(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 January 2006 (12.01.2006)

PCT

(10) International Publication Number
WO 2006/002841 A2

(51) International Patent Classification:
G01N 33/68 (2006.01)

(21) International Application Number:
PCT/EP2005/006866

(22) International Filing Date: 22 June 2005 (22.06.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0414233.7 25 June 2004 (25.06.2004) GB

(71) Applicant *(for all designated States except US)*: **BIOINVENT INTERNATIONAL AB** [SE/SE]; Sölvegatan 41, S-223 70 Lund (SE).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: **NILSSON, Fredrik** [SE/SE]; Trollenäsgatan 19, S-211 46 Malmö (SE).

(74) Agent: **PILKINGTON, Stephanie**; Eric Potter Clarkson, Park View House, 58 The Ropewalk, Nottingham NG1 5DD (GB).

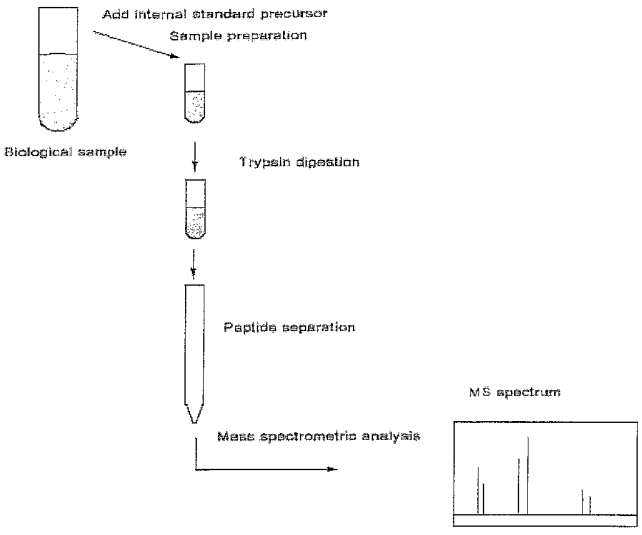
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, 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, 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:
— *without international search report and to be republished upon receipt of that report*

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.

(54) **Title:** ANALYSIS METHOD



(57) **Abstract:** A method for quantitative measurement using mass spectroscopy of a peptide or peptides generated from one or multiple analyte proteins by enzymatic or chemical cleavage, the method comprising the step of exposing together the sample to be analysed and one or multiple mass-labelled peptide internal standard precursors to enzymatic or chemical cleavage so as to generate from said precursor one or multiple mass-labelled peptide internal standards, wherein the mass-labelled peptide internal standard precursor comprises an extension of a least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom.

WO 2006/002841 A2

ANALYSIS METHOD

FIELD OF INVENTION

This invention provides methods and reagents for quantification of proteins and peptides from complex polypeptides mixtures such as plasma, tissues and cell cultures.

BACKGROUND OF INVENTION

Concentration levels of proteins and peptides in endogenous samples such as plasma and tissue are usually measured by immunochemical methods such as enzyme-linked immunosorbent assay (ELISA) or radio-linked immuno assay (RIA). These methods are well established, sensitive and generally very reliable, but requires access to specific antibodies and may suffer from problems involving cross-reactivity, poor reproducibility and denaturation of epitopes. Also, ELISA and RIA assays are as a rule developed for analysing one analyte per assay and are less suited for multiplex analysis.

Mass spectrometry (MS) is an established method for quantification of low molecular weight compounds in complex samples such as plasma and is often coupled to high-performance liquid chromatography. The main advantages with mass spectrometric detection include high resolution and sensitivity, allowing absolute quantification of low levels of analytes in complex samples.

The development of "soft ionisation" techniques such as matrix-assisted laser desorption (MALDI) and electrospray has made it possible to analyse

large biomolecules, such as peptides and proteins, using mass spectrometry. MS has mainly been used for qualitative analysis such as sequence analysis, protein identification and analysis of post-translational modifications.

As a quantitative tool, MS has been used for quantification of smaller peptides much in the same way as for low molecular weight compounds (1,2,3,4,5,6), that is as a detection method coupled to high-performance liquid chromatography or immuno-affinity chromatography and using an internal, mass-shifted standard peptide. Quantification of proteins is less straight-forward, one reason being that the resolution of intact proteins is usually insufficient to resolve different protein species in a complex sample or to distinguish the signal from a internal standard from that of the analyte.

Semi-quantitative applications include surface-enhanced laser desorption/ionisation (SELDI)(7), in which sub fractions of proteins are extracted from complex samples by affinity surfaces and analysed by linear MALDI-TOF mass spectrometry. By comparing spectra generated from different samples (e.g. plasma from healthy/sick patients) it is sometimes possible to detect changes in relative abundance of individual proteins.

Another approach to measure protein concentration is to fragment the proteins by enzymatic or chemical means into shorter peptides. The benefits of analysing peptides rather than intact proteins are that the peptides tend to display more homogenous physio-chemical properties and fall into a molecular mass range analysable by high-resolution mass spectrometric techniques such as reflectron MALDI-TOF MS or electrospray MS. On the other hand, the complexity of the sample is increased since a single protein generates many different peptides. Thus, analysis of proteins in complex

samples by fragmentation usually involves high-resolution separation steps prior to MS analysis to reduce the number of analytes introduced into the mass spectrometer. Such pre-fractionation of the fragmented samples may include multi-dimensional liquid chromatography and/or extraction of subgroups of peptides by affinity tags.

One way to quantify a given peptide generated from a specific protein is to add an internal standard with the same amino acid sequence but with a mass-shift obtained by chemical modification or introduction of stable isotopes. As an example, Barr et al (8), demonstrated how apolipoprotein A-1 was measured by trypsin fragmentation and addition of isotope-labelled internal standard peptides. More recently, the same approach was described in Gerber et al (9) although in this work MS/MS was used for quantification. This work of Gerber et al is also the subject of WO 03/016861. Gerber et al teaches that "ragged ends" and reactive amino acids should be avoided if possible by choice of protease and peptide, but if a peptide with two adjacent endoprotease cleavage sites (a ragged end) is to be analysed then an N-terminal extension can be added such that preference for either cleavage site is accounted for.

Another variation is to add an affinity label, such as biotin, to a subpopulation of the peptides, for example by employing a cysteine-specific chemistry. It is possible to measure relative difference in protein content between two different samples by introducing affinity tags with different masses. This method is called isotope-coded affinity tags (ICAT, 10) and has the advantage of introducing a mass-shift in the analytes in the same step as an affinity tag is added. However, this method will not give an absolute quantification of the protein or peptide of interest.

An internal standard should preferably be added to the sample to be analysed as early as possible during the analytical process so that all sample handling steps will affect the standard and analyte equally. The addition of a synthetic small peptide to a protein mixture prior to sample preparation may introduce several problems: i) the yield of the internal standard after sample preparation will not reflect the efficiency of the fragmentation step; ii) a small and maybe hydrophobic peptide can be partially or completely lost during sample preparation due to adsorption to other proteins or plastic sample vessels; and iii) small peptides may not necessarily co-purify with the protein to be analysed during certain types of sample preparations steps such as precipitation, ultra filtration and chromatography.

Thus, there is a need for a method for quantification of proteins and peptides, which is free from the above drawbacks.

SUMMARY OF THE INVENTION

The present invention provides a method for quantitative measurement of peptides generated from intact proteins by enzymatic or chemical cleavage, using mass-labelled peptide internal standard precursors. The mass-labelled peptide internal standard precursor contains an identical or almost identical (for example differing in one amino acid) amino-acid sequence to the analyte protein fragment but is extended at both the C-terminus and N-terminus with one or more amino acids.

The present invention also provides a kit that may be used for the quantification of proteins and peptides in a sample.

The present invention also provides a diagnostic method for quantifying one or more peptides or proteins in a sample.

DETAILED DESCRIPTION OF THE INVENTION

The term "mass-labelled peptide internal standard precursor" includes a synthetic peptide or a recombinant protein that contains one or several amino acid sequences which corresponds to sequences in one or several known or predicted proteins and which is labelled in such a manner that the internal standard precursor will yield, upon fragmentation by proteases or chemicals, fragments identical or almost identical to those generated from the protein(s) to be analysed except for a slightly different molecular mass. The chemical modification of the mass-labelled peptide internal standard precursor is preferably an incorporation of a stable isotope. In this case, the fragments from the internal standard precursor and from the analyte protein will be identical except for a slightly different molecular mass.

The term "polypeptide" includes a compound containing two or more amino acids, in which the amino acids are linked by peptide bonds.

"Protein" refers to any protein without limitation and preferably includes those comprised of at least more than 50 amino acids.

"Peptide" refers to shorter polypeptides and preferably includes those comprising 50 or less amino acids, for example between 4 and 50 amino acids, or between 10 and 24 amino acids.

"Analyte protein" and "analyte peptide" refers to the specific protein or peptide to be quantified.

A first aspect of the invention provides a method for quantitative measurement using mass spectroscopy of a peptide or peptides generated from one or multiple analyte proteins by enzymatic or chemical cleavage, the method comprising the step of exposing together the sample to be analysed and one or multiple mass-labelled peptide internal standard precursors to enzymatic or chemical cleavage so as to generate from said precursor one or multiple mass-labelled peptide internal standards, wherein the mass-labelled peptide internal standard precursor comprises an extension of at least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom.

An embodiment of this aspect of the invention is wherein the said N-terminal extension comprises at its C-terminal end an amino acid sequence of at least one, two, three, four or five amino acids identical to the amino acid sequence found in an analyte protein immediately N-terminal (prior to cleavage) to the sequence corresponding to the N-terminal-most mass-labelled peptide internal standard of the precursor.

A further embodiment of this aspect of the invention is wherein the said C-terminal extension comprises at its N-terminal end an amino acid sequence of at least one, two, three, four or five amino acids identical to the amino acid sequence found in an analyte protein immediately C-terminal (prior to

cleavage) to the sequence corresponding to the C-terminal-most mass-labelled peptide internal standard of the precursor.

A further embodiment of this aspect of the invention is wherein the amino acid sequence of at least one, two, three, four or five amino acids immediately N-terminal of a mass-labelled peptide internal standard of the precursor is identical to the amino acid sequence found in an analyte protein immediately N-terminal (prior to cleavage) to the sequence corresponding to that mass-labelled peptide internal standard.

A further embodiment of this aspect of the invention is wherein the amino acid sequence of at least two, three, four or five amino acids immediately C-terminal of a mass-labelled peptide internal standard of the precursor is identical to the amino acid sequence found in the analyte protein immediately C-terminal (prior to cleavage) to the sequence corresponding to that mass-labelled peptide internal standard.

The invention provides a method for quantitative measurement of peptides generated from intact proteins by enzymatic or chemical cleavage, using mass-labelled peptide internal standard precursors. The mass-labelled peptide internal standard precursor contains an identical or almost identical amino-acid sequence to the analyte protein fragment but is extended on both the C-terminus and N-terminus with one or more amino acids. The extensions may or may not be identical to the extension amino acids in the analyte protein adjacent to the generated fragment prior to fragmentation. Cleavage of the analyte protein and the mass-labelled peptide internal standard precursor will yield an analysable peptide from the protein and an internal standard peptide from the internal standard precursor, the only

difference between the two from an analytical perspective being a mass-shift of typically 2-10 Dalton (Fig 1)

By using a mass-shifted peptide internal standard precursor several advantages are achieved: i) variation in the yield of peptides after fragmentation of the protein(s) is compensated for as the internal standard precursor will fragment with the same yield as that of the analyte protein or analyte peptide; ii) the internal standard precursor may be designed to co-purify with the analyte protein or analyte peptide during various sample preparations steps performed prior to protein fragmentation; and iii) a single internal standard precursor may be designed to generate internal standard peptides for analysis of peptide fragments originating from several different analyte protein or analyte peptide.

The mass-labelled peptide internal standard precursors can be generated by peptide synthesis, during which one or more mass-labelled amino acids are incorporated. Preferably, the mass-label amino acids contain of one or more stable isotopes, including but not limited to ^{13}C, ^{15}N, ^{18}O, ^{2}H, and ^{34}S.

Peptides may be synthesised by the Fmoc-polyamide mode of solid-phase peptide synthesis as disclosed by Lu *et al* (1981) *J. Org. Chem.* **46**, 3433 and references therein. Temporary N-amino group protection is afforded by the 9-fluorenylmethyloxycarbonyl (Fmoc) group. Repetitive cleavage of this highly base-labile protecting group is effected using 20% piperidine in N,N-dimethylformamide. Side-chain functionalities may be protected as their butyl ethers (in the case of serine threonine and tyrosine), butyl esters (in the case of glutamic acid and aspartic acid), butyloxycarbonyl derivative (in the case of lysine and histidine), trityl derivative (in the case of cysteine) and 4-

methoxy-2,3,6-trimethylbenzenesulphonyl derivative (in the case of arginine). Where glutamine or asparagine are C-terminal residues, use is made of the 4,4'-dimethoxybenzhydryl group for protection of the side chain amido functionalities. The solid-phase support is based on a polydimethyl-acrylamide polymer constituted from the three monomers dimethylacrylamide (backbone-monomer), bisacryloylethylene diamine (cross linker) and acryloylsarcosine methyl ester (functionalising agent). The peptide-to-resin cleavable linked agent used is the acid-labile 4-hydroxymethyl-phenoxyacetic acid derivative. All amino acid derivatives are added as their preformed symmetrical anhydride derivatives with the exception of asparagine and glutamine, which are added using a reversed N,N-dicyclohexyl-carbodiimide/1-hydroxybenzotriazole mediated coupling procedure. All coupling and deprotection reactions are monitored using ninhydrin, trinitrobenzene sulphonic acid or isotin test procedures. Upon completion of synthesis, peptides are cleaved from the resin support with concomitant removal of side-chain protecting groups by treatment with 95% trifluoroacetic acid containing a 50% scavenger mix. Scavengers commonly used are ethanedithiol, phenol, anisole and water, the exact choice depending on the constituent amino acids of the peptide being synthesised. Trifluoroacetic acid is removed by evaporation *in vacuo*, with subsequent trituration with diethyl ether affording the crude peptide. Any scavengers present are removed by a simple extraction procedure which on lyophilisation of the aqueous phase affords the crude peptide free of scavengers. Reagents for peptide synthesis are generally available from Calbiochem-Novabiochem (UK) Ltd, Nottingham NG7 2QJ, UK. Purification may be effected by any one, or a combination of, techniques such as size exclusion chromatography, ion-exchange chromatography and (principally) reverse-phase high performance liquid chromatography. Analysis of peptides may be carried out using thin

layer chromatography, reverse-phase high performance liquid chromatography, amino-acid analysis after acid hydrolysis and by fast atom bombardment (FAB) mass spectrometric analysis.

The mass label may also consist of other chemical modifications of the amino acids including, but not limited to, acetylation, methylation, deamidation and carboxymethylation. If the mass label modification is other than incorporation of stable isotopes, it is important to avoid modifications that may alter the internal standard peptides behaviour during sample preparation, separation and ionisation.

Furthermore, the mass-labelled peptide internal standard precursor may be synthesised to have one or more replacement amino acid residues with reference to the analyte protein or peptide; for example, an alanine in replacement of a glycine. In such a circumstance, cleavage of the analyte protein or peptide and the mass-labelled peptide internal standard precursor will yield an analysable peptide from the protein and an internal standard peptide from the internal standard precursor, the only difference between them from an analytical perspective being a mass-shift due to the presence of alanine in the peptide internal standard relative to the presence of glycine in the analyte protein or peptide. An advantage of such an embodiment of the invention is that it may be easier to replace an alanine with a glycine rather than with an isotopic-labelled alanine during synthesis of the mass-labelled peptide internal standard precursor. Further suitable amino acid replacements can be readily appreciated by a person of skill in the art, for example valine and isoleucine may be interchanged.

The mass-labelled peptide internal standard precursor may have no, one, two, three, four or five replacement amino acid residues with reference to the analyte protein or peptide, or may differ in up to five, ten, fifteen, twenty or twenty-five percent of the total number of amino acid residues with reference to the analyte protein or peptide. Preferably the mass-labelled peptide internal standard precursor has no or one replacement amino acid residue with reference to the analyte protein or peptide.

If the analyte protein fragment contains any posttranslational modifications such as phosphorylation, the mass-labelled peptide internal standard may be synthesised to be identically modified.

The mass-labelled peptide internal standard precursors may also be generated by recombinant protein expression in the presence of mass-labelled amino acids as a part of a hybrid protein, in which case the hybrid protein as such can constitute the internal standard precursor or the hybrid protein may be processed to yield one or more peptide internal standard precursors.

In preparing a mass-labelled peptide internal standard precursor using recombinant protein expression, a DNA molecule is prepared that encodes the desired peptide using methods well known to those skilled in the art and exemplified by Sambrook *et al* (2001) "*Molecular Cloning, a Laboratory Manual*", 3rd edition, Sambrook *et al* (eds), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.

The DNA is then expressed in a suitable host to produce a mass-labelled peptide internal standard precursor. Thus, the DNA encoding the said

peptide may be used in accordance with known techniques, appropriately modified in view of the teachings contained herein, to construct an expression vector, which is then used to transform an appropriate host cell for the expression and production of the said peptide. Such techniques include those disclosed in US Patent Nos. 4,440,859 issued 3 April 1984 to Rutter *et al*, 4,530,901 issued 23 July 1985 to Weissman, 4,582,800 issued 15 April 1986 to Crowl, 4,677,063 issued 30 June 1987 to Mark *et al*, 4,678,751 issued 7 July 1987 to Goeddel, 4,704,362 issued 3 November 1987 to Itakura *et al*, 4,710,463 issued 1 December 1987 to Murray, 4,757,006 issued 12 July 1988 to Toole, Jr. *et al*, 4,766,075 issued 23 August 1988 to Goeddel *et al* and 4,810,648 issued 7 March 1989 to Stalker, all of which are incorporated herein by reference.

The DNA encoding a mass-labelled peptide internal standard precursor may be joined to a wide variety of other DNA sequences for introduction into an appropriate host. The companion DNA will depend upon the nature of the host, the manner of the introduction of the DNA into the host, and whether episomal maintenance or integration is desired.

Generally, the DNA is inserted into an expression vector, such as a plasmid, in proper orientation and correct reading frame for expression. If necessary, the DNA may be linked to the appropriate transcriptional and translational regulatory control nucleotide sequences recognised by the desired host, although such controls are generally available in the expression vector. Thus, the DNA insert may be operatively linked to an appropriate promoter. Bacterial promoters include the E.coli *lacI* and *lacZ* promoters, the T3 and T7 promoters, the *gpt* promoter, the phage λ PR and PL promoters, the *phoA* promoter and the *trp* promoter. Eukaryotic promoters include the

CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters and the promoters of retroviral LTRs. Other suitable promoters will be known to the skilled artisan. The expression constructs will desirably also contain sites for transcription initiation and termination, and in the transcribed region, a ribosome binding site for translation. (Hastings *et al*, International Patent No. WO 98/16643, published 23 April 1998)

The vector is then introduced into the host through standard techniques. Generally, not all of the hosts will be transformed by the vector and it will therefore be necessary to select for transformed host cells. One selection technique involves incorporating into the expression vector a DNA sequence marker, with any necessary control elements, that codes for a selectable trait in the transformed cell. These markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture, and tetracyclin, kanamycin or ampicillin resistance genes for culturing in *E.coli* and other bacteria. Alternatively, the gene for such selectable trait can be on another vector, which is used to co-transform the desired host cell.

Host cells that have been transformed by the recombinant DNA encoding a mass-labelled peptide internal standard precursor are then cultured for a sufficient time and under appropriate conditions known to those skilled in the art permit the expression of the polypeptide, which can then be recovered.

The a mass-labelled peptide internal standard precursor can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulphate or ethanol precipitation, acid extraction,

anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Many expression systems are known, including systems employing: bacteria (eg. *E.coli* and *Bacillus subtilis*) transformed with, for example, recombinant bacteriophage, plasmid or cosmid DNA expression vectors; yeasts (eg. *Saccaromyces cerevisiae*) transformed with, for example, yeast expression vectors; insect cell systems transformed with, for example, viral expression vectors (eg. baculovirus); plant cell systems transfected with, for example viral or bacterial expression vectors; animal cell systems transfected with, for example, adenovirus expression vectors.

Suitable amino acid sequences for use as a mass-labelled peptide internal standard may be determined by predicting fragmentation products of a protein based on the specificity of the protein and the analyte protein's amino acid sequence using *in silico* methods well known to those skilled in the art. The amino acid sequence or sequences suitable for use as peptide internal standard precursors may also be determined by fragmentation and mass spectrometric analysis of the analyte protein and the actual primary structure of the generated peptide fragments are determined. In another aspect, an unknown peptide discovered by, for example, differential analysis of fragmented proteins from e.g. healthy and sick patient may be identified and sequenced and the information used to design a mass-labelled peptide internal standard precursor. In this case (as an example), the extension sequence may not match flanking sequences in the analyte

protein, unless it is possible to identify (following sequencing) the polypeptide from which the unknown peptide was derived and thereby identify the flanking amino acids in the analyte protein.

The analyte protein can be any known protein or a hypothetical protein predicted by analysis of nucleic acid sequences. The analyte protein can also be previously unknown and its existence indicated by protein expression analysis of intact proteins or fragmented proteins using 2-dimensional electrophoresis, liquid chromatography, mass spectrometry or other analytical methods or combination thereof. Such methods are well known to those skilled in the art.

The mass-labelled peptide internal standard precursors may be designed to detect and measure modified proteins and the peptides, where the modifications include but are not limited to phosphorylation, glycosylation, oxidation, farnesylation, acetylation, ubiquination, lipidation, prenylation and sulfonation. They may also be designed to detect and measure known or predicted protein and peptide species generated by alternative splicing of mRNA, by specific or unspecific degradation of the protein *in vivo* or by variation due to single nucleotide polymorphisms.

A single internal standard precursor may be designed and synthesised to generate one, two or more internal standard peptides upon cleavage. Hence a further embodiment of this aspect of the invention is wherein a mass-labelled peptide internal standard precursor comprises more than one mass-labelled peptide internal standard. A further embodiment of this aspect of the invention is wherein a mass-labelled peptide internal standard precursor

comprises mass-labelled peptide internal standards for analysis of peptide fragments originating from multiple analyte proteins.

A heterogeneous sample of peptides or proteins may be extracted from a cell or tissue sample, or derived from fragmentation of a heterogeneous sample of peptides and proteins extracted form a cell or tissue sample, typically (but not necessarily) of human origin. The cell or tissue sample may be derived from normal or diseased tissue. The cell or tissue sample may be derived from tissues at various states of differentiation or activity. Additional appropriate sources of proteins and peptides include prokaryotes, eukaryotic cell lines, tissue materials from knockout mice and other animal models as well as transgenic plants and plant material. Methods of extracting proteins from such tissue samples are well known to those skilled in the art.

The mass-labelled peptide internal standard of use in this aspect of the invention can be any length that is suitable for the method of the invention and are typically of a length of between 4 and 50 amino acids, preferably a length of between 10 and 40 amino acids. The size of the mass-labelled peptide internal standard is dictated by the requirement for the standard peptide to have a minimum size such that it can be related to the analyte protein or peptide and a maximum size such that the standard peptide can be resolved using existing methods of mass spectrometry.

A further embodiment of this aspect of the invention is wherein the mass-labelled peptide internal standard precursor has a length of between 6 and 200 amino acids. As mentioned above the precursor may contain one or several amino acid sequences which corresponds to sequences in one or

several known or predicted proteins. Since the mass-labelled peptide internal standard precursor is extended at both the C-terminus and N-terminus with one or more amino acids then the minimum size of the mass-labelled peptide internal standard precursor is 6 amino acids.

A further embodiment of this aspect of the invention is wherein the mass-labelled peptide internal standard precursor co-purifies with the analyte protein or analyte proteins to be measured during sample preparation steps performed prior to exposure to enzymatic or chemical cleavage.

A further embodiment of this aspect of the invention is wherein the mass-labelled peptide internal standard precursor is combined with the sample to be analysed after sample collection but prior to any sample preparation or fractionation steps.

With reference to Figs 2 and 3 the mass-shifted peptide internal standard precursor is added to the sample to be analysed, preferably as soon as possible after sample collection and preferably prior to any sample preparation or fractionation steps. Hence the mass-shifted peptide internal standard precursor and the analyte protein or peptide are subjected to identical sample preparation or fractionation steps. Several different internal standard precursors may be added to allow quantification of multiple different protein and peptides simultaneously, and several internal standard precursors can be added for different fragments from the same protein to obtain redundant information. Alternatively, a single internal standard precursor generating more than one internal standard may be used.

The sample can then, if required, be prepared for analysis by removing substances that may interfere with the analysis and to enrich the analyte(s). Methods that may be used include solid phase extraction, liquid chromatography, precipitation, ultra filtration and purification using affinity-based techniques, as would be appreciated to those skilled in the art.

The proteins in the sample may also be chemically modified, for example by reduction and carboxymethylation of cysteines to break disulfide bridges and avoid formation of peptide dimers.

The method of the invention comprises the step of fragmenting the heterogeneous sample of proteins or peptides to produce a heterogeneous sample of peptide fragments.

The step of fragmenting of the heterogeneous sample of proteins, polypeptides or peptides may be achieved by any method known in the art. For example, chemical or enzymatic cleavage may be used. Numerous methods of chemical or enzymatic (i.e. protease directed) cleavage are known in the art. For example, proteases include trypsin, chymotrypsin, pepsin, thrombin, papain, bromelain, thermolysin, subsilisin, Factor Xa, *Staphylococcus aureus* protease and carboxypeptidase A. In a preferred embodiment, the fragmentation method will cleave proteins, polypeptides or peptides at defined locations. Enzymatic cleavage is typically sequence-directed, as shown in Table 1 below. Hence a further embodiment of this aspect of the invention is wherein the enzymatic or chemical cleavage is sequence-directed.

Chemical cleavage methods may also be sequence-directed e.g. cyanogen bromide fragmentation, which will cleave a protein or peptide on the C-terminal side of methionine.

Table 1

Enzyme	Preferred Site
trypsin:	R_1 = Lys, Arg
chymotrypsin	R_1 = Tyr, Phe, Leu, Ile, Val, Trp and His at high pH
pepsin	R_1 = Phe, Leu, many others
thrombin	R_1 = Arg
papain	R_1 = Arg, Lys, Phe-X (CO side of residue next to Phe)
bromelain	R_1 = Lys, Ala, Tyr, Gly
Staphylococcus aureus protease	R_1 = Glu, Asp
Factor Xa	R_1 = Ile-Glu-Gly-Arg
thermolysin	R_2 = Tyr, Phe, Leu, Ile, Val, Trp and His

Wherein R_1 and R_2 are defined according to the following formula:

N-terminal---NH-CHR$_1$-CO-NH-CHR$_2$-CO---C-terminal

Thus, for example, trypsin cleavage is a sequence-directed means of fragmentation, since cleavage is directed by the presence of arginine or lysine residues in a protein, polypeptide or peptide, and accordingly produces cleavage fragments that have, as their C-terminal residue, either an arginine or lysine.

As mentioned above, the mass-labelled peptide internal standard precursor comprises an extension of at least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal

to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom. Suitable amino acid residues for such extensions include hydrophilic or neutral amino acids, for example glycine, alanine, valine, leucine, isoleucine, serine and threonine.

Less suitable amino acid residues include cysteine (which may invoke disulphide bridge formation), proline (this may render the resulting peptide less flexible and influence the structure of the epitope as well as making the peptide less susceptible to trypsin cleavage) and hydrophobic or aromatic residues (as these may reduce the solubility of the peptide). In addition, if trypsin is used as a means of sequence-directed means of fragmentation (as discussed above) then trypsin may cleave to a lesser extent at the C-terminus of hydrophobic amino acid residues.

A further embodiment of this aspect of this invention is wherein the mass-labelled peptide internal standard precursor and sample comprising one or multiple analyte proteins are separated after enzymatic or chemical cleavage.

Fragmentation of a heterogeneous protein population into peptides will usually yield a highly complex mixture of peptides that requires separation to prior mass spectrometric analysis to reduce the complexity. Methods that may be used for peptide separation include liquid chromatography (in one or more separation dimensions), solid phase extraction and affinity capture, as would be appreciated by a person skilled in the art.

The separation can be coupled to mass spectrometric analysis either on-line (liquid chromatography coupled to electrospray mass spectrometry) or by

fractionation followed by analysis of the individual fraction by e.g. matrix assisted laser desorption/ionisation mass spectrometry (MALDI). The difference in abundance between the analyte peptide and the mass-labelled internal standard peptide can be measured either by single-stage mass spectrometry (MS) or multi-stage MS (MS/MS or MS^n). In a two-stage mass spectrometric analysis, the first mass separation is used to select the peptides to be analysed, the selected peptides are then fragmented and the fragments analysed in the second-stage mass separation.

MS/MS is a very powerful approach for analysing complex peptide mixtures as it has the potential of resolving peptides with identical molecular mass but with different amino acid composition.

The intensity of the signal obtained from a specific peptide by mass spectrometry is dependent on the concentration, molecular weight and ionisation characteristics of the peptide as well as quenching effects of other components in the sample. For two peptides that are identical apart from e.g. isotopic composition, the relative signal intensity will depend only on the concentrations, as all other factors should affect them equally (Fig 2).

A preferred method for separating peptides is 'Signature Peptide Capture' (SPC) as set out in PCT/EP2004/002566 and herein incorporated by reference.

SPC is a method for analysing a heterogeneous sample of peptides or proteins, or fragments thereof, the method comprising-

(a) separating a heterogeneous sample of peptides or proteins or fragments thereof into heterogeneous classes by binding the heterogeneous peptide or protein members of each class to a spaced apart defined location on an array, wherein peptides or proteins in each class have a motif common to that class; and

(b) characterising the peptides or proteins in each class.

Each heterogeneous class of peptides or proteins consists of all peptides or proteins in the heterogeneous sample that will bind to a specific binding molecule present on the array. The binding molecule is selected for its ability to bind a motif, rather than a particular protein or peptide, and so a binding molecule can bind different types of proteins and peptides containing the same motif. Preferably each binding molecule is specific for a given motif. Thus, a heterogeneous class of proteins and peptides bound by a given binding molecule in SPC typically comprises, as a mean average, at least two, more typically greater than two, such as 10, 20, 50, 100, 200, 500, 1000 or more, different types of protein or peptide.

Accordingly, proteins and peptides are classified by SPC based on their ability to be captured and retained by a specific binding molecule. A heterogeneous class of peptides or proteins will bind to specific binding molecule due to the presence of a motif common to all members of a particular class. The identity of the motif bound in each class of peptides is, therefore, a consequence of the binding specificity of the binding molecule that defines that class.

The method of the invention would allow for SPC to not only provide analysis for a heterogeneous sample of proteins or peptides, but also to quantify one or more components of the heterogeneous sample.

A further aspect of the invention is a method of aiding diagnosing whether a patient has or may be susceptible to developing a disorder characterised by the increase or decrease of levels of one or more polypeptides comprising obtaining a test sample from said patient and quantifying the levels of said one or more polypeptides using the method of the first aspect of the invention.

The method of the invention can be used in conjunction with other diagnostic methods to determine whether a patient has or may be susceptible to developing a disorder. This is because the method of the invention can be used to quantify protein or peptide levels in a test sample. Hence the diagnosis of diseases that result in the increase or decrease of levels of one or more polypeptides can be aided by employing the method of the invention.

By 'a test sample' we include a sample of a body fluid such as blood, serum, plasma, urine, cerebrospinal fluid, pleural fluid and semen.

Examples of disorders that may result in the increase or decrease of levels of one or more polypeptides include cancer, stroke, metabolic disorders, inflammation, mycocardial infarction or atherosclerosis.

A further aspect of the invention is a mass-labelled peptide internal standard precursor comprising an extension of at least one amino acid N-terminal to

the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom, wherein on exposure to enzymatic or chemical cleavage said precursor generates one or multiple mass-labelled peptide internal standards.

The term "mass-labelled peptide internal standard precursor" includes those peptides or proteins defined in relation to the first aspect of the invention. Hence the internal standard precursor will yield, upon fragmentation by proteases or chemicals, fragments identical or almost identical to those generated from a protein(s) to be analysed except for a slightly different molecular mass.

Methods of enzymatic or chemical cleavage of the mass-labelled peptide internal standard precursor are set out above in relation to the first aspect of the invention.

The amino acid sequence of the mass-labelled peptide internal standard precursor will necessarily vary depending on the target analyte protein or peptide, as would be appreciated by a person skilled in the art.

A further aspect of the invention provides a kit of parts comprising a mass-labelled peptide internal standard precursor as defined in above or in relation to the first aspect of the invention and an agent for fragmenting peptides.

A further aspect of the invention provides a kit of parts comprising a mass-labelled peptide internal standard precursor as defined in above or in relation to the first aspect of the invention and a test sample containing or to be tested for (for example thought to contain) the analyte protein or peptide. The kit may further comprise an agent for fragmenting peptides.

As mentioned above the mass-labelled peptide internal standard precursor comprises an extension of at least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom.

Suitable agents for fragmenting peptides are set out above in relation to the first aspect of the invention. The agent to be supplied as part of the kit will necessarily vary in accordance with the fragmentation properties of the mass-labelled peptide internal standard precursor, for example should the precursor give rise to the correct mass-labelled peptide internal standard after digestion with trypsin, then the kit would comprise trypsin.

The kit may also comprise further reagents for use in the method of the invention, for example an affinity array in a chip format, antibodies or affinity columns, as would be appreciated by a person skilled in the art.

Preferably, the kit would also comprise all the components required to perform SPC-mediated separation of fragmented peptide as discussed above in relation to the first aspect of the invention.

Hence the kit may also comprise an array having a number of different types of binding molecule each immobilised at a discrete region on a solid support (typically glass or a polymer). Each binding molecule has the ability to specifically bind a motif, rather than a particular protein or peptide, as discussed above in relation to the first aspect of the invention.

All documents referred to herein are, for the avoidance of doubt, hereby incorporated by reference.

The invention is now described by reference to the following, non-limiting, figures and examples.

Figure 1: Cleavage of an analyte protein and the mass-labelled peptide internal standard precursor to yield an analysable peptide from the protein and an internal standard peptide from the internal standard precursor.

Figure 2: Analysis of a biological sample and a mass-shifted peptide internal standard precursor.

Figure 3: Mass spectrometry analysis of analyte and internal standard peptides.

Example 1: Quantification of transferrin in human serum.

Transferrins are iron binding transport proteins, which are responsible for the transport of iron from sites of absorption and heme degradation to those of storage and utilization. Serum transferrin may also have a further role in stimulating cell proliferation. The plasma levels of transferring is increased

during pregnancy and iron deficiency and may increase due to imflammation, tumors, nephrosos and haemochromatosis. Normal plasma levels are 2-4 mg/ml.

An internal peptide standard precursor with the amino acid sequence "EPNNK EGYYGYTGAFR*CLVE" is synthesised, in which the arginine is isotopically labelled by incorporation of 6 ^{13}C atoms. The peptide is dissolved in phosphate buffer saline (PBS) pH 7.4 at a concentration of 0.5 mg/ml.

Sample preparation.

0.1 ml of the serum sample to be analysed is mixed with 0.9 ml of the internal peptide standard precursor solution. The sample is reduced by addition of 220 µl 75 mM dithiothreitol in PBS and incubated at 50 °C for 80 minutes. 60 µl of 1 M iodoacetic acid in PBS is added and the sample is incubated at 37 °C in darkness for 30 minutes. Remaining iodoacetic acid is reduced by addition of 220 µl of 75 mM dithiothreitol in PBS. The reduced and alkylated proteins are separated from excess reagents by gel filtration on a PD-10 column (Amersham Biotech) in accordance with the protocol provided by the manufacturer. The protein fraction is eluted in 3 ml PBS with an approximate protein concentration of 2.5 mg/ml. 150 ul trypsin solution (1 mg/ml trypsin in PBS) is added and the sample is incubated for 2 h at 37 °C. The digested sample is frozen at –80 °C if not analysed immediately.

Separation of peptides.

The peptides are separated by reverse-phase high performance liquid chromatography on a capillary column (100 µm x 10 cm) with C18

stationary phase at a flow rate of 2 µl/minute. 10 µl sample was loaded and the peptides eluted by a mobile phase gradient from 5 to 60 % phase B during 45 minutes (eluante A: 2 % acetonitril, 0.1 % trifluoric acetic acid, eluante B: 95 % acetonitril, 0.1 % trifluoric acid). The eluted peptides are collected in 2 µl fractions directly on a MALDI-TOF metal target and are allowed to dry.

To identify the fraction containing the relevant peptide, 0.9 ml internal peptide standard precursor solutions is prepared as described above but replacing the serum sample with 0.1 ml PBS. The peptides are separated as described and collected on a separate MALDI target. The MALDI target plate is scanned as bellow to localise the peptide on the plate.

Mass spectrometric analysis.

1 µl alpha-cyano-sinapinic acid solution (5 mg/ml in 1 % trifluoric acetic acid) is added to each fraction. The fractions are analysed using a MALDI-TOF mass spectrometer equipped with a reflector.

The analyte peptide from transferrin with the sequence "EGYYGYTGAFR" will generate a signal corresponding to a mass of 1284.57 dalton (1283.57 plus one proton), whereas the internal peptide standard will be detected as having a mass of 1280.57 dalton. The relative intensity of the signals will reflect the relative amounts of analyte and reference peptides.

1. Crowther J, Adusumalli V, Mukherjee T, Jordan K, Abuaf P, Corku N, Goldstein G, Tolan J "Determination of nanogram levels of peptide drug in rabbit and human plasma using high-performance chromatography coupled with electrospray ionisation mass spectrometry", Anal Chem. 1994 Jul 15;66(14):2356-61
2. Kobayashi N, Kanai M, Seta K, Nakamura K. "Quantitative analysis of synthetic human calcitonin by liquid chromatography-mass spectrometry" J Chromatogr B Biomed Appl. 1995 Oct 6;672(1):17-23.
3. Marquez CD, Weintraub ST, Smith PC. "Quantitative analysis of two opioid peptides in plasma by liquid chromatography-electrospray ionization tandem mass spectrometry." J Chromatogr B Biomed Sci Appl. 1997 Jun 20;694(1):21-30.
4. Carrascal M, Schneider K, Calaf RE, van Leeuwen S, Canosa D, Gelpi E, Abian J. "Quantitative electrospray LC-MS and LC-MS/MS in biomedicine" J Pharm Biomed Anal. 1998 Sep 1;17(6-7):1129-38.
5. Zheng K, Lubman DM, Rossi DT, Nordblom GD, Barksdale CM. "Elucidation of peptide metabolism by on-line immunoaffinity liquid chromatography mass spectrometry". Rapid Commun Mass Spectrom. 2000;14(4):261-9.
6. Kippen AD, Cerini F, Vadas L, Stocklin R, Vu L, Offord RE, Rose K. "Development of an isotope dilution assay for precise determination of insulin, C-peptide, and proinsulin levels in non-diabetic and type II diabetic individuals with comparison to immunoassay." J Biol Chem. 1997 May 9;272(19):12513-22.
7. Hutchens TW, Yip TT. Rapid Commun Mass Spectrom 1993, 7, 576-580
8. Barr JR, Maggio VL, Patterson DG Jr, Cooper GR, Henderson LO, Turner WE, Smith SJ, Hannon WH, Needham LL, Sampson EJ."Isotope dilution-mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I." Clin Chem. 1996 Oct; 42(10):1676-82
9. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. "Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A. 2003 Jun 10;100(12):6940-5.
10. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." Nat Biotechnol. 1999 Oct;17(10):994-9

CLAIMS

1. A method for quantitative measurement using mass spectroscopy of a peptide or peptides generated from one or multiple analyte proteins by enzymatic or chemical cleavage, the method comprising the step of exposing together the sample to be analysed and one or multiple mass-labelled peptide internal standard precursors to enzymatic or chemical cleavage so as to generate from said precursor one or multiple mass-labelled peptide internal standards, wherein the mass-labelled peptide internal standard precursor comprises an extension of at least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom.

2. The method of claim 1 wherein the said N-terminal extension comprises at its C-terminal end an amino acid sequence of at least one, two, three, four or five amino acids identical to the amino acid sequence found in an analyte protein immediately N-terminal (prior to cleavage) to the sequence corresponding to the N-terminal-most mass-labelled peptide internal standard of the precursor.

3. The method of claim 1 wherein the said C-terminal extension comprises at its N-terminal end an amino acid sequence of at least one, two, three, four or five amino acids identical to the amino acid sequence found in an analyte protein immediately C-terminal (prior to cleavage) to the sequence corresponding to the C-terminal-most mass-labelled peptide internal standard of the precursor.

4. The method of any one of claims 1 to 3 wherein the amino acid sequence of at least one, two, three, four or five amino acids immediately N-terminal of a mass-labelled peptide internal standard of the precursor is identical to the amino acid sequence found in an analyte protein immediately N-terminal (prior to cleavage) to the sequence corresponding to that mass-labelled peptide internal standard.

5. The method of any one of claims 1 to 4 wherein the amino acid sequence of at least two, three, four or five amino acids immediately C-terminal of a mass-labelled peptide internal standard of the precursor is identical to the amino acid sequence found in the analyte protein immediately C-terminal (prior to cleavage) to the sequence corresponding to that mass-labelled peptide internal standard.

6. The method of any one of the preceding claims wherein each mass-labelled peptide internal standard has a length of between 4 and 50 amino acids.

7. The method of any one of claims 1 to 5 wherein each mass-labelled peptide internal standard has a length of between 10 and 40 amino acids.

8. The method of any one of the preceding claims wherein the mass-labelled peptide internal standard precursor has a length of between 6 and 200 amino acids.

9. The method of any one of the preceding claims wherein a mass-labelled peptide internal standard precursor comprises more than one mass-labelled peptide internal standard.

10. The method of claim 9 wherein a mass-labelled peptide internal standard precursor comprises mass-labelled peptide internal standards for analysis of peptide fragments originating from multiple analyte proteins.

11. The method of any one of the preceding claims wherein the mass-labelled peptide internal standard precursor co-purifies with the analyte protein or analyte proteins to be measured during sample preparation steps performed prior to exposure to enzymatic or chemical cleavage.

12. The method of any one of the preceding claims wherein the mass-labelled peptide internal standard precursor is combined with the sample to be analysed after sample collection but prior to any sample preparation or fractionation steps.

13. The method of any one of the preceding claims wherein the enzymatic or chemical cleavage is sequence-directed.

14. A method of any one of the preceding claims wherein the mass-labelled peptide internal standard precursor and sample comprising one or multiple analyte proteins are separated after enzymatic or chemical cleavage.

15. The method of claim 14 wherein said separation is effected using signature peptide capture.

16. A method of aiding diagnosing whether a patient has or may be susceptible to developing a disorder characterised by the increase or decrease of levels of one or more polypeptides comprising obtaining a test sample from said patient and quantifying the levels of said one or more polypeptides using a method defined in any of the previous claims.

17. A mass-labelled peptide internal standard precursor comprising an extension of at least one amino acid N-terminal to the N-terminus of the N-terminal-most mass-labelled peptide internal standard generated therefrom, and an extension of at least one amino acid C-terminal to the C-terminus of the C-terminal-most mass-labelled peptide internal standard generated therefrom, wherein on exposure to enzymatic or chemical cleavage said precursor generates one or multiple mass-labelled peptide internal standards.

18. A kit of parts comprising a mass-labelled peptide internal standard precursor according to claim 17 or as defined in any one of claims 1 to 16 and an agent for fragmenting peptides.

19. A kit of parts comprising a mass-labelled peptide internal standard precursor according to claim 17 or as defined in any one of claims 1 to 16 and a test sample containing or to be tested for the analyte protein or peptide.

Figure 1

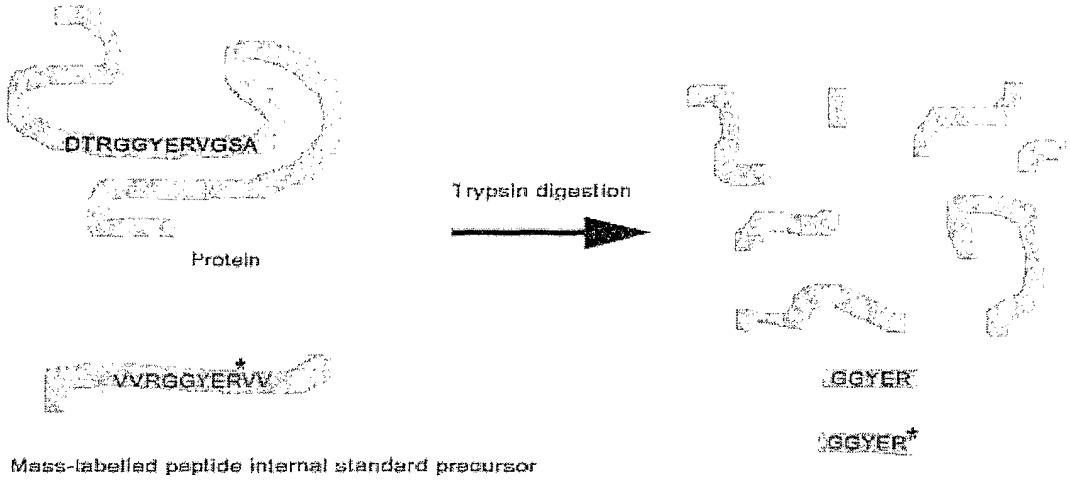

DTRGGYERVGSA
Protein
Trypsin digestion
VVRGGYERVV*
Mass-labelled peptide internal standard precursor
GGYER
GGYER*

Figure 2

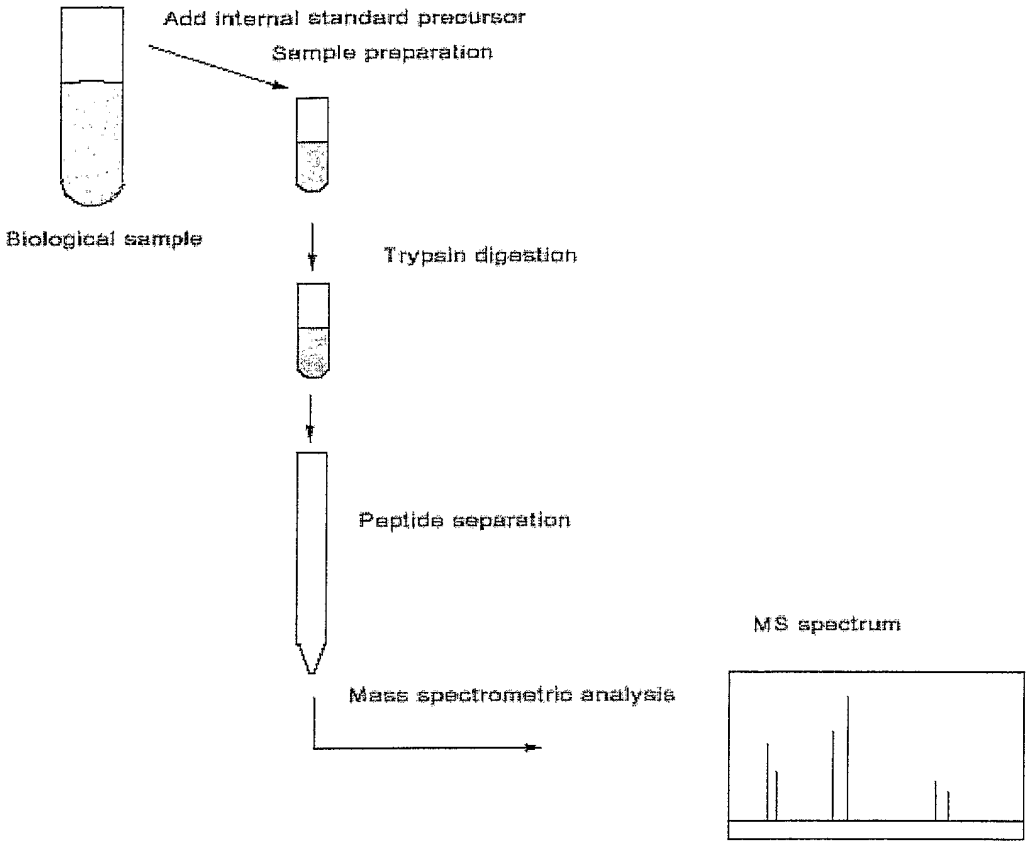

Add internal standard precursor
Sample preparation
Biological sample
Trypsin digestion
Peptide separation
MS spectrum
Mass spectrometric analysis

Figure 3