12.82
5	4.40	17.71
6	4.45	16.79
7	4.47	10.71
8	4.51	14.19
9	4.51	16.12
10	4.55	16.96
11	4.60	14.52
12	4.61	16.34
13	4.66	17.20
14	4.70	14.08
15	4.77	14.35
16	4.81	11.23
17	4.86	14.86
18	4.90	12.02
19	5.11	13.45
20	5.59	11.97
21	5.77	14.14
22	5.83	14.14
23	5.86	10.76
24	6.14	11.87

Table 12C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Saligalinthong)

Spot No. (L_ _)	pI	M.W. (kDa)
1	4.09	13.40
2	4.22	16.19
3	4.28	13.44
4	4.30	16.38
5	4.33	11.57
6	4.34	14.54
7	4.39	14.64
8	4.44	12.17
9	4.51	15.02
10	4.71	15.44
11	4.78	14.26
12	4.78	12.94
13	5.06	14.45
14	5.09	15.39
15	5.34	12.99

Table 13C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Payawan)

Spot No. (M_ _)	pI	M.W. (kDa)
1	4.16	12.35
2	4.32	13.57
3	4.47	13.98
4	4.51	13.42
5	4.53	14.37
6	4.63	13.63
7	4.68	15.57
8	5.03	13.72
9	5.65	14.16
10	5.67	13.84
11	6.07	12.00

Table 14C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Enlueng)

Spot No. (N_ _)	pI	M.W. (kDa)
1	4.11	13.21
2	4.35	10.74
3	4.39	14.28
4	4.44	17.19
5	4.63	14.28
6	4.73	11.21
7	4.81	14.95
8	4.86	12.26
9	5.08	13.58
10	5.43	13.53
11	5.66	12.21
12	5.85	10.53
13	5.92	11.64
14	6.15	11.63
15	6.15	12.95

Table 15C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Zingiber cassumunar*

Spot No. (O_ _)	pI	M.W. (kDa)
1	4.09	32.57
2	4.10	15.64
3	4.16	35.83
4	4.20	15.69
5	4.25	37.00
6	4.29	16.27
7	4.30	23.71
8	4.30	64.78
9	4.33	14.21
10	4.37	23.19
11	4.37	16.36
12	4.52	35.58
13	4.56	35.66
14	4.59	32.59
15	4.63	24.32
16	4.65	22.67
17	4.76	13.35
18	4.78	18.38
19	4.80	22.67
20	4.80	24.32
21	4.83	17.37
22	4.90	14.8
23	4.96	17.56
24	4.99	12.90
25	5.02	16.93
26	5.15	15.28
27	5.20	17.22
28	5.29	16.88
29	5.31	18.82
30	5.71	18.19
31	5.73	14.30
32	5.76	15.35
33	5.83	14.40
34	5.89	15.30
35	4.83-6.13	41.25-92.9

Table 16C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Zingiber ottensii*

Spot No. (P_ _)	pI	M.W. (kDa)
1	4.18	17.04
2	4.29	17.34
3	4.42	19.49
4	4.43	23.36
5	4.48	17.55
6	4.51	24.64
7	4.54	17.04
8	4.58	25.37
9	4.62	18.16
10	4.77	50.10
11	4.83	41.23
12	4.85	34.73
13	4.89	50.10
14	4.98	18.88
15	5.02	21.41
16	5.39	52.52
17	5.40	17.55
18	5.68	20.00
19	5.79	52.52
20	6.04	30.52
21	6.16	44.82

Table 17C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Zingiber officinale*

Spot No. (Q_ _)	pI	M.W. (kDa)
1	4.33	17.48
2	4.36	27.29
3	4.39	13.20
4	4.43	17.45
5	4.51	30.85
6	4.58	15.28
7	4.59	30.85
8	4.68	13.00
9	4.70	30.85
10	4.83	30.68
11	5.61	13.66

Table 18C Number of protein spots, pI and molecular weight (M.W.) in the 2-DE gel
of *Curcuma* sp. (Haroynang)

Spot No. (R_ _)	pI	M.W. (kDa)
1	4.01	13.70
2	4.06	14.95
3	4.10	16.52
4	4.29	11.11
5	4.32	13.72
6	4.40	16.33
7	4.60	13.70
8	4.70	14.26
9	4.77	12.31
10	4.82	15.32
11	4.83	19.49
12	4.98	19.54
13	4.99	15.09
14	5.21	30.69
15	5.52	30.69
16	5.93	11.76
17	6.14	12.21

Table 19C Number of protein spots, pI and molecular weight (M.W.) in 2-DE gel of
Curcuma mangga

Spot No. (S_ _)	pI	M.W. (kDa)
1	5.23	19.29
2	5.72	19.75

Appendix D
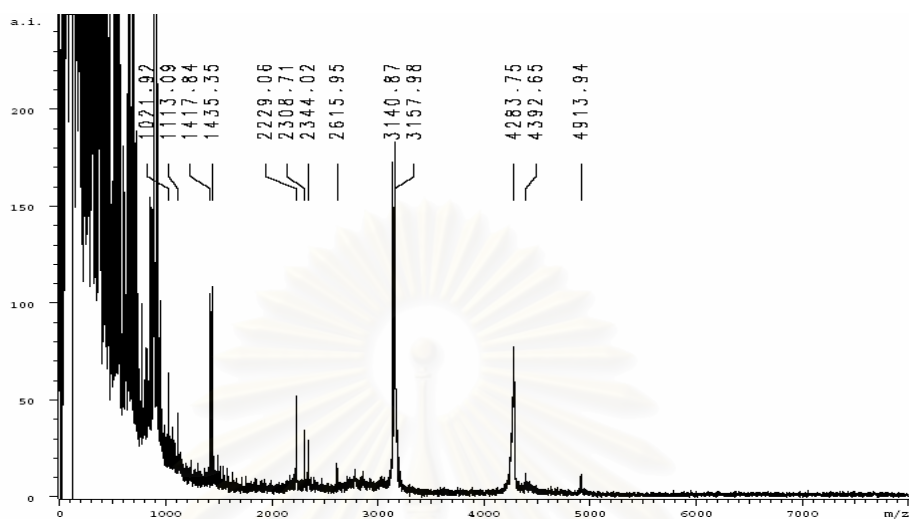
Figure 1D MALDI-TOF MS spectrum of tryptic fragment of spot A6
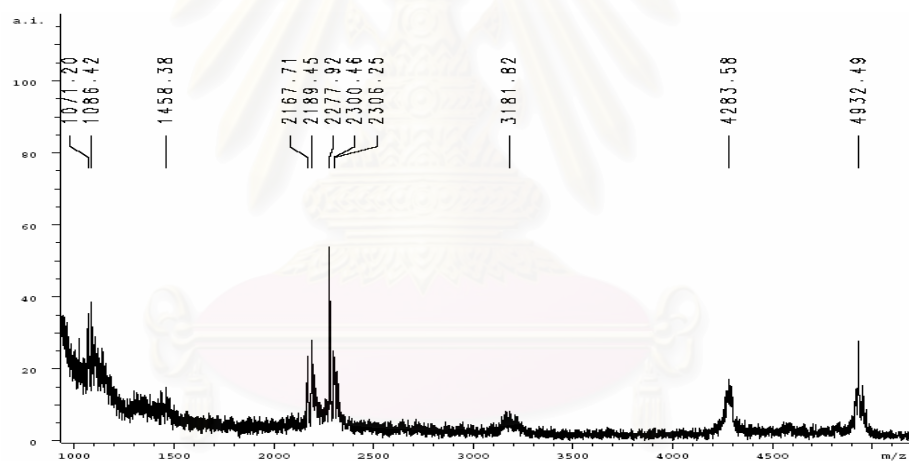
Figure 2D MALDI-TOF MS spectrum of tryptic fragment of spot A8
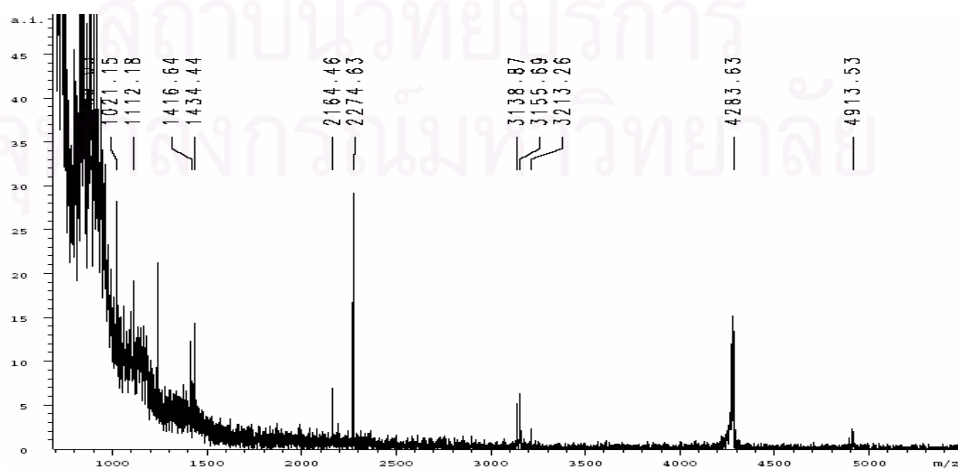
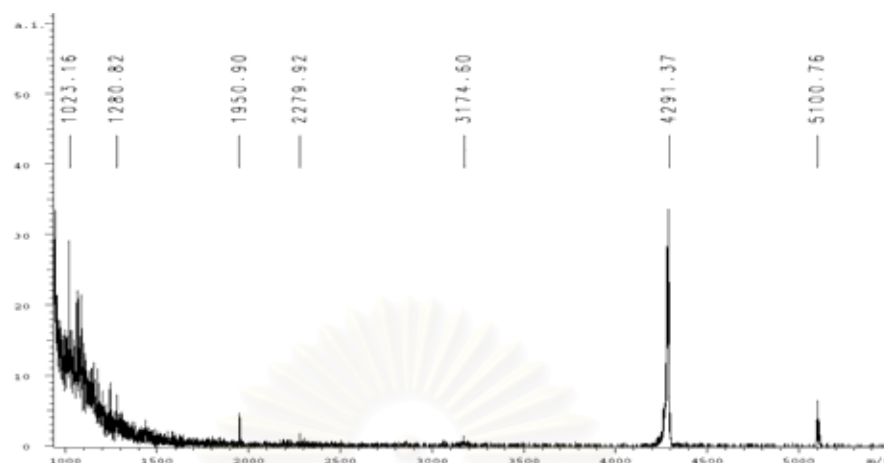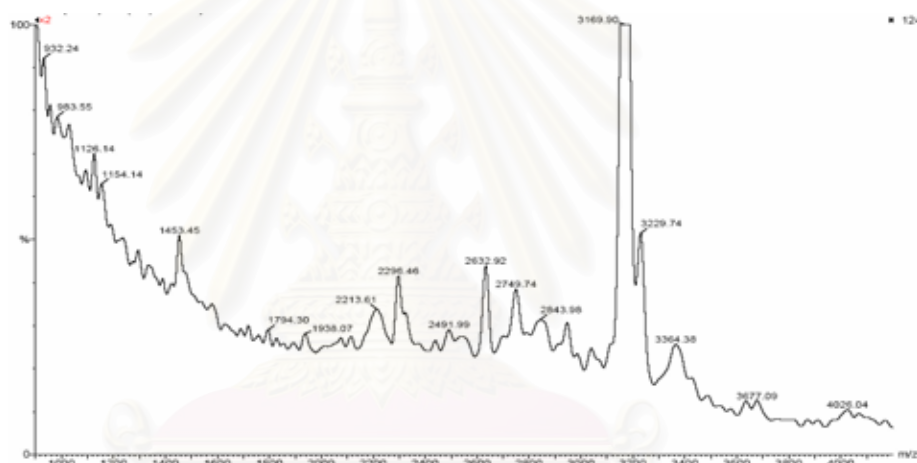**Figure 3D** MALDI-TOF MS spectrum of tryptic fragment of spot A9
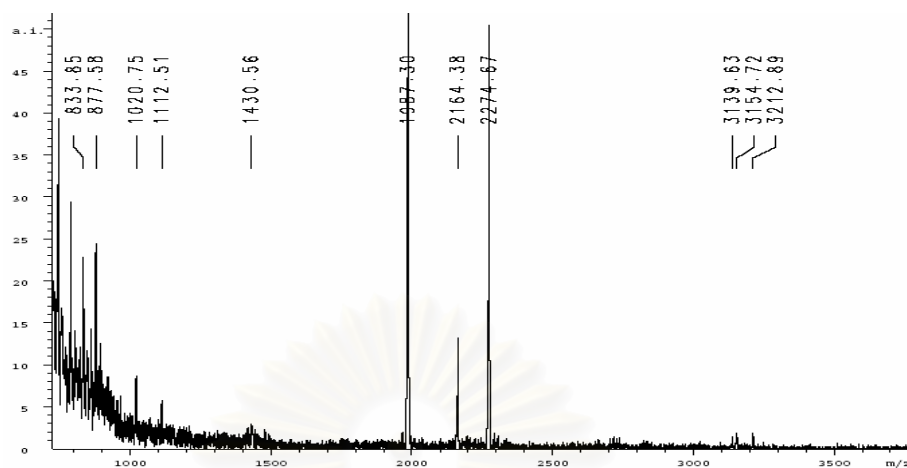
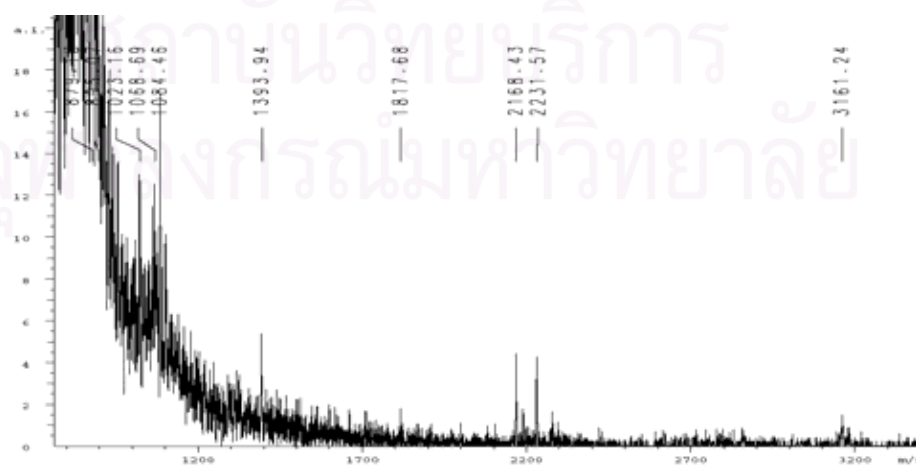**Figure 4D** MALDI-TOF MS spectrum of tryptic fragment of spot A10
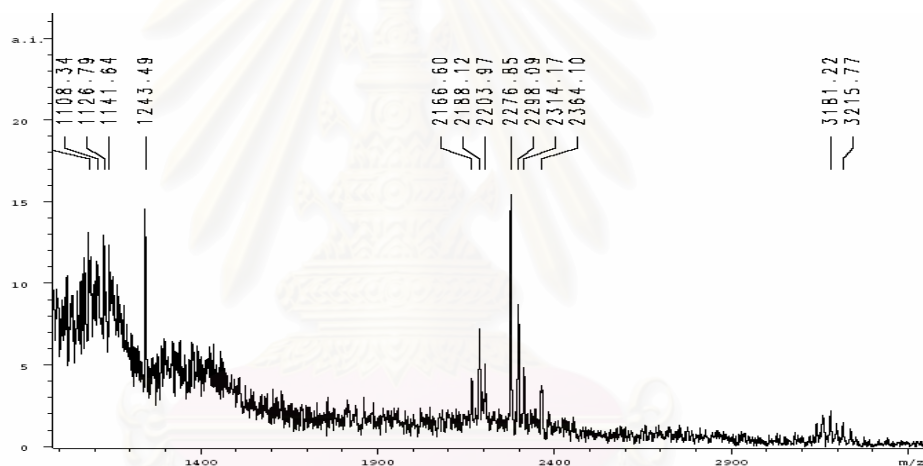
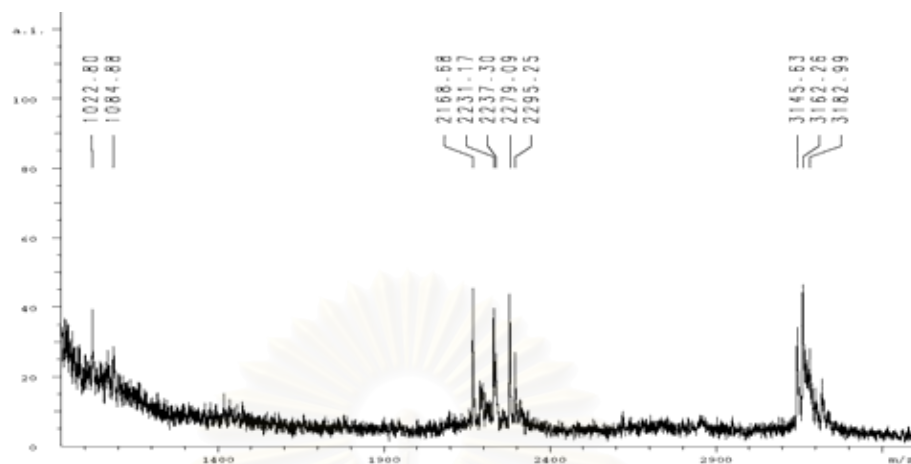**Figure 5D** MALDI-TOF MS spectrum of tryptic fragment of spot A13
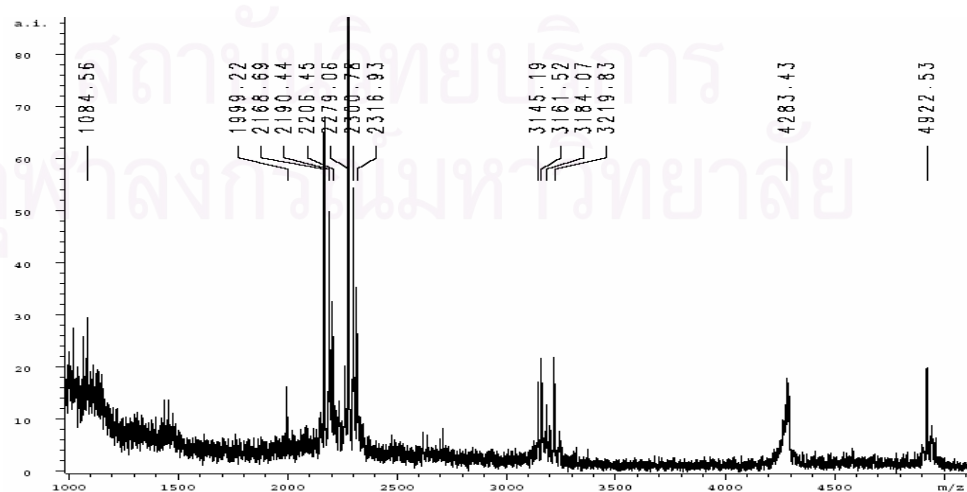
Figure 6D MALDI-TOF MS spectrum of tryptic fragment of spot A23
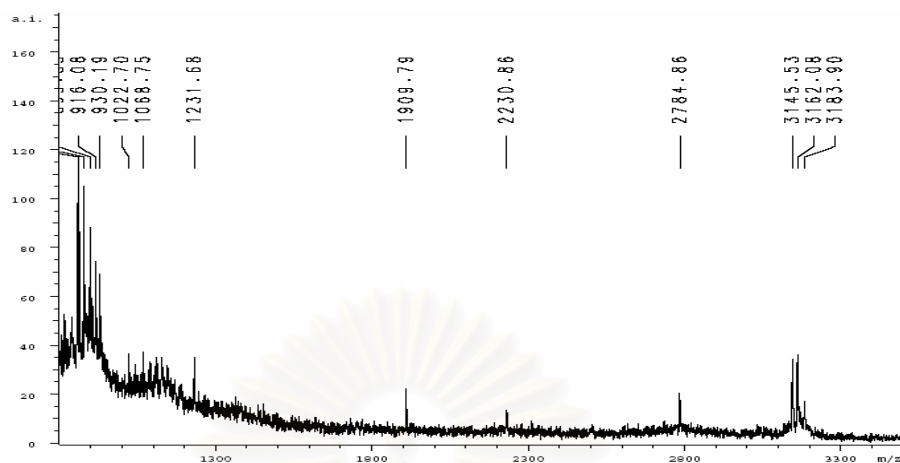
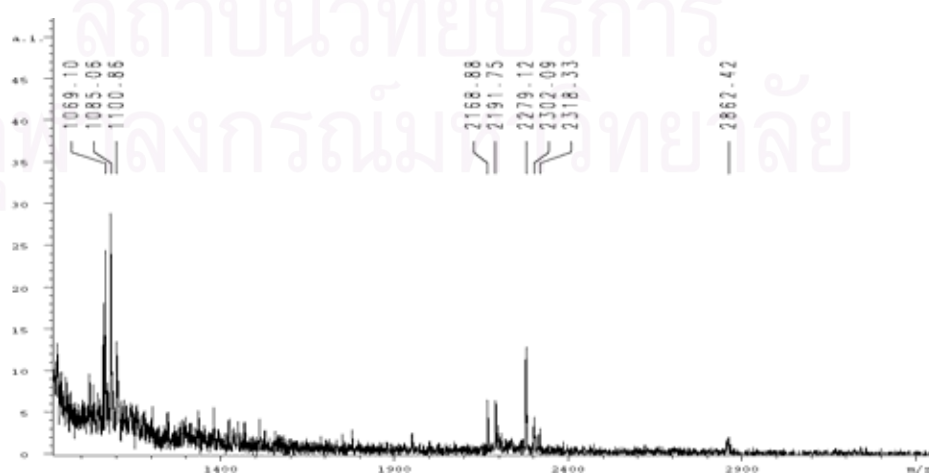**Figure 7D** MALDI-TOF MS spectrum of tryptic fragment of spot B2
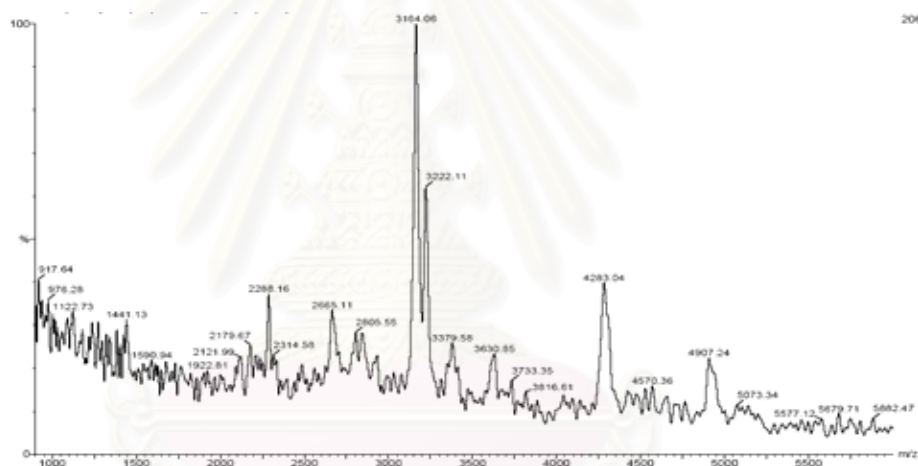
Figure 8D MALDI-TOF MS spectrum of tryptic fragment of spot B15
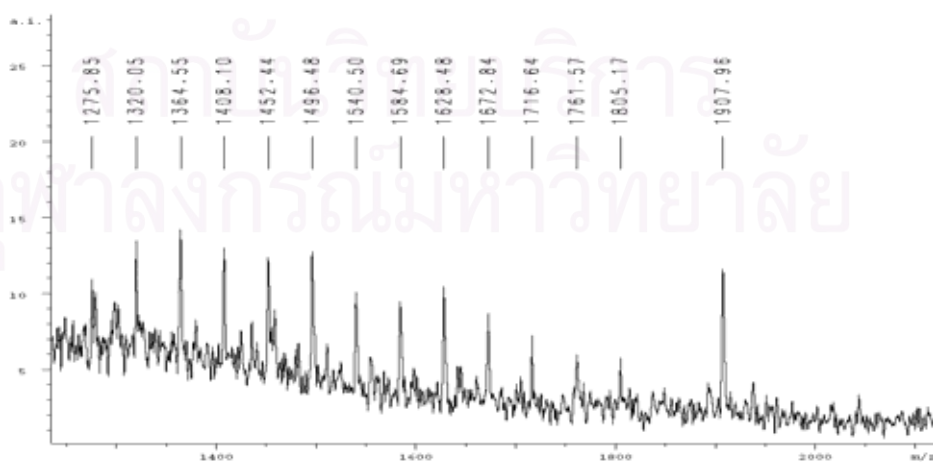
Figure 9D MALDI-TOF MS spectrum of tryptic fragment of spot D1
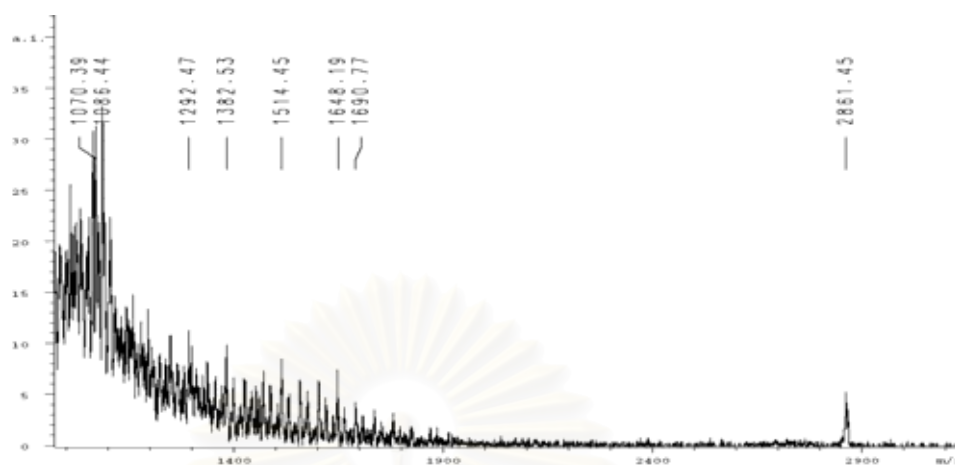
Figure 10D MALDI-TOF MS spectrum of tryptic fragment of spot D5
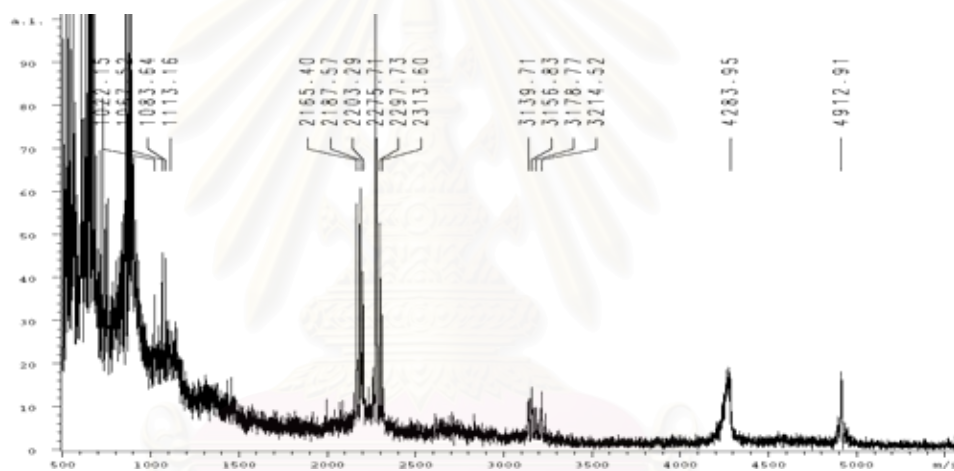
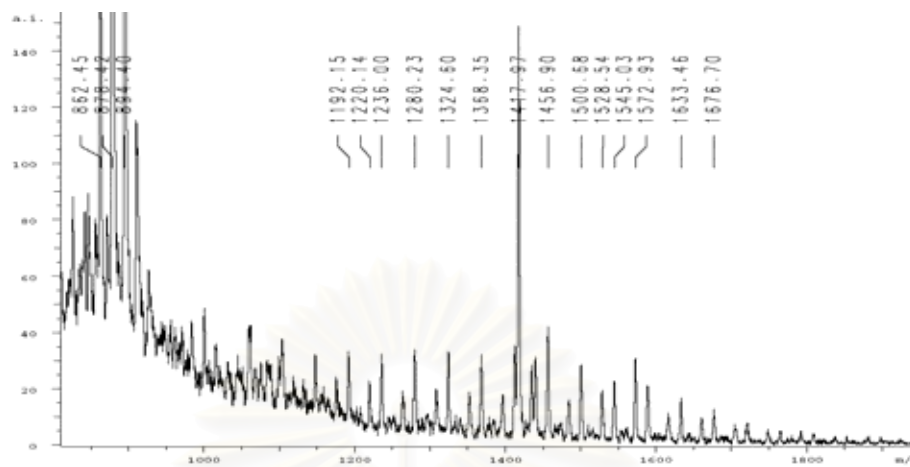**Figure 11D** MALDI-TOF MS spectrum of tryptic fragment of spot D12
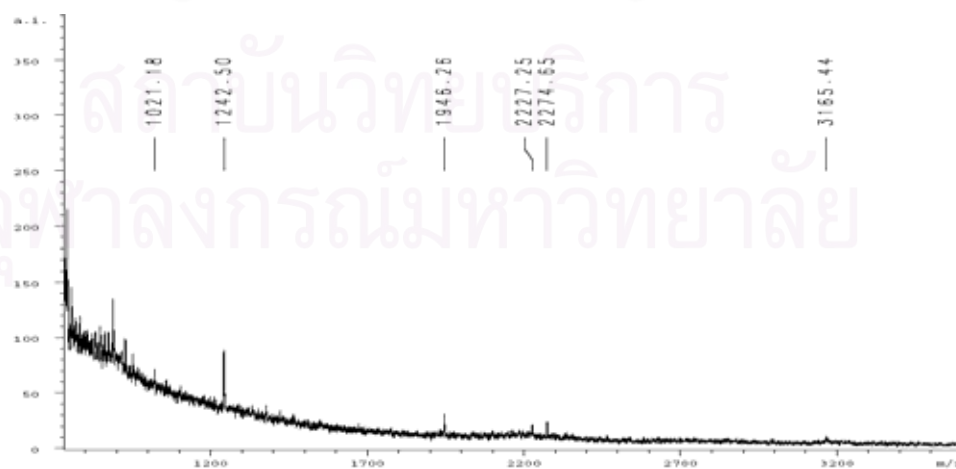
Figure 12D MALDI-TOF MS spectrum of tryptic fragment of spot E1

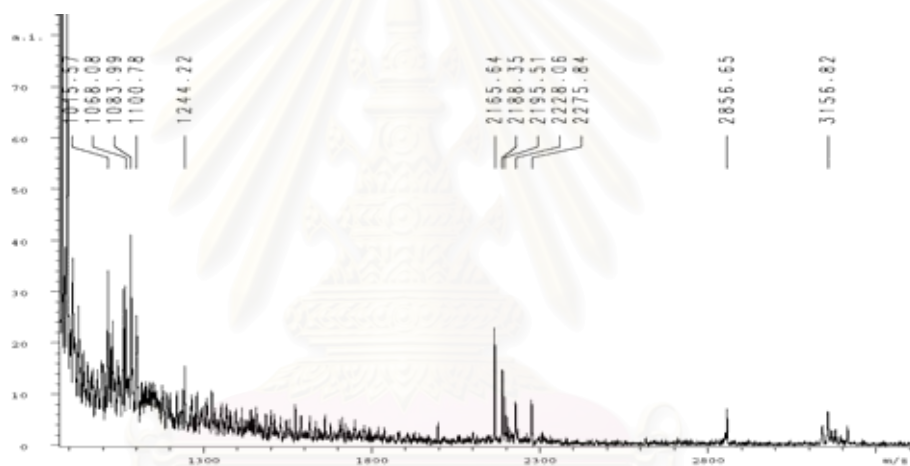**Figure 13D** MALDI-TOF MS spectrum of tryptic fragment of spot E12

Figure 14D MALDI-TOF MS spectrum of tryptic fragment of spot F5

Figure 15D MALDI-TOF MS spectrum of tryptic fragment of spot F15

Figure 16D MALDI-TOF MS spectrum of tryptic fragment of spot F16

Figure 17D MALDI-TOF MS spectrum of tryptic fragment of spot F18

Figure 18D MALDI-TOF MS spectrum of tryptic fragment of spot G31

Figure 19D MALDI-TOF MS spectrum of tryptic fragment of spot I7

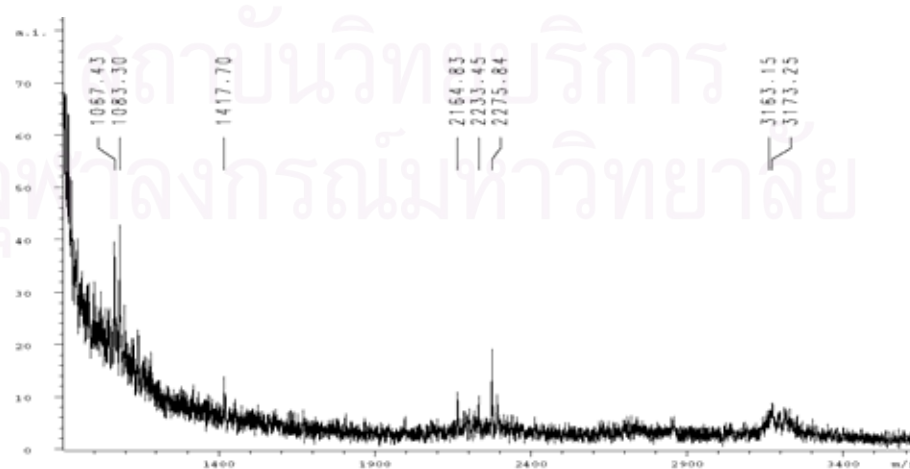**Figure 20D** MALDI-TOF MS spectrum of tryptic fragment of spot K3

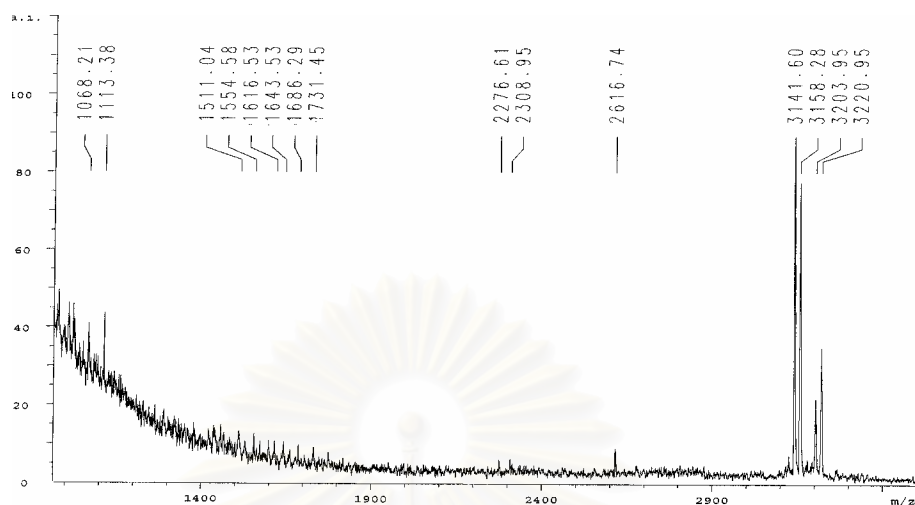**Figure 21D** MALDI-TOF MS spectrum of tryptic fragment of spot K5

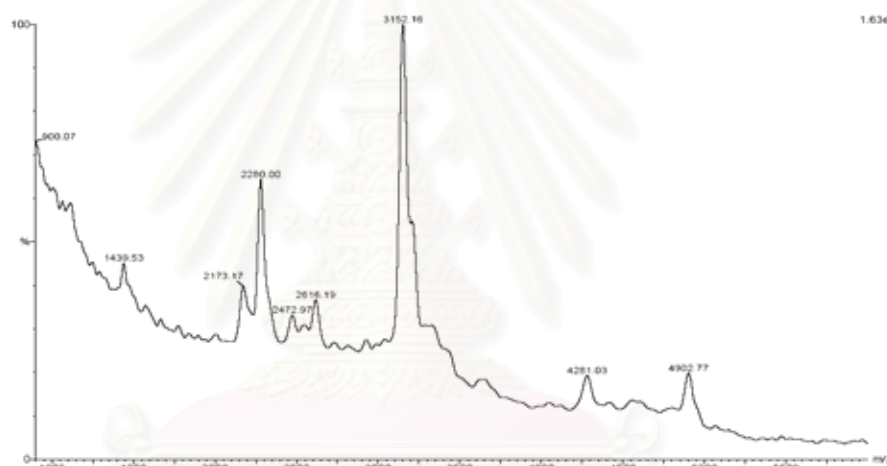**Figure 22D** MALDI-TOF MS spectrum of tryptic fragment of spot L10

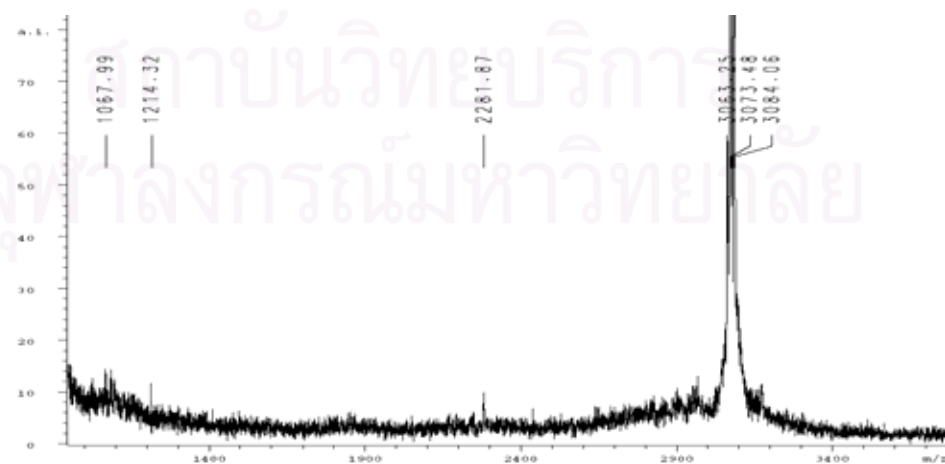**Figure 23D** MALDI-TOF MS spectrum of tryptic fragment of spot L15

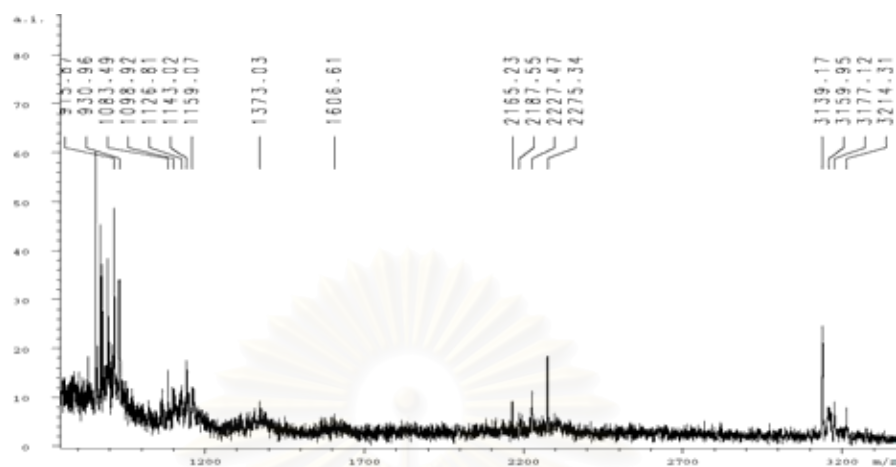**Figure 24D** MALDI-TOF MS spectrum of tryptic fragment of spot M8

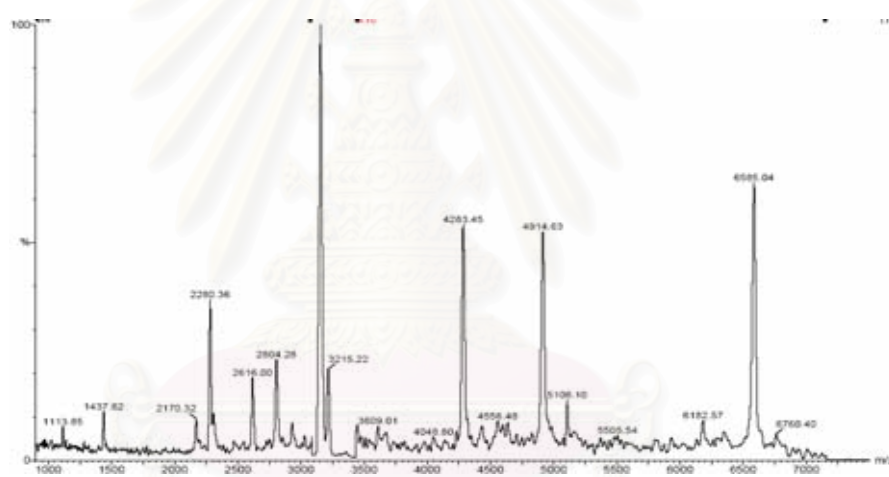**Figure 25D** MALDI-TOF MS spectrum of tryptic fragment of spot M11

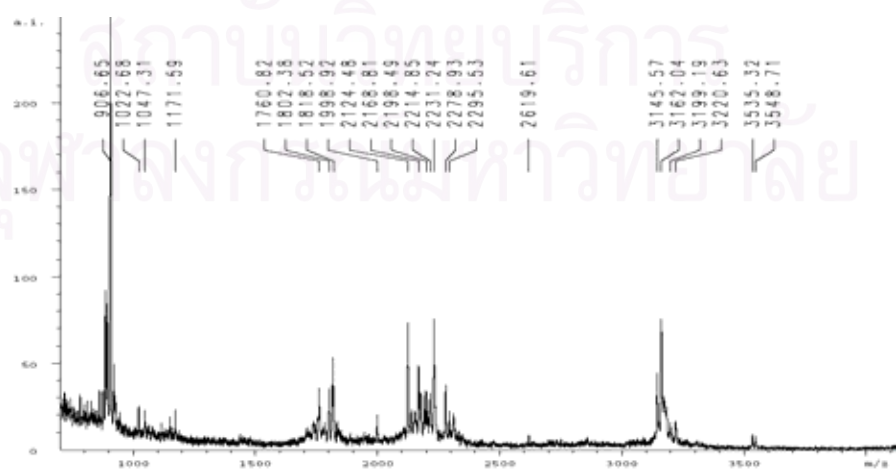**Figure 26D** MALDI-TOF MS spectrum of tryptic fragment of spot N3

Figure 27D MALDI-TOF MS spectrum of tryptic fragment of spot O18

113

Figure 28D MALDI-TOF MS spectrum of tryptic fragment of spot P3
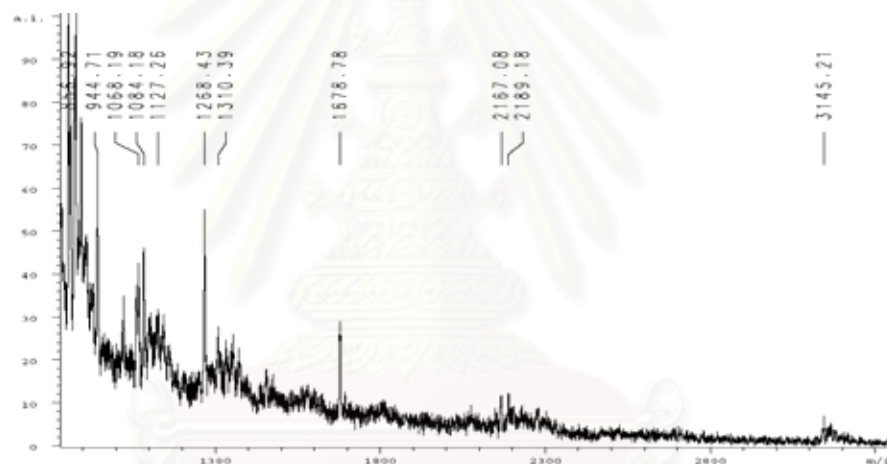
Figure 29D MALDI-TOF MS spectrum of tryptic fragment of spot P4
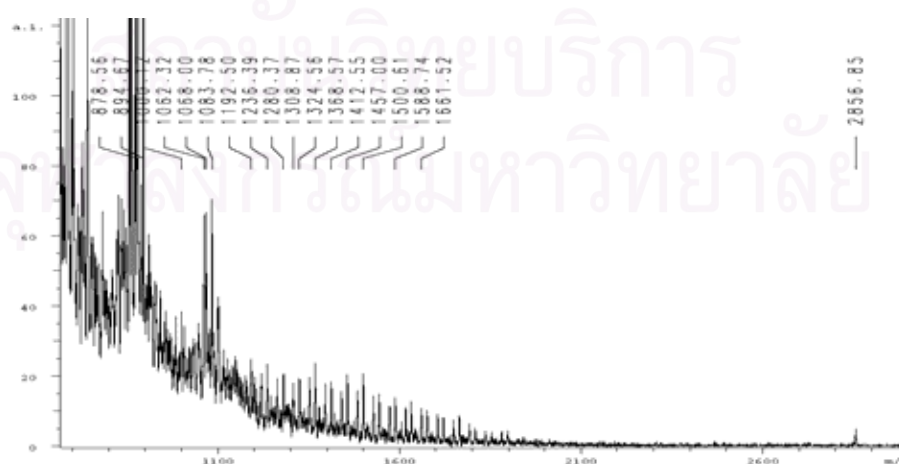
Figure 30D MALDI-TOF MS spectrum of tryptic fragment of spot R5
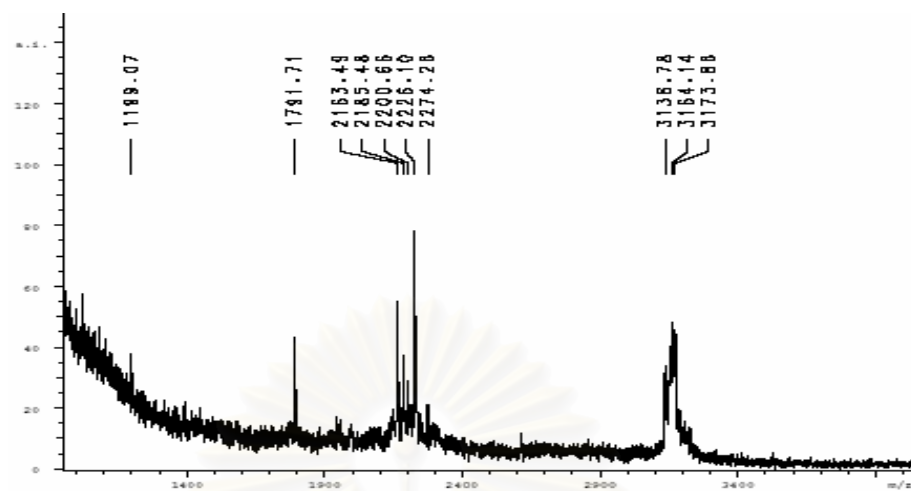
Figure 31D MALDI-TOF MS spectrum of tryptic fragment of spot R7

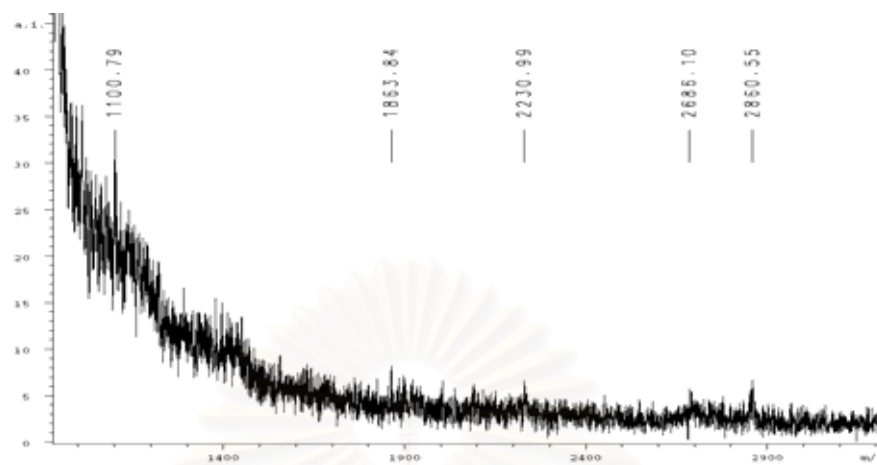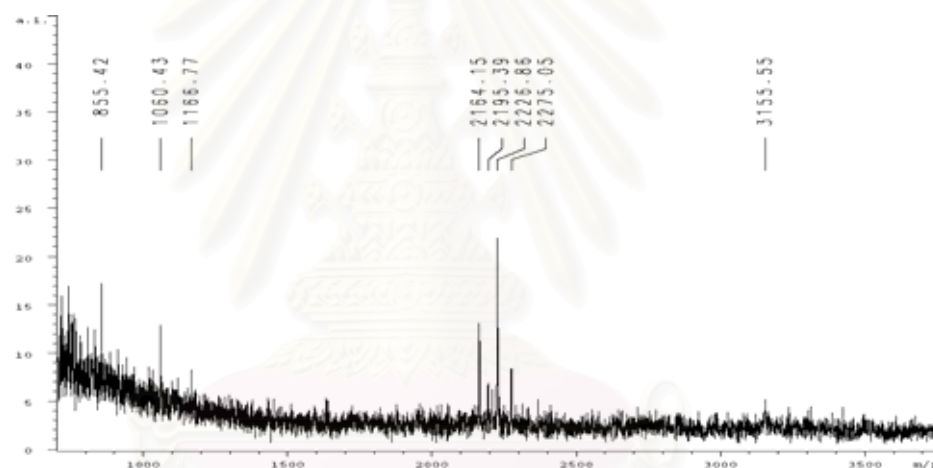**Figure 32** MALDI-TOF MS spectrum of tryptic fragment of spot R17
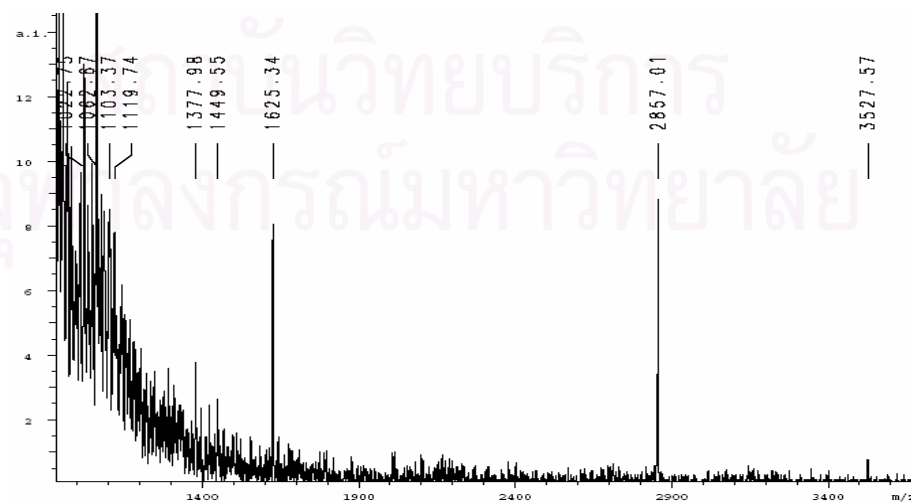
Figure 33D MALDI-TOF MS spectrum of tryptic fragment of spot Z1
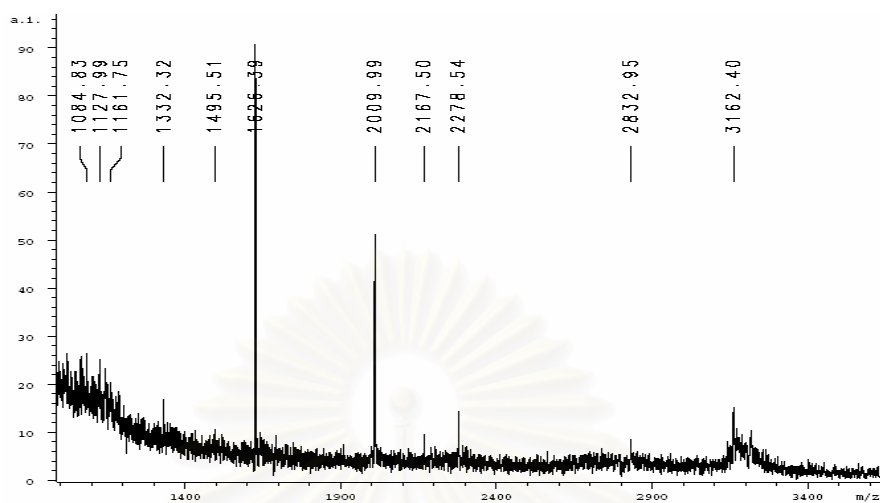
Figure 34D MALDI-TOF MS spectrum of tryptic fragment of spot Z2

Appendix E
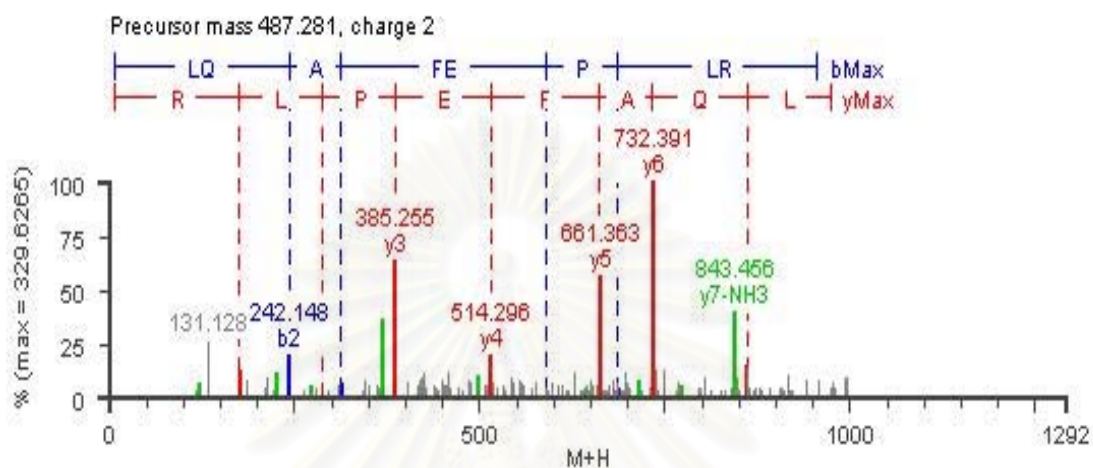
Figure 1E The product ion spectrum of precursor ion m/z of 487.28
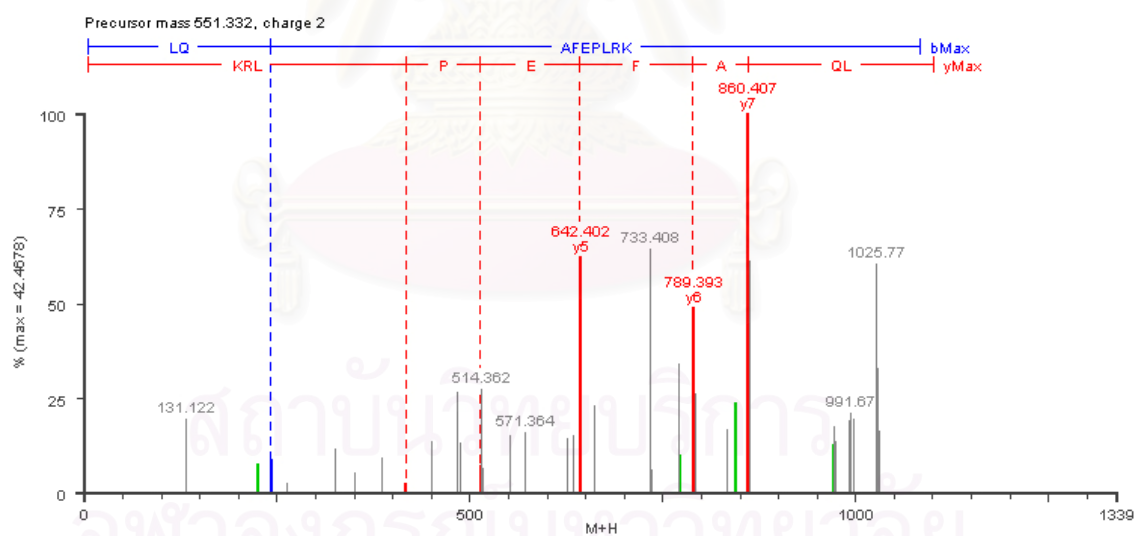
Figure 2E The product ion spectrum of precursor ion m/z of 551.33
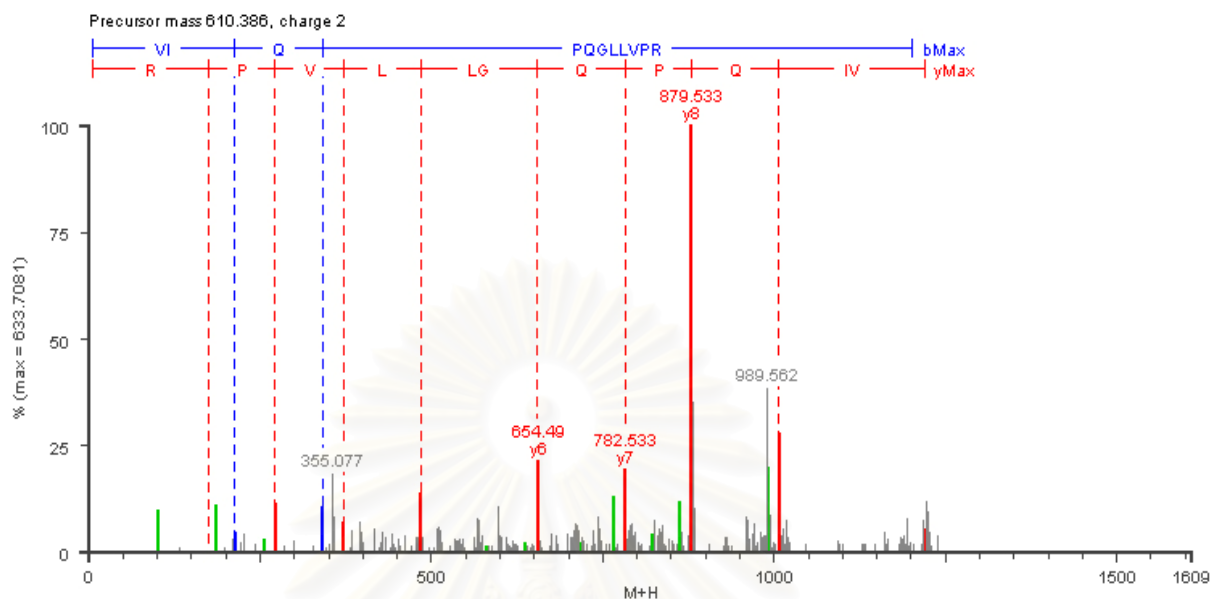
Figure 3E The product ion spectrum of precursor ion m/z of 610.38
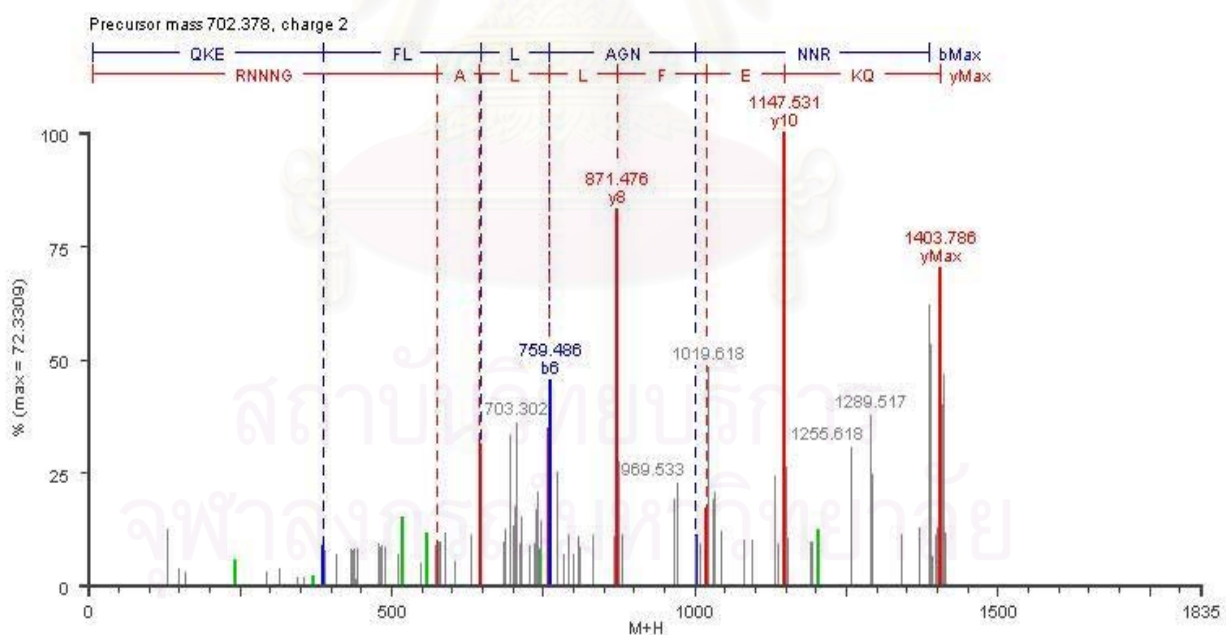
Figure 4E The product ion spectrum of precursor ion m/z of 702.37
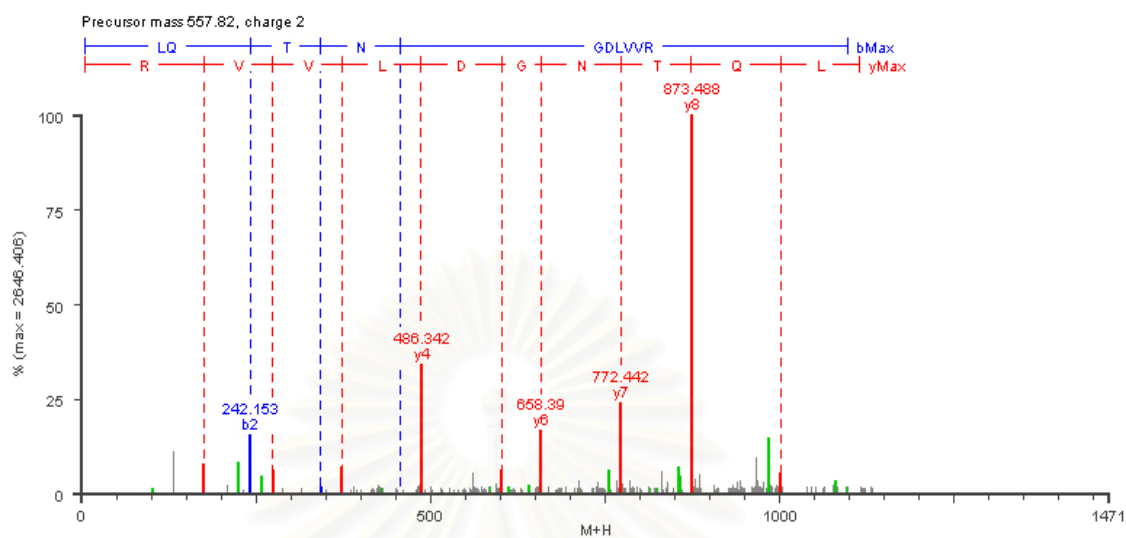
Figure 5E The product ion spectrum of precursor ion m/z of 557.82
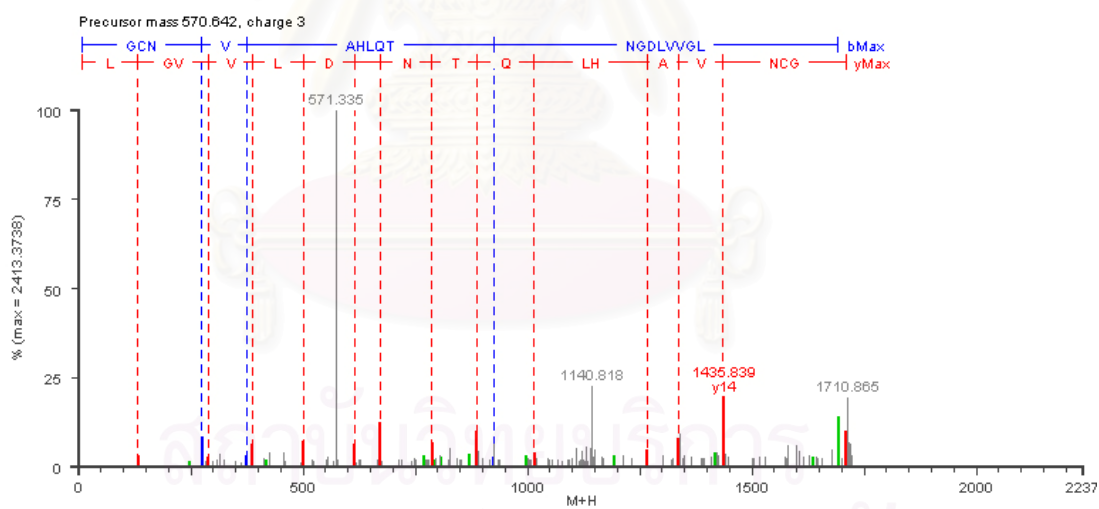
Figure 6E The product ion spectrum of precursor ion m/z of 570.642
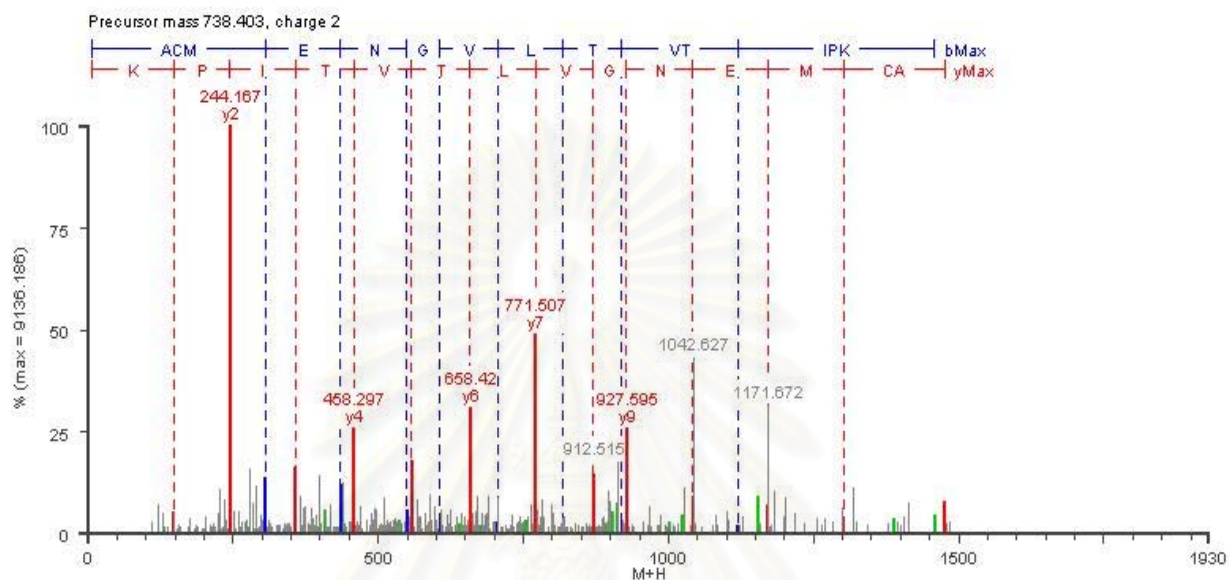
Figure 7E The product ion spectrum of precursor ion m/z of 738.403
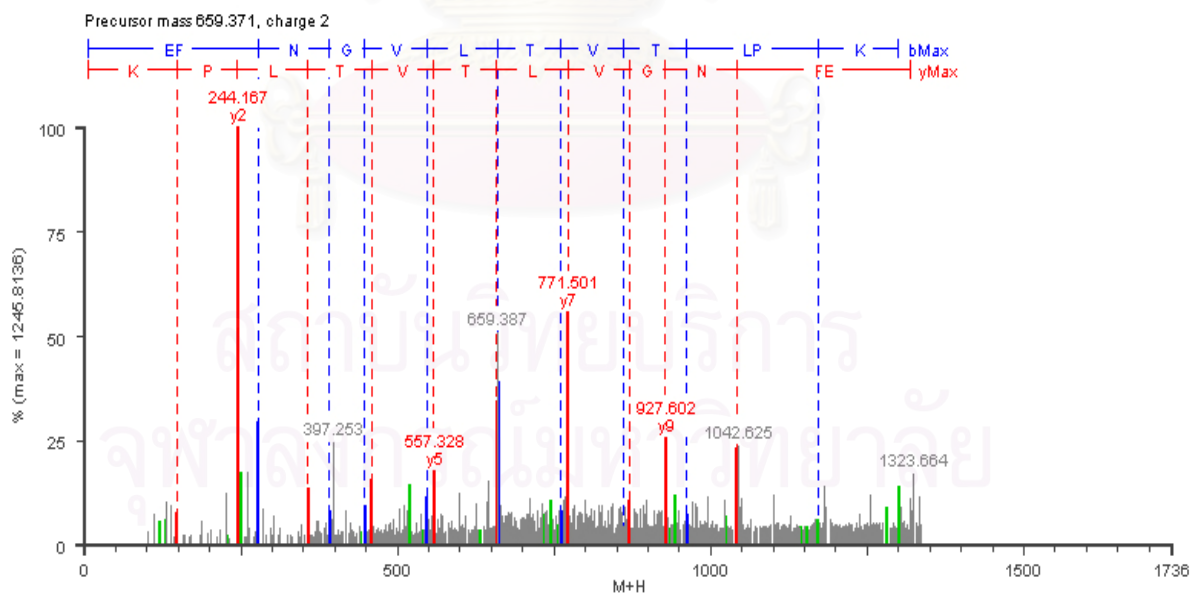
Figure 8E The product ion spectrum of precursor ion m/z of 659.371

Figure 9E The product ion spectrum of precursor ion m/z of 811.899

Figure 10E The product ion spectrum of precursor ion m/z of 820.393
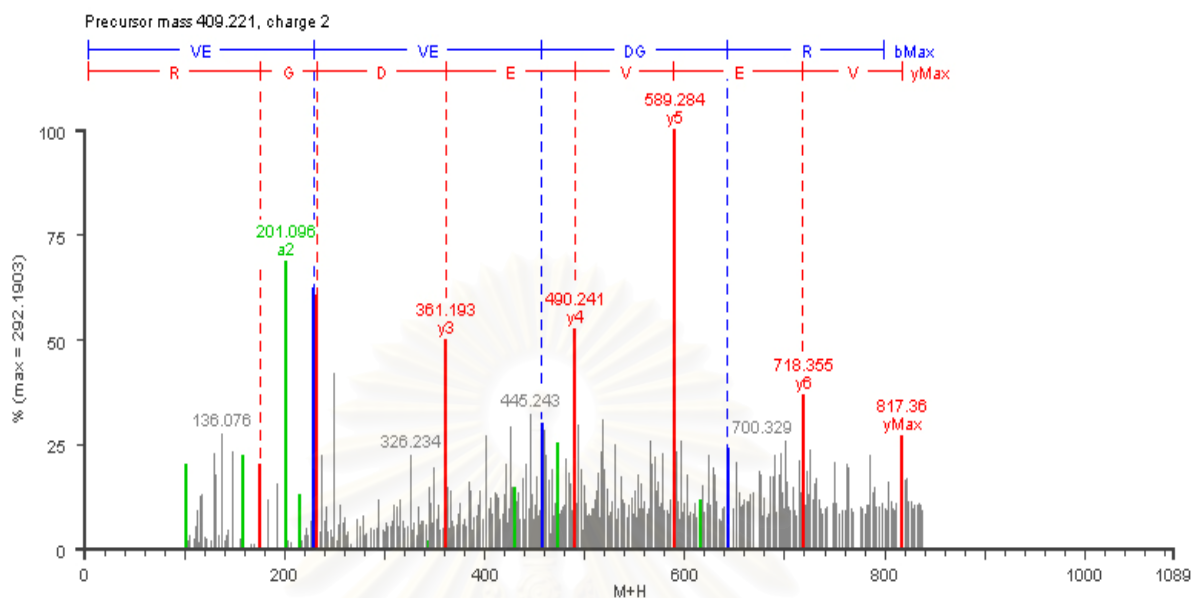
Figure 11E The product ion spectrum of precursor ion m/z of 409.221
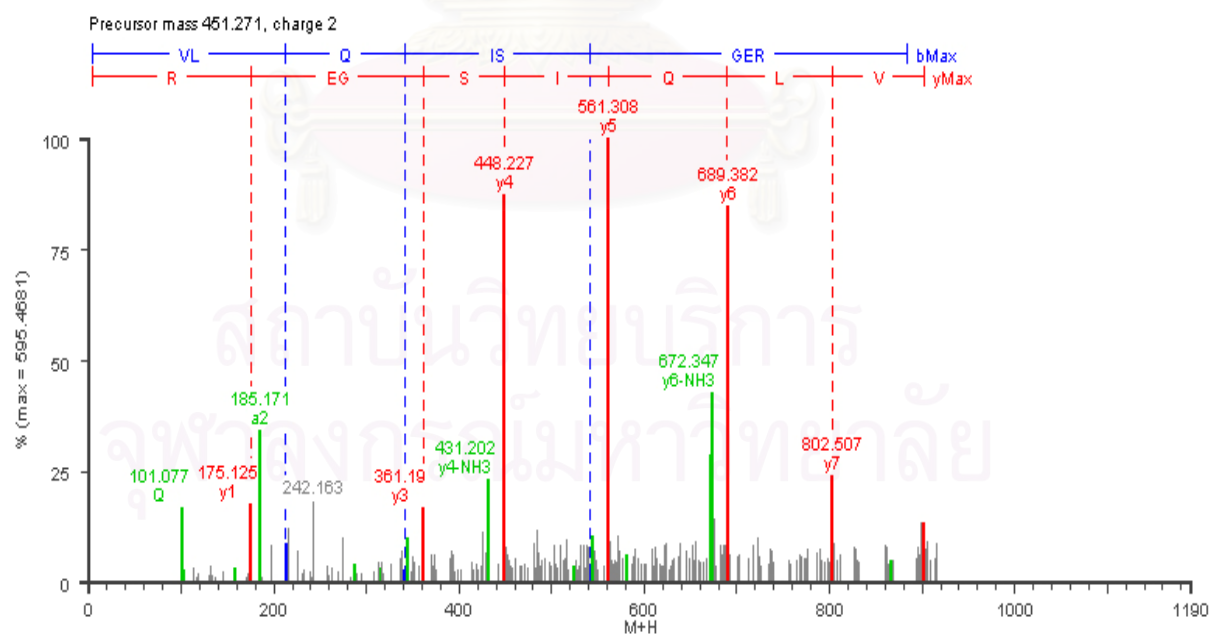
Figure 12E The product ion spectrum of precursor ion m/z of 451.271

Table 1E Amino acid sequences of peptide precursors using BLAST search of
protein spot A8

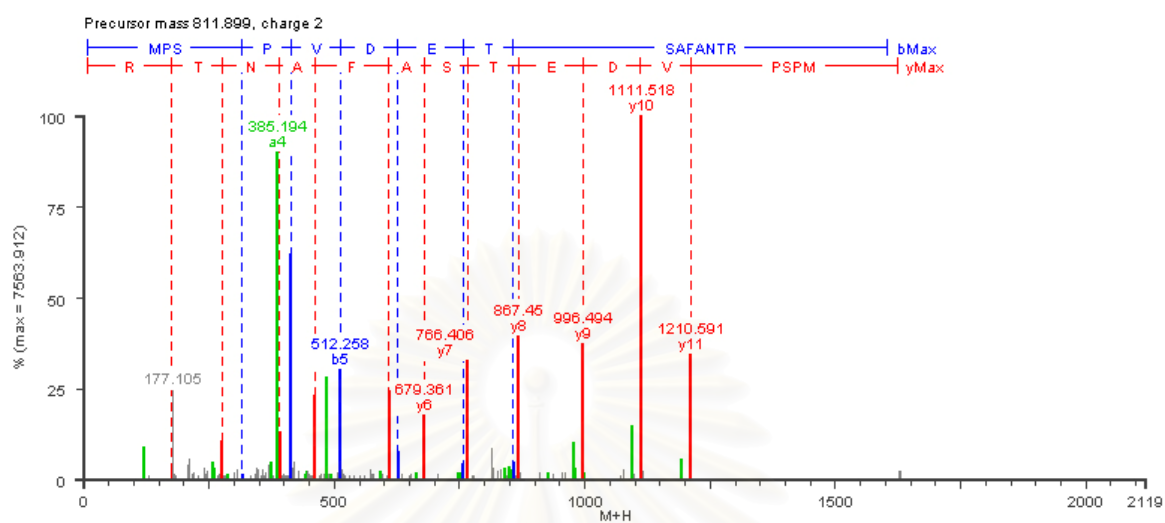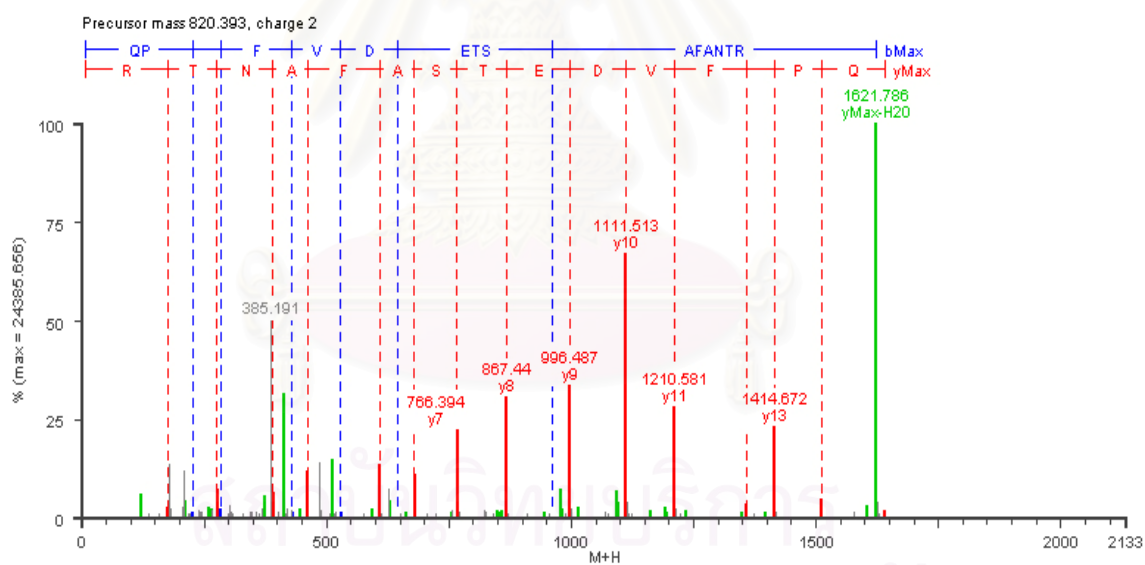| Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
445.271	2	888.526	(-)QPPRYTK(-), (-)QPPGVYTK(-)
453.788	2	905.56	(-)NKPGVYTK(-), (-)KNPGVYTK(-), (-)NKPRYTK(-)
539.306	2	1076.596	(-)APLLSDSMLM(-), (-)APLLSMAMLM(-), (-)APLLSDSSGCK(-), (-)APLLSDSSCGK(-), (-)APLLSSDMLM(-)
568.325	2	1134.634	(-)SSGGLYPDVLK(-), (-)SSLNYPDVLK(-), (-)SSLGGYPDVLK(-), (-)SSNLYPDVLK(-), (-)SSGLGYPDVLK(-)
568.826	2	1135.636	(-)VQNYVSGELK(-), (-)VQNYVSWLK(-), (-)VQNYVSGEKL(-), (-)VQYNVSWLK(-)
616.851	2	1231.686	(-)EENGGLFLQLL(-), (-)EENGGLFLLQL(-), (-)EENGGLLFQLL(-), (-)EENGGLLFLQL(-) (-)EENGGLLLQFL(-)
618.319	2	1234.622	(-)FSSSSGFFAGLE(-), (-)FSSSSGAFFGLE(-), (-)FSSSSGFFALEG(-), (-)FSSSSGFAFLDA(-),(-)FSSSSGAFFALD(-)
639.02	3	1914.037	(-)APLLSDSQRYFPEYTK(-), (-)APLLSDSQRYLMEYTK(-), (-)APLLSDSQRYMLEYTK(-), (-)APLLSDSQRYEMLYTK(-), (-)APLLSDSQRYMLETYK(-)
639.417	2	1276.818	(-)LALVEELPRHT(-), (-)LALDLELLQHL(-), (-)LALDLELVRVH(-), (-)LALDLEVLRVH(-),(-)LALDPLELLAPL(-)
648.359	3	1942.054	(-)APLLFLFEFEERFER(-), (-)APLLEMLDEFFLFSRD(-), (-)APLLDYLEEYLRERY(-), (-)APLLDYELELYRERY(-), (-)APLLDYYREERLELY(-)
651.918	2	1301.82	(-)LSDLDSLLVVTK(-), (-)LSDLDSLLVLSK(-), (-)LSDLALMLVVTK(-), (-)LSDLALMLVLSK(-),(-)LSDLALMLVTVK(-)
681.912	2	1361.808	(-)SQLLSSVTNKTGK(-), (-)QSLLSSVTNKTGK(-), (-)TNLLSSVTNKTGK(-), (-)SKLLSSVTSKKGK(-), (-)KSLLSSVTSKKGK(-)
688.923	2	1375.83	(-)SKLLSVSMLRRS(-), (-)SQLLSVSMLRRS(-), (-)SKLLSSMVLRRS(-), (-)SQLLSVSMLRSR(-), (-)KSLLSSMVLRRS(-)
690.902	2	1379.788	(-)APLLSDYARLYV(-), (-)APLLSDAYRLYV(-), (-)APLLSDYRALYV(-), (-)APLLSDSFRLYV(-),(-)APLLSDSRFLYV(-)
695.928	2	1389.84	(-)AQELSELFLLDL(-), (-)AQELSELFLDLL(-), (-)SQLLSQQYVLLV(-), (-)TNLLSQQYVLLV(-), (-)TNLLSQQYVVLL(-)
700.427	2	1398.838	(-)LQGLVSASTLAVLQ(-), (-)LQGLVSADLLNEK(-), (-)LQGLVSAGGAGTLQK(-), (-)LQGLVSTGTLGPEK(-)
715.724	3	2144.149	(-)APLLSDLYADDGGRNGRLNK(-), (-)APLLSDLYEGDGGRNGRLNK(-), (-)APLLSDLYEGGDGRNGRLNK(-), (-)APLLSDLYEDNGRNGRLNK(-) (-)APLLSDLYENDGRNGRLNK(-)
731.725	3	2192.152	(-)APLLSDEMRAMLAAGYLEGGK(-),(-)APLLSDEMRMLAAAGYLEGGK(-), (-)APLLSDEMRALMAAGYLEGGK(-),(-)APLLSDEMRLMAAAGYLEGGK(-), (-)APLLSDEMRLAMAAGYLEGGK(-)
731.935	2	1461.854	(-)FLTETLTHELMK(-)
739.927	2	1477.838	(-)FLTEGEFTFLFK(-), (-)FLTEGYLTFLFK(-), (-)FLTEGFETFLFK(-), (-)FLTEGYLTFLFQ(-)
745.931	2	1489.846	(-)LQGLVSSVGSGCGAKK(-), (-)LQGLVSSVGSGGCAKK(-), (-)LQGLVSSVGSGKCAGK(-),(-)LQGLVSSVGSGCKAGK(-), (-)LQGLVSSVGSGKCQK(-)
745.961	2	1489.906	(-)KSLKLSEGTLETGK(-), (-)QSLKLSEGTLETGK(-)
746.94	2	1491.864	(-)FLTKGGSFTFLFK(-), (-)FLTKSNFTFLFK(-)

Table 1E Amino acid sequences of peptide precursors using BLAST search of
protein spot A8 (continued)

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
747.424	2	1492.832	(-)LTYEKLLWLLVC(-), (-)LTYEKLLWVLLC(-), (-)LTYEKLLWLVLC(-), (-)LTYEKLLLVWLC(-), (-)LTYEKLLVLWLC(-)
758.453	3	2272.336	(-)SLHLSYLTALTLNNDLMLLK(-), (-)SLVHTYLTALTLNNDLMLLK(-), (-)SLHLSLYTALTLNNDLMLLK(-), (-)SLLHSYLTALTLNNDLMLLK(-), (-)SLLHSLYTALTLNNDLMLLK(-)
759.931	2	1517.846	(-)LQGLVSGQGNSAGALF(-), (-)LQGLVNGTGNSAGALF(-), (-)LVGGGTGNSAGALF(-), (-)LQGLVSANGNSAGALF(-), (-)LQGLVGNTGNSAGALF(-)
997.582	2	1993.148	(-)THNLERNQRGPPVFLLC(-), (-)THNLERNRQGPPVFLLC(-), (-)THNLERAAQKGPPVFLLC(-)
1083.665	2	2165.314	(-)RSVGTLAGLVPGTFAPLLGTAPA(-), (-)SRVGTLAGLVPGTFAPLLGTAPA(-), (-)RSVGTLAGLVPGTFAPLLGTAAP(-), (-)SRVGTLAGLVPGTFAPLLGTAAP(-)

Table 2E Amino acid sequence of peptide precursors using BLAST search of
protein spot A10

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
734.316	3	2199.925	(-)LGEDNSTSWGETGSGCDLNMK(-), (-)LGEDNSTSWGEASSGCDLNMK(-), (-)LGEDNSTSWGEASSGCDLNMK(-), (-)LGEDNSTSWWTGSGCDLNMK(-), (-)LGEDNSTSWMGVGSGCDLNMK(-), (-)LGEDNSTSWWASSGCDLNMK(-)
675.66	3	2023.957	(-)HVPCFNSNTLNNDVSHVK(-), (-)HVPSYNSNTLNNDVSHVK(-)
669.812	2	1337.608	(-)APSTYGGGLSDASR(-), (-)APSTYGGGLGESSR(-) (-)APSTYGGGLSSGER(-), (-)APSTYGVGASSGER(-)
831.894	2	1661.772	(-)EGGSSTDEVALNVSAAR(-), (-)WGSSTDEVALNDAAAR(-), (-)WGSSTDEVALNGEAAR(-) , (-)WGSSTDEVALNVSAAR(-), (-)WGSSTDEVALNVSAAR(-), (-)WGSSTDEVALNWAAR(-)
742.357	3	2224.048	(-)EAVHPSYNSNTLNQNVFYK(-), (-)EAVHPSYNSNTLNPMLFYK(-), (-)EAVHPSYNSNTLNNRAFYK(-), (-)EAVHPSYNSNTLNKCKEYK(-), (-)EAVHPSYNSNTLNKCKEYK(-), (-)EAVHPSYNSNTLNPPFFYK(-)
477.73	2	953.444	(-)THNAELGVN(-), (-)THNAELVGN(-), (-)THNAELNR(-), (-)THNAELNR(-), (-)THNAELNGV(-), (-)THNAELNVG(-)
679.339	2	1356.662	(-)HSATNSMAHVRF(-), (-)HSATNSMAHVFR(-), (-)HSATNSMAHFVR(-), (-)SHATNSMAHVFR(-), (-)HSATNSMAHVFR(-) , (-)SHATNSMAHVFR(-)
731.65	3	2191.927	(-)APLCHNDFCYVMDSDSTFK(-), (-)APLCHNDFYCVMDSDSTFK(-), (-)APLCHNDFMSDYDSDSTFK(-), (-)APLCHNDFYCVMDCWTFK(-), (-)APLCHNDFMDSYDSDSTFK(-), (-)APLCHNDFYCVMDCWTFK(-), (-)APLCHNDFMDSYDSDSTFK(-)

Table 3E Amino acid sequences of peptide precursors using BLAST search of

protein spot O3

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
427.775	2	853.534	(-)SVTLPLPK(-), (-)SVTLPLPK(-), (-)VSTLPLPK(-), (-)VSTLPLPK(-), (-)LSSLPLPK(-), (-)LSSLPLPK(-), (-)SLSLPLPK(-), (-)SVTLPLPK(-), (-)SVTLPLPK(-), (-)VSTLPLPK(-), (-)VSTLPLPK(-) ,(-)LSSLPLPK(-), (-)LSSLPLPK(-), (-)TWLPLPK(-), (-)WTLPLPK(-)
436.252	2	870.488	(-)NPVGESLR(-), (-)PVNGESLR(-), (-)VPNGESLR(-), (-)PNVGESLR(-)
445.261	2	888.506	(-)LNTGESLR(-), (-)NLTGESLR(-), (-)LNTGELSR(-)
452.798	2	903.58	(K)IYIVLRK(R), (K)IYIVLRK(R)
453.776	2	905.536	(K)NKPGVYTK(V), (K)NKPGVYTK(V), (R)NKPGVYTK(V), (K)NKPGVYTK(V), (K)NKPGVYTK(V)
460.308	2	918.6	(-)YLLVRQK(-), (-)YLLVQKR(-), (-)YLLVKQR(-), (-)YLLARKR(-), (-)YLLAKRR(-)
468.239	3	1401.694	(-)SDELGNLLYPPGK(-), (-)SDELGNLLYPGPK(-), (-)SDELGNLLYGPPK(-), (-)SDEGLNLLYPPGK(-), (-)SDEGLNLLYGPPK(-)
479.276	2	956.536	(-)NNVLLSAAR(-), (-)NNVLLNDR(-), (-)NNVLLDNR(-), (-)GGNVLLNDR(-), (-)NNVLLGGDR(-)
489.58	3	1465.717	(-)SDESFNLLAWER(-), (-)SDEAYNLLAWER(-) , (-)SDESFNLLAEWR(-), (-)SDEAYNLLADLDK(-),(-)SDEAYNLLAEWR(-)
495.257	2	988.498	(-)VEGTNGLSGR(-), (-)VEGTNGGSLR(-), (-)VGETNGVSAR(-), (-)VEGTNGVSAR(-)
507.793	2	1013.57	(-)GNNVLAATVR(-), (-)GNNVLAAEAR(-, (-)NGNVLAAEAR(-), (-)GNNVLAAAER(-), (-)GNNVLNEVR(-)
557.968	3	1670.881	(-)SDEGLNLLYKLGHGR(-), (-)SDEGLNLLYKLGHGR(-), (-)SDEGLNLLYKAVHGR(-), (-)SDEGLNLLYKVAHGR(-), (-)SDELGNLLYKGLHGR(-)
568.306	2	1134.596	(K)SSGTSYPDVLK(C)[a]
568.622	3	1702.843	(-)NNKGPGESSETNGLATK(-), (-)NNKGPWSSETNGLATK(-), (-)NNKPGWSSETNGLATK(-), (-)NNKPGVSSSETNGLATK(-), (-)NNKGPWSSETNGLTAK(-)
568.622	3	1716.862	(-)VVVDDDNGSLTNGLTAK(-), (-)LALDDDNGSLTNGLTAK(-) (-)LALDDDNLSGTNGLTAK(-), (-)VELADDNLSGTNGLTAK(-) (-)VELAGNDNNSTNGLGSR(-)
573.295	2	1152.584	(K)SSGTSYPDVLK(C)[a]
573.295	3	1791.022	(-)THAVVLYGPPLWTLPK(-), (-)PPDVVLYGPPLWTLPK(-), (-)RHDVLYGPPLWTLPK(-)
598.015	2	1199.64	(-)TNTPLLYGGHK(-), (-)TNTPLLYNHK(-), (-)NTTPLLYGGHK(-), (-)NTTPLLYNHK(-)
600.828	2	1231.648	(-)SGGGGGGGKERSAR(-), (-)SGNGGGGKERSAR(-), (-)SGGGNGGKERSAR(-)
616.832	3	1848.967	(-)CVGEQVLKYEDKVALR(-), (-)CVGEQVLKYEKDVALR(-), (-)MAGEQVLKYEDKVALR(-), (-)MAGEQVLKYEKDVALR(-), (-)AMGEQVLKYEKDVALR(-)
617.33	2	1263.658	(-)ESFNLLYNKH(-), (-)ESFNLLYGGKH(-), (-)ESFNLLYNHK(-), (-)ESFNLLYGGHK(-), (-)SEFNLLYNHK(-)
632.837	2	1356.772	(K)LNDLEDALQQAK(E)[a]
632.837	2	1378.694	(-)DESFNLLYGGKH(-), (-)DESFNLLYGGHK(-), (-)DESFNLLYNHK(-), (-)DESFNLLYPPGK(-), (-)DESFNLLYGPPK(-)
690.355	2	1382.762	(K)SLNNKFASFIDK(V)
690.355	2	1401.66	(-)SDELGNLLYGGHK(-),(-)SDELGNLLYNHK(-), (-)SDELGNLLGYGHK(-), (-)DSELGNLLYGGHK(-), (-)SPESTNLLYPPGK(-)

Table 3E Amino acid sequences of peptide precursors using BLAST search of
protein spot O3 (continued)

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
701.838	2	1428.888	(-)GPKDKYLLVLRK(-),(-)GPENKYLLVLQR(-), (-)GPKNEYLLVLQR(-), (-)GPQSRYLLVLQR(-), (-)GPQRSYLLVLQR(-)
715.452	2	1447.724	(-)DTPSFNLLVNSNK(-), (-)DTPSFNLLVGGSNK(-), (-)ESPSFNLLVNSNK(-), (-)DTNSEDLLYNHK(-), (-)DTNESDLLYNHK(-)
724.87	2	1449.702	(-)AMESMNLLYNHK(-), (-)CVESMNLLYNHK(-), (-)MAESMNLLYNHK(-), (-)SDESMNLLYNHK(-), (-)VCESMNLLYNHK(-)
725.859	2	1458.752	(-)CVNTTPLLYGGKPP(-), (-)CVTNTPLLYGGKPP(-), (-)AMTNTPLLYGGKPP(-), (-)CVNTTPLLYGGPPK(-), (-)CVTNTPLLYGGPPK(-)
730.384	2	1461.736	(-)VNNSFGGLLYNHK(-)[a]
730.384	2	1465.724	(-)SDESFNLLPTESK(-), (-)SDESFNLLPTTDK(-), (-)SDESFNLLPTDTK(-), (-)SDESFNLLPTSEK(-), (-)SDESFNLLVNNSK(-)
733.87	2	1474.81	(K)WELLQQVDTSTR(T)[a]
733.87	2	1479.742	(-)ETDSFLNLYGGHK(-), (-)ETDSFNLLYGGHK(-), (-)TEDSFNLLYGGHK(-) (-)ETDSFNLLYNHK(-), (-)ESESFNLLYGGHK(-)
740.879	2	1485.914	(-)GPSVPHNGLLVLQR(-), (-)LGLGPHNGLLVLQR(-), (-)VSPGPHNGLLVLQR(-), (-)GPSVKNYLLVLQR(-), (-)LNLPHNGLLVLQR(-)
743.965	2	1489.82	(K)LQGIVSWGSGCAQK(N)[a]
743.965	2	1506.746	(-)MAGESMNLLPSETK(-), (-)MQESMNLLPSETK(-), (-)MGAESMNLLPSETK(-), (-)MQESMNLLPDTTK(-), (-)QMESMNLLPSETK(-)
754.381	3	2272.276	(K)SIVHPSYNSNTLNNDIMLIK(L)[a]
754.381	2	1522.746	(-)SDESFLNLYGGKGH(-), (-)SDESFLNLYGGGHK(-), (-)SDESFLNLYGGHGK(-), (-)MAESFLNLYGGPPK(-), (-)CVESFLNLYGGPPK(-)
762.381	2	1603.912	(-)YNSNTLKVASALPAR(-), (-)YNSNTLAGVASALPAR(-), (-)YNSNTLQLGSALPAR(-), (-)YNSNTLQVASALPAR(-)
802.964	2	1778.868	(-)NNKAMGELMSTNRSAR(-), (-)NNKAMLSVMSTNRSAR(-), (-)NNKAMSVLMSTNRSAR(-), (-)NNKAMWLMSTNRSAR(-), (-)NNKAMVSLMSTNRSAR(-)
905.029	2	1808.042	(-)NRLVHRKRQLVVQY(-), (-)LRNVHRKRQLVVQY(-), (-)VQRVHKRRQLVVQY(-), (-)NRLVHKRRQLVVQY(-), (-)RNLVHKRRQLVVQY(-)
905.03	2	1808.044	(-)DPPVVLYGNPLVSTLPK(-), (-)DPPVVLYGNPLWTLPK(-), (-)DPPVVLYGNPLSVTLPK(-), (-)PPDVVLYGNPLVSTLPK(-), (-)PPDVVLYGNPLWTLPK(-)
912.546	2	1823.076	(-)ATHVVLYSVPLVSTLPK(-), (-)THAVVLYSVPLVSTLPK(-), (-)TAHVVLYSVPLVSTLPK(-), (-)HTAVVLYSVPLVSTLPK(-), (-)PDPVVLYSVPLVSTLPK(-)
919.037	2	1836.058	(-)THAVVLYGNVPKGSGLPK(-), (-)THAVVLYGNVPKSGGLPK(-), (-)THAVVLYGNVLAGAGGLPK(-), (-)THAVVLYGNVLAGAARPK(-), (-)THAVVLYGNVLQAARPK(-)

[a]Databank Serach (40)

Table 4E Amino acid sequences of peptide precursors using BLAST search of
protein spot O5

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
436.913	3	1307.716	(-)LEYMLELAAKK(-), (-)LEYMLELAKAK(-), (-)LEYMLELKAAK(-), (-)LEYLMELRNK(-)
445.268	2	888.52	(-)LNTGESLR(-)
451.235	1	450.227	(-)VTEC(-)
452.812	2	903.608	(K)IYIVLRK(R)[a]
479.291	2	956.566	(-)NNVLLDNR(-), (-)NNVLLTQR(-)
489.589	3	1465.744	(-)SDEFSKVLAVCLR(-), (-)SDEFSKVLALCVR(-), (-)SDESARFLAVCLR(-), (-)SDESARFLALVCR(-), (-)SDESARFLALAMR(-)
533.297	2	1064.578	(R)AQYEDVANR(S)[b]
533.799	2	1065.582	(K)YEDEINKR(T)[b]
545.811	2	1089.606	(K)VTMQNLNDR(L)[b], (R)VTMQNLNDR(L)b
559.317	2	1116.618	(-)PSAMGLYGGHK(-), (-)MSPAGLYGGHK(-), (-)SMPAGLYGGHK(-), (-)TERGLYGGHK(-), (-)AKMLLYGGHK(-)
568.319	2	1134.622	(K)SSGTSYPDVLK(C)[b]
573.839	2	1145.662	(-)QLLWESGSVK(-), (-)QLLWESKDK(-), (-)QLLWESGVSK(-)
577.317	2	1152.618	(K)SSGTSYPDVLK(C)[a]
587.863	2	1173.71	(K)TLNNDIMLIK(L)[a]
590.342	2	1178.668	(K)YEELQITAGR(H)[a]
615.314	2	1228.612	(-)APLDQGLNSSLD(-), (-)APLDQGLNSAED(-), (-)APLDQGLNSADE(-)
618.311	2	1234.606	(R)FSSSSGYGGGSSR(V)[a]
623.843	2	1245.67	(-)SEENLLNGGVSK(-), (-)SEENLLGGGRSK(-), (-)SEENLLNGRSK(-), (-)ESENLLNYHK(-), (-)SEENLLNYHK(-)
633.361	2	1264.706	(R)TAAENEFVTLK(K)[b]
639.014	3	1914.019	(-)APLLSDACASAQRGALEGGK(-), (-)APLLSDACASARQGALEGGK(-), (-)APLLSDACASAAGRQLEGGK(-), (-)APLLSDSSERDLQLEGGK(-), (-)APLLSDSSEVRQELEGGK(-)
639.872	2	1277.728	(K)YEQLCRIILK(D)[a]
651.905	2	1301.794	(R)SLDLDSIIAEVK(A)[a]
651.905	2	1301.794	(R)SLDLDSIIAEVK(A)[a]
654.878	2	1307.74	(K)NKYEDEVQQR(T)[b]
657.86	2	1313.704	(-)LVGGYDHGNGAVR(-)
670.886	2	1339.756	(K)SKAEAESLYQSK(Y)[a]
679.393	2	1356.77	(K)LNDLEDALQQAK(E)[a]
683.374	2	1364.732	(R)SQYEQLAEKNR(R)[a]
691.373	2	1380.73	(R)ALEESNYELEGK(I)[a]
692.394	2	1382.772	(K)TVNNKFASFIDK(V)[b]
695.894	2	1389.772	(K)QSLEASLAETEGR(Y)[a]
701.888	2	1401.76	(-)SDELGNLLYGGHK(-), (-)DSELGNLLYGGHK(-), (-)EDSLGNLLYGGHK(-), (-)DESLGNLLYGGHK(-), (-)SDELGNLLYNHK(-)
716.88	2	1431.744	(-)PSESMNLLTAGVLT(-), (-)PSESMNLLTAVGLT(-), (-)PSESMNLLTARLT(-), (-)PSESMNLLTNLLT(-), (-)GQQSMNLLTARLT(-)
724.882	2	1447.748	(-)ESPSFNLLPVSMK(-), (-)ESPSFNLLVSLDK(-), (-)ESPSFNLVPLAFK(-), (-)ESPSFNLVPLFAK(-), (-)EPSSFNLVPLAFK(-)
725.886	2	1449.756	(-)MAESMNLKSLRNG(-)
732.893	2	1463.77	(-)GSSNSLASSLYNHK(-), (-)SGSNSLASSLYNHK(-),
733.888	2	1465.76	(-)SDESFNLLYGPPK(-), (-)SDESFNLLYPGPK(-), (-)SSGNSLASSLYNHK(-), (-)SDESFNLLYGGHK(-), (-)SDESFNLLYNHK(-), (-)SDESFNLLYPPGKI(-)

Table 4E Amino acid sequences of peptide precursors using BLAST search of

protein spot O5 (continued)

Submitted Mass	Submitted Charge	Experimental Mass	Sequence (De Novo Sequencing)
737.054	3	2208.139	(-)APLLVCMLEWQMAGYLEGGK(-), (-)APLLSDMNKWQMAGYLEGGK(-),
			(-)APLLSDMNGTRQMAGYLEGGK(-)
738.426	2	1474.836	(K)WELLQQVDTSTR(T)[a]
742.439	3	2224.294	(-)TVVHPHLNSNTLNNPLTLLK(-), (-)SLVHPHLNSNTLNNPLTLLK(-),
			(-)LSVHPHLNSNTLNNPLTLLK(-), (-)SLVHPHLNSNTLNNPLLTLK(-),
			(-)VTVHPHLNSNTLNNPLTLLK(-)
745.925	2	1489.834	(K)LQGIVSWGSGCAQK(N)[a]
747.425	2	1492.834	(R)SQYEQLAEQNRK(D)[a]
748.916	2	1495.816	(-)LVGANVPGGGNAGAGGTK(-), (-)LVGANVPNGNAGAGGTK(-),
			(-)LVGANVPGGNAGGGKTK(-), (-)LVGANVPNNAGGGKTK(-),
			(-)LVGANVPGGNAGNKTK(-)
752.434	3	2254.279	(K)SIVHPSYNSNTLNNDIMLIK(L)[b]
758.432	3	2272.273	(K)SIVHPSYNSNTLNNDIMLIK(L)[a]
769.077	3	2304.208	(-)DAVVSKYFLVWAHFYGFQK(-), (-)DAVVSKYFLVWAHFYGQFK(-)
			(-)DAVVSKYFLVWAHVMHFQK(-), (-)DAVVSKYFLVWAHTEHFQK(-)
			(-)DAVVSKYFVLWAHFYGFQK(-)
803.671	4	3210.653	(K)APILSDSSCKSAYPGQITSNMFCAGYLEGGK(D)[a]
826.962	2	1651.908	(-)YNSNTLPPVYMLLK(-), (-)YNSNTLPMPVYLLK(-), (-)YNSNTLPMDLMLLK(-)
854.445	2	1706.874	(K)GSLGGGFSSGGFSGGSFSR(G)[b]
858.987	2	1715.958	(K)QISNLQQSISDAEQR(G)[a]
887.97	2	1773.924	(K)TPVTEAIATLKETLK(G)[a]
887.97	2	1773.924	(K)TPVTEAIATLKETLK(G)[a]
905.01	2	1808.004	(-)DPPVVLYGNPLVSTLPK(-), (-)HATVVLYGNPLVSTLPK(-), (-)AHTVVLYGNPLVSTLPK(-), (-)DPPVVLYGNPLWTLPK(-),
			(-)HTAVVLYGNPLVSTLPK(-)
958.02	2	1914.024	(-)APLLSGNADFASSSAHKGGK(-), (-)APLLSDAGNFASSASHKGGK(-)
			(-)APLLSGNADFGTSSAHKGGK(-), (-)APLLSGNADFASSASHKGGK(-)
			(-)APLLSGNADFASSSAHKNK(-)
958.02	2	1914.024	(-)APLLSGNADFASSSAHKGGK(-), (-)APLLSDAGNFASSASHKGGK(-)
			(-)APLLSGNADFGTSSAHKGGK(-), (-)APLLSGNADFASSASHKGGK(-)
			(-)APLLSGNADFASSSAHKNK(-)
964.182	3	2889.523	(K)APLITTLILITLLSMGGLPPLSGFMPK(W)[a]
1071.224	3	3210.649	(-)SAYPGKLTSPRPLLAMKPALPTTPGYLENK(-) (-)SAYPGKLTSPRPLLAMKPALPWLGYLENK(-), (-)SAYPGKLTSPRPLLAMKPALPGELGYLENK(-), (-)SAYPGKLTSPRPLLSDKPALPGELGYLENK(-)
1568.831	2	3135.646	(-)QQRESLLCLTRLLFLRLYLLNSVNCK(-), (-)QQRESLLLCTRLLFLRLYLLNSVNCK(-), (-)QQRESLLCLTRLLFLRLYLLNSNVCK(-), (-)QQRESLLLCLLLKYRLLYLLNSVCNK(-), (-)QQRESLLLCLKLLYRLLYLLNSVCNK(-)

[a]DataBank Search (40)
[b]AutoMed Query

VITA

Mr. Sophon Kaeothip was born on Octerber 31, 1980 in Bangkok. He obtained a Bachelor of Science, from Department of Chemistry, Faculty of Science, and Ramkhaheng University in 2000. In 2001, he was worked at Department of Pharmacognosy, Faculty of Pharmacy, and Mahidol University in position of Assistance researcher. He was admitted to the Master degree program in Organic chemistry at Chulalongkorn University in 2002.