METHOD FOR QUANTIFYING AN ANALYTE IN A BIOLOGICAL SAMPLE USING MASS SPECTROMETRY

Abstract: This invention provides methods for quantifying the amount of analyte in a biological sample, e.g., a clinical sample. The methods comprise creating a calibration series using a complex milieu of proteins comprising a large number of standard proteins and the analyte of interest. The calibration series is used to generate a standard curve for quantification of the analytes in a test biological sample.
METHOD FOR QUANTIFYING AN ANALYTE IN A BIOLOGICAL
SAMPLE USING MASS SPECTROMETRY

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of U.S. provisional application no. 60/779,246, filed March 3, 2006, which is incorporated by reference herein.

FIELD OF THE INVENTION

[0002] This invention relates to the field of diagnostic assays.

BACKGROUND OF THE INVENTION

[0003] Assays for an analyte is a sample need to have an acceptable level of accuracy and precision for the particular need at hand. In the setting of clinical diagnostics, a test result informs the health care worker about the condition of a patient. Therefore, unacceptably high coefficients of variance (CVs) in an assay can render that assay useless for clinical application.

[0004] Mass spectrometry (MS) is rapidly becoming a method of choice for analysis of proteins, peptides and other biological molecules. Laser desorption/ionization mass spectrometry methods such as matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) and surface enhanced laser desorption/ionization mass spectrometry (SELDI-MS) procedures are very sensitive analytical methods and are probably the MS procedures most compatible with biological samples. Further, the ability of these procedures to generate high-mass ions at high efficiency from sub-picomole quantities of biological macromolecules makes these techniques extremely useful for macromolecule analysis and macromolecule identification in biological samples. Analysis of peptide analytes in crude biological samples, such as blood, plasma, or serum, however offers special problems for mass spectrometry analysis.

[0005] Laser desorption/ionization methods of mass spectrometry ("LDI-MS") for protein analytes in complex sample mixtures, such as serum or urine, pose particular problems for quantititation. In particular, peak intensity of an analyte can vary based on the nature of the protein milieu in which it is found, variances in sample handling and variances in instrument performance. One way to overcome these problems is to introduce normalization standards into a sample against which the analyte can be compared. Another method is to construct a
standard curve of the analyte on which the quantity of the analyte can be measured. However, standard curves can suffer if the calibration series is not matched with the sample being tested.

[0006] The ability of LDI-MS to quantitate an analyte within a biological sample can be difficult. The laser desorption process is a complex and poorly understood event that depends on the interaction of a laser ionization source and co-crystals consisting of both protein sample and desorption matrix. hi the laser desorption event, ionization of any one analyte cannot occur independent of any other analyte exposed at the same time to the ionization source. When two samples containing an analyte as well as a number of other proteins, where the other proteins are vastly different in either concentration or composition, are ionized the relative abundance of ions formed for the analyte in each of the samples may not be correlated with their concentration. Thus, the ability to quantity the analyte becomes problematic.

[0007] The methods of this invention provide a method of standardizing an assay using a calibration curve with absolute standards achieved by minimizing interference from non-analyte (i.e., background) proteins naturally associated with the biological sample while simultaneously providing a complex set of invariant internal standards. This results in a normalized analyte value that can be quantified directly from clinical or test samples by comparison to the calibration curve.

**BRIEF SUMMARY OF THE INVENTION**

[0008] The present invention relates to methods for quantifying analytes in a biological sample using mass spectrometry.

[0009] In one aspect, this invention provides a method for generating a standard curve for a protein analyte quantity in a laser desorption/ionization (LDI) mass spectrometry assay, the comprising: a) providing a protein calibration series having at least three members in the series, the members each comprising: (i) a protein analyte, and (ii) a complex milieu of proteins that are not naturally associated with the analyte comprising at least three reference standards, wherein the concentration of analyte protein varies in the members of the series and the concentration of the complex milieu of proteins does not vary; b) subjecting each member of the series to a chromatographic step to enrich for the protein analyte and the proteins in the complex milieu; c) generating an LDI spectrum of the relative mass-to-charge
signals for the analyte protein and the reference standards in the members of the analyte calibration series, wherein the spectra comprise a preponderance of signals from the enriched complex milieu; and d) normalizing the analyte protein to the reference standards present in the series, thereby generating a standard curve.

[0010] The chromatographic step can be an ion exchange chromatography, cation exchange or anion exchange.

[0011] In other embodiments, the chromatographic step is selected from size exclusion chromatography, hydrophobic interaction chromatography, metal chelate interaction or mixed mode interaction.

[0012] In some embodiments, step (b) is performed on a chip used in the LDI determination, in some embodiments, step (b) is performed prior to application of the sample to a chip for LDI determination. In some embodiments, the chromatographic step is a size exclusion chromatography.

[0013] The biological sample can be, e.g., serum or urine.

[0014] The complex milieu of proteins can be a cellular lysate. In some embodiments, the cellular lysate is a bacterial lysate. In some embodiments, the cellular lysate is a yeast lysate.

[0015] In some embodiments, the methods for generating a standard curve for a protein analyte quantity in a laser desorption/ionization mass spectrometry assay further comprises generating a standard curve for a second protein analyte present in the biological sample.

[0016] In some embodiments, the analyte is a cancer marker. In some embodiments, the cancer marker is transthyretin.

[0017] Another aspect of the invention provides a method for determining the amount of a protein analyte in a test sample using a laser desorption/ionization (LDI) mass spectrometry assay, the method comprising: a) providing a test biological sample comprising (i) at least one protein analyte and (ii) a complex milieu of proteins that are not naturally associated with the analyte, wherein the complex milieu of proteins comprises at least three reference standards; b) subjecting the sample to a chromatographic step; c) generating an LDI spectrum of the relative mass-to-charge signal for the analyte protein, wherein the spectrum comprise a preponderance of signal from the complex milieu; d) normalizing the analyte protein to the reference standards present in the test biological sample; and e) determining the
amount of analyte in the sample by comparing the normalized value to a standard curve of
claim created for the calibration series.

[0018] The chromatographic step can be an ion exchange chromatography. In some
embodiments, the ion exchange chromatography is anion exchange. In some embodiments,
the ion exchange chromatography is cation exchange.

[0019] In other embodiments, the chromatographic step is selected from size exclusion
chromatography, hydrophobic interaction, metal chelate interaction or mixed mode
interaction.

[0020] In some embodiments, step (b) is performed on the chip used in the LDI
determination. In some embodiments, step (b) is performed prior to application of the test
biological sample to the chip for LDI determination.

[0021] The test biological sample can be, e.g., serum or urine.

[0022] The complex milieu of proteins can be a cellular lysate. In some embodiments, the
cellular lysate is a bacterial lysate. In some embodiments, the cellular lysate is a yeast lysate.

[0023] In a further embodiment, the methods of this invention comprise measuring a
second analyte present in the test biological sample and determining the amount of the second
analyte in the sample by comparing the normalized value to a standard curve for the second
analyte. In some embodiments, the analyte is a cancer marker. In some embodiments, the
cancer marker is transthyretin.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**Figure 1.** Mass spectrometry read-out for individual samples in the calibration series. Cal A
is the highest concentration of sample (i.e. lowest dilution), Cal F is the lowest concentration
of the sample (i.e. highest dilution) in the dilution series. *E. coli* lysate complex milieu
standard peaks are labeled in the Cal F read-out. The intensity of the peak corresponding to
transthyretin (as labeled in Cal A) is decreasing across the dilution series.

**Figure 2.** Calibration Curve. Plot of peak intensity ratios (analyte transthyretin peak
divided by average of three *E. coli* standard peaks from the complex milieu) against analyte
concentration as determined for each individual sample by ELISA. The Calibration curve is
the best fit line.
DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

[0024] "Normalization standards" as used herein refers to a plurality of molecules, e.g., protein, carbohydrates, oligosaccharides, selected among the members of the complex milieu against which the analyte protein is normalized.

[0025] A "calibration series" as used herein refers to a series of at least three samples comprising the protein analyte or protein analytes of interest where the concentration of the protein analyte or protein analytes of interest are present in a known and differing quantity. A calibration series can be generated, for instance, by preparing a dilution series of a biological sample whereby the concentration of protein analyte(s) in at least one of the dilutions was previously known and can thus be used to derive the protein analyte(s) concentration in the remaining dilution series.

[0026] In the context of this invention a "standard curve" is a best fit curve through points corresponding to a plot of normalized intensity of the protein analyte against the concentration of the analyte over a range of calibration samples. The plotted calibration series encompasses the dynamic range of the analyte so that the protein analyte concentration in a test biological sample, e.g., a clinical sample for diagnostic analysis, can be extrapolated from a known normalized mass spectrometry peak intensity.

[0027] "Chromatography" as used in the present invention refers to a step utilized to enrich for the protein analyte of interest and all other proteins, either non-analyte naturally associated proteins or complex milieu proteins that share a similar biochemical /biophysical characteristic such as charge, hydrophobicity, hydrophilicity, size, etc., from a biological sample. Chromatographic materials include, for example, ion exchange materials (anion or cation exchange), metal chelators (e.g., nitroacetic acid or iminodiacetic acid), immobilized metal chelates, hydrophobic interaction adsorbents, hydrophilic interaction adsorbents, dyes, simple biomolecules (e.g., nucleotides, amino acids, simple sugars and fatty acids) and mixed mode adsorbents (e.g., hydrophobic attraction/electrostatic repulsion adsorbents).

[0028] "Laser desorption/ionization agent" is any active cause which enables an analyte to be laser desorped/ionized. An example of a laser desorption/ionization agent is the addition of a matrix." Any suitable matrix may be used however it is preferred that the matrix is selected from the group consisting of sinapinic acid (SPA), a-cyano-4-hydroxycinnamic acid (CHCA), 2,5- dihydroxybenzoic acid (2,5-DHB), 2-(4-hydroxy phenylazo) benzoic acid
(HABA), succinic acid, 2,6-Dihydroxyacetophenone, Feralic acid, caffeic acid, 2,4,6-
trihydroxyacetophenone (THAP) and 3-hydroxypicolinic acid (HPA), Anthranilic acid,
Nicotinic acid, Salicylamide and mixtures thereof.

[0029] "Mass spectrometric mixture" is a mixture comprising the analyte which can be
laser desorped/ionized by a mass spectrometer.

[0030] "Mass spectrometrically analyzed" is analyzing a mass spectorometric mixture with
a mass spectrometer resulting in a mass spectrometric response. Generally, a mass
spectrometer is an instrument designed to determine the mass-to-charge ratio of ions. The
process requires the analyte molecules to be volatilized, and ionized. The ionized molecules
are then accelerated by an electric field into an analyzing device, which separates the ions by
virtue of a property dependent on the mass-to-charge ratio (m/z). Among these properties
are: deflection in the magnetic field, velocity after acceleration through a fixed electric
potential drop, cyclotron orbital frequency in a magnetic field, and trajectory in a radio
frequency quadrupole field. Because polypeptides are generally large molecules and cannot
be volatilized under common mass spectrometry conditions, volatization and ionization of
polypeptides involves a need for assistance. One common procedure of assistance is known
as matrix-assisted laser desorption/ionization or MALDI. Another common method that
involves matrix assistance is surface enhanced laser desorption/ionization (SELDI).

[0031] "Normalize" or any form of the word in the context of this invention refers to using
internal reference standards as constant values to adjust analyte values in order to control for
assay variability.

II. Introduction

[0032] The present invention provides methods for quantifying the amount of analyte in a
biological sample, e.g., a clinical sample used for diagnostic detection of disease markers.
The methods involve supplementing a calibration series and unknown sample with a complex
milieu comprising a large number of proteins to make all calibrants and samples substantially
similar. Additionally, at least three of the complex milieu proteins can be used as reference
standards for normalization, and the analyte of interest. The calibration series is used to
generate a standard curve which can be used to determine the amount of analyte present in a
test sample by extrapolating from the standard curve. The methods provide for near identical
background between the calibration series and the test sample and for the measurement of the
amount of analyte polypeptide or polypeptides of interest in the biological sample tested. The proteins are detected using mass spectrometry. In a preferred embodiment, the proteins are detected using MALDI (matrix assisted laser desorption/ionization) or SELDI (surface enhanced laser desorption ionization) mass spectrometry.

[0033] A traditional calibration series includes a series of solutions containing the analyte of interest at a variety of concentrations. The calibration series also may have normalization markers. One method of creating a calibration series for an analyte is to make a dilution series of pure analyte in a buffer. However, when the test sample is a complex mixture of proteins such as is found in most biological samples, such a series will not do because the backgrounds of test sample, on the one hand, and the calibration series, on the other hand, are not compatible. hi LDI-MS, background proteins can have a large impact on the measured intensity of an analyte, for example, due to ion suppression. This adds a level of variability to the measurement process.

[0034] The methods of the invention are therefore useful in the quantification of target analyte proteins in the diagnosis of disease states such as cancer, e.g., prostate, breast, lung, bladder, ovarian, colon, brain and kidney; cancer metastasis; diabetes, both juvenile and late-onset; autoimmune disease such as rheumatoid arthritis and multiple sclerosis; heart disease, e.g., myocardial infarction, atherosclerosis, and cardiomyopathy; cerebrovascular disease, e.g., stroke; renal disease; lung disease, e.g., emphysema and where the calibrants and unknown samples may have different non-analyte associated protein compositions and relative concentrations.

[0035] In a partial solution to this problem, the dilution series can be made by diluting a sample of the same type as test sample, which contains the analyte. For example, if the test sample is serum, the calibration series can be made by diluting a serum sample known to contain the analyte. Typically one should use a sample for the dilution series that is known to contain more of the analyte than the test sample is likely to have. This makes it likely that the measurement of the analyte in the test sample will fall between points on the standard curve, and not require extrapolation. However, even in this case, the dilution of analyte in the calibration series with a buffer will result in members of the series having backgrounds that are very different than the test sample. Therefore, a complex milieu of proteins is added to each member of the dilution series to make the backgrounds more similar.
[0036] The calibration series of this invention overcomes at least two problems that can be encountered in developing a quantitative LDI-MS assay: The need for normalization markers and the need for similar background between the test sample and each member of the calibration series. These problems are solved by introducing into the test sample and into the calibration series equal amounts of a complex milieu of proteins not normally found with analyte, and that co-purify with the analyte in a selected fractionation process. For example, if the sample is derived from a mammal, such as a human serum sample, then the complex milieu can be derived from a non-mammalian or a non-animal source, for example a bacterial source such as E. coli. The complex milieu of proteins provides both a constant background between test samples and members of the calibration series, as well as proteins (e.g., at least three) that can be used for normalization against the analyte.

[0037] The calibration series is used to generate a standard curve which can be used to determine the amount of analyte present in a test sample by extrapolating from the standard curve. The proteins are fractionated using mass spectrometry, hi a preferred embodiment, the proteins are fractionated using SELDI (surface enhanced laser desorption ionization) mass spectrometry.

[0038] The methods of the invention are therefore useful in the identification and quantification of target analyte proteins in the diagnosis of disease states such as cancer, (e.g., prostate, breast, lung, bladder, ovarian, colon, brain and kidney); cardiovascular disease (e.g., acute coronary syndrome, myocardial infarction, atherosclerosis and cardiomyopathy), neurodegenerative diseases (e.g., Alzheimer's disease); infectious diseases (e.g., viral, bacterial and parasitic), women's health areas (e.g., premature labor, pre-eclampsia); and autoimmune diseases (e.g., rheumatoid arthritis and multiple sclerosis) and cerebrovascular diseases (e.g., stroke; renal disease; lung disease, e.g., emphysema).

III. Mass spectrometry
[0039] This invention pertains to methods of using mass spectrometry, a method that employs a mass spectrometer to detect gas phase ions to detect and quantify analytes that can be used as disease markers from biological samples. Examples of mass spectrometers are time-of-flight, magnetic sector, quadrupole filter, ion trap, ion cyclotron resonance, electrostatic sector analyzer and hybrids of these.
In a preferred method, the mass spectrometer is a laser desorption/ionization mass spectrometer, i.e., laser desorption/ionization mass spectrometry, the analytes are placed on the surface of a mass spectrometry probe, a device adapted to engage a probe interface of the mass spectrometer and to present an analyte to ionizing energy for ionization and introduction into a mass spectrometer. A laser desorption mass spectrometer employs laser energy, typically from an ultraviolet laser, but also from an infrared laser, to desorb analytes from a surface, to volatilize and ionize them and make them available to the ion optics of the mass spectrometer. In one embodiment, the laser desorption/ionization mass spectrophotometric technique for use in the invention is "Matrix Assisted Laser Desorption/ionization" or "MALDI".

In another embodiment, the laser desorption/ionization mass spectrometric technique for use in the invention is "Surface Enhanced Laser Desorption and Ionization" or "SELDI," as described, for example, in U.S. Patents No. 5,719,060 and No. 6,225,047, both to Hutchens and Yip. This refers to a method of desorption/ionization gas phase ion spectrometry (e.g., mass spectrometry) in which an analyte (here, one or more of the biomarkers) is captured on the surface of a SELDI mass spectrometry probe. There are several versions of SELDI.

One version of SELDI is called "affinity capture mass spectrometry." It also is called "Surface-Enhanced Affinity Capture" or "SEAC". This version involves the use of probes that have a material on the probe surface that captures analytes through a non-covalent affinity interaction (adsorption) between the material and the analyte. The material is variously called an "adsorbent," a "capture reagent," an "affinity reagent" or a "binding moiety." Such probes can be referred to as "affinity capture probes" and as having an "adsorbent surface." For the purposes of the present invention, the capture reagent can be any material capable of binding an analyte while enriching for the component proteins of the complex milieu. The capture reagent may be attached directly to the substrate of the selective surface, or the substrate may have a reactive surface that carries a reactive moiety that is capable of binding the capture reagent, e.g., through a reaction forming a covalent or coordinate covalent bond. Epoxide and carbodiimidizole are useful reactive moieties to covalently bind polypeptide capture reagents such as antibodies or cellular receptors. Nitriloacetic acid and iminodiacetic acid are useful reactive moieties that function as chelating agents to bind metal ions that interact non-covalently with histidine containing
peptides. Adsorbents are generally classified as chromatographic adsorbents and biospecific adsorbents.

[0043] Another popular method of analyte ionization for mass spectrometry is electrospray (e.g., LC-MS).

IV. Analytes and Biological Samples

[0044] An analyte protein is a protein in a mixture of proteins, which analyte protein is the target of measurement in an assay. Typically, the analyte protein will be found in a biological sample.

[0045] The biological sample typically contains non-analyte, naturally associated proteins, which are protein components in the sample which are not the target of measurement in an assay. A biological sample can be obtained from any organism including viruses, prokaryotes or eukaryotes including animals, plants and fungi, and is preferably mammalian, for example, a human.

[0046] The analytes analyzed in accordance with the methods described herein are present in biological samples such as amniotic fluid, blood, cerebrospinal fluid, intraarticular fluid, intraocular fluid, lymphatic fluid, milk, perspiration plasma, saliva semen, seminal plasma, serum, sputum, synovial fluid, tears, umbilical cord fluid, urine, biopsy homogenate, cell culture fluid, cell extracts, cell homogenate, conditioned medium, fermentation broth, tissue homogenate and derivatives of these.

[0047] The analyte or analytes of the present invention are typically associated with disease and can be used as diagnostic markers. The levels of a biomarker in a biological sample decrease or increase in concentration in response to or as a result of disease or a propensity for a disease. For example, the analyte of interest can be a cancer biomarker. In one embodiment, the analyte is transthyretin. Transthyretin levels are abundant in normal serum, but are decreased in the serum of ovarian cancer patients. In order to be useful in a clinical setting, an assay for transthyretin quantification has to have a low coefficient of variance. In another embodiment, the analyte is Apo Al. In another embodiment the analyte is a pure protein spiked into a complex mixture of other molecules such as biomolecules, for example, proteins, carbohydrates or nucleic acids, or polyethylene glycol.

[0048] The biological sample in which the analyte is to be measured typically is a complex sample that contains an abundance of protein components, e.g., serum, plasma or urine.
Often, the biological material is serum or derived from serum. The biological sample is selected based on the analyte or analytes being measured and evaluated in the assay.

The biological sample can be used in the assays essentially in the state as it is obtained from the subject where it is subjected to little or no processing. In other embodiments, the sample can be processed prior to its use in preparing the calibration series (or the test series). For example, a preliminary extraction prior to preparing the calibration or test sample can be performed to remove at least a portion of the naturally occurring substances present in the biological sample.

**V. Calibration series—Preparation of biological samples and addition of standards**

**A. Sample dilutions**

A dilution series of the analyte in buffer is created to vary the amounts of the analyte polypeptide or polypeptides of interest across the dilution series. The polypeptide content of both analyte and non-analyte proteins at the highest dilutions (i.e. lowest concentration) in the dilution series is much lower than that of the test sample. The dilution series preferably is commenced at the upper limit of analyte detection and continued until the lower limit of detection is reached. In a preferred embodiment, the dilution series is created from a sample of the same type as the test sample (e.g., serum). However, a dilution series of pure analyte protein also can be prepared.

When the calibration series is made by diluting a sample, rather than spiking samples with the analyte, the buffer used to create the dilution series should be compatible with both the sample and the subsequent fractionation method.

The concentration of analyte in the calibration series should reflect the dynamic range of the detection method. Classically, this is about three orders of magnitude. However, it also could be about two orders of magnitude or one order of magnitude.

The level of the analyte or analytes of interest in the biological sample that is used for generating the calibration series is known, such that a standard curve can be generated. Methods of determine the level of analyte protein are well known in the art. These include, e.g., immunoassays, such as ELISA.

Non-analyte, naturally associated proteins generally create background noise in mass spectra. In the present invention, when the calibration series is prepared from a
complex biological sample, non-analyte, naturally associated proteins are diluted out in a
dilution step and/or relatively diminished in a subsequent chromatography step.

B. Complex milieu

[0055] The complex milieu is a diverse molecular mixture, preferably a protein mixture. The composition of the complex milieu is determined empirically and depends on the set of proteins that will be presented to the ionization source. A plurality of the proteins in the complex milieu must co-purify with the analyte and, therefore, must have similar physico-chemical properties as the analyte. Also, these co-purifying proteins should have sufficiently different masses than the analyte so that they can be resolved by the mass spectrometer and not cause ion suppression of the analyte during mass spectrometry. The proteins from the complex milieu presented to the ionization source preferably are in amounts so that the peak intensities are between the lowest and highest concentration of analytes test samples. The proteins presented to the ionization source need to include at least three proteins from the complex milieu that are resolved from the endogenous non-analyte and analyte proteins so that they can be used as normalization factors. Useful sources of the complex milieu are biological fluids such as serum, urine, cell lysates, tissue lysates, cell culture medium and digests of a single protein or a mixture of proteins. Biological fluids typically have hundreds or thousands of different protein species. The complex milieu also can be created by mixing diverse proteins "off-the-shelf" as long as the mixture meets the aforementioned requirements. It is also contemplated that the molecules of the complex milieu can comprise other biological molecules, such as nucleic acids and carbohydrates, as well as polymeric materials such as polyethylene glycol. The complex milieu could be a library of polypeptides or of small molecules compatible with the size of the analyte.

[0056] The complex milieu, e.g., cellular lysate, is added to each member of dilution series and to each test sample in equal amounts. Use of a complex milieu provides at least two advantages. Firstly, using the complex milieu allows for making the samples across the dilutions and among the calibration series and test/clinical samples as similar in background as possible. Secondly, members of the complex milieu, e.g. polypeptide components present in the bacterial or yeast lysates, are used as internal reference standards to normalize the analyte peak intensity resulting from the mass spectrum, thus providing an absolute standard of reference. At least three, often at least ten or more, reference standards are selected from the complex milieu. The present invention thus decreases, more preferably eliminates, variability among the sample analysis.
[0057] As can be appreciated by one of skill in the art, the amount of complex milieu added to the samples is measured carefully and is kept constant in the calibration dilution series and test samples analyzed.

[0058] The complex milieu is chosen such, as compared with proteins in the test sample, the components of the complex milieu are more likely to co-purify with the analyte during a subsequent fractionation step than proteins naturally occurring in the test sample. Typically these components have similar biochemical characteristics as the analyte protein (e.g. pi, hydrophobic index, metal chelate binding ability). This allows for enrichment of complex milieu proteins, relative to the naturally occurring non-analyte proteins present in the biological sample, in a chromatography step that also selects for the analyte polypeptide or polypeptides.

[0059] The complex milieu is also chosen such that the member proteins of the complex milieu do not interfere with the mass spectrophotometric peak of the analyte polypeptide or polypeptides (i.e. the analyte polypeptide or polypeptides and the members of the complex milieu do not have the same mass-to-charge ratio). Alternatively, the complex milieu can be pre-processed to remove the proteins of interfering masses by methods known in the art, e.g., chromatography such as size-exclusion chromatography.

[0060] A complex milieu can be from any source that will provide a mixture of multiple, usually at least 10 or more, proteins. Cellular lysates are often used. Examples of useful bacterial lysates that can serve as a complex milieu include but are not limited to *Escherichia, Enterobacter, Azotobacter, Erwinia, Klebsiella, Bacillus, Pseudomonas, Proteus* and *Salmonella*. Examples of useful yeast lysates that can serve as a complex milieu include but are not limited to *Saccharomyces cerevisiae, Candida albicans, Schizosaccharomyces pombe*. Cellular lysates from other eukaryotic cells may also be used. As noted above, the complex milieu is selected so that interfering peaks will not be present when the sample is analyzed. As appreciated by one in the art, the complex milieu.

[0061] Reference standards for use in normalizing the sample are selected from the members of the complex milieu. The reference standards are protein, peptide or polypeptide components. The complex milieu is added to the dilution series and test samples, so as to maintain the concentration of the proteins in the reference series invariant across the series. Reference standards should be easily identified as unique from the complex milieu, with a
signal-to-noise ratio greater than 3 and preferably greater than 5 and not be interfered by the protein analyte or non-analyte associated proteins.

[0062] In some embodiments additional purified proteins of known concentration can be added to the complex milieu of proteins to provide additional internal standards. Preferably, these proteins are not present in the test sample.

VI. Chromatographic processing of the calibration series

[0063] Following the preparation of the diluted biological sample which contains the analyte or analytes of interest and addition of the complex milieu across the sample series, fractionation, in particular chromatography, typically is performed. The samples can be subjected to ion exchange chromatography (anion, cation), affinity chromatography, or any other type of chromatography known in the art.

[0064] The chromatography step enriches for the protein analyte and the polypeptide components of the complex milieu that share a common biochemical/biophysical characteristic. The type of chromatography is chosen such that the chemical properties of the analyte polypeptide or polypeptides are very similar to those of the complex milieu components being enriched. In one embodiment the analyte polypeptide or polypeptides are negatively charged and E. coli lysate is used as a complex milieu. In this embodiment anion chromatography enriches for the negatively charged E. coli lysate component proteins, peptides and polypeptides and the analyte polypeptide or polypeptides in the analysis sample. It is apparent to those of skill in the art that other chemical properties such as for example, but not limited to size, charge or pi can be used to determine the best chromatography step and complex milieu to be combined with a particular analyte polypeptide or polypeptides.

[0065] Other types of chromatography can also be used. For example, if the analyte is positively charged it can be captured and complex milieu proteins enriched on a cation exchange material, such as CM Ceramic HyperD F resin.

[0066] In some embodiments of this invention, the chromatography step is performed directly on a SELDI chip. For example, the dilution series/test sample containing the complex milieu can be prepared, applied to an anion exchange MS chip, washed, and analyzed. For example, protein biochips produced by Ciphergen Biosystems, Inc. comprise surfaces having chromatographic or biospecific adsorbents attached thereto at addressable locations. Ciphergen ProteinChip® arrays include NP20 (hydrophilic); H4 and H50 (hydrophobic); SAX-2 and Q-10 (anion exchange); WCX-2 and CM-10 (cation exchange);

The other embodiments of this invention, the chromatography step is performed as a separate step on a chromatographic material (e.g. packed in a column and subjected to FPLC processing or chromatographic material commercially available in a fast column run by gravity or by centrifugation) followed by measurement by either MALDI or SELDI.

VII. Generating an LDI spectrum

Analysis of analytes by time-of-flight mass spectrometry generates a time-of-flight spectrum. The time-of-flight spectrum ultimately analyzed typically does not represent the signal from a single pulse of ionizing energy against a sample, but rather the sum of signals from a number of pulses. This reduces noise and increases dynamic range. This time-of-flight data is then subject to data processing. For instance, in Ciphergen's ProteinChip® software, data processing typically includes TOF-to-M/Z transformation to generate a mass spectrum, baseline subtraction to eliminate instrument offsets and high frequency noise filtering to reduce high frequency noise.
Preferably, the complex milieu is chosen so that after fractionation and analysis by LDI-MS, the spectrum generated comprises a preponderance of measurable peaks from the complex milieu, compared with measurable peaks from the sample used to create the dilution series. More preferably, measurable signals from proteins in the complex milieu comprise at least 75% and more preferably at least 90% of the total measurable peaks in the spectrum. In one embodiment, the sample presented to the ionization source produces a spectrum comprising between about 20 to 100 measurable signals from the complex milieu and about 10-20 signals from the biological or test sample.

The computer can transform the resulting data into various formats for display. The standard spectrum can be displayed, but in one useful format only the peak height and mass information are retained from the spectrum view, yielding a cleaner image and enabling biomarkers with nearly identical molecular weights to be more easily seen. In another useful format, two or more spectra are compared, conveniently highlighting unique biomarkers and biomarkers that are up- or down-regulated between samples. Using any of these formats, one can readily determine whether a particular biomarker is present in a sample.

Analysis generally involves the identification of peaks in the spectrum that represent signal from an analyte. Peak selection can be done visually, but software is available, e.g. as part of Ciphergen's ProteinChip® software package, that can automate the detection of peaks. In general, this software functions by identifying signals having a signal-to-noise ratio above a selected threshold and labeling the mass of the peak at the centroid of the peak signal. In one useful application, many spectra are compared to identify identical peaks present in some selected percentage of the mass spectra. One version of this software clusters all peaks appearing in the various spectra within a defined mass range, and assigns a mass (M/Z) to all the peaks that are near the mid-point of the mass (M/Z) cluster. For the purposes of the present invention, complex milieu peaks can be analyzed in this manner to provide the most invariable standards to be used for analyte peak normalization.

Software used to analyze the data can include code that applies an algorithm to the analysis of the signal to determine whether the signal represents a peak in a signal that corresponds to a biomarker according to the present invention. The software also can subject the data regarding observed biomarker peaks to classification tree or ANN analysis, to determine whether a biomarker peak or combination of biomarker peaks is present that indicates the status of the particular clinical parameter under examination. Analysis of the
data may be "keyed" to a variety of parameters that are obtained, either directly or indirectly, from the mass spectrometric analysis of the sample. These parameters include, but are not limited to, the presence or absence of one or more peaks, the shape of a peak or group of peaks, the height of one or more peaks, the log of the height of one or more peaks, and other arithmetic manipulations of peak height data.

VIII. Creating a standard curve.

A. Normalizing the analyte mass spectrometric peak

[0073] Data generated by desorption and detection of peptide peaks can be analyzed with the use of a programmable digital computer. The computer program analyzes the data to indicate the number of peptides detected, and optionally the strength of the signal and the determined molecular mass for each peptide detected. Data analysis can include steps of determining signal strength of a peptide and removing data deviating from a predetermined statistical distribution. The intensity of the analyte peak is normalized as a function of the intensity of at least one normalization peak. Preferably, the intensity of the analyte protein calibrant is normalized against a plurality of normalization proteins, e.g., at least three, preferably five, ten, or more reference standards peaks chosen from the complex milieu. The normalization function typically is a ratio of the analyte peak intensity to the sum of the normalization peak intensities.

B. Generating the standard curve.

[0074] After normalizing the calibrant peak intensity in each member of the calibration series, a calibration curve is created by plotting the normalized values against analyte concentration in each member of the series. (The concentration having been calculated based on the determination of the amount of analyte present in the biological sample used to prepare the dilutions for the calibration series).

[0075] The present invention internal reference standards comprise reference standards kept at a constant concentration. Another important aspect of normalization as described in the present invention involves minimizing background noise from the non-analyte components of the biological sample that may vary in relative concentration between samples. Furthermore, in the context of the present invention, normalizing further involves producing background peaks between the calibration series and the sample series that are generated from a constant concentration or as allowed by the limitations of the instrument.
IX. Analyzing the test sample - processing of the test sample and comparison to calibration series

[0076] The test sample refers to a biological sample containing an analyte or analytes of interest and from which the concentration of the analyte or analytes of interest is to be determined. Typically, the test sample is a clinical sample for diagnostic evaluation. Test samples are prepared similarly to the calibration series samples.

[0077] The test sample can be measured undiluted (e.g., when the concentration of the analyte is within the dynamic range of the standard curve) or can be diluted into a series. As appreciated by those of skill in the art, the buffer used for diluting the biological sample in the calibration series is typically also utilized for diluting the test series.

[0078] The test sample or samples can be subjected to the methods of this invention either at the same time as the calibration series or a previously generated calibration series can be utilized to determine the analyte or analytes concentration in the test sample. However, it is important that the test sample and the calibration series members undergo the same processing steps for analyte measurement, beginning with addition of the same amount of complex milieu proteins to each of the test samples and the calibration series members.

[0079] The test or diagnostic sample preferably is the same biological sample as used in the calibration series (e.g., test series biological sample is serum and the calibration series biological sample is serum). However, they could be different (e.g., the test sample is serum and the calibration series is a pure protein diluted in buffer).

[0080] In some embodiments, the test sample is essentially unprocessed. In other embodiments, the biological sample may be subjected to a preliminary processing step, e.g., an extraction step, prior to adding the complex milieu.

[0081] The same amount of the complex milieu used in creating the calibration series is added to each test sample. The importance of keeping the complex milieu constant across the series and among the compared series (e.g. calibration and test series) can be appreciated by one of the skill in the art.

[0082] The test samples are subjected to the same chromatography step or series of chromatographic steps used in the creation of the calibration series in order to enrich for the proteins in the complex milieu and generate a background that is similar to that of the calibration series.
The samples are analyzed by LDI-MS as described for the samples of the calibration series. The peak intensity corresponding to the analyte of interest is normalized to the reference standards (at least 3, typically 5, 10, or more) in the complex milieu that were used in generating the standard curve.

The calibration curve that was created based on the calibration series can them be used to extrapolate the concentration of an analyte polypeptide or polypeptides in the test sample. It is apparent to those of skill in the art that in order to obtain the serum concentration of the analyte polypeptide, the value has to be corrected for the dilution for which the mass spectrum was produced.

**EXAMPLES**

**Example 1. E. coli lysate preparation.**

This example illustrates the preparation of a complex milieu derived from *E. coli* lysates.

**Reagents**

LB broth from Gibco (Cat # 10855-021)

lysis buffer (IxPBS, 1% Triton). See appendix for buffer preparation.

Ultra BL21(DE3) competent cells from Edge Biosystems (Cat # 45363)

Coomassie plus™ protein assay Kit from Pierce (23236).

**Preparing the BL12 Cell Lysate**

BL21 cells are first seeded from a 50ul glycerol stock in 5 ml of LB in a 15ml tube and incubated with shaking at 37°C for 6 hrs. The culture is then diluted into 500 ml of warm LB and incubated at 37°C with shaking at 200 rpm for overnight. The culture is then centrifuged in Sorvoll Rotor (250ml bottles) at 5,500 rpm for 10min, the supernatant is discarded.

To lyse the cells, the pellet is resuspended in 10ml of cold lysis buffer (IxPBS, 1% Triton) followed by freezing the lysate at -80°C. The lysate is thawed and sonicated on ice 5 times for 30 seconds at medium intensity. Upon spinning the mixture at 8,000 RPM for 20 min, the supernatant is removed fresh tubes and the pellet is again resuspended in 10ml of lysis buffer and lysated again to extract more lysate. The second obtained supernatant is then mixed with the first obtained supernatant and the complex milieu mixture is aliquoted, and
stored at -80 °C. Protein concentration was measured by coomassie plus™ protein assay Kit from pierce.

**Evaluation of *E. coli* lysate for use in the transthyretin assay.**

1. Presence of 10 *E. coli* peaks for TT on QIO (for TT: 9127, 9246, 11003, 12024, 12180, 15376, 16148, 17473, 18441, 19553 approx) and the ratio with respect to TT.
2. Presence of 6 *E. coli* peaks for ApoAl in IMAC30 (for ApoAl: 7244, 7838, 9497, 11135, 12600, 18074 approx), and then the ratio with respect to ApoAl.
3. Protein concentration.
4. Make three batches and observe the variability.

**Example 2. Identification of transthyretin.**

[0088] This is an example of creating a standard curve for the transthyretin analyte from serum using *E. coli* lysate as the complex milieu.

Preparation of TT sample buffer containing *E. coli*

[0089] The TT sample buffer was prepared as follows: 2.4ml E.Coli stock solution (prepared as described above and determined to have a protein concentration of 6.5mg/ml) was mixed with 153.5ml 0.1M NaP pH7 buffer.

Preparation of calibration series.

The calibration series was prepared by diluting the starting serum sample as follows:

1. 48µl Intergen human serum + 5952µl TT sample buffer (0.1M NaP pH7 + E. Coli) = cal A (51.4 mg/dl).
2. 2ml cal A + 1ml TT sample buffer = cal B (34.3 mg/dl).
3. 2ml cal A + 2ml TT sample buffer = cal C (25.7 mg/dl).
4. 2ml cal C + 2ml TT sample buffer = cal D (12.85mg/dl).
5. 2ml cal D + 2ml TT sample buffer = cal E (6.425mg/dl).
6. TT sample buffer only = cal F (Omg/dl).
7. Add 100µl to each well of sample v-plate.

[0090] The test samples (clinical samples) were prepared by diluting the starting serum as follows:

5µl serum + 1245µl TT sample buffer (1:250 dilution).
24µl Intergen serum + 5976µl sample buffer (1:250 dilution).
Add 100µl to each well of sample v-plate.
Transthyretin Assay Protocol Performed on QlO SELDI Chips.

1. Equilibrate with 150µl QlO wash buffer (0.1M Na phosphate pH7 only, no BSA) two times on TECAN. Spin bioprocessors at 800rpm for 1min after the first addition only. Stand for 10min for each equilibration. Remove buffer.
2. Add 50µl samples/standards to each well. Spin bioprocessors at 800rpm for 1min.
3. Seal bioprocessors with tape. Incubate on Micromix (15,3,120) at room temperature for 2 hours. Remove samples.
4. Wash with 150µl QlO wash buffer 4 times, pumping up and down 10 times each.
5. Wash with 150µl water 1 time, pumping up and down 10 times each.
6. Remove bioprocessor top, flick chip cassette by hand to remove remaining water.
7. Air dry chips for 30min.
8. Add 0.75 µl Sinapinic acid solution two times using Biodot.
9. Air dry for 10min after first addition in Biodot chamber.
10. Air dry for 30min after second addition inside Biodot chamber.
11. 1 tube of 5mg sinapinic acid + 200 µl acetonitrile + 200 µl 1% TFA, mix to dissolve.
12. Read on PCS4000, focus at 14KDa, 2-200KDa mass range, collect 10 shots at A partitions and a total of 530 shots.

Comparison using BSA as a standard.

[0091] BSA was used as a reference standard to illustrate the advantages of using a complex milieu to provide reference standards.

Preparation of Sample Buffer (SB):
125mg of BSA was added to 500ml 0.1M NaP pH7 buffer.

Preparation of calibration series:
1. 48µl Intergen human serum + 5952µl SB (0.1M NaP pH7 + 0.25mg/ml BSA) = cal A (51.4 mg/dl).
2. 2ml cal A + 1ml SB = cal B (34.3 mg/dl).
3. 2ml cal A + 2ml SB = cal C (25.7 mg/dl).
4. 2ml cal C + 2ml SB = cal D (12.85mg/dl).
5. 2ml cal D + 2ml SB = cal E (6.425mg/dl).
6. SB sample buffer only = cal F (0mg/dl).
7. Add 100µl to each well of sample v-plate.

*Preparation of test/clinical samples:*

5µl serum + 1245µl SB (1:250 dilution).

24µl Inergen serum + 5976µl SB (1:250 dilution).

Add 100µl to each well of sample v-plate.

[0092] The above examples are provided to illustrate the invention but not to limit its scope. Other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent application cited herein are hereby incorporated by reference for all purposes.
WHAT IS CLAIMED IS:

1. A method for generating a standard curve for a protein analyte quantity in a laser desorption/ionization (LDI) mass spectrometry assay, the comprising:
   a) providing a protein calibration series having at least three members in the series, the members each comprising: (i) a protein analyte, and (ii) a complex milieu of proteins that are not naturally associated with the analyte comprising at least three reference standards, wherein the concentration of analyte protein varies in the members of the series and the concentration of the complex milieu of proteins does not vary;
   b) subjecting each member of the series to a chromatographic step to enrich for the protein analyte and the proteins in the complex milieu;
   c) generating an LDI spectrum of the relative mass-to-charge signals for the analyte protein and the reference standards in the members of the analyte calibration series, wherein the spectra comprise a preponderance of signals from the enriched complex milieu;
   d) normalizing the analyte protein to the reference standards present in the series, thereby generating a standard curve.

2. The method of claim 1, wherein the chromatographic step is an ion exchange chromatography.

3. The method of claim 2, wherein the ion exchange chromatography is anion exchange.

4. The method of claim 2, wherein the ion exchange chromatography is cation exchange.

5. The method of claim 1, wherein the chromatographic step is selected from hydrophobic interaction, metal chelate interaction or mixed mode interaction.

6. The method of claim 1, wherein (b) is performed on a chip used in the LDI determination.

7. The method of claim 1, wherein (b) is performed prior to application of the sample to a chip for LDI determination.
8. The method of claim 1, wherein the chromatographic step is a size exclusion chromatography.

9. The method of claim 1, wherein the biological sample is serum.

10. The method of claim 1, wherein the biological sample is urine.

11. The method of claim 1, wherein the complex milieu of proteins is a cellular lysate.

12. The method of claim 11, wherein the cellular lysate is a bacterial lysate.

13. The method of claim 11, wherein the cellular lysate is a yeast lysate.

14. The method of claim 1, further comprising generating a standard curve for a second protein analyte present in the biological sample.

15. The method of claim 1, wherein the analyte is a cancer marker.

16. The method of claim 15, wherein the cancer marker is transthyretin.

17. A method for determining the amount of a protein analyte in a test sample using a laser desorption/ionization (LDI) mass spectrometry assay, the method comprising:

   a) providing a test biological sample comprising (i) at least one protein analyte and (ii) a complex milieu of proteins that are not naturally associated with the analyte, wherein the complex milieu of proteins comprises at least three reference standards;
   b) subjecting the sample to a chromatographic step;
   c) generating an LDI spectrum of the relative mass-to-charge signal for the analyte protein, wherein the spectrum comprise a preponderance of signal from the complex milieu;
   d) normalizing the analyte protein to the reference standards present in the test biological sample; and
   e) determining the amount of analyte in the sample by comparing the normalized value to a standard curve of claim 1.
18. The method of claim 17, wherein the chromatographic step is an ion exchange chromatography.

19. The method of claim 18, wherein the ion exchange chromatography is anion exchange.

20. The method of claim 18, wherein the ion exchange chromatography is cation exchange.

21. The method of claim 17, wherein the chromatographic step is selected from hydrophobic interaction, metal chelate interaction or mixed mode interaction.

22. The method of claim 17, wherein (b) is performed on the chip used in the LDI determination.

23. The method of claim 17, wherein (b) is performed prior to application of the test biological sample to the chip for LDI determination.

24. The method of claim 17, wherein the chromatographic step is a size exclusion chromatography.

25. The method of claim 17, wherein the test biological sample is serum.

26. The method of claim 17, wherein the test biological sample is urine.

27. The method of claim 17, wherein the complex milieu of proteins is a cellular lysate.

28. The method of claim 27, wherein the cellular lysate is a bacterial lysate.

29. The method of claim 27, wherein the cellular lysate is a yeast lysate.

30. The method of claim 17, further comprising measuring a second analyte present in the test biological sample and determining the amount of the second analyte in the sample by comparing the normalized value to a standard curve for the second analyte.
31. The method of claim 17, wherein the analyte is a cancer marker.

32. The method of claim 31, wherein the cancer marker is transthyretin.
The graph shows a linear relationship between Transthyretin Concentration (mg/dL) and Transthyretin Peak Intensity (Average E coli calibrant peak intensities). The equation of the line is given as:

\[ y = 0.1162x - 0.2457 \]

with an R² of 0.9919.