PURIFICATION AND CHARACTERIZATION OF PROTEASE FROM *ARTOCARPUS INTEGER* LEAF

By

SITI BALQIS BINTI ZULFIGAR

Thesis submitted in fulfillment of the requirements for the Degree of Master of Science

July 2012

ACKNOWLEDGEMENTS

In this great opportunity, I would like to express my most sincere appreciation to Associate Professor Dr. Rosma Ahmad for her guidance, constructive comments, constant encouragement and invaluable advice throughout the duration of this Msc. study.

I am also grateful for the financial assistance from the Academic Staff Training Scheme of Universiti Sains Malaysia and the Ministry of Higher Education.

A note of gratitude goes to Mr. Zainoddin Osman, Mrs Najmah and Mr Azmaizan Yaakub for their support and assistance. I wish to state here that all encouragement and moral support rendered by all lab mates, Mr Mohd Fadzirul Hezly, Miss Nadiah Hassan, Miss Anis Najiha. Miss Tan Zi Ni, Mr Ho Chun Hoong and other postgraduate students will not go forgotten. They will always be greatly cherished.

Finally, I would like to express my indebtedness to my mother, Mrs Azizon Md Nor, father, Mr Zulfigar b Yassin, husband, Mr Saiful Syarizal Yaakob and mother in law, Mrs Aminah Ayub for their love that had and always be my greatest encouragement in life. Thank you for everything.

Siti Balqis Zulfigar

October 2011.

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOL / ABBREVIATION	CAPTION
%	percentage
ºC	Degree Celsius
Da	Dalton
k	kilo
m	mili
M	molar
µ	micro
h	hour
APS	ammonium persulfate
DIFP	diisopropylphosphofluoridate
EDTA	ethylenediaminetetraacetic acid
E-64	trans-epoxysuccinyll-leucylamido-(4-guanidino) butane
KCl	potassium chloride
NaOH	sodium hydroxide
$NaHCO_3$	sodium hydrogen carbonate
PMSF	phenylmethylsulfonylfluoride
PVPP	polyvinylpolypyrrolidone
SBTI	soybean trypsin inhibitor
SDS-PAGE	sodium duodecyl sulphate polyacrylamide gel electrophoresis
TCA	tricholoroacetic acid
Vc	column volume

PENULENAN DAN PENCIRIAN ENZIM PROTEASE DARI DAUN *ARTOCARPUS INTEGER*

ABSTRAK

Kehadiran enzim protease di dalam daun cempedak (*Artocarpus integer*) yang digunakan sebagai pelembut daging telah dikenalpasti apabila satu jalur pada 69 kDa didapati menerusi kaedah zimografi kasein dan kaedah elektroforesis gel poliakrilamida natrium dodesil sulfat (SDS PAGE). Penulenan enzim yang dilakukan menerusi kaedah pembezaan suhu dan fasa dengan menggunakan 6 % (i/i) Triton X-114, pemendakan dengan 44 % (b/i) ammonium sulfat dan kaedah kromatografi penurasan gel telah menghasilkan enzim yang ditulenkan sebanyak 12 kali ganda dan aktiviti spesifik yang bernilai 76.67 U/mg. Enzim protease yang telah ditulenkan ini mempunyai suhu optimal pada 40ºC dan pH optimal pada pH 10.0. Selain itu, enzim ini juga berkeupayaan menguraikan substrat semulajadi pada suhu dan pH optimal menurut urutan, gelatin < BSA < kasein. Ciri-ciri enzim dari kelas sistein telah dikenalpasti menerusi perencatan aktiviti oleh E-64 dan iodoasetamida pada kepekatan 1 mM yang menghasilkan baki aktiviti masing-masing sebanyak 30 $\pm$ 1.6 % dan 9 $\pm$ 0.8 %. Peningkatan aktiviti sehingga 212.1 % didapati setelah enzim ini diinkubasi bersama 5 mM sistein dan juga sebanyak 186.7 % apabila diinkubasi bersama 30 mM 2-merkaptoetanol. Secara keseluruhannya, aktiviti enzim ini dipengaruhi oleh kepekatan ion logam yang digunakan. Aktiviti enzim direncat dengan kehadiran ion Cu^{2+}, Mg^{2+}dan Mn^{2+} secara umumnya. Walau bagaimanapun, kadar perencatan oleh ion Mn^{2+} berkurangan apabila kepekatannya ditingkatkan kepada 10 mM. Ion Ca^{2+} didapati merencatkan aktiviti enzim pada kepekatan 1 mM dan bertindak sebagai pengaktif apabila kepekatannya ditingkatkan kepada 5 dan 10 mM. Kehadiran ion Zn^{2+} pula meningkatkan aktiviti enzim pada kepekatan 1 mM (108. 4 $\pm$ 3.3 %) dan

bertindak sebagai perencat pada kepekatan 5 dan 10 mM dengan aktiviti relatif masing-masing sebanyak 74.8 $\pm$ 4.1% and 64.3 $\pm$ 2.1 %. Enzim ini juga stabil sehingga suhu 70 ºC, dengan 80% daripada aktivitinya didapati masih kekal. Tenaga penyahaktifan dihitung dan memberikan nilai sebanyak 89 $\pm$ 11 kJ/mol.

PURIFICATION AND CHARACTERIZATIONS OF PROTEASE FROM *ARTOCARPUS INTEGER* LEAF.

ABSTRACT

The presence of a protease in *Artocarpus integer* leaves, which can be used as a meat tenderiser, was verified by the presence of a band at 69 kDa, using caseinolytic zymography and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE). Purification by temperature-phase partitioning with 6 % (v/v) Triton X-114, 44 % (w/v) ammonium sulphate precipitation and gel filtration chromatography yielded a preparation with a 12-fold increase in enzyme purity and a final specific activity of 76.67 U/mg. The purified protease was maximally active at 40ºC and at pH 10.0. The enzyme hydrolysed natural substrates in the order of gelatin < BSA < casein when tested at the optimal pH and temperature.The cysteinic nature of this enzyme was verified through the inhibition of the activity by E-64 and iodoacetamide at the concentration of 1 mM with the residual activity of 30 $\pm$ 1.6 % and 9 $\pm$ 0.8 % respectively. The enhancement of activity up to 212.1 % and 186.7 % were obtained after 15 min incubation with 5 mM cysteine and 30 mM 2-mercaptoethanol respectively. The activity of the enzyme was affected by the metal ions' concentrations. The enzyme was generally inhibited with the presence of Cu^{2+}, Mg^{2+} and Mn^{2+}. However, lower degree of inhibition of Mn^{2+} was detected when the concentration of the cation was increased to 10 mM. Ca^{2+} suppressed the activity at 1 mM and activated the enzyme at 5 and 10 mM. The presence of Zn^{2+} activated the activity of *A. integer* leaf protease at the low concentration of 1 mM (108. 4 $\pm$ 3.3 %) and suppressed the activity at the higher concentrations of 5 and 10 mM with 74.8 $\pm$ 4.1% and 64.3 $\pm$ 2.1 % relative activity respectively. The enzyme was stable at

temperatures up to 70 ºC, with 80 % of its activity intact and the energy of inactivation (E_{ia}) was calculated as 89 $\pm$ 11 kJ/mol.

CHAPTER 1

INTRODUCTION

1.1 Research background

Current market demand for proteases has contributed to an ever-increasing interest in discovering new enzymes with distinct substrate specificities. The quest for new enzyme sources often involves consideration of the ease of enzyme availability and the possible economic benefits that may arise from developing the potential enzyme source.

Plant extracts with high content of proteolytic enzymes have been used in the pharmaceutical industry as antitumors (Wald et al., 2001; Rajkapoor et al., 2007), treatment of digestive disorders (Mello et al., 2008), wound healer (Silva et al., 2003), treatment against osteoarthritis (Brien et al., 2004), modulation of immune response (Gaspani et al., 2002) and the list is non exhaustive. In the light of this matter, a number of reviews on the plant proteases as pharmaceutical and has been written (Maurer, 2001; Waller et al., 2001; Stepek et al., 2004; Salas et al., 2008; Chobotova, et al., 2010). Equally important are natural therapeutic agents produced from plants reportedly having little or no side effects with low toxicity (Moss et al., 1963). Given these points, it is understandable that the hunt for new enzymes with distinct specificities has become significant.

The leaves of *Artocarpus integer* are used as a meat tenderizer by rural Malay communities, suggesting the presence of strong proteolytic enzyme in this plant. This hypothesis was verified by previous work of Rosma et al. (2008) where *A. integer* leaves were screened together with jackfruit (*A. heterophyllus*) leaves, papaya (*Carica papaya*) leaves and unripe fruit, lemon grass (*Cymbopogon citrates*), galangal (*Alpinia galangal*), ginger (*Zingiber offinale*) and tumeric tuber (*Curcuma*

domestica) for their proteolytic activity. The experiment demonstrated that *A. integer* leaves contain the highest proteolytic activity with 1.70 ± 0.06 U/g, followed by papaya leaves with 0.97 ± 0.05 U/g, jackfruit leaves 0.34 ± 0.02 U/g, ginger 0.25 ± 0.07 U/g, lemon grass 0.25 ± 0.05 U/g, galangal 0.21 ± 0.03 U/g, unripe papaya fruit 0.14 ± 0.04 U/g and turmeric 0.11 ± 0.03 U/g. Further investigations were carried out by comparing the efficacy of the crude enzyme of the leaves of *A. integer* in meat tenderizing to a commercial protease. The study demonstrated positive results such as improved tenderness, augmented water holding capacity and increased protein solubility in the *A. integer*-treated sample. A remarkable deformation and disruption of muscle fibers and connective tissues in the *A. integer*-treated sample were observed through the scanning electron microscope (SEM) analysis thus strengthening the hypothesis of the presence of protease in the leaf.

The genus *Artocarpus*, which covers a vast area of South and Southeast Asia, New Guinea, the southern Pacific, Sri Lanka, India, Pakistan and Indo-China towards the Malaysian archipelago (Lemmens et al., 1995), has been used for various applications. *Artocarpus integer* (*Thunb. Merr*) syn. *A. chempeden* (*Lour*.) *Stokes* syn. *Polyphema champeden*, or commonly known as cempedak, is known for its edible fruit, and previous studies of this species have focused on the characterization of bioactive compounds and lectins derived from various parts of the plant (Lim et al., 1997; Boonlaksiri et al., 2000; Abdul Rahman et al., 2002). The studies regarding *Artocarpus* focused on the isolation and characterizations of bioactive compounds such as flavonoids, stilbenoids and arylbenzofurans from the barks, heartwoods (Parenti et al., 1998; Hakim et al., 2005; Syah et al., 2006), roots (Hakim et al., 1999) and aerial parts of the plant (Boonlaksiri et al., 2000) and also on lectins that are derived from various parts of the plant (Lim et al., 1997; Abdul Rahman et al., 2002).

These bioactive compounds have the capability of showing cytotoxicity against cancer cells, anthelmintic activity (Hakim et al., 2005) and antimalarial activity (Boonlaksiri, 2000).

Besides the discovery of *A. integer's* tenderizing capabilities, the commercial development of this protease is also supported by the availability of leaves from trimming and pruning methods used during industrial fruit cultivation. This agricultural activity removes excess leaves, twigs and small branches to avoid excessive tree growth and to divert nutrients more effectively to developing fruits (Yaacob, 1980), hence providing a continuous enzyme source. Given the ease of availability and abundant supply of the enzyme source, *A. integer* offers a new alternative for the profitable production of a beneficial enzyme that might replace the existing commercial proteases.

No literature however, addresses on the isolation and characterization of protease from *A. integer* to date. Therefore, it is highly desirable to obtain the fundamental knowledge of this enzyme, such as its stability, optimal working conditions and mechanistic class, as this information is essential before considering its potential uses in the industry.

1.2 Objectives of research

The present study was designed with the aim to achieve the following objectives:

1. To extract and purify the polypeptide with proteolytic activity from the leaves of *A. integer*.
2. To elucidate the properties of protease from *A. integer*.

CHAPTER 2

LITERATURE REVIEW

2.1 Description of *Artocarpus integer*

Artocarpus integer is known locally as cempedak (En); chempedak, campedak, baroh (Indonesia); chempedak (cultivated), bankong (wild) (Malaysia); sonekadat (Myanmar); champada (Thailand) or mit tó nù (Vietnam) (Jensen, 1995). The plant is evergreen; being able to grow up to 20 m tall with its bark is used as a source of timber (Jensen, 1995). Among the key characteristics of *A. integer* is the presence of hair on the leaves at the midrib, main veins and occasionally on the leaf margins (Rukayah, 1992). The leaves are obovate to elliptic, 5 – 25 cm long and 2.5–12 cm wide with cuneate to rounded base, pointed tip and 6 – 10 pairs of lateral veins curving forward and the leaf stalk is 1 – 3 cm long. *A. integer* is mostly appreciated for its seasonal fruit that is accompanied with a distinctive smell. They are cylindrical in shape with yellowish to brownish outer skin (Jensen, 1995). The fruit flesh can be eaten raw or cooked while the seeds can be eaten after roasting or boiling. The sturdy and durable barks are used for building construction, furniture and boats.

2.2 Distribution of *A. integer*

The genus *Artocarpus* covers the area of South and Southeast Asia, New Guinea, the southern Pacific, Sri Lanka, India, Pakistan and Indo-China towards the Malaysian archipelago (Lemmens et al., 1995) *Artocarpus integer (Thunb. Merr)* syn. *A. champeden* (*Lour*) *stokes* syn. *Polyphema champeden* spreads through the tropical regions of South East Asia where the greatest diversity are distributed

mainly in Indonesia, Malaysia, and the Philippines (Hakim et al., 2005). Fig 2.1 shows the distribution of *A. integer* that spreads through the South East Asian region.

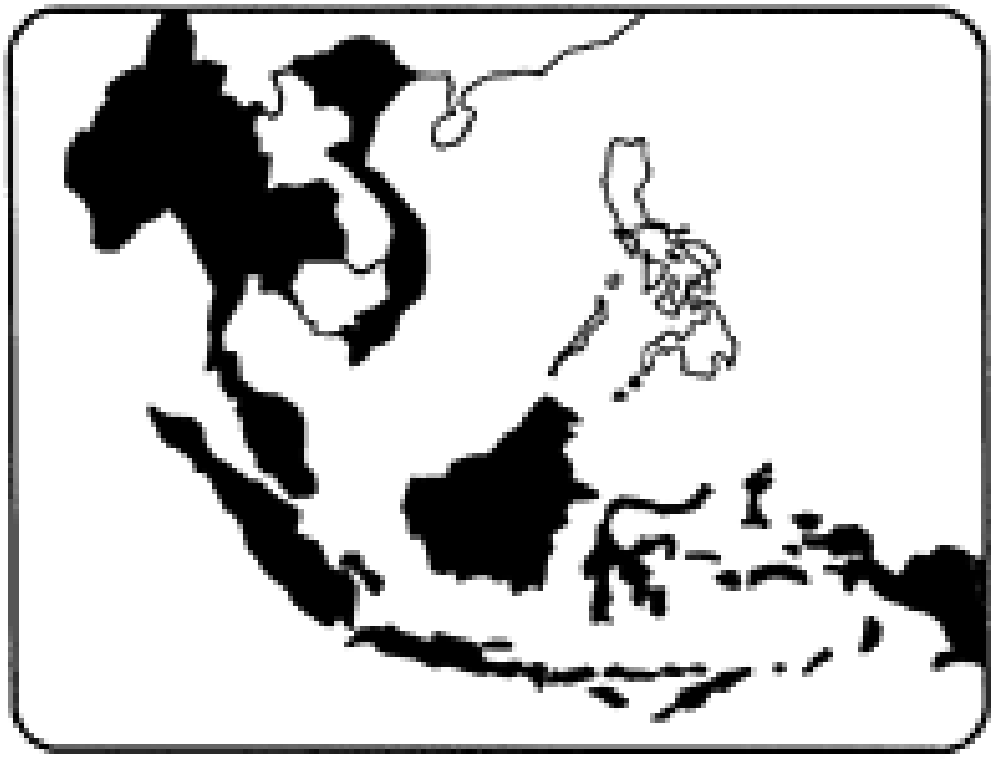

Fig. 2.1: The distribution of *A. integer* which spreads through South East Asia towards the Malaysian archipelago. Adapted from Jensen (1995).

2.3 Usage and potential of *A. integer*

Studies regarding *A. integer* primarily focused on the elucidation of the bioactive compounds such as flavonoids, stilbenoids and arylbenzofurans from the barks, heartwoods (Parenti et al., 1998; Hakim et al., 2005; Syah et al., 2006), roots (Hakim et al., 1999) and aerial parts of the plant (Boonlaksiri et al., 2000) and also on lectins that are derived from various parts of the plant (Lim et al., 1997; Abdul Rahman et al., 2002). These bioactive compounds have the capability of showing cytotoxicity against cancer cells, anthelmintic activity (Hakim et al., 2005) and antimalarial activity (Boonlaksiri, 2000).

Personal communication with rural Malay communities have led to the discovery of the traditional use of the leaves as meat tenderizer whereby the leaves are cooked together with meat for approximately 15 to 30 min in order to obtain the tenderizing effect. Rosma et al. (2008) who studied the capability of the crude extract as a meat tenderizer as compared to a commercial enzyme obtained positive results, such as improved tenderness, augmented water holding capacity and increased protein solubility. Remarkable deformation of muscle fibres and connective tissues were also observed.

In the agricultural industry especially in fruit cultivation, the method of trimming and pruning of the tree is essential in optimizing the production of fruits. This includes the activity of removing the non-productive parts such as the leaves and branches that grow excessively in order to divert the nutrients needed by the tree effectively (Yaacob, 1980). The work done by Bah et al. (2006) in characterizing cysteine protease in ten Malian medicinal plants demonstrated that the aerial section of plants showed the highest proteolytic activity compared to the other parts. Given the high content of protease in the leaves and abundant source of the enzyme, *A. integer* offers a new alternative in the production of profitable and beneficial enzyme that might replace the existing commercial or pharmaceutical proteases available.

2.4 Proteases

The global market value for industrial enzyme was estimated to be around USD 2.9 billion in 2008 with the growing average rate at 6 - 8 % (Danisco, 2012). Proteases or peptidases are a group of hydrolase enzymes that cleave peptide bonds. A report entitled ‘World Enzymes to 2011’ stated that the annual sales value of protease charted over 65 percent of total sales revenue generated by all industrial

enzymes combined (Freedonia, 2011 cited in Jaouadi et al, 2011, p. 445). Protease had found its application in numerous industries, may it be in the food, leather, diagnostics, waste management, detergent, silver recovery or textile and recently had been one of the major hit in the pharmaceutical area. The non-exhaustive list of pharmacological activities of plant proteases which include debridement of wounds from kiwifruit (Hafezi et al., 2009), treatment against digestive disorders in bromelain (Hale et al., 2005), role as anthelmintics (Stepek et al., 2004; Salas et al., 2008), antitumor activity (Maurer, 2001; Wald et al. 2001; Chobotova et al., 2010), treatment against osteoarthritis (Brien et al., 2004) and modulation of immune response (Gaspani et al., 2002), spark the interest in the hunt of new source of proteases.

2.4.1 General structure of protease

Proteases are generally composed of simple building structures which consist of amino acids with no protesthetic group or non protein material. Mihalyi (1978) summarized the general structure of protease as simple and constructed by a single unit which is devoid of allosteric mechanism. He also mentioned that the structure is generally near spherical, relatively small and compact compared to the other groups of enzymes. As most organisms are built of proteins, the proteolytic nature of proteases can be destructive even to the cells that manufactured them, thus most animal proteases are produced in an inactive form. These inactive precursors are called zymogens which require limited proteolysis in order to initiate a reaction (Mihalyi, 1978). However, nearly all plant and microbial proteases are produced in their active form where they are systematically segregated in order to provide spatial separation between the proteases and their substrates.

2.4.2 Classification of protease

Proteases are ubiquitous in nature and the efforts to classify this enzyme into smaller divisions have evolved through out the years. There are many approaches in classifying proteases into systematic structures which they are divided based on various common characteristics. The classification methods are based on several aspects which were summarized in the sub sections below:

2.4.2.1 Type of proteolytic reaction catalyzed

A scheme introduced by Bergmann and Fruton (1939) divides the proteases based on the location of attack of proteases on a polypeptide chain into exopeptidases and endopeptidases. The term exopeptidase refers to enzymes that split terminal amino acids of a polypeptide whereas endo refers to proteases that act on the internal bonds of a polypeptide chain (Fig. 2.2).

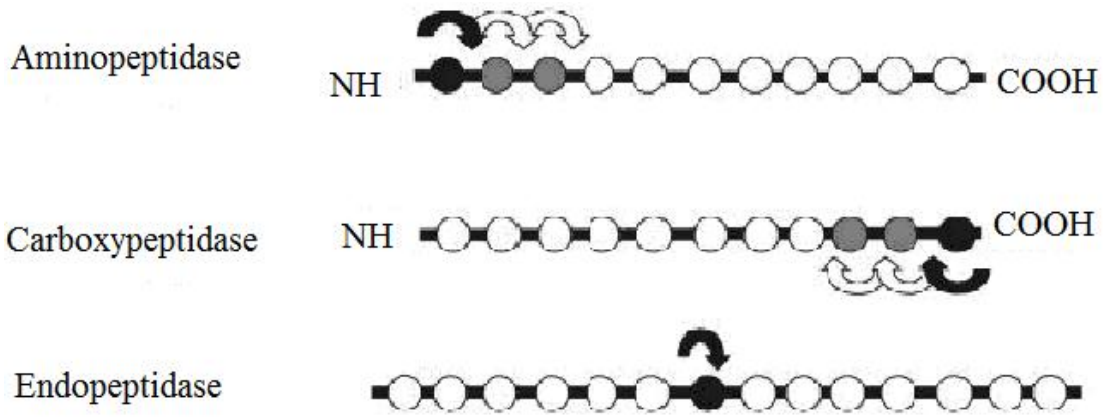

Fig. 2.2: Classification of proteases by reaction catalyzed. Peptides are represented as beads on string with each bead representing an amino acid and the string represents the peptide bonds. Black arrows indicated the first cleavage and white arrows show subsequent cleavages. The black amino acid indicated the first cleavage target and grey amino acids show the subsequent attacks. Adapted from Rawlings et al. (2007).

Exopeptidases can be further divided into two categories; the aminopeptidase which act on the N-terminus liberating free amino acids and carboxypeptidase which attacks the C-terminus (Barrett, 2000). Meanwhile endopeptidase refers to proteases that are able to cleave internal bonds in a polypeptide chain and they are capable to act on chains of any lengths.

2.4.2.2 Type of catalytic mechanism

In 1960, Hartley introduced a classification method based on the residue of the active site of the enzyme. This resolved to four main classes of proteases which include serine protease, cysteine/thiol protease, metal/metallo protease and aspartic/ acid protease. The classification scheme is recognized by the International Union of Biochemistry and Molecular Biology (IUBMB) where each family has a distinct characteristic set of functional amino acid residue in a configuration that forms the active site (Neurath, 1989). This method of classification is used in this dissertation as it is commonly applied by other researchers.

2.2.2.2. (a) Serine protease (EC 3.4.21)

Serine proteases are characterized by the presence of a serine residue in their active centre and specifically inhibited by diisopropylphosphofluoridate (DIFP) and also by other organophosphate derivatives (Mihalyi, 1978). Dunn (1989) proposed that the general mechanism of a serine protease involves the formation of covalent enzyme/ substrate complexes, with the serine residue being a strong nucleophilic amino acid in the active site. Therefore the mechanism of reaction of a serine protease involves a nucleophilic attack by the enzyme on the carbonyl group of the peptide bond.

Serine proteases include two distinct families that differ from each other in the amino acid sequence and the three dimensional structures. They are the mammalian serine proteases, for example chymotrypsin (EC 3.4.21.1), trypsin (EC 3.4.21.4) and elastase (EC 3.4.21.11) and the bacterial serine protease, for example subtilisin (EC 3.4.21.14) (Neurath, 1989). Although serine proteases were once thought to be rare in plants, recent discoveries of this class of enzyme in various plant species have proven otherwise.

Antao and Malcata (2005) presented a review on the plant serine proteases where they summarized that the molecular weights vary between 19 - 110 kDa with majority of it lies between 60 - 80 kDa. In addition, the optimum pHs are usually in the alkaline region that ranged between pH 7-11. The optimum temperature varies from 30 - 80 ºC with a majority of the proteases most effective at 20 - 50 ºC. The inhibitory mechanism of serine proteases involves the introduction of electrophilic groups such as DIFP, phenylmethylsulfonylfluoride (PMSF), soybean trypsin inhibitor (SBTI) which chemically modify the active site of the enzyme (Dunn, 1989).

2.2.2.2. (b) Cysteine protease (EC 3.4.22)

The mechanism of a cysteine protease is similar to the serine protease which in this instance, the attacking nucleophile is the sulphur atom of a cysteine side chain instead of a serine residue (Dunn, 1989). These proteases are widely distributed in living organisms which include plants (papain (EC 3.4.22.2), actinidain (EC 3.4.22.14), bromelain (EC 3.4.22.32)), mammalian (lysosomal cathepsins (EC 3.4.22.38), cytosolic calpains (EC 3.4.22.51)), viruses (coronavirus (EC 3.4.22.69) and bacteria (Staphylococcal cysteine protease). Papain which is extracted from the

latex of *Carica papaya* is the best studied member of the family of cysteine protease (Drenth et al., 1968). Cysteine proteases are often associated with plants with majority of plant proteases extracted belong to this class of enzyme. However their activity is easily affected by oxidation or heavy metal ions thus reducing their catalytic activity.

The molecular weight of cysteine proteases are mostly distributed between 20 - 50 kDa (Rao et al., 1998). Fahmy et al. (2004) summarized the molecular weight of plant cysteine proteases into the following groups, with barley ranged between 29 - 37 kDa (Koehler & Ho, 1988; Poulle & Jones, 1988; Zhang & Jones, 1996), maize cysteine protease (12-36 kDa) (Abe et al., 1977; De Barros & Larkins, 1990) and resting wheat grains (40-50 kDa) (Dominguez & Cejudo, 1995). Their optimum pHs are usually distributed at the neutral pH. Inhibition by iodoacetate and p-choloromercuric benzoate is among the key characteristic that define a cysteine protease. Furthermore, compounds that cause alkylation of the cysteine sulphydryl group such as iodoacetate, p-chloromecuric benzoate, N-ethylamide were reported to inhibit the activity of the enzyme (Uchikoba et al., 2002; Fahmy et al., 2004 Singh et al., 2010).

2.2.2.2 (c) Aspartic protease (EC 3.4.23)

Unlike cysteine and serine proteases, the mechanism of aspartic protease does not involved the nucleophilic attack by the functional group of the enzyme, hence there is no covalent intermediate formed between the enzyme and the substrate. Dunn (1989) summarized that the mechanism of this protease relies more upon the general acid/general base catalysis attack of a water molecule which is devoid of the aggressive nucleophilicity as in serine and cysteine protease. The enzyme has two

aspartic residues that are involved in the catalytic activity and show preferences on the cleavage of the peptide bonds between hydrophobic amino acid residues (Frazao et al., 1999). Aspartic enzymes were reported to be produced initially in the form of zymogens and the mechanisms of activation of these precursors were studied and elucidated (Van den Hazel et al., 1997; Varon et al., 2006).

Rennet, which has chymosin as its active component is probably the most widely used aspartic protease which is obtained from the stomach of calves. However this enzyme is rather expensive and scarce in nature (Vioque et al., 2000). Therefore, a more economical method of recombinant enzymes in microbial hosts is produced as an alternative to the existing enzyme (Pitts et al., 1993; Vega-Hernandez et al.,2004; Kappeler et al., 2006). Aspartic proteases were also discovered in plants for example in *Ficus racemosa* (Devaraj et al., 2008) and *Momordica charantia L.* (Wang et al., 2008), in nematodes (Jolodar et al., 2004) and fungi (Flentke et al., 1999). Aspartic proteases are characterized by being most active in the acidic pH (pH 2 - 3) (Dunn, 1989) and the molecular weights of the protease are mostly distributed in the range of 30 - 45 kDa (Rao et al., 1998). In addition to that, pepstatin is recognized as a potent inhibitor to aspartic proteases.

2.2.2.2 (d) Metallo-protease (EC 3.4.24))

The catalytic activity of metallo-protease relies on the presence of divalent cations in the active site most commonly zinc but may be replaced by other cations such as nickel or copper (Hooper, 1994). The mechanism of metallo-protease is similar to aspartic protease through the perspective of the non-involvement of nucleophilic attack. Carboxypeptidase A is recognized as one of the most studied

metallo protease in the aspect of its kinetic, spectroscopic and crystallographic method of analysis (Dunn, 1989).

The optimum pH of metallo-protease normally distributed in the range of pH 7-8 (Bolumar et al., 2001) The protease is also depicted as the most unstable group compared to the other class of protease and expected to undergo autolysis above pH 9 and below pH 6 (Salleh et al., 2006). Chelating agents such as ethylenediaminetetraacetic acid (EDTA) has the capability of binding cations hence inhibiting the enzyme's activity. However the usage of a more specific agent such as 1, 10-phenanthroline is more specific towards the inhibition of metallo-proteases.

2.4.2.3 Structural relationships

A more recent method of enzyme classification was developed by Rawlings and Barrett (1993). The classifications of the proteases are based on the structural features that are believed to reflect the evolutionary relationships and the system is known as *MEROPS.* This scheme grouped the enzymes based on the similarities of amino acid sequence and closely related peptidases are grouped into a family. Families that show evidence of common origins will then be grouped in a clan (Barrett, 2000).

2.5 Plant proteases

Proteases are ubiquitous in nature and they are isolated from various sources of organisms which include animals, plants, microorganisms and viruses (Neurath, 1989). The interest received by plant proteases have increased over the pass two decades demonstrating the significance of this enzyme in the industry. Proteolytic enzymes from plant proteases have been receiving special attention due to their

property of being active over a wide range of temperature and pH (Uhlig, 1998). In context with the present matter, most of the proteases derived from plants are classified as cysteine with rare occurrence of plant serine protease (Asif-Ullah et al., 2006; Singh et al., 2010) and aspartate proteases (Frazao et al., 1999; Devaraj et al., 2008; Wang et al., 2008). However recent discoveries of serine and aspartate proteases are encouragingly increasing and these enzymes possess properties that make them distinct and unique. In the light of this matter, serine proteases have no requirement for any co-factors and more stable towards oxidative reactions, in contrast with cysteine proteases which are known to be easily affected by air oxidation or metal ions (Kaneda et al., 1997).

Enzymes from plant extracts have had a long history especially in the ethno-pharmaceutical applications when the enzymes were administered as traditional medicine. These plant derived proteases have become increasingly important in the modern pharmaceutical industry as an alternative to synthetic drugs (Salas et al., 2008). Unlike synthetic drugs, protease such as bromelain was reported to have a very low oral toxicity with few or no side-effects. As an illustration to this subject, study by Moss et al., (1963) on mice, rats and rabbits discovered that no LD_{50} could be determined with oral doses up to 10 g/kg and no immediate toxic reactions were detected from the intake of bromelain. Such findings stimulate researches on discovering novel plant proteases with promising capacities.

Papain and bromelain are prominent among plant proteases and they are used extensively in the industrial and medical applications which will be further elaborated in Section 2.5.5. The extensive studies on these enzymes have provided users with the information and knowledge that allows the incorporation of the enzymes in the development of their products. For example, Enzyme Development

Corporation (EDC) had developed various enzyme products based on these two enzymes such as Enzeco®, Panol®, Liquipanol® and many others (EDC, 1999). However, these prominent enzymes have their own limitations and weaknesses that limit their capacity and applications. Therefore the searches of new plant proteases are often targeted to provide alternatives to the existing plant proteases. The knowledge on the basic properties and the limitations of the prominent plant proteases (papain, bromelain and ficin) are essential to allow comparison between the new proteases and the existing ones (Section 2.5.1 - 2.5.3)

Distinct and unique characteristics of the newly discovered proteases are usually highlighted and often compared to the properties of the existing commercial proteases. For example, in 2003, procerain, a cysteine protease from the latex of *Calotropis porcera* was discovered to be stable over a wide range of pH and temperature (Dubey & Jagannadham, 2003). This protease was able to withstand extreme pH ranging from pH 3-12 and able to retain its activity up to 70 ºC. Besides that, Tripathi et al. (2011) had discovered Streblin, a thermostable serine protease isolated from the latex of *Streblus asper*. This enzyme was able to recover 100 % of its activity at 15 - 85 ºC after 15 min and it had retain its complete activity in 40 % (v/v) ethanol, 45 % (v/v) acetonitrile and 70 % (v/v) dioxan. Meanwhile Kim et al (2007) had discovered proteases isolated from ginger rhizome that are able to hydrolyze native collagen, a property that is only discovered in microbial proteases. These enzymes, GP2 and GP3 are able to hydrolyze the native triple helical collagen and there was no report on similar properties in other plant proteases before.

2.5.1 Papain (EC 3.4.22.2)

Papain is the most extensively investigated cysteine protease extracted from the latex of the tropical tree of papaya (*Carica papaya*). The term papain referred both to the crude dried latex as well as to the crystalline proteolytic enzymes (Glazer & Smith, 1971). The production of crude papain reaches 500 tons per year and it is estimated to worth approximately USD 15 million (Rao, 1998). The molecular weight of papain is approximately 21 kDa (Glazer & Smith, 1971) and this enzyme is catalytically active in the pH range of 6.5 - 7.8 when tested against hemoglobin as the substrate. However the value slightly shifted when tested against other natural substrates such as casein (pH 7.5 - 8.0) and gelatin (pH 4.8 - 5.2) (Uhlig, 1998). Papain will be inactivated when exposed to oxidizing substances such as peroxides (H_2O_2) and oxygen (O_2) or exposed to heavy metal ions such as Hg^{2+}, Fe^{2+} and Cu^{2+}. The activity can be revived by various thiol compounds such as cysteine and glutathione (Kimmel & Smith, 1957) as well as reducing agents such as sodium borohydrate (Glazer & Smith, 1965) and 2,3-dimercaptopropanol (Kimmel & Smith, 1954). Papain is a nonspecific protease that is able to cleave general proteins, peptides, amides and esters. However the usage of synthetic substrate, benzoyl-L-*p*-nitroanilide (BAPA) (Uhlig, 1998) or benzoyl-L-argininamide (BAA) (Johansen & Ottesen, 1968) is used for analytical purposes.

Papain is stable at elevated temperatures when heated at the near neutral pH. The sturdiness of this enzyme was further investigated and it was discovered that this enzyme is stable when heated up to 70 ºC for 90 min. However papain was reported to be irreversibly inactivated when heated at elevated temperatures at acidic pH of less than pH 4 (Schwimmer, 1981). The enzyme was also reported to remain fully active when exposed to 9 M urea solution. Native papain were reported to maintain

its conformation when exposed to high concentrations of organic solvents such as methanol (0 - 50 %, v/v), ethylene glycol (0 - 40%, v/v) and dioxane (0 - 30 %, v/v) (Barel & Glazer, 1969) with only small changes were detected under spectroscopic studies.

Chemical modifications of papain stretched the enzyme's functionality and stability allowing wider applications of this protease in various fields. Khaparde and Singhal (2001) investigated the stability of succinic anhydride-modified-papain and native papain towards detergents. The succinylation of papain had shifted the optimum pH from 6 to 8 resulting to a higher specific activity of the modified papain in detergent liquor than that of native papain. Furthermore, the activity of the modified papain was comparable with the proteolytic activity of microbial proteases that are currently used in commercial detergents formulations. Such finding offers an inexpensive alternative to alkaline proteases that are presently used in detergents. In 2006, Sangeetha and Abraham carried out a similar study when they investigated the stability of modified papain in alkaline medium. The papain was chemically modified with dicarboxylic anhydrides of citraconic, pthalic and maleic acids.

Papain was known to catalyze the synthesis of amides, esters and macromolecules in a variety of reaction media such as aqueous buffers (Nitsawang et al., 2006), organic media (Stevenson & Storer., 1991; Teruhiko et al., 2006), biphasic organic-aqueous media (Fan et al., 2001; Szabo et al., 2009) and ionic liquids (Lou et al., 2006; Bian et al., 2012). The sturdiness of the enzyme is often tested in various reaction media in order to recognize the potential of the enzyme in numerous applications. Khmelnitsky and Rich (1999) summarized in their review that the use of organic solvent as media in enzyme reaction had increased the solubility of organic substrates. In relation to that, Roy et al., (2005) figured that the stability of

papain depends on the polarity of the medium when the highest stability of the enzyme was discovered in less polar medium such as ethanol. However, no stabilizing effect was observed with the usage of aprotic medium such as tetrahydrofuran (THF). Papain retained almost 100% of its activity after 24 h of incubation with acetonitrile (ACN) and a more compact structure was detected after the incubation. However the incubation in methanol caused 80% decreased in the activity and the active site was discovered to be altered and complete inactivation was discovered in media containing dihydrofuran (DHF). Llena-Suster et al. (2012) mentioned that the inclusion of organic solvents in an aqueous-organic system resulted to a more rigid and compact structure regardless of the organic: aqueous ratio tested.

The usage of immobilized enzymes results to a more cost effective production over the use of free enzymes in terms of reusability, rapid termination of reactions, controlled product formation and also the ease of enzyme removal from reaction mixture (Li et al., 2007). The extensive use of papain in pharmaceuticals and domestics industries, coupled with good stability of the enzyme had triggered the development of immobilizations of papain in recent years. Current advances on the immobilizations of papain are mostly focused on the application of various carriers and also on the use of variety immobilization techniques. For example, Wang et al. (2011) developed a protocol to prepare porous-cross-linked enzyme aggregates (p-CLEA) on papain where the capability of the immobilized enzyme to catalyze macromolecules such as bovine serum albumin (BSA) and ovalbumin was tested. The immobilized enzyme was discovered to possess great efficiency in catalyzing macromolecules when 93.5 % of the activity was obtained, relative to the activity of free papain. Meanwhile, Vasconcellos et al. (2011) investigated the use of chitosan

microparticles as carrier which allows the controlled release of papain. The promising outcome from the analysis increases the potential of this controlled-released papain to be developed as wound treatment in the future. However, these examples of papain immobilizations are non-exhaustive as numerous studies on various applications and techniques regarding this subject are continuously expanding (Li et al., 2007; Bayramoglu et al., 2011; Bian et al., 2012).

2.5.2 Bromelain (EC 3.4.22.4/5)

Any protease extracted from the family of Bromeliaceae is known as bromelain. However, the term bromelain is often associated with the protease that is extracted from the pineapple (*Ananas comosus*) either from the fruit (EC 3.4.22.4) or from the stem (EC 3.4.22.5) (Henicke, 1953). Bromelain belongs to the cysteine group thus it resembles papain in its mechanism of activation and inactivation. Recent statistics demonstrated that bromelain and papain made up almost 95 % of meat tenderizers sold in the United States (Amid et al., 2011). However researches in the development of bromelain as anti-cancer agents and therapeutic components are aggressively expanding (Tysnes et al., 2001; Stepek et al., 2004; Secor Jr. et al., 2009).

Although papain and bromelain share some common characteristics, the molecule of bromelain is slightly larger at 33 kDa (Ota et al., 1964). The catalytic activity of the enzyme is optimal at pH 6-8 with hemoglobin and casein as the substrate and at pH 5 when tested with gelatin. Meanwhile, the optimal working temperatures are ranged between 50 - 55 ºC for casein and 60 ºC for gelatin (Uhlig, 1998). A more specific study that emphasized on the mechanism of hydrolysis on different substrates was carried out by Corzo et al. (2012) in a series of

experiments. They discovered that the optimum working conditions for the hydrolysis of albumin occurred at 55ºC and pH 7.5, casein and sodium carbonate at 59 ºC and at pH 7.7 and pH 6.5 respectively, and hemoglobin was discovered to be optimally active at 37 ºC and pH 2.9. Denaturation occurred at low pH when the molecule structure of bromelain begins to unfold as the pH is lowered. Haq et al. (2002) observed the loss of secondary structure as the pH was lowered and the transition to the loss of structure begin to occur at pH 4.5. The far-UV-CD spectra however, indicated retention of some native-like secondary structure at pH 2.0 with further decrease of pH led to the formation of acid-unfolded state. This partially folded conformation revealed the concealed residues such as tryptophan where such condition can be taken advantage to enhance the ability to bind fluorescent probes to the enzyme. However, subsequent experiments discovered that the inclusion of salt to the partially folded structure resulted to the formation of a new folded structure somewhat different to the native enzyme. This salt-induced structure significantly reduced the exposure of the hydrophobic residues compared to the previous partially folded structure of bromelain at pH 2.0. This is due to the increase of the ionic strength which resulted to the refolding of the enzyme (Haq et al., 2005).

Meanwhile, Hale et al. (2005) discussed on the relationship between the concentration of the enzyme and the effect on the proteolytic activity. The concentrated bromelain solution (> 50 mg/ml) was discovered to be more resistant towards inactivation compared to the dilute solutions. In addition, the catalytic activity of the protease remained stable for at least 1 week when incubated in room temperature. The relatively stable concentrated bromelain as compared to the dilute solutions suggests that the consumption of the enzyme in concentrated form for *in*

vivo analysis is preferred to maximize the effect of the proteolytic activity (Hale et al., 2005).

Various therapeutic effects associated with bromelain urges the need to discover effective methods of purification and separation of the enzyme. Although the purification procedure had been well developed by Rowan et al. (1988), the purification strategies of bromelain has greatly evolved and documented through out the years ever since. For example, Yin et al. (2011) purified the crude pineapple fruit with high-speed counter-current chromatography (HSCCC) through a reversed micellar system. A large scale procedure of this method was carried out producing 3.01g of bromelain from 5.00 g of crude extract after 200 minutes run. Meanwhile, Kumar and Hemavathi (2011) used a similar chromatography method with concavalin A (Con A) was used as an affinity ligand. The optimized conditions of the extraction resulted to an enzyme preparation being purified by 12.32 fold with the yield of 185.6%. The efficiency of drying methods were analyzed based on two common procedures which include the spray drying and the freeze drying method (Devakate, 2009). It was observed that the freeze drying method was more efficient with almost 96 % of the activity was still intact compared to the spray drying method which retained only 50-70% of the activity.

The huge potential of bromelain in medical industries had triggered the effort of cloning the genes that expressed the desired characteristics. Bulk quantities of the enzyme required either in *in vivo* analysis or to be used as a final product makes it rather tedious to purify the enzyme direct from the plant. The cloning of the gene into suitable hosts such as *Escherichia coli* allows a more systematic and controllable production of the enzyme. Amid et al. (2011) cloned the gene encoding stem bromelain from *A. comosus* into BL21-AI *E. coli* when positive outcome was

observed when the DNA sequencing the amplified gene displayed a high level of homology with the corresponding gene of bromelain from the database. The purified recombinant protein was discovered to exhibit optimum activity at pH 4.6 and at 45º C.

2.5.3 Ficin (EC 3.4.22.3)

The name ficin or ficain is used to describe the protease from the crude latex from different species of *Ficus* and it is can also be used to refer crystalline proteases prepared from the plants' latex (Glazer & Smith, 1971). The genus *Ficus* belongs to the same family of Moraceae with *Artocarpus*, however they comprised over 2000 species of tropical and sub tropical trees, making them one of the largest genus in the Moraceae family. The most extensively investigated ficins are the cysteine proteases extracted from the latex of *F. glabrata* and *F. carica* (Devaraj et al., 2008). Jones and Glazer (1970) separated the crude latex enzyme into five isoforms with the molecular weight that ranged from 25 - 26 kDa. Although cysteine proteases discovered in these two plants were discovered to exist in several forms (Kramer & Whitaker, 1964; Englund et al., 1968), the effort to isolate and purify the individual forms was only carried out in 2011 by Azarkan and his research team. The five ficin isoforms which were fractionated with SP-Sepharose chromatography column were chemically modified with monomethoxypolyethylene glycol. The modification allowed enhanced purification of these isoforms when five highly purified fractions were obtained from the procedure. These isoforms exhibited different specific amidase activities against synthetic substrates suggesting that there might be some differences in the active sites among the cysteine proteases. Meanwhile, ficin purified from

Shandong fig trees was discovered to have a half life of more than 1 hour at 65ºC and the Michaelis constant (K_m) of this protease was determined to be 1.56 mg/ml (Huang et al., 2008).

The optimum pH for ficin's catalytic activity is at pH 7.5 when tested with liquefaction of gelatin and pH 4 - 7 for the hydrolysis of elastin. They also have a reasonably good stability over the pH range of 3.5 - 9.0 (Whitaker, 1957). The resistance of the enzyme molecule from unfolding was also investigated through denaturation by urea and guanidine hydrochloride (GuHCl) (Devaraj et al., 2011). The molecule of ficin was resistant towards unfolding by urea at neutral pH. However, as the pH was lowered, the protein became susceptible and begins to unfold. Meanwhile, ficin lost its molecule structure in the presence of 4M GuHCl even in neutral conditions.

The study on the energy of activation (E_a) of ficin under high hydrostatic pressure decreased discovered that the E_a decreased from 139 to 43 kJ/mol (Katsaros et al., 2009). The investigation also revealed that the inactivation rates of the enzyme at higher pressures were higher than the corresponding thermal rates at ambient pressure. In relation to that, Katsaros et al. (2009) concluded that ficin demonstrated a high thermal and pressure stability suitable for intense applications and process conditions.

2.5.4 Occurrence of plant proteases and their physiological functions

Proteases have multiple roles in regulating a plant's physiological functions which cover the structure of a plant, the interactions with the environment, the biochemistry processes, the cell and also the molecular biology of the plant. Proteases also occur in various organs, such as in leaves (Huangpu & Graham, 1995;

Popovic et al., 1998), seeds (Mohamed Ahmad et al., 2009), fruits (Wang et al., 2008), latex (Dubey & Jagannadham, 2003), rhizomes (Hashim, 2011) and the full list of examples can be viewed from the MEROPS site (http://merops.sanger.ac.uk). A study by Bah et al. (2006) on the characterizations of cysteine protease from 10 Malian medicinal plants had discovered that the aerial section of the plants showed the highest proteolytic activity compared to the other parts. Some of the roles of protease in plants are listed below based on a review by Antao and Malcata (2005).

2.5.4.1 Microsporogenesis

Microsporogenesis refers to the formation of four haploid microspores through meiosis. This process occurred in angiosperm for the formation of microsporocyte that will eventually develop into pollen that is the male gametophyte. The elucidation of the structure of microsporocyte led to the discovery of amino acid sequence, termed LIM9 that is similar to the serine protease from *Bacillus* sp. (Kobayashi et. al., 1994). The N-terminal sequencing of the microsporocyte extract of *Lilium longiforum* cv. Echantment bulb by Taylor et al. (1997) had also discovered a similar protein which shared the same homology as LIM9. Taylor et al., (1997) extended their discussion of the role of LIM9 on the anther development and also the possible role of this protein as the developmental regulator during microsporogenesis. LIM9 assists the secretion of β-1,3-glucanase during the destruction of the cell wall after the completion of meiosis of the microspores. β-1,3-glucanase is secreted into the locule of the plant to assist in the liberation of the microspores (Steiglitz & Stern, 1973; Engleke et al., 2002; Popova et al., 2008). The LIM9 is secreted during this event to create gaps in the microsporocyte cell wall and