CHAPTER 2

THEORETICAL AND LITERATURE REVIEWS

Theoretical background

2.1 Snake venom[2]

Snake venom is an aqueous solution which consists of many substances and is produced by salivary glands, which are located at the back of the snake's head (behind the eyes). These glands produce and store a cocktail of venoms which travel through ducts to either fangs or grooves in the teeth (back-fanged snakes). The proteinaceous nature of snake venom was established by Napoleon Bonaparte's brother, Lucien in 1843. Proteins are the major portion of venom's dry weight (90%). Other portions are composed of nonprotein components, which are divided into inorganic and organic constituents, are as follows:

1) Inorganic constituents
    (1) Metal content calcium, magnesium, sodium, iron etc.
    (2) Biological significance
2) Organic constituents
    (1) Amino acids and small peptides
    (2) Nucleotides and related compounds
    (3) Carbohydrates
    (4) Lipids
    (5) Biogenic amines

The major portion is a cocktail of hundreds, sometimes thousands, of different proteins and enzymes. There are approximately 20 types of toxic enzymes found in snake poisons. Each of these enzymes has its own special function. Some aid in the digestive process, while others specialize in paralizing the prey. Scientists believe they

have identified the following chemicals from snake venom and the specific purpose of each as follows:

1) cholinesterase attacks the nervous system, relaxing muscles to the point where the victim has very little control.
2) amino acid oxidase plays a part in digestion and the triggering of other enzymes, (is responsible for venom's characteristic light yellowish coloring.)
3) hyaluronidase causes other enzymes to be absorbed more rapidly by the victim.
4) proteinase plays a large part in the digestive process, breaking down tissues at an accelerated rate. (causes extensive tissue damage in human victims)
5) adenosine triphosphatase believed to be one of the central agents resulting in the shock of the victim and immobilizing smaller prey. (probably present in most snakes.)
6) phosphodiesterase accounts for the negative cardiac reactions in victims, most notably a rapid drop in blood pressure.

Due to the complicated components, snake venom has its complicated pharmacological effect and toxicological effect. Different snake venom has different venomous properties. Therefore, the symptoms of snakebite possess different characteristics. Their main characteristics are as followed

1) Blood toxin
2) Myotoxin
3) Nerve toxin

2.2 Mechanism of neurotoxin

In general, neurotoxic venom affect to the neuromuscular junction which contains the membrane of the axon terminal of the motor neuron, also called presynaptic membrane, or presynaptic site; the motor end-plate of muscle cell, also called

postsynaptic membrane, or postsynaptic site; and the space between them called synaptic cleft. Here the events of presynaptic membrane, synaptic cleft, and postsynaptic membrane is delineated:

1) Presynaptic membrane

The neurotransmitter, acetylcholine is synthesized in the cytosol of the axon terminal, and must be taken up by the synaptic vesicle. Acetylcholine release is triggered by the arrival of an action potential in the axon terminal. The depolarization of the presynaptic membrane causes voltage-gated calcium channels in the active zones to open. The extracellular concentration of calcium is far higher than the intracellular concentration of calcium, so calcium ions will flood the axon terminal as long as the calcium channels are open. The resulting elevation in the internal calcium ion concentration is the signal that causes neurotransmitter to be released from synaptic vesicle. The vesicle releases their contents by a process called exocytosis. The membrane of the synaptic vesicle fuses to the presynaptic membrane at the active zone, allowing the contents of the vesicle to spill out into the synaptic cleft. The venom which generates this action is king cobra and cobra.

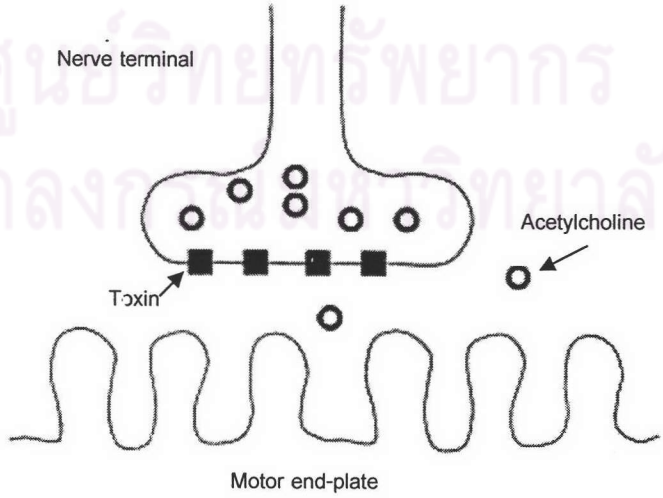


Figure 2.1 Pre-synaptic block.

2) Postsynaptic membrane

Acetylcholine released in the synaptic cleft affect the muscle cell by binding to thousands of specific receptor proteins that are embedded in the motor end-plate. The binding of neurotransmitter to the receptor is like inserting a key in a lock; this causes conformation changes in the protein. Receptor proteins also called Acetylcholine-gated sodium ion channels. They are membrane-spanning proteins consisting of five subunits that come together to form a pore between them. In the absence of neurotransmitter, the pore is closed. When neurotransmitter binds to a specific site on the extracellular region of the channel, the pore is open due to the conformation changes. The resulting elevation in the internal sodium ion concentration depolarizes the muscle cell from resting membrane potential, and eventually causes muscle to contract. The venom which generates this action is banded kraits.

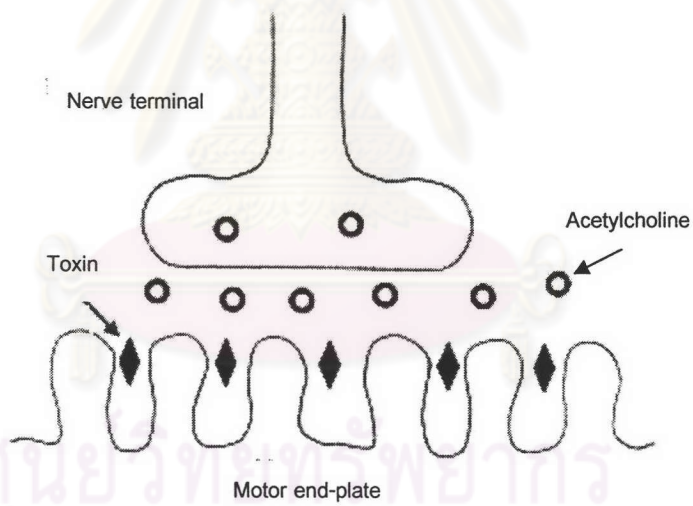


Figure 2.2 Post-synaptic block.

Interestingly, Malayan kraits venom has both actions in their venom activity[3,4].

2.3 Bungarus candidus

The kraits (Genus *Bungarus*)[5] are found in Southeast Asia, Thailand, China and South Indonesia. They are members of the Family *Elapidae*, and are all of secretive habits and placid disposition. The kraits do not bite readily but the venom is generally highly lethal, so they are considered extremely dangerous. Malayan krait (*Bungarus*

candidus) snakes are considered to be the smallest in *Bungarus* species which have alternating black and white bands on their triangle-shape bodies.

Figure 2.3 Malayan krait (*Bungarus candidus*).

In Thailand, it is found in the north, east and south (Jintakune et al., 1990) part. Since these snakes are nocturnal, most bites occur at night or early in the morning. Ptosis, muscular weakness, difficultly in breathing, pupillary dilation, and tachycardia are important symptoms of this snake bite patients[1].

2.4 Separation Technique

2.4.1 Sodium dodecyl sulfate polyacrylamide gel electrophoresis[6]

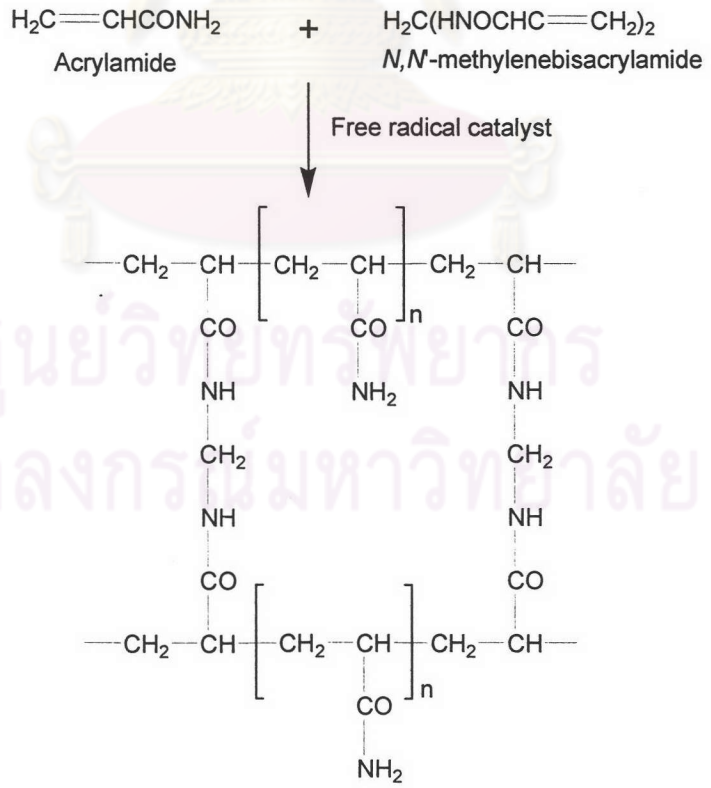
Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) is the most widely used method for analyzing protein mixtures qualitatively. It is particularly useful for monitoring protein purification and, because the method is based on the separation of proteins according to size. The method can also be used to determine the relative molecular mass of proteins.

2.4.1.1 Components of SDS-PAGE

There are two major important components are as follows:

1) Polyacrylamide[6]

Cross-linked polyacrylamide gels are formed from the polymerization of acrylamide monomer in the presence of smaller amounts of *N, N'*-methylenebisacrylamide (normally referred to as bis-acrylamide). Note that bis-acrylamide is essentially two acrylamide molecules linked by a methylene group, and is used as cross-linking agent. Acrylamide monomer is polymerized in a head-to–tail fusion into long chains and occasionally a bis-acrylamide molecule is built into the growing chain, thus introducing a second site for chain extension. Preceding in this way a cross-linked matrix of fairly well-defined structure is formed.



Scheme 2.1 The formation of a polyacrylamide gel from acrylamide and bis-acrylamide.

The polymerization is initiated by the addition of ammonium persulfate and the base *N, N, N', N'*-tetramethylenediamine (TEMED). TEMED catalyses the decomposition of the persulfate ion to give a free radical (i.e. a molecule with an unpaired electron)

Free radicals are highly reactive species due to the presence of an unpaired electron that needs to be paired with another electron to stabilize the molecule. Therefore, free radical reacts with acrylamide monomer molecule, forming a single bond by sharing its unpaired electron with one from the outer shell of the monomer molecule. This therefore produces a new free radical molecule, which is equally reactive and will attack a further monomer molecule. In this way long chains of acrylamide are built up, being cross-linked by the introduction of the occasional bis-acrylamide molecule into the growing chain. Oxygen removes free radicals and therefore all gel solutions are normally degassed prior to use.

Polyacrylamide gels are transparent and flexible, yet relatively strong and resilient. They are chemically inert and are compatible with numerous buffers, salts and detergents. Overall, protein mobility through polyacrylamide gels is proportional to the pore size, which is a function of both the acrylamide concentration (%T) and that of the bis-acrylamide cross-linker (%C). In general, the pore size is inversely proportional to %T[7].

$$\%T = \frac{\text{acrylamide (g) + bis-acrylamide (g)}}{\text{100 ml}} \times 100\% \quad (1)$$

$$\%C = \frac{\text{bis-acrylamide (g)}}{\text{acrylamide (g) + bis-acrylamide (g)}} \times 100\% \quad (2)$$

The composition of any given polyacrylamide gel is described by two parameter, %T and %C. The %T value represents the total concentration of

monomer to produce the gel, and %C is the percentage of the total monomer which is the cross-linking agent. For any given total monomer concentration, the effective pore size, stiffness, brittleness, light scattering, and swelling properties of the polyacrylamide gel vary with proportion of cross-linker used[8].

2) Sodium dodecyl sulfate[7]

Sodium dodecyl sulfate (SDS) is an anionic detergent that is used to denature proteins, giving them all the same conformational properties, and to prevent protein interactions during electrophoresis. Sodium dodecyl sulfate also masks the intrinsic protein charge and gives all proteins a similar net negative charge, so that electrophoretic migration is towards the anode. In general, sodium dodecyl sulfate coats proteins with a uniform negative charge and constant charge-to-mass ratio. Hence, in an electric field SDS-coated proteins experience the same field strength and migrate at identical intrinsic rates towards the anode. Proteins of different size are subjected to different degrees of sieving by the gel, however, and the mobility of a protein is in inversely proportional to its size.

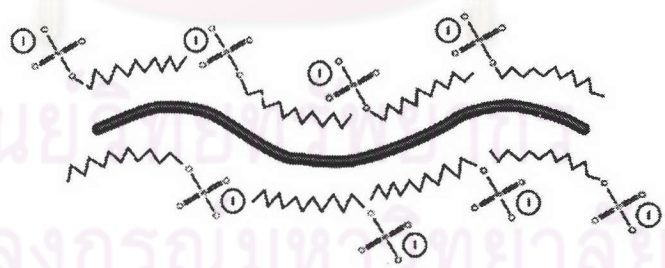


Figure 2.4 The action of dodecyl sulfate in denaturing proteins.

Uniform dodecyl sulfate binding is critical for the appropriate electrophoretic migration, and it generally approximates one sodium dodecyl sulfate molecule per two amino acid residues, or 1.4 g SDS/g protein. However, the exact shape of sodium dodecyl sulfate-coated proteins remains unclear.

2.4.1.2 One-dimensional gel electrophoresis[6]

The gel of one-dimensional gel electrophoresis has two parts. The first part is stacking gel. The stacking gel is used to concentrate the protein sample into a sharp band before it enters the main separating gel. The stacking gel has a very large pore size, which allows the proteins to move freely and concentrate under the effect of electric field. Then negatively charged protein-SDS complexes continue to move towards the anode through the second part which is separating gel. The smaller proteins are more easily pass through the pores of gel, whereas large proteins are successively retarded by frictional resistance due to the sieving effect of gel.

2.4.1.3 Two-dimensional gel electrophoresis[9]

The first dimension: Isoelectric focusing

Isoelectric focusing is performed in a pH gradient. Proteins are amphoteric molecules with acidic and basic buffering groups. Those proteins become protonated or deprotonated depending on the pH environment. In basic environment the acidic groups become negatively charged, in acidic environment the basic groups become positively charged. The net charged of a protein is the sum of all negative or positive charges of the amino acid side chains. When a protein is placed at certain in pH value of the gradient, and an electric field is applied, it will start to migrate towards the electrode of the opposite sign of its net charge. Because it migrates inside a gradient, it will arrive at a pH value of its isoelectric point (pI). At the pI, the net charge is zero and it stops migrating. When it diffused away (above or below its pI), it will become charged again and migrate back to its pI. This is called the focusing effect, which results in very high resolution.

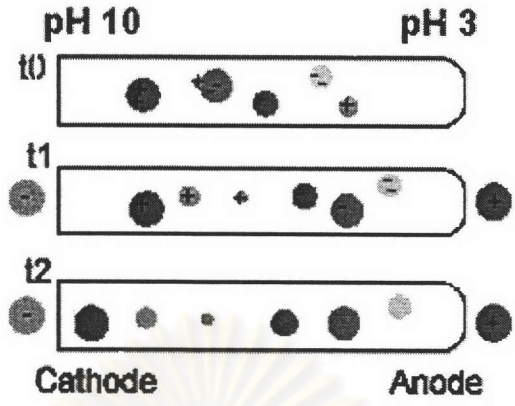


Figure 2.5 Migration of proteins to their pI in the pH gradient of an isoelectric focusing gel.

The second dimension: SDS-PAGE

With SDS-PAGE, the polypeptides are separated according to their molecular weights. All proteins will migrate towards the anode. The electrophoretic mobility of proteins treated with SDS and dithiothretitol (DTT) depends largely on the molecular weight of the protein. At a certain polyacrylamide percentage there is an approximately linear relationship between the logarithm of the molecular weights and the relative migration distance of the SDS-polypeptide complexes of a certain molecular weight range. The molecular weights of the sample proteins can be estimated with the help of co-migrated standards with known molecular weights.

2.4.2 Ion-exchange Chromatography[10]

An ion exchanger consists of a matrix with charged groups. A cation exchanger contains negative charged and anion exchanger positive charges. Weak ion exchangers contain weakly acidic or basic groups and strong ion exchangers a strong acid or base. The ionization of weak ion exchangers is pH dependent and is characterized by their pKa values, whereas the charged of strong ion exchangers is constant in the pH range for protein chromatography.

Ion-exchange chromatography mainly exploits differences in charge, but it is not always valid for proteins, since them posses an asymmetrical distribution of charges and a region where one type of charged dominates can orient the molecule toward an ion exchanger of opposite sign. A protein can be retarded on an ion exchanger at its isoelectric point (pI) when the net charge is zero or even adsorb to a cation exchanger above its pI. However, separation not only depends on differences in net or local charge but also hydrogen bonding and hydrophobic interaction affect to the chromatography process. The electrostatic attraction between a protein and an ion exchanger is influenced by the following:

1) Ionic strength one type of buffer ion competes with protein for the ionic sites on the absorbent. Buffer ions of the opposite charge tend to bind to the charged groups of the protein that interact with the ion exchanger. They compete with the adsorbent for the ionic sites on the protein. The influence of buffer ions is weak at low ionic strength, and therefore the electrostatic attraction is stronger and the protein is adsorbed to the ion exchanger. The more increasing of ionic strength, the more decreasing of the electrostatic attraction, and when the ionic strength reaches a critical value the protein is desorbed and elutes from the column. Although, there is still some attraction between the protein and the ion exchangers, and this delays the protein elution from the column. A further increase in ionic strength eventually abolishes the electrostatic attraction between the protein and the absorbent, and the proteins elutes without being retarded. Adsorption at low ionic strength and desorption by increasing the ionic strength is the most common method used for ion-exchanger chromatography of proteins.

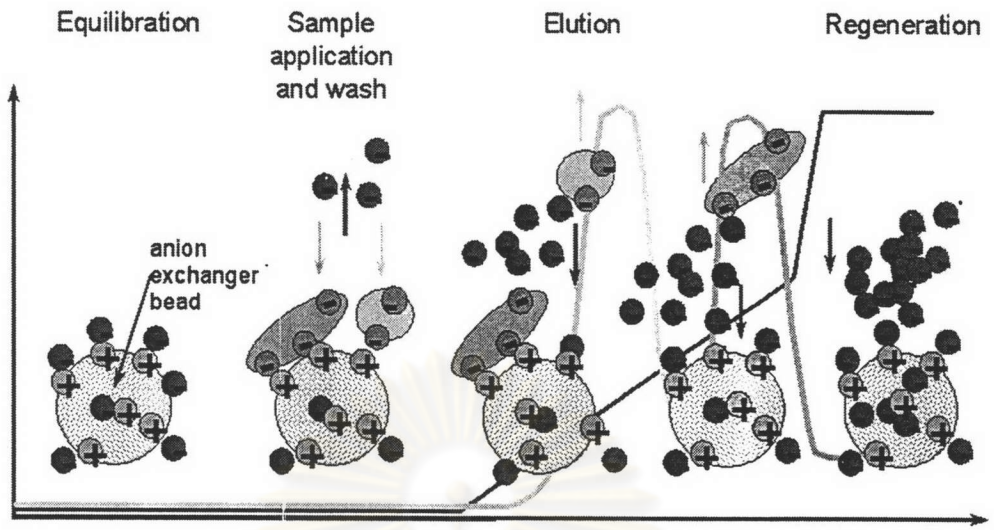


Figure 2.6 The action of ion in anion exchanger .

2) pH a change of pH alters the net charge of the protein and the charge density of a weak ion exchanger. Below the pI the net charge of a protein is positive and cation exchangers are used; at the pI, the net charged is zero and ion-exchange chromatography cannot be used; and at pH values above the pI, when the net charged is negative, anion exchangers are used. This is a normal case, but as pointed out above there are exceptions with proteins containing clusters of positive or negative charges. The pKa values of common functional groups of ion exchangers are given in table1-1. Weak cation exchangers are usually used at pH > pKa and anion exchangers at pH < pKa. Under these conditions the exchangers have a large number of ionic sites and there fore a good protein-binding capacity.

Table 2.1 Media used in ion exchange chromatography

Exchanger	Formula	pKa
Caion exchanger :		
CM; carboxymethyl	MX-OCH_2COOH	3.5-4.0
Orthophosphate	MX-OPO_3H	3.0 & 6.0
SE; sulfoxyethyl	MX-$OCH_2CH_2SO_3H$	2.0
SP; sulphopropyl	MX-$CH_2CH_2CH_2SO_3H$	2.0-2.5
S; sulphonate (Pharmacia mono S)	MX-CH_2SO_3H	2.0
Anion exchanger :		
DEAE; diethylaminoethyl	MX-$OC_2H_4NH^+(C_2H_5)_2$	9.0-9.5
QAE; quaternary aminoethyl, diethyl-(2-hydroxypropyl)-aminoethyl	MX-$OC_2H_4N^+(C_2H_5)_2$ $CH_2CH(OH)CH_3$	-
QA; trimethylhydroxypropyl	MX-$OCH_2CH(OH)CH_2N^+(CH_3)_3$	-
Q; quaternary amine (Pharmacia mono Q)	MX-$CH_2N^+(CH_3)_3$	-

MX = matrix

2.5 Identification technique

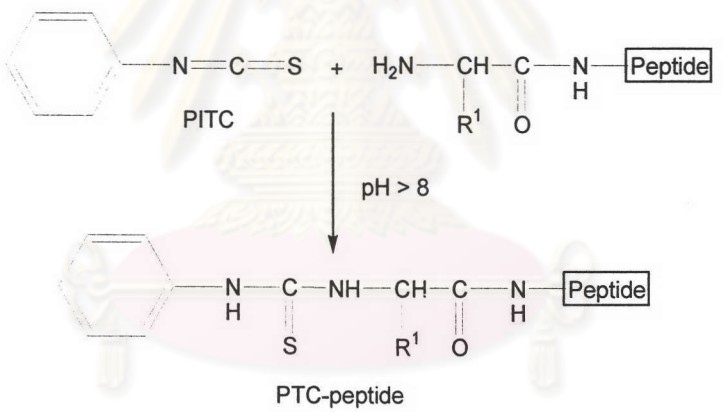
2.5.1 Edman degradation[11]

The sequential degradation of proteins and peptides was studied by the Swedish scientist Pehr Edman from 1949 and became known as Edman degradation (Edman 1949) which cleaves the N-terminal amino acid from a peptide or protein backbone and prepares the derivatized residue for its identification. In this way the amino acid sequence of protein is determined by repetitive chemical reaction. Edman degradation comprises three individual steps: the coupling, the cleavage and the conversion. Finally, a derivatized amino acid, the PTH amino acid, is produced and identified.

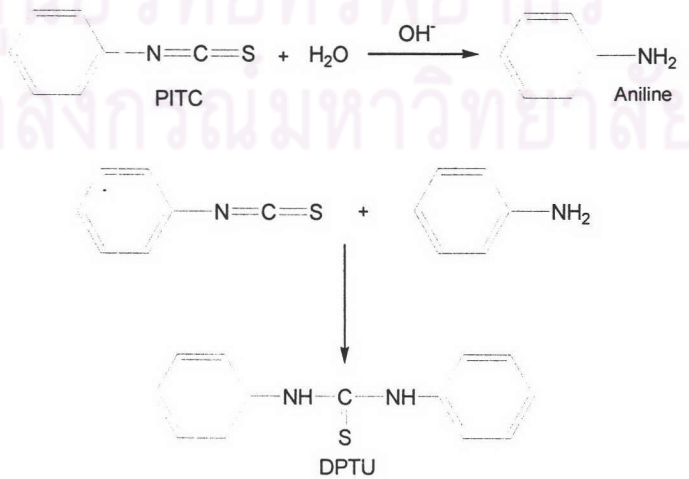
2.5.1.1 Coupling, cleavage and conversion

The first step couples the Edman reagent phenylisothiocyanate (PITC) to the free N-terminal amino group of a peptide chain. The reaction takes place within 15-30 minute at a basic pH and a temperature of 40-55 ^{0}C in a very high yield. The resulting phenylthiocarbamyl peptide is washed by some hydrophobic solvent to get rid of excessive reagent and reaction by-product.

An unprotonated amine group is required for the coupling reaction and alkaline conditions of at least pH9 are suitable. A further increase in pH would support reactivity but also increase side reactions, e.g. the alkali-catalyzed hydrolysis of PITC, giving aniline. Additionally, the free amine group of aniline reacts with PITC, yielding diphenylthiourea (DPTU), the most intensive by-product of Edman degradation.

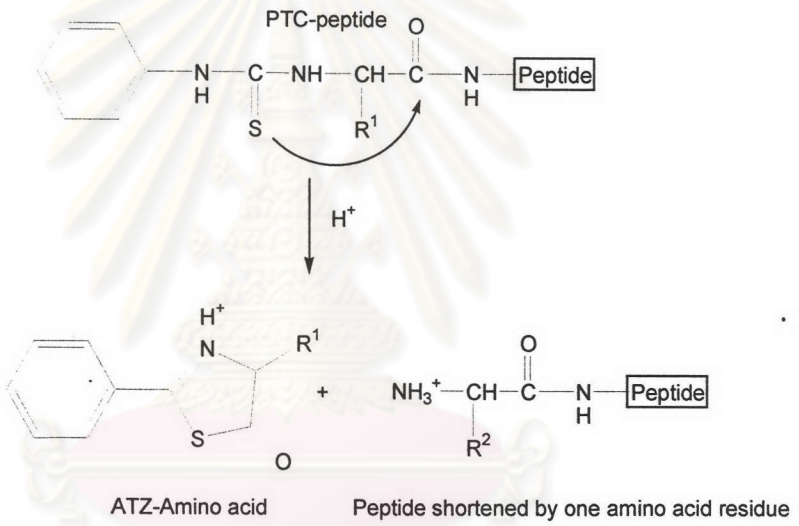


Scheme 2.2 The coupling reaction.



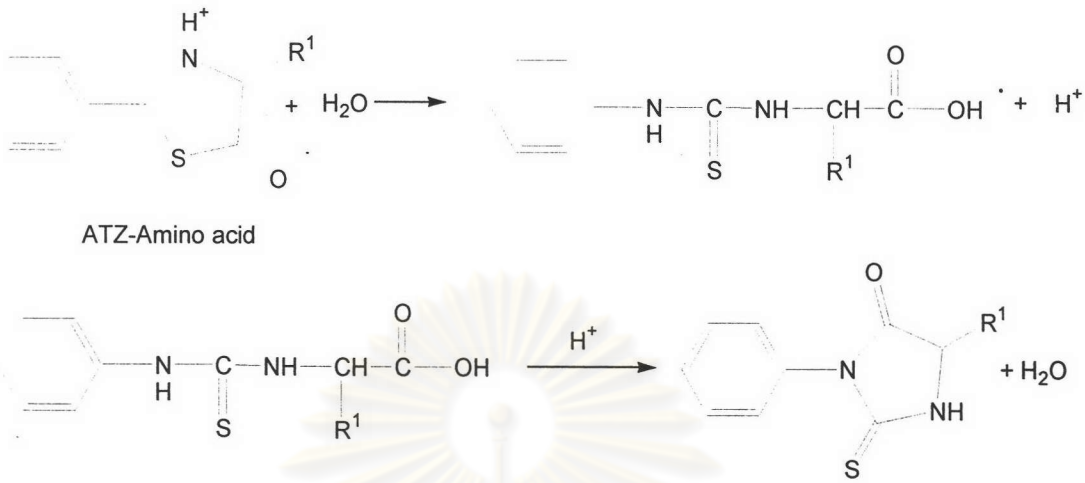
Scheme 2.3 Formation of the by-product DPTU during Edman degradation.

The dried PTC peptide is treated with dehydrated acid, e.g. trifluoroacetic acid. The first amino acid is cleaved as heterocyclic derivative, an anilinothiazolinone (ATZ) amino acid, after the nucleophilic attack of the sulphor atom to carbonyl group of the first peptide bond. This emphasizes the importance of the sulphor atom in PITC; its nucleophilicity is needed for the cyclization. The corresponding oxygen-containing reagent, the phenyisocyanate, also has the ability to couple to the amino group, but cyclization and cleavage is not possible. In conclusion, any sulphor or oxygen exchange must be avoided and this achieved by inert gas atmosphere during the sequencing reaction.



Scheme 2.4 The cleavage reaction.

The small, hydrophobic ATZ amino acid has a significantly different solubility from that of the hydrophilic peptide. It is extracted by a non-polar solvent, e.g. chlorobutane or ethylacetate. In a following step, called the conversion, the instable ATZ amino acid is rearranged, yielding the stable PTH derivative. The shortened peptide comprises a new N-terminus and after drying it can be subjected to another reaction cycle.



ATZ-Amino acid

Scheme 2.5 The conversion step.

The unstable ring structure of the ATZ amino acids is opened by the influence of the watery acid and an increased temperature. Rearrangement takes place, yielding the more stable PTH amino acid.

2.5.1.2 Identification of the PTH amino acids

PTH amino acids are analyzed with a limit in the upper femtomole range after separation by reverse phase high performance chromatography (RP-HPLC). Chromatographic identification and quantitation of the UV signals is done by reference to the retention times and the absorbance of a PTH standard run. The PTH amino acids display characteristic UV spectra with an absorbance maximum at 269 nm.

Another approach is using mass spectrometric identification of amino acid derivatives. Coupling reagent like 3-[4'(ethylene-N,N,N-trimethylamino)phenyl]-2-isothiocyanate result in products which are extremely sensitive for mass analysis (Aebersold 1992). However, once more the quantitative yield for coupling and cleavage is troublesome.

2.5.2 Mass spectrometry[7]

The molecular weight is a highly specific characteristic of a molecule and is often the first physiochemical property that is measured of a protein. Traditionally, SDS-PAGE or gel permeation chromatography is used for molecular weight measurements. Recently, mass spectrometric techniques have been developed that permit mass determination of intact proteins with an accuracy far superior to the above mentioned methods.

2.5.2.1 Historical Perspective[12]

Sir Joseph J. Thomson (1856-1940) conceptualized the idea of mass spectrometry (MS) in 1897 through his cathode ray tube experiment. He measured the mass-to-charge ratio (m/z) of the negatively charged cathode ray particles by passing the collimated beam through crossed electric and magnetic fields. The birth of mass spectrometry is credited to his work on the analysis of positive rays with a parabola mass spectrograph in the early part of the twentieth century.

2.5.2.2 Basic concept of mass spectrometry

Mass spectrometry is an analytical technique that measures the masses of individual molecules and atoms. The first essential step in mass spectrometry analysis is to convert the analyze molecules in to gas-phase ionic species because one can experimentally manipulate the motion of ions, and to detect them (which are not possible with neutral species). The excess energy transferred to molecule during the ionization event leads to fragmentation. Next, a mass analyzer separates these molecular ion and their charged fragments according to their m/z ratio. The ion current due to these mass-separated ions is detected by a suitable detector and displayed in the form of a mass spectrum. To enable the ions to move freely in space without colliding or interaction with other species, each of these steps is carried out under high vacuum (10^{-4}–10^{-8} torr).

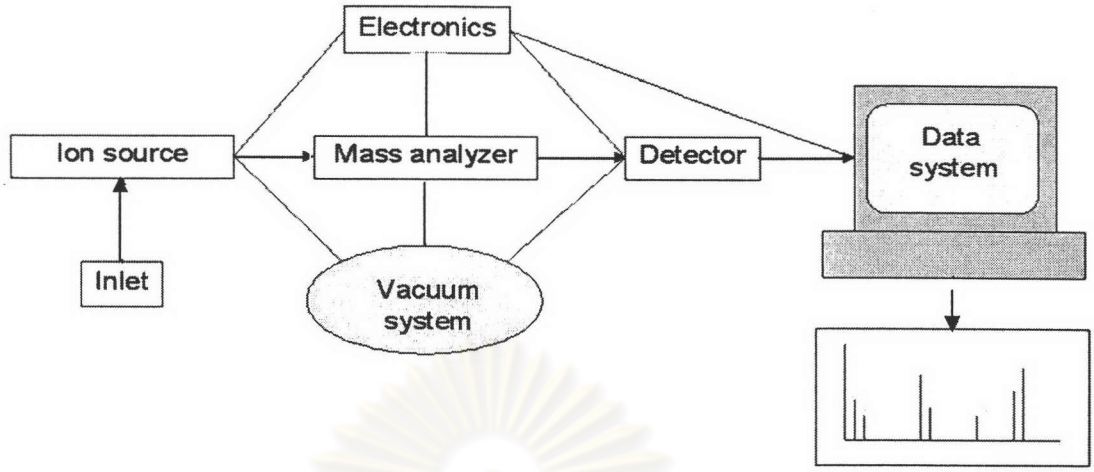


Figure 2.7 Basic component of a mass spectrometer.

Thus, a mass spectrometer consists of several essential functional units. These units are

1. An inlet system to transfer a sample to the ion source
2. A vacuum system to maintain a very low pressure in the mass spectrometer
3. An ion source to convert the neutral sample molecules into gas phase ions
4. A mass analyzer to separate and mass-analyze ionic species
5. A detector to measure the relative abundance of the mass-resolved ions
6. Electronics to control the operation of various units
7. A data system to record, process, store, and display the data

A mass spectrometer can be effectively split in to three constituent parts; the area of production, the ion source; the area of ion separation, the mass analyzer; and the area of ion detection, the detector.

2.5.2.3 The ion source

The ion source is the region of mass spectrometer where the gas phase ions are produced from sample molecule, the area of ion production. Various procedures are used to form gas phase ions from molecules, depending on the physical state of the analytes. Choices are available as to the types of ions produced (positively and negatively charged, radical cations and protonated molecules, etc.) and the degree,

which these ions are internally excited. Internally excited molecular ions dissociate to produce fragment ions, which may reveal details molecular structure. On the other hand, an intact molecular ion (such as the protonated molecule or radical anion) provides information on molecular weight.

The two ionization methods most commonly used to volatized and ionize the proteins or peptides for mass spectrometric analysis are:

1) Matrix-Assisted Laser Desorption/ionization (MALDI)

2) Electrospray ionization (ESI)

1) Matrix-Assisted Laser Desorption/ionization (MALDI)

Karas and Hillenkamp introduced matrix-assisted laser desorption/ionization (MALDI) in 1988 as a technique that could readily ionize biomolecules in a very sensitive manner. MALDI is a pulse ionization technique which utilizes the energy from a laser to desorb and ionize the analyte molecules in the presence of a light absorbing matrix. As this is a pulsed ionization technique, generating packets of ion with each laser pulse, a pulsed analyzer is typically required for separation and resolution of the ions. Consequently, MALDI is routinely coupled with a time-of-flight analyzer (TOF)[9]

Mechanism of matrix-assisted laser desorption ionization[11]

The matrix is believed to serve three major functions:

(1) Absorption of energy from the laser light. The matrix molecules absorb the energy from the laser light and transfer it into excitation energy of the solid system. Thereby an instantaneous phase transition of a small volume (some molecular layer) of the sample to gases species is induced. In this way the analyte molecules are desorbed together with matrix molecules, with limited internal excitation.

(2) Isolation of the biomolecules from each other. The biomolecules are incorporated in a large excess of matrix molecules, strong intermolecular forces are thereby reduced (matrix isolation). Incorporation of analyte into matrix crystal taking place upon evaporation of the solvents forms an essential prerequisite for successful MALDI analysis, it moreover provides an in-situ cleaning of the sample and is the reason for a high tolerance against contaminants.

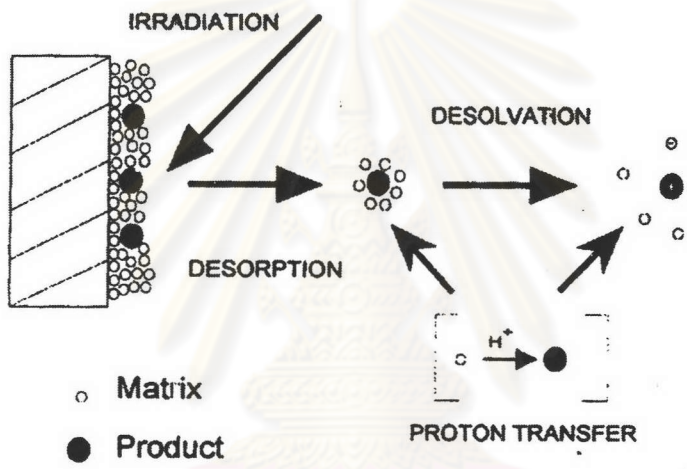
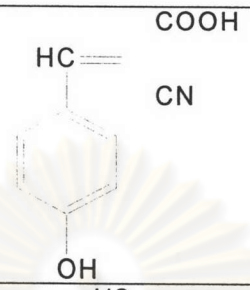
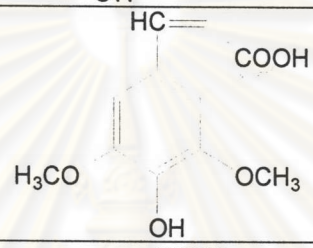
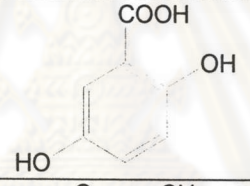
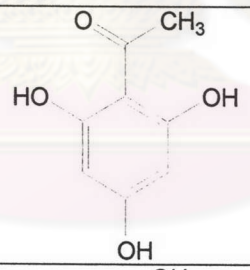
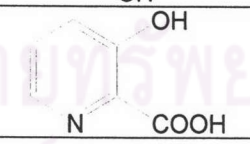


Figure 2.8 Diagram of the MALDI process.

(3) Ionization of the biomolecules. An active role of the matrix in the ionization of the analyte molecules by photoexcitation or photoionization of matrix molecules, followed by proton transfer to the analyte molecules is likely though not proven unequivocally to date.

A wide range of matrices for bio-mass spectrometry applications have been adopted for use with UV laser (the typical wavelength is 337 nm) as shown in Table 2.2

Table 2.2 Common MALDI matrices used in biological applications[9].

Matrix	Matrix structure	Application
α-cyano-4-hydroxycinnamic acid		Peptide analysis and Protein digests. Analyte < 10 kDa
Sinapinic acid (4-hydroxy-3,5-dimethylcinnamic acid)		Analysis of large polypeptides and proteins > 10 kDa
2,5-dihydroxybenzoic acid (2,5 DHB)		Protein digests and proteins Oligosaccharide released from glycoproteins
2,4,6-trihydroxyacetophenone (THAP)		Oligonucleotides < 3kDa
3-hydroxy picolinic acid		Oligonucleotides > 3kDa

Sample Preparation

The preparation of the sample for MALDI analysis requires utmost care. The homogeneity of the sample-matrix mixture is a critical factor to obtain good sample ion yields. Fortunately, MALDIMS is somewhat more tolerant of impurities, buffers, salts and mixtures. Several techniques have emerged for the sample preparation. These techniques include the

(1) Dried-droplet technique

(2) Fast evaporation technique

(3) Sandwich matrix technique

(4) Spin-dry technique

(5) Seed-layer technique

2) Electrospray ionization (ESI)[12]

Electrospray ionization (ESI) is an atmospheric pressure ionization technique applicable to a wide range of compounds that are present in liquid matrices. The emergence of ESI represents a significant advance in the capabilities of mass spectrometry for the characterization of large biomolecules. It has also become the most widely used interface to combine HPLC with mass spectrometry. The wide popularity of ESI in these fields is due to its continuous-flow operation, tolerance to different types of solvent, acceptance of wide solvent flow rates, and ability to generate intact multiply charged ions of fragile chemical and biochemical species. ESI uses a novel concept first introduced by Dole in 1968 for the generation of gas phase ions from electrically charged liquid droplet, which are produced by electrospraying the solution of an analyte at atmospheric pressure. The coupling of ESI to mass spectrometry was achieved in mid-1980s by Fenn and co-workers and Aleksandrov et al.. Several research groups, notably of Fenn, Henion, and Smith, later developed the applications of ESIMS for the analysis of biomolecules. With ESIMS, the molecular masses of these biopolymers in the mass range of over 100kDa can be determined with an accuracy of > 0.01%

Principle

The basic principle of electrospray ionization can be explained through a simple schematic of a typical ESI source as shown in Figure 2.9.

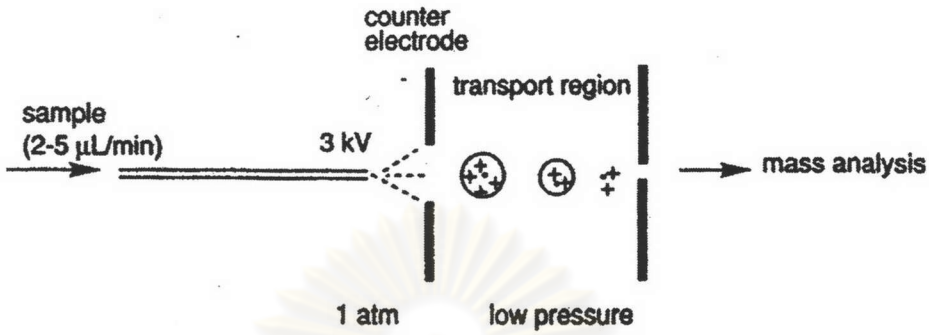


Figure 2.9 A pictorial depiction of the concept of electrospray ionization.

The heart of the ESI source is a stainless-steel capillary tube through which a solvent flows continuously at the rate of 2-5 µL/min. The solvent consists of a mixture (usually 1:1) of water and an organic solvent (typically methanol, acetonitrile, or isopropanol) and typically contains <1% acetic acid or another suitable acid or another suitable acid. A solution of the sample is injected into this solvent stream. A potential difference of 3-4 kV between the tips of the capillary and walls of the surrounding atmospheric pressure region produces an electrostatic field sufficiently strong to disperse the emerging solution into a fine mist of charged droplets. A flow of hot-bath gas, usually nitrogen, is added to the interface to assist in the evaporation of the solvent from those charged droplets, decreasing its diameter. Consequently[11], the charged density on its surface increases until the so-called Rayleigh limit is reached, at which the coulomb repulsion becomes of the same order as the surface tension. The resulting instability, sometimes called a coulomb explosion, tears the droplet apart, producing charged daughter droplets that also evaporate. This sequence of event repeats and finally produces droplets so small that the combination of charge density and radius of curvature at the droplet surface produces an electric field intense enough to finally desorb ions from the droplets into the ambient gas phase. After that, the ions are transported from the atmospheric pressure region to the high vacuum region of the mass analyzer.

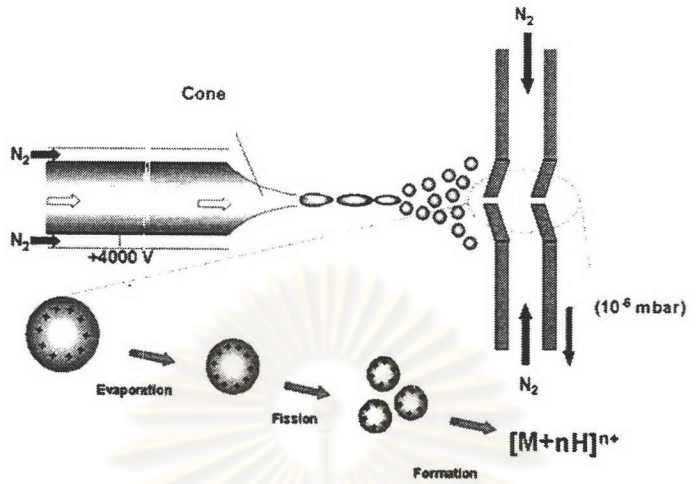


Figure 2.10 Mechanism of evaporation and exploration of droplets.

2.5.2.4 Mass Analyzer

Ions can be separated on the basis of their mass-to-charge ratios using electric or magnetic fields arranged so as to spread them in time or space. The two mass analyzers most commonly used for mass spectrometric analysis are:

(1) Time-of-flight (TOF)

(2) Quadrupole (Q)

Time-of-flight (TOF)[12]

A time-of-flight (TOF) mass spectrometer is one of the simplest mass-analyzing devices. Since the 1990s, it has reestablished itself as a mainstream technique and is becoming increasingly useful in meeting the demands of contemporary research in biomedical sciences.

Principle of operation

A TOF mass spectrometer behaves as a velocity spectrometer, on which ions are separated on the basis of their velocity differences. A short pulse of ions, after exiting the source, is dispersed in time by allowing it to drift in a field-free region of a long flight tube. The principle behind the mass analysis is that after acceleration to a constant kinetic energy (equal to zV, where z is the charge on the ion and V the accelerating potential), ions travel at velocities, v, that are an inverse function of the square root of their m/z values:

$$v = \left(\frac{2zV}{m}\right)^{1/2} \qquad (3)$$

The lighter ions travel faster and reach the detector placed at the end of the flight tube (of length L) earlier than do the heavier ones. Thus, a short pulse of ions is dispersed into packets of isomass ions. Therefore, mass analysis of ions that enter the flight tube can be accomplished by determining their time of arrival given by

$$t = \frac{L}{v} = L\left(\frac{m}{2zV}\right)^{1/2} \qquad (4)$$

In order to convert the time spectrum into a mass spectrum, the instrument is mass calibrated by measuring the flight times of two different known mass ions.

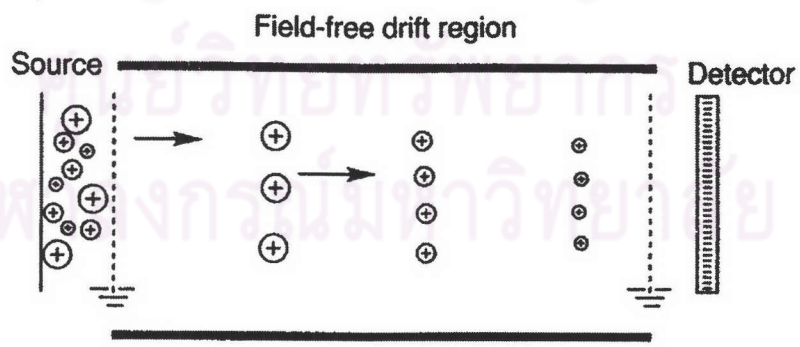


Figure 2.11 Principle of the mass separation by TOF.

Quadrupole (Q)[12]

The quadrupole mass spectrometer is the most widely used type of mass spectrometer. The mass separation in this instrument is accomplished solely by using

electric fields. Quadrupoles are dynamic mass analyzer, which means that ion trajectories are controlled by a set of time-dependent forces that are generated by applying direct current (dc) and radiofrequency (rf) potentials to a set of electrodes.

Principle

A quadrupole mass analyzer is two-dimensional quadrupole field device. As shown in Figure 2.12, it consists of four accurately aligned parallel rods that are arranged symmetrically in a square array. The field within the square array is created by electrically connecting opposite pairs together. Ions are injected at one end of the quadrupole structure in the direction of the quadrupole rods (z direction). The separation of different *m/z* ions is accomplished through the criterion of path stability within the quadrupole field. In other woeds, at a given set of operating parameters, ions of very narrow *m/z* range have stable trajectories (i.e., their motion is confined within the field-defining electrodes), whereas the remainder of the ions will have unstable trajectories (i.e., the amplitude of their motion exceeds the boundaries of the electrodes). To obtain a mass spectrum, the quadrupole field is varied to force other ions to follow the stable path. Thus, its function is analogous to a variable narrowband filter.

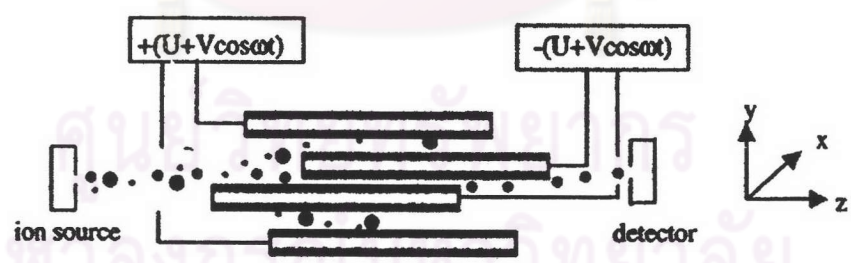


Figure 2.12 A quadrupole mass analyzer.

The quadrupole field is created by supplying a positive direct current (dc) potential *U* and a superimposed radiofrequency (rf) potential $V\cos\omega t$ (i.e., $U- V\cos\omega t$) to one pair of rods (where ω is the angular frequency related to the frequency f, in hertz, by $\omega = 2\pi f$, *V* the amplitude of rf voltage, and *t* the time). The other opposing pair of rods receives a dc potential of $-U$ and an rf potential of magnitude $V\cos\omega t$, but

out of phase by 180 [i.e., -(U- $V\cos\omega t$)]. This arrangement creates an oscillating field such that the potential Φ at any point within the rods is given by

$$\Phi_{(x,y)} = \Phi_0 \frac{x^2-y^2}{r_0^2} \quad (5)$$

Where Φ_0, the applied potential, is equal to U- $V\cos\omega t$, r_0 the inscribed radius between the rods (i.e., one-half the distance between the opposite electrodes), and x and y the distances from the center of the field. It is obvious from this equation that the potential is zero at the center of the square array along the z axis (i.e., where x and y are equal to zero). The potential is also zero when the value of x and y are equal.

The motion of an ion in the x and y directions is described using a quadratic equation of the form

$$\frac{d^2u}{d^2\xi} + (a_u - 2q_u\cos 2\xi)u = 0 \quad (6)$$

which is commonly known to mathematicians the *Mathieu equation*, where u represents the transverse displacement in the x and y directions from the center of the field, is equal to $\omega t/2$, and the dimensionless variables a and q are given by

$$a_u = a_x = -a_y = \frac{8eU}{m\omega^2 r_0^2} \quad (7)$$

$$q_u = q_x = -q_y = \frac{4eU}{m\omega^2 r_0^2} \quad (8)$$

The solution to the Mathieu equation shows trajectories for some ion are stable and for others unstable in the quadrupole field.

In order to obtain a mass spectrum, the quadrupole field is changed by simultaneously scanning U and V, while keeping their ratio and f constant. Less commonly, the applied frequency is changed at fixed values of U and V. This way, ions of different m/z can be brought into the stability region, and transmitted along the length of quadrupole field.

2.5.2.5 Detector

After ions were separated by a mass analyzer, they reach at a detector for the detection of their mass and abundance. A detector measures the electric current in proportion with the number of ions striking it. Sensitivity, accuracy, resolution, and response time are the most important characteristics of any detector. These are major instruments for ion detection:

1) Photomultiplier
2) Secondary electron multiplier
3) Post acceleration Detector

2.5.3 Tandem mass spectrometry[9]

Tandem mass spectrometry or MS/MS was first used in the late 1960s. Since that time, the applications and popularity of this technique continue grow. This technique is performed with instrument capable of selecting ions of particular m/z value and subjecting the selected ions to fragmentation within the mass spectrometer. Generally, these experiments are performed successfully on two types of instrument where analyzers are in series (tandem in space) such as the triple quadrupole and hybrid quadrupole-TOF and secondly instruments which employ ion trapping mechanisms such as the quadrupole ion trap and FT-ICR analyzer (tandem in time).

Considering[13], an instrument that combines in sequence two mass spectrometers separated by a collision cell, MS/MS analysis can be carried out in either of three scan modes, all represented in Figure 2.13

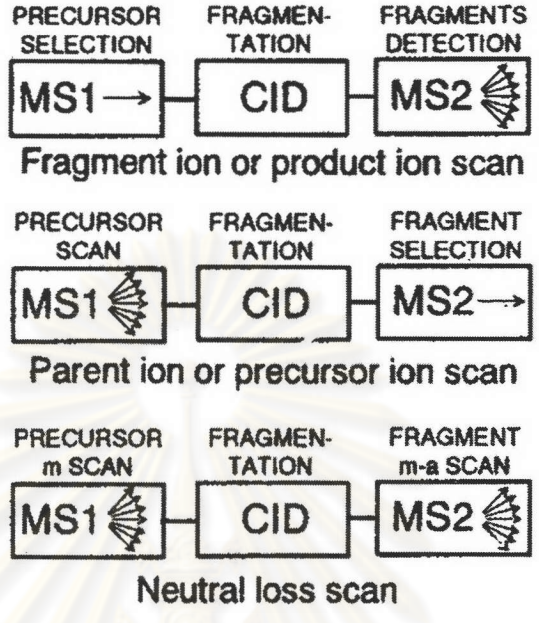


Figure 2.13 Main processes in tandem mass spectrometry (MS/MS).

2.5.3.1 **Product ion scan**[9] (daughter scan) the first part of the analyzer, MS1, is used to specifically select the ion of interest. The precursor ion is allowed into the collision cell where it will undergo CID. Here, the peptide precursor ions collide with molecules of the collision gas (typically argon, helium[14], xenon or nitrogen) causing the precursor ion to fragment which yielding a distribution of fragment ions, or product ions. The product ions are scanned by the third analyzer, MS2, and then detect at the detector producing a product ion spectrum.

As a results, of the low energy collisional fragmentation the peptide precursor ion fragments predictably at each peptide amide bond along the peptide backbone yielding a distribution of product ions in complimentary ion series forming a ladder which is indicative of the peptide sequence[15].

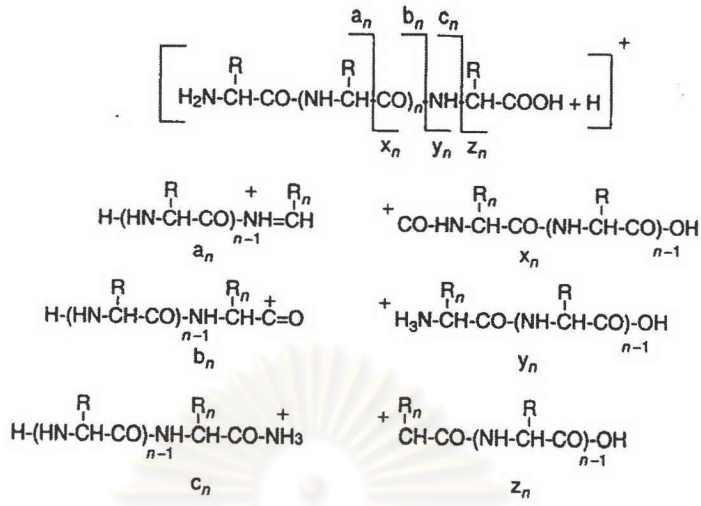


Figure 2.14 The nomenclature of the common peptide fragment ions.

The peptide is fragmented forming two complimentary ion series[9]

- The N-terminal ion series or a, b and c ion series. The ions of the N-terminal ion series will contain the N-terminal amino acid and extensions from this residues (Figure 2.13)

- The C-terminal ion series or x, y and z ion series. The ions of the C-terminal ion series will contain the C-terminal amino acid and extensions from this residues (Figure 2.13)

The structure of the internal amino acids shown in this figure have the form –NH-CH(R)-CO-, where R is the side chain of the amino acid. This structure is referred to as the the "residue structure" of an amino acid, and the formula weight of this moiety is defined as its residue mass. Table 2.3 lists the residue masses of the 20 amino acids and a selected number of modified amino acids. These residue structures and the masses are central to the interpretation of product ion spectra because they provide the means for distinguishing the difference amino acids. The N-terminal amino acid has the structure NH_2-CH(R)-CO-, which includes an additional H relative to the residue structure. The C-terminal amino acid has the structure –NH-CH(R)-COOH, which includes an additional OH relative to the residue structure.

These structural differences relative to the residue masses in Table 2.3, which allow recognition of the N- and C-terminus of the peptide sequence.

Table 2.3 The residue masses of the 20 genetically encoded amino acids and selected modified amino acids[15].

Amino acid	One-letter code	Residue mass (Da)	Immonium ion (m/z)
Glycine	G	57.02	30
Alanine	A	71.04	44
Serine	S	87.03	60
Proline	P	97.05	70
Valine	V	99.07	72
Threonine	T	101.05	74
Cystine	C	103.01	76
Leucine	L	113.08	86
Isoleucine	I	113.08	86
Asparagine	N	114.04	87
Aspartate	D	115.03	88
Glutamine	Q	128.06	101
Lysine	K	128.09	101
Glutamate	E	129.04	102
Methionine	M	131.04	104
Histidine	H	137.06	110
Phenylalanine	F	147.07	120
Arginine	R	156.10	129
Tyrosine	Y	163.06	136
Tryptophan	W	186.08	159

The product ion MS/MS spectrum can be used to identify a protein by manually determining the peptide sequence and using the determined sequence to search the sequence databases or by automatically correlating the native product ion MS/MS spectra with the sequence database.

2.5.3.2 **Precursor ion scan**[9] (parent scan) consists in choosing a product ion (or daughter ion) and determining the precursor ions (or parent ions).

2.5.3.3 **Neutral loss scan**[9] consists in selecting a neutral fragment and detecting all the fragmentation leading to the loss of that neutral. With a reactive gas (CAR), adduct ions can also be observed.

Hybrid mass analyzers (Quardrupole-time of flight)[9]

These instruments have rapidly become the standard instrument for MS/MS applications in proteomic fields. By combing a mass filtering quadrupole analyzer and a collision cell with a non-scaning reflectron TOF analyzer, the user is able to acquire MS and, most notably MS/MS data with high mass accuracy, resolution and sensitivity.

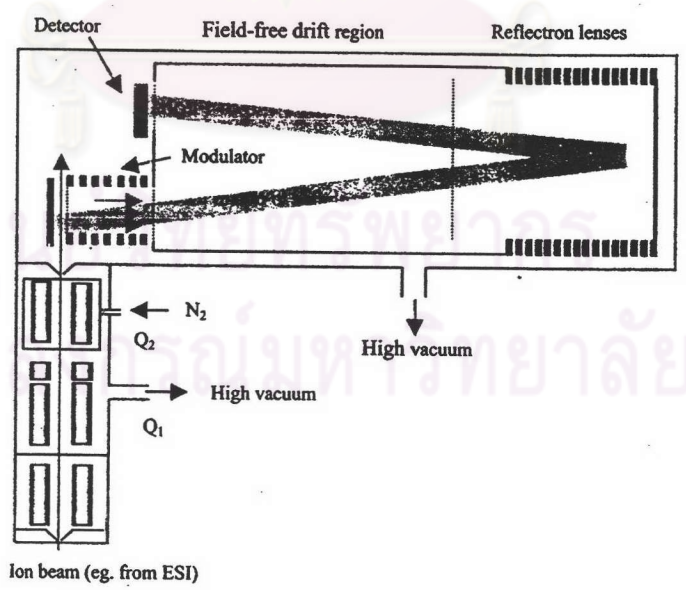


Figure 2.15 Schematic diagram of quadrupole/time of flight hybrid tandem mass spectrometer[12].

2.6 Protein identification by database searching[9]

Mass spectrometry has become the preferred method of choice for protein identification. However, the development of several software programs and continuous updating of sequence databases have been crucial to the success of mass spectrometry in this field.

Generally, MS data can be used in four approaches for protein identification

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 peptides.

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

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

As a rule, programs for protein identification by mass spectra use freely available databases containing amino acid sequences of protein. The most widely known database was shown below [16].

- SWISS-PROT is a database of annotated protein sequence; it also contains additional information on function of the protein, its domain structure, posttranslational modification, etc.;

- TrEMBL is a supplement to SWISS-PROT, which contains all protein sequences, translated from nucleotide sequences of the EMBL database;

- NCBInr (National Center of Biotechnological Information) is a database containing sequences translated from DNA sequences of GenBank and also sequences from PDB, SWISS-PROT and PIR database

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

The existing algorithms and corresponding programs may be subdivided into three main groups:

- program using proteolytic peptide fingerprint for protein identification (PepIdent, MultiIdent, Profound)
- program additionally operating with MS/MS spectra (MASCOT, MS-Fit)
- program operating with MS/MS (SEQUEST, PepFrag, MS-Tag)

2.7 Literature Reviews

The snake *Bungarus candidus*, Malayan krait is common in Southeast Asia and ranges from Thailand to Malaysia and Indonesia. In 1983, Warrel *et al.*[4] reported that *Bungarus candidus* caused severe neurotoxic envenomtion in man and Tan N-H. *et al.*[5,17] described that this snake venom had high acetylcholinesterase and high hyaluronidase activities in 1990. There are 22 proteins in Swiss-Prot and TrEMBL database[18] were shown below.

In 2000, Kuhn P. *et al.* found a high-resolution structure of bucandin[19, 20], a novel neurotoxin that enhances acetylcholine release from nerve terminals, was determined by ^{1}H-NMR spectroscopy and molecular dynamics. This 63-amino-acid polypeptide belongs to a family of three-finger toxins that incorporate five disulphide bridges and was purified to homogeneity by two-step methods which were gel filtration and high performance liquid chromatography.

In 2002, Watanabe L. *et al.* [21] described bucain is a three-finger toxin, structurally homologous to snake-venom muscarinic toxins. These proteins have molecular masses of approximately 6000-8000 Da. Bucain was crystallized in two crystal forms by the hanging-drop vapour-diffusion technique.

Parvathy VR. *et al.* (2000) reported candoxin[22, 23] which is a novel three-finger toxin from *Bungarus candidus* and was purification by consecutive gel filtration and reverse phase HPLC and determination of the N-terminal amino acid sequence by Edman degradation.

In the literature, Tsai I. H. *et al.* (2002) found phospholipase A2 (PLA2). The PLA(2) cDNAs from the venom gland of *B. candidus* (Indonesia origin) were amplified by the polymerase chain reactions (PCR) and cloned. The primers used were based on the cDNA sequences of several homologous *B. multicinctus* venom PLA(2)s. In addition to the A-chains of beta-bungarotoxins, a novel *B. candidus* PLA(2) was cloned and its full amino acid sequence deduced. Having totally 125 amino acid residues, the PLA(2) contains a pancreatic loop and is 61% identical to the acidic PLA(2) of king cobra venom.

From the TrEMBL database[18], Tsai I.H. et al. (2001) reported 16 proteins. They are Kunitz inhibitor c (Fragment), Kunitz inhibitor b (Fragment), Kunitz inhibitor a (Fragment), Beta bungaratoxin B2b chain (Fragment), Beta bungaratoxin B2a chain (Fragment), Beta bungaratoxin B1 chain (Fragment), Beta bungaratoxin A2 chain (Fragment), Beta bungaratoxin A1 chain (Fragment), Weak toxin 3, Weak toxin 2, Weak toxin 1, Candiduxin 2, Candiduxin 1, Alpha bungaratoxin, Kappa 1b bungaratoxin and Kappa 1a bungaratoxin.

In 2003, Kuch U. *et al.*[18] reported Alpha-bungarotoxin (A31) precursor (Fragment) and Cytochrome b.

Nawarak J. et al.[25] (2003) reported that proteins component from ten species of snake venoms in the Elapidae and Viperidae families which had been separated by

different chromatographic methods including RP-HPLC, SDS-PAGE, 2-DE and mass spectrometry. The protein patterns of snake venom from various techniques shows a large quantity of low-molecular-mass basic proteins for the Elapidae family and neutral proteins of different molecular masses for the Viperidae family.