ISOLATION AND CHARACTERIZATION OF PROTEIN FRACTIONS FROM CHICKPEA (*Cicer arietinum* L.) AND OAT (*Avena sativa* L.) SEEDS USING PROTEOMIC TECHNIQUES

By

Yu-Wei Chang

Department of Food Science and Agriculture Chemistry
Macdonald Campus, McGill University, Montreal

July, 2010

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

Suggested short title:

CHARACTERIZATION OF CHICKPEA AND OAT PROTEINS

Dedicated to my beloved wife, Hsiao-Chien, to my adorable daughters, Anya and Miya, and to my father and mother, Shin-Lieh and Shen.

ABSTRACT

Chickpea (*Cicer arietinum* L.) and oat (*Avena sativa* L.) seeds are important sources of protein ingredients with potential nutritional, functional and bioactive properties. Protein fractions were prepared from chickpea and oat using sequential extractions with distilled water (albumins), NaCl solution (globulins) and NaOH solution (glutelins), respectively. Molecular characteristics of individual protein fractions were investigated using polyacrylamide gel electrophoresis (Native- and SDS-PAGE, and 2-DGE) in combination with reversed-phase high performance liquid chromatography (RP-HPLC). Tryptic peptide sequences were identified using proteomic techniques including 1D trypsin in-gel digestion, liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS) analysis and Mascot MS/MS ion search. Sequence similarity and potential bioactivity of proteins were examined using BLAST and BIOPEP analysis, respectively.

Native-PAGE results showed chickpea and oat globulin fractions (C-Gb and O-Gb) contained proteins corresponding to legumin (11S) and avenalin (12S), respectively. SDS-PAGE revealed that chickpea albumin and globulin fractions (C-Ab and C-Gb) showed protein bands with MWs related to legumin (11S) and pea vicilin (7S) while chickpea glutelin fraction (C-Gt) showed protein bands with MWs related to rice glutelin; oat protein fractions (O-Ab, O-Gb and O-Gt) showed protein bands with MWs related to oat 12S globulin (avenalin). α- and β-subunits of globulin and glutelin fractions from chickpea and oat were identified with estimated MWs ranging from 31 to 45 kDa and from 21 to 31 kDa, respectively. *In vitro* chickpea albumin, globulin and glutelin

hydrolysates showed DH of 22.8%, 28.6% and 28.8%, respectively; SDS-PAGE revealed that legumin α- and β-subunits from chickpea globulin fraction (C-Gb) were hydrolyzed.

The identified tryptic peptides from chickpea and oat protein fractions showed sequence homology that corresponded to chickpea legumin α- and β-subunit (NCBI accession number: gi|6273402; theoretical mass 56,216 Da), chickpea provicilin precursor (NCBI accession number: gi|82173888; theoretical mass 51,390 Da) and 12S seed storage globulin 1 (NCBI accession number: gi|134918; theoretical mass 58,508 Da). Chickpea and oat glutelin fractions (C-Gt and O-Gt, respectively) exhibited large sequences homology with origin to chickpea legumin and oat 12S globulin. BLAST analysis of sequence alignments gave approximately 30% similarity of AAs sequences from chickpea legumin, oat 12S globulin 1 and rice glutelin precursor; the oat 12S globulin 1 and rice glutelin precursor exhibited relatively high sequence homology of 63%. *In silico* BIOPEP results showed that chickpea legumin and provicilin precursor contained 177 and 133 potential ACE-inhibitory peptides within their primary sequences, respectively; ficain and proteinase K were the suggested proteases that could theoretically release greater numbers of predicted ACE-inhibitory peptides (34 and 35, respectively) from these two proteins.

RÉSUMÉ

Les semences du pois chiche (*Cicer arietinum* L.) et de l'avoine (*Avena sativa* L.) sont d'importantes sources d'ingrédients protéiques dont les propriétés nutritionnelles, fonctionnelles et bioactives démontrent un grand potentiel. Les fractions protéiques ont été préparées à partir du pois chiche et de l'avoine par extraction séquentielle avec de l'eau distillée (albumine), une solution de NaCl (globuline) et une solution de NaOH (glutelines), respectivement. Les caractéristiques moléculaires des fractions de protéines individuelles ont été examinées par électrophorèse en gel de polyacrylamide (non dénaturante et SDS-PAGE, et 2-DGE) en combinaison avec la chromatographie en phase liquide à haute performance en phase inversée. Les séquences de peptides tryptiques ont été identifiées par des techniques protéomiques telles que la digestion de trypsine en gel unidimensionnelle, l'analyse chromatographique en phase liquide couplée à la spectrométrie de masse en tandem avec ionisation de type électrospray (LC-ESI-MS/MS), et la recherche d'ions MS/MS (Mascot). Les similarités séquentielles et la bioactivité potentielle des protéines ont été examinées sous analyse par BLAST et BIOPEP, respectivement.

Les résultats de l'électrophorèse non dénaturante en gel de polyacrylamide démontrent que les fractions de globulines du pois chiche et d'avoine (C-Gb et O-Gb) contiennent des protéines correspondant aux légumines (11S) et avenaline (12S), respectivement. La SDS-PAGE révèle que les fractions d'albumine et de globuline de pois chiche (C-Ab et C-Gb) montrent des bandes protéiques ayant des poids moléculaires reliés à la légumine (11S) et le viciline de pois (7S) alors que la fraction de glutéline de pois chiche (C-Gt) montre des bandes protéiques avec des poids moléculaires reliés à la

glutéline de riz; les fractions protéiques d'avoine (O-Ab, O-Gb et O-Gt) montrent de bandes protéiques avec des poids moléculaires reliés à la globuline d'avoine 12S (avenaline). Les sous-unités α- et β- des fractions de globuline et glutéline du pois chiche et de l'avoine des poids moléculaires estimés entre 31 à 45 KDa et entre 21 à 31 KDa, respectivement. L'albumine *in vitro* de pois chiche, les hydrolysates de globuline et de glutéline montrent un DH de 22.8%, 28.6% et 28.8%, respectivement; LA SDS-PAGE a révèle que les sous-unités α-légumine et β-légumine des fractions de la globuline de pois chiche (C-Gb) ont été hydrolysées.

Les peptides tryptiques identifiés des fractions protéiques de pois chiche et d'avoine montrent une homologie de séquence qui correspond aux sous-unités α- et β- de la légumine de pois chiche (numéro d'enregistrement de NCBI: gi|6273402; masse théorique 56,216 Da), au précurseur de proviciline de pois chiche (numéro d'enregistrement de NCBI: gi|6273402; masse théorique 56,216 Da) et au 12S semence de globuline de stockage 1 (numéro d'enregistrement de NCBI: gi|134918; masse théorique 58,508 Da). Les fractions de glutéline du pois chiche et d'avoine (C-Gt et O-Gt, respectivement) démontrent une grande homologie de séquences avec l'origine de pois chiche et de la globuline 12S de l'avoine. L'analyse BLAST d'alignement de séquences donne approximativement 30% de similitude avec les alignements d'acides aminés de la légumine de pois chiche, de 12S globuline 1 d'avoine et du précurseur de glutéline de riz; la 12S globuline 1 d'avoine et le précurseur de glutéline de riz démontre une homologie de séquence relativement élevée de 63%. Les résultats de BIOPEP *in silico* montrent que la légumine de pois chiche et le précurseur de proviciline contiennent 177 et 133 peptides inhibiteurs de l'ACE dans leurs séquences primaires, respectivement; ficain et protéinase

K sont suggérés comme étant les protéases pouvant théoriquement libérer le plus grand nombre prévu de peptides inhibiteurs de l'ACE (34 et 35, respectivement) de ces deux protéines.

ACKNOWLEDGEMENTS

I am heartily grateful to my supervisor, Dr. Inteaz Alli, whose patience, encouragement, support and advice directed me towards the completion of this thesis; an extended thank is expressed to his family members for their friendship as well. However it would not have been possible without the help from other people who contributed their effort to this work.

I would like to thank Dr. Yasuo Konishi for the use of LC-ESI-MSMS machine located in the Biotechnology Research Institute, and Dr. Edmund Ziomek and Ms. Beata Usakiewicz for their continuous assistance, technical support of gaining MS-related knowledge and invaluable friendship.

I would like to thank Dr. Selim Kermasha and Dr. Salwa Karboune for kindly allowing me to use their facilities.

I would like to thank my colleagues, Dr. Anwer Ayad, Dr. Lam Chu, Dr. Mario Rivero-Huguet, Lotthida Inthanavong and Sun-Hung Park for their friendship.

I would like to thank our department secretaries, Leslie Ann LaDuke and Diane Chan-Hum, and lab technician, Eby Noroozi for their support and friendship.

I would like to thank my parents and parents-in-law for their moral and financial support, and continuous encouragement throughout this study.

The profoundest gratitude goes to my adored wife Hsiao-Chien; thank you so much for giving us two lovely girls, Anya and Miya to enrich our life and for supporting me "almost everything". I love you all.

CLAIMS TO ORIGINAL RESEARCH

1. This is the first comparative study on molecular characteristics of chickpea and oat sequential protein fractions (albumins, globulins and glutelins) using electrophoresis (Native- and SDS-PAGE) in combination with RP-HPLC chromatography; it is also the first study to characterize chickpea and oat protein fractions (globulins and glutelins) using 2-DGE electrophoresis.

2. This is the first study to identify the tryptic peptide sequences from in-gel digests (1D SDS-PAGE gel) of chickpea, oat and rice protein fractions using the proteomic techniques (trypsin in-gel digestion, LC-ESI-MS/MS analysis, MassLynx and ProteinLynx softwares, and Mascot MS/MS ion search).

3. This is the first study to investigate the sequence similarities of chickpea legumin, oat 12S globulin 1 and rice glutelin precursor protiens using BLAST (NCBI) tool.

4. This is the first study to report the predicted ACE inhibitory peptides embedded within chickpea legumin and provicilin precursor proteins sequences using BIOPEP analysis.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER 2

CHAPTER 3

CHAPTER 4

ABBREVIATIONS

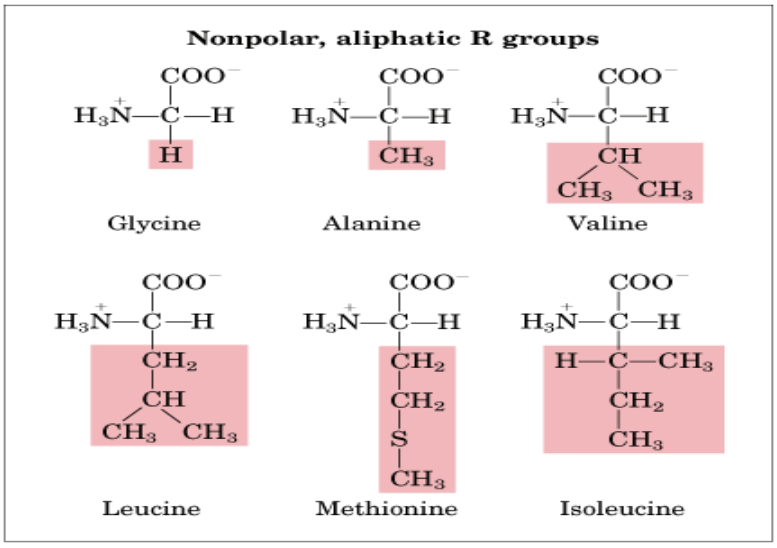
AA	Amino Acid
ACE	Angiotensin-I-Converting Enzyme
BIOPEP	Bioactive Peptides
BLAST	Basic Local Alignment Search Tool
BSA	Bovine Serum Albumin
C-Ab	Chickpea Albumin Fraction
C-Gb	Chickpea Globulin Fraction
C-Gt	Chickpea Glutelin Fraction
2-DGE	Two-Dimensional Gel Electrophoresis
DH	Degree of Hydrolysis
EMBL	European Molecular Biology Laboratory
EST	Express-Sequence-Tag
ExPASy	Expert Protein Analysis System
FTICR	Fourier Transform Ion-Cyclotron Resonance
GM	Genetic Modified
IEF	Isoelectric Focusing
IPG	Immobilized pH Gradient
IT	Ion Traps
kDa	Kilodalton
LC-ESI-MS/MS	Liquid Chromatography Electrospray Ionization Tandem Mass Sepctrometry
MALDI-TOF	Matrix-Assisted Laser Desorption-Ionization Mass Spectrometry Time of Flight
MW	Molecular Weight
m/z	Mass-to-Charge Ratio
NCBI	National Center for Biotechnology Information
NMR	Nuclear Magnetic Resonance
O-Ab	Oat Albumin Fraction
O-Gb	Oat Globulin Fraction

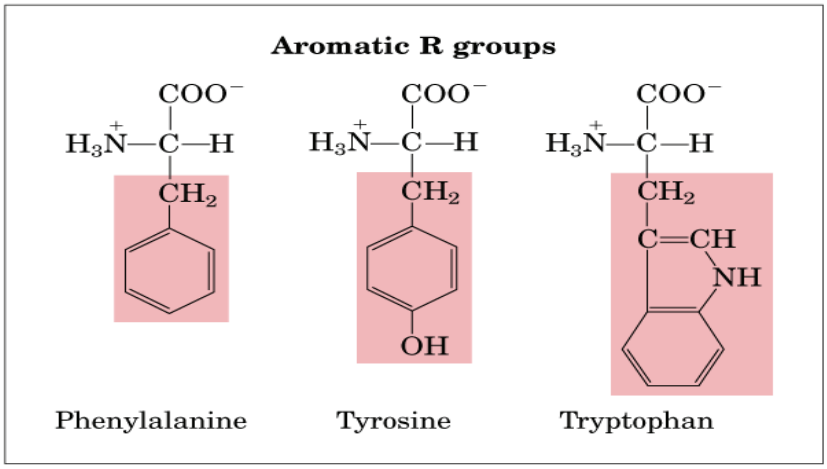
O-Gt	Oat Glutelin Fraction
PAGE	Polyacrylamide Gel Electrophoresis
PDB	Protein Data Bank
PIR	Protein Information Resource
QTOF	Quardrupole-Time-Of-Flight
RP-HPLC	Reverse phase high performance liquid chromatography
Rt	Retention time
R-Gt	Rice Glutelin Fraction
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
XRC	X-Ray Crystallography

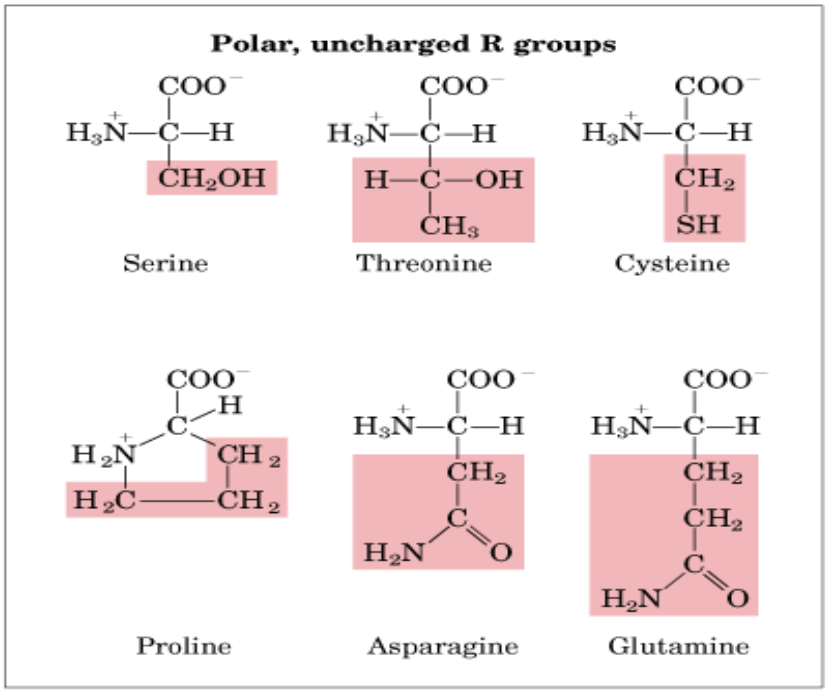
LIST OF AMINO ACIDS ABBREVIATION AND THEIR STRUCTURES

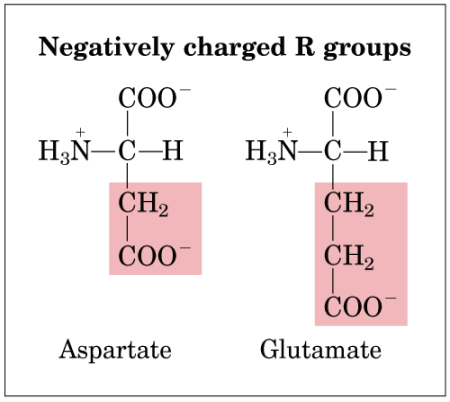
Amino acid	Three Letter Code	Single Letter Code (MW)[a]
Alanine	Ala	A (89.09)
Arginine	Arg	R (174.20)
Asparagine	Asn	N (132.11)
Aspartic Acid	Asp	D (133.10)
Cysteine	Cys	C (121.15)
Glutamic Acid	Glu	E (147.12)
Glutamine	Gln	Q (146.14)
Glycine	Gly	G (75.06)
Histidine	His	H (155.15)
Isoleucine (EAA)	Ile	I (131.17)
Leucine (EAA)	Leu	L (131.17)
Lysine (EAA)	Lys	K (146.18)
Methionine (EAA)	Met	M (149.21)
Phenylalanine (EAA)	Phe	F (165.19)
Proline (EAA)	Pro	P (115.13)
Serine	Ser	S (105.09)
Threonine	Thr	T (119.11)
Tryptophan (EAA)	Trp	W (204.22)
Tyrosine	Tyr	Y (181.18)
Valine (EAA)	Val	V (117.14)

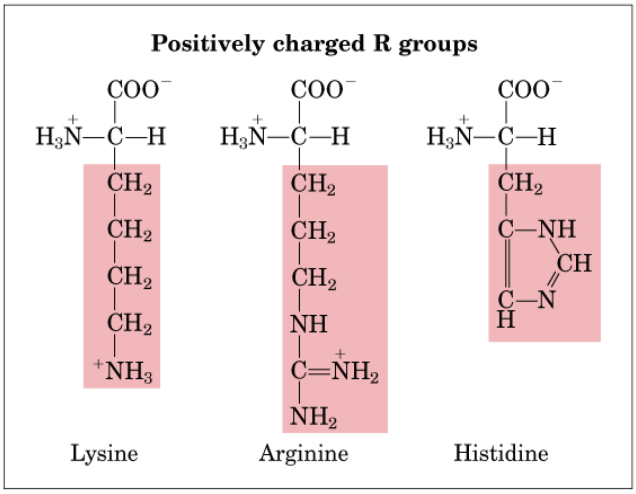
EAA: Essential Amino Acid; [a] (average monoisotopic mass in Da)




Aromatic R groups
Phenylalanine
Tyrosine
Tryptophan


Polar, uncharged R groups
Serine
Threonine
Cysteine
Proline
Asparagine
Glutamine


Negatively charged R groups
Aspartate
Glutamate


Positively charged R groups
Lysine
Arginine
Histidine

CHAPTER 1

INTRODUCTION

1.1 General Introduction

Chickpea (*Cicer arietinum* L.) and oat (*Avena sativa* L.) are important crops in Canada with the production of 0.22 and 4.6 million tonnes, respectively (FAOSTAT, 2007). Because of their relatively high protein contents (14.9-24.6% for chickpea, 12.4-24.5% for oat), both seeds can be good sources of protein concentrates for food applications (Duke, 1981; Lásztity, 1996). Chickpea is one of the few legumes containing a glutelin (cereal-like protein) whereas oat is unique among cereals with a relatively high amount of globulin (legume-like protein) as the major storage protein and a minor component of glutelins (Kulp & Ponte, 2000; Lásztity, 1996); the globulin proteins of both chickpea and oat have been studied by previous researchers (Kaur & Singh, 2007; Robert et al., 1985; Shotwell et al., 1988; Singh & Jambunathan, 1982; Yust et al., 2003). However, there is only little information on the molecular characterization of albumin and glutelin, including peptide sequences, amino acid sequences, biological properties and functional properties. Furthermore, oat is genetically regarded as the intermediate crop between legume and cereal (Robert et al., 1985; Zhao et al., 1983). Therefore, it is of great interest to study the relationship of molecular characteristics of oat globulins and glutelins in comparison to chickpea. Glutelins are present mainly in cereals. In rice, glutelins are the major protein and comprise up to 80% of total protein content (Villareal & Juliano, 1978).

Currently, proteomics offers novel technologies and empowers an immense potential for commercial exploitation such as the identification of new diagnostic marker,

therapeutic drugs and vaccine targets in connection with medical and clinical industrial (Carbonaro, 2004); it allows rapid study of the partial peptide sequences of proteins instead of tedious sequencing procedure. The main aspect of this novel approach is the use of the combined techniques including 2D gel, enzymatic in-gel digestion, mass spectrometry, and internet database search to identify the peptides or proteins. With respect to plant seed research, proteomics approach has been applied for studying wheat kernel (Amiour et al., 2002), sesame seed (Beyer et al., 2002b), hazelnut (Beyer et al., 2002c), soybean (Gianazza et al., 2003), and green coffee bean (Maria et al., 2005) over the past decade. The concepts and techniques of proteomics in relation to food proteins have resulted in use of the term "food proteomics" with potential applications to food areas such as food quality control (Kjærsgård et al., 2006; Martinez & Friis, 2004), food allergy (Beyer et al., 2002a; Natale et al., 2004), food safety (Kuiper et al., 2002) and bioavailability of food proteins (Carbonaro, 2004).

1.2 Rationale and Objectives of Study

Chickpea and oat globulins share some similarities of their glutelin proteins; however information on molecular characteristics of their glutelins is limited. In this research, glutelins and globulins from chickpea and oat protein isolates were characterized and compared using proteomic approaches.

The overall objectives were to study molecular characteristics of albumin, globulin and glutelin proteins from chickpea and oat seeds using proteomic techniques. The specific objectives of this research were:

1. To isolate albumin, globulin and glutelin proteins from chickpea and oat seeds by sequential extractions.
2. To determine the molecular characteristics of isolated albumin, globulin and glutelin proteins using native, sodium dodecyl sulfate and 2-dimensional gel electrophoresis (Native-, SDS-PAGE and 2-DGE), and reversed-phase high performance liquid chromatography (RP-HPLC).
3. To perform enzymatic digestion for chickpea isolates with digestive enzyme (trypsin) to evaluate the level of protein digestion.
4. To characterize protein fractions and identify their peptides sequences using in-gel enzymatic digestion and tandem mass spectrometry (LC-ESI/MS/MS), followed by database search on the "ExPASy" protein sever (http://ca.expasy.org/) using Mascot software and to study the homology of chickpea and oat protein fractions by aligning their proteins sequences using BLAST analysis.
5. To investigate the predicted biological properties of identified peptides from chickpea fractions *in silico* using the bioactive peptides (BIOPEP) database.

CHAPTER 2

LITERATURE REVIEW

2.1 General Characterization of Chickpea and Oat Proteins

The protein content of chickpea seeds ranges from 14.9 to 24.6% with an average of 21.5% (Duke, 1981). According to Osborne's classification, the storage proteins of chickpea seed have been fractionated into globulin (salt soluble; 56%), albumin (water soluble; 12%), prolamin (alcohol soluble; 2.8%), glutelin (acid/alkali soluble; 18.1%), and residual proteins (Chavan et al., 1989; Singh & Jambunathan, 1982). Oat groats possess the highest protein content (12.4 to 24.5%) among cereals (Lásztity, 1996). Similar to chickpea protein fractions, oat proteins are classified to four classic fractions, albumins (10-20%), globulins (12-55%), prolamins (12-14%) and glutelins (23-54%) (Haard, 1999). Table 2.1 shows the protein distribution from legume (soybean and chickpea) and cereal (oat and rice) seeds. The major storage proteins of soybean and rice are the globulins and glutelins; chickpea and oat seeds share similar protein distribution with globulins as the major storage proteins followed by glutelins, albumins and prolamins (Haard, 1999; Lásztity, 1996; Singh & Jambunathan, 1982; Steenson & Sathe, 1995; Wolf, 1970). The amino acid content of chickpea and oat proteins is shown in Table 2.2; chickpea and oat albumins have the highest lysine content; both globulins are relatively deficient in methionine and cysteine while their albumins and glutelins contain slightly higher amount of those amino acids (Lásztity, 1996; Singh & Jambunathan, 1982).

Table 2.1: Comparison of protein fractions distribution from legume (soybean and chickpea) and cereal (oat and rice) seeds (Haard, 1999; Singh & Jambunathan, 1982; Steenson & Sathe, 1995; Wolf, 1970)

Protein fractions	Soybean	Chickpea	Oat	Rice
Globulin (Salt-soluble)	78%	56%	12-55%	13.8%
Glutelin (Acid/alkali-soluble)	0%	18.1%	23-54%	74.5%
Albumin (Water-soluble)	22%	12%	10-20%	5.9%
Prolamin (Alcohol-soluble)	0%	2.8%	12-14%	5.8%

2.2 Molecular Characteristic of Chickpea and Oat Protein Fractions

2.2.1 Globulin Proteins

Globulins represent about 50% of chickpea seed proteins and are composed of two major groups, characterized by their sedimentation coefficients, the 11S (legumin) and the 7S (vicilin) proteins (Casey et al., 1993). Unlike other cereals, oat is the only cereal that the predominant storage protein is soluble in salt solution and characterized as globulin. The oat globulin fraction is also called “avenalin”; the protein proportion (12-55%) of oat globulin may vary differently depending on the cultivars. Oat globulins are a mixture of different polypeptides identified and separated as 12S (the major component), 7S and 3S proteins from the oat grain extracted by 1 M NaCl (Burgess et al., 1983).

Table 2.2: Average amino acid composition of isolated protein fractions from chickpea and oat seeds (Lásztity, 1996; Singh & Jambunathan, 1982)

Amino acid	Albumin		Globulin		Glutelin		Prolamin	
	C	O	C	O	C	O	C	O
Lysine	10.8	8.3	6.4	4.9	6.8	3.1	2.6	5.2
Histidine	2.3	3.1	2.6	3.3	2.9	1.6	2.6	3.1
Arginine	5.6	5.3	10.7	8.5	6.8	5.0	4.8	9.1
Aspartic acid	13.8	12.5	12.7	9.8	10.1	4.2	10.3	10.8
Threonine	5.4	5.5	3.5	3.8	5.7	2.3	2.2	4.8
Serine	5.2	6.3	5.2	4.3	5.6	3.2	1.9	4.7
Glutamic acid	18.4	15.1	15.2	19.5	16.6	36.1	17.7	19.1
Proline	-	5.8	-	5.4	-	11.3	-	8.1
Glycine	5.4	6.5	3.7	5.8	4.7	3.0	3.1	4.0
Alanine	5.3	7.1	4.3	6.5	4.9	3.8	2.3	4.1
Cysteine	3.5	1.7	1.0	1.8	1.4	3.1	0.6	1.4
Valine	4.5	7.3	4.2	5.1	5.7	5.5	2.1	4.4
Methionine	1.8	2.2	0.8	1.9	1.2	3.3	0.9	1.5
Isoleucine	5.1	4.1	4.4	4.7	5.4	3.8	2.3	4.3
Leucine	9.8	8.9	7.5	7.0	9.1	10.3	1.6	7.1
Tyrosine	4.2	2.6	2.9	2.4	3.7	2.1	2.3	4.8
Phenylalanine	5.1	7.8	6.1	6.1	4.4	7.5	3.4	7.1
Tryptophan	-	1.7	-	1.4	-	1.6	-	1.9

Data were in determined in g amino acid/100 g protein; C: chickpea, O: oat.

2.2.1.1 11S and 12S Globulins

Chickpea 11S (legumin) and oat 12S globulin are the major globulin; structurally, they are oligomeric proteins with six quaternary monomer units. The molecular weight (MW) of chickpea legumin is about 320-400 kDa (Casey et al., 1993) whereas the oat 12S globulin was first characterized with a MW of 322 kDa (Peterson, 1978). Both 11S and 12S globulins are heaxmers consisting of six 54-60 kDa subunits. Figure 2.1 represents the theoretical structure of a globulin hexamer. Under reducing condition in SDS-PAGE separation, the individual subunit, a monomer in native form, gives α- and β-subunits with MWs of about 32 and 22 kDa; these two subunits are connected by a single disulfide bond (Lásztity, 1996; Sánchez-Vioque et al., 1999).

Three chickpea globulin proteins have been sequenced by genomic approach (Table 2.3; UniProt Knowledgebase of ExPASy Proteomics Server available at http://ca.expasy.org/ (Gasteiger et al., 2003)). Mandaokar and Koundal (1997) studied the nucleotide sequence of chickpea legumin and submitted the complete amino acids (AAs) sequence (Figure 2.2) to EMBL/GenBank databases; despite the function of being a storage protein, this legumin (UniProtKB/Swiss-Prot ID: Q9SMJ4) is also considered a novel amylase inhibitor (Hao et al., 2009). The sequenced chickpea legumin is composed of a signal peptide at the AA position 1-21, an acidic polypeptide chain (known as α-subunit) at AA position 22-311, and a basic polypeptide chain (known as β-subunit) at AA position 312-496; a disulfide bond is observed at AA position 31↔64 and a interchain (between α- and β-chain) is observed at AA position 107↔318 (Mandaokar & Koundal, 1997).

Several primary structures of 11S and 12S oat globulins have been sequenced. Table

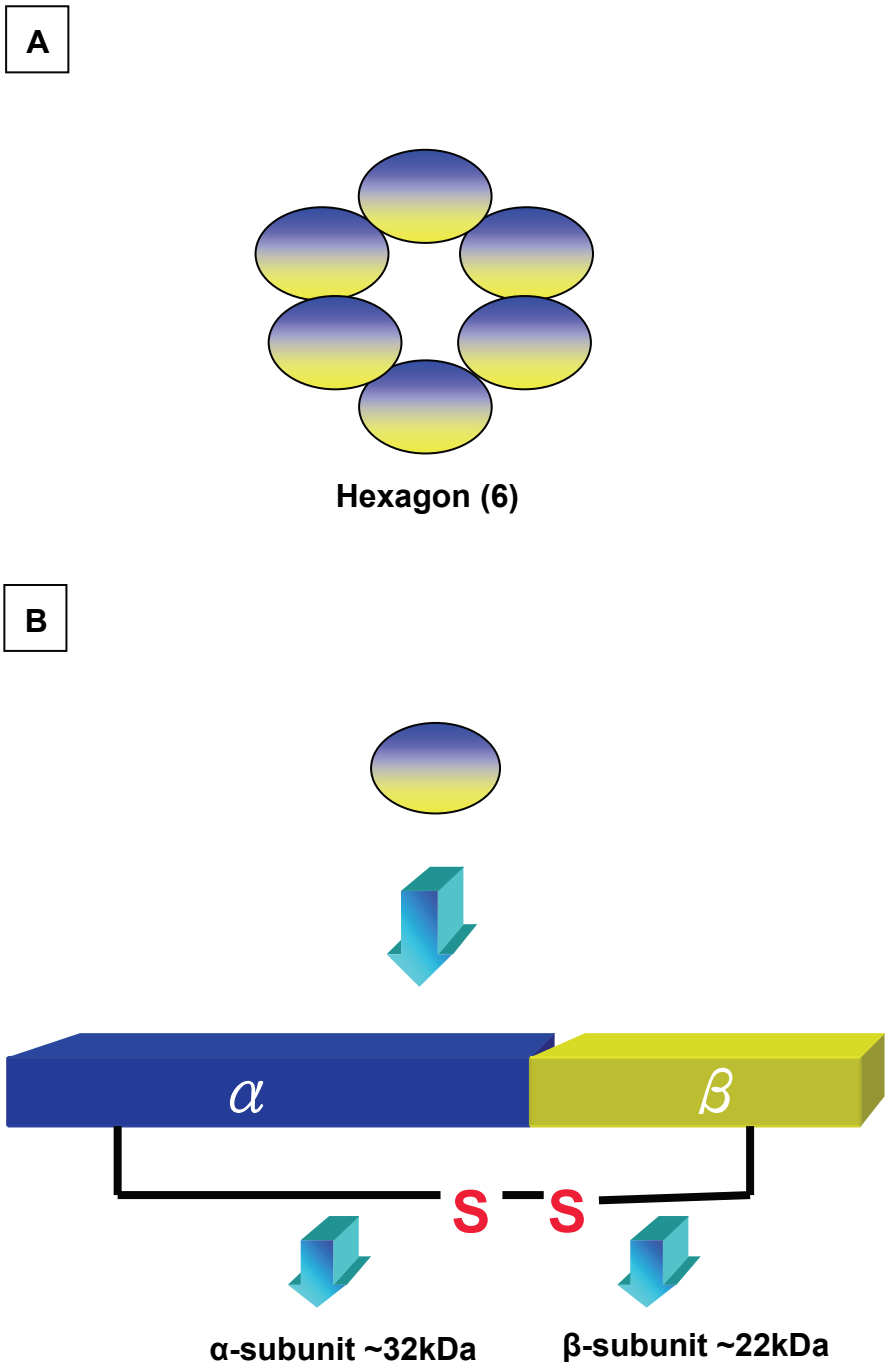


Figure 2.1: (A) theoretical structure of a hexamer from globulin and glutelin proteins, (B) their α- and β-subunits of an individual monomer (Rehm, 2006; Sánchez-Vioque et al., 1999)

Table 2.3: Summary of sequenced chickpea globulin proteins available in UniProt Knowledgebase (UniProtKB) of ExPASy Proteomics Server (http://ca.expasy.org/) (Gasteiger et al., 2003)

Accession number	Protein name	Amino acid length	Reference
Q9SMJ4	Legumin	496	Mandaoka & Koundal, 1997
O04219	Legumin	132	Koundal & Shasany, 1997
Q304D4	Provicilin	453	Yasothornsrikul et al., 1990

```
        10         20         30         40         50         60
MAKLLALSLS FCFLLFGTCF ALRDQPQQNE CQLEHLNALK PDNRIKSEGG LIETWNPSNK
        70         80         90        100        110        120
QFACAGVALS RATLQPNSLL QTFLHQRSPE IFIQQGNGYF GMVFPGCVET FEEPRESEQG
       130        140        150        160        170        180
EGSKFSDSHQ KVNRFREGDI IAVPTGVVFW MFNDQDTPVI AVSLIDTSSF QNQLDQMPRR
       190        200        210        220        230        240
FYLAGNHEQE FLRYQQEGSE EEENEGGNIF SGFKRDFLED ALNVNRRIVN KLQGRNEDEE
       250        260        270        280        290        300
KGAIVKVKGG LSITTPPEKE PRQKRGSRQE EDEDEDEKRQ PHRHSRQDED EDEKRQPHHH
       310        320        330        340        350        360
SRGGSKSQRD NGFEETICTA RLHQNIGSSS SPDIYNPQAG RIKTVTSFDL QALRFLKLSA
       370        380        390        400        410        420
EFGSLHKNAM FVPHYNLNAN SILYALKGRA RLLYALNCKG NSVFDGELEA GRALIVPQNF
       430        440        450        460        470        480
AIAAKSLSDR FSYVAFKTND RALINVCQKK LLQLLSIWKE MRPGSSSSTA PFHFLFHPAV
       490
TQTTKQQLDL VPNQYE
```

Figure 2.2: Amino acids sequence of legumin (UniProtKB/Swiss-Prot ID: Q9SMJ4) from chickpea seeds (Mandaokar & Koundal, 1997)

Table 2.4: Summary of sequenced 11S and 12S oat proteins available in UniProt Knowledgebase (UniProtKB) of ExPASy Proteomics Server (http://ca.expasy.org/) (Gasteiger et al., 2003)

Accession number	Protein name	Amino acid length	Reference
P14812	12S seed storage globulin 2	518	Schubert et al., 1990
P12615	12S seed storage globulin 1	518	Shotwell et al., 1988
Q38781	Oat storage protein 12S globulin	313	Walburg & Larkins, 1986
Q38780	11S globulin	527	Tanchak et al., 1995
Q38779	11S globulin	551	Tanchak et al., 1995

```
        10         20         30         40         50         60
MATTRFPSLL FYSYIFLLCN GSMAQLFGQS FTPWQSSRQG GLRGCRFDRL QAFEPLRQVR
        70         80         90        100        110        120
SQAGIIEYFD EQNEQFRCAG VSVIRRVIEP QGLLLPQYHN APGLVYILQG RGFTGLTFPG
       130        140        150        160        170        180
CPATFQQQFQ PFDQAQFAEG QSQSQNLKDE HQRVHHIKQG DVVALPAGIV HWCYNDGDAP
       190        200        210        220        230        240
IVAVYVFDVN NNANQLEPRQ KEFLLAGNNK REQQFGQNIF SGFSVQLLSE ALGISQQVAQ
       250        260        270        280        290        300
KIQSQNDQRG EIIRVSQGLQ FLKPFVSQQG PVEHQAYQPI QSQEEQSTQY QVGQSPQYQE
       310        320        330        340        350        360
GQSTQYQPGQ SWDQSFNGLE ENFCSLEARQ NIENPKRADT YNPRAGRITH LNSKNFPTLN
       370        380        390        400        410        420
LVQMSATRVN LYQNAILSPY WNINAHSVMH MIQGRARVQV VNNHGQTVFN DILRRGQLLI
       430        440        450        460        470        480
IPQHYVVLKK AEREGCQYIS FKTNPNSMVS QIAGKTSILR ALPVDVLANA YRISRQEAQN
       490        500        510
LKNNRGEEFD AFTPKFTQTG SQSYQDEGES SSTEKASE
```

Figure 2.3: Amino acids sequence of 12S seed storage globulin 2 (UniProtKB/Swiss-Prot ID: P14812) from oat seeds (Schubert et al., 1990)

2.4 summarizes the information on these proteins; an example of amino acids sequence from 12S oat seed storage protein 2 (UniProtKB/Swiss-Prot ID: P14812) is shown in Figure 2.3. This protein is composed of 518 amino acids and similar to 11S seed storage protein (globulin) family; essentially, the sequence consists of a signal peptide at the AA position 1-24, an acidic polypeptide chain (α-subunit) at AA position 25-317, and a basic polypeptide chain (β-subunit) at AA position 318-518; a disulfide bond is observed at AA position 45↔78 and a interchain (between α- and β-chain) is observed at AA position 126↔314 (Schubert et al., 1990).

2.2.1.2 3S and 7S Globulins

Chickpea 7S (vicilin) globulin, and oat 3S and 7S globulins are also storage proteins; their structure and conformation have been studied much less than 11S and 12S globulins. Guéguen (1991) reported MWs of the pea 7S globulin, 50, 35, 33, 19, 15, and 13 kDa based on SDS-PAGE characterization, which are comparable to chickpea 7S vicilin reported by Chang et al. (2009). Oat 3S and 7S globulins are minor components of oat storage proteins; Burgess et al. (1983) characterized the major subunit of oat 7S globulin with MW of 55 kDa; some minor components of MW 65 kDa were also observed. Oat 3S globulin was found to have at least two major subunits with MWs of about 15 and 21 kDa (Burgess et al., 1983; Lásztity, 1996).

2.2.2 Glutelin Proteins

Glutelins belong to the 11–12S globulin family; the proportion of glutelins is around 18.1% of the total proteins in chickpea and 23-54% in oat proteins, respectively (Haard,

1999; Singh & Jambunathan, 1982). Structurally (Figure 2.1), glutelins comprise poplypeptides of approximately 55 kDa, and are post-translationally cleaved to give α-subunits (~33 kDa in oats, 28–31 kDa in rice) and β-subunits (~23 kDa and 20–22 kDa, respectively) linked by single disulphide bonds (Shotwell, 1999; Takaiwa et al., 1999).

2.2.2.1 Molecular Characterization of Glutelins

Wen & Luthe (1985) reported that the α-subunits of rice glutelins ranged from 28.5 to 30.8 kDa and the β-subunits ranged from 20.6 to 21.6 kDa using SDS-PAGE characterization; the electrofocusing on polyacrylamide gels showed that the isoelectric points of the α- and β-subunits were 6.5 to 7.5 and 9.4 to 10.3, respectively. Steeson & Sathe (1995) characterized glutelin proteins from Basmati rice by SDS-PAGE. The fraction was comprised of 13 different MWs with major subunits estimated at 15.8, 19.5, 23.4 and 29.5 kDa. Agboola et al. (2005) reported Australian rice glutelin subunits 15.8, 21.4 and 38.5 kDa from SDS-PAGE characterization and 12.7, 21.4, 29.5 and 39.8 kDa from SDS capillary electrophoresis.

2.2.2.2 Homology of Glutelins

A comparison of the amino acid composition (Wen & Luthe, 1985) of rice glutelin subunits with those of the 11S proteins from eight other plant species indicated that there is more similarity between the β-subunits than that of α-subunits of several diverse plant species. Zhao et al., (1983) showed that a partial amino acid sequence of the purified MW 22 kDa rice glutelin subunit was homogenous to the β-subunit of pea legumin. Robert et al. (1985) also reported that rice glutelin, like oat 12S globulin, is in fact a legumin-like

protein due to the similarity of molecular characteristics. Higuchi & Fukazawa (1987) reported a comparison of the amino acid sequence of rice glutelin A subunit precursor with those of glycinin $A_{1a}B_{1b}$ subunit and glycinin A_3B_4 subunit; overall 32% of the amino acid positions are identical in both rice glutelin and soybean glycinin; in results of regions that show identities are dispersed throughout both molecules, the similarity is not due to convergent evolution, but to divergence evolution from a common ancestral gene (Higuchi & Fukazawa, 1987).

2.3 Characteristics of Chickpea and Oat Protein Concentrates

The use of plant protein concentrates (such as cereal, legume and oilseed proteins) as protein supplements in various food formulations has become progressively more attractive than those of animal proteins due to their superior functional properties, low flavor profile, relative freedom from anti-nutrition factors (protease inhibitors, α-amylase inhibitors, oligosaccharides, phytic acid, saponins, phenolic compounds, and tannins), indigestible carbohydrate and low cost of production (Damodaran, 1996; Maiti & Wesche-Ebeling, 2001; Singh et al., 2008). Protein concentrates can not only be chemically modified to improve the functional properties but also can be biochemically modified as protein hydrolysates, potentially health promoting ingredients, using protease enzymes to generate bioactive peptides (Goulet et al., 1987; Yust et al., 2003).

2.3.1 Biological Properties of Chickpea and Oat Protein Hydrolysates

Biological properties from chickpea hydrolysates have been characterized as reduced antigenic (Clemente et al., 1999), angiotensin I-converting enzyme (ACE) inhibitory

(Yust et al., 2003), antifungal (Ye et al., 2002), and antioxidant and free radical-scavenging activities (Arcan & Yemenicioglu, 2007; Li et al., 2008). Ye et al. (2002) reported two chickpea antifungal peptides, Arietin (GVGYKVVVTTTAAAD--DDDVV) and Cicerin (ARCENFADSYRQPPISSSQT), with MW 8.2 and 5.6 kDa using gel filtration chromatography and SDS-PAGE.

Cheung et al. (2009) reported the ACE inhibitory activity of hydrolysates from oat proteins by *in silico* and *in vitro* analyses. Based on *in silico*, over 150 di- or tri-peptides was predicted from oat 12S and 11S proteins; thermolysin (EC 3.4.24.27) was chosen using BIOPEP database as the enzyme would theoretically produce the most ACE-inhibitory peptides from oat proteins; the strong ACE-inhibitory activity was confirmed by various hydrolysis condition (Cheung et al., 2009).

Angiotensin Ⅰ-Converting Enzyme (ACE, EC 3.4.15.1) plays an important role in regulating human blood pressure in the rennin-angiotensin system; essentially ACE elevates blood pressure by not only converting the inactive decapeptide angiotensin I to the vasoconstrictor octapeptide angiotensin II but inactivating the vasodilating nonapeptide, bradykinin; as a result, the inhibition of ACE could effectively exert an antihypertensive effect of decreasing angiotensin II as well as increasing bradykinin (Gohlke et al., 1994; Pihlanto-Leppäläa et al., 1998). Currently, chemically synthesized ACE inhibitors including captopril, the first ACE inhibitor developed from snake venom (Ondetti et al., 1971), served as potent antihypertensive drugs; however, they may trigger adverse side effect. Therefore it is safer and more economical to investigate ACE-inhibitors derived from natural food resources such as milk proteins (FitzGerald &

Meisel, 2000; Pihlanto-Leppäläa et al., 1998), buckwheat proteins (Ma et al., 2006), oat proteins (Cheung et al., 2009) and chickpea proteins (Yust et al., 2003).

2.4 Proteomics for Characterization of Proteins

In 1994, the term "proteomics" was introduced to the scientific community. Proteomics has become increasingly prevalent as PCR (polymerase chain reaction) was at the end of 1980s; nevertheless, unlike the PCR, one of the major weaknesses of proteomics is the absence of an amplification method for the preparation of trace proteins (Rehm, 2006). Over the past two decades, proteomics has been used not only for the identification and characterization of the proteins but also for the sequence, structure, abundance localization, post-translational modification, interaction and biochemical function of proteins or peptides (Twyman, 2004). The potential applications of proteomics to study food including food quality, food allergy and bioavailability of food proteins have lately been demonstrated (Carbonaro, 2004). Since food proteins are fundamental and essential components for human consumption, a more in-depth study on food proteomics needs to be developed to build up a comprehensive understanding of how food proteins interact with molecules nutritionally, functionally, biologically and physiologically to human. Without doubt, applications of proteomics technologies to the study of foodstuffs have become more and more important to health and nutritional issues (Kussmann et al., 2006).

2.5 Methodologies and Technologies in Proteomics

2.5.1 Strategies for Proteomic Analysis

The analysis of proteins involves separating protein mixtures into individual components; there is no single method suitable for the separation of all proteins (Twyman, 2004). Currently, the three major steps are the separation of complex protein mixtures, the identification and characterization of the separated polypeptides, and bioinformatics which is databases searching (Blackstock & Mann, 2000; Görg, 2000; Kvasnička, 2003). Figure 2.4 shows the general proteomic approach by gel electrophoresis (GE) and mass spectrometry (MS) (Crawford et al., 2000). In first step for sample separation, the protein samples are separated by 1-dimensional SDS gel electrophoresis or 2-DGE, the desirable stained spots are excised, subjected to in-gel digestion with trypsin and the resulting peptides then separated by high performance liquid chromatography. From the second step, an eluting peptide is ionized by electrospray ionization then enters the mass spectrometer passing through quadrupole 1 (Q1) and is fragmented in a collision cell (Q2); the fragmented ions of the molecules are divided into daughter ions (B and Y series; shown in Figure 2.5) in order to generate amino acid sequence information. The fragment-ion tandem-MS spectrum is collected (Q3); this contains sequence-specific information. For the last step, the tandem MS spectrum from selected ionized peptide gives information to identify the peptide comparing the acquired results obtained from MS and the theoretical results from a database search (Gygi & Aebersold, 2000).

2.5.2 Two-Dimensional Gel Electrophoresis (2-DGE)

Presently, two-dimensional gel electrophoresis (2-DGE) is the only technique that can separate hundreds or even thousands of proteins at a time. The first dimension separates proteins according to their charge under an immobilized pH gradient (IPG). The

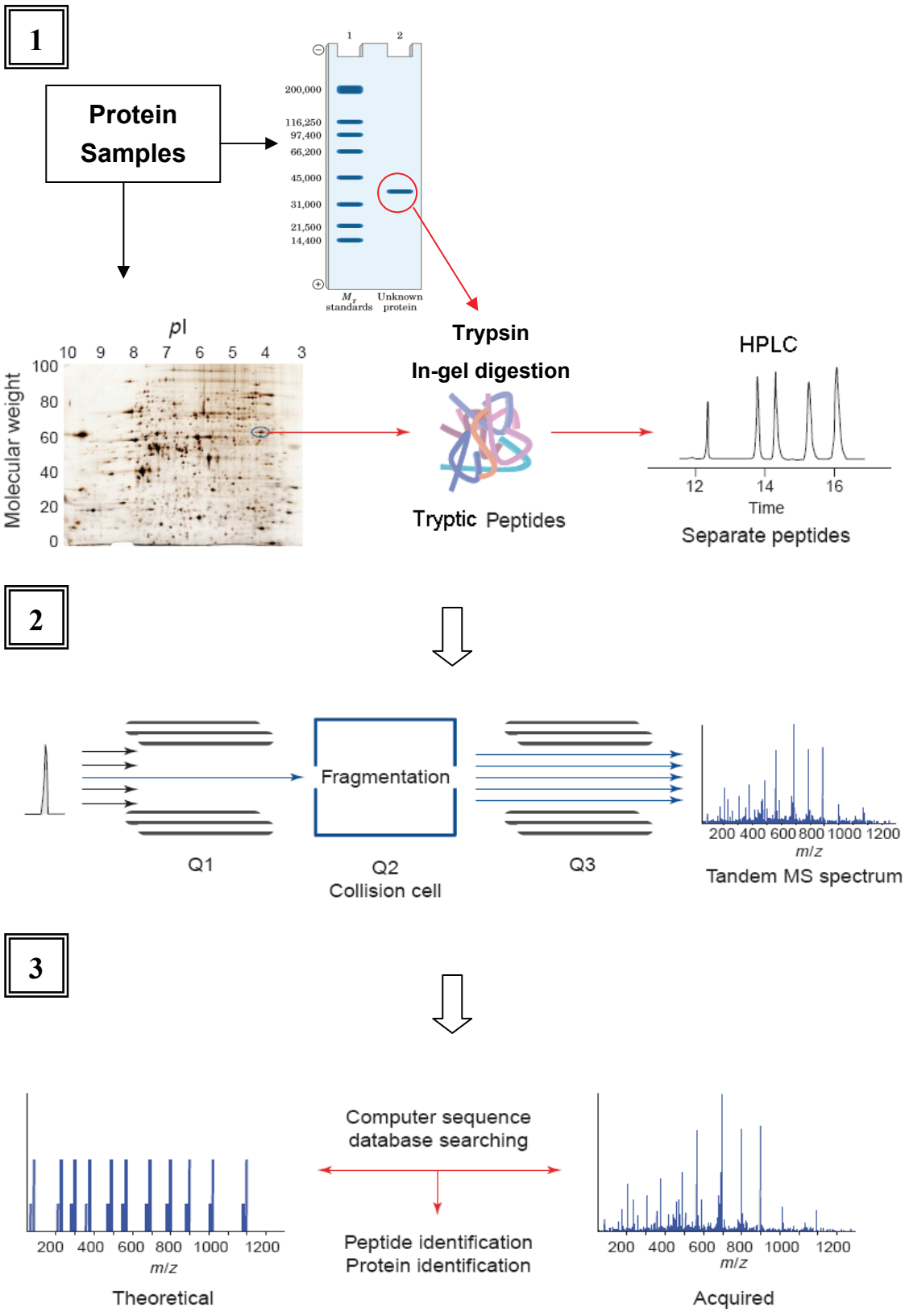


Figure 2.4: General proteomic approach by 1D SDS-PAGE or two-dimensional gel electrophoresis (2-DGE) and mass spectrometry (MS) (Gygi & Aebersold, 2000)

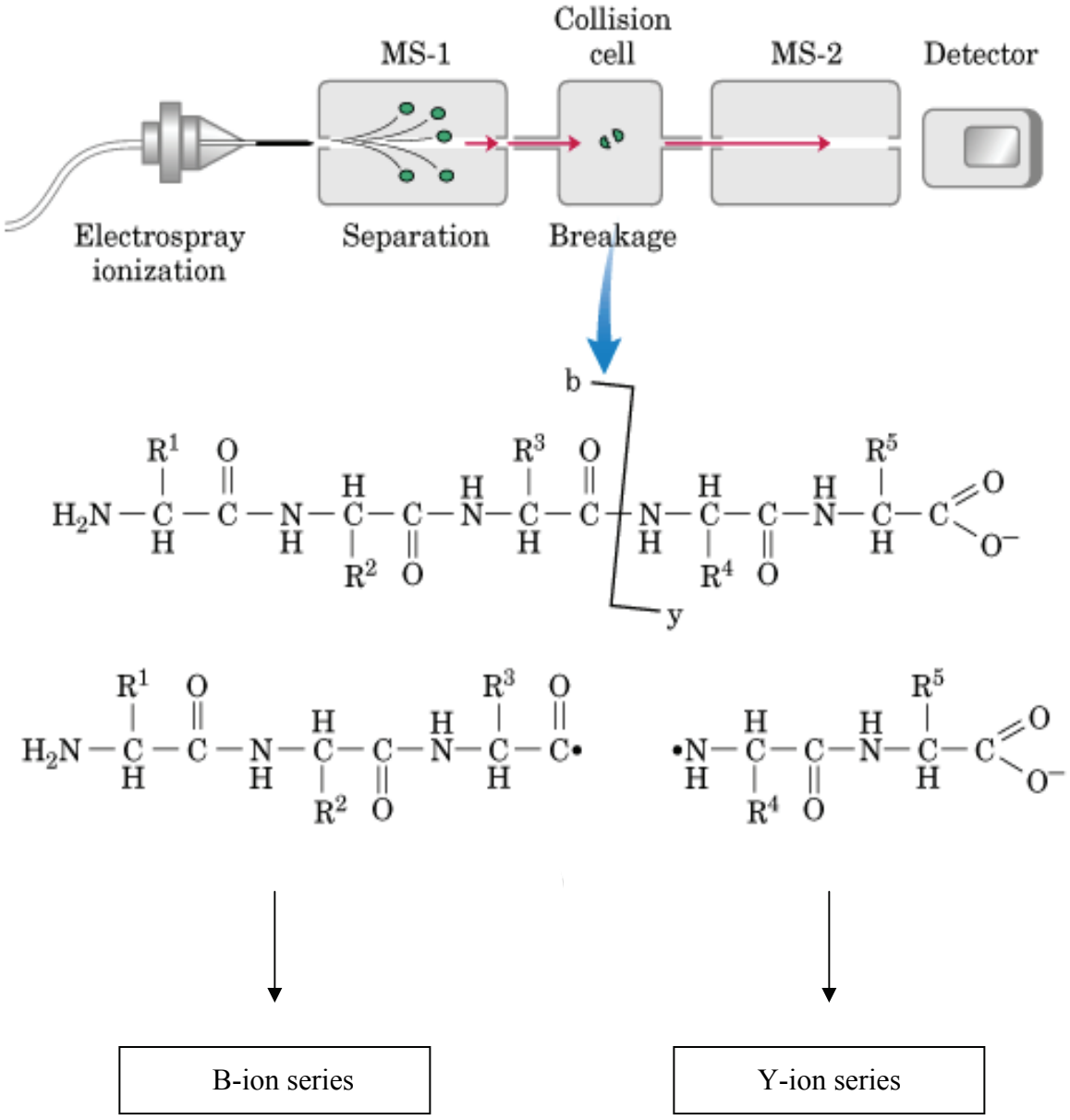


Figure 2.5: Generation of daughter ions by cleaving the amide bond by collision dissociation from ESI/MS/MS analysis where R represents the variable aminoacid side chain (Twyman, 2004)

second dimension separation in 2-DGE is generally conducted by the standard SDS-PAGE and separates the proteins based on molecular mass. However, in order to meet the requirements of proteomic analysis, several conditions or limitations which need to be addressed are: (a) maximum reproducibility of the 2-D maps for the reliable detection of changes in proteome expression of a particular cell, tissue or fluid, and to possibly establish image-based 2-D reference maps and to contrast protein databases, (b) in order to detect the maximum number of spots and to avoid the accumulation of multiple proteins in a spot or the cross-contamination of spots in 2D gels, (c) maximum loading capacity for the micro-preparative separation and purification of proteins for subsequent spot identification; zoom-in gels (with IPGs of pH 4-5, 4.5-5.5, 5-6, etc.) are the methods of choice for micro-preparative separation, and (d) simplified or automated procedures for high-throughput analysis to identify proteins (Görg, 2000; Twyman, 2004).

Herbert et al. (2001) reported that 2-DGE continues to deliver high quality protein resolution and dynamic range for proteomic researchers in terms of the reinforcement of detecting low-abundance proteins, membrane proteins, alkaline proteins and the characterization of protein isoforms.

2.5.3 Protease Digestion

This involves cleaving proteins into peptides at specific peptide bonds using different proteases and is most commonly used to receive sequenceable peptides. One of the most predominant enzymes used in proteomics is trypsin which can cleave proteins at the carboxyl terminal side of lysine and arginine residues. Some considerations for the

use of trypsin are as follows: Arg-Pro or Lys-Pro sites are trypsin resistant; trypsin slowly attacks peptide bindings between a basic and an acidic amino acid; the pH optimum of trypsin lies between 8 and 9 and the optimum ratio of enzyme to substrate is 1:50 to100; calcium ions inhibit the self-digestion of trypsin (Blackstock, 2000; Rehm, 2006).

2.5.4 In-Gel Protein Digestion

In-gel digestion of proteins is performed after the separation of 2-DGE before analyses by mass spectrometry. Individual protein spots are treated by specific enzymes to produce fragments (peptides) which can be used for protein identification in MS (Kvasnička, 2003). Two essential requirements for in-gel digestion prior to MS analysis are the efficient peptide recovery and minimal contamination.

In 2-DGE analysis, various staining methods can be used to visualize proteins. Excellent recoveries are usually obtained with Coomassie Blue or colloidal Coomassie, followed by destaining with 10% acetic acid and rising into water. More sensitive staining methods that show good recovery for in-gel digestion include reverse-staining with Zn^{2+} -SDS- imidazole, and SyproRuby fluorescence staining; the latter is highly sensitive when visualized with fluorescence image analysis. Silver staining recipes that use aldehyde fixation give the highest sensitivity but result in lowest recovery from in-gel digestion among the methods described above (Resing & Ahn, 2004).

2.5.5 Determining Protein Sequences by Chemical Degradation

2.5.5.1 Compete Hydrolysis

Proteins can be completely broken down into their constituent amino acids by boiling in highly concentrated hydrochloric acid for 24-72 hours. The amino acids can then be labelled with an agent such as ninhydrin or fluorecamine, separated by HPLC and detected as they elute from the column using a panel of standard amino acid as a reference. The sensitivity of fluorecamine labelling can be as little as 1ng amino acid. The main drawback of this method is that it shows the amino acid composition of a protein and it does not reveal the sequence since all the peptide bonds in the proteins are broken (Twyman, 2004).

2.5.5.2 Edman Degradation

Edman degradation, one of several methods for sequencing protein from either N-terminal or C-terminal ends using broad-specificity exopeptidase, was developed and automated by Pehr Edman between 1960 and 1967. Edman degradation involves labelling the N-terminal amino acid of a protein or peptide with phenyl isothiocyanate. Mild acid hydrolysis results in cleavage of the peptide immediately adjacent to this modified residue, but leaves the rest of the protein intact. The terminal amino acid (or rather its phenylhyronation derivative) can then be identified by chromatography, and the procedure is repeated on the next residue and the next, thus building up a longer sequence. However, without automation, Edman degradation can sequence a small peptide (ten residues) in about 24 hours and a larger peptide (30-40 residues) in 3 days. It is not suitable for sequencing proteins larger than 50 residues in a single run because each cycle

of degradation is less than 100% efficient. With this regard, after a large numbers of cycles there will be a mixed population of molecules in the analyte (Twyman, 2004).

2.5.6 Mass Spectrometry

As a strongest competitor of Edman degradation method, mass spectrometry (MS) does provide a much more powerful technique for protein identification and characterization. More importantly, due to its versatility, it can be coupled with many techniques (such as HPLC, 2-DGE) and offers a promising future for proteomics (Rehm, 2006; Smith, 2002). There are three reasons for this reported by Blackstock (2000). The first is the development of new methods for the ionization of proteins and peptides, especially matrix-assisted laser desorption-ionization (MALDI) and electrospray ionization (ESI) (Shown in Table 2.5); combine with time-of-flight (TOF) and new hybrid spectrometers, these offer high sensitivity and selectivity. The second is the rapid growth of databases: over 2×10^9 bases of human genomic sequence are now available and software has been written to search protein, express-sequence-tag (EST) and genome databases with MS data. The third is that MS provides detailed information on post-translational modification, in particular phosphorylation and glycosylation (Blackstock, 2000).

Mass spectrometers comprise three principal components: a source of ions (to generate ions from the sample), a mass analyzer (to separate the ions according to their mass/charge (m/z) ratios), and a detector. For proteomics, MALDI and ESI are widely used and analysers range from simple (e.g. time of flight, TOF) to complex (e.g. Fourier transform ion-cyclotron resonance, FTICR). Both can provide information on molecular

Table 2.5: Mass spectrometry performance in proteome measurements (Simth, 2000)

Instrument type	Detection limit (moles)[a]	Dynamic range[b]	Resolving power	Mass accuracy[c]
IT	10^{-16}-10^{-14}[d]	~10^2	~10^3	300-1000 ppm
QTOF	10^{-15}-10^{-13}[d]	~10^2	~10^4	5-40 ppm[e]
FTICR	10^{-19}-10^{-15}	~10^3	~10^5	<1[f]-10 ppm

[a]Approximate range of minimum concentrations (detection limits) for peptide tandem mass spectrometry (MS-MS) identification. 1 fetomole = 10^{-15} moles; 1 attomole = 10^{-18} moles.

[b]The difference in intensity between the most and least abundant peaks for a spectrum acquired in -1 s in combination with liquid chromatography separations.

[c]Mass measurement accuracy in parts per million (ppm). The greater the accuracy of mass measurements, the greater the ability to distinguish peptides and make correct identications.

[d]While ion traps (IT) provide greater sensitivity for MS-MS compared to quardrupole-time-of-flight (QTOF), the opposite is true for mass spectrometry (MS) measurements.

[e]Accuracy can vary significantly with peak intensities; peaks that are distorted due to high signal intensities, or poorly defined due to low intensities, yield worse accuracies.

[f]Better mass accuracies are achievable with Fourier transform ion cyclotron resonance (FTICR) using internal calibration.

mass and fragment ions. Furthermore, mass spectrometer generate both mass and sequences information for the peptides digested by specific enzyme. Some examples include MOALDI-TOF and LC-ESI mass spectrometers which measure mass only; they are used for simple peptide mixtures; LC-ESI mass spectrometers are more suitable for the analysis of complex protein sample; tandem mass spectrometers are the most powerful tool for peptide sequencing; combination of either LC-ESI or MALDI on

tandem mass spectrometers gives both peptide mass and sequence; FTICR mass spectrometers improve the mass accuracies and sensitivity (Appendix 1) (Blackstock, 2000; Twyman, 2004).

2.5.7 Structural Proteomics

Both protein sequence and structures can be used to predict protein function and how proteins interact. In reality, it is difficult to obtain the structural data other than sequences. The two major techniques for solving protein structures are X-ray crystallography (XRC) and nuclear magnetic resonance (NMR) spectroscopy. More than 98% of structures in the Protein DataBank (PDB, http://www.rcsb.org) have been solved using one of these two methods (Twyman, 2004). Pechkova & Nicolini, (2004) reported that a new approach (protein nanocrytallography) to structural proteomics that can produce and characterize diffracting, stable and radiation-resistant crystals of miniscule dimensions using nanotechnology which is based on protein thin-film template.

2.5.8 Bioinformatics in Proteomics

Bioinformatics is an interdisciplinary research area at the interface between computer science and biological science; it involves the technology that uses computers for storage, retrieval, manipulation, and distribution of information corresponded to biological macromolecules such as DNA, RNA, and proteins (Xiong, 2006). The 2-DGE and mass spectrometry are the most prevalent technologies within proteomics. Both methods resolve protein and peptides by physical parameters that can be used in protein identification. 2-DGE separates proteins by size and charge. With carefully standardized

conditions, position on the gel can be easy to identify some proteins. Using SWISS-2DPAGE, a database on the ExPASy server (Table 2.6), many standardized gel images and tool can help to predict or identify the key proteins in comparison to experimental 2-DGE images. With MS, a large amount of data is generated for highly accurate mass and sequence information of peptide fragments isolated from proteolytic digests of gel-separated proteins (Crawford et al., 2000). To deal with this considerable data, resources for the identification of protein by peptide masses and sequence include Swiss-2DPAGE, ProFound, Peptldent, MASCOT, PepFrag, Protein Prospector, ExPAsy, NCBI/BLAST Tools, SWISS-PROT, PIR and BIOPEP; the URL links of these resources are shown in Table 2.6.

Identifying proteins by peptide mass requires access to a protein-sequence database. Crawford et al. reported the most commonly used databases are SWISS-PROT, TrEMBL and the non-redundant collection (nr) of protein sequences at the US National Center for Biotechnology Information (NCBI). SWISS-PORT is an annotated collection of protein sequences on the ExPASy serve; TrEMBL is a large collection of predicted protein sequences given automatic annotation until they can fully annotated and then entered into SWISS-PORT; the NCBI nr database contains translational protein sequences from the entire collection of DNA sequences kept at GenBank, and also the protein sequences in the Protein Data Bank (PDB), SWISS-PORT and Protein Information Resource (PIR) databases (Beavis & Fenyo, 2000; Crawford et al., 2000); the NCBI/BLAST tool is used for analyzing the similarities of nucleotides and proteins sequences or their alignments (Altschul et al., 1997).

Table 2.6: Useful web tools for protein databases (Crawford et al., 2000; Iwaniak et al., 2005)

Stie name	URL	Information available
Swiss-2DPAGE	http://ca.expasy.org/ch2d/	2DPAGE image analysis
ProFound	http://prowl.rockefeller.edu/	Peptide mass mapping and sequencing
Peptldent	http://ca.expasy.org/tools/peptident.html	Peptide mass mapping and sequencing
MASCOT	http://www.matrixscience.com/	Peptide mass mapping and sequencing
PepFrag	http://www.proteometrics.com	Peptide mass mapping and sequencing
Protein Prospector	http://prospector.ucsf.edu/	Peptide mass mapping and sequencing
ExPAsy	http://ca.expasy.org/	Protein Server
NCBI/BLAST Tools	http://www.ncbi.nlm.nih.gov/BLAST/	Sequence Databases
SWISS-PROT	http://www.expasy.ch/sprot/	Sequence Databases
PIR	http://pir.georgetown.edu/pirwww/	Sequence Databases
BIOPEP	http://www.uwm.edu.pl/biochemia/index_en.php	Bioactive Peptides Databases

Beavis & Fenyo (2000) suggested that future developments for data search engines to identify proteins by experimentally determined masses will be shaped into the following points including improvements in the signal-processing algorithms used to generate mass spectra, the automation of protein-identification searches and the rational storage of results in large databases, the development of data-dependent search engines

that can guide the data-gathering in real time, and post-processing using statistical data from complied databases of theoretical and experimentally determined properties.

2.5.8.1 BIOPEP Database

A relatively new database of biologically active peptide sequences, BIOPEP (http://www.uwm.edu.pl/biochemia), has been established and applied for classification of numerous food proteins as potential sources of bioactive peptides. This database has collected information on 2123 peptides representing 48 types of reported bioactivities, their EC_{50} values and source of origin; the usefulness of this tool is to analyze protein samples *in silico* by predicting the potential bioactive peptides cleaved from various proteases (Iwaniak & Dziuba, 2009; Iwaniak et al., 2005); Cheung et al. (2009) reported the predicted ACE-inhibitory peptides from oat proteins using BIOPEP analysis (Section 2.3.1).

2.6 The Scope and Applications of Food Proteomics

Korhonen et al. (1998) surveyed the impact of processing on bioactive proteins and peptides which revealed how proteins may be altered by the treatments during the processing and storage such as heat treatment, fermentation, ultra high pressure, extrusion, pH modification, enzymatic hydrolysis and membrane separation. He also suggested that the bioactive proteins or peptides may not possess allergic or toxic effects after these treatments. However, their interactions during processing or storage with other proteins, sugars, and lipids need to be studied with a view to generate possible formation of toxic, allergic or carcinogenic substances. By applying proteomic techniques to this idea, it is

likely to develop the analytical methodologies to detect the presence of allergens or toxicities or even the modified proteins (Wiseman, 2004).

2.6.1 Novel Functional Bioactivity from Food Proteins

Mostly known bioactive proteins or peptides are from common foods such as milk or egg proteins (Korhonen et al., 1998). The proteomic approach addresses the identification of bioactive proteins or peptides and requires another complementary study such as metabonomics to see how bioactive proteins or peptides interact with molecules in human biological system (Kussmann et al., 2006).

2.6.2 Food Safety Evaluation

The combination of pathogen-specific markers with the detection on DNA microarrays can provide a sensitive, quick and informative detection of pathogens in food and food production sites than doing classical culture methods (Abee et al., 2004). Screening of potential changes in the physiology of genetically modified (GM) food crop can address the safety concern of GM foods by understanding possible altered expression of macro- and micro-nutrients with specific toxins or allergens. Comparative protein analysis of GM food crops and counterparts with similar genetic background can be analysed through proteomic techniques (Kuiper et al., 2002).

2.6.3 Food Allergy

The identification of allergic proteins using proteomic analysis based on the recognition of specific post-translational modification and digestion-resistant peptide

features, allows the experimental data to compare with previously characterized allergenic proteins (Carbonaro, 2004; Kvasnička, 2003). Application of proteomics in food allergens (Beyer et al., 2002a) indicated four allergens in sesame seeds with the immunolabeling performance with 20 patients' sera. Two are well known allergens, 2S albumin precursor and 7S globulin (sesame seed) from seed storage proteins; Natale et al. (2004) used similar proteomic approaches to identify bovine milk allergens. Recent study demonstrated the proteomic characterization of Kunitz trypsin inhibitor, which is known as an anti-nutritional/allergen, in wild and cultivated soybean (Natarajan et al., 2006).

2.6.4. Food Quality and Authenticity

Proteomics offers a great tool to identify the specific gene products that can be involved in meat quality alterations (Carbonaro, 2004). Martinez & Friis (2004) reported on the comparison of two-dimension (2D) maps of the intact and treated samples to identify fish species and muscle tissues, and characterize the post-mortem changes in arctic and tropical species, and study the effect of some additives during the processing of fish muscle. Proteomic analysis of amphiphilic proteins of hexaploid kernels was performed using 2-DE gels, MALDI-TOF and ESI-MS/MS to obtain useful information about the protein components which can be linked to bread wheat quality and particularly to kernel hardness (Amiour et al., 2002).

2.6.5 Food Processing

A proteome reference map of major soluble proteins from *Medicago Sativ.* (alfalfa) leaves and stems has been established industrial process proteomics (Incamps et al., 2005).

This study showed the variation of the protein patterns at different steps of industrial-scale processing by using 2-DGE; the process induced significant changes in chemical modifications, proteolytic events, heat-shock protein responses. In complex fermentations by different microoganism and complex substrates, the quality of proteome or metabolome of starter culture can be used to predict the quality of the fermented end-product; substantial cost can be saved if the bad starter culture batches can be expelled (Carbonaro, 2004).

CHAPTER 3

ISOLATION AND CHARACTERIZATION OF CHICKPEA AND OAT PROTEIN FRACTIONS

3.1 Justification

Chickpea and oat globulin fractions are the predominant proteins in chickpea and oats, and have been extensively studied in comparison to other fractions such as albumin and glutelin fractions. This chapter addresses the first, second and third research objectives: (1) to isolate individual protein fractions according to Osborne protein classification, (2) to characterize their molecular properties using native, sodium dodecyl sulfate and 2-dimensional gel electrophoresis (Native, SDS-PAGE and 2-DGE), and reversed-phase high performance liquid chromatography (RP-HPLC) and (3) to perform enzymatic digestion for chickpea isolates with digestive enzyme (trypsin) in vitro simulating the human gastric condition and to evaluate the level of protein digestion.

3.2 Materials and Methods

3.2.1 Materials

Chickpea (Kabuli type) seeds and oat seeds were purchased from Emerson Milling Inc (Emerson, MB, Canada), and rice grains (NuPak, USA) were purchased locally in Montreal; seeds were ground using a micro sample mill (Braun, KSM2 Type 4041, CA) and stored in air-tight plastic containers at room temperature until they were used. All chemical reagents were analytical grade.

3.2.2 Extraction of Chickpea, Oat and Rice Protein Isolates

3.2.2.1 Sequential Extraction of Proteins from Chickpea and Oat Flour

Albumin, globulin, and glutelin protein isolates were extracted sequentially according to the procedure of Agboola et al. (2005) with few modifications. Figure 3.1 illustrates the scheme of sequential extraction. A sample (50 g) of chickpea and oat flour was mixed with distilled water (200 mL). The mixture was allowed to stand with intermittent stirring for 2 h and centrifuged (8000×g, 20 min). The supernatant was filtered with fine glass wool and then precipitated by addition of 1 M HCl solution to pH 4.1 with subsequent centrifugation (8000×g, 20 min); this albumin protein isolate (chickpea albumin (C-Ab); oat albumin (O-Ab)) was obtained after lyophilization of the precipitate and the supernatant was discarded. The residue 1 was mixed with 5% NaCl solution (200 mL) and stirred for 2 h then centrifuged (8000×g, 20 min); the supernatant was filtered with fine glass wool and then precipitated by addition of 1 M HCl solution to pH 4.3 with subsequent centrifugation (8000×g, 20 min); this globulin protein isolate (chickpea globulin (C-Gb); oat globulin (O-Gb)) was recovered after lyophilization of the

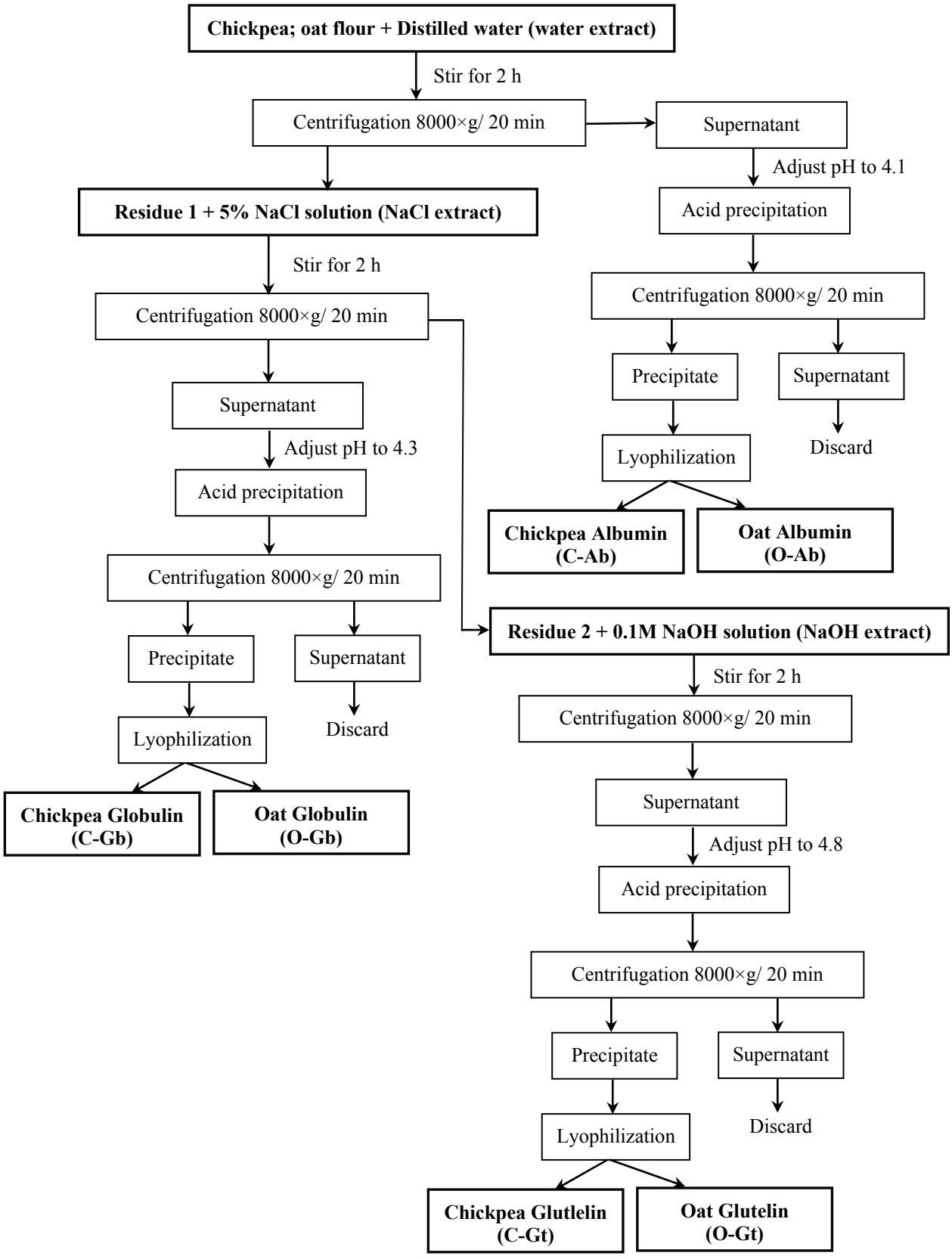


Figure 3.1: Scheme of sequential protein extractions for chickpea and oat flour

precipitate and the supernatant was discarded. The residue 2 was mixed with 0.1M NaOH solution (200 mL) and stirred for 2 h then centrifuged (8000×g, 20 min); the supernatant was filtered with fine glass wool and then precipitated by addition of 1 M HCl solution to pH 4.8 with subsequent centrifugation (8000×g, 20 min); this glutelin protein isolate (chickpea glutelin (C-Gt); oat glutelin (O-Gt)) was recovered after lyophilization of the precipitate and the supernatant was discarded. All the freeze-dried protein isolates were store at -20 ℃ until further use.

3.2.2.2 Extraction of Rice Glutelin Isolate

Rice glutelin isolate (R-Gt) was prepared by the extraction of using 0.02% NaOH solution followed by acid precipitation (1M HCl to pH 4.8); the precipitate was lyophilized and designated as rice glutlein isolate (Chandna & Matta, 1990); it was used as a reference protein since there is little information on chickpea and oat glutelins.

3.2.3 Determination of Total Soluble Protein Contents and Yields

The total soluble protein contents of chickpea, oat and rice protein isolates were determined by a modified Lowry assay (Bensadoun & Weinstein, 1976). A standard curve (Figure 3.2) was prepared using 2-100 μg of bovine serum albumin (BSA). All analyses were conducted in triplicate. Protein yield was calculated on the basis of the weight of protein fractions isolated and on the weight of the protein content of chickpea, oat flour.

$$\text{Yield\%} = \frac{\text{Weight of protein fractions} \times 100}{\text{Weight of samples} \times \text{Protein content (chickpea or oat flour) \%}} \%$$

3.2.4 PolyAcrylamide Gel Electrophoresis (PAGE)

3.2.4.1 Native-PAGE

Native-PAGE was performed according to the method of Davis (1964). A Mini-Protean Ⅱ Electrophoresis Cell unit (Bio-Rad, Hercules, CA) was used with a 4% acrylamide stacking gel and a 8% separation gel. Sample solutions (20 μL) were prepared from 8-20 mg of freeze-dried protein precipitates dissolved in 1 mL sample buffer (distilled water, 1.5 M Tri-HCl pH 8.8, glycerol and 1% bromophenol blue), and injected into each sample well. High molecular weight calibration kit (Amersham Bioscience, UK) was used; the standard protein markers were thyroglobulin (669,000 Da), ferritin (440,000 Da), catalase (232,000 Da), lactate dehydroxygenase (140,000 Da) and albumin (66,000 Da). The migration of proteins was carried out for approximately 3.5 h at constant current (7.5 mA/gel). Gels were fixed with fixing solution (water: methanol: acetic acid/700 mL: 200 mL: 100 mL) for 30 min and then stained with Coomassie Brilliant Blue R-250 for 1 h. The stained gels were destained by frequently changing the fixing solution until the excess stain disappeared.

3.2.4.2 Sodium Dodecyl Sulfate-PAGE (SDS-PAGE)

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was carried out using the technique reported by Laemmli (1970) with 4% stacking gel and 12% resolution gel. Sample solutions (20 μL) were prepared from 10 mg of freeze-dried protein precipitates dissolved in 1 mL sample buffer (distilled water, 0.5 M Tri-HCl pH 6.8, glycerol 10% SDS, 1% bromophel blue and β- mercaptoethanol) heated at 98 ℃ for 10 min, then applied to the sample wells. The standard protein marker (broad range

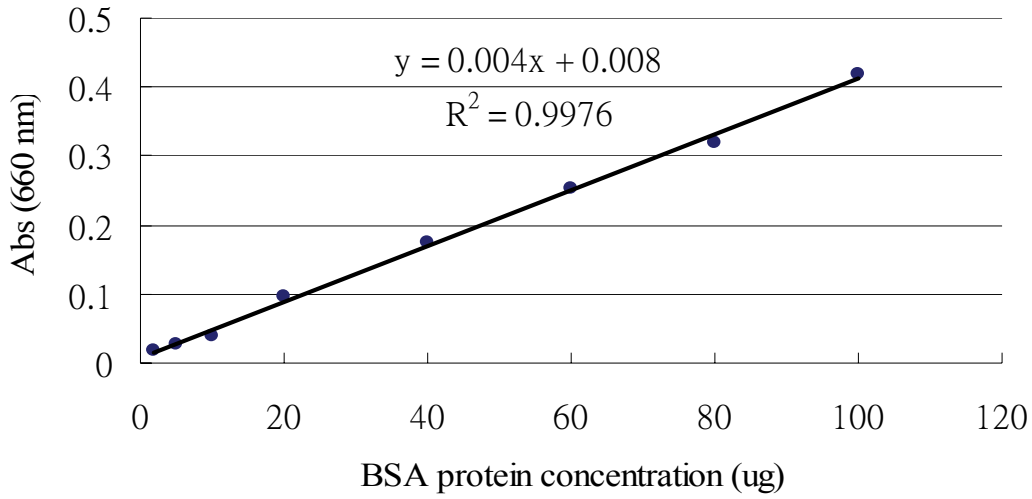


Figure 3.2: Standard curve of BSA protein concentration for Lowry assay

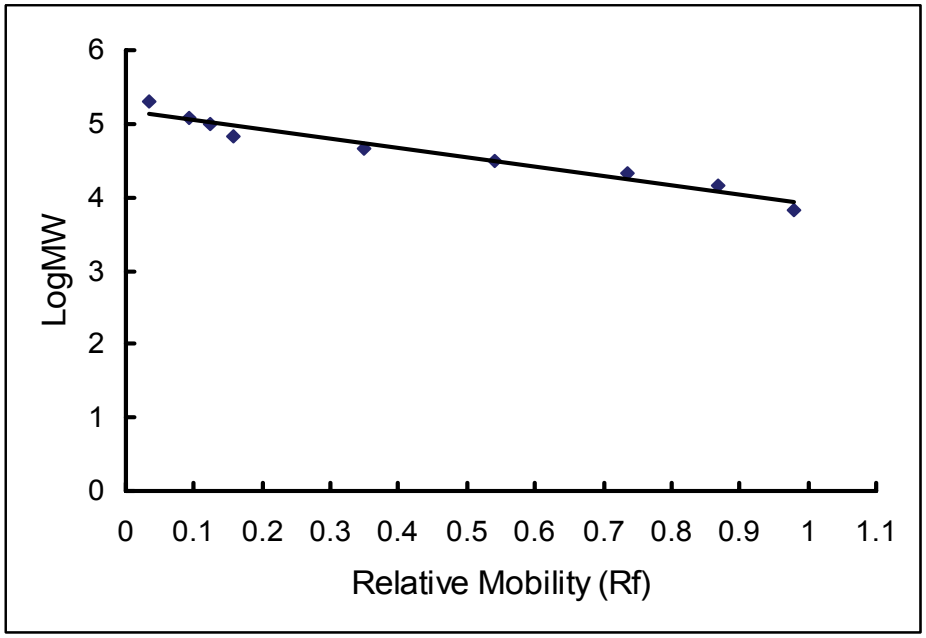


Figure 3.3: Standard curve generated by plotting the log of molecular weights of broad range standard protein markers (myosin MW 200 kDa, β-galactosidase MW 116.2 kDa, phosphorylase b MW 97.4 kDa, serum albumin MW 66.2 kDa, ovalbumin MW 45.0 kDa, carbonic anhydrase MW 31.0 kDa, trypsin inhibitor MW 21.5 kDa, lysozyme MW 14.4 kDa, and aprotinin MW 6.5 kDa) vs. the relative mobility (Rf); this curve was used to determine molecular weights of unknown proteins in SDS-PAGE gels

molecular weight, Bio-Rad Hercules, CA) which contained myosin (200,000 Da), β-galactosidase (116,250 Da), phosphorylase b (97,400 Da), serum albumin (66,200 Da), ovalbumin (45,000 Da), carbonic anhydrase (31,000 Da), trypsin inhibitor (21,500 Da), lysozyme (14,400 Da) and aprotinin (6,500 Da), was used to prepared a standard curve for molecular weight estimation (Figure 3.3). Electrophoretic migration was monitored at constant voltage (120V) for 1.5 to 2 h. Gels were fixed with fixing solution (water: methanol: acetic acid/700 mL: 200 mL: 100 mL) for 30 min and then stained with Bio-safe stain (Bio-Rad) for 2 h. The stained gels were destained by changing the fixing solution until the excess stain disappeared.

3.2.4.3 Two-Dimensional Gel Electrophoresis (2-DGE)

The first dimension isoelectric focusing (IEF) separation was performed as described by Righetti & Bossi (1990) with few modifications. The IPG gel strip's dimensions were as follow: 70 × 3.0 ×0.5 mm, pH range: 3–10, on a polyacrylamide gel matrix) using an Amersham Pharmacia Biotec IPGPhor unit (GE Healthcare, Uppsala, Sweden). Separation on the SDS-PAGE was carried out with Bio-Rad electrophoresis unit (Bio-Rad Laboratories, Richmond, CA) using separating gel (12.5%) and SDS-buffer as described by Laemmli (1970). The standard protein marker (Precision Plus Protein Prestained Standards; Bio-Rad Laboratories, Richmond, CA) was used with MWs ranging from 250 to 15 kDa.

3.2.4.3.1 Trichloroacetic Acid (TCA)/Acetone Precipitation of Proteins

Samples (50-100 μg) of chickpea and oat protein fractions were prepared by TCA/acetone precipitation which is employed to selectively separate proteins in the sample from contaminating species such as salts, detergents, dye excess, et., that would otherwise interfere with the 2-DGE results. Protein samples were precipitated with cold 25% TCA and kept on ice for 15 min then spun at 10,000 rpm, 5 min in Eppendorf tubes. The supernatant was removed and the protein pellet was washed with 12.5% TCA and spun again using the same condition, followed by acetone wash, then dried in a speed vac concentrator (Savant, USA). The purified protein samples were stored at -20 ℃.

3.2.4.3.2 IPG Strip Rehydration

Protein samples were dissolved in 150 μL of 2% rehydration buffer pH 3-10 (20 mM dithiothreitol (DTT), 8 M urea, 2% cholamido-propyl-dimethylammonio-propane sulfonate (CHAPS), 0.5% (v/v) IPG buffer (pH 3-10) and 0.002% bromophenol blue); the mixed samples (125 μL) were carefully loaded on to IPG strips in order not to generate any bubbles. Subsequently the IPG gel strips with a nonlinear pH range (3–10) were rehydrated for 10 h at 20°C into the strip holder covered with mineral oil.

3.2.4.3.3 Isoelectric Focusing (IEF) Run

After rehydration, strip holder was positioned on the IPGphor and isoelectric focusing was run with the initial voltage limited to 500 V for 30 min, and then stepped up to 1000 V for 30 min and finished at 5000V for 1 h and 45 min.

3.2.4.3.4 IPG Strip Equilibration and Running Gels

After isoelectric focusing run, the IPG gel strip was prepared for transfer to the second dimension by soaking with gentle agitation for 15 min in an equilibration solution (10 mL) containing 50 mM tris–HCl pH 8.8, 6 M urea, 30% glycerol, 2% SDS, 0.002% bromophenol blue and 100 mg dithiothreitol (DTT). The equilibrated IPG gel strip was embedded at the top of the SDS-PAGE gel in molten 2% (w/v) agarose in electrode buffer using tris–glycine pH 8.3 as buffer for the second dimension. SDS-PAGE gels (12.5%) were run at 50-60V per gel for 90-120 min.

3.2.4.3.5 Silver Stain of 2-DGE Gels

2-DGE gel was placed in fixing solution (50 mL; 50% ethanol, 5% acetic acid in water) for 30 min then washed for 10 min in 50% ethanol (50 mL) and for 2 × 10 min in water. The 2-DGE gel was placed in 0.02% sodium thiosulphate (50 mL) for 3 min then washed for 2 × 3 min in water. For staining step, the gel was incubated in 0.1% silver nitrate solution (50 mL) for 30 min. After staining, the gel was rinsed for 2 min in water then a small amount (~5 mL) of developer (0.04% formalin in 2% sodium carbonate) was added into the gel with gently swirl and the solution was discarded; the developer solution was added again and discarded until the protein spots appeared, then a stop solution (50 mL; 5% acetic acid in water) was incubated for 5 min.

3.2.5 Separation of Chickpea and Oat Protein Fractions by Reversed-Phase High Performance Liquid Chromatography (RP-HPLC)

A quantity (10-30 mg) of the lyophilized chickpea and oat protein fraction was dissolved in 1 mL of the trifluoroacetic acid (TFA) solution (0.1%, water:acetonitrile/9:1), then filtered through a membrane filter (0.45 μm, Osmonics Inc., USA). The filtrates were subjected to RP-HPLC using the procedure reported by Alli et al. (1994) with modifications; the equipment was a Beckman model liquid chromatography system (Beckman, USA). Samples injected manually using a 100 μL loop (Life Science, USA). Separation was performed using a guard column (RP-HPLC Guard Column 4.6 × 50 mm, J. T. Baker, USA) followed by a reversed phase analytical column (C4, 5 micron pore size, 4.6 × 250 mm; Restek Co., USA) and a gradient solvent system was with a flow rate of 0.5 mL/min. Solvent A was 0.1% trifluoroacetic acid (TFA) in deionized water and B consisted of 0.1% trifluoroacetic acid (TFA) in acetonitrile/deionized water (70:30). A linear gradient elution (30% to 70% solvent B 30 min) was used for separation. Elution profiles were detected by a programmable detector module (model 166) at 220nm and fractions which showed relatively high response were collected using Water Fraction Collector (NE, USA), dried in a speed-vac concentrator (Savant, USA). Chromatographic data were analyzed by Beckman Gold System (version V810, USA) then translated from print (PRN) format to Microsoft Excel© worksheet. Fractions obtained from RP-HPLC were subjected to SDS-PAGE analysis.

3.2.6 Preparation of Chickpea Protein Hydrolysates

3.2.6.1 Trypsin Hydrolysis

Chickpea protein fractions (C-Ab, C-Gb and C-Gt) were prepared as described in Section 3.2.2.1; trypsin (EC 3.4.21.4; porcine pancreas, 13,000-20,000 BAEE units/mg protein, Sigma, Canada) hydrolysis of chickpea protein fractions was performed to simulate human gastric condition according to the methods reported by previous researchers (Chen et al., 1996; Pihlanto-Leppälä et al., 1994) with modifications. An enzyme/substrate ratio of 1:100; 2% (w/v) solution of protein isolates in 10 mM phosphate buffer at pH 8 for tryptic hydrolysis. 0.1% (w/v) trypsin in distilled water was prepared; the enzyme-substrate mixture was incubated for 3 h at 37 ℃ and the reaction was stopped by cooling the samples to 0 ℃. Aliquot (25 mL) of hydrolysates was centrifuged (2000×g, 10 min) and the supernatant was lyophilized. Hydrolysates (50 μL) were collected at 30 min intervals and used to measure the degree of hydrolysis.

3.2.6.2 Degree of Hydrolysis Analysis

The degree of hydrolysis (DH) was determined using the methods reported by Adebiyi et al. (2008) and Church et al. (1983) with the *o*-phthaldialdehyde (OPA) agent. 50 μL hydrolysates were mixed with 2 mL OPA reagent in a disposable cuvette for 2 min at room temperature. The absorbance was measured at 340 nm (UV/visible spectrophotometer, Ultrospec 2100 pro; GE, Canada). The DH was calculated on the basis of percentage using the following equation:

$(\mathrm{Mw}\triangle_{340\ \mathrm{nm}})/\ (d \times\ \varepsilon \times P) \times 100,$

Where Mw is the averaged molecular weight of amino acids, 120, $\triangle_{340\,nm}$ the absorbance at 340 nm, *d* the dilution factor, 1/41, ε the constant value 6000 $_M{}^{-1}cm^{-1}$ and *P* is the protein concentration mg mL^{-1}, 0.1%.

3.2.7 SDS-PAGE Characterization of Chickpea Protein Hydrolysates

Samples of lyophilized chickpea protein hydrolysates were submitted to 12% SDS-PAGE characterization described in Section 3.2.4.2.

3.3 Results and Discussion

3.3.1 Total Soluble Protein Contents and Yields of Chickpea and Oat Protein Fractions

Table 3.1 summarizes the protein contents and yields of chickpea and oat protein fractions. The protein contents of chickpea protein fractions ranged from 27.5-93.8%; the protein fraction C-Gb (93.8% protein) is comparable to values of reported chickpea isolates extracted by sodium hydroxide solution (84% protein, Paredes-López et al., 1991; 74% protein, Sánchez-Vioque et al., 1999); the relatively low protein content (27.5%) of fraction C-Gt can conceivably be because of this protein isolate was extracted the last during the sequential extraction procedure thus it may retain more carbohydrates than other protein fractions. Fraction C-Gb showed a relatively low protein yield (16.3%) similar to fraction C-Gt (13.1%) and the fraction C-Ab showed the highest protein yield (40%); the combined protein yield of the three chickpea fractions was 69%; Sánchez-Vioque et al. (1999) reported relatively low soluble nitrogen (~20%) at pH 5 in

Table 3.1: Total soluble protein contents and yields of chickpea and oat isolates

Protein Fractions[a]	Protein Content %[b]	Yield %[b]
Chickpea		
C-Ab	43.06 ± 0.91	40.05 ± 1.03
C-Gb	93.87 ± 1.63	16.31 ± 0.77
C-Gt	27.50 ± 1.50	13.17 ± 0.39
Oat		
O-Ab	67.40 ± 2.51	3.54 ± 0.34
O-Gb	65.70 ± 2.04	23.97 ± 0.89
O-Gt	77.27 ± 1.10	38.32 ± 1.14

[a] C-Ab: chickpea albumin fraction, C-Gb: chickpea globulin fraction, C-Gt: chickpea glutelin fraction, O-Ab: oat albumin fraction, O-Gb: oat globulin fraction, O-Gt: oat glutelin fraction

[b] Data are means and standard deviations of triplicate determinations

which the fraction C-Gb was extracted in that range (pH 5.95); this may explain fraction C-Gb had low protein yield (13%).

The protein contents of oat protein fractions shared similar results (65.7-77.2% protein); these results agree with those reported by Pernollet et al. (1982). The combined protein yield of the three oat fractions was 65% which is similar to the value of combined chickpea protein yield (69%) and lower than oat protein yield of 90% reported by Ma and Harwlakar (1984); the oat albumin fraction (O-Ab) which showed lowest protein yield (3.5%) was lower than to the protein content (11.7%) of oat albumin fraction determined by Ma and Harwlakar (1984).

3.3.2 PAGE Characterization

3.3.2.1 Native-PAGE Characterization of Chickpea, Oat and Rice Protein Fractions

Glutlein and albumin fractions from chickpea and oats have not been well characterized at the molecular level. Figures 3.4 and 3.5 show the Native electrophoretic patterns of chickpea, rice and oat protein fractions. Two major protein bands were observed from chickpea albumin (lane 1; C-Ab) with MW ranged between 669 kDa and 140 kDa; a minor band appeared below 60 kDa. These results suggest that chickpea albumin fractions contain proteins bands comparable to reported chickpea globulins, the 11S legumin (320-400 kDa) and 7S vicilin (145-190 kDa) (Casey et al., 1993), and 2S albumin (20-26 kDa) (Clemente et al., 2000). This indicates that chickpea water-soluble fraction could be a mixture of globulin and albumin proteins. The chickpea globulin fraction (lane 2; C-Gb) shows a major and a minor protein band corresponding mainly to

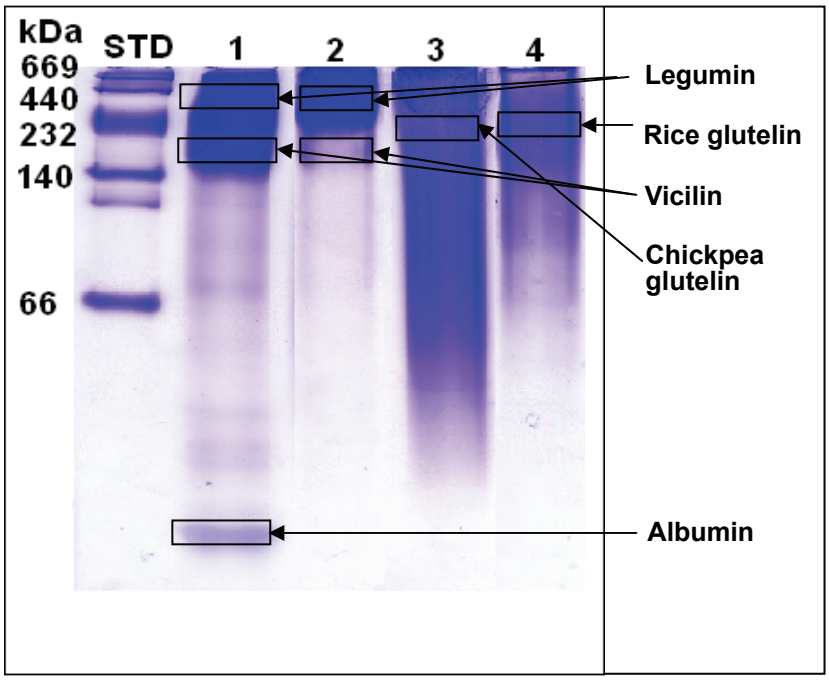


Figure 3.4: 8% Native-PAGE of chickpea protein isolates from sequential extractions. STD: Standard protein markers; (1) chickpea albumin fraction (C-Ab); (2) chickpea globulin fraction (C-Gb); (3) oat glutelin fraction (C-Gt); (4) rice glutelin fraction (R-Gt)

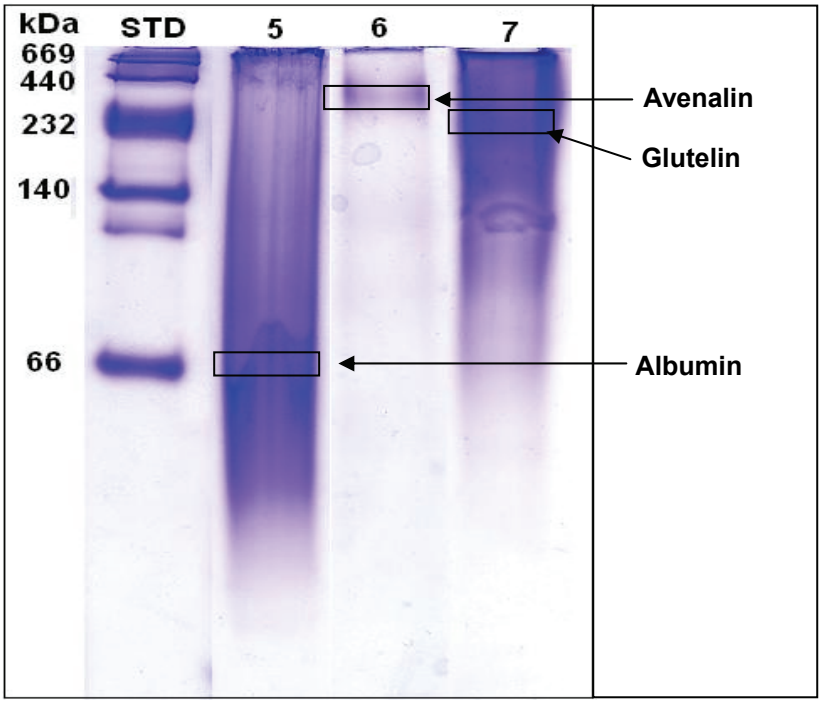


Figure 3.5: 8% Native-PAGE of oat protein isolates from sequential extractions. STD: Standard protein markers; (5) oat albumin fraction (O-Ab); (6) oat globulin fraction (O-Gb); (7) oat glutelin fraction (O-Gt)

chickpea 11S legumin and a minor presence of 7S vicilin; in comparison to fraction C-Ab, fraction C-Gb demonstrated more homogenous protein bands of globulins. In lanes 3 and 4, chickpea glutelin fraction (C-Gt) showed different and dispersed profiles of protein bands compared to rice glutelin fraction (R-Gt) with protein bands observed between MW 140 to 440 kDa. Oat albumin fraction (Figure 3.5, lane 5; O-Ab) exhibited one major band around 66 kDa. The oat globulin (known as avenalin) fraction (lane 6; O-Gb), which showed more homogenous protein band of globulin similar to chickpea fraction C-Gb, was also reported by Peterson (1978). In lane 6, the observed oat glutelin fraction was similar to that of rice glutelin fraction (lane 4; R-Gt).

3.3.2.2 SDS-PAGE Characterization of Chickpea, Oat and Rice Protein Fractions

Figures 3.6 and 3.7 show protein bands of fractions C-Ab, C-Gb, C-Gt, R-Gt, O-Ab, O-Gb, and O-Gt separated by 12% SDS-PAGE. The molecular weights of protein bands are summarized in the Table 3.2. 15 protein bands were observed from chickpea albumin fraction (lane 1; C-Ab) and the MWs of protein bands were 93.6 (1A), 63.6 (1B), 55.3 (1C), 51.0 (1D), 41.7 (1E), 38.9 (1F), 36.3 (G), 28.5 (1H), 25.9 (1I), 24.3 (1J), 21.0 (1K), 19.0 (1L), 16.5 (1M), 8.0 (N) and 5.3 (O) kDa; in lane 2 (Figure 3.6), 12 protein bands were labelled from chickpea globulin fraction (C-Gb), and the MWs of protein bands were 93.6 (2A), 63.6 (2B), 56.0 (2C), 51.0 (2D), 40.7 (2E), 38.0 (2F), 34.6 (2G), 29.2 (2H), 26.5 (2I), 25.2 (2J), 22.4 (2K) and 19.2 (2L) kDa; 5 protein bands were observed from chickpea glutelin fraction (C-Gt), and the MWs of protein bands were 200 (3A), 51.0 (3B), 38.1 (3C), 26.2 (3D) and 17.1 kDa (3E). As shown in Table 3.2, these results

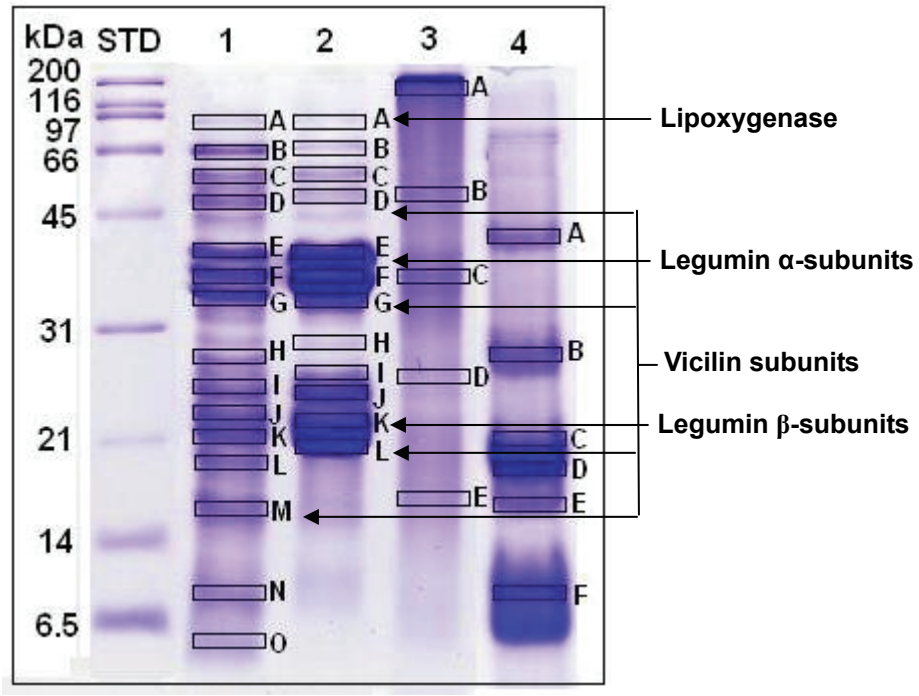


Figure 3.6: 12% SDS-PAGE of chickpea protein isolates from sequential extractions. STD: Standard protein markers; (1) chickpea albumin fraction (C-Ab); (2) chickpea globulin fraction (C-Gb); (3) oat glutelin fraction (C-Gt); (4) rice glutlein fraction (R-Gt)

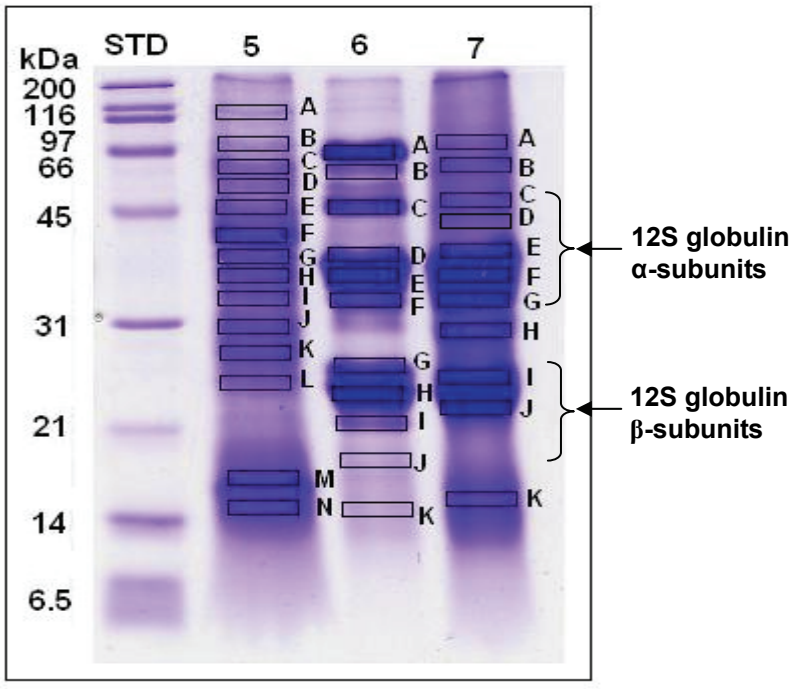


Figure 3.7: 12% SDS-PAGE of oat protein isolates from sequential extractions. STD: Standard protein markers; (5) oat albumin fraction (O-Ab); (6) oat globulin fraction (O-Gb); (7) oat glutelin fraction (O-Gt)

Table 3.2: Protein bands MWs from chickpea, rice, and oat protein fractions

Protein	MWs reported by previous researchers (KDa)	MWs of chickpea and rice protein fractions obtained in present work (Fig. 3.6)				Protein	MWs reported by previous researchers (KDa)	MWs of oat protein fractions obtained in present work (Fig. 3.7)		
		C-Ab	**C-Gb**	**C-Gt**	**R-Gt**			**O-Ab**	**O-Gb**	**O-Gt**
Lipoxygenase	92[1]	93.6 (1A)	96.3 (2A)			12S globulin	50-70[7]	69.6 (5B), 62.3, (5C)	65.8 (6A), 60.9 (6B)	69.3 (7A), 62.5 (7B)
Legumin (11S)	46.5-39.8[2] (α-subunit)	41.7 (1E), 38.9 (1F)	40.7 (2E), 38.0 (2F)			α-subunit	32-43[7]	45.7 (5E), 43.2 (5F), 39.3 (5G), 36.9 (5H), 33.7 (5I)	45.9 (6C), 39.8 (6D), 36.8 (6E), 33.2 (6F)	46.2 (7C), 44.8 (7D), 36.7 (7F), 33.5 (7G),
	24.3-25.3[2] (β-subunit)	25.9 (1I), 24.3 (1J)	26.5 (2I), 25.2 (2J)							
Pea vicilin (7S)	50[3]	51.0 (1D)	51.0 (2D)			β-subunit	19-25[7]	25.8 (5L), 17.2 (5M)	24.1 (6H), 22.0 (6I), 18.5 (6J)	26.2 (7I), 23.5 (7J)
	35[3]	36.3 (1G)	34.6 (2G)							
	33[3]					7S globulin	65[8]			
	19[3]	19.0 (1L)	19.8 (2L)				55[8]	55.2 (5D)		
	15[3]	16.5 (1M)				3S globulin	21[8]			
	13[3]						15[8]	15.0 (5N)	15.1 (6K)	
Rice glutelin	55[4]	55.3 (1C)	56.0 (2C)	51.0 (3B)	48.6 (4A)					
	38.5[5]			38.1 (3C)						41.5 (7E)
	21.4[5]	21.0 (1K)	22.4 (2K)		21.5 (4C), 19.0 (4D)					
	15.8[5]			17.1 (3E)	16.3 (4E)					16.1 (7K)
Albumin (2S)	12[6], 10[6]	8.0 (1N)								
Unknown		63.6 (1B), 28.5 (1H), 5.3 (1O)	63.6 (2B), 29.2 (2H)	200 (3A), 26.2 (3D)	8.0 (4F), 28.6 (4B)			98.2 (5A), 30.9 (5J), 28.5 (5K)	27.8 (6G)	30.7 (7H)

1: Clemente et al., 2000; 2: Sánchez-Vioque et al., 1999; 3: Guéguen, 1991; 4: Takaiwa et al., 1999; 5: Agboola et al., 2005; 6: Vioque et al., 1999; 7: Lásztity, 1996; 8: Burgess et al., 1983

suggest that protein bands 1A and 2A are comparable to the chickpea lipoxygenase (92 kDa) (Clemente et al., 2000); protein bands 1E, 1F, 1I, 1J, 2E, 2F, 2I, and 2J are likely the legumin α- and β-subunits (Chang et al., 2009; Sánchez-Vioque et al., 1999); protein bands 1D, 1G, 1L, 1M, 2D, 2G, and 2L are comparable to 7S pea vicilin (Guéguen, 1991); protein bands 1C, 1K, 2C, and 2K showed similar MWs to rice glutelin (Agboola et al., 2005; Takaiwa et al., 1999); protein band 1N with MW 8.0 kDa could be a polypeptides from 2S albumin (Vioque et al., 1999). The SDS-PAGE electrophoretic patterns of albumin and globulin fractions (C-Ab and C-Gb) agree with other studies on chickpea albumin and globulin fractions (Kaur & Singh, 2007; Liu et al., 2008). In lane 3, most of the observed protein bands (3B, 3C, and 3E) from fraction C-Gt are comparable to rice glutelin (Agboola et al., 2005; Takaiwa et al., 1999). Protein bands of chickpea glutelin fraction were different from protein bands of rice glutelin fraction (R-Gt) as shown in lane 4 (Figure 3.6).

For oat protein fractions, 14 protein bands were observed from oat albumin fraction (lane 5; O-Ab) and the MWs of protein bands were 98.2 (5A), 69.6 (5B), 62.3 (5C), 55.2 (1D), 45.7 (5E), 43.2 (5F), 39.3 (5G), 36.9 (5H), 33.7 (5I), 30.9 (5J), 28.5 (5K), 25.8 (5L), 17.2 (5M), and 15 (5N) kDa; 11 protein bands were observed from oat globulin fraction (lane 6; O-Gb) and the MWs of protein bands were 65.8 (6A), 60.9 (6B), 45.9 (6C), 39.8 (6D), 36.8 (6E), 33.2 (6F), 27.8 (6G), 24.1 (6H), 22.0 (6I), 18.5 (6J), and 15.1 (6K) kDa; 11 protein bands were observed from oat glutelin fraction (lane 7; O-Gt) and the MWs of protein bands were 69.3 (7A), 62.5 (7B), 46.2 (7C), 44.8 (7D), 41.5 (7E), 36.7 (7F), 33.5 (7G), 30.7 (7H), 26.2 (7I), 23.5 (7J), and 16.1 (7K) kDa. Lásztity (1996) reported that the oat 12S globulin with MW ranging from 50-70 kDa and the comparable protein bands

(5B, 5C, 6A, 6B, 7A and 7B) were observed in all three oat protein fractions as 12S globulin; protein bands 5E, 5F, 5G, 5H, 5I, 6C, 6D, 6E, 6F, 7C, 7D, 7F and 7G were comparable to the α-subunit (32-43 kDa) from oat 12S globulin along with the protein bands 5L, 5M, 6H, 6I, 6J, 7I and 7J were comparable to β-subunits (19-25 kDa) from oat 12S globulin reported by the same author. Furthermore oat albumin fraction (O-Ab) and globulin fraction (O-Gb) showed protein bands 5D, and 5N and 6K that were in relation to 7S (55 kDa) and 3S (15 kDa) globulin, respectively (Burgess et al., 1983); the oat glutelin fraction showed similar protein bands (7E and 7K) to rice glutlein subunits (38.5 and 15.8 kDa) (Agboola et al., 2005).

3.3.2.3 2-DGE Characterization of Chickpea and Oat Protein Fractions

Chickpea 11S legumin shares similar molecular characteristics with oat 12S globulin (Burgess et al., 1983); structurally, the proteins resemble the rice glutelin as hexamers with α- and β-subunits (Section 2.2.2). Figure 3.8 shows the 2-DGE electrophoretic patterns of chickpea globulin fraction (gel A; C-Gb) and chickpea glutelin fraction (gel B; C-Gt); at least 90 protein spots were identified for fraction C-Gb and 66 protein spots were identified for chickpea glutelin fraction.

In gel A, most of protein spots from chickpea globulin fraction (C-Gb) were located between pH 4 to pH 8 with MWs between 37-75 kDa (protein spots 10-40); a group of acidic α-subunits and other α-subunits were observed in protein spots 41-46 (~pH 3.5-4.0) and 47-58 (~pH 5.5-8.0) with MW ranging from 30-37 kDa, and are comparable to protein bands 2 E (40.7 kDa), 2F (34.6 kDa) and 2G (34.6 kDa) in Figure 3.6. For chickpea β-subunits of globulin, the protein spots (60-75 and 80-88) were observed

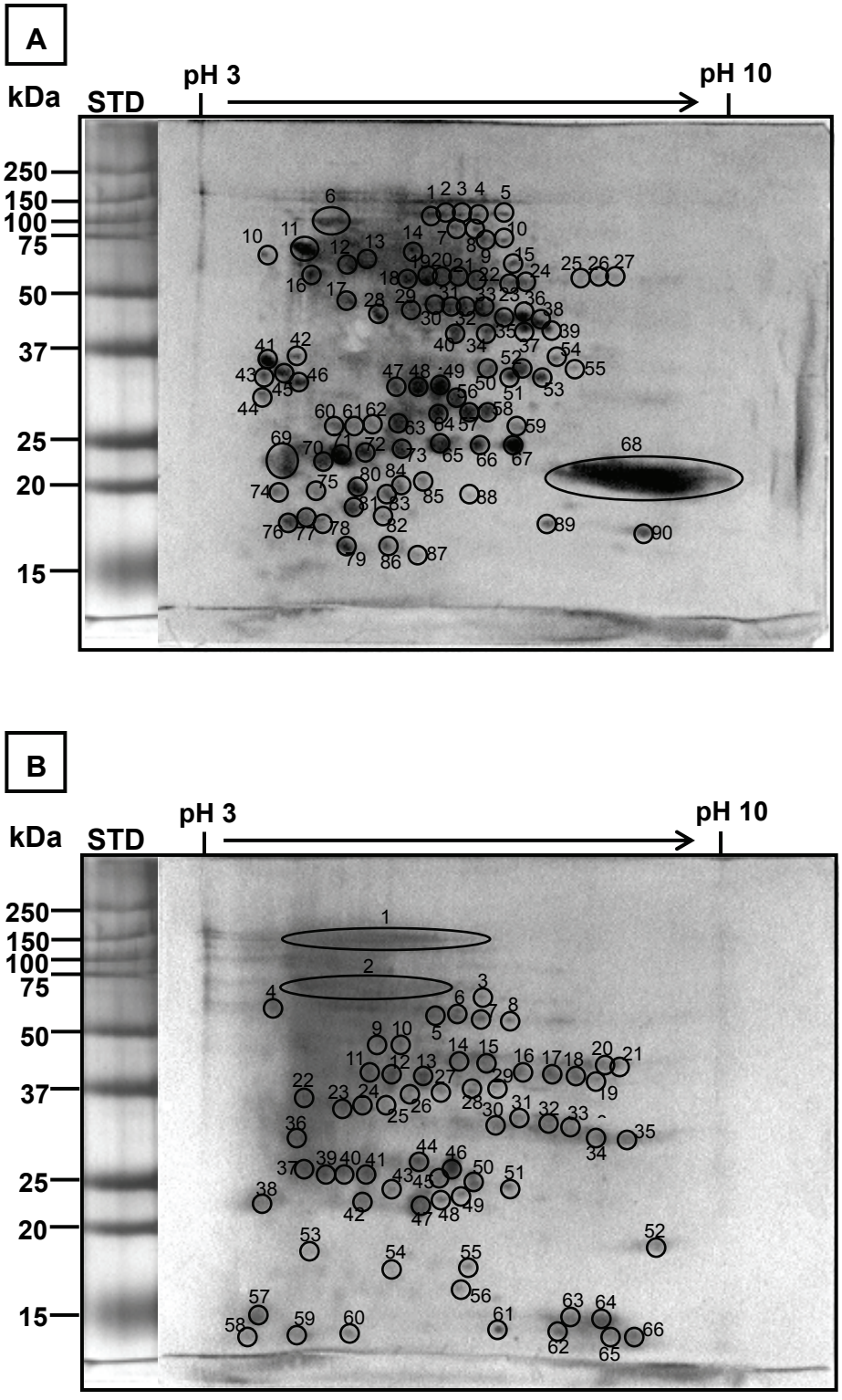


Figure 3.8: 2-DGE electrophoretic patterns of chickpea protein fractions. (A) globulin fraction (C-Gb), (B) glutelin fraction (C-Gt). First dimension: non-linear pH 3-10 IPG strip; second dimension: 12.5% SDS-PAGE gel. STD: Standard protein markers

between pH 3.5 to pH 7.0 with MWs ranging from 20 to 25 kDa or above, and are comparable to protein bands 2J (25.2 kDa), 2K (22.4 kDa) and 2L (19.8 kDa).

In gel B (Figure 3.8), two protein spots 1 and 2 were identified between pH 4.0 to pH 6.5 with relatively high MW (75-250 kDa); this agrees with the observation of protein band 3A (200 kDa; Figure 3.5) in 1D-SDS gel separation. In comparison to the chickpea glutelin fraction (lane 3; Figure 3.6), the protein spots 4-8 were detected between pH 4.0 to pH 7.0 with MW ranging from 50 to 70 kDa and are comparable to protein band 3B (51.0 kDa); the two groups of protein spots (9-35 and 36-51) can be chickpea glutelin α- and β-subunits with the MWs ranging from 37 to 50 kDa and around 25 kDa, respectively, and they are comparable to protein bands 3C (38.1 kDa) and 3D (26.2 kDa) (Figure 3.5); the other protein spots (52-66) showed scattered pH range between 3 to 10 with MWs between 15 to 20 kDa.

Figure 3.9 shows the 2-DGE electrophoretic patterns of oat globulin (gel C; O-Gb) and glutelin (gel D; O-Gt) fractions; at least 32 protein spots were identified for fraction O-Gb, and 25 protein spots were identified for oat glutelin fraction.

In gel C, the protein spots 3 and 4 gave protein clusters observed in pH 4.0 to pH 5.0 with MWs between 50 to 75 kDa; they are comparable to protein bands 6A (65.8 kDa) and 6B (60.9 kDa) as subunits of oat 12S globulin proteins in Figure 3.7. Protein spots 11-19 were identified between pH 4.0 to pH 8.0 with MWs ranging from 30 to 50 kDa, and are related to protein bands 6C (45.9 kDa), 6D (39.8 kDa), 6E (36.8 kDa) and 6F (33.2 kDa) as α-subunits of oat 12S globulin in Figure 3.7; protein spots 20-29 were observed between pH 4.5 to pH 7.0 with MWs ranging from 20 to 25 kDa, and are related to protein bands 6I (22.0 kDa) and 6H (24.1 kDa) as β-subunits of oat 12S globulin in

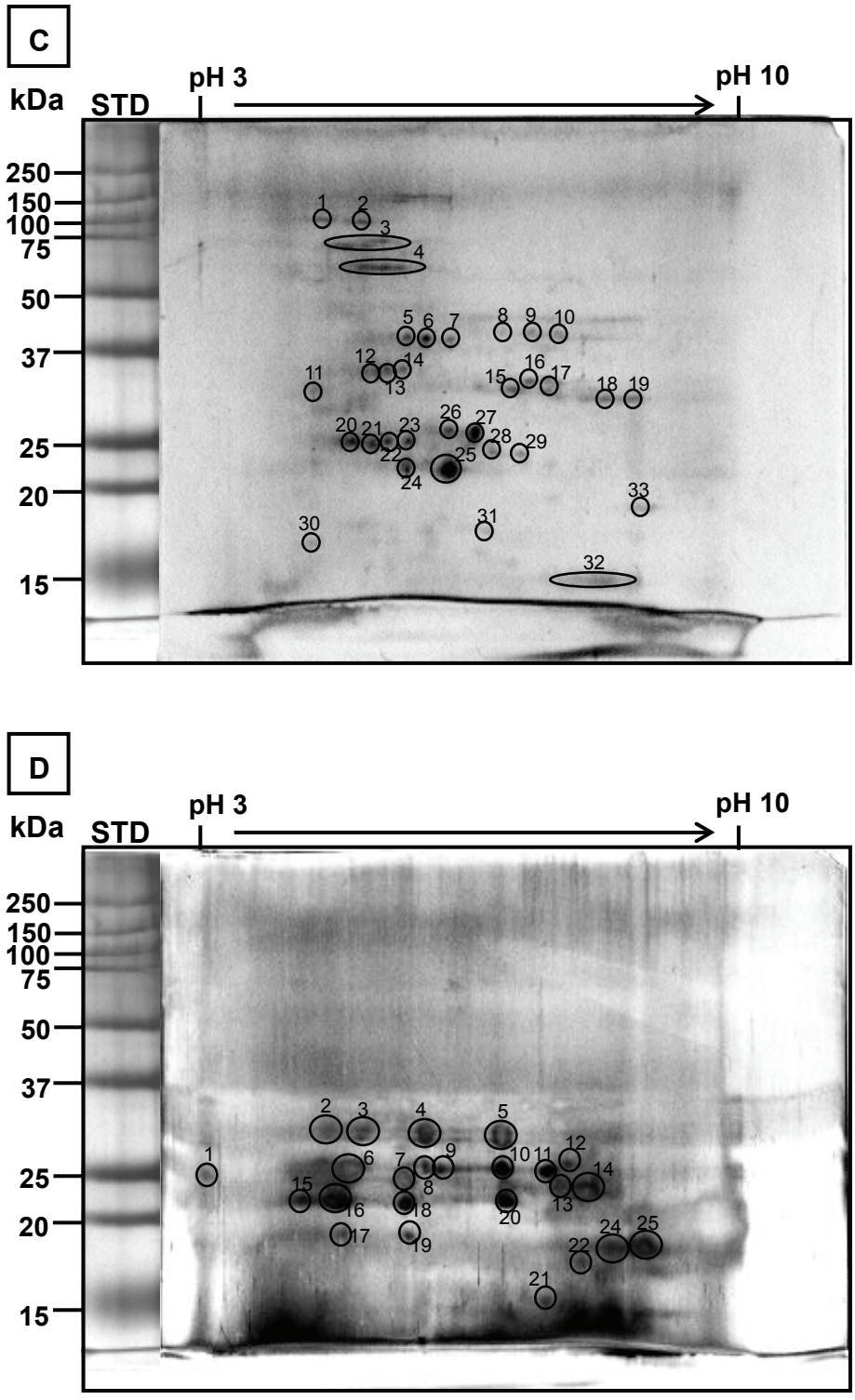


Figure 3.9: 2-DGE electrophoretic patterns of oat protein fractions. (C) globulin fraction (O-Gb), (D) glutelin fraction (O-Gt). First dimension: non-linear pH 3-10 IPG strip; second dimension: 12.5% SDS-PAGE gel. STD: Standard protein markers

Figure 3.7; these results are in agreement with previously reported 2-DGE results giving the separation of oat globulin α- and β-subunits between pH 4.5 to pH 7.0 and with MWs determined at 40 and 20 kDa, respectively (Walburg & Larkins, 1983).

In gel D, the oat glutelin fraction showed separation of protein spots observed at pH 3.0 to pH 8.5 with MWs between 15 to 37 kDa. Protein spots 2-5 are likely the α-subunits of oat glutelin fraction and protein spots 1 and 6-19 are likely the β-subunits of oat glutelin fraction.

3.3.3 RP-HPLC Characterization of Chickpea and Oat Protein Fractions

Figure 3.10 (A, B and C) and Figure 3.11 (D, E and F) show the RP-HPLC chromatograms of chickpea and oat albumin, globulin and glutelin fractions. The collected RP-HPLC fractions and their retention times (Rt) are summarized in Table 3.3; chickpea albumin fraction (C-Ab) gave four major RP-HPLC fractions: A1 (Rt: 19.3 min), A2 (Rt: 24.3 min), A3 (Rt: 25.5 min) and A4 (Rt: 26.4 min); chickpea globulin fraction (C-Gb) gave three major RP-HPLC fractions: B1 (Rt: 26.8 min), B2 (Rt: 27.3 min) and B3 (Rt: 28.6 min); chickpea glutelin fraction (C-Gt) gave one minor RP-HPLC fraction, C1 (Rt: 29.7 min) and one major RP-HPLC fraction, C2 (Rt: 34.8 min); oat albumin fraction (O-Ab) gave two major RP-HPLC fractions: D1 (Rt: 26.8 min) and D2 (Rt: 31.5 min); oat globulin fraction (O-Gb) gave three major RP-HPLC fractions: E1 (Rt: 27.7 min), E2 (Rt: 29.5 min) and E3 (Rt: 34.5 min); oat glutelin fraction (O-Gt) gave one minor RP-HPLC fraction, F1 (Rt: 29.9 min) and one major RP-HPLC fraction, C2 (Rt: 34.9 min). Chang et al. (2009) reported chickpea isoelectric precipitate and

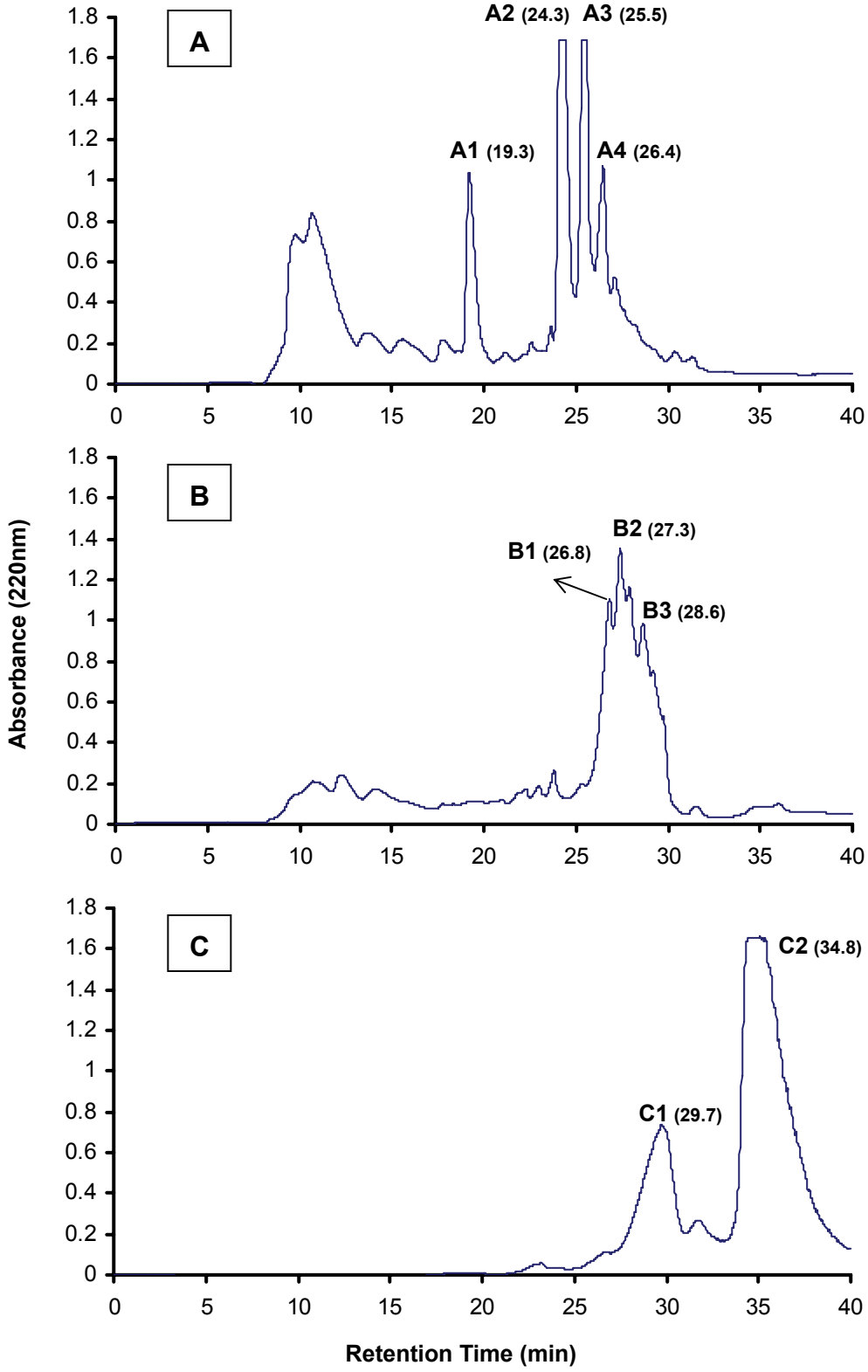


Figure 3.10: RP-HPLC chromatograms of chickpea protein fractions; A: chickpea albumin, B: chickpea globulin, and C: chickpea glutelin

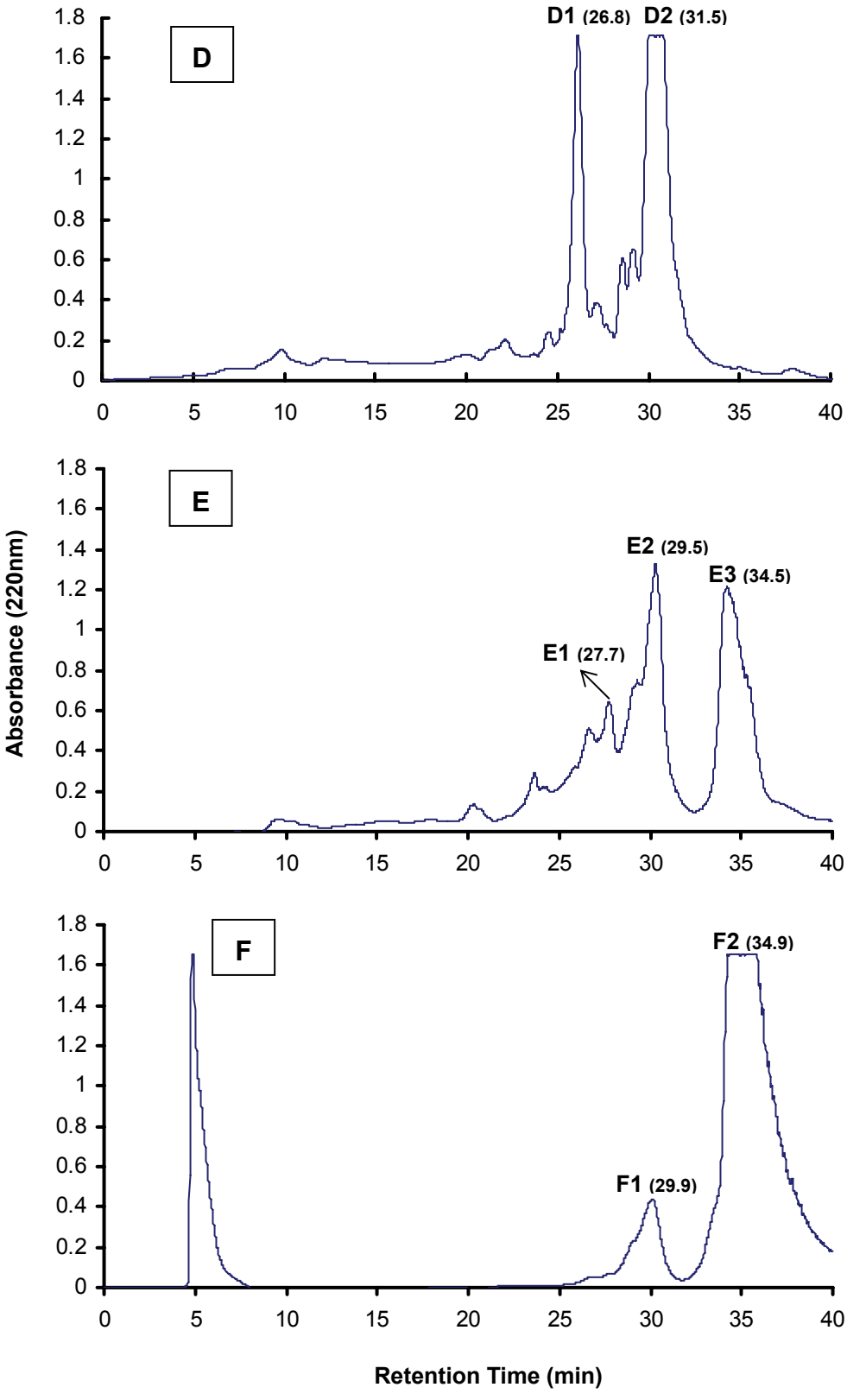


Figure 3.11: RP-HPLC chromatograms of oat protein fractions; D: oat albumin, E: oat globulin, and F: oat glutelin

cryoprecipitate gave four major RP-HPLC fractions from each isolate within the Rt: 20 to 40 min; Yust et al. (2003) reported the alcalse hydrolysates of chickpea legumin were separated to peptides by RP-HPLC C18 column; oat prolamin fractions were separated and identified using RP-HPLC (C18 column) separation and PAGE characterization to differentiate oat cultivars (Lookhart, 1985).

3.3.4 SDS-PAGE Characterization of Chickpea and Oat RP-HPLC Fractions

RP-HPLC fractions from chickpea and oat protein fractions were collected and submitted to 12% SDS-PAGE separation; Figures 3.12 and 3.13 show the SDS-PAGE electrophoretic patterns of chickpea and oat RP-HPLC fractions; the estimated MWs of RP-HPLC are presented in Table 3.3. RP-HPLC Fractions (A1-A4) from chickpea albumin fraction (C-Ab) gave five protein bands with MWs at 28.5 (A3a, A4a), 19.0 (A3b, A4b), 16.5 (A2a), 14.0 (A1a, A2b, A3c) and 12.5 (A4c) kDa; fractions A2, A3 and A4 showed protein bands with MWs 16.5 and 19.0 kDa that correspond to 7S pea vicilin subunits (Guéguen, 1991). Fractions B1-B3 from chickpea globulin fraction (C-Gb) showed six protein bands with MWs at 43.6 (B1a), 40.7 (B1b), 34.6 (B2a, B3a), 22.4 (B1c, B2b, B3b), 17.5 (B1b, B2c, B3c), 14.2 (B3d) and 12.8 (B1e, B2d) kDa; fraction B1 showed separated protein bands (B1a, B1b) with similar MWs that are related to α-subunits of chickpea legumin (Sánchez-Vioque et al., 1999); fractions B2 and B3 showed separated protein bands (B2a, B3a) with similar MWs that are related to 7S pea vicilin (Guéguen, 1991). Fractions C1 and C2 from chickpea glutelin fraction (C-Gt) showed four protein bands with MWs at 51.0 (C2a), 38.1 (C1a, C2b), 26.2 (C2c) and 17.1

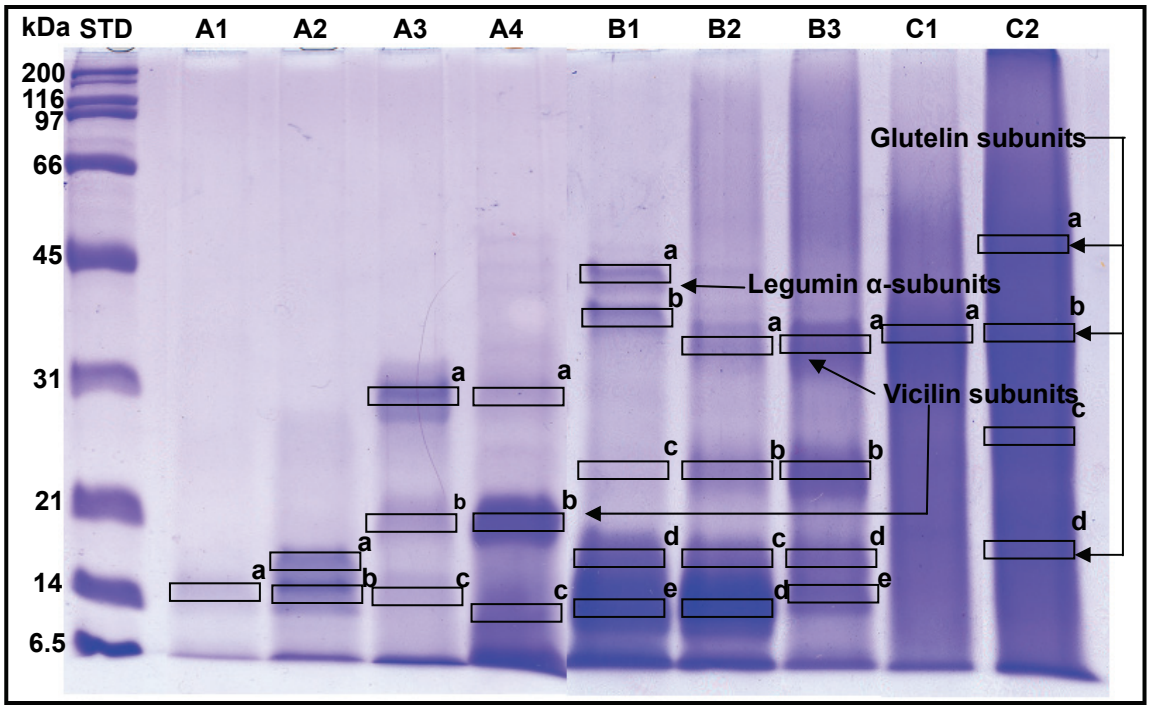


Figure 3.12: 12% SDS-PAGE of RP-HPLC fractions from chickpea protein fractions. STD: Standard protein markers; (A1-A4) albumin fractions; (B1-B3) globulin fractions; (C1 & C2) glutelin fractions

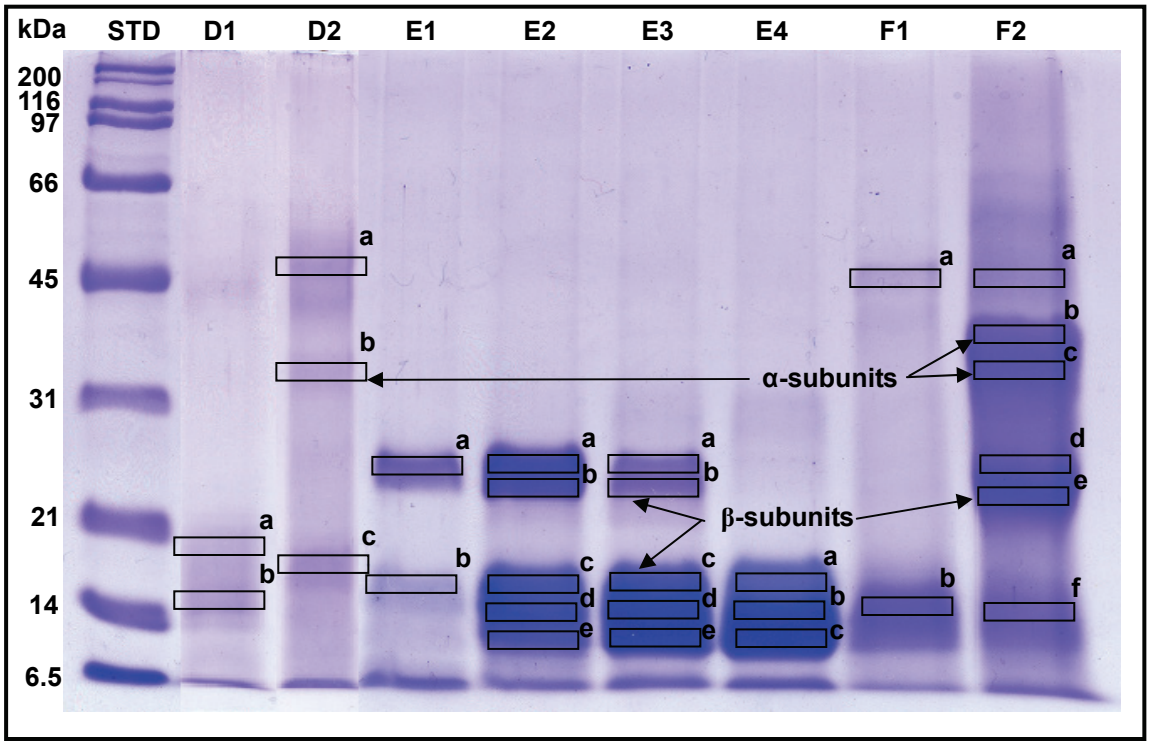


Figure 3.13: 12% SDS-PAGE of RP-HPLC fractions from chickpea protein fractions. STD: Standard protein markers; (D1 & D2) albumin fractions; (E1- E4) globulin fractions; (F1 & F2) glutelin fractions

Table 3.3: Summary of estimated MWs of chickpea, and oat fractions obtained from RP-HPLC

Protein fraction	RP-HPLC Fractions	Rt (min)	Estimated MWs of protein bands from SDS-PAGE (kDa) (Fig. 12 & Fig 13)	Reported subunits (kDa)
C-Ab	A1	19.3	14.0 (A1a)	Legumin (11S) 46.5-39.8[1] 24.3-25.3[1]
	A2	24.3	16.5 (A2a), 14.0 (A2b)	
	A3	25.5	28.5 (A3a), 19.0 (A3b), 14.0 (A3c)	
	A4	26.4	28.5 (A4a), 19.0 (A4b), 12.5 (A4c)	Pea Vicilin (7S) 50[2] 35[2] 33[2] 19[2] 15[2] 13[2]
C-Gb	B1	26.8	43.6 (B1a), 40.7 (B1b), 22.4 (B1c), 17.5 (B1d), 12.8 (B1e)	
	B2	27.3	34.6 (B2a), 22.4 (B2b), 17.5 (B2c), 12.8 (B2d)	
	B3	28.6	34.6 (B3a), 22.4 (B3b), 17.5 (B3c), 14.2 (B3d)	
C-Gt	C1	29.7	38.1 (C1a)	Albumin (2S) 12[3] 10[3]
	C2	34.8	51.0 (C2a), 38.1 (C2b), 26.2 (C2c), 17.1 (C2d)	
O-Ab	D1	26.8	19.8 (D1a), 15.0 (D1b)	12 globulin 50-70[4] α-subunits 32-43[4] β-subunits 19-25[4]
	D2	31.5	45.7 (D2a), 33.7 (D2b), 19.0 (D2c)	
O-Gb	E1	24.1	27.8 (E1a), 18.5 (E1b)	
	E2	27.7	27.8 (E2a), 24.1 (E2b), 18.5 (E2c), 15.1 (E2d), 12.6 (E2e)	
	E3	29.5	27.8 (E3a), 24.1 (E3b), 18.5 (E3c), 15.1 (E3d), 12.6 (E3e)	7S globulin 65[5] 55[5]
	E4	34.5	18.5 (E4a), 15.1(E4b), 12.6 (E4c)	
O-Gt	F1	29.9	16.1 (F1a)	3S globulin 21[5] 15[5]
	F2	34.9	41.5 (F2a), 36.7 (F2b), 33.5 (F2c), 26.2 (F2d), 23.5 (F2e), 16.1 (F2f)	

1: Sánchez-Vioque et al., 1999; 2: Guéguen, 1991; 3: Vioque et al., 1999; 4: Lásztity, 1996; 5: Burgess et al., 1983

(C2d) kDa; these two fractions showed protein bands (C2a, C1a, C2b, C2d) with similar MWs that are related to subunits of rice glutelin (51.0, 38.1 and17.1 kDa) as shown in Table 3.2 (Agboola et al., 2005; Takaiwa et al., 1999).

Fractions D1 and D2 from oat albumin fraction (O-Ab; Figure 3.13) gave four protein bands with MWs at 45.7 (D2a), 33.7 (D2b), 19.8 (D1a), 19.0 (D2c) and 15 (D1b) kDa; the separated protein band (D2a) showed similar MWs that are related to α-subunits of oat 12S globulin while the separated protein bands (D1b) is related to oat 3S globulin in Table 3.3 (Lásztity, 1996). Fractions E1-E4 from oat globulin fraction (O-Gb) gave five protein bands with MWs at 27.8 ((E1a, E2a, E3a), 24.1 (E2b, E3b), 18.5 (E1b, E2c, E3c, E4a), 15.1 (E2b, E3b,E4b) and 12.6 (E2e, E3e, E4c) kDa; the separated protein bands (E2b, E3b, E1b, E2c, E3c, E4a) showed similar MWs that are related to β-subunits of oat 12S globulin while protein bands (E2b, E3b, E4b) showed similar MWs that are related to oat 3S globulin (15.1 kDa) in Table 3.3 (Lásztity, 1996). Fractions F1 and F2 from oat glutelin fraction (O-Gt) gave six protein bands with MWs at 41.5 (F2a), 36.7 (F2b), 33.5 (F2c), 26.2 (F2d), 23.5 (F2e) and 16.1 (F1a, F2f); the separated protein bands (F2b, F2c) showed similar MWs that are related to α-subunits of oat 12S globulin, and the separated protein bands (F2d, F2e) showed similar MWs that are related to β-subunits of oat 12S globulin shown in Table 3.3 (Lásztity, 1996); the separated protein bands (F2a, F1a, F2f) showed similar MWs that are related to rice glutelin subunits in Table 3.2 (Agboola et al., 2005).

3.3.5 Characterization of Chickpea Protein Hydrolysates

Figure 3.14 shows the degree of tryptic hydrolysis of chickpea albumin, globulin and glutelin hydrolysates with degree of hydrolysis (DH) of 22.8%, 28.6% and 28.8%, respectively; the DH of chickpea globulin and glutelin hydrolysates is comparable to the DH (27%) of alcalase hydrolysate from chickpea legumin reported by Yust et al. (2003); however, Barbana & Boye (2010) reported relatively high DH (~78%) of chickpea protein hydrolysates digested by alcalase / flavourzyme, and DH (~40%) digested by papain. Figure 3.15 shows that the separated protein bands of chickpea protein hydrolysates (lane Ab-H, Gb-H and Gt-H) with lower MWs (below 35 kDa for Ab-H and Gb-H; below 21 kDa for Gt-H) compared to unhydrolyzed protein fractions; most protein bands (above 31 kDa) from chickpea albumin and globulin fractions (C-Ab and C-Gb) were hydrolyzed in particular the legumin α-subunits and some of the legumin β-subunits from chickpea globulin fractions.

3.4 Conclusion

The results of chickpea and oat total soluble protein yields suggest the sum of protein yields from albumin, globulin and glutelin fractions about 70%. Native- and SDS-PAGE results show similarities of chickpea and oat globulin fractions (C-Gb, O-Gb) such as similar protein bands observed from Native-PAGE (Figures 3.4 and 3.5) and comparable MWs of their α- and β-subunits (Table 3.2) from SDS-PAGE; oat protein fractions (O-Ab, O-Gb and O-Gt) demonstrate more protein bands with MWs related to α- and β-subunits of 12S globulin as compared to chickpea protein fractions (C-Ab, C-Gb and C-Gt). Chickpea and oat protein fractions were separated at Rt: 20-40 min by

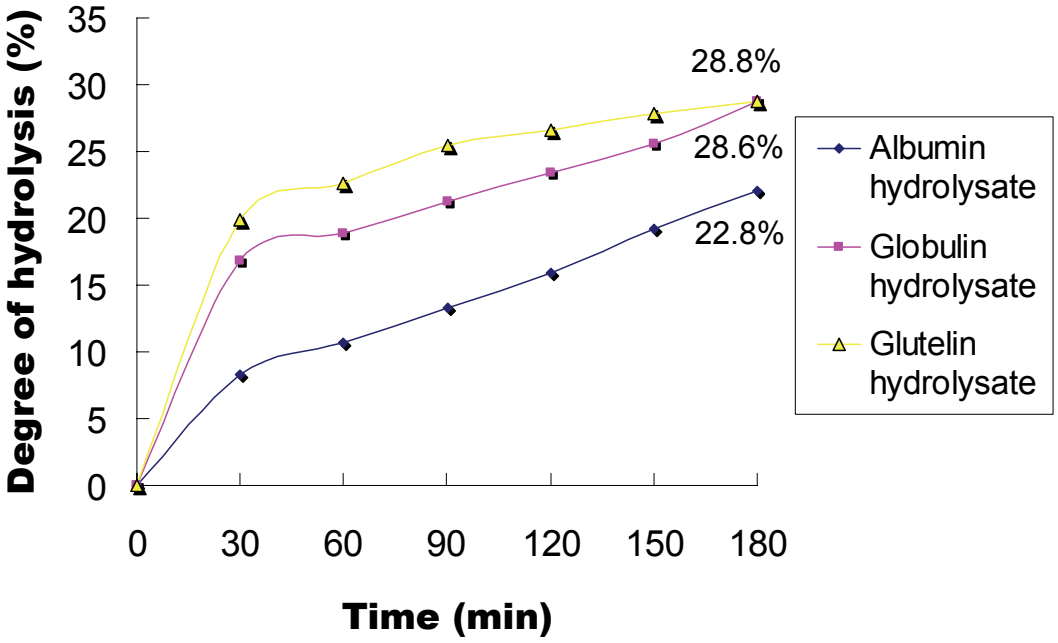


Figure 3.14: Degree of tryptic hydrolysis of chickpea albumin (◆), globulin (■)and glutelin (△) hydrolysates

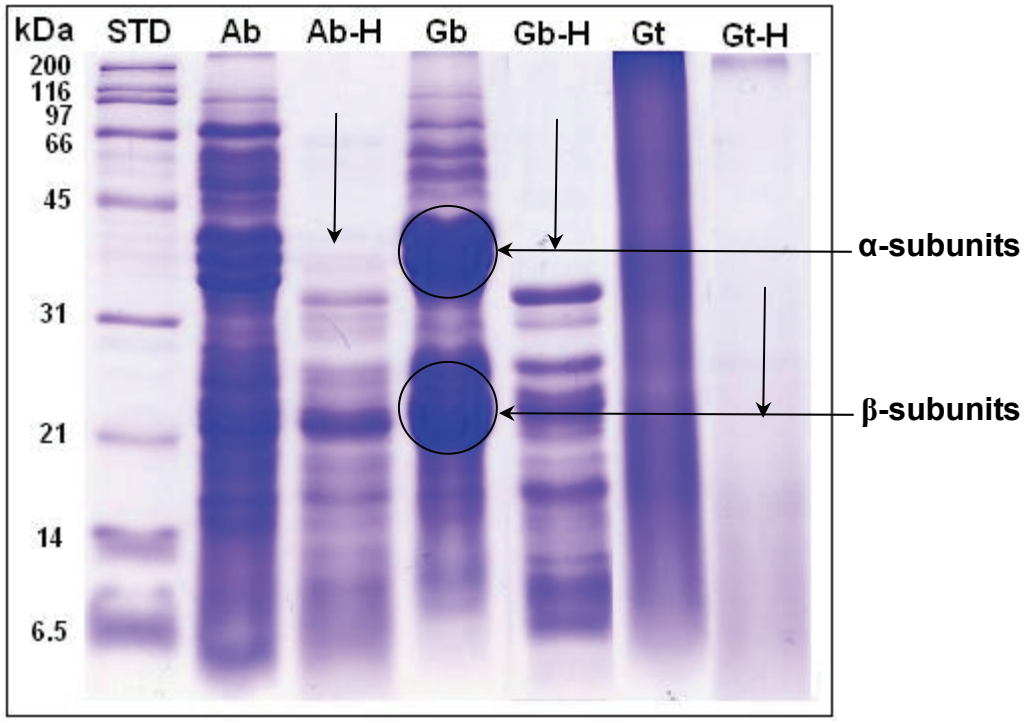


Figure 3.15: 12% SDS-PAGE electrophoretic pattern of chickpea protein fractions and hydrolysates. STD: Standard protein markers; (Ab) albumin fraction; (Ab-H) albumin hydrolysate; (Gb) globulin fraction; (Gb-H) globulin hydrolysate; (Gt) glutelin fraction; (Gt-H) glutelin hydrolysate

RP-HPLC; the results of SDS-PAGE of RP-HPLC fractions from chickpea and oat protein fractions show separated protein bands with MWs related to subunits of chickpea legumin, pea vicilin, oat globulins and rice glutelins. Chickpea albumin, globulin and glutelin fractions were hydrolyzed by trypsin enzyme.

CHAPTER 4

IDENTIFICATION OF PARTIAL TRYPTIC PEPTIDE SEQUENCES FROM CHICKPEA, OAT AND RICE PROTEIN FRACTIONS USING PROTEOMIC TECHNIQUES

4.1 Justification

There is relatively little information about the amino acid sequences of tryptic peptides of chickpea and oat protein fractions (albumin and glutelin) available. Several plant seeds (wheat, coffee bean, hazelnut, soybean and sesame) have been studied using proteomics (Chapter 2; Section 2.5.1). In this chapter, proteomic techniques, including 1D in-gel trypsin digestion, liquid chromatography electrospray ionization tandem mass spectrometry (LC-ESI-MS/MS), data analysis (MassLynx and ProteinLynx softwares) and on-line Mascot MS/MS ion search were used to partially identify the tryptic peptide sequences of chickpea, rice and oat protein fractions; the BLAST (Basic Logic Alignment Search Tool) analysis was used to align the protein sequences and to study the sequence similarity. In addition, the protein sequences from the identified chickpea legumin and provicilin precursor were examined with various proteases *in silico* to predict the ACE-inhibitory peptides by BIOPEP analysis.

4.2 Materials and Methods

4.2.1 Materials

The protein sequences of chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402) and provicilin precursor (NCBI accession number: gi|82173888) were available at NCBI (http://www.ncbi.nlm.nih.gov/guide/) and used for the BIOPEP analysis.

4.2.1.1 Experimental Approach

Figure 4.1 illustrates the schematic workflow chart for experimental proteomic approach to identify tryptic peptides fragments from chickpea, oat and rice protein fractions prepared in Sections 3.2.2.1 and 3.2.2.2; separated protein bands (1A-1O from chickpea albumin fraction (C-Ab); 2A-2L from chickpea globulin fraction (C-Gb); 3A-3E from chickpea glutelin fraction (C-Gt); 4A-4F from rice glutelin fraction (R-Gt); 5A-5N from oat albumin fraction (O-Ab); 6A-6K from oat globulin fraction (O-Gb); 7A-7K from oat glutelin fraction (O-Gt)) from 12% SDS-PAGE gels (Figure 3.6 for chickpea and rice protein fractions and Figure 3.7 for oat protein fractions; Section 3.3.2.2) were excised, destained and digested with trypsin enzyme; subsequently the tryptic peptide fragments were analyzed and identified by LC-ESI-MS/MS, ProteinLynx and MassLynx analysis, and Mascot MS/MS ion search.

4.2.2 Gel Destain and In-Gel Tryptic Digestion

The gel destain and in-gel trypsin digestion protocols are complied from the methods reported by Rosenfeld et al. (1992) and Shevchenko et al. (1996). Selected protein bands

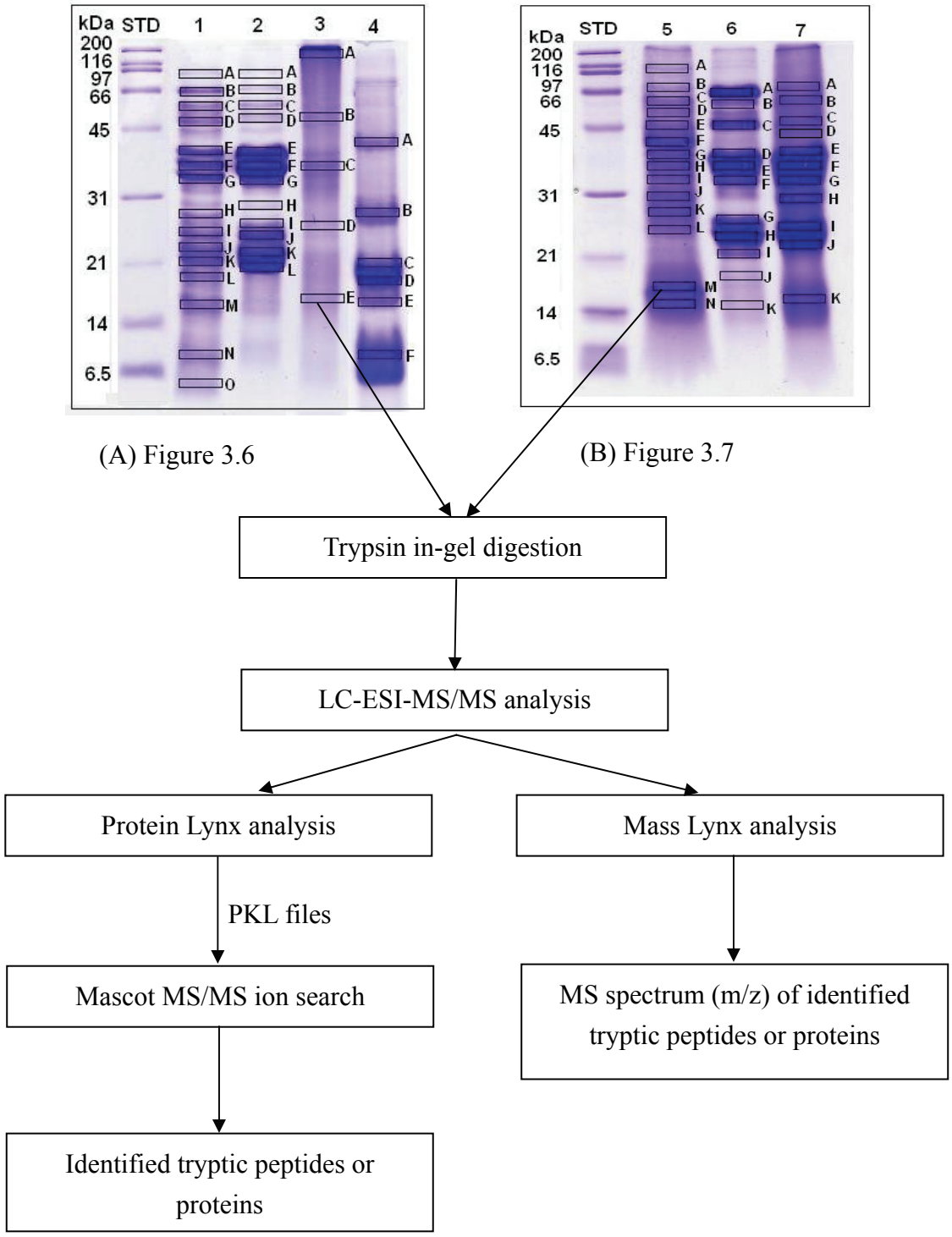


Figure 4.1: Experimental proteomic approach for identifying partial tryptic peptide sequences of chickpea, oat and rice protein fractions; (A) & (B): STD (standard protein markers), lane 1 (chickpea albumin fraction; C-Ab), lane 2 (chickpea globulin fraction; C-Gb), lane 3 (chickpea glutelin fraction; C-Gt), lane 4 (rice glutelin fraction; R-Gt), lane 5 (oat albumin fraction; O-Ab), lane 6 (oat globulin fraction; O-Gb), lane 7 (oat glutelin fraction; O-Gt)

of 1D SDS-APGE gels (Figures 3.6 and 3.7) from chickpea, rice and oat protein fractions were excised manually using 250 µL pipette tips then transferred to sterilized Eppendorf tubes (1.5 mL). The gel pieces were destained twice with 200 µL 100 mM NH_4HCO_3/50% acetonitrile for 2 × 45 min at 37 ℃. To break disulfide linkages in the proteins, the gel pieces were treated with 50 µL 10 mM DTT in 50 mM NH_4HCO_3 and incubated for 30 min at 56 ℃ and centrifuged (13,000 rpm, 5 min). After removal of supernatant, the gel pieces were alkylated with 100 µL 55 mM iodoacetamide in 50 mM NH_4HCO_3 and incubated for 10 min at room temperature. The gel pieces then dehydrated and rehydrated with 100 µL 100% CH_3CN and 50 µL 50 mM NH_4HCO_3; the dehydration step was repeated with CH_3CN and the gel pieces were air-dried at room temperature. For trypsin (modified porcine trypsin, sequencing grade, Promega, Madison, WI) digestion, 30 µL of trypsin working solution (20 µg trypsin/3 mL 50 mM NH_4HCO_3) was added to each sample including a blank sample (with non-protein gel pieces from SDS-PAGE) as a negative control and samples were incubated for 4 h at 37 ℃.

After incubation, the samples were centrifuged (13,000 rpm, 1 min) and the supernatants were transferred to Eppendorf tubes; 30 µL of extraction buffer (90% CH_3CN/0.5 M Urea) was added to the remaining gel pieces and supernatants were recovered by centrifugation (13,000 rpm, 1 min). The supernatants were combined with the respective original supernatants. This gel extraction procedure was repeated twice. The supernatants were vacuum-dried at room temperature and the dried samples were resuspended in 60 µL 0.2% formic acid (Anachemia, N.Y.) and stored at -20 ℃ prior to LC-ESI-MS/MS analysis.

4.2.3 Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry (LC-ESI-MS/MS)

Sample solution (1 μL) of tryptic in-gel digests prepared in Section 4.2.2 was injected into the LC-ESI-MS/MS (Waters Micromass QTOF Ultima Global, Micromass, Manchester, UK); the hybrid mass spectrometer was coupled to a Waters CapLC system, operated at a flow rate of 6 μL/min. A splitter placed before the column delivered a final flow rate of 0.3 μL/min. The guard column comprised Waters Symmetry 300 NanoEase C18, 5 μm; the separation was achieved using a Waters Atlantis dC18, 3 μm, 75 μm × 50 mm column. Operating procedures were based on the method described by Tezcucano Molina et al. (2007) with some modifications. A binary gradient of solvent B (acetonitrile: 0.1% formic acid) and solvent A (water: 0.1% formic acid) was increased from 5 to 50% in 25 min. After LC separation, elutes passed a nanoflow electrospray source, operated in positive ionization mode (+ESI), at 3.80 kV; source temperature was 80 ℃, desolvation temperature was 150 ℃. The TOF was monitored at an acceleration voltage of 9.1 kV, a cone voltage of 100 V, and a collision energy of 10 eV. For the MS survey mass range, m/z, was 400-1990 and for MS/MS was 50-1990, scanned continuously over the chromatographic run. The mass spectrometer was tuned and calibrated with [Glu]-Fibrinopeptide B (Sigma Chemicals; St. Louis MO). Instrumental control and data analysis were manipulated by using MassLynx V4.0 and ProteinLynx Global Server 2.1 software packages (Waters Corporation, United Kingdom, 2005).

4.2.4 Tandem MS Data Analysis of Proteins and Peptides Identification

MS/MS raw data were firstly converted to PKL files by ProteinLynx software using *de novo* sequencing parameter in combination with Mascot MS/MS ion search (Perkins et

al., 1999). MS/MS data were searched against NCBI viridiplantae (green plant) entries (Pruitt et al., 2005); search parameters used were carobxamidomethyl cysteine as fixed modification, oxidation as variable modification, ± 1.2 Da peptide mass tolerance and ± 0.6 Da fragment mass tolerance and the enzyme entry was set as trypsin. All peptide masses were obtained as monoisotopic masses.

The Mascot ion score is -10*Log (P), where P is the probability that the observed match is a random event. Individual ions scores > 45 indicate identity or extensive homology ($p<0.05$). Protein scores are derived from ions scores as a non-probabilistic basis for ranking protein hits (Matrix Science, London, United Kingdom).

4.2.4.1 Sequence Coverage (%) of Identified Tryptic Peptides

Based on Mascot MS/MS ion search results (Section 4.2.4), the sequence coverage of protein hits was expressed in percentage (%) indicating the sequence homology of identified tryptic peptides from chickpea, oat and rice protein fractions to corresponding protein hits.

4.2.5 Basic Local Alignment Search Tool (BLAST) Analysis of Protein Sequences Alignment

Identified protein sequences of chickpea, oat and rice protein fractions were aligned using the BLAST analysis which is available at NCBI website (*http://blast.ncbi.nlm.nih.gov/Blast.cgi*); two or more protein sequences in FASTAT format, a text-based format for representing either nucleotide sequences or peptide sequences, can be aligned, and generally a pairwise alignment will be displayed on

demand and a series of informative data can be generated such as scores, percentage of identities, positives and gaps (Altschul et al., 1997; Altschul et al., 2005).

4.2.6 BIOPEP Analysis

The prediction of the potential antihypertensive or ACE-inhibitory peptides from chickpea proteins was carried out using BIOPEP database (available at http://www.uwm.edu.pl/biochemia/index_en.php), (Iwaniak et al., 2005) and performed according to the approach reported by Cheung et al. (2009) with some modifications. Firstly, the protein sequences of chickpea legumin alpha and beta subunit [Cicer arietinum] (NCBI accession number: gi|6273402) and provicilin precursor (NCBI accession number: gi|82173888), acquired from NCBI database of protein sequences (http://www.ncbi.nlm.nih.gov/guide/) (Pruitt et al., 2005), were examined using "profiles of potential biological activity" tool of the BIOPEP analysis; the potential ACE-inhibitory peptides sequences within the sequences of chickpea legumin and provicilin precursor proteins were examined and matched up against the published peptides sequences of demonstrating either *in vitro* or *in vivo* ACE-inhibitory activity that is available in the BIOPEP database. Secondly, the protein sequences of chickpea legumin and provicilin precurcor were directed into the "enzymatic action" tool which is available in the BIOPEP database to predict the theoretical peptides sequences cleaved by 27 proteases (chymtryosin A, trypsin, pepsin, proteinase K, pancreatic elastase, prolyl oligopeptidase, V-8 protease, thermolysin, chymotrypsin C, plasmin, cathepsin G, clostripain, chymase, papain, ficain, leukocyte elastase, metridin, thrombin, pancreatic elastase Ⅱ, bromelain, glutamyl endopeptidase Ⅱ, oligopeptidase B, calpain, glycyl endopeptidase,

oligopeptidase F, proteinase P1, prolidase L. lactis s. cremoris H61) with specific cleaved locations. Thirdly, the theoretical peptides obtained from various proteases were submitted to "search for active fragments" option. Finally, a list of potential bioactive peptides was displayed and only peptides containing potential ACE-inhibitory activity were selected for further analysis; in addition, except the action of single enzymes, a combined action of pepsin, trypsin and chymotrypsin A enzymes was investigated to mimic the gastrointestinal digestion process.

4.3 Results and Discussion

4.3.1 Identification of Tryptic Peptide Sequences with Origin from Chickpea and Rice Protein Fractions

4.3.1.1 Identified Tryptic Peptide Sequences with Origin from Chickpea Legumin

Based on the results from Mascot MS/MS ion search (Section 4.2.4) for the in-gel tryptic digests of each protein bands from chickpea, rice and oat protein fractions, all the identified tryptic peptides from chickpea, rice and oat protein fractions were listed as double or triple charged peptides; the information including peptide sequence, protein origin, MWs of peptides and Mascot scores on these tryptic peptides were collected and organized for investigation.

Figure 4.2 illustrates how the representative doubly charged peptide, ALIVPQNFAIAAK (MW 1,354.79 Da) listed from Mascot MS/MS ion search with origin from chickpea legumin, from protein band 1J (chickpea albumin fraction, C-Ab; lane 1, Figure 3.6) was identified using the mass spectrum from LC-ESI/MS/MS analysis;

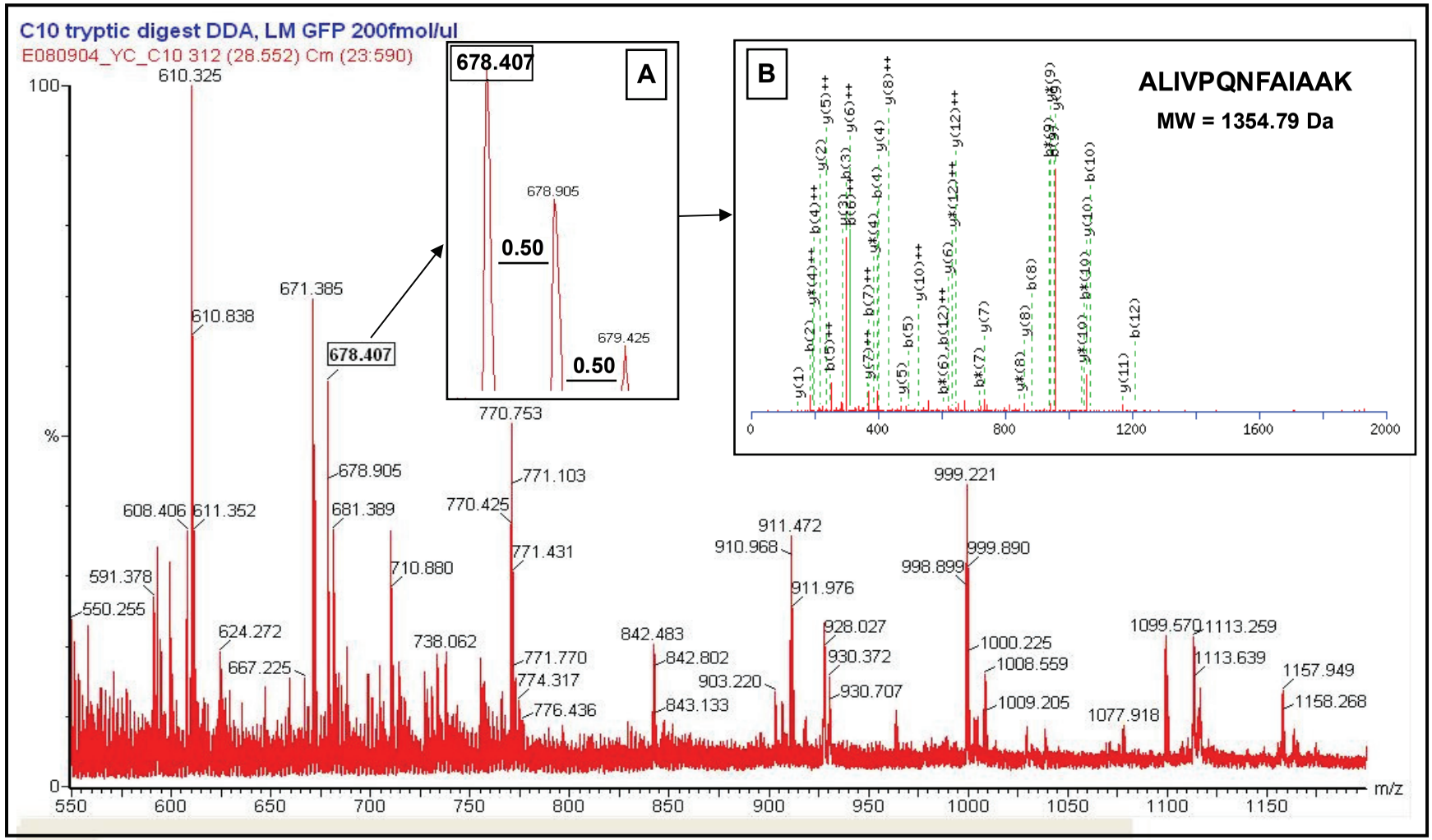


Figure 4.2: LC-ESI-MS spectra (m/z region 550 to 1200) of chickpea protein band 1J, m/z 678.407 signal in bold type. Insert A shows the identified doubly charged signal by the difference of 0.50 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide ALIVPQNFAIAAK, calculated MW 1354.79 Da

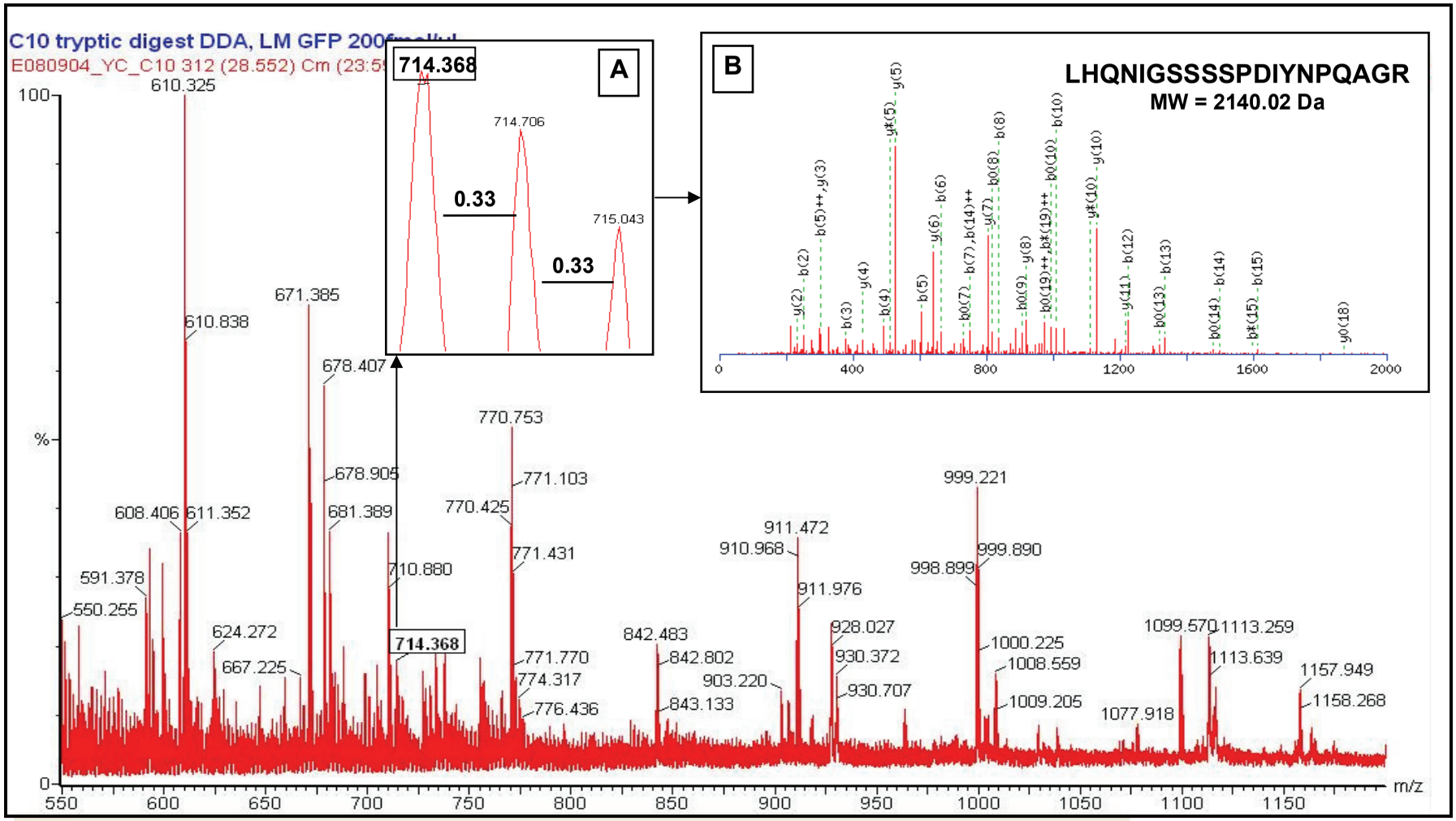


Figure 4.3: LC-ESI-MS spectra (m/z region 550 to 1200) of chickpea protein band 1J, m/z 714.368 signal in bold type. Insert A shows the identified triply charged signal by the difference of 0.33 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide LHQNIGSSSSPDIYNPQAGR, calculated MW 2140.02 Da

the observed signal with m/z 678.407 was doubly charged (difference of 0.5 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra. Similarly, Figure 4.3 represents a triply charged peptide LHQNIGSSSSPDIYNPQAGR (MW 2,140.02 Da) from the same protein band (1J); the observed signal with m/z 714.368 was triply charged (difference of 0.33 between adjacent signals as shown in insert A); the insert B represents the final result of identified peptide from LC-ESI-MS/MS spectra.

Table 4.1 summarizes the identified tryptic peptides of protein bands (Figure 3.6) from chickpea albumin (C-Ab, lane 1), globulin (C-Gb, lane 2) and glutelin (C-Gt, lane3) fractions with the origin from chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402; theoretical mass 56,216 Da). Structurally, chickpea legumin (~360 kDa) is a hexamer which is composed of six monomers (αβ subunits; 55-60 kDa) and each monomer is linked by a disulfide bond (Casey et al., 1993); the chickpea legumin α- and β-subunits (496 AAs) is composed of a signal peptide at the AA position 1-21, an acidic polypeptide chain (α-subunit) at AA position 22-311, and a basic polypeptide chain (β-subunit) at AA position 312-496; a disulfide bond is observed at AA position 31↔64 and a interchain (between α- and β-chain) is observed at AA position 107↔318 (Mandaokar & Koundal, 1997). In Table 4.1, 10 tryptic peptides from chickpea albumin, globulin and glutelin fractions (C-Ab, C-Gb and C-Gt, respectively) were identified as legumin α-subunit and 9 were identified as legumin β-subunit; most of the identified tryptic peptides with origin from chickpea legumin were present in all chickpea protein fractions except the peptides, RDFLEDALNVNR, NAMFVPHYNLNANSILYALK and NAMFVPHYNLNANSILYALKGR which were

Table 4.1: Identified tryptic peptides of SDS-PAGE protein bands (Figure 3.6) from chickpea albumin (lane 1), globulin (lane 2) and glutelin (lane 3) fractions with origin from chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402; theoretical mass 56,216 Da)

Peptide sequences[a]	MW (Da)	Identified in SDS-PAGE protein bands (Fig. 3.6)
Legumin α-subunit[b]		
IKSEGGLIETWNPSNK $_{45-60}$	1771.91	1C (55.3 kDa), 2C (56.0 kDa), 2F (38.0 kDa), 2G (34.6 kDa)
SEGGLIETWNPSNK $_{47-60}$	1530.73	1F (38.9 kDa), 2B (63.6 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 2J (25.2 kDa), 2L (19.2 kDa), 3C (38.1 kDa)
RFYLAGNHEQEFLR $_{180-193}$	1778.88	1F (38.9 kDa), 2C (56.0 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 2L (19.8 kDa), 3C (38.1 kDa)
FYLAGNHEQEFLR $_{181-193}$	1622.78	1C (55.3 kDa), 1F (38.9 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2G (34.6 kDa), 3B (51.0 kDa)
YQQEGSEEEENEGGNIFSGFK $_{194-214}$	2376.99	1C (55.3 kDa), 1F (38.9 kDa), 2E (40.7 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 3C (38.1 kDa)
YQQEGSEEEENEGGNIFSGFKR $_{194-215}$	2533.09	2B (63.6 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2G (34.6 kDa)
RDFLEDALNVNR $_{215-227}$	1460.73	1E (41.7 kDa)
DFLEDALNVNR $_{216-226}$	1304.63	1C (55.3 kDa), 1F (38.9 kDa), 1K (21.0 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 2H(29.2 kDa), 2I (25.2 kDa), 3A (200 kDa), 3B (51.0 kDa), 3C (38.1 kDa)
DFLEDALNVNRR $_{216-227}$	1460.73	1E (41.7 kDa), 1F (38.9 kDa), 2B (63.6 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 3B (51.0 kDa), 3C (38.1 kDa)
NEDEEKGAIVK $_{236-246}$	1230.60	1C (55.3 kDa), 1E (41.7 kDa), 1F (38.9 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2F (38.0 kDa), 2G (34.6 kDa), 3C (38.1 kDa)
Legumin β-subunit[c]		
LHQNIGSSSSPDIYNPQAGR $_{322-341}$	2140.02	1C (55.3 kDa), 1E (41.7 kDa), 1J (24.3 kDa), 1K (21.0 kDa), 2B (63.6 kDa), 2C (56.0 kDa), 2E (40.7 kDa), 2G (34.6 kDa), 2H(29.2 kDa), 2I (26.5 kDa), 2J (25.2 kDa), 2K (22.4 kDa), 3A (200 kDa), 3B (51.0 kDa), 3C (38.1 kDa)
FLKLSAEFGSLHK $_{355-367}$	1475.81	1J (24.3 kDa), 1K (21.0 kDa), 2C (56.0 kDa), 2K (22.4 kDa), 3A (200 kDa), 3E (17.1 kDa)
LSAEFGSLHK $_{358-367}$	1087.56	1C (55.3 kDa), 2J (25.2 kDa), 2K (22.4 kDa), 2L (19.8 kDa), 3A (200 kDa), 3C (38.1 kDa)
NAMFVPHYNLNANSILYALK $_{368-387}$	2308.16	1J (24.3 kDa), 1K (21.0 kDa)
NAMFVPHYNLNANSILYALKGR $_{368-389}$	2521.28	1J (24.3 kDa), 1K (21.0 kDa)
GNSVFDGELEAGR $_{400-412}$	1349.62	1K (21.0 kDa), 2I (26.5 kDa), 2L (19.8 kDa), 3A (200 kDa), 3E (17.1 kDa)
ALIVPQNFAIAAK $_{413-425}$	1354.79	1J (24.3 kDa), 1K (21.0 kDa), 2C (56.0 kDa), 2J (25.2 kDa), 2K (22.4 kDa), 3B (51.0 kDa), 3C (38.1 kDa)
SLSDRFSYVAFK $_{426-437}$	1418.71	1J (24.3 kDa), 1K (21.0 kDa), 2B (63.6 kDa), 2I (26.5 kDa), 2J (25.2 kDa), 2K (22.4 kDa), 3A (200 kDa), 3E (17.1 kDa)
FSYVAFK $_{431-437}$	860.44	1J (24.3 kDa), 1K (21.0 kDa), 2C (56.0 kDa), 2J (25.2 kDa), 2K (22.4 kDa), 2L (19.8 kDa), 3B (51.0 kDa), 3C (38.1 kDa)

[a] Amino acids are presented in single letter code within peptide sequences; [b] Legumin α-subunit: AAs position located between 22-311; [c] Legumin β-subunit: AAs position located between 312-496

found in chickpea albumin fraction (C-Ab; lane 1, Figure 3.6). Proteins bands 1C (55.3 kDa), 2B (63.6 kDa), 2C (56.0 kDa) and 3B (51.0 kDa) showed tryptic peptides present in both legumin α- and β-subunits; this suggests these protein bands can be the legumin parent (monomer) protein consisting of α-subunit (~33 kDa) and β-subunit (~20 kDa) (Mandaokar & Koundal, 1997). Protein bands 1F (38.9 kDa), 2F (38.0 kDa), 2G (34.6 kDa) showed tryptic peptides mostly from the legumin α-subunit; protein bands 1J (24.3 kDa), 1K (21.0 kDa), 2J (25.2 kDa), 2K (22.4 kDa) and 3E (17.1 kDa) showed tryptic peptides mostly from the legumin β-subunit; this suggests these protein bands (1F, 2F and 2G; 1J, 1K, 2J, 2K and 3E) can be the legumin α- and β-subunits, respectively; SDS-PAGE MWs of protein bands (1F, 2F; 2J) corresponded to MWs of legumin α- and β-subunits, respectively, reported by Sánchez-Vioque et al. (1999) (Section 3.3.2.2). Protein bands 1E (41.7 kDa), 2E (40.7 kDa), 2H (29.2 kDa), 2I (26.5 kDa), 2L (19.8 kDa) and 3C (38.1 kDa) showed tryptic peptides mostly from either the legumin α- or β-subunits but with minor presence of their counterparts of either α- or β-subunits.

Figure 4.4 shows the AAs sequence of chickpea legumin with the tryptic peptides from chickpea protein fractions (C-Ab, C-Gb and C-Gt) identified during the proteomic analysis; the solid line represents the theoretical peptides cleaved by trypsin and the broken line represents the experimental tryptic peptides identified from in-gel digestion. Theoretically 59 cutting sites at either arginine (R) or lysine (K) of legumin AAs sequence are available for tryptic hydrolysis; two cutting sites at AAs position 40-41 (KP) and 462-463 (RP) are not available due to tyrpsin does not react to praline at C-terminal; 25 cutting sites appeared in identified tryptic peptides derived from in-gel digestion.

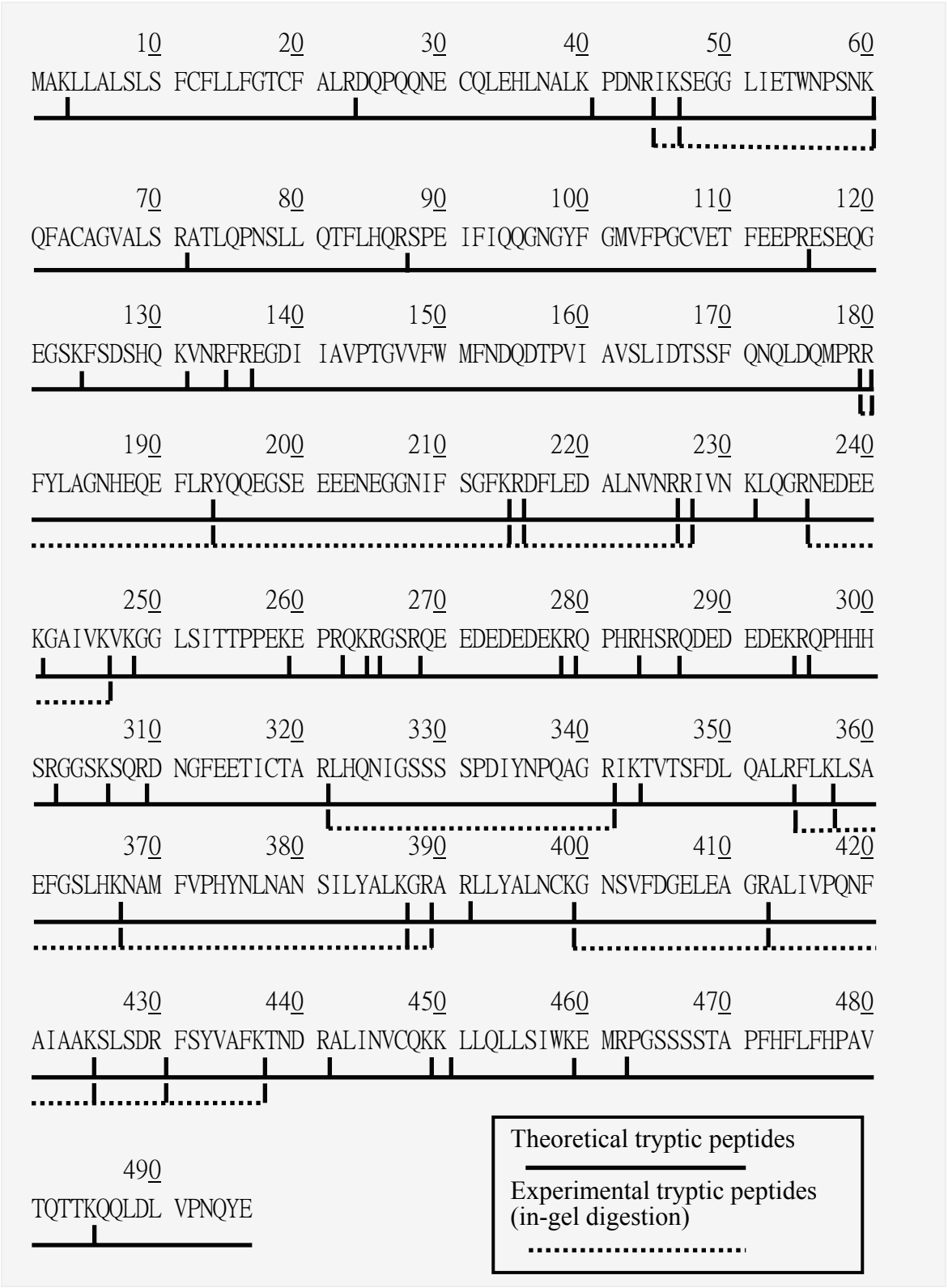


Figure 4.4: Amino acids sequence of identified tryptic peptides with origin from chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402; theoretical mass 56,216 Da)

Based on the AAs sequence, 155 AAs from identified tryptic peptides out of 496 total AAs contributes 31% sequence coverage.

4.3.1.2 Identified Tryptic Peptide Sequences with Origin from Chickpea Provicilin Precursor

Figures 4.5 and 4.6 give the representative examples of how the doubly and triply charged tryptic peptides of protein band 1H (28.5 kDa) from chickpea albumin fraction (C-Ab; lane 1, Figure 3.6) were identified using the MS spectra during the LC-ESI/MS/MS analysis, respectively; Figure 4.5 shows the peptide SVSSESEPFNLR (MW 1,350.64 Da) with the detected signal (m/z 676.309) which was doubly charged (difference of 0.5 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra; similarly, Figure 4.6 shows the peptide KSVSSESEPFNLR (MW 1,478.73 Da) with the detected signal (m/z 493.912) which was triply charged (difference of 0.33 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra.

Tryptic peptides identified with origin from chickpea provicilin precursor [Cicer arietinum] (NCBI accession number: gi|82173888; theoretical mass 51,390 Da) were only found in protein bands 1G (36.3 kDa), 1H (28.5 kDa) and 1M (16.5 kDa) from chickpea albumin fraction (C-Ab); these tryptic peptides are listed in Table 4.2 and two peptides sequences, $EQIEELSK_{224\text{-}231}$ and $SRNPIYSNK_{252\text{-}260}$, appeared frequently in protein bands 1G, 1H and 1M.

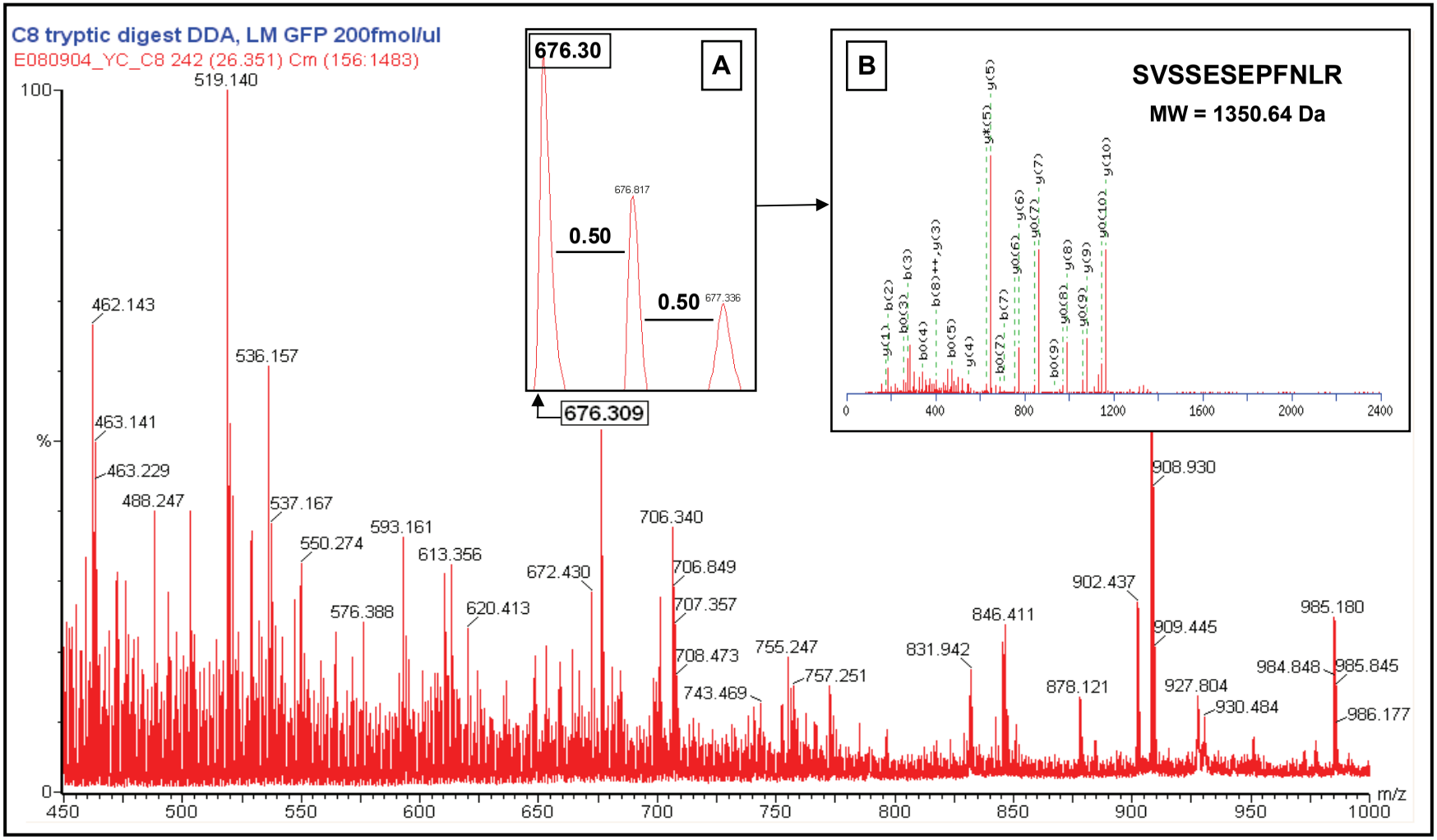


Figure 4.5: LC-ESI-MS spectra (m/z region 450 to 1000) of chickpea protein band 1H, m/z 676.309 signal in bold type. Insert A shows the identified doubly charged signal by the difference of 0.50 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide SVSSESEPFNLR, calculated MW 1350.64 Da

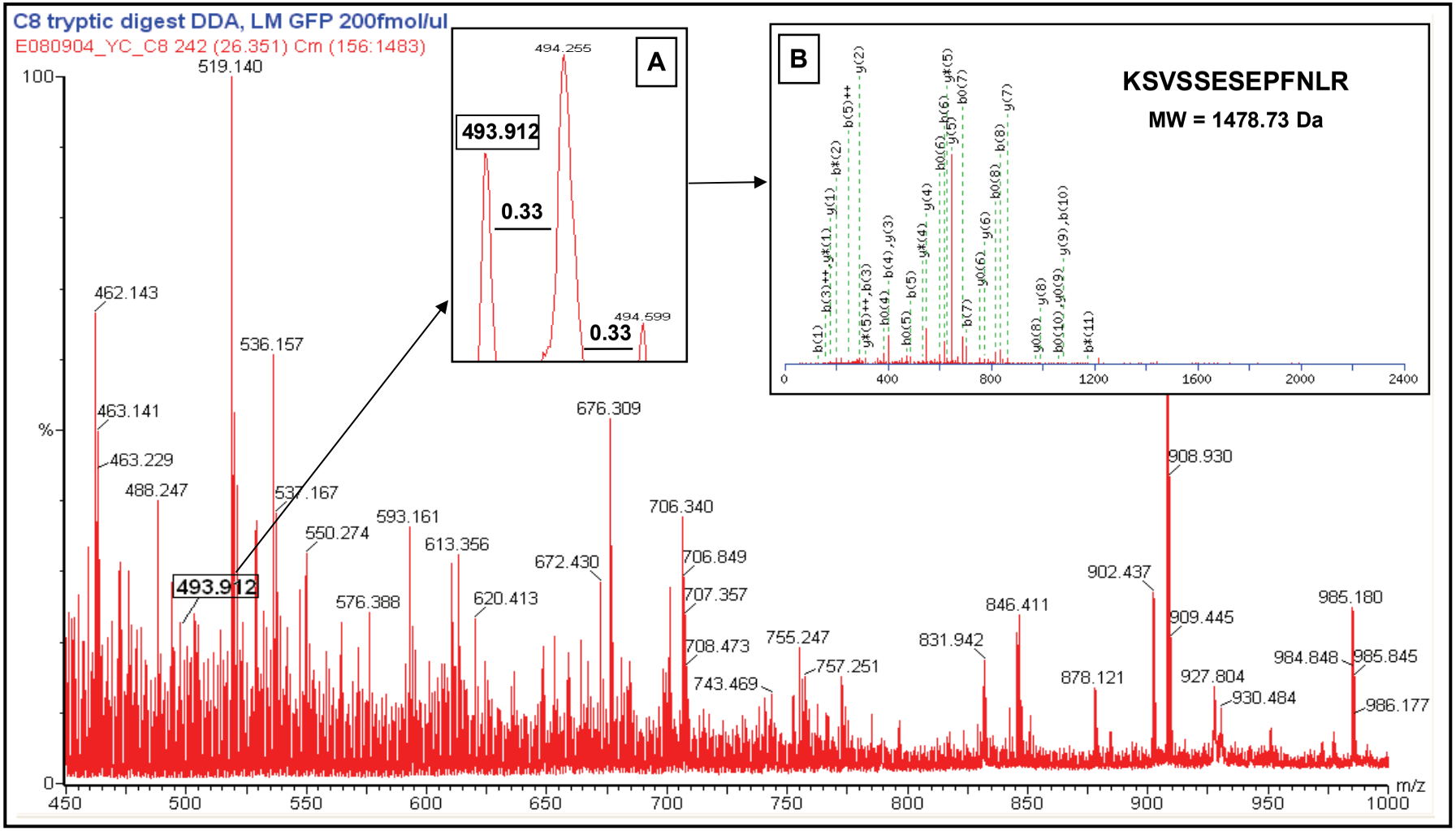


Figure 4.6: LC-ESI-MS spectra (m/z region 450 to 1000) of chickpea protein band 1H, m/z 493.912 signal in bold type. Insert A shows the identified triply charged signal by the difference of 0.33 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide KSVSSESEPFNLR, calculated MW 1478.73 Da

Table 4.2: Identified tryptic peptides of SDS-PAGE protein bands (Figure 3.6) from chickpea albumin (lane 1) fraction with origin from provicilin precursor [Cicer arietinum] (NCBI accession number: gi|82173888; theoretical mass 51,390 Da)

Peptide sequences[a]	MW (Da)	Identified in SDS-PAGE protein bands (Fig. 3.6)
VLLEEQEQKPK $_{186-196}$	1339.73	1H (28.5 kDa)
DRQQSQSQSQQEADVIVK $_{203-220}$	2073.00	1G (36.3 kDa)
QQSQSQSQQEADVIVK $_{205-220}$	1801.88	1G (36.3 kDa), 1H (28.5 kDa)
ISREQIEELSK $_{221-231}$	1330.70	1G (36.3 kDa), 1H (28.5 kDa)
EQIEELSK $_{224-231}$	974.49	1G (36.3 kDa), 1H (28.5 kDa), 1M (16.5 kDa)
KSVSSESEPFNLR $_{239-251}$	1478.73	1G (36.3 kDa), 1H (28.5 kDa)
SVSSESEPFNLR $_{240-251}$	1350.64	1M (16.5 kDa), 1H (28.5 kDa)
SRNPIYSNK $_{252-260}$	1077.55	1G (36.3 kDa), 1H (28.5 kDa), 1M (16.5 kDa)
NPIYSNK $_{254-260}$	834.42	1G (36.3 kDa)
YGNFFEITPEKNPQLQDLDISLNSVEINEGSLLLPHFNSR $_{261-300}$	4587.29	1H (28.5 kDa)
NPQLQDLDISLNSVEINEGSLLLPHFNSR $_{272-300}$	3261.66	1G (36.3 kDa), 1H (28.5 kDa)
ATVILVVNEGK $_{301-311}$	1141.67	1M (16.5 kDa)
NFLAGEEDNVISQIQRPVK $_{386-404}$	2156.12	1M (16.5 kDa)
EVAFPGSAEEVDR $_{405-417}$	1404.65	1M (16.5 kDa)
QSHFANAQPQQK $_{424-435}$	1382.66	1M (16.5 kDa)

[a] Amino acids are presented in single letter code within peptide sequences

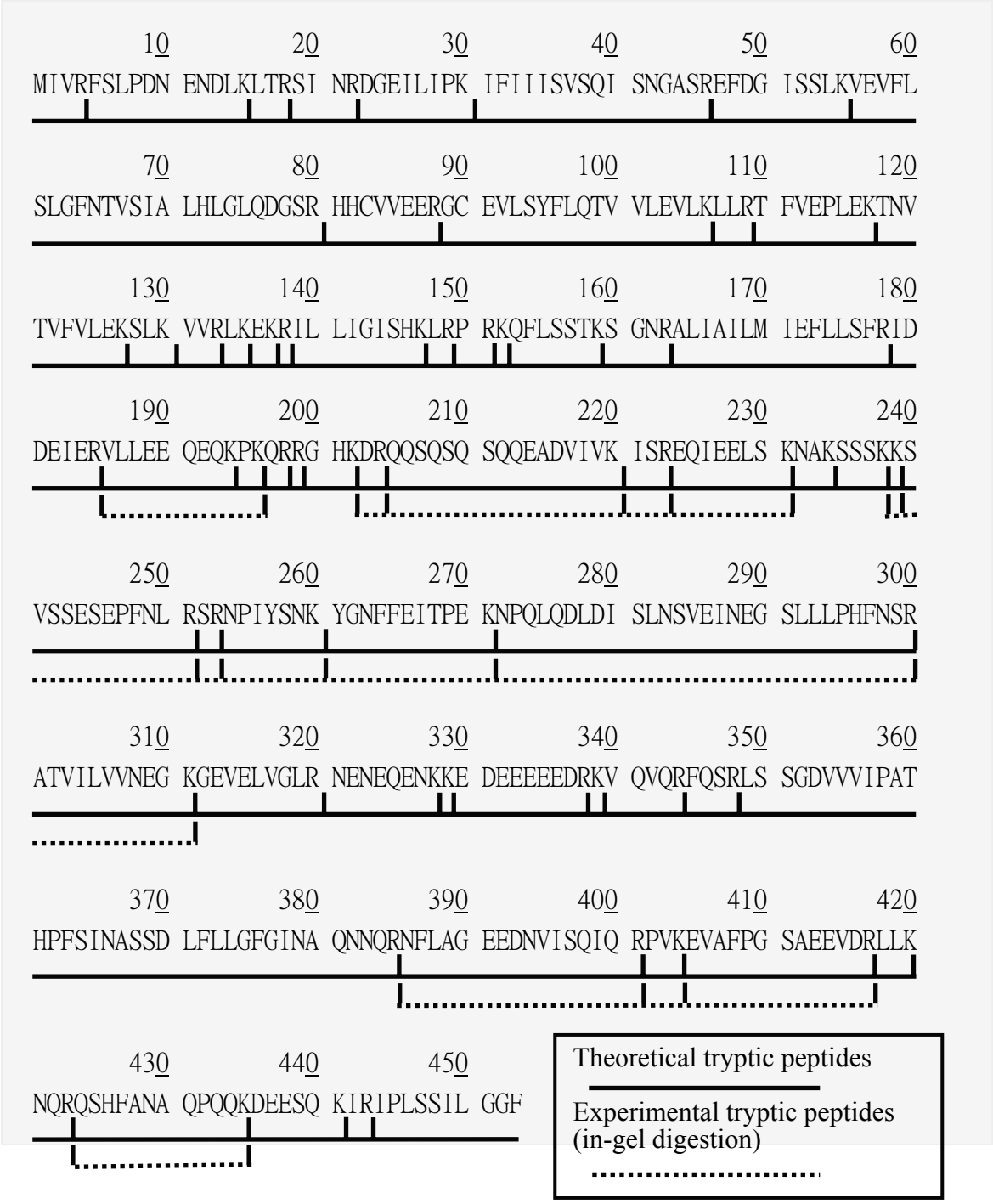


Figure 4.7: Amino acids sequence of identified tryptic peptides with origin from chickpea provicilin precursor [Cicer arietinum] (NCBI accession number: gi|82173888; theoretical mass 51,390 Da)

Figure 4.7 presents the AAs sequence of chickpea provicilin precursor with the identified tryptic peptides from chickpea albumin fraction (C-Ab); the solid line and the broken line represents the theoretical and experimental (in-gel digestion) tryptic peptides, respectively. In theory, 60 cutting sites at either arginine (R) or lysine (K) of chickpea provicilin precursor AAs sequence are available for tryptic hydrolysis; one cutting site at AAs position 149-150 (RP) is not available due to tyrpsin does not react to praline at C-terminal; 21 cutting sites appeared in identified tryptic peptides derived from in-gel digestion. 157 AAs are accountable from identified tryptic peptides; it gives 34% sequence coverage out of 453 total AAs from chickpea provicilin precursor.

4.3.1.3 Identified Tryptic Peptide Sequences Corresponding to Other Legume Proteins (Lentil and Pea)

Figures 4.8 and 4.9 give the representative examples of how the doubly and triply charged peptides were identified in protein band 1D (51.0 kDa) from chickpea albumin fraction (C-Ab; lans 1, Figure 3.6) using the MS spectra during the LC-ESI/MS/MS analysis, respectively; Figure 4.8 shows the peptide LPAGTIAYLANR (MW 1,258.70 Da) with the detected signal (m/z 630.338) which was doubly charged (difference of 0.5 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra. Similarly, Figure 4.9 shows the peptide NFLAGEEDNVISQVERPVK (MW 2,143.09 Da) with the detected signal (m/z 715.359) which was triply charged (difference of 0.33 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra.

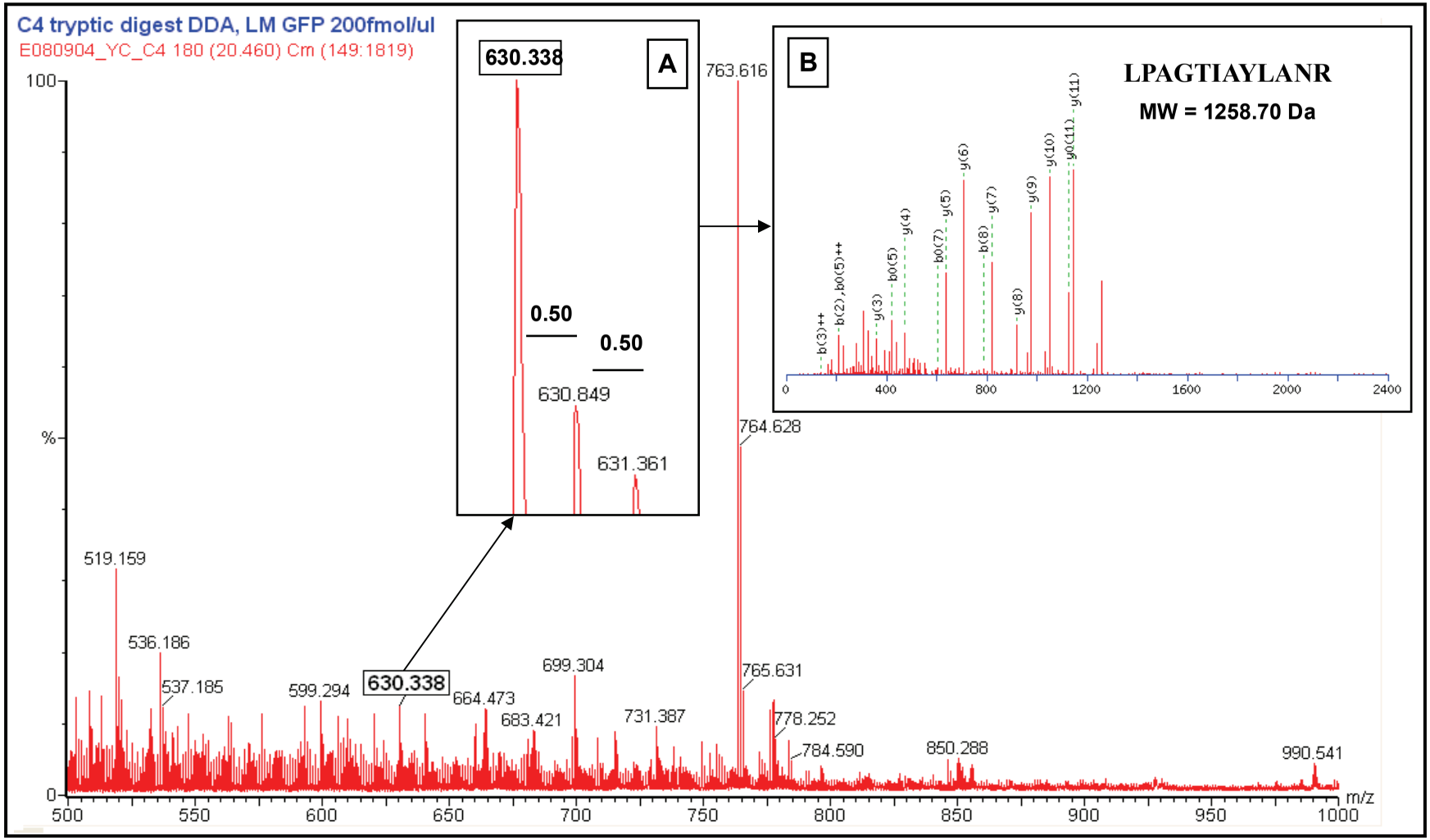


Figure 4.8: LC-ESI-MS spectra (m/z region 500 to 1000) of chickpea protein band 1D, m/z 630.338 signal in bold type. Insert A shows the identified doubly charged signal by the difference of 0.50 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide LPAGTIAYLANR, calculated MW 1258.70 Da

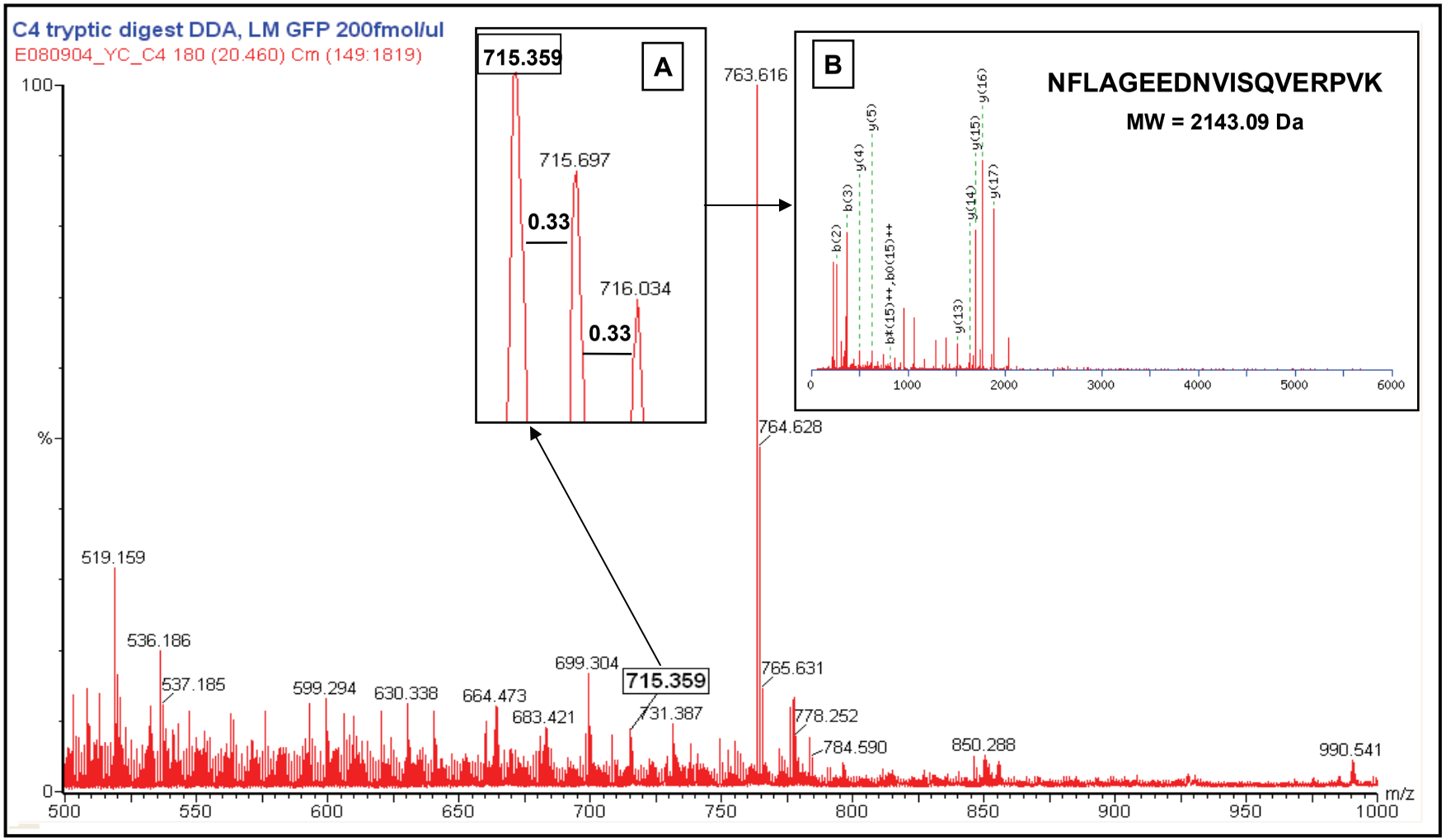


Figure 4.9: LC-ESI-MS spectra (m/z region 500 to 1000) of chickpea protein band 1D, m/z 715.359 signal in bold type. Insert A shows the identified triply charged signal by the difference of 0.33 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide NFLAGEEDNVISQVERPVK, calculated MW 2143.09 Da

Table 4.3 summarizes the results of tryptic peptides from chickpea albumin and globulin fractions (C-Ab and C-Gb) identified with similarity to peptides from other legumes such as lentil and pea proteins; protein bands 1D (51.0 kDa) and 2D (51.0 kDa) showed tryptic peptides corresponding to pea vicilin 47 kDa protein and pea vicilin precursor, respectively, along with other protein bands 1G (36.3 kDa), 1H (28.5 kDa), and 1M (16.5 kDa) that showed tryptic peptides with origin from chickpea provicilin precursor (Table 4.2; Section 4.3.1.2); these protein bands with comparable SDS-PAGE MWs to chickpea vicilin (7S) subunits reported by Chang et al. (2009), showed their tryptic peptide sequences correspond to pea vicilins and chickpea provicilin precursor; in addition, the tryptic peptides identified from protein bands 1D and 2D exhibited the sequence coverage, 15% and 13%, to pea vicilin and vicilin precursor, respectively; all these evidences suggest that protein bands 1D, 1G, 1H and 1M from chickpea albumin fraction, and protein band 2D from chickpea globulin fraction are related to vicilin family. Furthermore, protein bands 1B (63.6 kDa) and 1I (25.9 kDa) manifested the identified tryptic peptides with 13% sequence coverage corresponding to allergens (Allergen Len c 1.0101; Lens culinaris) from lentil vicilin reported in Spain (López-Torrejón et al., 2003); these two protein bands 1B (63.6 kDa) and 1I (25.9 kDa) also show matched SDS-PAGE MWs to the chickpea allergens identified at 70, 64, 35 and 26 kDa (Patil et al., 2001); several allergen proteins of 7S globulin family have been also reported from other plant seeds such as soybean (You et al., 2008), sesame (Beyer et al., 2002a), coconut (Benito et al., 2007). Additionally, protein band 1A (93.6 kDa) from chickpea albumin fraction (C-Ab) which is related to chickpea lipoxygenase (92 kDa) from an albumin fraction reported by Clemente et al. (2000) (Section 3.3.2.2) showed tryptic peptides (sequence

Table 4.3: Summary of tryptic peptides identified from chickpea albumin and globulin fractions (C-Ab and C-Gb) corresponding to other legume proteins (lentil and pea)

Protein bands excised from SDS-PAGE (Fig. 3.6)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
Fraction C-Ab[b]							
1A (93.6 kDa)	SVSLQLISATK $_{65-75}$	1145.66	10	Lipoxygenase [Pisum sativum]	gi/493730	97284	330
	NLFEGGIK $_{317-324}$	876.47					
	TDGEQVLK $_{345-352}$	888.45					
	SAWMTDEEFAR $_{364-374}$	1341.56					
	DTMNINALAR $_{559-568}$	1117.55					
	DSSSPYGLR $_{619-627}$	980.45					
	EVVEKGHGDLK $_{676-686}$	1209.63					
	HASDEVYLGQR $_{782-792}$	1273.60					
	ENPHWTSDSK $_{793-802}$	1199.52					
1B (63.6 kDa)	SKIFENLQNYR $_{37-47}$	1410.72	13	Allergen Len c 1.0101 [Lens culinaris]	gi/29539109	47798	316
	NSFNLER $_{88-94}$	878.42					
	VSREQIEELSK $_{200-210}$	1316.69					
	EQIEELSK $_{203-210}$	974.49					
	SVSSESEPFNLR $_{219-230}$	1350.64					
	SRNPIYSNK $_{231-239}$	1077.55					
	NPIYSNK $_{233-239}$	834.42					
	FFEITPEK $_{243-250}$	1009.51					
1D (51.0 kDa)	SKIFENLQNYR $_{60-70}$		15	Vicilin 47kD protein [Pisum sativum]	gi|297170	49592	333
	NSFNLER $_{111-117}$	878.42					
	LPAGTIAYLANR $_{123-134}$	1258.70					
	EQIEELSK $_{226-233}$	974.49					
	SRNPIYSNK $_{254-262}$	1077.55					
	NFLAGEEDNVISQVERPVK $_{388-406}$	2143.09					

Table 4.3 (Continued)

Protein bands excised from SDS-PAGE (Fig. 3.6)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
1I (25.9 kDa)	SKIFENLQNYR $_{37-47}$	1410.72	13	Allergen Len c 1.0101 [Lens culinaris]	gi/29539109	47798	174
	GDTIKLPAGTIAYLANR $_{95-111}$	1772.97					
	EQIEELSK $_{203-210}$	974.49					
	SVSSESEPFNLR $_{219-230}$	1350.64					
1L (19.0 kDa)	SKIFENLQNYR $_{63-73}$	1410.72	4	Cupin, RmlC-type [Medicago truncatula]	gi|87162567	54170	122
	AILTVLNPNDR $_{103-113}$	1224.68					
1N (8.0 kDa)				ND			
1O (5.3 kDa)				ND			
Fraction C-Gb[c]							
2A (93.6 kDa)	DEAFGHLK $_{264-271}$	915.44	6	Seed lipoxygenase-3	gi|126405	97568	370
	SSDFLTYGLK $_{272-281}$	1129.56					
	EHLEPNLEGLTVEEAIQNK $_{408-426}$	2162.08					
	EHLEPNLEGLTVEEAIQNKK $_{408-427}$	2290.18					
	HASDELYLGER $_{780-790}$	1288.60					
	LAEIEKK $_{812-818}$	829.49					
2D (51.0 kDa)	SKIFENLQNYR $_{64-74}$	1410.72	13	Vicilin precursor [Pisum sativum]	gi|137582	52257	222
	NSFNLER $_{115-121}$	878.42					
	SVSSESEPFNLR $_{246-257}$	1350.64					
	NFLAGDEDNVISQIQRPVKELAFPGSAQ EVDR $_{396-427}$	3541.78					
	ELAFPGSAQEVDR $_{415-427}$	1417.68					

[a] Amino acids are presented in single letter code within peptide sequences; [b] Fraction C-Ab: chickpea albumin fraction; [c] Fraction C-Gb: chickpea globulin fraction; ND: not deteced

coverage 10%) corresponding to pea (*Psium sativum*) lipoxygenase; it was also found that protein band 2A (93.6 kDa) from fraction C-Gb also showed tryptic peptides homologous to pea seed lipoxygenase-3.

4.3.1.4 Identified Tryptic Peptide Sequences with Origin from Rice Proteins

Figures 4.10 and 4.11 give the representative examples of how the doubly and triply charged peptides were identified in protein band 4B (28.6 kDa) from rice glutelin fraction (R-Gt; lane 4, Figure 3.7) using MS spectra during the LC-ESI/MS/MS analysis; Figure 4.10 shows the peptide QKEFLLAGNNNR (MW 1,402.73 Da) with the detected signal (m/z 702.371) which was doubly charged (difference of 0.5 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra; similarly, Figure 4.11 shows the peptide RVIEPQGLLVPR (MW 1,375.83 Da) with the detected signal (m/z 459.283) which was triply charged (difference of 0.33 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra.

Table 4.4 summarizes the tryptic peptides of protein bands from chickpea glutelin (C-Gt) and rice glutelin fraction (R-Gt) corresponding to rice proteins; the protein band 3A (200 kDa) from the chickpea glutelin fraction (C-Gt) showed tryptic peptide sequence, ALPVDVIANAYR, corresponds to rice glutelin C precursor; protein bands 4A (48.6 kDa) and 4B (28.6 kDa) from rice glutelin fraction (R-Gt) showed tryptic peptides with origin from rice glutelin precurosor (15% sequence coverage) and from rice glutelin (33% sequence coverage), respectively; also the protein bands 4C (21.5 kDa), 4D (19.0 kDa)

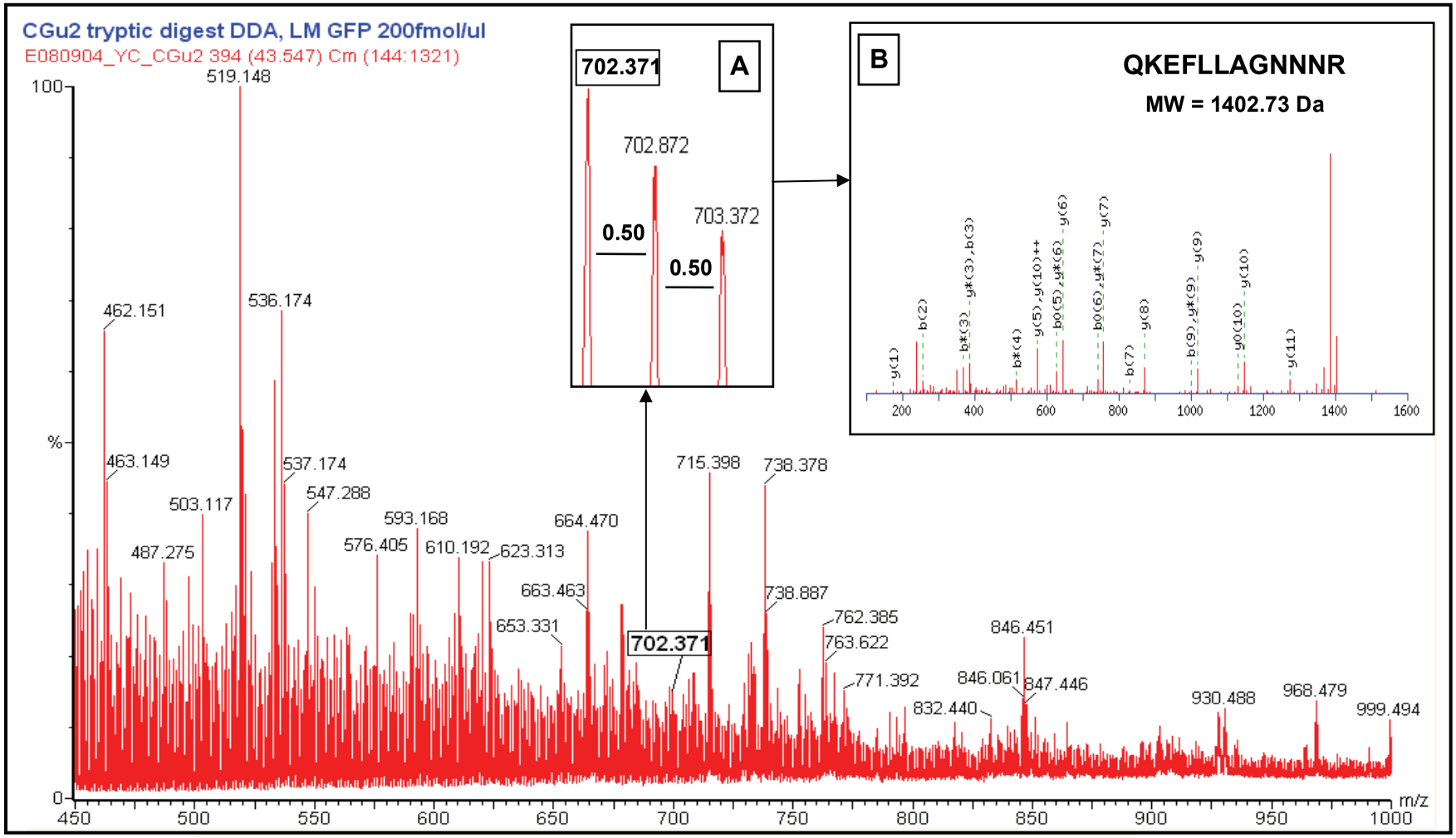


Figure 4.10: LC-ESI-MS spectra (m/z region 450 to 1000) of rice protein band 4B, m/z 702.371 signal in bold type. Insert A shows the identified doubly charged signal by the difference of 0.50 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide QKEFLLAGNNNR, calculated MW 1402.73 Da

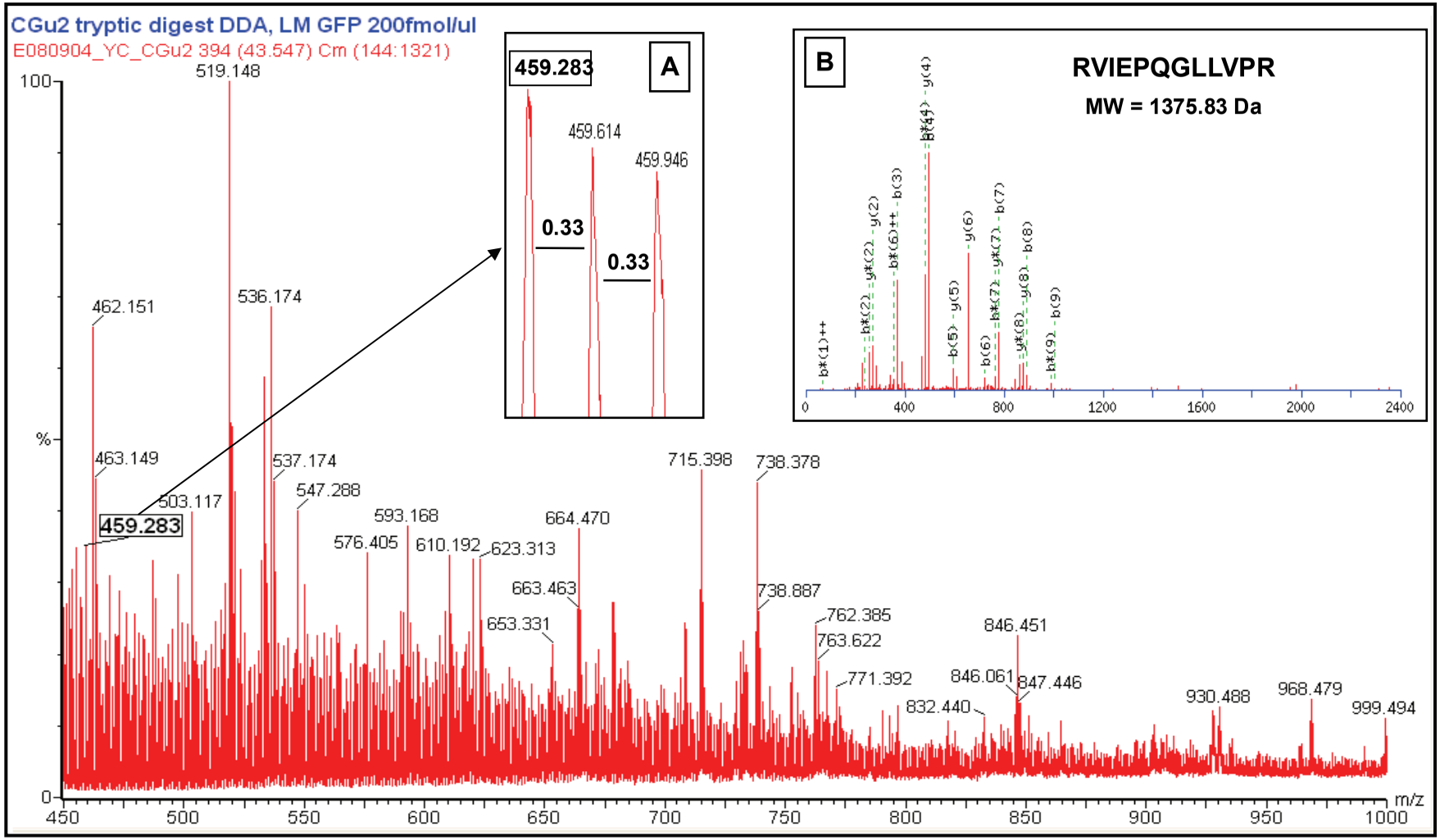


Figure 4.11: LC-ESI-MS spectra (m/z region 450 to 1000) of rice protein band 4B, m/z 459.283 signal in bold type. Insert A shows the identified triply charged signal by the difference of 0.33 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide RVIEPQGLLVPR, calculated MW 1375.83 Da

Table 4.4: Summary of tryptic peptides identified from chickpea and rice glutelin fractions (C-Gt and R-Gt) corresponding to rice proteins

Protein bands excised from SDS-PAGE (Fig. 3.6)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
Fraction C-Gt[b]							
3D (26.2 kDa)	ALPVDVIANAYR	1300.71	2	Glutelin C precursor [Oryza sativa (japonica cultivar-group)]	gi|37993736	54855	59
Fraction R-Gt[c]							
4A (48.6 kDa)	LQAFEPLR $_{50-58}$	972.53	15	Glutelin precursor [Oryza sativa (japonica cultivar-group)]	gi|62546207	56383	241
	VIQPQGLLVPR $_{87-97}$	1218.74					
	QKEFLLAGNNNR $_{196-207}$	1402.73					
	ISSVNSQKFPILNLIQMSATR $_{329-349}$	2362.26					
	TNANAFVSHLAGK $_{426-437}$	1328.68					
	ALPVDVVANAYR $_{442-453}$	1286.69					
4B (28.6 kDa)	FDRLQAFEPLR $_{47-57}$	1390.73	33	Glutelin [Oryza sativa (japonica cultivar-group)]	gi|31455453	36038	634
	VIEPQGLLVPR $_{86-97}$	1219.72					
	RVIEPQGLLVPR $_{87-97}$	1375.83					
	YSNTPGMVYIIQGR $_{98-111}$	1613.78					
	DEHQKIHQFR $_{145-154}$	1336.66					
	QKEFLLAGNNNR $_{196-207}$	1402.73					
	SIEQHSGQNIFSGFNNELLSEALGVNALVAK $_{215-245}$	3285.66					
	LQGQNDQRGEIIR $_{247-259}$	1525.79					
4C (21.5 kDa)	EVQDSPLDACR $_{39-49}$	1288.57	36	19 kDa globulin precursor [Oryza sativa (japonica cultivar-group)]	gi|20159	21492	350

Table 4.4 (Continued)

Protein bands excised from SDS-PAGE (Fig. 3.6)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
	EVQDSPLDACRQVLDR $_{39-54}$	1899.91					
	FQPMFR $_{62-67}$	840.39					
	RPGALGLR $_{68-75}$	838.51					
	MQCCQQLQDVSR $_{76-87}$	1551.65					
	QYAAQLPSMCRVEPQQCSIFAAGQY $_{162-186}$	2918.31					
4D (19.0 kDa)	EVQDSPLDACR $_{39-49}$	1288.57	38	19 kDa globulin precursor [Oryza sativa (japonica cultivar-group)]	gi|20159	21492	552
	EVQDSPLDACRQVLDR $_{39-54}$	1899.91					
	ERFQPMFR $_{60-67}$	1125.53					
	FQPMFR $_{62-67}$	840.39					
	RPGALGLR $_{68-75}$	838.51					
	MQCCQQLQDVSR $_{76-87}$	1551.65					
	MQCCQQLQDVSR ECR $_{76-90}$	1996.83					
	QYAAQLPSMCRVEPQQCSIFAAGQY $_{162-186}$	2918.31					
4E (16.3 kDa)	QCQREVQDSPLDACR $_{35-49}$	1860.82	39	19 kDa globulin precursor [Oryza sativa (japonica cultivar-group)]	gi|20159	21492	465
	EVQDSPLDACR $_{39-49}$	1288.57					
	EVQDSPLDACRQVLDR $_{39-54}$	1899.91					
	ERFQPMFR $_{60-67}$	1125.53					
	FQPMFR $_{62-67}$	840.39					
	RPGALGLR $_{68-75}$	838.51					
	MQCCQQLQDVSR $_{76-87}$	1551.65					
	MQCCQQLQDVSR ECR $_{76-90}$	1996.83					
	QYAAQLPSMCRVEPQQCSIFAAGQY $_{162-186}$	2918.31					
4F (8.0 kDa)	VSYAVEVGR $_{112-120}$	978.47	1	Os11g0546500 [Oryza sativa]	gi|115485845	53213	55

[a] Amino acids are presented in single letter code within peptide sequences; [b] Fraction C-Gt: chickpea glutelin fraction; [c] Fraction R-Gt: rice glutlein fraction

and 4E (16.3 kDa) contained tryptic peptides which showed about 37% sequence coverage corresponding to 19 kDa rice globulin precursor; these results suggest that SDS-PAGE protein bands of rice glutelin fraction are related to rice glutelin and globulin proteins which is in agreement with the results from SDS-PAGE characterization of rice glutelin fraction (R-Gt) (Section 3.3.2.2).

4.3.2 Identification of Tryptic Peptide Sequences from Oat Protein Fractions

4.3.2.1 Identified Tryptic Peptide Sequences with Origin from Oat 12S Globulin

Figures 4.12 and 4.13 give the representative examples of how the doubly and triply charged peptides were identified from protein band 6F (33.2 kDa) were detected using the MS spectrum during the LC-ESI/MS/MS investigation; Figure 4.12 shows the peptide LQAFEPLR (MW 972.53 Da) with the detected signal (m/z 478.251) which was doubly charged (difference of 0.5 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra. Similarly, Figure 4.13 shows the peptide IQSQNDQRGEIIR (MW 1,555.80 Da) with the detected signal (m/z 519.603) which was triply charged (difference of 0.33 between adjacent signals as shown in insert A); the insert B represents the final result of the identified peptide from LC-ESI-MS/MS spectra.

Table 4.5 summarizes the results of identified tryptic peptides of protein bands (Figure 3.7) from oat albumin (O-Ab; lane 5), globulin (O-Gb; lane 6) and glutelin (O-Gt) fractions with origin from oat 12S seed storage globulin 1; oat 12S globulin 1 protein is composed of 518 amino acids and it belongs to 11S seed storage protein (globulin) family;

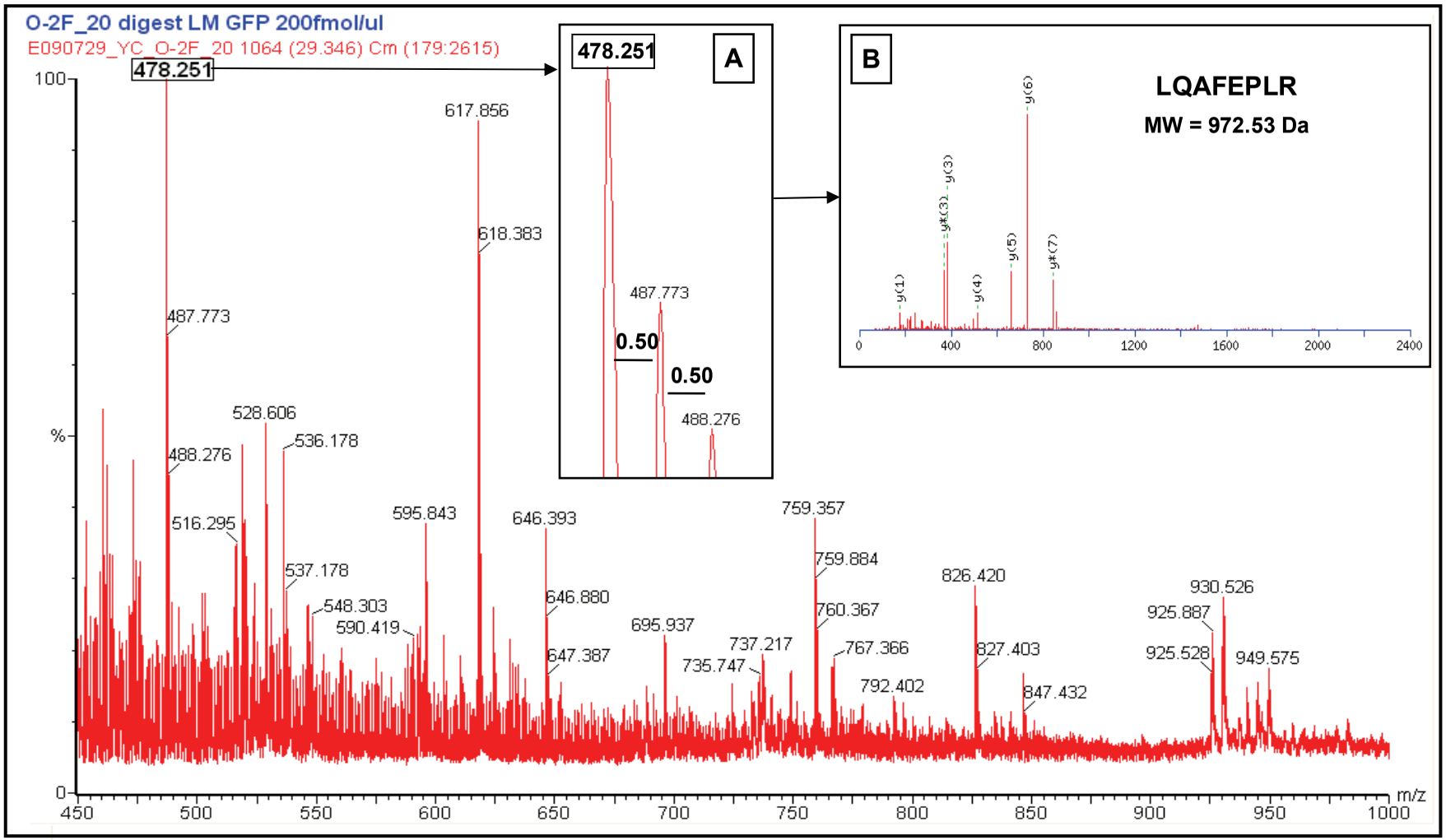


Figure 4.12: LC-ESI-MS spectra (m/z region 450 to 1000) of oat protein band 6F, m/z 478.251 signal in bold type. Insert A shows the identified doubly charged signal by the difference of 0.50 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide LQAFEPLR, calculated MW 972.53 Da

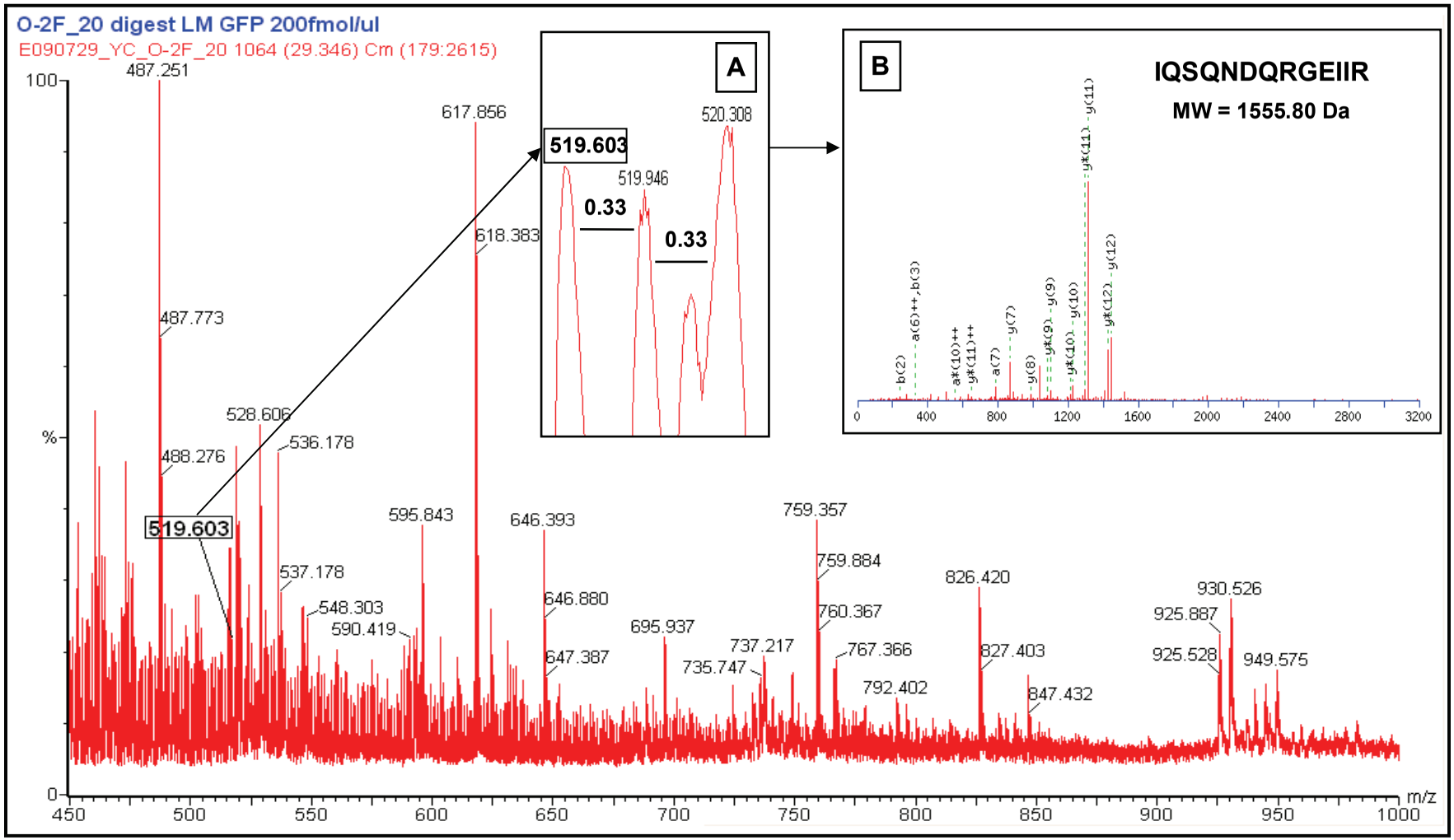


Figure 4.13: LC-ESI-MS spectra (m/z region 450 to 1000) of oat protein band 6F, m/z 519.603 signal in bold type. Insert A shows the identified triply charged signal by the difference of 0.33 between signals. Insert B represents the fragmentation of LC-ESI-MS/MS spectra of the peptide IQSQNDQRGEIIR, calculated MW 1555.80 Da

Table 4.5: Identified tryptic peptides of protein bands (Figure 3.7) from oat albumin (lane 5), globulin (lane 6), glutelin (lane 7) fractions with origin from oat 12S seed storage globulin 1 (NCBI accession number: gi|134918; theoretical mass 58,508 Da)

Peptide sequences[a]	MW (Da)	Identified in SDS-PAGE protein bands (Fig. 3.7)
α-subunit[b]		
LQAFEPLR $_{50-57}$	972.53	6B (60.9 kDa), 6D (39.8 kDa), 6E (36.8 kDa), 6F (33.2 kDa), 6K (15.1 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7C (46.2 kDa), 7D (44.8 kDa), 7F (36.7 kDa), 7G (33.5 kDa), 7H (30.7 kDa), 7E (41.5 kDa)
LQAFEPLRQVR $_{50-60}$	1355.76	6E (36.8 kDa)
SQAGITEYFDEQNEQFR $_{61-77}$	2060.90	6D (39.8 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7F (36.7 kDa), 7G (33.5 kDa)
RVIEPQGLLLPQYHNAPGLVYILQGR $_{86-111}$	2943.64	6D (39.8 kDa), 6F (33.2 kDa), 7F (36.7 kDa)
VIEPQGLLLPQYHNAPGLVYILQGR $_{87-111}$	2787.54	6D (39.8 kDa), 6E (36.8 kDa), 6F (33.2 kDa), 6K (15.1 kDa), 7F (36.7 kDa), 7G (33.5 kDa), 7E (41.5 kDa)
SQNLKDEHQR $_{144-153}$	1253.61	6B (60.9 kDa), 6D (39.8 kDa), 6E (36.8 kDa), 6F (33.2 kDa), 7F (36.7 kDa), 7E (41.5 kDa)
EQQFGQNIFSGFSVQLLSEALGISQQAAQK $_{212-241}$	3252.64	6D (39.8 kDa), 6E (36.8 kDa), 7E (41.5 kDa)
IQSQNDQRGEIIR $_{242-254}$	1555.80	6D (39.8 kDa), 6E (36.8 kDa), 6F (33.2 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7D (44.8 kDa), 7F (36.7 kDa), 7H (30.7 kDa)
β-subunit[c]		
QNIENPK $_{330-336}$	841.42	5L (25.8 kDa), 6B (60.9 kDa), 6G (27.8 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7I (26.2 kDa), 7J (23.5 kDa), 7E (41.5 kDa)
RADTYNPR $_{337-344}$	991.48	5L (25.8 kDa), 6B (60.9 kDa), 6G (27.8 kDa), 6H (24.1 kDa), 6I (22.0 kDa), 6K (15.1 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7C (46.2 kDa), 7D (44.8 kDa), 7F (36.7 kDa), 7H (30.7 kDa), 7I (26.2 kDa), 7J (23.5 kDa), 7E (41.5 kDa)
ADTYNPR $_{338-344}$	835.38	5L (25.8 kDa), 7D (44.8 kDa)
VQVVNNHGQTVFNDILR $_{398-414}$	1952.02	6G (27.8 kDa), 6H (24.1 kDa), 7D (44.8 kDa), 7I (26.2 kDa)
ALPVDVLANAYR $_{461-472}$	1300.71	5L (25.8 kDa), 6B (60.9 kDa), 6G (27.8 kDa), 6H (24.1 kDa), 6I (22.0 kDa), 6K (15.1 kDa), 7A(69.3 kDa), 7B (62.5 kDa), 7C (46.2 kDa), 7D (44.8 kDa), 7F (36.7 kDa), 7H (30.7 kDa), 7I (26.2 kDa), 7J (23.5 kDa), 7E (41.5 kDa)
NNRGEEFGAFTPK $_{483-495}$	1465.69	6G (27.8 kDa), 6H (24.1 kDa), 7D (44.8 kDa), 7I (26.2 kDa)
GEEFGAFTPK $_{486-495}$	1081.50	6G (27.8 kDa), 7D (44.8 kDa), 7I (26.2 kDa)

[a] Amino acids are presented in single letter code within peptide sequences; [b] α-subunit: AAs position located between 25-317; [c] β-subunit: AAs position located between 318-518

thus it resembles chickpea legumin (11S) in many profiles as described in Section 2.2.1.1. Essentially, the AAs sequence of this protein consists a signal peptide at the AA position 1-24, an acidic polypeptide chain (α-subunit) at AA position 25-317, and a basic polypeptide chain (β-subunit) with AA position 318-518; a disulfide bond is observed at AA position 45↔78 and an interchain (between α- and β-subunit) is observed at AA position 121↔314 (Shotwell et al., 1988). In Table 4.5, most of the protein bands from oat globulin and glutelin fractions (O-Gb and O-Gt) showed tryptic peptides with origin from oat 12S seed storage globulin 1; as for oat albumin fraction (O-Ab), there was only one protein band 5L (25.8 kDa) that typtic peptides corresponding to this protein; 8 tyrptic peptides from oat globulin and glutelin fractions are related to α-subunits of oat 12S globulin 1 while 7 tyrptic peptides from oat albumin, globulin and glutelin fractions are related to β-subunits of this protein. Protein bands 6B (60.9 kDa), 7A (69.3 kDa) and 7B (62.5 kDa) which showed tryptic peptides related to oat 12S globulin α- and β-subunits; these protein bands are likely the parent protein (monomers) of oat 12S globulin 1 protein; this result agrees to the protein bands (5B, 5C, 6A, 6B, 7A and 7B) with comparable SDS-PAGE MWs to reported 12S globulin (Lásztity, 1996) (Section 3.3.2.2). Protein bands 6D (39.8 kDa), 6E (36.8 kDa), 6F (33.2 kDa) and 7G (33.5 kDa) which showed the tryptic peptides strictly with origin from oat 12S globulin 1 protein are related to α-subunit (AAs position between 25-317; Table 4.5) of this protein; protein bands 5L (25.8 kDa), 6G (27.8 kDa), 6H (24.1 kDa), 6I (22.0 kDa), 7I (26.2 kDa) and 7J (23.5 kDa) are corresponded to the β-subunits of this protein in result of showing the tryptic peptides only relating to β-subunits (AA position in between 318-518). Protein bands 6K (15.1 kDa), 7C (46.2 kDa), 7D (44.8 kDa), 7E (41.5 kDa), 7F (36.7 kDa) and

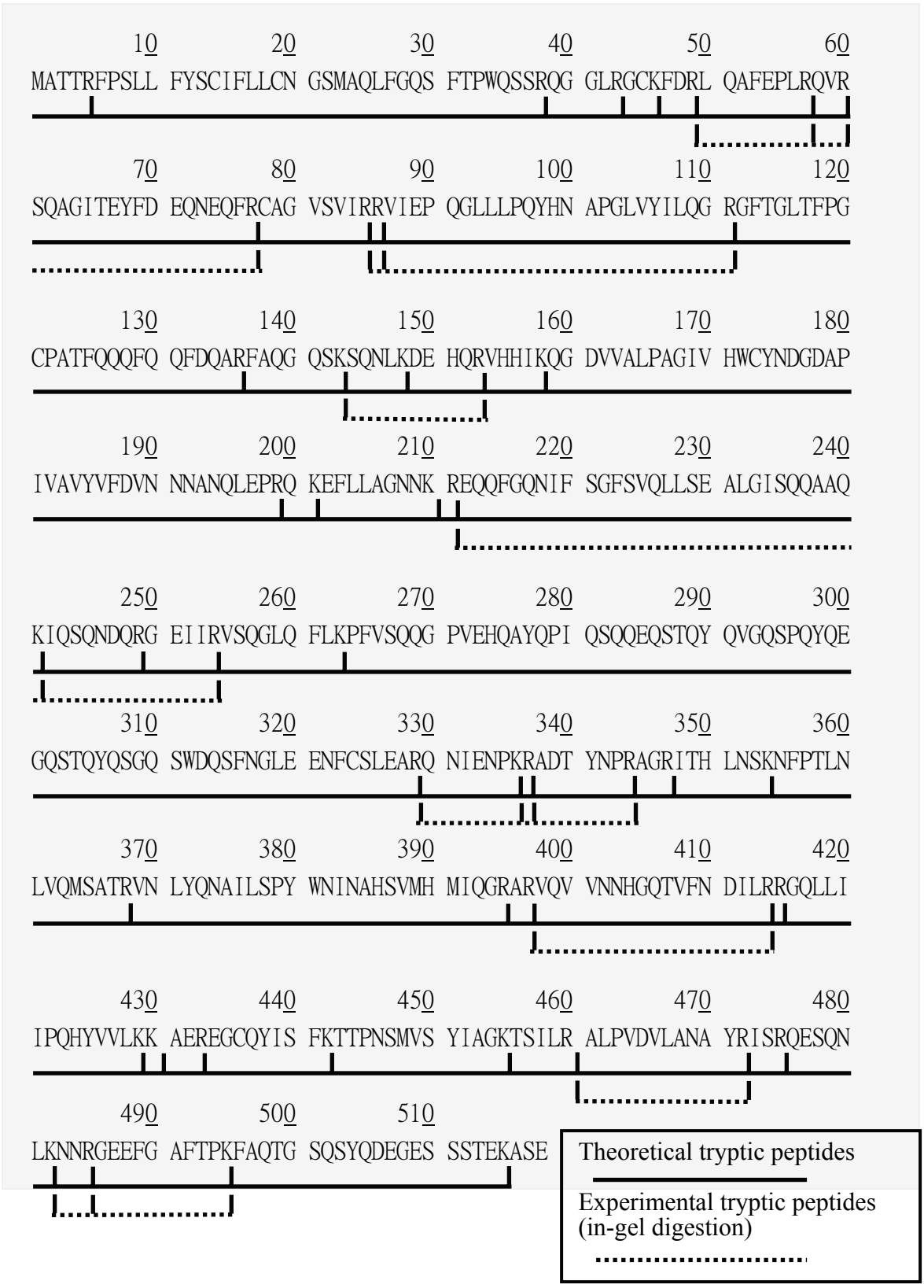


Figure 4.14: Amino acids sequence of identified tryptic peptides with origin from oat 12S seed storage globulin 1 (NCBI accession number: gi|134918; theoretical mass 58,508 Da)

7H (30.7 kDa) showed tryptic peptides present in both α- and β-subunits of oat 12S globulin1 protein.

Figure 4.14 shows the AAs sequence of oat 12S seed storage globulin1 with the tryptic peptides identified from oat protein fractions during the proteomic analysis; the solid line represents the theoretical peptides cleaved by trypsin and the broken line represents the experimental tryptic peptides identified from in-gel digestion. In theory, 47 cutting sites at either arginine (R) or lysine (K) of oat 12S globulin 1 AAs sequence are available for tryptic hydrolysis; one cutting site at AAs position 262-263 (KP) is not available due to tyrpsin does not react to praline at C-terminal; 23 cutting sites appeared in identified tryptic peptides derived from in-gel digestion. Based on the AAs sequence, 161 AAs from identified tryptic peptides out of 518 AAs is accounted for 21% sequence coverage.

4.3.2.2 Identified Tryptic Peptide Sequences Corresponding to Other Cereal Proteins (Wheat, Barley, Rice)

Table 4.6 represents the partial tryptic peptide sequences identified with origin from other cereal proteins (wheat, barley, rice). Protein bands 5A (98.2 kDa), 5B (69.6 kDa), 5M (17.2 kDa), 5N (15.0 kDa), 6A (65.8 kDa), 6C (45.9 kDa) and 6J (18.5 kDa) did not show tryptic peptides corresponding to any protein hits during the Mascot MS/MS ion search; however the unknown tryptic peptide sequences are still available in the Mascot search results for future investigation. As shown in Table 4.6, most protein bands from oat albumin fractions (O-Ab) showed tryptic peptide sequences corresponding to other cereal proteins such as wheat (5C and 5 K), barley (5D, 5E, 5F and 5H), rice (5G, 5I and

Table 4.6: Summary of tryptic peptides identified from oat albumin, globulin and glutelin fractions (O-Ab, O-Gb and O-Gt) corresponding to other cereal proteins (wheat, barley and rice)

Protein bands excised from SDS-PAGE (Fig. 3.7)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
Fraction O-Ab[b]							
5A (98.2 kDa)				ND			
5B (69.6 kDa)				ND			
5C (62.3 kDa)	IVLTIIR $_{176-182}$	826.56	2	Adenosylhomocysteinase (Wheat)	gi/417745	53402	54
	WVFPETK $_{378-384}$	905.46					
5D (55.2 kDa)	IVTEDFLPLPSK $_{172-183}$	1357.74	2	UTP--glucose-1-phosph-ate uridylyltransferase (Barley)	gi/6136111	51612	64
5E (45.7 kDa)	ADVQSADFQTKPAEAVAQVISWVEK $_{23-47}$	2716.37	11	Barley protein Z homolog [Avena fatua]	gi/4868129	31094	58
	VEVGQFK $_{163-169}$	805.43					
5F (43.2 kDa)	IGINGFGR $_{6-13}$	832.45	10	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic (Barley)	gi/120666	36663	170
	AASFNIIPSSTGAAK $_{203-217}$	1433.75					
	AGIALNDNFVK $_{301-401}$	1160.61					
5G (39.3 kDa)	LQAFEPLR $_{50-57}$	972.53	3	Glutelin precursor [Oryza sativa Japonica Group]	gi/27803592	56800	50
	IIEPLGLLLPR $_{87-97}$	1233.74					
5H (36.9 kDa)	TLLFGEK $_{36-42}$	806.45	10	Glyceraldehyde-3-phosphate dehyfrogenase, cytosolic (Barely)	gi/120668	33215	111
	VPTVDVSVVDLTVR $_{205-218}$	1497.84					
	AGIALNDNFVK $_{269-279}$	1160.61					

Table 4.6 (Continued)

Protein bands excised from SDS-PAGE (Fig. 3.7)	Identified peptide sequences[a]	MW (Da)	Sequence coverage (%)	Protein Origin	Accession number from NCBI database	Theoretical Mass (Da)	Mascot score
5I (33.7 kDa)	LQAFEPLR $_{43-50}$	972.54	4	Hypothetical protein OsI_06564 [Oryza sativa Indica Group]	gi/125538806	54682	67
	QKEFLLAGNNQR $_{189-200}$	1416.77					
5J (30.9 kDa)	VSFGENFSPAR $_{199-209}$	1209.57	4	Hypothetical protein [Oryza sativa]	gi/14192878	27403	70
5K (28.5 kDa)	VLLTLEEK $_{25-32}$	943.55	10	Dehydroascorbate reductase [Triticum aestivum] (Wheat)	gi/28192421	23343	149
	WIADSDVITQVIEEK $_{69-83}$	1744.88					
5M (17.2 kDa)				ND			
5N (15.0 kDa)				ND			
Fraction O-Gb[c]							
6A (65.8 kDa)				ND			
6C (45.9 kDa)				ND			
6J (18.5 kDa)				ND			
Fraction O-Gt[d]							
7K (16.1 kDa)	ALPVDVLANAYR $_{429-440}$	1300.71	2	Glutelin C precursor [Oryza sativa]	gi/37993736	54855	76

[a] Amino acids are presented in single letter code within peptide sequences; [b] Fraction O-Ab: oat albumin fraction; [c] Fraction O-Gb: oat globulin fraction; [d] Fraction O-Gt: oat glutelin fraction; ND: not deteced

5J) with relatively low Mascot score ranging from 50 to 216; this may explain that the AAs sequence of oat albumin has not been resolved. A notable tryptic peptide sequence, LQAFEPLR appeared in protein bands 5I (33.7 kDa) and 5G (39.3 kDa) from oat albumin fractions and corresponded to hypothetical protein of rice and glutelin precursor of rice, respectively; it was related to α-subunit from oat 12S globulin 1 (Table 4.5); this suggests that rice and oat proteins share common peptide sequence and possibly they have evolved from a common ancestral gene; Higuchi & Fukazawa (1987) reported rice glutelin and soybean glycinin have evolved from a common ancestral gene. Similar observation was found in peptide sequence ALPVDVLANAYR; this peptide appeared in protein bands 3D (26.2 kDa), 4A (48.6 kDa) and 7K (16.1 kDa) from chickpea glutelin, rice glutelin and oat glutelin fractions which were corresponded to rice glutelin C precursor; furthermore it was also related to β-subunit from oat 12S globulin 1 (Table 4.5); all these clues suggest that chickpea, rice and oat proteins may have evolved from a common ancestral gene (Robert et al., 1985).

4.3.3 Suggested Homology of Chickpea, Oat and Rice Proteins

Homology assessment was conducted using BLAST analysis (Basic Local Alignment Search Tool); three previously identified proteins from proteomic analysis, chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402), oat 12S seed storage globulin 1 (NCBI accession number: gi|134918) and rice glutelin precursor [Oryza sativa (japonica cultivar-group)] (NCBI accession number: gi|62546207) were chosen and their protein sequences from NCBI protein database were used for the homology analysis based on the result that common tryptic peptides

(LQAFEPLR and ALPVDVLANAYR) were found from chickpea, oat and rice protein fractions as discussed in Section 4.3.2.2; these proteins also share similar molecular characteristics in terms of molecular weights (~50 kDa) identified in SDS-PAGE , the AAs sequence lengths (~500 AA); particularly, many protein bands from chickpea and oat glutelin fractions (C-Gt, O-Gt, respectively) showed identified tryptic peptide sequences identical to chickpea legumin and oat 12S globulin 1, respectively (Table 4.1 and Table 4.5).

Figure 4.15 shows the protein sequences alignment of chickpea legumin and oat 12S globulin; line A represents the AAs sequence of chickpea legumin from AA position 15 to 450; line B represents the AAs sequence of oat 12S globulin from AA position 26 to 455; line C represents the aligned sequence with matched amino acids and "+" represents conservative substitution of amino acids that share similar properties; the gap, "-", was used in order to aligned both sequences successively. Similar interpretation was applied to Figure 4.16 and Figure 4.17; Figure 4.16 shows the protein sequences alignment between chickpea legumin (AAs position from 30-445) and rice glutelin precursor (AAs position from 44-431); Figure 4.17 shows the protein sequences alignment between oat 12S globulin (AAs position from 1-498) and rice glutelin precursor (AAs position from 1-479). The results from the BLAST analysis of these three identified proteins are presented in Table 4.7; the "identities" represents the percentage of matched AAs within the total length of the aligned sequences in other words the higher the identities, the more homology between two aligned AAs sequences; the "positives" represents the numbers of matched AAs and the substitutions "+" within the total length of the aligned sequences; the "gaps" represents the numbers of "-" within the total length of the aligned sequences

```
A: 15   LFGTCFALRDQPQQNE---CQLEHLNALKPDNRIKSEGGLIETWNPSNKQFACAGVALSR  71
C:      LFG  F       +Q     C+ + L A +P  +++S+ G+ E ++  N+QF CAGV++ R
B: 26   LFGQSFTPWQSSRQGGLRGCKFDRLQAFEPLRQVRSQAGITEYFDEQNEQFRCAGVSVIR  85

A: 72   ATLQPNSLLQTFLHQRSPEIFIQQGNGYFGMVFPGCVETFEEPRESE--------QGEGS  123
C:        ++P  LL    H     ++I QG G+ G+ FPGC  TF++  +           Q +
B: 86   RVIEPQGLLLPQYHNAPGLVYILQGRGFTGLTFPGCPATFQQQFQQFDQARFAQGQSKSQ  145

A: 124  KFSDSHQKVNRFREGDIIAVPTGVVFWWMFNDQDTPVIAVSLIDTSSFQNQLDQMPRRFYL  183
C:         D HQ+V+  ++GD++A+P G+V W +ND D P++AV + D ++  NQL+   + F L
B: 146  NLKDEHQRVHHIKQGDVVALPAGIVHWCYNDGDAPIVAVYVFDVNNNANQLEPRQKEFLL  205

A: 184  AGNHEQEFLRYQQEGSEEEENEGGNIFSGFKRDFLEDALNVNRRIVNKLQGRNEDEEKGA  243
C:      AGN+++E            +  G NIFSGF    L +AL ++++   K+Q +N+  ++G
B: 206  AGNNKRE-----------QQFGQNIFSGFSVQLLSEALGISQQAAQKIQSQND--QRGE  251

A: 244  IVKVKGGL-------SITTPPEKEPRQKRGSRQEEDEDEDEKRQPHRHSRQDEDEDEKRQ  296
C:      I++V  GL       S   P E +  Q   S+QE+       + P          + ++ Q
B: 252  IIRVSQGLQFLKPFVSQQGPVEHQAYQPIQSQQEQSTQYQVGQSP---------QYQEGQ  302

A: 297  PHHHSRGGSKSQRDNGFEETICTARLHQNIGSSSSPDIYNPQAGRIKTVTSFDLQALRFL  356
C:         +  G S  Q  NG EE  C+    QNI +    D YNP+AGRI  + S +   L  +
B: 303  STQYQSGQSWDQSFNGLEENFCSLEARQNIENPKRADTYNPRAGRITHLNSKNFPTLNLV  362

A: 357  KLSAEFGSLHKNAMFVPHYNLNANSILYALKGRARLLYALNCKGNSVFDGELEAGRALIV  416
C:      ++SA   +L++NA+  P++N+NA+S+++ ++GRAR +  +N  G +VF+  L  G+ LI+
B: 363  QMSATRVNLYQNAILSPYWNINAHSVMHMIQGRAR-VQVVNNHGQTVFNDILRRGQLLII  421

A: 417  PQNFAIAAKSLSDRFSYVAFKTNDRALINVCQKK  450
C:      PQ++ +  K+  +   Y++FKT   ++++   K
B: 422  PQHYVVLKKAEREGCQYISFKTTPNSMVSYIAGK  455
```

Figure 4.15: BLAST analysis for protein sequences alignment, A: chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402), B: oat 12S seed storage globulin 1 (NCBI accession number: gi|134918), C: aligned sequence, "+"conserved substitutions, "-"gap

```
A: 30   ECQLEHLNALKPDNRIKSEGGLIETWNPSNKQFACAGVALSRATLQPNSLLQTFLHQRSP  89
C:      EC+ + L A +P  +++SE G+ E ++  N+ F C G  + R  +QP  LL
B: 44   ECRFDRLQAFEPLRKVRSEAGVTEYFDEKNELFQCTGTFVIRRVIQPQGLLVPRYSNTPG  103

A: 90   EIFIQQGNGYFGMVFPGCVETFEEPRESEQGEGS----KFSDSHQKVNRFREGDIIAVPT  145
C:       ++I QG G  G+ FPGC  T+++  +    +G      KF D HQK+++FR+GD++A+P
B: 104  LVYIIQGRGSMGLTFPGCPATYQQQFQQFSSQGQSQSQKFRDEHQKIHQFRQGDVVALPA  163

A: 146  GVVFWMFNDQDTPVIAVSLIDTSSFQNQLDQMPRRFYLAGNHEQEFLRYQQ-EGSEEEEN  204
C:      GV  W +ND D  V+A+ + D ++  NQL+   + F LAGN+     R QQ  GS  E++
B: 164  GVAHWFYNDGDASVVAIYVYDINNSANQLEPRQKEFLLAGNNN----RVQQVYGSSIEQH  219

A: 205  EGGNIFSGFKRDFLEDALNVNRRIVNKLQGRNEDEEKGAIVKVKGGLSITTPPEKEPRQK  264
C:       NIF+GF  + L +AL +N      +LQ +N  +++G IV VK GL +  P
B: 220  SSQNIFNGFGTELLSEALGINTVAAKRLQSQN--DQRGEIVHVKNGLQLLKP--------  269

A: 265  RGSRQEEDEDEDEKRQPHRHSRQDEDEDEKRQPHHHSRGGSKSQRDNGFEETICTARLHQ  324
C:       +  ++ E    + Q  ++S Q +                  S R NG EE  CT +
B: 270  --TLTQQQEQAQAQYQEVQYSEQQQT----------------SSRWNGLEENFCTIKARV  311

A: 325  NIGSSSSPDIYNPQAGRIKTVTSFDLQALRFLKLSAEFGSLHKNAMFVPHYNLNANSILY  384
C:      NI + S  D YNP+AGRI +V S     L  +++SA   +L++NA+  P +N+NA+S++Y
B: 312  NIENPSRADSYNPRAGRISSVNSQKFPILNLIQMSATRVNLYQNAILSPFWNVNAHSLVY  371

A: 385  ALKGRARLLYALNCKGNSVFDGELEAGRALIVPQNFAIAAKSLSDRFSYVAFKTNDRALI  444
C:       ++G++R+    N  G +VFDG L  G+ LI+PQ++A+  K+  +   Y+A KTN  A +
B: 372  MIQGQSRVQVVSNF-GKTVFDGVLRPGQLLIIPQHYAVLKKAEREGCQYIAIKTNANAFV  430

A: 445  N
C:      +
B: 431  S
```

Figure 4.16: BLAST analysis for protein sequences alignment, A: chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402), B: rice glutelin precursor [Oryza sativa (japonica cultivar-group)] (NCBI accession number: gi|62546207), C: aligned sequence, "+"conserved substitutions, "-"gap

```
A: 1    MATTRFPSLLFYSCIFLLCNGSMAQLFGQSFTPWQSSRQGGLRGCKFDRLQAFEPLRQVR  60
C:      MATT F     Y C  LLC GSMAQLF  S  PW S RQG  R C+FDRLQAFEPLR+VR
B: 1    MATTIFSRFSIYFCAMLLCQGSMAQLFNPSTNPWHSPRQGSFRECRFDRLQAFEPLRKVR  60

A: 61   SQAGITEYFDEQNEQFRCAGVSVIRRVIEPQGLLLPQYHNAPGLVYILQGRGFTGLTFPG  120
C:      S+AG+TEYFDE+NE F+C G  VIRRVI+PQGLL+P+Y N PGLVYI+QGRG  GLTFPG
B: 61   SEAGVTEYFDEKNELFQCTGTFVIRRVIQPQGLLVPRYSNTPGLVYIIQGRGSMGLTFPG  120

A: 121  CPATFQQQFQQFDQARFAQGQSKSQNLKDEHQRVHHIKQGDVVALPAGIVHWCYNDGDAP  180
C:      CPAT+QQQ     Q   +QGQS+SQ  +DEHQ++H  +QGDVVALPAG+ HW YNDGDA
B: 121  CPATYQQQ----FQQFSSQGQSQSQKFRDEHQKIHQFRQGDVVALPAGVAHWFYNDGDAS  176

A: 181  IVAVYVFDVNNNANQLEPRQKEFLLAGNNKR---------EQQFGQNIFSGFSVQLLSEA  231
C:      +VA+YV+D+NN+ANQLEPRQKEFLLAGNN R          EQ   QNIF+GF  +LLSEA
B: 177  VVAIYVYDINNSANQLEPRQKEFLLAGNNNRVQQVYGSSIEQHSSQNIFNGFGTELLSEA  236

A: 232  LGISQQAAQKIQSQNDQRGEIIRVSQGLQFLKPFVSQQGPVEHQAYQPIQSQQEQSTQYQ  291
C:      LGI+  AA+++QSQNDQRGEI+ V  GLQ LKP ++QQ       YQ +Q  ++Q
B: 237  LGINTVAAKRLQSQNDQRGEIVHVKNGLQLLKPTLTQQQEQAQAQYQEVQYSEQQ-----  291

A: 292  VGQSPQYQEGQSTQYQSGQSWDQSFNGLEENFCSLEARQNIENPKRADTYNPRAGRITHL  351
C:                    Q+   W    NGLEENFC+++AR NIENP RAD+YNPRAGRI+ +
B: 292  ---------------QTSSRW----NGLEENFCTIKARVNIENPSRADSYNPRAGRISSV  332

A: 352  NSKNFPTLNLVQMSATRVNLYQNAILSPYWNINAHSVMHMIQGRARVQVVNNHGQTVFND  411
C:      NS+ FP LNL+QMSATRVNLYQNAILSP+WN+NAHS+++MIQG++RVQVV+N G+TVF+
B: 333  NSQKFPILNLIQMSATRVNLYQNAILSPFWNVNAHSLVYMIQGQSRVQVVSNFGKTVFDG  392

A: 412  ILRRGQLLIIPQHYVVLKKAEREGCQYISFKTTPNSMVSYIAGKTSILRALPVDVLANAY  471
C:      +LR GQLLIIPQHY VLKKAEREGCQYI+ KT  N+ VS++AGK S+ RALPVDV+ANAY
B: 393  VLRPGQLLIIPQHYAVLKKAEREGCQYIAIKTNANAFVSHLAGKNSVFRALPVDVVANAY  452

A: 472  RISRQESQNLKNNRGEEFGAFTPKFAQ  498
C:      RISR++++++KNNRGEE GAFTP+F Q
B: 453  RISREQARSIKNNRGEEHGAFTPRFQQ  479
```

Figure 4.17: BLAST analysis for protein sequences alignment, A: oat 12S seed storage globulin 1 (NCBI accession number: gi|134918), B: rice glutelin precursor [Oryza sativa (japonica cultivar-group)] (NCBI accession number: gi|62546207), C: aligned sequence, "+"conserved substitutions, "-"gap

Table 4.7: Summary of protein sequences alignments for chickpea legumin, oat 12S globulin 1 and rice glutelin precursor

Protein sequence aligments	Identities	Positives	Gaps	Scores
Chickpea legumin vs. oat 12S globulin1	141/454 (31%)	243/454 (53%)	42/454 (9%)	245 bits (625)
Chickpea legumin vs. rice glutelin precursor	143/421 (33%)	230/421 (54%)	38/421 (9%)	240 bits (613)
oat 12S globulin1 vs. rice glutelin precursor	321/507 (63%)	389/507 (76%)	37/507 (7%)	636 bits (1640)

Method: Compositional matrix adjust (Altschul et al., 2005)

(Altschul et al., 2005). Score or bit score is a value calculated from the number of gaps and substitutions associated with each aligned sequence; normally, the higher the score, the more significant the alignment (Altschul et al., 1997). In Table 4.7, chickpea legumin and oat 12S globulin 1 show 31% identities; chickpea legumin and rice glutlein precursor show 33% identities; this result shows that chickpea legumin, oat 12S globulin 1 and rice glutelin precursor share about 30% sequence homology; as shown in common AA regions (Figure 4.15 and Figure 4.16), the AA identities are dispersed throughout these three protein sequences, therefore the similarity is not due to convergent evolution, but to divergence evolution from a common ancestral gene; evidence is in agreement with Higuchi & Fukazawa (1987) who reported a overall 32% of the amino acid positions that are identical in both rice glutelin and soybean glycinin, and with Robert et al. (1985) who reported that rice glutelin and oat globulin share homology as legumin-like protein based on the similarities of their α- and β-subunits group. Oat 12S globulin 1 and rice glutelin precursor show 63% identities; this evidence suggests there is a substantial similarity of protein sequences between oat 12S globulin and rice glutelin precursor; Wen & Luthe reported a similarity of the amino acid composition of subunits from rice glutlein and 11S oat globulins.

4.3.4 Prediction of Potential ACE-Inhibitory Peptides from Chickpea Proteins

The protein sequences of chickpea legumin and proviclin precursor were submitted to BIOPEP analysis using the "profiles of potential biological activity" tool; only those potential bioactive peptides with ACE-inhibitory activity were collected for further

investigation. The results of prediction of potential ACE-inhibitory peptides from chickpea legumin and provicilin are summarized in Table 4.8. Chickpea legumin and provicilin showed 177 and 133 potential ACE-inhibitory peptides, respectively, within their primary sequences; the predicted peptides corresponded to the peptide sequences published and found in various sources including whey protein (Abubakar et al., 1998; FitzGerald & Meisel, 1999), sardine (Matsufuji et al., 1994a), sake (Saito et al., 1994), corn endosperm (Miyoshi et al., 1995), milk protein (Meisel, 1998; van Platerink et al., 2008), soybean hydrolysate (Wang & Gonzalez de Mejia, 2005), synthetic dipeptide (Cheung et al., 1980; Cushman, 1981), porcine myosin (Meisel et al., 2006), ovalbumin (Meisel et al., 2006), red algae (Meisel et al., 2006), wakame (Meisel et al., 2006), pea vicilin (Meisel et al., 2006), garlic (Meisel et al., 2006), anchovy (Murray & FitzGerald, 2007) and Alaskan pollack skin (Byun & Kim, 2002). The total number of predicted ACE-inhibitory peptides represents the sum of frequency of the specific peptides occurring in the protein sequence using the BIOPEP analysis. Chickpea legumin and provicilin give a total of 92 and 72 peptides, respectively, with 55 of these peptides are common to both proteins (Table 5.1); among those common peptides, sequences RF, KR, RA, AG, KG, GS, GG, EG, KL, LQ, LN and EK are most frequent.

Figures 4.18 and 4.19 illustrate the predicted ACE-inhibitory peptides and identified typtic peptides (Section 4.3.1) within the primary sequence of chickpea legumin and provicilin precursor, respectively; as an example, the peptide sequence FCFLLFGTCF at AA position 11-20 of chickpea legumin (Figure 4.18) gives the predicted ACE-inhibitory peptides, FCF, CF, LLF, LF, FG, GT, and CF. Information on the predicted

Table 4.8: Summary of predicted ACE-inhibitory peptide sequences in chickpea proteins using BIOPEP "profile of potential biological activity" tool

ID	Sequence		Number of peptides[a]		Location[b]		Origin	Reference
BIOPEP database	Legumin	Provicilin precursor	Legumin	Provicilin precursor	Legumin	Provicilin precursor		
3257	RL	RL	2	3	[321-322], [391-392]	[133-134], [348-349], [417-418]	Whey protein (lactokinins)	(FitzGerald & Meisel, 1999)
3258	-	IR	-	1	-	[442-443]	Whey protein (lactokinins)	(FitzGerald & Meisel, 1999)
3361	LKL	LKL	1	2	[356-358]	[14-16], [105-107]	Sardine	(Ukeda et al., 1992)
3370	AVP	-	1	-	[142-144]	-	Casein (casokinins)	(Meisel, 1993)
3380	RY	-	1	-	[193-194]	-	Sardine	(Matsufuji et al., 1994)
3381	LY	-	2	-	[383-384], [393-394]	-	Sardine	(Matsufuji et al., 1994)
3383	IY	IY	1	1	[334-335]	[256-257]	Sardine	(Matsufuji et al., 1994)
3384	VF	VF	3	2	[103-104], [148-149], [403-404]	[58-59], [122-123]	Sardine	(Matsufuji et al., 1994)
3385	MF	-	2	-	[151-152], [370-371]	-	Sardine	(Matsufuji et al., 1994)
3406	TAP	-	1	-	[469-471]	-	Human β-Casein	(Kohmura et al., 1989)
3489	RF	RF	4	2	[134-135], [180-181], [354-355], [430-431]	[4-5], [344-345]	Sake and sake lees	(Saito et al., 1994)
3494	HY	-	1	-	[374-375]	-	Sake and sake lees	(Saito et al., 1994)
3502	FP	FP	1	1	[104-105]	[408-409]	Whey protein	(Abubakar et al., 1998)
3507	-	IPA	-	1	-	[357-359]	Whey protein	(Abubakar et al., 1998)
3532	GY	-	1	-	[98-99]	-	Sake and sake lees	(Saito et al., 1994)
3537	PR	PR	3	1	[114-115], [178-179], [261-262]	[150-151]	Sake and sake lees	(Saito et al., 1994)
3542	LQP	-	1	-	[74-76]	-	Corn endosperm	(Miyoshi et al., 1995)
3543	-	LRP	-	1	-	[148-150]	Corn endosperm	(Miyoshi et al., 1995)
3550	YL	-	1	-	[182-183]	-	Milk protein (bovine b-Lg)	(Meisel, 1998)
3551	LF	LF	2	1	[15-16], [475-476]	[371-372]	Milk protein (bovine b-Lg)	(Meisel, 1998)
3553	-	YG	-	1	-	[261-262]	Milk protein (bovine b-Lg)	(Meisel, 1998)
3556	FY	-	1	-	[181-182]	-	α-zein	(Yano et al., 1996)
3573	-	AFP	-	1	-	[407-409]	β-Casein	(Kohmura et al., 1990)
3713	-	LLP	-	1	-	[293-295]	α-zein	(Yano et al., 1996)

Table 4.8 (Continued)

ID	Sequence		Number of peptides[a]		Location[b]		Origin	Reference
BIOPEP database	Legumin	Provicilin precursor	Legumin	Provicilin precursor	Legumin	Provicilin precursor		
7513	-	PL	-	2	-	[114-115], [445-446]	Alaskan pollack skin	(Byun & Kim, 2002)
7544	IW	-	1	-	[457-458]	-	Finnish cereals	(Loponen, 2004)
7549	LKP	-	1	-	[39-41]	-	Casein (casokinins)	(Meisel, 1993)
7558	VK	VK	2	2	[245-246], [247-248]	[219-220], [403-404]	Soybean hydrolysate	(Wang & Gonzalez de Mejia, 2005)
7562	IA	IA	3	2	[141-142], [160-161], [422-423]	[69-70], [166-167]	Soybean hydrolysate	(Wang & Gonzalez de Mejia, 2005)
7581	-	IP	-	3	-	[28-29], [357-358], [444-445]	Synthetic dipeptide	(Cheung et al., 1980)
7582	RP	RP	1	2	[462-463]	[149-150], [401-402]	Synthetic dipeptide	(Cheung et al., 1980)
7583	AF	AF	1	1	[435-436]	[407-408]	Synthetic dipeptide	(Cheung et al., 1980)
7584	AP	-	1	-	[470-471]	-	Synthetic dipeptide	(Cheung et al., 1980)
7585	LA	LA	2	1	[5-6], [183-184]	[388-389]	Synthetic dipeptide	(Cushman, 1981)
7586	KR	KR	4	1	[214-215], [264-265], [278-279], [294-295]	[137-138]	Synthetic dipeptide	(Cushman, 1981)
7587	VP	-	4	-	[143-144], [372-373], [416-417], [419-412]	-	Synthetic dipeptide	(Cheung et al., 1980)
7588	RA	RA	4	2	[71-71], [389-390], [412-413], [411-412]	[163-164], [300-301]	Synthetic dipeptide	(Cushman, 1981)
7589	YA	-	2	-	[384-385], [394-395]	-	Synthetic dipeptide	(Cushman, 1981)
7590	AA	-	1	-	[423-424]	-	Synthetic dipeptide	(Cushman, 1981)
7591	GF	GF	2	3	[212-213], [312-313]	[63-64], [375-376], [452-453]	Synthetic dipeptide	(Cheung et al., 1980)
7592	FR	FR	1	1	[135-136]	[177-178]	Synthetic dipeptide	(Cushman, 1981)
7593	IF	IF	2	1	[91-92], [209-210]	[31-32]	Synthetic dipeptide	(Cheung et al., 1980)
7594	-	VG	-	1	-	[317-318]	Synthetic dipeptide	(Cheung et al., 1980)
7595	IG	IG	1	1	[326-327]	[142-143]	Synthetic dipeptide	(Cheung et al., 1980)
7596	-	GI	-	3	-	[50-51], [143-144], [377-378]	Synthetic dipeptide	(Cheung et al., 1980)
7597	GM	-	1	-	[101-102]	-	Synthetic dipeptide	(Cheung et al., 1980)
7598	GA	GA	1	1	[242-243]	[43-44]	Synthetic dipeptide	(Cheung et al., 1980)

Table 4.8 (Continued)

ID	Sequence		Number of peptides[a]		Location[b]		Origin	Reference
BIOPEP database	Legumin	Provicilin precursor	Legumin	Provicilin precursor	Legumin	Provicilin precursor		
7599	GL	GL	2	2	[50-51], [250-251]	[74-75], [318-319]	Synthetic dipeptide	(Cheung et al., 1980)
7600	AG	AG	4	1	[65-66], [184-185], [339-340], [410-411]	[389-390]	Synthetic dipeptide	(Cheung et al., 1980)
7601	-	GH	-	1	-	[200-201]	Synthetic dipeptide	(Cheung et al., 1980)
7602	HL	HL	1	1	[35-36]	[72-73]	Synthetic dipeptide	(Cushman, 1981)
7603	GR	-	4	-	[234-235], [340-341], [388-389], [411-412]	-	Synthetic dipeptide	(Cheung et al., 1980)
7604	KG	KG	4	1	[241-242], [248-249], [387-388], [399-400]	[311-312]	Synthetic dipeptide	(Cheung et al., 1980)
7605	FG	FG	3	1	[16-17], [100-101], [362-363]	[376-377]	Synthetic dipeptide	(Cheung et al., 1980)
7606	DA	-	1	-	[220-221]	-	Synthetic dipeptide	(Cushman, 1981)
7607	GS	GS	7	3	[122-123], [198-199], [266-267], [304-305], [327-328], [363-364], [464-465]	[78-79], [290-291], [410-411]	Synthetic dipeptide	(Cheung et al., 1980)
7608	GV	-	2	-	[66-67], [146-147]	-	Synthetic dipeptide	(Cheung et al., 1980)
7611	-	GK	-	1	-	[310-311]	Synthetic dipeptide	(Cheung et al., 1980)
7612	GT	-	1	-	[17-18]	-	Synthetic dipeptide	(Cheung et al., 1980)
7615	GE	GE	2	3	[120-121], [406-407]	[24-25], [312-313], [390-391]	Synthetic dipeptide	(Cheung et al., 1980)
7616	GG	GG	4	1	[49-50], [206-207], [249-250], [303-304]	[451-452]	Synthetic dipeptide	(Cushman, 1981)
7617	QG	-	3	-	[95-96], [119-120], [233-234]	-	Synthetic dipeptide	(Cheung et al., 1980)
7618	SG	SG	1	2	[211-212]	[160-161], [351-352]	Synthetic dipeptide	(Cheung et al., 1980)
7619	-	LG	-	4	-	[62-63], [73-74], [374-375], [450-451]	Synthetic dipeptide	(Cheung et al., 1980)
7620	GD	GD	1	1	[138-139]	[352-353]	Synthetic dipeptide	(Cheung et al., 1980)
7621	TG	-	1-		[145-146]	-	Synthetic dipeptide	(Cheung et al., 1980)

Table 4.8 (Continued)

ID	Sequence		Number of peptides[a]		Location[b]		Origin	Reference
BIOPEP database	Legumin	Provicilin precursor	Legumin	Provicilin precursor	Legumin	Provicilin precursor		
7622	EG	EG	5	2	[48-49], [121-122], [137-138], [197-198], [205-206]	[289-290], [309-310]	Synthetic dipeptide	(Cheung et al., 1980)
7623	EA	EA	1	1	[409-410]	[214-215]	Synthetic dipeptide	(Cheung et al., 1980)
7624	NG	NG	2	1	[97-98], [311-312]	[42-43]	Synthetic dipeptide	(Cushman, 1981)
7625	PG	PG	2	1	[105-106], [463-464]	[409-410]	Synthetic dipeptide	(Cheung et al., 1980)
7628	-	VR	-	2	-	[3-4], [132-133]	Milk protein	(Gómez-Ruiz et al., 2007)
7644	ITT	-	1	-	[253-255]	-	Porcine myosin	(Meisel et al., 2006)
7648	FCF	-	1	-	[11-13]	-	Ovalbumin	(Meisel et al., 2006)
7649	LRY	-	1	-	[192-194]	-	Red algae	(Meisel et al., 2006)
7654	NKL	-	1	-	[230-232]	-	Wakame	(Meisel et al., 2006)
7680	QK	QK	3	3	[130-131], [263-264], [448-449]	[193-194], [434-435], [440-441]	Pea vicilin	(Meisel et al., 2006)
7681	DG	DG	1	3	[405-406]	[23-24], [49-50], [77-78]	Soy	(Meisel et al., 2006)
7683	NF	NF	1	2	[419-420]	[263-264], [386-387]	Garlic	(Meisel et al., 2006)
7684	SY	SY	1	1	[432-433]	[94-95]	Garlic	(Meisel et al., 2006)
7685	SF	SF	3	1	[10-11], [169-170], [347-348]	[176-177]	Garlic	(Meisel et al., 2006)
7691	-	KY	-	1	-	[260-261]	Wakame	(Meisel et al., 2006)
7692	KF	-	1	-	[124-125]	-	Wakame	(Meisel et al., 2006)
7693	KL	KL	4	3	[3-4], [231-232], [357-358], [450-451]	[15-16], [106-107], [147-148]	Wakame	(Meisel et al., 2006)
7698	NK	NK	2	2	[59-60], [230-231]	[259-260], [327-328]	Wakame	(Meisel et al., 2006)
7741	RR	RR	2	1	[179-180], [226-227]	[198-199]	Synthetic dipeptide	(Sentandreu & Toldr, 2007)
7742	AR	-	2	-	[320-321], [390-391]	-	Synthetic dipeptide	(Sentandreu & Toldr, 2007)
7751	CF	-	2	-	[12-13], [19-20]	-	Shark meat hydrolysate	(Wu et al., 2008)
7807	LLF	-	1	-	[14-16]	-	Caprine b-Lg	(Murray & FitzGerald, 2007)
7810	KP	KP	1	1	[40-41]	[194-195]	Anchovy and bonito	(Murray & FitzGerald, 2007)
7823	FAL	-	1	-	[20-21]	-	Micro algae	(Murray & FitzGerald, 2007)
7826	EI	EI	1	4	[90-91]	[25-26], [182-183], [266-267], [286-287]	Milk hydrolysate	(van Platerink et al., 2008)

Table 4.8 (Continued)

ID	Sequence		Number of peptides[a]		Location[b]		Origin	Reference
BIOPEP database	Legumin	Provicilin precursor	Legumin	Provicilin precursor	Legumin	Provicilin precursor		
7827	IE	IE	1	3	[52-53]	[171-172], [183-184], [226-227]	Milk hydrolysate	(van Platerink et al., 2008)
7828	-	EV	-	6	-	[57-58], [91-91], [103-104], [313-314], [405-406], [414-415]	Milk hydrolysate	(van Platerink et al., 2008)
7829	VE	VE	1	5	[108-109]	[56-57], [85-86], [112-113], [285-286], [314-315]	Milk hydrolysate	(van Platerink et al., 2008)
7831	LQ	LQ	5	3	[74-75], [80-81], [232-233], [350-351], [452-453]	[75-76], [97-98], [275-276]	Milk hydrolysate	(van Platerink et al., 2008)
7832	LN	LN	4	1	[36-37], [222-223], [377-378], [396-397]	[282-283]	Milk hydrolysate	(van Platerink et al., 2008)
7833	PT	-	1	-	[144-145]	-	Milk hydrolysate	(van Platerink et al., 2008)
7834	TQ	-	1	-	[481-482]	-	Milk hydrolysate	(van Platerink et al., 2008)
7836	PP	-	1	-	[256-257]	-	Milk hydrolysate	(van Platerink et al., 2008)
7837	PQ	PQ	3	2	[26-27], [337-338], [417-418]	[273-274], [432-433]	Milk hydrolysate	(van Platerink et al., 2008)
7840	EK	EK	4	4	[240-241], [258-259], [277-278], [293-294]	[116-117], [126-127], [136-137], [270-271]	Milk hydrolysate	(van Platerink et al., 2008)
7841	KE	KE	2	3	[259-260], [459-460]	[135-136], [329-330], [404-405]	Milk hydrolysate	(van Platerink et al., 2008)
7842	HP	HP	1	1	[477-478]	[361-362]	Milk hydrolysate	(van Platerink et al., 2008)
7843	PH	PH	3	1	[281-282], [297-298], [373-374]	[295-296]	Milk hydrolysate	(van Platerink et al., 2008)
7844	HK	HK	1	2	[366-367]	[146-147], [201-202]	Milk hydrolysate	(van Platerink et al., 2008)
8185	TF	TF	2	1	[82-83]	[110-111]	Wheat bran	(Nogata et al., 2009)
8193	AI	AI	2	1	[243-244], [421-422]	[167-168]	Soy sauce	(Nakahara et al., 2009)
	92[c]	72[c]						
			177[d]	133[d]				

[a] The number of peptides represents the frequency of theoretical ACE-inhibitory peptides which predicted within the protein sequences of chickpea legumin and provicilin investigated during the BIOPEP analysis

[b] Location represents the amino acids (AAs) position of predicted ACE-inhibitory peptides from chickpea legumin and provicilin

[c] The number represents the sum of predicted ACE-inhibitory peptides from chickpea legumin and provicilin

[d] The number represents the sum of the frequency of predicted ACE-inhibitory peptides from chickpea legumin and provicilin

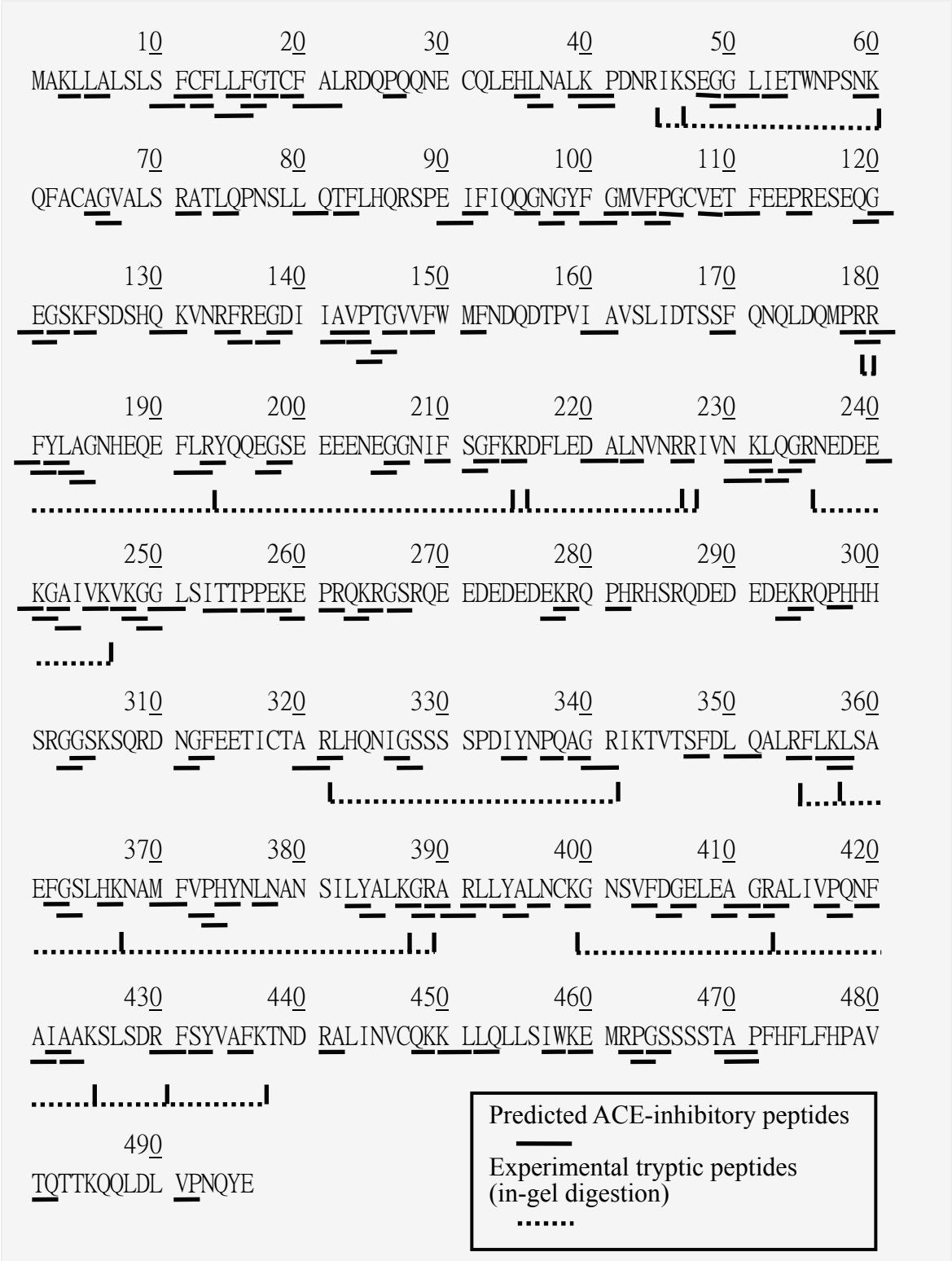


Figure 4.18: Predicted ACE-inhibitory peptides sequences and experimental tryptic peptide sequences with origin from chickpea legumin α- and β-subunit [Cicer arietinum] (NCBI accession number: gi|6273402; theoretical mass 56,216 Da)

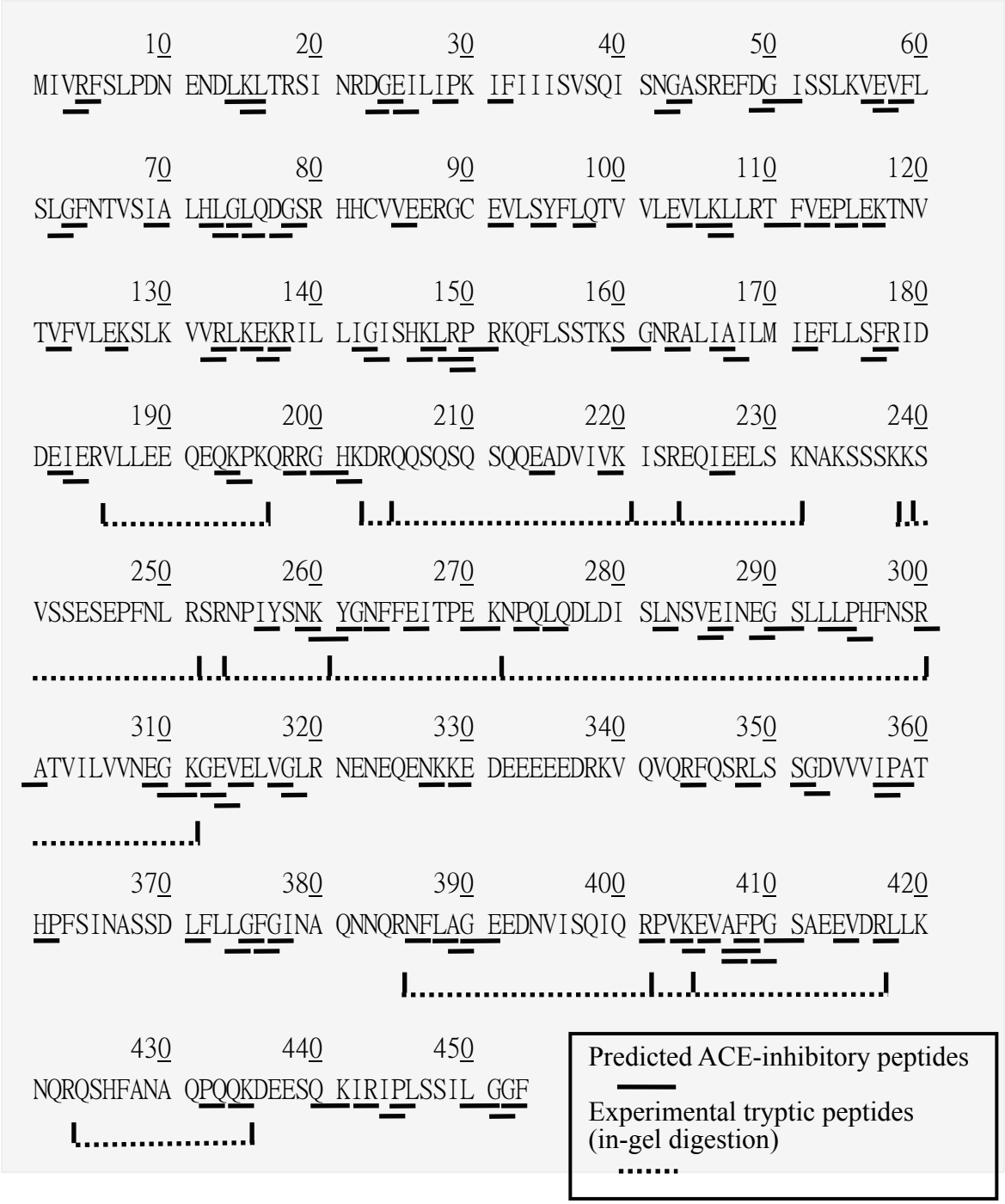


Figure 4.19: Predicted ACE-inhibitory peptides sequences and experimental tryptic peptides with origin from chickpea provicilin precursor [Cicer arietinum] (NCBI accession number: gi|82173888; theoretical mass 51,390 Da)

Table 4.9: Summary of identified tryptic peptides with origin from chickpea legumin and provicilin precursor, and their potential ACE-inhibitory peptides predicted from BIOPEP analysis

Peptide sequences (Figs. 4.1 and 4.2)	Potential ACE-inhibitory peptides predicted from BIOPEP analysis (Figs. 4.18 and 4.19)
Legumin α-subunit	
IKSEGGLIETWNPSNK $_{45-60}$	EG (1), GG (1), GL (1), IE (1), NK (1)
SEGGLIETWNPSNK $_{47-60}$	EG (1), GG (1), GL (1), IE (1), NK (1)
RFYLAGNHEQEFLR $_{180-193}$	RF (1), FY (1), YL (1), LA (1), AG (1), FLR (1)
FYLAGNHEQEFLR $_{181-193}$	FY (1), YL (1), LA (1), AG (1), FLR (1)
YQQEGSEEEENEGGNIFSGFK $_{194-214}$	EG (2), GS (1), GG (1), IF (1), SG (1), GF (1)
YQQEGSEEEENEGGNIFSGFKR $_{194-215}$	EG (2), GS (1), GG (1), IF (1), SG (1), GF (1), KR (1)
RDFLEDALNVNR $_{215-227}$	DA (1), LN (1)
DFLEDALNVNR $_{216-226}$	DA (1), LN (1)
DFLEDALNVNRR $_{216-227}$	DA (1), LN (1), RR (1)
NEDEEKGAIVK $_{236-246}$	EK (1), KG (1), GA (1), AI (1), VK (2)
Legumin β-subunit	
LHQNIGSSSSPDIYNPQAGR $_{322-341}$	IG (1), GS (1), IY (1), PQ (1), AG (1), GR (1)
FLKLSAEFGSLHK $_{355-367}$	LKL (1), KL (1), FG (1), GS (1), HK (1)
LSAEFGSLHK $_{358-367}$	FG (1), GS (1), HK (1)
NAMFVPHYNLNANSILYALK $_{368-387}$	MF (1), VP (1), PH (1),HY (1), LN (1), LY (1), YA (1)
NAMFVPHYNLNANSILYALKGR $_{368-389}$	MF (1), VP (1), PH (1),HY (1), LN (1), LY (1), YA (1), KG (1), GR (1)
GNSVFDGELEAGR $_{400-412}$	VF (1), DG (1), GE (1), EA (1), AG (1), GR (1)
ALIVPQNFAIAAK $_{413-425}$	VP (1), PQ (1), NF (1), AI (1), IA (1), AA (1)
SLSDRFSYVAFK $_{426-437}$	RF (1), SY (1), AF (1)
FSYVAFK $_{431-437}$	SY (1), AF (1)
Provicilin precursor	
VLLEEQEQKPK $_{186-196}$	QK (1), KP (1)
DRQQSQSQSQQEADVIVK $_{203-220}$	EA (1), VK (1)
QQSQSQSQQEADVIVK $_{205-220}$	EA (1), VK (1)
ISREQIEELSK $_{221-231}$	IE (1)
EQIEELSK $_{224-231}$	IE (1)
KSVSSESEPFNLR $_{239-251}$	-
SVSSESEPFNLR $_{240-251}$	-
SRNPIYSNK $_{252-260}$	IY (1), NK (1)
NPIYSNK $_{254-260}$	IY (1), NK (1)
YGNFFEITPEKNPQLQDLDISLNSVEIN EGSLLLPHFNSR $_{261-300}$	YG (1), NF (1), EI (2), EK (1), PQ (1), LQ (1), LN (1), VE (1), EG (1), GS (1), LLP (1), PH (1)
NPQLQDLDISLNSVEINEGSLLLPHFNS R $_{272-300}$	PQ (1), LQ (1), LN (1), VE (1), EG (1), GS (1), LLP (1), PH (1)
ATVILVVNEGK $_{301-311}$	EG (1), GK (1)
NFLAGEEDNVISQIQRPVK $_{386-404}$	NF (1), LA (1), AG (1), GE (1), RP (1), VK (1)
EVAFPGSAEEVDR $_{405-417}$	EV (1), AF (1), AFP (1), FP (1), PG (1), GS (1), EV (1)
QSHFANAQPQQK $_{424-435}$	PQ (1), QK (1)

ACE-inhibitory peptides is useful in the evaluation of these chickpea protein being potential sources of ACE-inhibitory peptides (Cheung et al., 2009).

Table 4.9 summarizes the predicted ACE-inhibitory peptides from tryptic peptide sequences identified (Sections 4.3.1.1 and 4.3.1.2); predicted ACE-inhibitory peptides EG, GG, GL, IE, NK, RF, FY, YL, LA, AG, FLR, GS, IF, SG, GF, KR, DA, LN, RR, EK, KG, GA, AI and VK are related to the tryptic peptides with origin from legumin α-subunit while predicted ACE-inhibitory peptides IG, GS, IY, PQ, AG, GR, LKL, KL, FG, HK, MF, VP, PH, HY, LN, LY, YA, KG, VF, DG, GE, EA, NF, AI, AA, RF, SY and AF are related to legumin β-subunit; peptides RF, AG, GS, LN, KG, AI are common in both legumin α- and β-subunits among these predicted peptides. Predicted ACE-inhibitory peptides QK, KP, EA, VK, IY, NK, YG, NF, EI, PQ, LQ, LN, VE, EG, GS, LLP, PH, GK, LA, AG, GE, RP, EV, AF, AFP, FP and PG are related to the tryptic peptides with origin from chickpea proviclin precursor; peptides EA, VK, IY, NK, NF, PQ, LN, EG, PH, LA, AG, GS, GE and AF are common in tryptic peptides with origin from both chickpea legumin and provicilin precursor.

4.3.5 Prediction of Potential ACE-Inhibitory Peptides Generated from Various Proteases

Table 4.10 shows that the predicted ACE-inhibitory peptides released from a selection of proteases demonstrate the ability to generate theoretical ACE-inhibitory peptides for chickpea legumin and provicilin precursor proteins; 4 of the proteases (prolyl oligopeptidase, thrombin, oligopeptidase F and prolidase L. lactis s. cremoris H61) showed no matches of peptides with ACE-inhibitory activity. The number represents the

Table 4.10: Predicted ACE-inhibitory peptides to be released from chickpea proteins using BIOPEP "enzyme action" tool

Protease[a]	Number of peptides		Peptide sequence[b]
	Legumin	Provicilin Precursor	
Bromelain	8	4	VK (1), IA (1), LA (1), RA (2), FG (2), EG (1), NG (1), SG (1), EK (1), HK (1)
Cathepsin G	8	9	VF (1), RF (1), YL (1), VK (1), GM (1), SF (2), CF(1), HK (1), GF (2), IF (1), GL (1), HL (1), EK (3)
Calpain	1	0	VK (1)
Chymase	6	0	MF (1), LF (1), LRY (1), SY (1), CF (1), LLF (1)
Chymotrypsin A	7	8	RY (1), MF (1), RF(1), SY(2), SF (2), KL (3), CF (1), GF (2), GL (1), HL (1)
Chymotrypsin C	11	15	RY (1), HY (1), RP (3), VP (1), HL (2), GE (1), KL (3), KP (2), IE (4), KE (3), IY (1), IP (1), GL (1), SY (1), VE (1)
Clostripain	2	1	FR (1), AR (1), PR (1)
Ficain	15	19	RL (3), RY (1), IA (3), RA (3), IG (2), FG (3), QG (1), EG (4), EA (1), HK (2), IPA (1), HL (1), SG (1), SY (1), EV (4), EK (3)
Glutamyl endopeptidase Ⅱ	2	2	LN (1), KE (1), IE (1), VE (1)
Glycyl endopeptidase	2	2	EG (1), NG (1), KG (1), FG (1)
Leukocyte elastase	4	5	IA (3), RA (2), GV (1), IPA (1), EV (2)
Metridin	7	8	RY (1), MF (1), RF (1), SY (2), SF (2), KL (3), CF (1), GF (2), GL (1), HL (1)
Oligopeptidase B	5	3	VK (1), FR (1), GR (1), QK (1), AR (1), IR (1), PR (1), EK (1)
Pancreatic elastase	15	12	RL (3), RA(2), YA (2), KG (3), FG (2), QG (1), EG (1), EA (1), NG (1), NKL (1), KL (3), PL (1), HL (1), EI (2), EV (3)
Pancreatic elastase Ⅱ	7	7	VF (1), RF (1), YL (1), GM (1), SF (2), KL (3), CF (1), GF (2), GL (1), HL(1)
Papain	18	11	VA (2), LA (1), PA (1), PR (2), FY (1), IA (1), LA (1), FR (1), FG (3), DA (1), QG (1), NG (1), PG (1), QK (3), IR (1), PR (1), SG (1), LG (1), DG (1), NK (1), IE (1), VE (2)
Pepsin	5	7	RF (1), YL (1), SF (2), KL (3), CF (1), GF (2), GL (1), HL (1)
Plasmin	5	3	VK (1), FR (1), GR (1), QK (1), AR (1), IR (1), PR (1), EK (1)
Proteinase K	15	20	RY (1), RF (2), HY(1), RP (2), AF (2), GM (1), AI (2), NKL (1), SY (2), SF (2), KL (3), CF (1), KP (1), EI (3), HP (1), RL (1), HL (1), GF (2), GI (2), GL (2), EV (2)
Proteinase P1	3	4	EG (1), KF (1), NK (1), RL (1), RF (1), DG (1), LN (1)
Thermolysin	15	17	AP (2), LY (2), FY (1), VK (2), AG (1), FG (2), AR (2), LQ (2), LN (3), IR (1), LH (1), IP (2), FR (1), VG (1), LG (3), VR (2), IE (1), VE (3)
Trypsin	5	3	VK (1), FR (1), GR (1), QK (1), AR (1), IR (1), PR (1), EK (1)
V-8 protease	1	2	KE (1), IE (1), VE (1)
Pepsin + trypsin + chymotrypsin A	9	14	MF (1), VK (1), GR (1), QK (1), SY (2), SF (2), AR (1), CF (1), HK (1), IR (1), PR (1), GF (2), IF (1), GL (1), HL (1), NF (1), EK (3), TF (1)

[a] Proteases (prolyl oligopeptidase, thrombin, oligopeptidase F, prolidase L. lactis s. cremoris H61) did not show any ACE-inhibitory peptides were not included in Table 4.10

[b] The number in the parentheses represents the sum of frequency of ACE-inhibitory peptides which occurred by the theoretical enzymatic cleavage of the chickpea proteins investigated during the BIOPEP analysis

theoretical peptides released by proteases from both chickpea legumin and provicilin precursor (in parentheses); proteases chymotrypsin C (26), ficain (34), pancreatic elastase (27), papain (29), proteinase K (35) and thermolysin (32) yielded a much greater number than other proteases; basically this result is agreed to the prediction from oat proteins (11S, 12S globulin and avenin) reported by Cheung et al. (2009); nevertheless the proteinase K gave the highest number of ACE-inhibitory peptides for chickpea legumin and provicilin precursor proteins whereas the thermolysin gave the most predicted peptides (34) for oat proteins. In addition, the combined action of digestive proteases (pepsin, trypsin and chymotrypsin A) gave a total of 24 predicted peptides with ACE-inhibitory activity which is comparable to the predicted peptides released from those proteases (chymotrypsin C, ficain, pancreatic elastase, papain, proteinase K and thermolysin; Table 4.10).

It has been reported that hydrophobic amino acids such as alanine (A), isoleucine (I), leucine (L), methionine (M), phenylalanine (F), tryptophan (W), tyrosine (Y), valine (V) and proline (P) within the ACE-inhibitory peptide sequences can raise greater effect on inhibition of ACE specifically when these AA are located at C-terminal positions (Saito et al., 1994); proteases ficain and proteinase K gave the liberation of containing largely the hydrophobic AA at the C-terminal positions such as leucine (L), tyrosine (Y), alanine (A), isoleucine (I), valine (V) and praline (P) (Table 5.3).

4.4 Conclusion

Tryptic peptide sequences from chickpea, oat and rice protein fractions were identified using the proteomic techniques; the tryptic peptides of chickpea and oat

globulins fractions (C-Gb and O-Gb) with origin from α- and β-subunits of chickpea legumin and oat 12S globulin 1 were identified in accordance with SDS-PAGE results in Chapter 3. The results of tryptic peptides from chickpea albumin fraction (C-Ab) showed peptide sequences corresponding to pea vicilins, pea vicilin precursor and lentil allergens (vicilins) in which these three proteins belong to 7S vicilin family; it suggests this protein fraction contain certain degree of chickpea vicilins and potential allergen proteins that are associated with 7S vicilins. Chickpea and oat glutelin fractions (C-Gt and O-Gt) showed similarities of tryptic peptides to their globulins. BLAST analysis suggests that chickpea legumin, oat 12S globulin 1 and rice glutelin precursor may share a common ancestral gene. Chickpea legumin and provicilin proteins are potential sources of releasing ACE-inhibitory peptides with enzymatic hydrolysis based on the theoretical peptides predicted using the BIOPEP analysis while chickpea legumin showed greater number of predicted ACE-inhibitory peptides than provicilin precursor; the chickpea globulin fraction (C-Gb), containing largely the legumin proteins identified in both SDS-PAGE and LC-ESI/MS/MS results, thus it would be a great source in production of potential ACE-inhibitory peptides as well as the oat globulin fractions (O-Gb) identified and surveyed previously.

GENERAL CONCLUSIONS

Chickpea and oat proteins have the potency to be bioactive ingredients or nutraceutical components as demanded for raising applications in food and nutraceutical industrials; knowledge of molecular characteristics of a food protein is fundamental before it can achieve widespread use. This research studied the molecular characteristics of sequential protein fractions (albumins, globulins and glutelins) from chickpea and oat seeds; the tryptic peptide sequences from these protein fractions were also investigated using proteomic techniques.

The results obtained from the present work have established valuable molecular information on chickpea and oat proteins. Native- and SDS-PAGE results gave well defined protein bands for both chickpea and oat protein fractions (C-Ab, C-Gb, C-Gt and O-Ab, O-Gb, O-Gt, respectively); both globulin fractions (C-Gb and O-Gb) showed molecular similarities especially the observed α- and β-subunits in SDS-PAGE results; RP-HPLC analysis also confirmed the separated protein bands with MWs related to α- and β-subunits in SDS-PAGE. Chickpea protein hydrolysates (Ab-H, Gb-H and Gt-H) from trypsin proteolysis showed degradation of protein bands observed in SDS-PAGE along with degree of hydrolysis ranging 22-28%.

The proteomic techniques, employing 1D SDS-PAGE, in-gel tyrptic digestion, LC-ESI/MS/MS analysis, and online Mascot MS/MS ion search established an approach of identifying tryptic peptides sequences from chickpea, oat and rice protein fractions; chickpea and oat protien fractions (globulins and glutelins; C-Gb, C-Gt and O-Gb, O-Gt) showed most tryptic peptides corresponding to chickpea legumin and oat 12S globulin1 proteins, respectively; chickpea albumin fraction (C-Ab) showed tryptic peptides

corresponding to vicilins (lentil and pea; 7S globulins) while oat albumin fraction (O-Ab) showed tryptic peptides corresponding to wheat, barley and rice proteins. BLAST analysis results showed that chickpea legumin shared 30% sequence similarities to oat 12S globulin and rice glutelin and suggested these three proteins may have evolved from a common ancestral gene. *In silico* BIOPEP analysis showed large numbers of di- or tri-peptides of chickpea legumin and proviclin precursor proteins with predicted ACE-inhibitory activity; some of these predicted were also found in the tryptic peptides identified from chickpea protein fractions; this suggests chickpea protein fractions (C-Ab, C-Gb and C-Gt) could be potential sources with ACE-inhibitory peptides in particular the globulin fraction (C-Gb) which was identified mostly related to chickpea legumin.

REFERENCES

Abee, T., van Schaik, W., & Siezen, R. J. (2004). Impact of genomics on microbial food safety. *Trends in Biotechnology*, *22*(12), 653-660.

Abubakar, A., Saito, T., Kitazawa, H., Kawai, Y., & Itoh, T. (1998). Structural Analysis of New Antihypertensive Peptides Derived from Cheese Whey Protein by Proteinase K Digestion. *Journal of Dairy Science*, *81*(12), 3131-3138.

Adebiyi, A. P., Adebiyi, A. O., Ogawa, T., & Muramoto, K. (2008). Purification and characterisation of antioxidative peptides from unfractionated rice bran protein hydrolysates. *International Journal of Food Science & Technology*, *43*(1), 35-43.

Agboola, S., Ng, D., & Mills, D. (2005). Characterisation and functional properties of Australian rice protein isolates. *Journal of Cereal Science*, *41*(3), 283-290.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389-3402.

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schäffer, A. A., & Yu, Y.-K. (2005). Protein database searches using compositionally adjusted substitution matrices. *FEBS Journal*, *272*(20), 5101-5109.

Amiour, N., Merlino, M., Leroy, P., & Branlard, G. (2002). Proteomic analysis of amphiphilic proteins of hexaploid wheat kernels. *Proteomics*, *2*(6), 632-641.

Arcan, I., & Yemenicioglu, A. (2007). Antioxidant activity of protein extracts from heat-treated or thermally processed chickpeas and white beans. *Food Chemistry*, *103*(2), 301-312.

Barbana, C., & Boye, J. I. (2010). Angiotensin I-converting enzyme inhibitory activity of chickpea and pea protein hydrolysates. *Food Research International*, *In Press, Accepted Manuscript* DOI: 10.1016/j.foodres.2010.1005.1003.

Beavis, R. C., & Fenyo, D. (2000). Database searching with mass-spectrometric information. *Trends in Biotechnology*, *18*(Supplement 1), 22-27.

Benito, C., González-Mancebo, E., Díaz de Durana, M. D. A., Tolón, R. M., & Fernández-Rivas, M. (2007). Identification of a 7S globulin as a novel coconut allergen. *Annals of Allergy, Asthma & Immunology*, *98*(6), 580-584.

Bensadoun, A., & Weinstein, D. (1976). Assay of proteins in the presence of interfering materials. *Analytical Biochemistry*, *70*(1), 241-250.

Beyer, K., Bardina, L., Grishina, G., & Sampson, H. A. (2002a). Identification of four major sesame seed allergens by 2D-proteomics and Edman sequencing -- Seed storage proteins as common food allergens. *Journal of Allergy and Clinical Immunology*, *109*(1, Supplement 1), S288-S288.

Beyer, K., Bardina, L., Grishina, G., & Sampson, H. A. (2002b). Identification of sesame seed allergens by 2-dimensional proteomics and Edman sequencing: Seed storage proteins as common food allergens. *Journal of Allergy and Clinical Immunology*, *110*(1), 154-159.

Beyer, K., Grishina, G., Bardina, L., Grishin, A., & Sampson, H. A. (2002c). Identification of an 11S globulin as a major hazelnut food allergen in hazelnut-induced systemic reactions. *Journal of Allergy and Clinical Immunology*, *110*(3), 517-523.

Blackstock, W. (2000). Trends in automation and mass spectrometry for proteomics. *Trends in Biotechnology*, *18*(Supplement 1), 12-17.

Blackstock, W., & Mann, M. (2000). Proteomics comes of age: the quest for function. *Trends in Biotechnology*, *18*(Supplement 1), 1-2.

Burgess, S. R., Shewry, P. R., Matlashewski, G. J., Altosaar, I., & Miflin, B. J. (1983). Characteristics of oat (Avena sativa L.) seed globulins. *Journal of Experimental Botany*, *34*(10), 1320-1332.

Byun, H.-G., & Kim, S.-K. (2002). Structure and activity of angiotensin I-converting enzyme inhibitory peptides derived from Alaskan pollack skin. *Journal of Biochemistry and Molecular Biology*, *35*(2), 239-243.

Carbonaro, M. (2004). Proteomics: present and future in food quality evaluation. *Trends in Food Science & Technology*, *15*(3-4), 209-216.

Casey, R., Domoney, C., & Smith, A. M. (1993). Peas : genetics, molecular biology and biotechnology. In: R. Casey, & D. R. Davies, *Biotechnology in agriculture series, 10* (pp. 121-164). Wallingford, Oxon, UK: CAB International.

Chandna, M., & Matta, N. K. (1990). Characterization of pearl millet protein fractions. *Phytochemistry*, *29*(11), 3395-3399.

Chang, Y.-W., Alli, I., Molina, A., Konishi, Y., & Boye, J. (2009). Isolation and characterization of chickpea (*Cicer arietinum* L.) seed protein fractions. *Food and Bioprocess Technology*, DOI: 10.1007/s11947-11009-10303-y.

Chavan, J. K., Kadam, S. S., & Salunkhe, D. K. (1989). Chickpea. In: D. K. Salunkhe, & S. S. Kadam, *CRC handbook of world food legumes: nutritional chemistry, processing technology, and utilization*, vol. I (pp. 247-288). Boca Raton, Fla.: CRC Press.

Chen, H.-M., Muramoto, K., Yamauchi, F., & Nokihara, K. (1996). Antioxidant activity of designed peptides based on the antioxidative peptide isolated from digests of a soybean protein. *Journal of Agricultural and Food Chemistry*, *44*(9), 2619-2623.

Cheung, H. S., Wang, F. L., Ondetti, M. A., Sabo, E. F., & Cushman, D. W. (1980). Binding of peptide substrates and inhibitors of angiotensin-converting enzyme. Importance of the COOH-terminal dipeptide sequence. *Journal of Biological Chemistry*, *255*(2), 401-407.

Cheung, I. W. Y., Nakayama, S., Hsu, M. N. K., Samaranayaka, A. G. P., & Li-Chan, E. C. Y. (2009). Angiotensin-I converting enzyme inhibitory activity of hydrolysates from oat (Avena sativa) proteins by in silico and in vitro analyses. *Journal of Agricultural and Food Chemistry*, *57*(19), 9234-9242.

Church, F. C., Swaisgood, H. E., Porter, D. H., & Catignani, G. L. (1983). Spectrophotometric assay using o-phthaldialdehyde for determination of proteolysis in milk and isolated milk proteins. *Journal of Dairy Science*, *66*(6), 1219-1227.

Clemente, A., Vioque, J., Sánchez-Vioque, R., Pedroche, J., Bautista, J., & Millán, F. (2000). Factors affecting the *in vitro* protein digestibility of chickpea albumins. *Journal of the Science of Food and Agriculture*, *80*(1), 79-84.

Clemente, A., Vioque, J., Sanchez-Vioque, R., Pedroche, J., & Millan, F. (1999). Production of extensive chickpea (*Cicer arietinum* L.) protein hydrolysates with reduced antigenic activity. *Journal of Agricultural and Food Chemistry*, *47*(9), 3776-3781.

Crawford, M. E., Cusick, M. E., & Garrels, J. I. (2000). Databases and knowledge resources for proteomics research. *Trends in Biotechnology*, *18*(Supplement 1), 17-21.

Cushman, D. W. (1981). Evolution of a new class of antihypertensive drugs In: Z. P. Horovitz, *Angiotensin converting enzyme inhibitors: mechanisms of action and clinical implications* (p. 19): Urban & Schwarzenberg.

Damodaran, S. (1996). Functional properties. In: S. Nakai, & H. W. Modler, *Food proteins: properties and characterization* (pp. 167-234). New York, N.Y.: VCH.

Duke, J. A. (1981). *Cicer arietinum* L., *Handbook of legumes of world economic importance* (pp. 52-57). New York: Plenum Press.

FAOSTAT (2007). Food and agricultural commodities production. Accessed: April 13th, 2010. http://faostat.fao.org/site/339/default.aspx

FitzGerald, R. J., & Meisel, H. (1999). Lactokinins: Whey protein-derived ACE inhibitory peptides. *Nahrung/Food*, *43*(3), 165-167.

FitzGerald, R. J., & Meisel, H. (2000). Milk protein-derived peptide inhibitors of angiotensin-I-converting enzyme. *British Journal of Nutrition*, *84(Suppl. 1)*, 33-37.

Gómez-Ruiz, J. Á., Ramos, M., & Recio, I. (2007). Identification of novel angiotensin-converting enzyme-inhibitory peptides from ovine milk proteins by CE-MS and chromatographic techniques. *ELECTROPHORESIS*, *28*(22), 4202-4211.

Görg, A. (2000). Advances in 2D gel techniques. *Trends in Biotechnology*, *18*(Supplement 1), 3-6.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., & Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, *31*(13), 3784-3788.

Gianazza, E., Eberini, I., Arnoldi, A., Wait, R., & Sirtori, C. R. (2003). A Proteomic Investigation of Isolated Soy Proteins with Variable Effects in Experimental and Clinical Studies. *The Journal of Nutrition*, *133*(1), 9-14.

Gohlke, P., Linz, W., Scholkens, B. A., Kuwer, I., Bartenbach, S., Schnell, A., & Unger, T. (1994). Angiotensin-converting enzyme inhibition improves cardiac function. Role of bradykinin. *Hypertension*, *23*(4), 411-418.

Goulet, G., Ponnampalam, R., Amiot, J., Roy, A., & Brisson, G. J. (1987). Nutritional value of acylated oat protein concentrates. *Journal of Agricultural and Food Chemistry*, *35*(4), 589-592.

Guéguen, J. (1991). Pea and fafabean proteins. In: B. J. F. Hudson, *Developments in food proteins. Vol. 7* (pp. 35-78). London: Elsevier applied science.

Gygi, S. P., & Aebersold, R. (2000). Using mass spectrometry for quantitative proteomics. *Trends in Biotechnology*, *18*(Supplement 1), 31-36.

Haard, N. F. (1999). Cereals: Rationale for fermentation. In: N. F. Haard, *Fermented cereals: a global perspective*, vol. 1 (pp. 1-27). Rome: Food and Agriculture Organization of the United Nations.

Hao, X., Li, J., Shi, Q., Zhang, J., He, X., & Ma, H. (2009). Characterization of a novel legumin & alpha-amylase inhibitor from chickpea (Cicer arietinum L.) seeds. *Bioscience, Biotechnology, and Biochemistry*, *73*(5), 1200-1202.

Herbert, B. R., Harry, J. L., Packer, N. H., Gooley, A. A., Pedersen, S. K., & Williams, K. L. (2001). What place for polyacrylamide in proteomics? *Trends in Biotechnology*, *19*(10), S3-S9.

Higuchi, W., & Fukazawa, C. (1987). A rice glutelin and a soybean glycinin have evolved from a common ancestral gene. *Gene*, *55*(2-3), 245-253.

Incamps, A., Florence, H.-J., Chagvardieff, P., Rambourg, J.-C., Dedieu, A., Linares, E., & Quéméneur, E. (2005). Industrial process proteomics: Alfalfa protein patterns during wet fractionation processing. *Biotechnology and Bioengineering*, *91*(4), 447-459.

Iwaniak, A., & Dziuba, J. (2009). Motifs with potential physiological activity in food proteins-BIOPEP database. *ACTA Scientiarum Polonorum, Technologia Alimentaria*, *8*(3), 59-85.

Iwaniak, A., Dziuba, J., & Niklewicz, M. (2005). The BIOPEP database - a tool for the in silico method of classification of food proteins as the source of peptides with antihypertensive activity. *Acta Alimentaria*, *34*(4), 417-425.

Kaur, M., & Singh, N. (2007). Characterization of protein isolates from different Indian chickpea (*Cicer arietinum* L.) cultivars. *Food Chemistry*, *102*(1), 366-374.

Kjærsgård, I. V. H., Nørrelykke, M. R., & Jessen, F. (2006). Changes in cod muscle proteins during frozen storage revealed by proteome analysis and multivariate data analysis. *6*(5), 1606-1618.

Kohmura, M., Nio, N., & Ariyoshi, Y. (1990). Inhibition of angiotensin-converting enzyme by synthetic peptide fragments of various beta-caseins. *Agricultural and Biological Chemistry*, *54*(4), 1101-1102.

Kohmura, M., Nio, N., Kubo, K., Minoshima, Y., Munekata, E., & Ariyoshi, Y. (1989). Inhibition of angiotensin-converting nnzyme by synthetic peptides of human β-casein. *Agricultural and Biological Chemistry*, *53*, 2107-2114.

Korhonen, H., Pihlanto-Leppälä, A., Rantamäki, P., & Tupasela, T. (1998). Impact of processing on bioactive proteins and peptides. *Trends in Food Science & Technology*, *9*(8-9), 307-319.

Koundal, K. R., & Shasany, A. K. (1997). Protein Knowledgebase (UniProtKB). Accessed: March 8th, 2010. http://www.uniprot.org/uniprot/O04219

Kuiper, H. A., Noteborn, H. P. J. M., Kok, E. J., & Kleter, G. A. (2002). Safety aspects of novel foods. *Food Research International*, *35*(2-3), 267-271.

Kulp, K., & Ponte, J. G. (2000). *Handbook of cereal science and technology*. New York: Marcel Dekker.

Kussmann, M., Raymond, F., & Affolter, M. (2006). OMICS-driven biomarker discovery in nutrition and health. *Journal of Biotechnology*, *124*(4), 758-787.

Kvasnička, F. (2003). Proteomics: general strategies and application to nutritionally relevant proteins. *Journal of Chromatography B*, *787*(1), 77-89.

López-Torrejón, G., Salcedo, G., Martín-Esteban, M., Díaz-Perales, A., Pascual, C. Y., & Sánchez-Monge, R. (2003). Len c 1, a major allergen and vicilin from lentil seeds: Protein isolation and cDNA cloning. *Journal of Allergy and Clinical Immunology*, *112*(6), 1208-1215.

Lásztity, R. (1996). Oat proteins. *The chemistry of cereal proteins* (pp. 275-292). Boca Raton, Fla.: CRC Press.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, *227*(5259), 680-685.

Li, Y., Jiang, B., Zhang, T., Mu, W., & Liu, J. (2008). Antioxidant and free radical-scavenging activities of chickpea protein hydrolysate (CPH). *Food Chemistry*, *106*(2), 444-450.

Liu, L. H., Hung, T. V., & Bennett, L. (2008). Extraction and characterization of chickpea (*Cicer arietinum*) albumin and globulin. *Journal of Food Science*, *73*(5), C299-C305.

Lookhart, G. L. (1985). Identification of oat cultivars by combining polyacrylamide gel electrophoresis and reversed-phase high-performance liquid chromatography. *Cereal Chemistry*, *62*(5), 345-350.

Loponen, J. (2004). Angiotensin converting enzyme inhibitory peptides in Finnish cereals: a database survey. *Agricultural and Food Science*, *13*, 39-45.

Ma, C. Y., & Harwalkar, V. R. (1984). Chemical characterization and functionality assessment of oat protein fractions. *Journal of Agricultural and Food Chemistry*, *32*(1), 144-149.

Ma, M.-S., Bae, I. Y., Lee, H. G., & Yang, C.-B. (2006). Purification and identification of angiotensin I-converting enzyme inhibitory peptide from buckwheat (*Fagopyrum esculentum* Moench). *Food Chemistry*, *96*(1), 36-42.

Maiti, R. K., & Wesche-Ebeling, P. (2001). The chickpea crop. In: R. K. Maiti, & P. Wesche-Ebeling, *Advances in chickpea science* (pp. 1-32). Enfield, NH, USA: Science Publishers.

Mandaokar, A. D., & Koundal, K. R. (1997). NCBI: National Center for Biotechnology Information. Accessed: March 8th, 2010. *http://www.ncbi.nlm.nih.gov/protein/CAB60140.1*

Maria, T. G.-A., Natascia, C., Lello, Z., Corrado, C., Carlo, I., & Pier, G. R. (2005). Two-dimensional mapping as a tool for classification of green coffee bean species. *Proteomics*, *5*(3), 710-718.

Martinez, I., & Friis, J. T. (2004). Application of proteome analysis to seafood authentication. *Proteomics*, *4*(2), 347-354.

Matsufuji, H., Matsui, T., Seki, E., Osajima, K., Nakashima, M., & Osajima, Y. (1994a). Angiotensin I-converting enzyme inhibitory peptides in an alkaline proteinase hydrolysate derived from sardine muscle. *Bioscience, Biotechnology, and Biochemistry*, *58*, 2244-2245.

Matsufuji, H., Matsui, T., Seki, E., Osajima, K., Nakashima, M., & Osajima, Y. (1994b). Angiotensin I-converting enzyme inhibitory peptides in an alkaline proteinase hydrolysate derived from sardine muscle. *Bioscience, Biotechnology,and Biochemistry*, *58*, 2244-2245.

Meisel, H. (1993). Casokinins as bioactive peptides in the primary strucure of casein. In: *Food proteins: structure and functionality* (pp. 67-75). Weinheim; New York: VCH.
Meisel, H. (1998). Overview on Milk Protein-derived Peptides. *International Dairy Journal*, *8*(5-6), 363-373.

Meisel, H., Walsh, D. J., Murray, B., & FitzGerald, R. J. (2006). ACE inhibitory peptides. In: Y. Mine, & F. Shahidi, *Nutraceutical proteins and peptides in health and disease* (pp. 269-315). Boca Raton: CRC/Taylor and Francis.

Miyoshi, S., Kaneko, T., Ishikawa, H., Tanaka, H., & Maruyama, S. (1995). Production of bioactive peptides from corn endosperm proteins by some proteases. *Annals of the New York Academy of Sciences*, *750*(Enzyme Engineering XII), 429-431.

Murray, B. A., & FitzGerald, R. J. (2007). Angiotensin converting enzyme inhibitory peptides derived from food proteins: biochemistry, bioactivity and production. *Current Pharmaceutical Design*, *13*(8), 773-791.

Nakahara, T., Sano, A., Yamaguchi, H., Sugimoto, K., Chikata, H., Kinoshita, E., & Uchida, R. (2009). Antihypertensive effect of peptide-enriched soy sauce-like seasoning and identification of its angiotensin I-converting enzyme inhibitory substances. *Journal of Agricultural and Food Chemistry*, *58*(2), 821-827.

Natale, M., Bisson, C., Monti, G., Peltran, A., Garoffo, L. P., Valentini, S., Fabris, C., Bertino, E., Coscia, A., & Conti, A. (2004). Cow's milk allergens identification by two-dimensional immunoblotting and mass spectrometry. *Molecular Nutrition & Food Research*, *48*(5), 363-369.

Natarajan, S., Xu, C., Bae, H., & Bailey, B. A. (2006). Proteomic and genomic characterization of Kunitz trypsin inhibitors in wild and cultivated soybean genotypes. *Journal of Plant Physiology*, *164*(6), 756-763.

Nogata, Y., Nagamine, T., Yanaka, M., & Ohta, H. (2009). Angiotensin I converting enzyme inhibitory peptides produced by autolysis reactions from wheat bran. *Journal of Agricultural and Food Chemistry*, *57*(15), 6618-6622.

Ondetti, M. A., Williams, N. J., Sabo, E., Pluscec, J., Weaver, E. R., & Kocy, O. (1971). Angiotensin-converting enzyme inhibitors from the venom of Bothrops jararaca. Isolation, elucidation of structure, and synthesis. *Biochemistry*, *10*(22), 4033-4039.

Patil, S. P., Niphadkar, P. V., & Bapat, M. M. (2001). Chickpea: A major food allergen in the Indian subcontinent and its clinical and immunochemical correlation. *Annals of Allergy, Asthma, & Immunology*, *87*(2), 140-145.

Perkins, D. N., Pappin, D. J. C., Creasy, D. M., & Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, *20*(18), 3551-3567.

Pernollet, J. C., Kim, S. I., & Mosse, J. (1982). Characterization of storage proteins extracted from Avena sativa seed protein bodies. *Journal of Agricultural and Food Chemistry*, *30*(1), 32-36.

Peterson, D. M. (1978). Subunit structure and composition of oat seed globulin. *Plant Physiology*, *62*(4), 506-509.

Pihlanto-Leppälä, A., Antila, P., Mäntsälä, P., & Hellman, J. (1994). Opioid peptides produced by in-vitro proteolysis of bovine caseins. *International Dairy Journal*, *4*(4), 291-301.

Pihlanto-Leppäläa, A., Rokka, T., & Korhonen, H. (1998). Angiotensin I converting enzyme inhibitory peptides derived from bovine milk proteins. *International Dairy Journal*, *8*(4), 325-331.

Pruitt, K. D., Tatusova, T., & Maglott, D. R. (2005). NCBI reference sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Research*, *33*(suppl_1), D501-504.

Rehm, H. (2006). Chapter 8: Subunits. *Protein biochemistry and proteomics* (pp. 191-200). Amsterdam: Academic Press.

Resing, k. A., & Ahn, N. G. (2004). Protein identification by in-gel digestion and mass spectrometry. In: D. W. Speicher, *Proteome analysis : interpreting the genome* (pp. 163-179). Amsterdam; Boston: Elsevier.

Righetti, P. G., & Bossi, A. (1997). Isoelectric focusing in immobilized pH gradients: an update. *Journal of Chromatography B: Biomedical Sciences and Applications*, *699*(1-2), 77-89.

Robert, L. S., Nozzolillo, C., & Altosaar, I. (1985). Homology between rice glutelin and oat 12 S globulin. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, *829*(1), 19-26.

Rosenfeld, J., Capdevielle, J., Guillemot, J. C., & Ferrara, P. (1992). In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Analytical Biochemistry*, *203*(1), 173-179.

Sánchez-Vioque, R., Clemente, A., Vioque, J., Bautista, J., & Millán, F. (1999). Protein isolates from chickpea (*Cicer arietinum* L.): chemical composition, functional properties and protein characterization. *Food Chemistry*, *64*(2), 237-243.

Saito, Y., Wanezaki, K., Kawato, A., & Imayasu, S. (1994). Structure and activity of angiotensin I converting enzyme inhibitory peptides from sake and sake lees. *Bioscience, Biotechnology, and Biochemistry*, *58*(10), 1767-1771.

Schubert, R., Baumlein, H., Czihal, A., & Wobus, U. (1990). Genomic sequence of a 12S seed storage protein gene from oat (Avena sativa L. cv. 'Solidor'). *Nucleic Acids Research*, *18*(2), 377.

Sentandreu, M. A., & Toldr, F. (2007). Evaluation of ACE inhibitory activity of dipeptides generated by the action of porcine muscle dipeptidyl peptidases. *Food Chemistry*, *102*(2), 511-515.

Shevchenko, A., Wilm, M., Vorm, O., & Mann, M. (1996). Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Analytical Chemistry*, *68*(5), 850-858.

Shotwell, M. A., Afonso, C., Davies, E., Chesnut, R. S., & Larkins, B. A. (1988). Molecular Characterization of Oat Seed Globulins. *Plant Physiol.*, *87*(3), 698-704.

Singh, G. D., Wani, A. A., Kaur, D., & Sogi, D. S. (2008). Characterisation and functional properties of proteins of some Indian chickpea (*Cicer arietinum*) cultivars. *Journal of the Science of Food and Agriculture*, *88*(5), 778-786.

Singh, U., & Jambunathan, R. (1982). Distribution of seed protein fractions and amino acids in different anatomical parts of chickpea (*Cicer arietinum* L.) and pigeonpea (*Cajanus cajan* L.). *Qualitas Plantarum Plant Foods for Human Nutrition*, *31*, 347-354.

Smith, R. D. (2002). Trends in mass spectrometry instrumentation for proteomics. *Trends in Biotechnology*, *20*(12), s3-s7.

Steenson, D. F., & Sathe, S. K. (1995). Characterization and Digestibility of Basmati Rice (*Oryza sativa L. var. Dehraduni*) Storage Proteins. *Cereal Chemistry*, *72*(3), 275-280.

Takaiwa, F., Ogawa, M., & Okita, T. W. (1999). Rice glutelin. In: P. R. Shewry, & R. Casey, *Seed proteins* (pp. 401-425). Dordrecht; Boston: Kluwer Academic.

Tanchak, M. A., Giband, M., Potier, B., Schernthaner, J. P., Dukiandjiev, S., & Altosaar, I. (1995). Genomic clones encoding 11S globulins in oats (Avena sativa L.). *Genome*, *38*(4), 627-634.

Tezcucano Molina, A. C., Alli, I., Konishi, Y., & Kermasha, S. (2007). Effect of dephosphorylation on bovine casein. *Food Chemistry*, *101*(3), 1263-1271.

Twyman, R. M. (2004). Principles of proteomics. *Advanced text* (pp. 1-168). New York: BIOS Scientific Publishers.

Ukeda, H., Matsuda, H., Osajima, K., Matsufuji, H., Matsui, T., & Osajima, Y. (1992). Peptides from peptic hydrolyzate of heated sardine meat that inhibit angiotensin I converting enzyme. *Nippon Nogeikagaku Kaishi*, *66*, 25-29.

van Platerink, C., Janssen, H.-G., & Haverkamp, J. (2008). Application of at-line two-dimensional liquid chromatography–mass spectrometry for identification of small hydrophilic angiotensin I-inhibiting peptides in milk hydrolysates. *Analytical and Bioanalytical Chemistry*, *391*(1), 299-307.

Villareal, R. M., & Juliano, B. O. (1978). Properties of glutelin from mature and developing rice grain. *Phytochemistry*, *17*(2), 177-182.

Vioque, J., Sánchez-Vioque, R., Clemente, A., Pedroche, J., Bautista, J., & Millán, F. (1999). Purification and partial characterization of chickpea 2S albumin. *Journal of Agricultural and Food Chemistry*, *47*(4), 1405-1409.

Walburg, G., & Larkins, B. A. (1983). Oat seed globulin: Subunit characterization and demonstration of its synthesis as a precursor. *Plant Physiology*, *72*(1), 161-165.

Walburg, G., & Larkins, B. A. (1986). Isolation and characterization of cDNAs encoding oat 12S globulin mRNAs. *Plant Molecular Biology*, *6*(3), 161-169.

Wang, W., & Gonzalez de Mejia, E. (2005). A new frontier in soy bioactive peptides that may prevent age-related chronic diseases. *Comprehensive Reviews in Food Science and Food Safety*, *4*(4), 63-78.

Wen, T.-N., & Luthe, D. S. (1985). Biochemical characterization of rice glutelin. *Plant Physiology*, *78*(1), 172-177.

Wiseman, A. (2004). Novel food-proteomics by control of bioprocess-enzymology outcomes? *Trends in Food Science & Technology*, *15*(1), 44-45.

Wolf, W. J. (1970). Soybean proteins: Their functional, chemical, and physical properties. *Journal of Agricultural and Food Chemistry*, *18*(6), 969-976.

Wu, H., He, H.-L., Chen, X.-L., Sun, C.-Y., Zhang, Y.-Z., & Zhou, B.-C. (2008). Purification and identification of novel angiotensin-I-converting enzyme inhibitory peptides from shark meat hydrolysate. *Process Biochemistry*, *43*(4), 457-461.

Xiong, J. (2006). *Essential bioinformatics*. New York: Cambridge University Press.

Yano, S., Suzuki, K., & Funatsu, G. (1996). Isolation from a-zein of thermolysin peptides with angiotensin I-converting enzyme inhibitory activity. *Bioscience, Biotechnology, and Biochemistry*, *60*(4), 661-663.

Ye, X. Y., Ng, T. B., & Rao, P. F. (2002). Cicerin and arietin, novel chickpea peptides with different antifungal potencies. *Peptides*, *23*(5), 817-822.

You, J., Li, D., Qiao, S., Wang, Z., He, P., Ou, D., & Dong, B. (2008). Development of a monoclonal antibody-based competitive ELISA for detection of [beta]-conglycinin, an allergen from soybean. *Food Chemistry*, *106*(1), 352-360.

Yust, M. M., Pedroche, J., Giron-Calle, J., Alaiz, M., Millan, F., & Vioque, J. (2003). Production of ace inhibitory peptides by digestion of chickpea legumin with alcalase. *Food Chemistry*, *81*(3), 363-369.

Zhao, W.-M., Gatehouse, J. A., & Boulter, D. (1983). The purification and partial amino acid sequence of a polypeptide from the glutelin fraction of rice grains; homology to pea legumin. *FEBS Letters*, *162*(1), 96-102.

APPENDIX I

Mass spectrometry glossary:

Glossary

Data-dependent scanning
The mass-spectrometer software makes a choice in real time about which ion to fragment and optimizes the parameters to do this based on charge state and mass. A human operator could not make these decisions in the time available during the elution of a chromatographic peak.

Electrospray ionization (ESI)
Applying a voltage to a fine needle containing a dilute solution of peptides or proteins results in a spray of droplets, typically containing just a few molecules. Repeated breakup of the droplets caused by evaporation eventually leads to the release of intact peptide or protein ions into the gas phase, from where they are sampled into the mass spectrometer. Coupling the needle to a low-flow-rate chromatography system permits on-line analysis of complex protein or peptide mixtures. The electrospray source operates at atmospheric pressure, making interfacing relatively simple.

Matrix-assisted laser desorption–ionization (MALDI)
A dilute protein or peptide sample is mixed with a large excess of UV-absorbing matrix (typically a cinnamic-acid derivative) and allowed to dry to a small spot. A pulsed laser (usually at 337 nm) illuminates the spot in a vacuum and protein or peptide ions are entrained in the resulting gas plume and extracted by an electrical field into the mass analyser.

Nanoelectrospray ionization (nESI)
A very-low-flow-rate version of electrospray that allows many tandem mass-spectrometry experiments to be carried out over a long time (40–60 min μl^{-1}). The method has been used for the detailed study of phosphorylation using specific precursor-ion scans.

Peptide-mass fingerprint (PMF)
A signature of a protein, based on measuring the masses of the tryptic peptides and comparing them against a database of tryptic peptides created by *in silico* cleavage of the proteins in a database.

Post-source decay (PSD)
The fragmentation of ions from a MALDI source by virtue of their internal energy in the field-free region of the flight path. The fragmentation cannot easily be controlled and the fragmentation pattern can be difficult to interpret.

Precursor-ion scan
An operating mode of a triple quadrupole mass spectrometer that is often used to select ions that are diagnostic of peptide phosphorylation (mass:charge ratio = 79), for example.

Product-ion scan
An operating mode of a triple quadrupole mass spectrometer that generates fragment ions by colliding a selected precursor ion with gas (often argon) in a collision cell. The fragment masses describe the peptide sequence.

Tandem mass spectrometry
Two integrated mass analysers, the first for ion selection. The selected ion is fragmented by collision with gas and the second analyser measures the fragment-ion masses. The two analysers might differ (e.g. quadrupole time-of-flight) or be the same (e.g. triple quadrupole, the third quadrupole being the gas collision cell).

Triple quadrupole
One version of a tandem mass spectrometer, valuable for precursor scans.

Quadrupole time-of-flight
State-of-the-art tandem mass spectrometer. The time-of-flight second analyser offers better sensitivity and resolution than a quadrupole.

(Blackstock, 2000)