

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 March 2006 (23.03.2006)

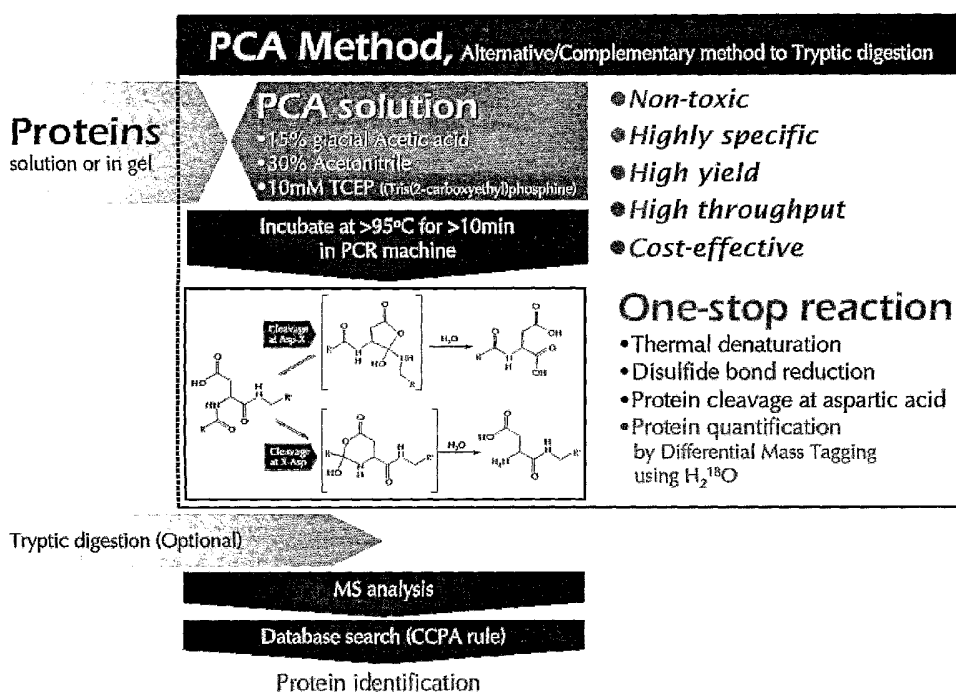
PCT

(10) International Publication Number
WO 2006/031063 A1

- (51) International Patent Classification⁷: **G01N 33/483**, C07K 1/12, H01J 49/00
- (21) International Application Number:
PCT/KR2005/003042
- (22) International Filing Date:
14 September 2005 (14.09.2005)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/610,306 15 September 2004 (15.09.2004) US
- (71) Applicant (for all designated States except US): **SIG-MOL, Inc.** [KR/KR]; Dept. Life Science, 116, Pohang Univ. of Sci. & Tech., San 31, Hyoja-Dong, Nam-Gu, Pohang, Kyungbuk 790-784 (KR).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **KWON, Joseph** [KR/KR]; 111/505 Uiduk Samsung APT., Yugeum-Ri, Kang dong-Myun, Kyungju, Kyungbuk 780-911 (KR). **LEE, Taehoon** [KR/KR]; #26 Steelhouse, 969-5, Jigok-Dong, Nam-Gu, Pohang, Kyungbuk 790-835 (KR).
- (74) Agent: **YOU ME PATENT & LAW FIRM**; Seolim Bldg., 649-10, Yoksam-dong, Kangnam-ku, Seoul 135-080 (KR).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— with international search report

[Continued on next page]

(54) Title: PROTEIN CLEAVAGE AT ASPARTIC ACID USING CHEMICAL REAGENTS



(57) Abstract: The present invention relates to the methods of identifying and quantifying polypeptides in a given sample by mass spectrometric analysis. More specifically, the invention provides the methods for sample preparation for proteomic analysis: the methods for the fragmentation of proteins into peptides with the specific cleavage rule (cleavage at amino-terminal or carboxyl-terminal of aspartic acid), which are suitable for the analysis by mass spectrometry apparatus.



For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

PROTEIN CLEAVAGE AT ASPARTIC ACID
USING CHEMICAL REAGENTS

CROSS REFERENCE TO RELATED APPLICATION

5 The present application claims priority to and the benefit of United States provisional application No. 60/610,306 filed in the United State Patent and Trademark Office on September 15, 2004, the entire content of which is incorporated hereinto by reference.

10 **FIELD OF INVENTION**

 The present invention provides a method of processing proteins for identification and quantification. To be specific, the invention relates to a kit and an apparatus for processing proteins into peptides by using chemical reagents, and to
15 the methods of using them in broad range of proteomics researches.

BACKGROUND OF THE INVENTION

 Although complete genomic sequencing can provide useful information for
20 the prediction of genes in a given species, the sequences alone do not explain the mechanisms underlying biological and pathophysiological processes, because neither the quantity nor the molecular details of the translated protein product such as structure, functional activity, state of post-translational modification can be precisely predicted. The new discipline of proteomics aims to unravel biochemical
25 information at the molecular level. Therefore, the understanding of the proteome in a given context is essential to assess the physiological state of a cell or organism.

 At present, a number of techniques have been developed to address the growing need to identify proteins more quickly and accurately through mass-
30 accurate methods such as mass spectrometry. The platforms used for proteomic analysis involve the integration of two broad practices, including separation and identification of proteins in a sample. Two-dimensional gel electrophoresis (2-DE)

or liquid chromatography (LC) of one or more types are used for the separation of proteins prior to mass spectrometry. In 2-DE, the proteins placed in a gel migrate depending largely on molecular weight and isoelectric point, thus generating a characteristic gel pattern. Mass spectrometry is generally used for the identification of proteins. In mass spectrometry, proteins or peptides are ionized and ionized species are subject to electric and/or magnetic fields in a vacuum. Their molecular weights can be deduced from the travel path of the ions. The identities of proteins or peptides can be disclosed by a peptide fingerprinting method or de novo sequencing using mass spectrometry (MS) or tandem MS (MS/MS).

In a widely used strategy for sample preparation in proteomic research, the proteins are enzymatically cleaved into their constituent peptides prior to MS analysis to enhance the likelihood that at least some of the protein is sufficiently ionized so as to be detected by generating the peptides compatible for an accurate analysis in general MS apparatus which has a limited ranges of mass measurement. Peptide mass mapping of proteins separated by polyacrylamide gel electrophoresis with enzymatic digestion has been a routine procedure for protein characterization. The most frequently used protease is trypsin because of its well-defined specificity and the appropriate size of tryptic peptides for mass spectrometric analysis. A few other commercially available proteases such as Lys-C, Glu-C, and Asp-N are discriminated against certain substrates. Having a lot of strong points, tryptic digestion is not a method of choice for hydrophobic or very basic proteins. In addition, protease autolysis products (for example, trypsin; 261.14, 514.32, 841.50, 905.50, 1005.48, 1044.56, 1468.72, 1735.84, 1767.79, 2157.02, 2210.10, 2282.17, 3012.32, 4474.09, 4488.11 Da) sometimes interfere with spectrum interpretation. Furthermore, some buffers required for efficient and specific proteolysis may generate chemical noises, thus requiring additional purification before mass spectrometric analysis.

As an alternative to enzymatic digestion, new attempts of protein analysis that use acid hydrolysis have recently been made. Several different approaches have been reported that pertain to efficient cleavage of protein. Bark et al. developed a high-temperature proteolytic digestion method using thermolysin.

Gobom et al. suggested vapor-phase acid hydrolysis using pentafluoro propionic acid (PFPA), but three different types of cleavages, including sequence ladder, were observed, thus leading to the unwanted drawbacks of increased spectrum complexity. Aiqun Li et al. suggested chemical cleavage at aspartic acid with 2(v/v)% formic acid, but unpredictable formylated fragments, which make it very difficult to identify the proteins using peptide-mass fingerprinting (PMF), were generated. Moreover, the formic acid method does not show clear cleavage rule with sequencing data.

In other points of view, various analytical approaches have been reported that utilize bottom-up proteomics and stable isotope labeling to perform relative quantification of proteins. These methods can be broadly classified as either (1) metabolic, where the isotope label is incorporated during protein synthesis, (2) amino acid-specific, where the stable isotope label is applied only to peptides containing a specific amino acid, such as cysteine in the case of the ICATTM method, and (3) global labeling methods where the label is applied to every peptide in a given proteome. As used herein an "isotope tag" refers to a chemical moiety having suitable chemical properties for incorporation of an isotope, allowing the generation of differentially tagged polypeptides in two samples. The isotope tag also has an appropriate composition to allow incorporation of a stable isotope at one or more atoms. A particularly useful stable isotope pair is hydrogen and deuterium, which can be readily distinguished using mass spectrometry, for example, ¹³C, ¹⁵N, ¹⁷O, ¹⁸O or ³⁴S. Amino acid-specific methods such as ICATTM have the advantage of reducing sample complexity but have the disadvantage of discriminating against proteins with low number of cysteines. Proteins can be also isotopically labeled at the C-termini of the tryptic peptides. One method of global labeling inserts an isotopic label via the molecule of water that is incorporated into peptides during cleavage of amide backbones by enzyme. While chymotrypsin and Asp-N incorporate only one ¹⁸O atom, trypsin, Glu-C or Lys-C can incorporate two ¹⁸O atoms into the C-termini of the resulting peptides. Moreover, ¹⁸O labels in the carboxylate groups of peptides and amino acids are resistant to back exchange. Thus under common conditions for liquid chromatography, electrospray ionization (ESI), and matrix-assisted laser

desorption/ionization (MALDI), covalent bonds between oxygen atoms and carbonyl carbon in a C-terminal carboxylate group are stable. The practice of ^{18}O -labeling is receiving increasing attention as a preferred method for heavy isotope labeling and several examples of its application have been published. It is common practice and considered advantageous to use the highest enrichment of H_2^{18}O as possible in order to achieve the highest degree of labeling of each proteolytic fragment for a quantitative application of proteomics.

BRIEF DESCRIPTION OF DRAWING

Figure 1 is a flow chart for the identification of proteins by proteomic analysis using protein cleavage at aspartic acid (PCA) method.

Figures 2A, 2B, and 2C are MALDI-MS and MALDI-TOF/TOF analysis of BSA (Bovine Serum albumin) by PCA method. Spectra were obtained utilizing an Applied Biosystems 4700 Proteomics AnalyzerTM: Panel A; MALDI mass spectra (mass range 800-4300Da); Panel B; Database search results by PMF with MASCOTTM; Panel C; For MS/MS(MALDI-TOF/TOF) analyses of the 1723.86 Da monoisotopic peak, spectra were obtained by the accumulation of 5000 consecutive laser shots at a collision energy of 1kV with air serving as the collision gas.

Figures 3A and 3B are MALDI-MS analysis of Ubiquitin by PCA method. Spectra were obtained utilizing Applied Biosystems 4700 Proteomics AnalyserTM: Panel A; MALDI mass spectra (mass range: 400-4300Da); Panel B; Database search results by PMF (peptide mass fingerprint) with MASCOTTM.

Figures 4A, 4B and 4C are database search results with MASCOTTM: Protein sample; (BSA) Bovine Serum Albumin; Panel A; Database search result with MALDI-TOF mass list (above S/N 10) of BSA digested through PCA method. [Sequence coverage: 40%]; Panel B; Database search result with MALDI-TOF mass list (above S/N 10) of BSA digested by trypsin. [Sequence coverage: 32%]; Panel C; Database search result with MALDI-TOF mass list (above S/N 10) obtained by combination PCA method and tryptic digestion. PCA/trypsin method's database search result shows high sequence coverage [87%]

Figure 5 shows detergent-free and Chaotropic reagent-free chemical digestion of membrane proteins. De-lipidated membrane protein aggregates

obtained from mouse brain lipid raft were digested in PCA solution (15(v/v)% of acetic acid (pH 2.0), 30(v/v)% of acetonitrile, and TCEP (10mM)) were incubated at 99°C for 4hrs by using PCR machine. The chemically digested membrane proteins were analyzed by SDS-PAGE, and visualized by CBB-staining.

Figure 6 shows UV chromatogram from a reverse-phase HPLC separation of the hydrolysate obtained from PCA reaction of de-lipidated membrane protein aggregates. About 1mg of the protein hydrolysate was injected. Pooled fractions were indicated by the circled number.

Figures 7A, 7B and 7C are MALDI-TOF/TOF detection of ubiquitin by PCA method: Panel A; MS spectra of ubiquitin processed by PCA method; Panel B; MS spectra of ubiquitin obtained by PCA-DMT (Differential Mass Tagging) method; Panel C and D; 1:1 and 1:2 mixture of PCA/PCA-DMT cleaved ubiquitin.

Figure 8 shows the result of LC-MS/MS experiment using the hydrolysates obtained from sequential digestion of mouse brain lipid raft by PCA and Trypsin: Panel A; Total ion chromatogram of tryptic hydrolysate of each pooled fraction in Figure 4. Each pooled fraction was indicated by circled number ;Panel B; Extracted ion chromatogram at a marked time in panel A ; Panel C; Tandem MS/MS result of the ion circled in panel B.

Figure 9 is an assignment of the mass of peptide fragments of ubiquitin obtained by PCA method [sequence coverage is 100%]

Figures 10A and 10B are an assignment of the mass of observed peptide fragments of BSA (Bovine Serum Albumin) by PCA method. The peptide mass list obtained from MALDI-TOF and the identities of some peptide were verified from de novo sequencing by tandem MS.

Figures 11A and 11B are the selected list of mouse brain lipid raft proteins identified by LC-MS/MS of the hydrolysates obtained from sequential digestion by PCA and Trypsin.

SUMMARY OF THE INVENTION

An objective of this invention is to provide the method, kit, and apparatus for acid hydrolysis of proteins, which can guarantee the strict specificity of cleavage at aspartyl residue without the production of unpredictable modification

of the peptides.

The invention provides the optimal composition of reagents for acid hydrolysis of proteins, referred to herein as protein cleavage at aspartic acid (PCA), for protein identification and quantification. The identification of proteins in a given sample can be achieved by de novo sequencing of the peptides generated or by peptide mass fingerprinting from MS analysis results by accommodating newly developed rules of fragmentation.

The present invention includes the designing of apparatus for PCA, which is developed for incubating the solution above 95°C with minimizing the loss of vapor pressure by heating the lid as well as bath simultaneously in the same temperature. The present method provides handy and simple procedure for processing of proteins prior to MS analysis comprising of just a few hours of incubation and sample dry. Furthermore, PCA can be used in combination with tryptic digestion to generate the peptides suitable for tandem MS analysis in order to get enough information for the detailed structural analysis of proteins.

The invention further provides methods for quantifying proteins in a sample by adopting the concept of ^{18}O -labeling of proteins using H_2^{18}O during hydrolysis for comparative proteomics.

DETAILED DESCRIPTION

The present invention provides a polypeptide hydrolyzing composition comprising an acid component, water miscible organic solvent and a reducing agent. The acid component is trifluoroacetic acid, phosphoric acid, propionic acid, HCl, o-iodobenzoic acid, glacial acetic acid, or any acid having buffering capacity near pH 2. Preferably, the acid component can be a mixture of trifluoroacetic acid, phosphoric acid, propionic acid, HCl, and o-iodobenzoic acid.

pH of hydrolyzing solution at time of reaction is in the range of 1.5 to 2.5. The hydrolyzing solution comprises at least 2 to 30(v/v)% glacial acetic acid. The hydrolyzing solution comprises 15(v/v)% glacial acetic acid, pH 2.0.

The water miscible organic solvent is Acetonitrile, DMF (Dimethyl formamide), DMSO (Dimethylsulfoxide), THF (Tetrahydrofurane), or an alcohol. The alcohol is methanol or ethanol. For example, the water miscible organic solvent

is at least 5-70(v/v)% Acetonitrile, preferably 30(v/v)% Acetonitrile.

The reducing agent is TCEP (Tris(2-carboxyethyl)phosphine), DTT (Dithiothreitol), or beta-Mercaptoethanol. The reducing agent is phosphine compound which can work at acidic pH range (1.5-2.5) such as TCEP (Tris(2-carboxyethyl)phosphine). The reducing agent is at least 1mM-1M TCEP (Tris(2-carboxyethyl)phosphine) or DTT (Dithiothreitol). The reducing agent is at least 10 mM TCEP (Tris(2-carboxyethyl)phosphine) or DTT (Dithiothreitol).

For example, the composition comprises about 2-30(v/v)% acetic acid of the hydrolyzing solution, about 5-70(v/v)% acetonitrile of the hydrolyzing solution, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine). The composition comprises about 15(v/v)% of acetic acid of the hydrolyzing solution, acetonitrile at an amount of about 5-70(v/v)% of acetonitrile of the hydrolyzing solution, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine). The composition comprises about 15(v/v)% of acetic acid of the hydrolyzing solution, about 30(v/v)% of acetonitrile of the hydrolyzing solution, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine). The composition comprises about 15(v/v)% of acetic acid of the hydrolyzing solution, about 30(v/v)% acetonitrile of the hydrolyzing solution, and 10mM of TCEP ((Tris(2-carboxyethyl)phosphine).

The composition may not include the water miscible organic solvent and a reducing agent. The composition may not include the water miscible organic solvent or the reducing agent.

The composition of the present invention can include a detergent which is OBG (octyl-beta-glucopyranoside) or SDS(Sodium dodecyl sulfate).

In addition, the present invention provides a method for hydrolyzing a polypeptide at an aspartic acid amino acid residue comprising contacting the polypeptide with a hydrolyzing solution of the present invention to obtain polypeptide fragments having aspartic acid residue at the N- or C- terminus and optionally determining amino acid sequence of resultant polypeptide fragments.

In an embodiment of the present invention, a method of determining amino acid sequence of a polypeptide comprises:

(i) hydrolyzing the polypeptide with the composition to obtain polypeptide fragments having aspartic acid residue at the N- or C-terminal ends of the fragments;

(ii) determining sequence of resultant polypeptide fragments; and

(iii) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the full sequence of the polypeptide.

5 The sequence of polypeptide fragments is determined through mass spectrometry. Water is labeled with deuterium, or tritium, or ^{17}O , or ^{18}O labeled water. The hydrolysis of the polypeptide is carried out in a reaction temperature in the range of about 75 to 150°C. The container material for hydrolysis reaction is made of plastic which is made of polyethylene, polypropylene, high density
10 polyethylene, or low density polyethylene. The reaction heat is created by micro wave, or ultra sonic wave.

 The sequence of the polypeptide can be determined by a database search using a modified cleavage rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-
15 terminal ends. The database search is carried out with PCA database menu which has a cleavage rule and modification rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

 In an embodiment of the present invention, a method of determining amino
20 acid sequence of a polypeptide comprises:

(i) hydrolyzing the polypeptide with protease(s) to obtain polypeptide fragments;

(ii) hydrolyzing the composition obtained in step (1) with the composition to obtain polypeptide fragments having aspartic acid residue at the N- or C-
25 terminal ends of the fragments;

(iii) determining sequence of resultant polypeptide fragments;

(iv) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the full sequence of the polypeptide.

30 The sequence of the polypeptide can be determined by a database search using a modified cleavage rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends. The database search is carried out with PCA database menu which

has a cleavage rule and modification rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

In an embodiment of the present invention, a kit for hydrolyzing polypeptide comprises (i) a container containing an acid solution and water; (ii) a container containing a water miscible organic solvent and a reducing agent. The reducing agent can be omitted.

The present invention is further explained in more detail with reference to the following examples. These examples, however, should not be interpreted as limiting the scope of the present invention in any manner.

1. Protein cleavage at Aspartic acid (PCA) using chemical reagents

This invention (PCA) provides a superior method for the proteomic analysis of proteins than any other digestion method. As shown in Figure 1, the proteins in-solution or in-gel can efficiently be cleaved to generate peptides using PCA solution, and the mass pattern and the sequence of amino acid of the resulting peptides can be analyzed by mass spectrometer. Proteins dissolved in solvents or in gel band are incubated at the temperature higher than 95°C (Incubation at 99.9°C is preferable) for more than 10 minutes in the presence of PCA solution.

1-1. The composition of PCA solution

1. Acid: trifluoroacetic acid, phosphoric acid, propionic acid, HCl, o-iodobenzoic acid, glacial acetic acid, Formic acid,

(Acid such as acetic acid that has a buffering capacity near pH 2 but do not bring about any unexpected modifications during reaction is preferable.)

2. Water miscible organic solvent: Acetonitrile, DMF (Dimethyl formamide), DMSO (Dimethylsulfoxide), THF (Tetrahydrofurane), any kinds of alcohol such as methanol and ethanol,

(The amount of acetonitrile is preferably 30(v/v)%.)

3. Reducing agent for disulfide bond: TCEP (Tris(2-carboxyethyl)phosphine), DTT (Dithiothreitol),

(Reducing agent such TCEP that can work at acidic pH range is preferable.)

In terms of yield, efficiency and specificity, reaction in the presence of 15% acetic acid (2.62mM), 30(v/v)% acetonitrile and TCEP (10mM) is optimal for protein cleavage at aspartic acid and reduction of disulfide bond in proteins.

5

The sample was then cooled to room temperature and the reaction solution was dried in the same reaction tube without transferring to a new tube for speed-vac dry. The dried peptide extract was diluted with appropriate volume of 0.1(v/v)% TFA. Desalting process for removing TCEP oxide (TCEPO) and other salts which can interfere with mass analysis can be done by the passage through μ -C18 ZipTips. The resulting peptides are analyzed by mass spectrometer.

10

1-2. Sample Test with BSA and Ubiquitin

15

1-2-1. BSA

20

25

30

Bovine Serum Albumin (BSA, Calbiochem Catalog No. 126609) is used to verify the usefulness of PCA method for sample preparation. The reaction was carried out at 99.9°C for 2hrs in the presence of PCA solution (15(v/v)% of acetic acid (pH 2.0), 30(v/v)% of acetonitrile, and 10mM TCEP). Any optional modifications except for pyro-glu E (N-term) and pyro-glu Q (N-term) were not observed in the mass spectra obtained. These optional modifications do not interfere with the identification of proteins by MASCOT™. During the course of reaction, 57 of peaks with satisfying signal-to-noise ratio are observed, which can be assigned to expected products generated by the cleavage of BSA at aspartyl residues. Of the BSA sequence, 43% was recovered directly by the accommodation of PCA method (Figure 2). More cysteine-containing peptides are retrieved with the inclusion of TCEP (Tris(2-carboxyethyl)phosphine) as a reducing agent instead of DTT (Dithiothreitol) in the PCA solution. Lots of DTT (dithiothreitol) are needed for disulfide bonds cleavage between peptides in the acidic hydrolysis condition, because the optimal pH for reducing reaction of DTT (dithiothreitol) is slightly basic. TCEP (Tris(2-carboxyethyl)phosphine) is a choice of reducing agent which can work at the range of pH optimal for acid hydrolysis of proteins.

1-2-2. Ubiquitin

We used ubiquitin (Sigma Catalog No. U6253) to test the feasibility of the PCA method for small size proteins. The reaction was carried out at 99.9°C for 2hrs in the presence 15(v/v)% of acetic acid having pH 2.0 and 30(v/v)% of acetonitrile. For ubiquitin, TCEP is not included in the reaction mixture due to the lack of disulfide bond in ubiquitin. Cleavage of ubiquitin with the PCA method resulted in perfect coverage, 100% (Figure 3). Tryptic digestion retrieved about 82% of sequence of ubiquitin, but another variant of ubiquitin was picked up by MASCOTTM.

2. PCA in combination with enzymatic digestion

We have explored the usefulness of PCA method in combination with enzymatic digestion. Trypsin is a common choice of enzyme widely used in the digestion of proteins. Cleavage at aspartyl residue (D) can be useful in dealing with the samples for which enzymes have poor accessibility for proteolytic attack and amino acids for cleavage are scarce or lacking.

2-1. Sample Test with BSA for multiple digestion

BSA (bovine serum albumin, 10 pmole) was dissolved in 50mM ammoniumbicarbonate solution to make a concentration of 1-5 μ M. Aliquots from protein solutions were thermally denatured by incubating at 90°C for 20min. Following incubation, the proteins were transferred to an ice-water bath to quench the denaturation process. Thermally denatured protein samples were enzymatically digested with sequencing-grade modified trypsin at 37°C for 12hrs. Digested sample in 50mM ammoniumbicarbonate was transferred into reaction tube and then heated at 99.9°C for 30min to remove ammoniumbicarbonate by evaporation. Following heating, the dried samples dissolved in PCA solution (15(v/v)% of acetic acid (pH 2.0), 30(v/v)% of acetonitrile, and TCEP (10mM)) were heated at 99.9°C for 2hrs. The sample was then cooled to room temperature and the reaction solution was dried at 99.9 °C in the same reaction tube. The dried peptide extract was diluted with 5 μ L 0.1(v/v)% TFA and desalted with μ -C18 ZipTips. The peptides bound to the ZipTip were eluted out sequentially in 5 μ L

each of 20(v/v)%, 50(v/v)% and 80(v/v)% ACN in 0.1(v/v)% HAc (acetic acid) solution or just 50(v/v)% ACN/0.1(v/v)% HAc. The sequentially peptide solution was dried to reduce its volume to about 5 μ L and a matrix solution was added for MALDI-TOF analysis. The purpose of using ZipTip is to remove salts from the sample prior to MS detection. This approach was applied to characterization of proteins such as BSA (bovine serum albumin) resulting in improved identification with high sequence coverage (87%) comparing with 40% and 32% obtained by PCA method and tryptic digestion, respectively as shown in Figure 4.

2-2. Sequential digestion of membrane proteins with chemical and enzymatic method

Proteomic analysis of membrane proteins has been indispensable and challenging subject for understanding of diverse signaling networks and for discovering targets for a given disease. Most of researchers have relied on classical proteomic method consisted of two-dimensional electrophoresis for separating proteins prior to enzymatic digestion and mass spectrometric analysis, while only few studies have been reported by accommodation of gel-free LC-MS/MS technologies. However, even gel-free LC-MS/MS technologies depend on detergents or chaotropic reagents to isolate the integral proteins from membrane fraction. Hereby, we present novel gel-free and detergent-free shotgun proteomic method for analyzing membrane proteins. By the application of our proprietary technology (CCPA method, Chemical Cleavage of Proteins at Aspartic acid), delipidated aggregate of mouse lipid raft proteins were successfully digested into peptides.

2-2-1. Lipid raft preparation from mouse brain

Lipid raft was selected for proving the usefulness of chemical digestion, because lots of proteomic studies have been performed to elucidate the composition of proteins in lipid raft. Lipid raft was prepared as described in Proteomics 2004, 4, 3536–3548. Mouse brains were homogenized 20 times with a tight Dounce homogenizer (Kontes, Vineland, NJ, USA) in the lysis buffer (1(v/v)% Triton X-100, 25 mM HEPES, pH 6.5, 150 mM NaCl, 1 mM EDTA, 1 mM PMSF, and protease cocktail (Roche Molecular Biochemicals, Indianapolis, IN, USA)), and

incubated at 40°C for 20 min. The extract was mixed with 2.5 M sucrose, transferred to an SW41 centrifuge tube, and overlaid with 30(w/v)% sucrose solution and 5(w/v)% sucrose solution containing 25 mM HEPES, pH 6.5, and 150 mM NaCl. The discontinuous sucrose gradients were centrifuged for 18 h at 40°C in an SW41 rotor at 39 000 rpm. The gradient was fractionated into 12 fractions from the bottom to the top. The lipid raft fractions were washed with washing buffer (25 mM HEPES, pH 7.4, 150 mM NaCl) by ultracentrifugation (20 000 rpm, 30 min, 47°C), and suspended with 50 mM sodium bicarbonate. By the treatment with acetone and successive washing steps, the precipitate of membrane was obtained. The resulting protein precipitated was dried.

2-2-2. PCA reaction of de-lipidated membrane proteins

The dried material dissolved in PCA solution (15(v/v)% acetic acid (pH 2.0), 30(v/v)% acetonitrile, and TCEP (10mM)) were incubated at 99°C for 4hrs by using PCR machine. The digestion of membrane proteins into peptides were confirmed by the visualization of proteins or peptides following Tris-Tricine gel electrophoresis as shown in Figure 5.

2-2-3. Separation of PCA-digested membrane proteins

Chemically digested membrane proteins were separated by using reverse-phase column chromatography (Chromolith-C18, Merck) as shown in Figure 6. The fractions from column were pooled to bring about 9 fractions. Nine fractions were dried separately.

2-2-4. Tryptic digestion of PCA-digested membrane proteins

The dried 9 fractions were incubated with 1/50 amount of Trypsin in the presence of 50mM Ammonium Bicarbonate at 37°C for 12 hrs. The resulting peptides were analyzed by LC-MS/MS.

3. PCA-DMT (Differential Mass Tagging)

For the quantitative analysis of the samples, two samples are labeled with a chemically identical but isotopically different tagging by using PCA solution containing ^{16}O -water or ^{18}O -water. During the course of reaction, the exchange of

one or two atoms of oxygen can take place, thereby bring about the difference of molecular mass of 2 or 4 Da which can be discriminated in MS analysis. In PCA-DMT method, relative quantification for protein can be achieved by comparing the signal intensities of monoisotopic mass obtained by the chemical digestion of protein in the presence of ^{16}O -water with those obtained by the inclusion of ^{18}O -water in PCA solution.

3-1. Sample Test with Ubiquitin

Proteins dissolved in water or dried gel band are heated at 99.9°C for 2hrs in the presence of PCA solution (15(v/v)% acetic acid (pH 2.0), 30(v/v)% acetonitrile, and 10mM TCEP). For differential mass tagging, ^{16}O -water was replaced by ^{18}O -water. The sample was then cooled to room temperature and the reaction solution was dried at 99.9°C in reaction tube and dissolved with $5\mu\text{L}$ 0.1(v/v)% TFA, and salt was removed by desalting process using $\mu\text{-C18}$ ZipTips. The peptides bound to the ZipTip were eluted out sequentially in $5\mu\text{L}$ each of 20(v/v)%, 50(v/v)% and 80(v/v)% ACN in 0.1 (v/v) HAc(acetic acid) solution or just 50(v/v)% ACN/0.1(v/v)% HAc. The sequentially peptide solution was dried to reduce its volume to about $5\mu\text{L}$ and a matrix solution was added for MALDI-TOF analysis. One of the examples is shown in Figure 7. Accurate mass measurement of the monoisotopic peak (MH^+ (obs) 1528.83, MH^+ (calc) 1528.80) allowed this peptide to be assigned as QQRLIFAGKQLED (Pyro-glu(N-term)) and the shift of molecular mass by 2 or 4 was also observed indicating incorporation of one or two ^{18}O atoms in the carboxylate group at the C-termini. Mass spectra was obtained in a mass accuracy of 10 ppm using the MALDI-TOF (ABI 4700 analyzerTM).

When m/z value (value of x-axis in mass spectra) of monoisotopic mass is described as M_0 , the series of isotopic mass is herein presented as M_1 , M_2 , M_3 , M_4 in ascending order according to the difference in molecular mass. The signal intensity (value of y-axis in mass spectra) of a given m/z is described as $S_{16}(M_0)$, $S_{16}(M_1)$, $S_{16}(M_2)$, $S_{16}(M_3)$, and $S_{16}(M_4)$. Number 16 in this description indicates the atomic mass of oxygen atom of water molecule used in the reaction of chemical cleavage. According to this, signal intensity of monoisotopic mass obtained through the chemical cleavage in the case of using ^{18}O -water is described as

S18(M0), and that obtained by mixing the products prepared from separate chemical cleavage reaction in the presence of 16O-water or 18O-water is presented as S16+18(M0). K1 indicates the coefficient for the incorporation of one atom of oxygen in chemical cleavage reaction done in the presence of 18O-water, and K2 for the incorporation of two atoms of oxygen. By calculating the sum of K1 and K2, relative comparison of the amounts of proteins in two different samples can be achieved. In mathematical terms, K1 and K2 can be extracted from the formula as followed.

$$S16+18(Mx) = S16(Mx) + K1*S16(Mx-2) + K2*S16(Mx-4)$$

[m/z value should not be less than 0]

$$S16+18(M2) = S16(M2) + K1*S16(M0) + K2*S16(M-2) = S16(M2) + K1*S16(M0)$$

$$K1 = (S16+18(M2) - S16(M2))/S16(M0)$$

$$S16+18(M4) = S16(M4) + K1*S16(M2) + K2*S16(M0)$$

$$K2 = (S16+18(M4) - S16(M4) - K1*S16(M2))/S16(M0)$$

From the mass spectra in Figure 7, mixing ratio of the peptides obtained by the chemical cleavage reaction of ubiquitin in the presence of 16O-water and 18O-water can be calculated according to the formula described above.

$$K1 = (5963.09 - 0.49389*8251.56)/8251.56 = 0.2287$$

$$K2 = (4966.49 - (0.06472*8251.56 + 0.49389*1887.73))/8251.56 = 0.4242$$

$$K1 + K2 = 0.6529$$

The ratio of the peptides obtained by the chemical cleavage reaction of ubiquitin in the presence of 16O-water and 18O-water is calculated as 1:0.65. The difference of calculated ratio from theoretical one can be due to the usage of 18O-water containing around 5% of 16O-water, the possible contamination of 16O-water in other reagents such TCEP and sample (Ubiquitin), and incorporation of 16O-water from air during the course of reaction.

4. Sample Preparation for MALDI MS Analysis.

The dried peptide extract was suspended in 5μL 0.1(v/v)% TFA and desalted with μ-C18 ZipTips. The peptides bound to the ZipTip were eluted out

sequentially in 5 μ L each of 20(v/v)%, 50(v/v)% and 80(v/v)% acetonitrile in 0.1(v/v)% acetic acid solution or just 50(v/v)% acetonitrile/0.1(v/v)% acetic acid. The sequentially peptide solution was dried to reduce its volume to about 5 μ L and a matrix solution was added. The purpose of using ZipTip is to remove salts from the samples prior to MS detection, but it does not completely remove the n-OG, a nonionic detergent. The two-layer sample deposition method with α -CHCA as matrix was used in the MALDI-MS analysis. The first layer was prepared as a 20mg/mL α -CHCA solution in 20(v/v)% methanol/acetone α -CHCA and the second layer with a saturating solution of matrix in 30(v/v)% methanol/water. The second layer was added to the Zip-Tipped peptide mixture for producing the ratio of matrix to analyte to 4:1 and the mixture vortexed. After 0.5 μ L of the first layer was deposited on the sample probe and air-dried, 0.5 μ L of the second layer was deposited on top of the first layer, allowed to air-dry and washed twice with 1 μ L water.

5. MALDI-Mass Spectrometry

A MALDI-TOF mass spectrometer with 4700 Proteomics AnalyzerTM (Applied Biosystems) was used to acquire the mass spectra. The matrix solution was a 10mg of α -CHCA in 50(v/v)% acetonitrile in water. Aliquots of 0.5 μ L of the peptide mixture and 0.5 μ L of the matrix solution were mixed on the sample plate and air-dried prior to analysis.

6. MALDI Tandem Mass Spectrometry (MS/MS)

All MS/MS data from the TOF/TOF (4700 Proteomics AnalyzerTM) was acquired using the default 1kv MS/MS method following manufacturer's instruction. MS/MS data acquisition from the plates (LC-MALDI plates) on which the LC eluent had been spotted by the ProbotTM was performed in a four step process. First, MS spectra were recorded from each of the six calibration spots, and the default calibration parameters of the instrument and the appropriate model of plate model were updated. Second, MS spectra were recorded for all 144 sample spots on that plate. Each spectrum was generated by accumulating the data from 750 laser shots using the newly updated default calibration settings. Third, the 144 MS spectra were analyzed using the Peak Picker software supplied with the

instrument. Spectral peaks that met the threshold criteria and were not on the exclusion list were included in the acquisition list for the MS/MS portion of the experiment. The threshold criteria were set as follows: mass range: 650 to 4000Da; minimum cluster area: 500; minimum signal-to-noise (S/N): 10; Peaks/spot: 30; maximum precursor gap: 200ppm; maximum fraction gap: 4. A mass filter excluding matrix cluster ions was applied. An XML file was generated that contains the list of the precursor masses selected for MS/MS and their corresponding spot numbers. Lastly, the list was imported into the 4700 Explorer software batch editor, and MS/MS spectra were recorded using air as the collision gas with 1kV collision energy setting. During MS/MS data acquisition, a method with a stop condition was used. In this method, a minimum of 750 shots (6 sub-spectra accumulated from 125 laser shots each) and a maximum of 2000 shots were allowed for each spectrum. The accumulation of additional laser shots was halted whenever at least 10 ions with a S/N of at least 10 were present on the accumulated MS/MS spectrum, in the region from m/z 400 to 90% of the precursor mass.

7. LC-MS/MS analysis of double-digested lipid raft proteome

PCA/Trypsin-cleaved peptides were eluted from the nano LC system, Agilent 1100 (Agilent), and ions were sprayed directly into the orifice of a QSTAR-XL quadrupole time-of-flight (TOF) hybrid MS (PE-Sciex, Thornhill, Ontario, Canada). The results of LC-MS/MS experiments for the hydrolysate obtained from sequential digestion of mouse brain lipid raft by PCA and Trypsin were as shown in Figure 8. Proteins were identified by LC/MS/MS by information-dependent acquisition of fragmentation spectra for multiply charged peptides that were then searched against the Human International Protein Index database (<ftp://ftp.ebi.ac.uk/pub/databases/IPI/current/MOUSE>) by using MASCOT (Matrix Science, London). The following search parameters were used in all MASCOT searches: maximum of two missed CCPA-trypsin cleavage, cysteine carbamidomethylation, methionine oxidation, and a maximum 0.2-Da error tolerance in the MS and 0.1-Da in the MS/MS data. Significant matches with the highest MOWSE scores were considered potential identification. All other hits were manually verified by using accepted rules for peptide fragmentation in a quadrupole-TOF hybrid MS. The selected subset of mouse brain lipid raft proteins

was listed in Figure 11. LC-MS/MS analysis of resulting peptides proteins brought about the identification of lots of bona fide membrane proteins such as G-proteins, adhesion molecules, channels/transporter, signaling proteins (CAM-Kinase and phosphodiesterase), and flotillin (a kind of raft marker protein).

5

8. Mass Spectra Interpretation and Database Searching.

The rule by which peptides are generated from protein was modified following instructions in the user's manual of MASCOTTM so that PCA cleavage rule at both Asp-X and X-Asp was reflected in the MASCOTTM program. The modifications of peptides which is observed in tryptic digestion, such as methionine oxidation, Pyro-glu E(N-term), Pyro-glu Q(N-term) were considered for database searching. The algorithm developed in this invention can be realized for any other programs developed for the analysis of mass spectra.

10

9. Kit for PCA Proteomics

Kit for PCA method can be supplied as a set of solution (solution A and solution B) and container designed for optimal heat transfer and for minimizing sample loss during chemical reaction (thin-wall tube for PCR reaction) is highly recommended. The reaction should be carried out in the heating apparatus designed for minimizing the loss of vapor pressure by heating the lid as well as bath simultaneously in the same temperature such as PCR machine for better performance.

15

20

Total reaction mixture (150 μ L); Solution A (45 μ L) + solution B (105 μ L)

A. Solution A; 10mM TCEP (Tris(2-carboxyethyl)phosphine) in Acetonitrile

B. Solution B; 21.4(v/v)% glacial acetic acid (acid) in Water

Mixture of peptide can be included as internal mass standard. For differential mass tagging, 21.4(v/v)% acetic acid in ^{18}O -Water can be used.

25

The method described in this invention can be modified in numerous ways by a specialist with a full understanding of the fundamental principles. Therefore, the invention is not restricted to the aforementioned examples. Other types of mass

instrument can be used for identification and quantification of proteins and peptide with specific cleavage rule and modification rule on the basis of database search algorithm.

5 The cited references are as follows:

1. Aiqun Li et al. Anal Chem. 2001, 73, 5395-5402, Chemical Cleavage at Aspartyl Residues for Protein Identification

2. Peter Roepstorff et al. Anal Chem. 1999, 71, 919-927, Use of Vapor-Phase Acid Hydrolysis for Mass Spectrometric Peptide Mapping and Protein
10 Identification

3. Zee-Yong Park et al. Anal Chem. 2000, 72, 2667-2670, Thermal Denaturation: A useful Technique in Peptide Mass Mapping

4. Steven L. Cohen et al. Anal. Chem. 1996, 68, 31-37, Influence of Matrix Solution Conditions on the MALDI-MS Analysis of Peptides and Proteins

5. Bart A. van Montfort et al. J. Mass Spectrom. 2002, 37, 322-330, Improved
15 in-gel approaches to generate peptide maps of integral membrane proteins with matrix-assisted laser desorption/ionization time-of flight mass spectrometry

6. J.otte et al, J. Agric. Food Chem. 2000, 48, 2443-2447, Identification of Peptides in Aggregates Formed during Hydrolysis of b-Lactoglobulin B with a Glu
20 and Asp Specific Microbial Protease

7. Adrienne Kishiyama et al, Anal. Chem. 2000, 72, 5431-5436, Cleavage and identification of Proteins: A Modified Aspartyl-Prolyl Cleavage

8. Cornelia Koy et al. Proteomics 2003, 3, 851-858, Matrix-assisted laser desorption/ionization-quadrupole ion trap-time of flight mass spectrometry
25 sequencing resolves structures of unidentified peptides obtained by in-gel tryptic digestion of haptoglobin derivatives from human plasma proteomes.

9. Melanie Lin et al. Rapid Commun. Mass Spectrum. 2003, 17, 1809-1814, Intact protein analysis by matrix-assisted laser desorption/ionization tandem time-of-flight mass spectrometry.

10. Xudong Yao et al. Anal. Chem. 2001, 73, 2836-2842, Proteolytic ¹⁸O
30 Labeling for Comparative Proteomics: Model Studies with Two Serotypes of Adenovirus.

11. Marcus Bantscheff et al, Rapid Commun. Mass Spectrum. 2004, 18, 869-

876, Femtomol sensitivity post-digest ^{18}O labeling for relative quantification of differential protein complex composition.

12. Kenneth L. Johnson et al. J Am Soc Mass Spectrom 2004, 15, 437-445, A Method for Calculating $^{16}\text{O}/^{18}\text{O}$ Peptide Ion Ratios for the Relative Quantification of Proteomes.

17. Y.Karen Wang et al, Anal. Chem. 2001, 73, 3742-3750, Inverse ^{18}O Labeling Mass Spectrometry for the Rapid Identification of Marker/Target Proteins.

18. Schnolze, M.; Jedrzejewski, P.; Lehmann, W.D. Electrophoresis 1996, 17, 945-953,

19. Methods in ENZYMOLOGY Vol 91, Enzyme Structure Part 1, 324-332, Cleavage at Aspartic Acid

20. Methods in ENZYMOLOGY Vol 4, Enzyme Structure, 255-263, Cleavage at Aspartic Acid

21. Gargi Choudhary et al. Journal of Proteome Research 2003, 2, 59-67

WHAT IS CLAIMED IS:

1. A polypeptide hydrolyzing composition comprising water and at least an acid component selected from the group consisting of trifluoroacetic acid, phosphoric acid, propionic acid, HCl, o-iodobenzoic acid, glacial acetic acid, and an acid having buffering capacity near pH 2.
2. The composition according to claim 1, wherein the composition further comprises at least one selected from the group consisting of a water miscible organic solvent and a reducing agent.
3. The composition according to claim 1, wherein the acid component is a mixture of trifluoroacetic acid, phosphoric acid, propionic acid, HCl, and o-iodobenzoic acid.
4. The composition according to claim 1, wherein the pH of the hydrolyzing composition at time of reaction is in the range of 1.5 to 2.5
5. The composition according to claim 1, wherein the hydrolyzing composition comprises at least 2 to 30(v/v)% of glacial acetic acid.
6. The composition according to claim 1, wherein the hydrolyzing composition comprises 15(v/v)% of glacial acetic acid having pH 2.0.
7. The composition according to claim 2, wherein the water miscible organic solvent is Acetonitrile, DMF (Dimethyl formamide), DMSO (Dimethylsulfoxide), THF (Tetrahydrofuran), or an alcohol.
8. The composition according to claim 7, wherein the alcohol is methanol or ethanol.
9. The composition according to claim 2, wherein the water miscible organic solvent is Acetonitrile.

10. The composition according to claim 2, wherein the water miscible organic solvent is at least 5-70(v/v)% of Acetonitrile.

5 11. The composition according to claim 2, wherein the water miscible organic solvent is 30(v/v)% of Acetonitrile.

12. The composition according to claim 2, wherein the reducing agent is TCEP (Tris(2-carboxyethyl)phosphine), DTT (Dithiothreitol), or beta-
10 Mercaptoethanol.

13. The composition according to claim 2, wherein the reducing agent is a phosphine compound which can work at an acidic pH range (1.5-2.5) such as TCEP (Tris(2-carboxyethyl)phosphine).

15 14. The composition according to claim 2, wherein the reducing agent is at least 1mM-1M TCEP (Tris(2-carboxyethyl)phosphine) or DTT (Dithiothreitol).

20 15. The composition according to claim 2, wherein the reducing agent is at least 10 mM TCEP (Tris(2-carboxyethyl)phosphine) or DTT (Dithiothreitol).

16. The composition according to claim 2, wherein the composition comprises acetic acid at an amount of about 2-30(v/v)% of the hydrolyzing composition, acetonitrile at an amount of about 5-70(v/v)% of the hydrolyzing
25 composition, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine).

17. The composition according to claim 2, wherein the composition comprises acetic acid at an amount of about 15(v/v)% of the hydrolyzing composition, acetonitrile at an amount of about 5-70(v/v)% of the hydrolyzing
30 composition, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine).

18. The composition according to claim 2, wherein the composition comprises acetic acid at an amount of about 15(v/v)% of the hydrolyzing

composition, acetonitrile at an amount of about 30(v/v)% of the hydrolyzing composition, and about 1mM-1M of TCEP ((Tris(2-carboxyethyl)phosphine).

19. The composition according to claim 2, wherein the composition
5 comprises acetic acid at an amount of about 15(v/v)% of the hydrolyzing composition, acetonitrile at an amount of about 30(v/v)% of the hydrolyzing composition, and 10mM of TCEP ((Tris(2-carboxyethyl)phosphine).

20. The composition according to claims 1 or 2, wherein the
10 composition further comprises a detergent.

21. The composition according to claim 20, wherein the detergent is OBG (octyl-beta-glucopyranoside) or SDS (Sodium dodecyl sulfate).

22. A method for hydrolyzing a polypeptide at an aspartic acid amino
15 acid residue comprising contacting the polypeptide with a hydrolyzing composition according to any one of claims 1 to 19 to obtain polypeptide fragments having aspartic acid residues at the N- or C- terminus and optionally determining the amino acid sequence of resultant polypeptide fragments.

23. A method for hydrolyzing a polypeptide at an aspartic acid amino
20 acid residue comprising contacting the polypeptide with a hydrolyzing composition according to claim 20 to obtain polypeptide fragments having aspartic acid residues at the N- or C- terminus and optionally determining the amino acid
25 sequence of resultant polypeptide fragments.

24. A method of determining the amino acid sequence of a polypeptide comprising:

(i) hydrolyzing the polypeptide with the composition according to
30 any one of claims 1 to 19 to obtain polypeptide fragments having aspartic acid residues at the N- or C-terminal ends of the fragments;

(ii) determining the sequence of resultant polypeptide fragments;

and

(iii) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the full sequence of the polypeptide.

5 25. A method of determining the amino acid sequence of a polypeptide comprising:

(i) hydrolyzing the polypeptide with the composition according to claim 20 to obtain polypeptide fragments having aspartic acid residues at the N- or C-terminal ends of the fragments;

10 (ii) determining the sequence of resultant polypeptide fragments; and

(iii) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the full sequence of the polypeptide.

15 26. The method according to claims 22-25, wherein the sequence of polypeptide fragments is determined through mass spectrometry.

20 27. The method according to claim 26, wherein the water is labeled with deuterium, or tritium, or ^{17}O or ^{18}O labeled water.

25 28. The method according to any one of claims 22-25, wherein the hydrolysis of the polypeptide is carried out by heating to about 75 to 150°C reaction temperature.

29. The method according to claim 28, wherein the reaction heat is created by micro wave or ultrasonic wave.

30 30. The method according to claim 28, wherein hydrolysis of the polypeptide is carried out under a reaction temperature ranging from about 95 to 105°C in a PCR(Polymerase Chain Reaction) machine.

31. The method according to claim 30, wherein hydrolysis of the

polypeptide is carried out by heating bath including the hydolyzing composition and the polypeptide, and heating lid of the PCR machine at about 95 to 105°C reaction temperature.

5 32. The method according to any one of claims 25-29, wherein the container material for hydrolysis reaction is made of plastic.

 33. The method according to claim 32, wherein the plastic is made of polyethylene, polypropylene, high density polyethylene, or low density
10 polyethylene.

 34. The method according to any one of claims 25-29, wherein the method comprises determining the sequence of the polypeptide by a database search using a modified cleavage rule incorporating polypeptide fragments having
15 aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

 35. The method according to any one of claims 25-29, wherein the database search is carried out with a PCA database menu which has a cleavage rule
20 and modification rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

 36. A method of determining amino acid sequence of a polypeptide comprising:

25 (i) hydrolyzing the polypeptide with protease(s) to obtain polypeptide fragments;

 (ii) hydrolyzing the composition obtained in step (1) with the composition according to any one of claims 1 to 19 to obtain polypeptide fragments having aspartic acid residues at the N- or C-terminal ends of the
30 fragments;

 (iii) determining the sequence of resultant polypeptide fragments;

 (iv) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the

full sequence of the polypeptide.

37. A method of determining amino acid sequence of a polypeptide comprising:

- 5 (i) hydrolyzing the polypeptide with the composition according to claim 20 to obtain polypeptide fragments having aspartic acid residues at the N- or C-terminal ends of the fragments;
- (ii) hydrolyzing the composition obtained in step (1) with protease(s) to obtain polypeptide fragments;
- 10 (iii) determining the sequence of resultant polypeptide fragments;
- (iv) determining the sequence of the polypeptide by matching and connecting the sequences of the polypeptide fragments so as to obtain the full sequence of the polypeptide.

15 38. The method according to any one of claims 36 and 37, comprising determining the sequence of the polypeptide by a database search using a modified cleavage rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

20 39. The method according to claim 38, wherein the database search is carried out with a PCA database menu which has a cleavage rule and modification rule incorporating polypeptide fragments having aspartic acid residues at either the N- or C-terminal ends or both the N- and C-terminal ends.

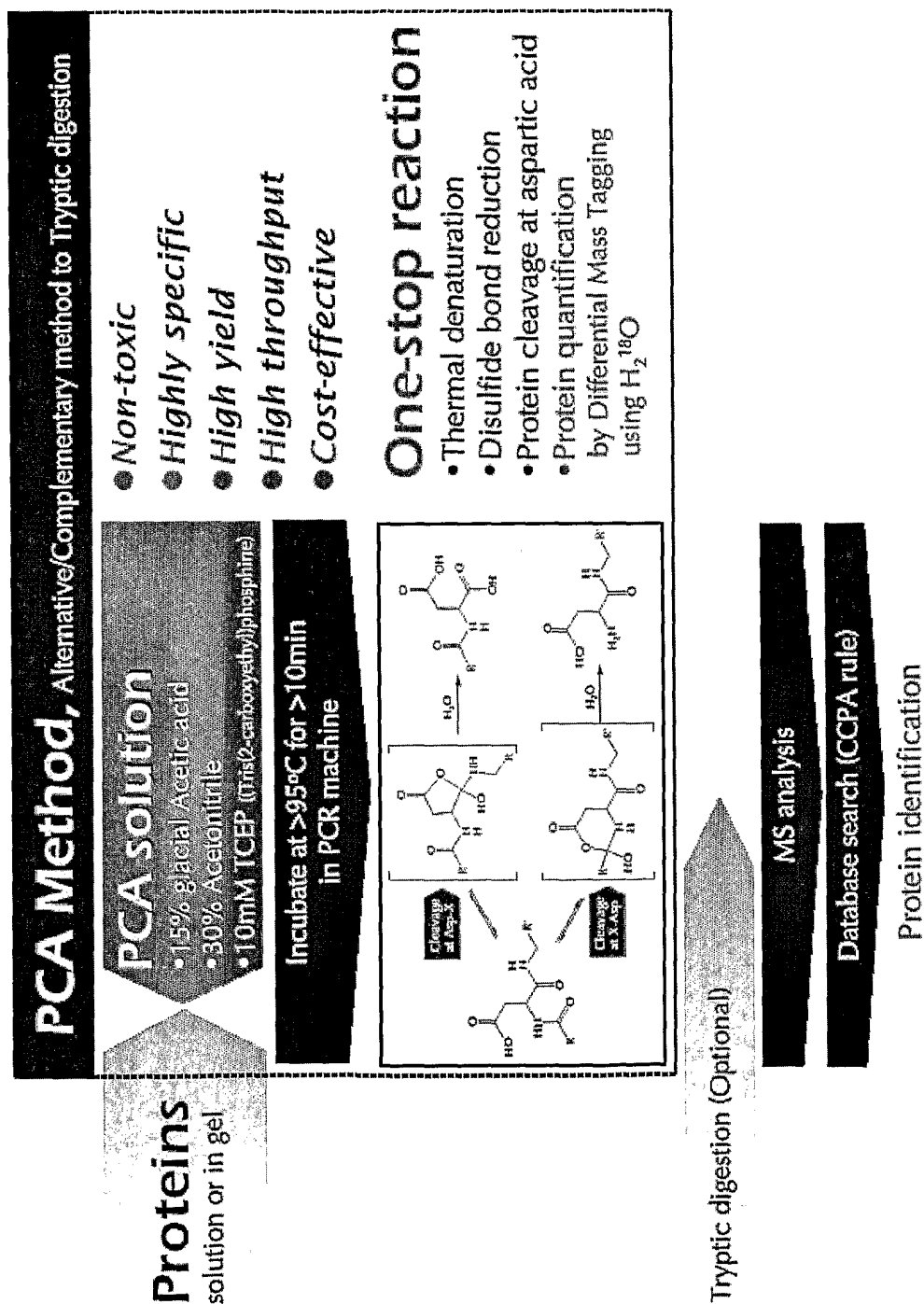
25 40. A kit for hydrolyzing polypeptide comprising: (i) a first container containing an acid solution and water; (ii) a second container containing a water miscible organic solvent, wherein the acid solution is at least one selected from the group consisting of trifluoroacetic acid, phosphoric acid, propionic acid, HCl, o-iodobenzoic acid, glacial acetic acid, and any acid having buffering capacity near

30 pH 2.

41. The kit according to claim 40, wherein the second container further contains a reducing agent.

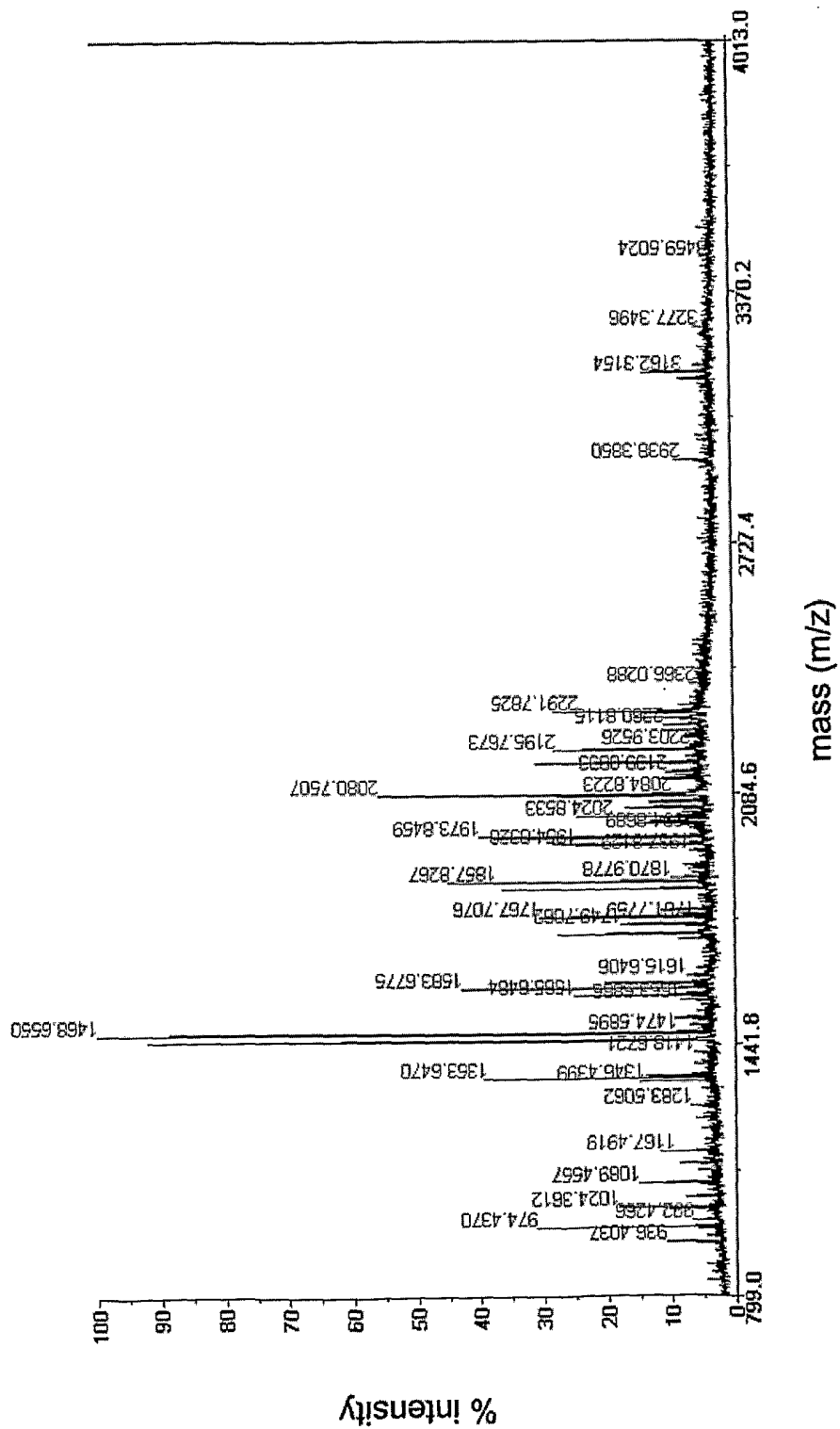
1/17

Fig. 1



2/17

Fig. 2a



3/17

Fig. 2b

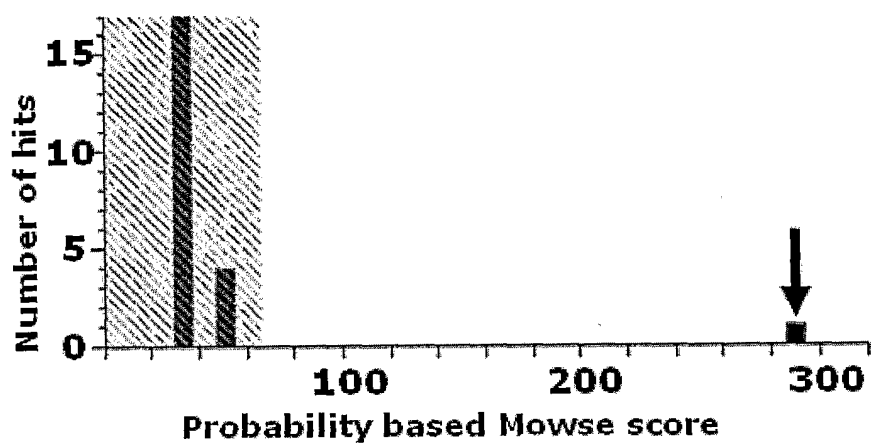
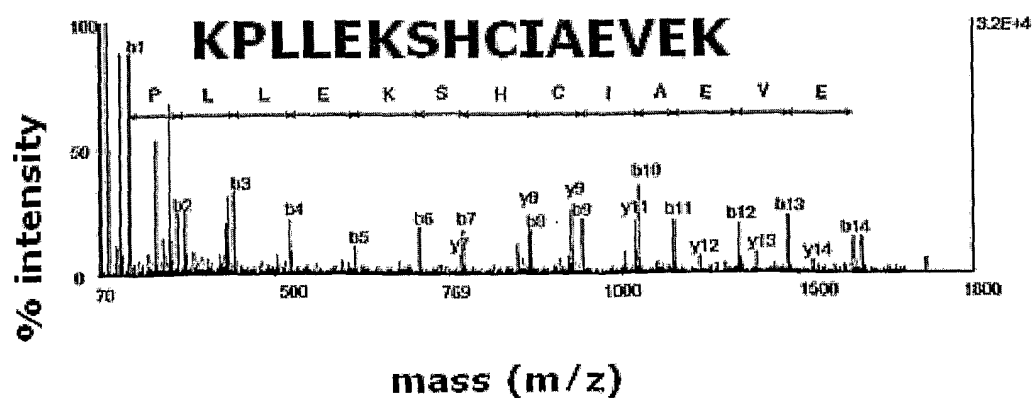
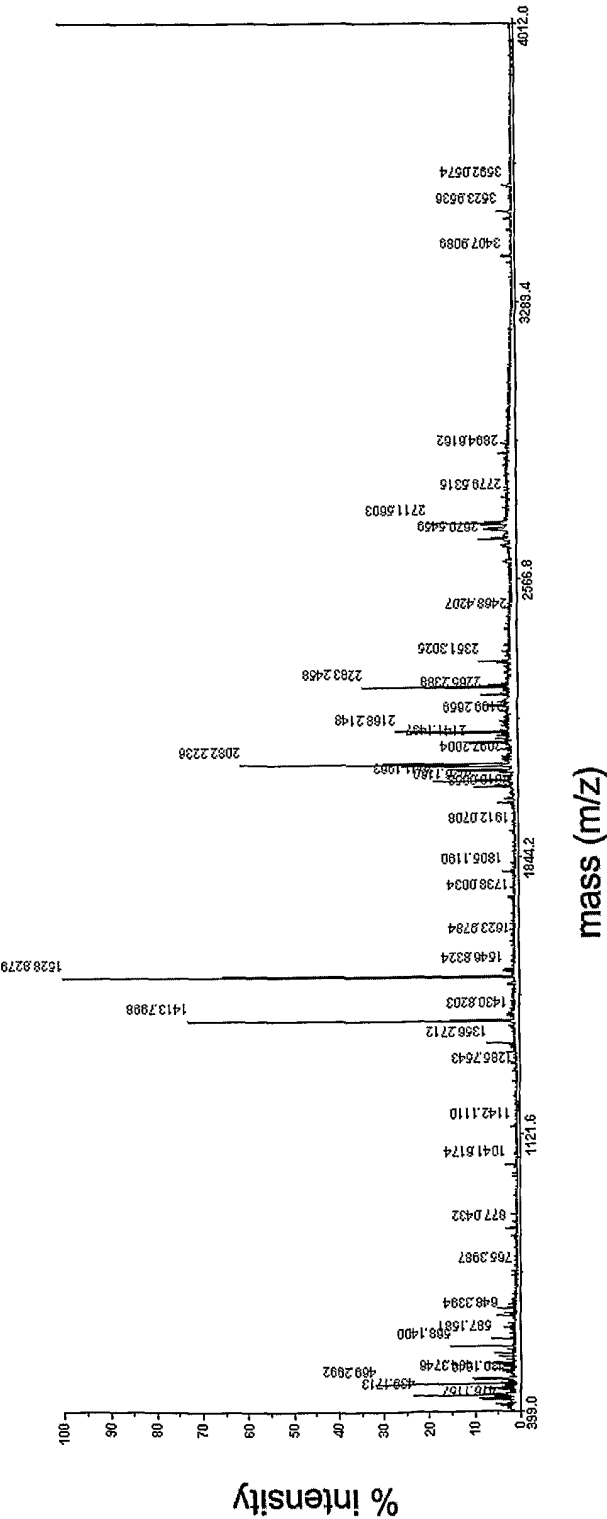


Fig. 2c



4/17

Fig. 3a



5/17

Fig. 3b

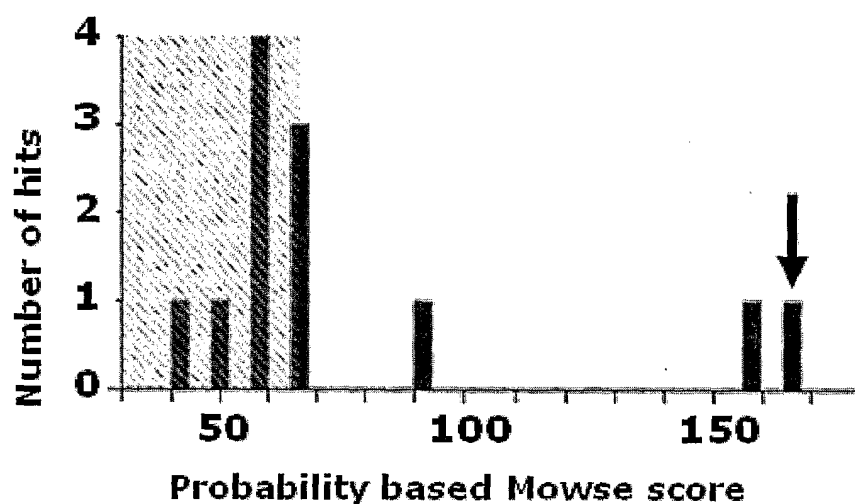


Fig. 4a

Number of mass values searched: 168

Number of mass values matched: 57

Sequence Coverage: 40%

Matched peptides shown in **Bold Red**

```

1  MKWVTFISLL LLFSSAYSRG VFRDTHKSE IAHRFKDLGE EHFKGLYLIA
51  FSQYLQQCPF DEHYKLYNEL TEFAKTCYAD ESHAGCEKSL HTLFGDELCK
101 YASLRETYGD NADCCEKQEP ERNECFLSHK DDSPDLPLK PDPNTLCDEF
151 KADEKKFWGK VLYEIARRHP VFYAPELLVY ANKYNGVFQE CCQAEDKGAC
201 LLPKIETMRE KVLTSARQR LRCASIQKFG ERALKAWSYA RLSQKFPKAE
251 FVEVTKLYTD LTKYHKECCH GDLLECADDR ADLAKYICDN QDTISSKLKE
301 CCDKPLLEKS HCIAEYEKDA IPENLPPLTA DFAEDKDYCK NYQEAKDAFL
351 GSFLYEVSRR HPEYAVSVLL RLAKEYEATL EECCAADDPH ACYSTVFDKL
401 KHLYDEPQNL IKQNCQFEK LGEYGFQNAL IVRVTRKYPQ VSTPTLVEVS
451 RSLGKVGTRC CTKPESERMP CTEDYLSLIL NRLCYLHEKT PVSEKVTGCC
501 TESLYNRRPC FSALTPDETY YPKAFDEKLF TFHADICTLP DTEKQIKKQT
551 ALVELLKHKP KATEEQLKTV MENFVAFVDK CCAADDKEAC FAYEGPKLYY
601 STQTALA
  
```

6/17

Fig. 4b

Number of mass values searched: 47

Number of mass values matched: 17

Sequence Coverage: 32%

Matched peptides shown in **Bold Red**

1 MKWYTFISLL LFFSSAYSRG VFRDTHKSE **I**AHRFKDLGE EQFKGLYLIA
 51 FSQYLQQCPF DEHVKLYNEL TEFAKTCYAD ESHAGCEKSL HTLFGDELCK
 101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPLK PDPNTLCDEF
 151 KADEKKFWGK **Y**LYEIARRHP **Y**FYAPELLYY **A**NKYNGYFQD **C**CQAEDKGAC
 201 LLPKIETMRE KYLASSARQR LRCASIQKFG **E**RA**LK**A**W**SYA **R**LSQKFPKAE
 251 FVEYTKLYTD LTKVHKECCH GDLLCADDR ADLAKYICDN QDTISSKLKE
 301 CCKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
 351 **G**SFLYEYSRR **H**PEYAYSVLL **R**LAKEYEATL **E**ECCA**K**DDPH **A**CYSTVFDKL
 401 KHLVDEPQNL **I**KQNCQFEK **L**GEYGFQNAL **I**YRYTRKYPQ **Y**STPTLYEYS
 451 RSLGKYGTRC CTKPESERMP CTEDYLSLIL **N**RLCYLHEKT PYSEKYTKCC
 501 TESLYNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
 551 **A**LYELLKHKP KATEEQLKTY MENFYAFYDK CCAADDKEAC FAVEGPKLVV
 601 STQTALA

Fig. 4c

Number of mass values searched: 185

Number of mass values matched: 87

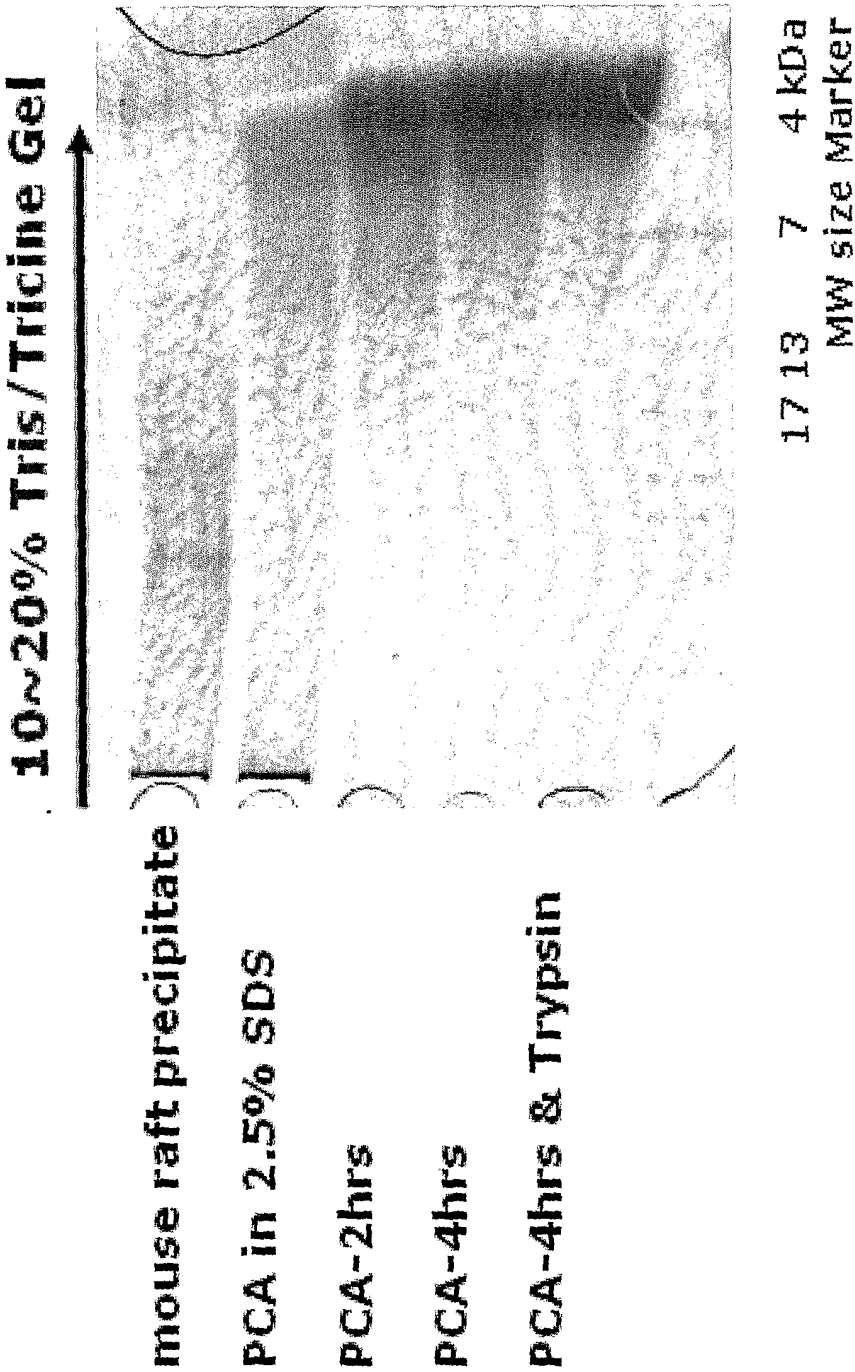
Sequence Coverage: 87%

Matched peptides shown in **Bold Red**

1 **M**K**W**YTFISLL LFFSSAYSRG VFRDTHKSE **I**AHRFKDLGE **E**HF**K**GLYLIA
 51 FSQYLQQCPF DEHVKLYNEL TEFAKTCYAD ESHAGCEKSL HTLFGDELCK
 101 VASLRETYGD MADCCEKQEP ERNECFLSHK DDSPDLPLK PDPNTLCDEF
 151 KADEKKFWGK **Y**LYEIARRHP **Y**FYAPELLYY **A**NKYNGYFQE **C**CQAEDKGAC
 201 LLPKIETMRE KYLTSSARQR LRCASIQKFG **E**RA**LK**A**W**SYA **R**LSQKFPKAE
 251 FVEYTKLYTD LTKVHKECCH GDLLCADDR ADLAKYICDN QDTISSKLKE
 301 CCKPLLEKS HCIAEVEKDA IPENLPPLTA DFAEDKDVCK NYQEAKDAFL
 351 **G**SFLYEYSRR **H**PEYAYSVLL **R**LAKEYEATL **E**ECCA**K**DDPH **A**CYSTVFDKL
 401 KHLVDEPQNL **I**KQNCQFEK **L**GEYGFQNAL **I**YRYTRKYPQ **Y**STPTLYEYS
 451 RSLGKYGTRC CTKPESERMP CTEDYLSLIL **N**RLCYLHEKT PYSEKYTKCC
 501 TESLYNRRPC FSALTPDETY VPKAFDEKLF TFHADICTLP DTEKQIKKQT
 551 **A**LYELLKHKP KATEEQLKTY MENFYAFYDK CCAADDKEAC FAVEGPKLVV
 601 STQTALA

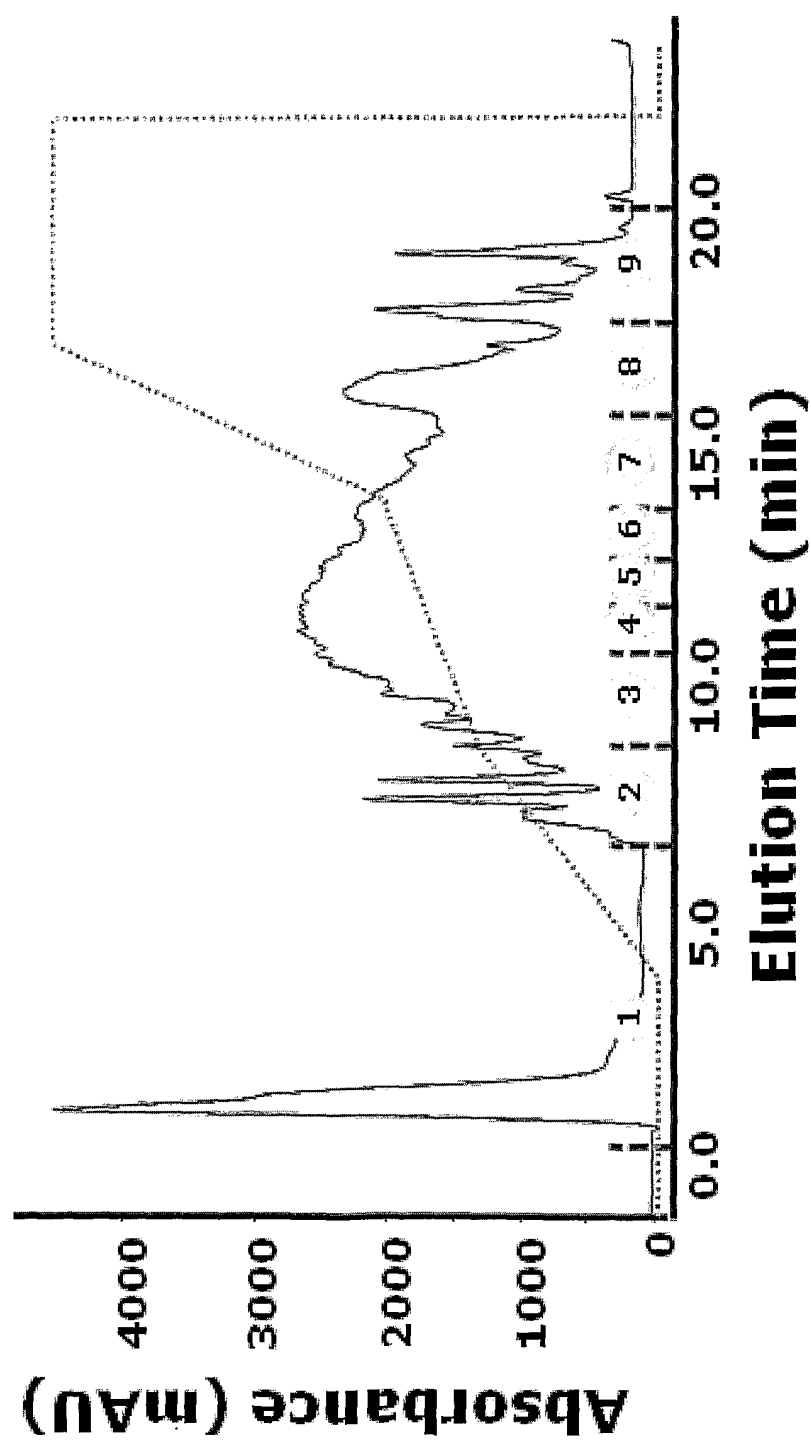
7/17

Fig. 5



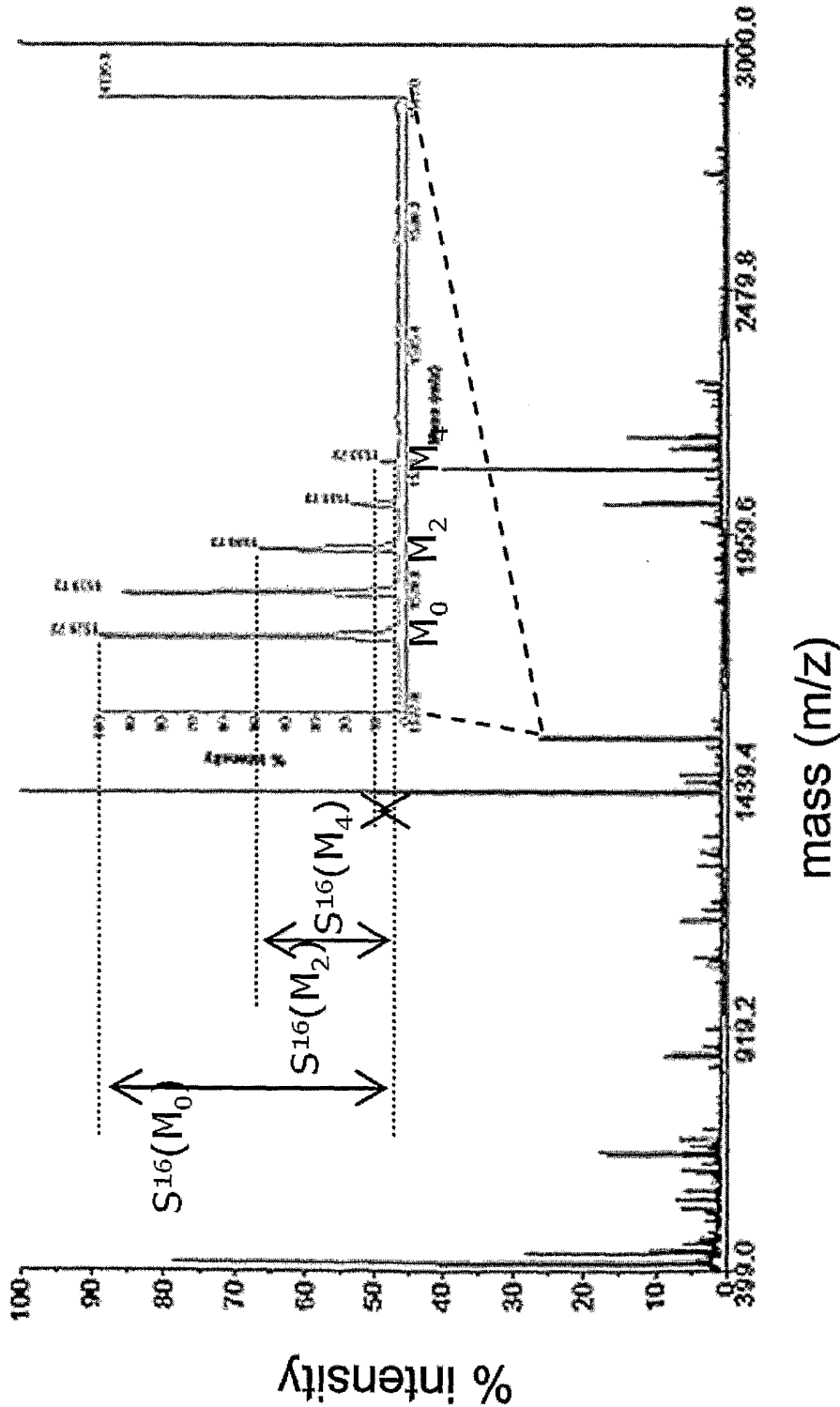
8/17

Fig. 6



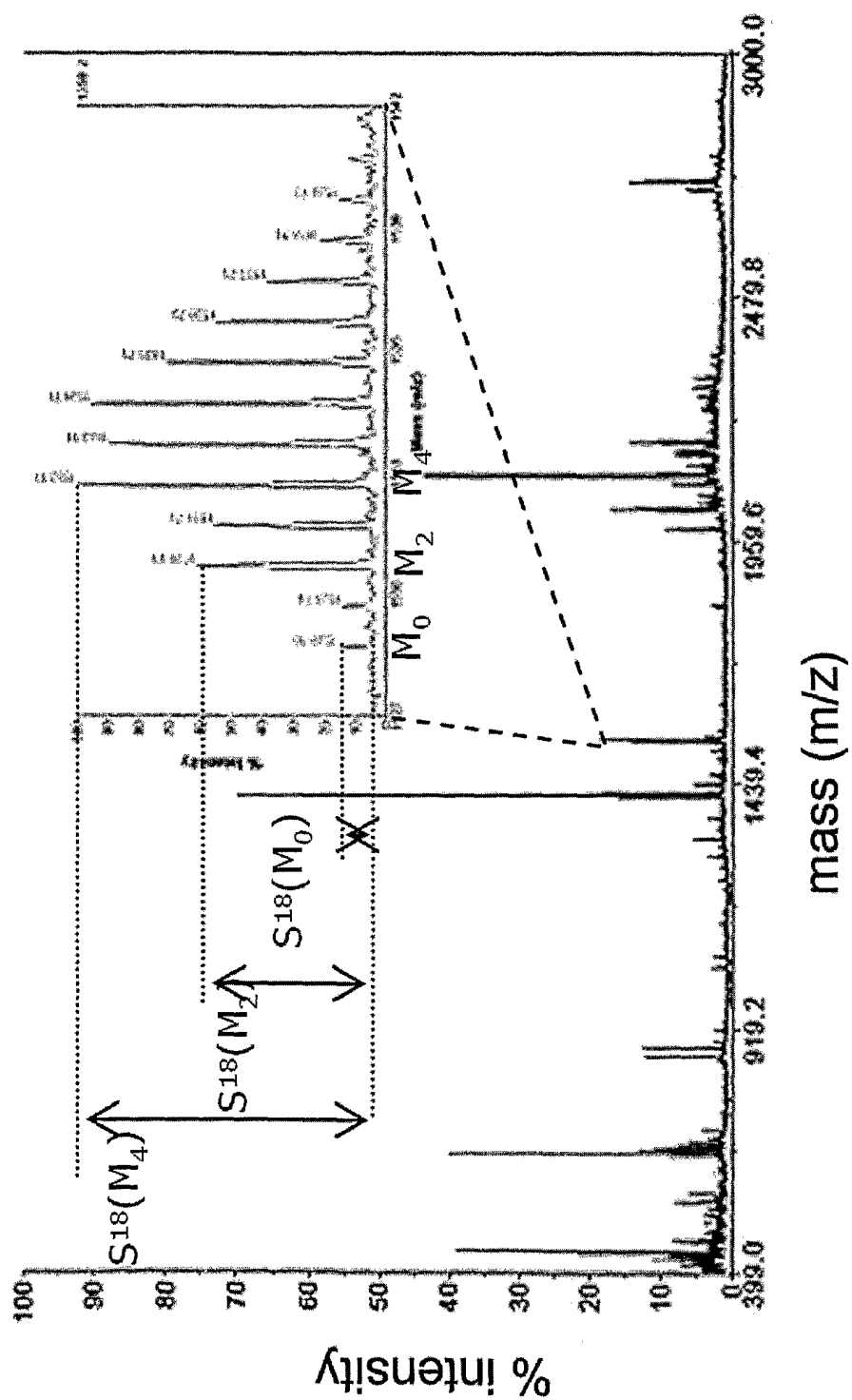
9/17

Fig. 7a



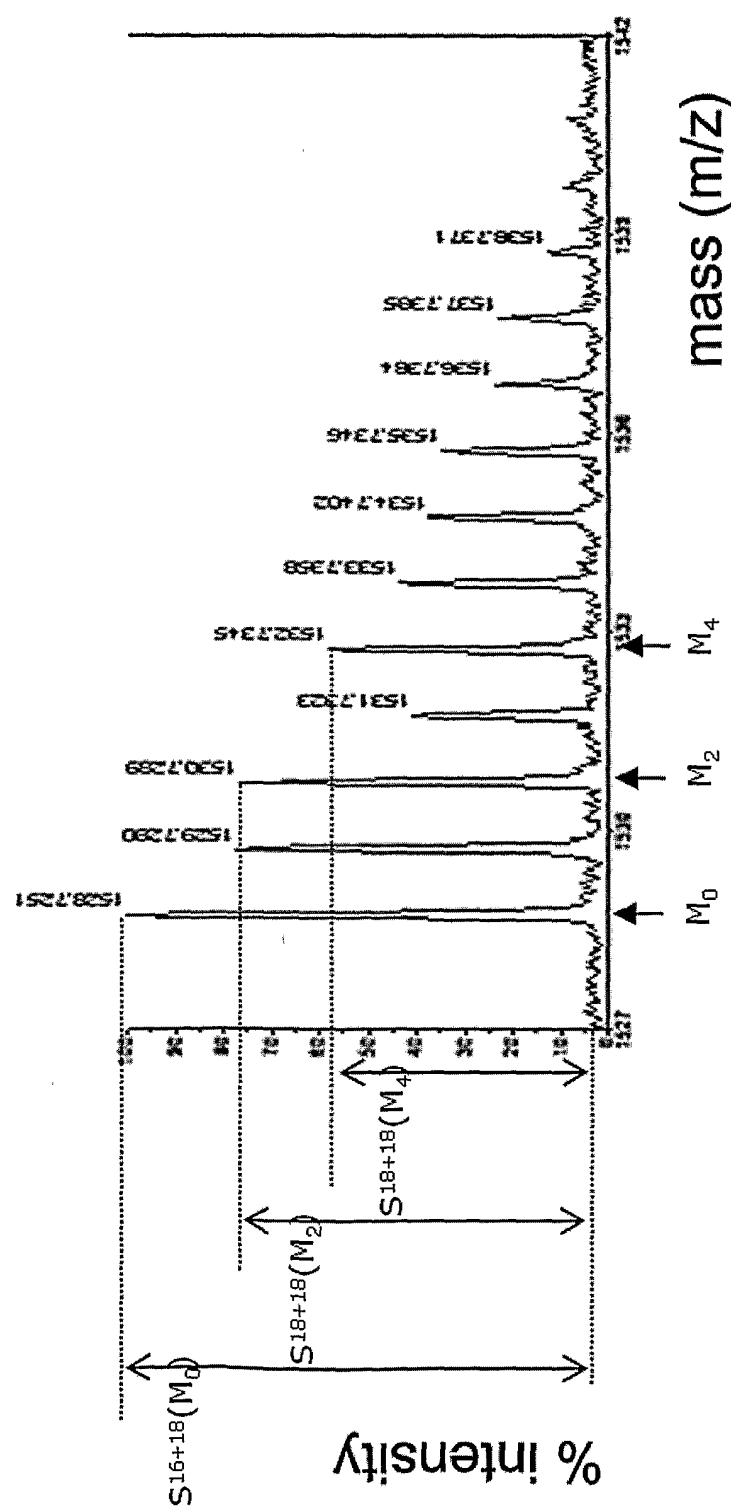
10/17

Fig. 7b



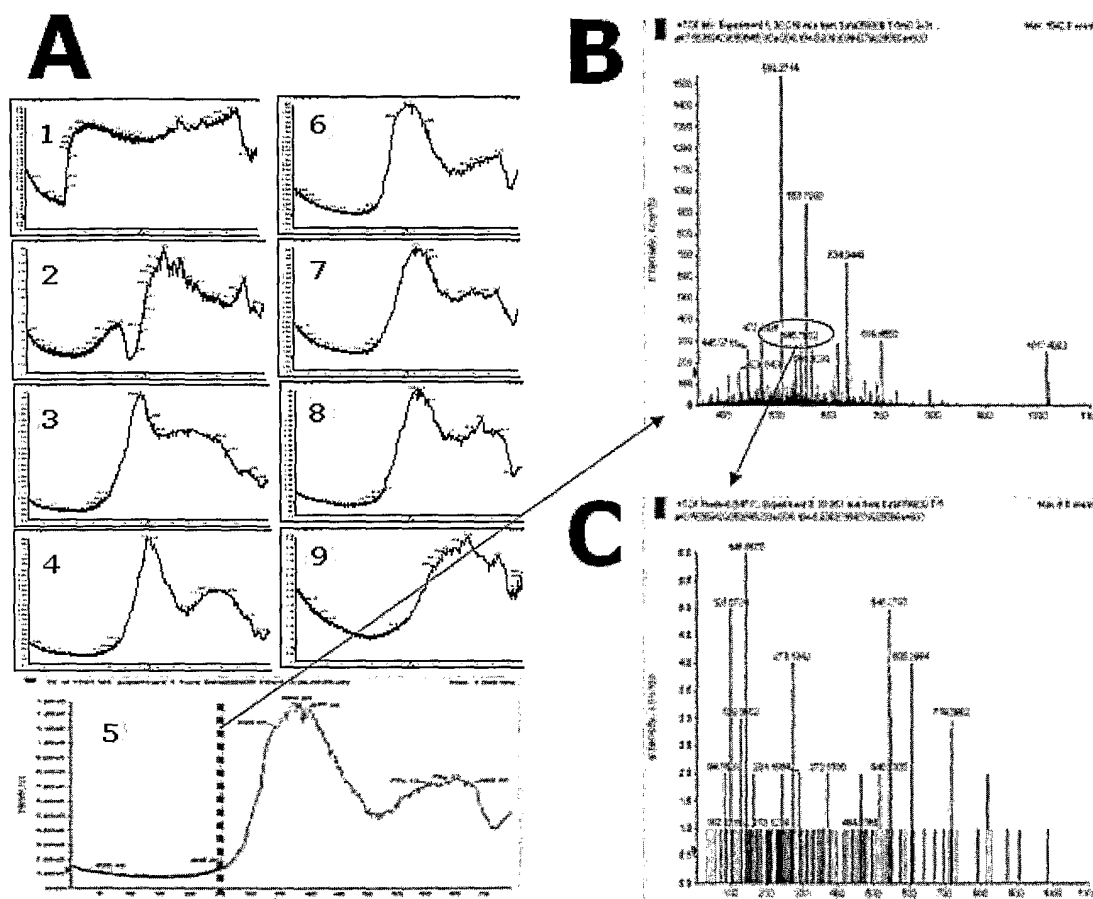
11/17

Fig. 7c



12/17

Fig. 8



13/17

Fig. 9

P-glu Q : Pyro-glu Q (N-term)

Start	End	Sequence	Modification	MH ⁺ (obs)	MH ⁺ (calc)	Delta
1	20	MQIFVKLTGKTITLEVPS		2235.31	2235.22	0.09
1	21	MQIFVKLTGKTITLEVPSD		2350.31	2350.25	0.06
1	31	MQIFVKLTGKTITLEVPSDTIENWKAKIQ		3475.02	3474.91	0.11
1	32	MQIFVKLTGKTITLEVPSDTIENWKAKIQD		3589.98	3589.93	0.05
22	32	TIENWKAKIQD		1258.72	1258.69	0.03
22	39	TIENWKAKIQDKEGIPPD		1995.1	1995.07	0.03
22	51	TIENWKAKIQDKEGIPPDQQRLLIFAGKQLE		3406.95	3406.86	0.09
22	52	TIENWKAKIQDKEGIPPDQQRLLIFAGKQLED		3522	3521.89	0.11
33	39	KEGIPPD		755.38	755.39	-0.01
33	51	KEGIPPDQQRLLIFAGKQLE		2167.21	2167.18	0.03
33	52	KEGIPPDQQRLLIFAGKQLED		2282.25	2282.21	0.04
33	58	KEGIPPDQQRLLIFAGKQLEDGRTLSD		2911.61	2911.52	0.09
40	51	QQRLLIFAGKQLE	P-glu Q	1413.8	1413.78	0.02
40	51	QQRLLIFAGKQLE		1430.82	1430.8	0.02
40	52	QQRLLIFAGKQLED	P-glu Q	1528.83	1528.8	0.03
40	52	QQRLLIFAGKQLED		1545.84	1545.83	0.01
40	58	QQRLLIFAGKQLEDGRTLSD	P-glu Q	2158.17	2158.12	0.05
53	57	GRTLS		533.32	533.3	0.02
53	58	GRTLSD		648.34	648.32	0.02
53	74	GRTLSDYNIQKESTLHLYRLR		2612.53	2612.46	0.07
59	74	YNIQKESTLHLYRLR		1983.18	1983.14	0.04

14/17

Fig. 10a

Ox-M : Oxidation M						
P-glu E : Pyro-glu E (N-term)						
start	end Peptide	Modification	MH ⁺ (calc)	MH ⁺ (obs)	Delta	Ion Score in MS/MS
1	12 DTHKSEIAHRRK(D)		1468.67	1468.7	0.03	88
1	13 DTHKSEIAHRRK(D)		1583.7	1583.72	0.02	88
2	13 (D)THKSEIAHRRK(D)		1353.73	1353.68	-0.05	51
38	55 (D)EHVKLVNELTEFAKTCVA(D)	P-glu E	2013.04			148
38	56 (D)EHVKLVNELTEFAKTCVAD	P-glu E	2128.07			118
57	71 (D)ESHAGCEKSLHTLFG(D)	P-glu E	1597.74	1597.67	-0.07	
57	71 (D)ESHAGCEKSLHTLFG(D)		1615.75	1615.7	-0.05	
57	72 (D)ESHAGCEKSLHTLFGD	P-glu E	1712.76	1712.68	-0.08	94
57	72 (D)FAVEGPKLWSTQTALA(D)		1730.77	1730.71	-0.06	58
57	89 (D)ESHAGCEKSLHTLFGDELCKVASLRETYGD MAD	P-glu E	3594.62	3594.4	-0.22	
73	85 (D)ELCKVASLRETYG(D)	P-glu E	1450.73	1450.67	-0.06	73
73	86 (D)ELCKVASLRETYGD	P-glu E	1565.76	1565.7	-0.06	57
73	88 (D)ELCKVASLRETYGDMAD(D)	P-glu E	1767.83	1767.74	-0.09	42
73	88 (D)ELCKVASLRETYGDMAD(D)	P-glu E, Ox-M	1783.83	1783.83	0	
73	89 (D)ELCKVASLRETYGDMAD	P-glu	1882.86	1882.77	-0.09	
73	89 (D)ELCKVASLRETYGDMAD	P-glu E, Ox-M	1898.85	1898.78	-0.07	30
76	85 (D)KVASLRETYG(D)		1123.6			27
76	86 (D)KVASLRETYGD		1238.6			
76	86 (D)EPQNLIKQNC	P-glu E	1283.6	1282.55	-1.05	24
76	88 (D)KVASLRETYGDMAD(D)		1440.7			
73	88 (D)ELCKVASLRETYGDMAD(D)		1785.84	1785.76	-0.08	
90	106 (D)CCEKQEPERNECFLSHK(D)		2079.9	2079.79	-0.11	
90	107 (D)CCEKQEPERNECFLSHKD		2194.92	2194.81	-0.11	
90	108 (D)CCEKQEPERNECFLSHKDD		2309.95	2309.84	-0.11	
90	111 (D)CCEKQEPERNECFLSHKDDSPD		2609.06	2608.92	-0.14	
90	118 (D)CCEKQEPERNECFLSHKDDSPDLPKLPD		3400.55	3400.39	-0.16	
108	118 DSPDLPKLPD		1224.64	1224.6	-0.04	
109	117 (D)SPDLPKLP(D)		994.59	994.55	-0.04	
109	118 (D)SPDLPKLPD		1109.61	1109.57	-0.04	41
111	118 DLPKLPD		925.53			42
111	129 DLPKLPDPPNTLCDEFKAD		2159.06			
112	124 (D)LPKLPDPPNTLQD		1453.76	1453.71	-0.05	
119	128 (D)PNTLCDEFKA(D)		1137.52	1137.47	-0.05	63
119	129 (D)PNTLCDEFKAD		1252.54	1252.5	-0.04	

15/17

Fig. 10b

Ox-M : Oxidation M P-glu E : Pyro-glu E (N-term)							
start	end	Peptide	Modification	MH ⁺ (calc)	MH ⁺ (obs)	Delta	Ion Score in MS/MS
249	265	(D) LLECADDRADLAKYTCD		1926.89	1926.81	-0.08	
255	265	DRADLAKYTCD		1282.6			
255	268	DRADLAKYTCDNQD		1639.73	1639.65	-0.08	
256	264	RADLAKYTIC		1052.55	1052.48	-0.07	
256	265	DRADLAKYTIC(D)		1167.58			22
256	268	(D) RADLAKYTCDNQD		1524.7	1524.64	-0.06	
257	266	(D) RADLAKYTCD		1167.58	1167.54	-0.04	28
269	294	(D) ITSSKLUKCCDKPILLESKSHCIAVEK(D)		2931.49	2931.33	-0.16	
269	295	(D) ITSSKLUKCCDKPILLESKSHCIAVEK(D)		3046.52	3046.34	-0.18	
280	294	(D) KPLEKSHCIAVEK(D)		1723.93	1723.86	-0.07	145
280	295	(D) KPLEKSHCIAVEK(D)		1838.96	1838.87	-0.09	74
282	294	(D) LLESKSHCIAVEK(D)		1498.8			93
296	312	(D) AIPENLPPLTADFAEDK(D)		1840.93	1840.87	-0.06	70
308	322	(D) FAEDKDVCKNYQEAK(D)		1787.82	1787.75	-0.07	
308	323	(D) FAEDKDVCKNYQEAK(D)		1902.85	1902.79	-0.06	
312	322	(D) KDVCNKYQEAK(D)		1325.64	1325.57	-0.07	
314	323	(D) WCKNYQEAK(D)		1197.55	1197.5	-0.05	
365	373	(D) PHACTSTVF(D)		1024.45	1024.4	-0.05	64
365	374	(D) PHACTSTVF(D)		1139.48	1139.43	-0.05	82
365	380	(D) PHACTSTVFDKHLV(D)		1857.96	1857.86	-0.1	99
365	381	(D) PHACTSTVFDKHLV(D)		1972.99	1972.89	-0.1	111
369	380	(D) PHACTSTVFDKHLVDEPQNLKQNC(D)		3255.59	3255.43	-0.16	
369	381	(D) YSTVFDKHLV(D)		1449.7			69
375	381	(D) KKLHLVD		1564.8			50
494	501	(D) JETYPKAF(D)	P-glu E	852.52			43
494	502	(D) JETYPKAF(D)		936.48			47
494	510	(D) JETYPKAFDEKLFTFHA(D)	P-glu E	1051.5			39
494	510	(D) JETYPKAFDEKLFTFHA(D)		2025			88
503	510	(D) JETYPKAF(D)	P-glu E	2043.01	2042.92	-0.09	54
503	510	(D) JETYPKAF(D)		974.5	974.47	-0.03	58
503	511	(D) EKLFTHAD	P-glu E	992.51	992.48	-0.03	
503	511	(D) EKLFTHAD		1089.53	1089.49	-0.04	58
503	517	(D) EKLFTHADICTLPD	P-glu E	1107.54	1107.49	-0.05	58
556	583	(D) KCCAAADKFAVEGPKLWSTQTALA		1731.83	1731.72	-0.11	
562	583	DKEACFAVEGPKLWSTQTALA		2868.39	2868.25	-0.14	68
563	583	(D) KEACFAVEGPKLWSTQTALA(D)		2277.17	2277.07	-0.1	
				2162.15	2162.03	-0.12	164

16/17

Fig. 11a

SwissProt-ID	Name	Sequence
O08917	Flotillin-1	GEAEFAIGAR AQQVAVQEQEIR VQVQVVER VQVTSQEYSAR AVDLIPWMEYEF DGEYVVEVR PPIIEGNMESAK ATSVALTWSR ETPAASEAPSSAAK SKPSSAEPAAPSSK AEGAGTEEEGTPK KTEAPAAAGPEAK AGEASAEESTGAA GAAPEEGEAK ESEPQAAADATEVK AEEK AGEASAEESTGAADGAAPEEGEAK SSASFSTTAVSAR YETSGIGEAR VLVVDQLSMR NGITEENLNK GTQSEATVNVK GLGEISAATEFK IVPSQGEISVGESK INENTGSVSVTR KDLHPNTD EFEGEEEEYLEILGITR HLTPLGR YLATASTMDHAI TTHYGSLPQK TPPPSQGK GLSATVTGGQK TQVWHESFQGR YLAEVATGEK VTLSNQPYIK LSAASGGTLDR
P12960	Contactin-1	
Q91XV3	BASP1 (NAP-22)	
O08599	Syntaxin binding protein (Unc-18 homolog)	
P13595	N-CAM 180	
Q9WTR5	Cadherin-13	
Q8BLK3	Limbic system-associated membrane precursor (LSAMP)	
P04370	MBP	
P60202	Myelin proteolipid protein (PLP)	
P20917	Myelin-associated glycoprotein precursor (Siglec-4a)	
P61982	14-3-3 gamma	
P01831	Thy-1 membrane glycoprotein precursor	
Q88998	Noelin precursor	

17/17

Fig. 11b

SwissProt-ID	Name	Sequence
P67778	Prohibitin	EFTEAVEAK
Q05186	Reticulocalbin 1	LDGDLTATR
		AAGAEYAQQEVVK
P16330	2',3'-cyclic-nucleotide 3'-phosphodiesterase	VQPVQTGLDILDILQQVK
		GGSQGEAVGELPR
		LSISALFVTPK
		GKPVPIHGSR
Q92284	cGMP-dependent 3',5'-cyclic phosphodiesterase	SEEQSLQVVEK
P11798	CAM-Kinase II alpha	PGMTAFEPEALGNLVEGLD
		GAILTTMLATR
Q9WTX5	S-phase kinase-associated protein 1A	QETV
P27601	G-protein alpha-13	FTEEEEAQVR
		ILLGAGESGK
		DVKDTILHD
		LLLKGAGESGK
P18872	G-protein G(o) alpha-1	MEDTEPFSAELLSAMMR
		AAAYIQTFESK
		SGIQECFNR
		DGISAAKQVK
		LLVSASQDGK
P62874	G-protein beta subunit 1	LLLKGAGESGK
P62880	G-protein beta subunit	EFLRISTASGD
P08752	G-protein G(i) alpha-2	MQEDIEMTK
Q9DC51	G-protein G(i) alpha-3	FVLPVGGATINM
Q8CGK7	G-protein G(olf) alpha	GTEFGGSIYQK
P63094	G-protein G(s) alpha	VTQSNFAVGKY
P63094	G-protein G(s) alpha	PTASISAK
P43006	Sodium-dependent glutamate/aspartate transporter	GTEFGGSIYQK
P43006	Sodium-dependent glutamate/aspartate transporter	PEAVLAEASSR
Q60932	VDAC-1	
Q60930	VDAC-2	
Q60931	VDAC-3	
P97441	Zinc transporter 3	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR 2005/003042

A. CLASSIFICATION OF SUBJECT MATTER

IPC⁷: G01N 33/483, C07K 1/12, H01J 49/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC⁷: C07KDocumentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, PAJ, Internet, Pubmed

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Li A, et al. "Chemical cleavage at aspartyl residues for protein identification." Anal Chem. 2001 Nov 15;73(22):5395-402. <i>*pages 5397 last paragraph, page 5399 top paragraph*.</i> --	1-41
X	De Boni S, et al. "Analysis of aspartyl peptide degradation products by high-performance liquid chromatography and high-performance liquid chromatography-mass spectrometry." J Chromatogr A. 2004 Jan 2;1022(1-2):95-102. <i>*page 97, last paragraph column 1 to first paragraph column 2.*</i> ----	1-41

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
21 December 2005 (21.12.2005)Date of mailing of the international search report
30 December 2005 (30.12.2005)Name and mailing address of the ISA/ AT
Austrian Patent Office
Dresdner Straße 87, A-1200 ViennaAuthorized officer
GÖRNER W.

Facsimile No. +43 / 1 / 534 24 / 535

Telephone No. +43 / 1 / 534 24 / 558

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR 2005/003042

Patent document cited
in search report

Publication
date

Patent family
member(s)

Publication
date

A

none