Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2782209 A1 2011/06/03

(21) 2 782 209

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2010/10/29
(87) Date publication PCT/PCT Publication Date: 2011/06/03
(85) Entrée phase nationale/National Entry: 2012/05/29
(86) N° demande PCT/PCT Application No.: EP 2010/066508
(87) N° publication PCT/PCT Publication No.: 2011/064068
(30) Priorité/Priority: 2009/11/30 (US61/265,029)

(51) Cl.Int./Int.Cl. *H01J 49/00* (2006.01)

(71) Demandeur/Applicant:
PHYSIKRON SA, FR

(72) Inventeur/Inventor:
SCIGOCKI, DAVID, FR

(74) Agent: NORTON ROSE CANADA S.E.N.C.R.L.,S.R.L./LLP

(54) Titre : PROCEDE DE SPECTROMETRIE DE MASSE EN TANDEM MULTIPLEXEE
(54) Title: MULTIPLEXED TANDEM MASS SPECTROMETRY METHOD

(57) Abrégé/Abstract:
The invention concerns a method for multiplexed tandem mass spectrometry of a sample to be analysed containing at least two precursors, wherein at least two simplified multiplexed MS-MS spectra are obtained each from at least two selected precursors of the sample, the method comprising: (d) for each selected precursor generating an individual MS-MS spectrum from the simplified multiplexed MS-MS spectrum by selecting fragment ions of the simplified multiplexed MS-MS spectrum, the fragment ions are potential fragment ions obtained from the precursor; (e) submitting each individual MS-MS spectrum of step (d) to a real and a decoy database searches using a scoring process without score threshold condition or low score threshold condition for identifying candidate precursors and their fragment ions; (f) producing real individual MS-MS spectra from identified candidate precursors resulting from the real database search of step (e); and producing decoy individual MS-MS spectra from identified candidate precursors resulting from the decoy database search of step (e); (g) submitting the real and decoy individual MS-MS spectra to a further scoring process with a score threshold condition for determining a score for each real and decoy individual MS-MS spectra.

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 June 2011 (03.06.2011)

PCT

(10) International Publication Number
WO 2011/064068 A1

(51) International Patent Classification:
H01J 49/00 (2006.01)

(21) International Application Number:
PCT/EP2010/066508

(22) International Filing Date:
29 October 2010 (29.10.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/265,029 30 November 2009 (30.11.2009) US

(71) Applicant *(for all designated States except US)*: **PHYSIKRON** [FR/FR]; 278, avenue du Maréchal Juin, F-92100 Boulogne-Billancourt (FR).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **SCIGOCKI, David** [FR/FR]; 20, avenue Séverine, F-06460 Saint Vallier de Thiey (FR).

(74) Agent: **LE FORESTIER, Eric**; Cabinet Regimbeau, 20, rue de Chazelles, F-75847 Paris Cedex 17 (FR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

— *of inventorship (Rule 4.17(iv))*

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

(54) Title: MULTIPLEXED TANDEM MASS SPECTROMETRY METHOD

(57) Abstract: The invention concerns a method for multiplexed tandem mass spectrometry of a sample to be analysed containing at least two precursors, wherein at least two simplified multiplexed MS-MS spectra are obtained each from at least two selected precursors of the sample, the method comprising: (d) for each selected precursor generating an individual MS-MS spectrum from the simplified multiplexed MS-MS spectrum by selecting fragment ions of the simplified multiplexed MS-MS spectrum, the fragment ions are potential fragment ions obtained from the precursor; (e) submitting each individual MS-MS spectrum of step (d) to a real and a decoy database searches using a scoring process without score threshold condition or low score threshold condition for identifying candidate precursors and their fragment ions; (f) producing real individual MS-MS spectra from identified candidate precursors resulting from the real database search of step (e); and producing decoy individual MS-MS spectra from identified candidate precursors resulting from the decoy database search of step (e); (g) submitting the real and decoy individual MS-MS spectra to a further scoring process with a score threshold condition for determining a score for each real and decoy individual MS-MS spectra.

WO 2011/064068 A1

Title

Multiplexed tandem mass spectrometry method

Field of the invention

The invention relates to the general field of mass spectrometry.

State of the art

By way of a reminder, mass spectrometry (MS), whatever its type, generally includes steps used to analyze the molecules present in a sample by measuring the mass of these molecules after they have been ionised in an ion source, accelerated and injected into a mass spectrometer.

A mass spectrometer generates a mass spectrum of the various molecules contained in the analysed sample, as a function of the mass–to–charge ratio (m/z) value of the generated ions.

In particular, tandem mass spectrometry (MS-MS) is well known as powerful tool for identifying and characterising molecules, and is generally used when the primary mass spectrum does not allow the identification of the generated ions.

Tandem mass spectrometers are generally composed of two mass spectrometers operating sequentially in space and separated by a dissociation device, or a single mass analyzer operating sequentially in time.

It generally includes steps required to generate, by means of the first mass spectrometer, a primary mass spectrum (MS) of the ionized molecules (called precursor ions) present in the sample to analyse, to perform a step for the selection of a precursor mass in the primary (MS) mass spectrum, for example via a mass selection window, and then to fragment, i.e. to dissociate by means of a dissociation device, the precursor ions of said selected precursor mass, so as to generate a mass spectrum described as the dissociation (MS-MS) mass spectrum of the fragments ions generated by the dissociation of the precursor ions, by means of the second mass spectrometer.

These steps are repeated for each selected precursor mass of the primary MS spectrum generating as many MS-MS spectra as selected precursor masses.

The precursor mass selection, generally implemented to generate each MS-MS spectrum, limits the acquisition debit of the tandem mass spectrometer, as the MS-MS spectra are generated one after the other.

It also significantly increases the amount of samples used to generate the MS-MS spectra compared with MS spectrum production, the remaining unselected precursor ions provided by the ion source being actually eliminated for the generation of the MS-MS spectrum of the selected primary ions.

Besides this first limitation in throughput due to the successive precursor mass selections, a second limitation is the possible selection of more than one precursor per mass selection window for producing each MS-MS spectrum.

This inadvertently multiplexed mass selection is due to the width of the mass selection window used to produce the MS-MS spectra which. The mass selection window is broader than the resolution of the mass spectrometer, especially for high resolution mass spectrometers. The width of the mass selection window is broader compared with MS resolution because of the MS ion selection devices used for the precursor mass selection in tandem mass spectrometers.

The fragment ions of the plurality of selected precursor ions increase the complexity of the produced MS-MS spectrum, and generally decrease the identification efficiency of the precursor that was aimed at by the precursor mass selection by the analysis of the MS-MS spectrum.

Simplified MS and MS-MS spectra are, thus, commonly produced from the peaks by different techniques such as deisotoping, de-charging, calibration, etc., in which the MS and MS-MS mass spectra are used for the final analysis leading to the identification of the precursors.

The simplified MS and MS-MS spectra are generally a list of mass-to-charge ratio m/z values and corresponding maximum intensity values corresponding to the peaks of the MS and MS-MS spectra. Ion charges are also used when they are determined.

The above limitations of standard tandem mass spectrometry are especially serious in protein analysis (proteomics) of complex mixtures of peptides obtained from digested proteins ("Bottom up"' proteomics), using liquid chromatography (LC) coupled with tandem mass spectrometers (LC-MS-MS) with, for example, Electrospray Ionisation ion sources (ESI ion sources).

In "Bottom up" proteomics, the mixture of proteins to be analysed is cleaned and digested with cleavage reagents such as trypsin, cyanogen bromide, or the like, to produce peptides before the LC separation.

This approach involves the LC separation of the peptides, and for each LC peak, the production of the primary MS spectrum of the peptides after their ionization (precursor ions) followed by the dissociation of selected peptides and the production of their MS-MS spectra with the tandem mass spectrometer, and the identification by protein sequence database searching of the selected peptides (and their parent proteins) with the produced MS-MS spectra.

During an LC-MS-MS acquisition with a sample containing a small number of proteins, each peptide (precursor) in the MS spectrum can be selected to produce a corresponding MS-MS spectrum.

But in complex protein sample analysis, the MS-MS throughput of the LC-MS-MS method is clearly limited by the time needed to successively acquire the MS-MS spectrum of each selected precursor of the MS spectrum, within the limited elution duration of the LC peak, which typically lies between 1 to 30 seconds.

Therefore, only a portion of the peptides (and parent proteins) can be identified during the LC-MS-MS analysis of a complex mixture of proteins.

The most common approach used to select the limited number of precursors to produce the corresponding MS-MS spectra after each LC peak is the “data dependant” analysis in which the most intense MS peaks of MS spectrum are automatically first selected for MS-MS.

Generally, the database search is carried out by using the simplified MS and MS-MS spectra described above. The database search can also be performed with a pre-treatment of the MS-MS spectra such as a “sequence tagging” in which only small parts of the amino acid sequences (“Tags”) are produced or with ”De Novo Sequencing” in which the complete amino acid sequence is directly calculated from the MS-MS spectra.

The database search is commonly performed by automatic computer search using scoring methods such as with Mascot or Sequest search tool, or the like.

Many protein databases such as Swissprot, NCBInr, MSDB, or the like, can be used for the automatic computer search.

During the database search, the proteins of the database are electronically digested (“In silico digestion”) with the same cleavage reagent used by the user for the LC-MS-MS data production. A peptide list comprising peptides corresponding to each digested protein is produced. A sub-list of potential peptide candidates is selected for each experimental MS precursor selected during the LC-MS-MS data production, within the MS accuracy chosen by the user.

All the possible peptide fragmentation patterns of each potential peptide candidate are calculated to produce a corresponding theoretical MS-MS spectrum as function of the parameters chosen by the user for the LC-MS-MS analysis (MS, and MS-MS accuracy, fragmentation energy, tandem mass spectrometer used, type of fragment ions produced, etc.).

The fragment ions of each experimental MS-MS spectrum are then compared with the fragment ions of the theoretical MS-MS spectra.

A list of identified peptide candidates (and corresponding proteins) is generated with corresponding identification scores for each MS-MS

spectrum submitted to the database search. The highest score corresponds normally to the best candidate identification.

A final list of identified protein candidates combining all the identified peptides with the highest score identification (normally the best identified peptide candidate of each MS-MS spectrum) of the complete LC-MS-MS acquisition of the analyzed sample is produced, after the selection of a peptide score threshold by the user.

The final list of peptide candidates (and corresponding proteins) comprises positive identifications with scores above score threshold. This final list does not only contain the true positive identifications of peptide candidates (and corresponding proteins), but also false positive identifications of peptide candidates.

The identifications below the score threshold are false negative and true negative identifications.

Many reasons give rise to the undesirable false positive and true negative identifications such as poor quality MS-MS spectra, selection of peptides corresponding to protein with post translational modification (PTM) not including in the search parameters, etc.

The protein composition of the analyzed sample is generally unknown, or only partially known, by the user. Therefore the number of false positive identifications in the final peptides (and corresponding proteins) list cannot be determined individually but by using statistic methods such as decoy database searches.

The decoy database is built from the real database. The proteins of the decoy database are obtained by reversing or randomising the amino acid sequences of the proteins of the real database. The decoy database search is performed using identical search parameters as in the real database search.

The positive identifications of the real database searches give the number of true positive plus false positive identifications, and the positive identifications of the decoy database searches using the same search

parameters and score threshold conditions give an estimation of the number of false positive identifications in the real database searches.

The confidence level in the peptide (and corresponding protein) identifications is given by the FDR (False Discovery Rate) value defined by the ratio of the number of positive identifications of the decoy database searches divided by the number of positive identifications of the real database searches. Lower the FDR is, higher the confidence level of identification is.

The user can decrease the FDR value by simply increasing the score threshold. More sophisticated analyses can be used such as selecting positive protein identifications for which at least two different peptides have been identified

In LC-MS-MS of complex samples of proteins, more than one precursor are very often selected inadvertently with a mass selection window around the mass of the given precursor that is aimed at for producing an MS-MS spectrum.

The fragment ions of the plurality of selected precursors increase the complexity of the produced MS-MS spectrum, and can decrease the identification score obtained by the database search using scoring methods for the given precursor.

Furthermore, the database search is generally performed only for the given MS precursor, the peak of which is the most intense, and the other selected precursors are not considered.

Different solutions have been proposed to increase the MS-MS throughput of tandem mass spectrometry by simultaneously producing several MS-MS spectra.

A first solution is the simultaneous hardware production of several MS-MS spectra, each MS-MS spectrum corresponding to a standard MS-MS spectrum of a single precursor selected in the MS spectrum. The MS-MS spectra which are produced simultaneously are spatially (MS-MS) and temporally (MS) separated [1] [2].

Another solution is the production of multiplexed MS-MS spectra produced from a plurality of precursors selected in the MS spectrum per multiplexed MS-MS spectrum. The fragment ions of the selected precursors are deliberately mixed.

Individual MS-MS spectra, each corresponding to a single selected precursor can be produced from the analysis of the multiplexed MS-MS spectrum by using different methods of fragment-precursor identifications [3] [4] [5] [6] [7].

In the references [3] [4] [5] [6] [7], all the methods of fragment-precursor identifications use comparison of at least two (or more) multiplexed MS-MS spectra of the same plurality of precursors. These MS-MS spectra are successively produced with a modification of one experimental parameter of the used tandem mass spectrometer between two successive MS-MS acquisitions.

All the solutions described above [1] [2] [3] [4] [5] [6] [7] are hardware solutions. They depend on the type of tandem spectrometers used, and cannot be extended to other existing tandem mass spectrometers.

Purely software solutions for analyzing deliberately or inadvertently multiplexed MS-MS spectra have also been proposed [8] [9] [10] for “Bottom up” proteomics. These solutions [8] [9] [10] do not specifically depend on the type of tandem mass spectrometer, and need the production of only one multiplexed MS-MS spectrum of the plurality of selected precursors for fragment-precursor identifications. But high accuracy for both MS and MS-MS are needed for these methods.

The precursor-fragment identification method of reference [8] consists in submitting the multiplexed MS-MS spectra with the mass-to-charge values and charges of the plurality of selected precursors to database searches, without any previous algorithmic analysis of the multiplexed MS-MS spectra.

This MS-MS multiplexed method is limited by MS-MS accuracy and the number of detected fragments [8]. It can be efficiently used only for

tandem mass spectrometers with high MS-MS accuracy such as FT-MS (Fourier Tranform Mass Spectrometers).

The identification scores of the plurality of selected precursors of the multiplexed MS-MS spectrum analyzed by database searches using scoring methods decrease when the number of selected precursors increases.

This decreasing score effect is worse with a large intensity dynamic range in the MS spectrum between the plurality of selected precursors of the analyzed multiplexed MS-MS spectrum, because the existing scoring methods generally select the most intense peaks of the multiplex MS-MS spectrum for the database searches.

For example, when a small intensity peak of the MS spectrum is selected with a larger one, the database search of the corresponding multiplexed MS-MS spectrum using scoring methods can only identify the precursor corresponding to the larger intensity peak of the MS spectrum with a good score, and will produce low score or no identification for the precursor corresponding to the smaller one.

The multiplexed MS-MS methods of the references [9] [10] enable database searches with existing scoring methods using an algorithmic fragment filter for precursor-fragment identifications, before the submission to database searches.

The algorithmic fragment filter used [9] [10] is based on the identification of the complementary fragment ion pairs or multiplets in the multiplexed MS-MS spectrum corresponding to different dissociation pathways of each selected precursor. The sum of the masses of the fragment pairs or multiplets within the MS-MS accuracy equals to the mass of the corresponding selected precursor.

The multiplexed MS-MS methods of the references [9] [10] can be efficiently used only by high MS-MS accuracy tandem mass spectrometers so that the number of false complementary fragment ion pair identifications is limited. The false complementary fragment pair or multiplet identifications decrease the identification scores of the database searches.

The identification scores obtained with the multiplexed MS-MS methods of the references [9] [10] are also limited by the number of fragment MS-MS peaks identified by the software fragment filter used, because only a portion of the fragment ions of each selected precursor forms fragment pairs and can be identified in the corresponding multiplexed MS-MS spectrum.

The number of MS-MS spectra successively produced by using MALDI (Matrix Assisted Laser Desorption Ionisation) ion sources is not limited by the elution time as with LC-ESI-MS-MS, but by the ablation of the surface of the target by the laser shots.

The limitations of tandem mass spectrometry described before do not only relate to applications in “Bottom up” proteomics, but also concern “Top Down” proteomics using undigested proteins, and small molecule applications such as in metabolomics, or in identification of pollutants.

Summary of the invention

An aim of the invention is therefore to overcome the drawbacks of the state of the art as presented above to increase the MS-MS throughput of tandem (MS-MS) spectrometry using multiplexed MS-MS spectra and real and decoy database searches with scoring methods to improve precursor identifications.

In particular, one aim of the invention is to propose a method of multiplexed tandem (MS-MS) mass spectrometry compatible with all tandem mass spectrometers.

The present provides for this purpose a method as recited in claim 1.

This method enables identification of a plurality of precursors, which simultaneously selected to produce a multiplexed MS-MS spectrum, after an identification of the corresponding fragment ions.

A first plurality of individual MS-MS spectra corresponding to each selected precursor is produced with or without previous fragment filtering from the multiplexed MS-MS spectrum.

The first plurality of individual MS-MS spectra is then submitted to database searches without score threshold condition.

Each individual MS-MS spectrum is sent to first real and decoy database searches using scoring methods without score threshold condition.

All the positive identifications of the first real and decoy database searches are used to construct corresponding corrected real and decoy MS-MS spectra.

More specifically, the fragment ions of the multiplexed MS-MS spectrum are compared to fragment ions of theoretical real and decoy MS-MS spectra calculated from the identified precursors, for which a positive identification has been obtained from the first real and decoy database searches.

Then, each of the corrected real MS-MS spectra is sent to a second real database search using scoring methods with score threshold condition and each of the corrected decoy MS-MS spectra is sent to a second decoy database search using scoring methods with score threshold condition.

An FDR (False Discovery Rate) value, which gives an estimation of false positive identifications of the real database search, is determined using the positive identifications above the score threshold of both second real and decoy database searches.

Others features are presented in the dependent claims as well as the other independent claims.

Brief description of the drawings

Other aspects, aims and advantages of the invention will more clearly appear from the following description of the invention, which is provided by way of a non–limiting example and with reference to the appended drawings in which:

- Figure 1 is a flow chart of a preferred method of implementation of the multiplexed tandem spectrometry method of the invention,

- Figure 2 is an example of a simplified MS spectrum of peptides from a LC peak produced by a LC-MS-MS acquisition of Escherichia Coli protein sample, where the mass-to-charge ratio m/z values in Dalton (Da) are on the abscissa axis and the corresponding maximum intensity values on the ordinate axis,

- Figure 3 shows the simplified multiplexed MS-MS spectrum produced by the dissociation of two precursors selected from the MS spectrum corresponding to Figure 2, where the mass-to-charge ratio m/z values in Dalton (Da) of each MS-MS peak are on the abscissa axis and the corresponding maximum intensity values on the ordinate axis, and

- Figure 4 shows, a block diagram of one example of a mass spectrometer suitable for the implementation of embodiments of the method of the invention.

Description of preferred embodiments of the invention

First of all, it is recalled that what is meant by a multiplexed dissociation mass (MS-MS) spectrum is a dissociation mass (MS-MS) spectrum produced with a plurality of precursors selected in the primary (MS) mass spectrum where the fragment ions of the selected precursors are mixed.

The peaks of the individual MS-MS spectra that would be obtained if each of the selected precursors were analysed separately from the other are consequently mixed in the generated multiplexed MS-MS spectrum.

Referring to Figure 1 in particular, in the method of the invention, implemented with whichever mass spectrometer, the first step (a1) comprises supplying a primary (MS) mass spectrum for precursors after they have been ionized. The precursors are obtained from molecules that are to be identified.

The primary mass spectrum can be obtained, as known by the skilled person, by the ionization of the molecules to be identified in a ion source of charged ions, and acceleration with a substantially electric field, before their

injection into the tandem mass spectrometer, in order to generate the primary (MS) mass spectrum of precursor, without dissociation, wherein said MS spectrum contains primary ions peaks.

The primary MS spectrum can also be obtained by reading it from a database, such as a third-party database, in which it was previously saved.

As known by the person skilled in the art, in step (a2) a simplified MS spectrum is generally produced containing a list of mass-to-charge ratio m/z and corresponding maximum intensity values of each peak of the primary MS spectrum. Ion charge values are also added to the list when they can be determined.

Steps (a1) and (a2) can be jointly referred to hereafter as step (a).

In step (b), a plurality of precursors are deliberately or inadvertently selected from the primary MS spectrum, and the mass-to-charge ratio (m/z) values and charge values of each of the selected precursors are determined from the primary MS spectrum of the step (a) or from a mass selection window used.

In step (c1), the plurality of selected precursor ions are dissociated into fragment ions in the tandem mass spectrometer and a multiplexed MS-MS spectrum of the plurality of selected precursors is produced with the fragment ions by the tandem mass spectrometer and comprises peaks corresponding to detection of one or more fragments of the selected precursors.

In step (c2), a simplified multiplexed MS-MS spectrum is produced as a list of mass-to-charge ratio values m/z and the corresponding maximum intensity values of peaks of the multiplexed MS-MS spectrum. Possibly, ion charge values, when they are known, are added to the list.

The multiplexed MS-MS spectrum can also be obtained by reading it from a database, such as a third-party database, in which it was previously saved.

Steps (c1) and (c2) can be jointly referred to hereafter as step (c).

In step (d), a plurality of individual MS-MS spectra are produced from the multiplexed MS-MS spectrum of step (c). Each individual MS-MS spectrum corresponds to only one precursor selected from the MS spectrum.

Each individual MS-MS spectrum comprises mass-to-charge ratio (m/z) values, the corresponding maximum intensity values, and charge values (when determined) of the simplified multiplexed MS-MS spectrum, the mass-to-charge ratio values m/z and charge values (when determined) values corresponding to the only one precursor selected from the MS spectrum.

The individual MS-MS spectra of step (d) can also be produced after filtering the fragment ions of the simplified multiplexed MS-MS spectrum.

Without filtering of fragment ions, for each selected precursor, the individual MS-MS spectrum of step (d) is identical and corresponds exactly to the simplified multiplexed MS-MS spectrum of step (c).

With filtering of fragment ions, only the fragment ions of the simplified multiplexed MS-MS spectrum selected by the fragment filter are used to produce each individual MS-MS spectrum of step (d).

Filtering techniques become more useful for the method of the invention to clarify the individual MS-MS spectra produced in step (d) as the number of selected MS precursors increases.

The method of the invention is compatible with all possible techniques of fragment ion filtering dependent or not on the precursor mass. Non limiting examples of fragment ion filtering are “sequence tagging”, ”De Novo Sequencing” or the complementary fragment pair and multiplet technique [9] [10].

In step (e), each of individual MS-MS spectrum of step (d) is submitted to a first real database search using scoring method without score threshold condition, and to a corresponding first decoy database search using the same search parameters as for the first real database search.

In the sense of the invention, without score threshold condition should be understood that all identifications obtained by the database searches are taken into consideration without considering the scores obtained for each identifications.

The results of the real and decoy database searches identify candidate precursors. These candidate precursors will be further confirmed or infirmed as later described.

As a variant, the scoring method is carried out with low score threshold condition, this low score threshold condition is lower that conventionally used score threshold condition, such as lower than 10 or more advantageously lower than 5.As known by the person skilled in the art, decoy database search is generally used in proteomics applications to estimate the number of false positive peptides (and corresponding proteins) identifications among the positive identifications of the first real database search.

The confidence level in the peptide and corresponding protein identifications is given by the FDR (False Discovery Rate) value. The FDR value is defined by the ratio of the number of positive identifications from the decoy database search divided by the number of positive identifications from the real database search. Lower the FDR is; higher the confidence level of identifications is.

Unlike standard analysis, the method of the invention in the steps following the step (e) uses all the positive identifications of the database search including the ones normally rejected below the threshold score values used in standard analysis.

In step (f), for the individual MS-MS spectra for which the real, respectively decoy, database search produces positive identifications, real, respectively decoy, individual MS-MS spectra are produced from these positive identifications. The real and decoy individual MS-MS spectra can be referred to as “corrected” individual MS-MS spectra.

A real individual MS-MS spectrum comprises the mass-to-charge ratio (m/z) values and corresponding maximum intensity values of fragment ions of a candidate precursor resulting from the real database search of step (e).

A decoy individual MS-MS spectrum comprises the mass-to-charge ratio (m/z) values and corresponding maximum intensity values of fragment ions of a candidate precursor resulting from the decoy database search of step (e).

In a first embodiment of step (f), for producing a corrected individual MS-MS spectrum, a list of mass-to-charge ratio m/z values is computed from the candidate precursor identified in step (e). The mass-to-charge ratio m/z values correspond to theoretical fragment ions of the candidate precursor. Then all the fragment ions of the multiplexed MS-MS spectrum, of which the mass-to-charge ratio (m/z) value is comprised in the list, are selected to produce the corrected individual MS-MS spectrum. Thus, the selection is done within the instrumental MS-MS accuracy.

In a second embodiment of step (f), a real, respectively decoy, individual MS-MS spectrum is produced by selecting fragment ions in the simplified multiplexed MS-MS spectrum, which match the fragment ions of the candidate precursor, the fragment ions of the candidate precursor being identified in step (e) using the real, respectively decoy, database search.

This second embodiment of step (f) reduces the duration of the corrected individual MS-MS spectra production of this step. However, some fragment ions can be ignored in the identification of the first database search due to parameters of search algorithms used such as too low MS-MS peak intensity, compared with the previous calculated comparison.

Two different sets of corrected individual MS-MS spectra, corresponding respectively to the real and decoy database search results of step (e), are produced in step (f).

In a first embodiment of step (g), the two sets of corrected individual MS-MS spectra of step (f) and the corresponding precursor m/z values and charge values are submitted to real and decoy database searches using scoring methods with identical score threshold conditions, and identical search parameters, both in the real and decoy database searches.

That is, the set of real, respectively decoy, individual MS-MS spectra is submitted to a second real, respectively decoy, database search.

The database searches of step (g) can be performed with the same scoring method and databases used in step (e), or with other scoring methods and/or databases. The best result is obtained by using the same scoring method and databases for steps (e) and (g).

The database searches of step (g) are not standard but specific to the method of the invention. Indeed real and decoy database searches do not use the same set of individual MS-MS spectra, but two different sets of individual MS-MS spectra each one corresponding respectively to the results of the real and decoy database search of step (e).

A standard database search method using the same set of individual real MS-MS spectra produced from the positive identifications of the first real database search of step (e) for subsequent real and decoy database searches underestimates false positive identifications of the second real database search due to bias effects.

The correct statistical estimation of the false positive identifications is obtained with step (g) of the method with the two different sets of corrected individual MS-MS spectra for the real and decoy database searches.

In a second embodiment of step (g), this step is performed using scoring methods with a score threshold on the two sets of corrected individual MS-MS spectra of step (f) and the identification results of the first real and decoy database searches of step (e), without second real and decoy database searches.

A non-limiting example of such a scoring method is the production of an identification score for each corrected individual MS-MS spectrum. The

identification score is obtained by dividing the number of fragment ions of the corrected individual MS-MS spectrum by the number of all theoretically possible fragment ions determined from the candidate precursor identified in step (e).

The second embodiment of step (g) avoids second database searches, thus shortening the process.

Back to the first embodiment, in step (h) the precursors of the multiplexed MS-MS spectrum are identified by using the positive identification results of the real database searches of step (g) which are above a chosen score threshold, and the number of false positive identifications are estimated with the number of positive identifications of the decoy database searches of step (g) above the score threshold. Score identification threshold conditions and search parameters are identical for the real and decoy database searches.

In the second embodiment, i.e. without second database search of step (g), in step (h) the positive precursor identifications are obtained by selecting identifications above the score threshold used for the scoring method of step (g) with the set of real individual MS-MS spectra.

The false positive identifications are estimated by selecting identifications above the same score threshold used for the scoring method of step (g) with the set of decoy individual MS-MS spectra.

In the first embodiment, in step (i) the FDR (False Discovery Rate) value, which gives the confidence level of precursor positive identifications of real database searches, is determined by the ratio of the number of positive identifications of the decoy database search of step (h) divided by the number of positive identifications of the real database search of step (h).

In the second embodiment, in the step (i), the FDR (False Discovery Rate) value, which gives the confidence level of precursor positive identifications is determined by the ratio of the number of positive identifications obtained in step (h) with the set of decoy individual MS-MS

spectra divided by the number of positive identifications obtained in step (h) with the set of real individual MS-MS spectra.

As in standard analysis, steps (e) to (i) of the method of the invention can successively be carried out with different scoring methods and with different databases by using Mascot, Sequest, X!Tandem, or others. The precursor positive identifications obtained by the different search tools can be combined to increase the number of precursor positive identifications.

The FDR value can be selected by the user simply by choosing the corresponding score threshold, or by using more complex conditions, such as for example in “Bottom up” proteomics using LC-MS-MS data, the combination a score threshold value for peptide identifications, and at least two peptides identified per protein for protein identifications.

It is understood that the concrete implementation of the method of the invention can be typically achieved by a digital computer such as a DSP (Digital Signal Processor) executing the appropriate programs.

More practically, the present invention can be embodied in the form of a software module that is added to any existing tandem mass spectrometry device, and interfaced with the other software of this equipment.

In any case, the person skilled in the art will understand that production of the primary MS spectrum and of the multiplexed MS-MS spectra obtained with tandem mass spectrometry provides the possibility of identifying the selected precursors by using the method of the invention.

Compared with standard analysis using one precursor per multiplexed MS-MS spectrum produced, the MS-MS throughput and the corresponding precursor identifications of the method are increased proportionally to the number of precursors selected for each multiplexed MS-MS spectrum.

As a non-limited example, if three precursors are selected in average per multiplexed MS-MS spectrum produced, the final MS-MS throughput is improving by a factor three by using the method of the invention.

Steps (f) to (i) of this method transform a significant proportion of true negative identifications (scores below the score threshold) obtained with the standard MS-MS method into true positive identifications (scores above the score threshold).

The method of the invention does not depend on the mass spectrometry technique used to measure the mass-to-charge ratio m/z values of the primary and fragment ions, and the mass-to-charge ratio m/z values can be measured using time-of-flight, deflection in a magnetic field, frequency, etc.

The method of the invention is compatible with all types of tandem mass spectrometers, and can be performed both at low and high MS and MS-MS resolution and accuracy.

As in standard analysis, when considering the same number of multiplexed MS-MS spectra produced, the method of the invention produces more precursor positive identifications at higher MS and MS-MS resolution and accuracy compared with lower MS and MS-MS resolution and accuracy, due to the lower false positive identifications produced by the database searches.

It should be noted that in all the present description, mass-to-charge ratio (m/z) values can be replaced with mass values and vice versa.

<u>Components and operation of tandem mass spectrometers implementing the method of the invention</u>

Now will be described in greater detail, and by way of non-limiting examples some preferred tandem mass spectrometer components and operations implementing the multiplexed tandem mass spectrometry method of the invention.

A non limited example of tandem mass spectrometer suitable for the implementation of the method of the invention is shown in the Figure 4.

The analysis of complex sample with tandem mass spectrometers generally requires separation techniques 1 of the molecules of the sample before the introduction into the tandem mass spectrometer.

After the separation phase the molecules of the analyzed sample are introduced in the ion source 2 to be ionized.

The primary ions are introduced into the mass spectrometer 5 to produce the primary MS spectrum after their ionization in the ion source 2.

After the production of each MS spectrum, the primary ions of interest are selected as precursors in the MS spectrum by the precursor mass selector 3 to produce the multiplexed MS-MS spectra.

The selected primary ions are fragmented in the dissociation device 4 to produce the fragment ions used to produce the multiplexed MS-MS spectra.

The fragment ions are introduced into the mass spectrometer 5 to produce the multiplexed MS-MS spectra.

The method of the invention can be implemented with all existing tandem mass spectrometers known by the person skilled in the art, composed of two mass spectrometers operating sequentially in space separated by a dissociation device or a single mass analyzer operating sequentially in time.

The existing tandem mass spectrometers with spatial separation which can be used with method of the invention are Q-q-MS tandem mass spectrometers, where Q is a quadrupolar mass spectrometer used as precursor MS selector 3, q is the dissociation device 4, generally a multipolar waveguide containing gas using CID (Collision Induced Dissociation) dissociation technique, and MS is a TOF (Time of Flight) mass spectrometer 5 using orthogonal injection system (OTOF), or a quadrupolar (Q) mass spectrometer 5, or a FT-ICR (Fourier Transform Ion Cyclotron Resonance) mass spectrometer 5 that uses a static magnetic field, or a linear Ion Trap (IT) mass spectrometer 5.

The MS and the multiplexed MS-MS spectra are produced in the second mass spectrometer used (Q, TOF, IT, or FT-ICR).

The first quadrupolar Q is used for the selection of the precursor ions in the MS spectrum to produce the multiplexed MS-MS spectra after the dissociation of the selected primary ions in multipolar waveguide q by CID (Collision Induced Dissociation), or another technique of fragmentation.

Other tandem mass spectrometers with spatial separation which can be used with the method of the invention are MALDI-TOF-TOF, equipped with MALDI (Matrix Assisted Laser Desorption Ionization) ion source, and composed of a first linear TOF (Time-Of-Flight) mass spectrometer with a Bradbury-Nielson temporal gate used as MS selector 3, a collision cell for dissociation by using high kinetic energy CID 4, and a second axial TOF mass spectrometer with reflectron (RTOF) 5.

The MS and MS-MS spectra are produced in the second RTOF mass spectrometer. The Bradbury-Nielson temporal gate is used for the selection of the precursor ions in the MS spectrum after TOF separation in the first linear TOF mass spectrometer, and the selected precursor ions are dissociated in the collision cell by high kinetic energy CID, to produce the multiplexed MS-MS spectra of the selected precursor in the second RTOF mass spectrometer.

The existing single tandem mass spectrometers operating sequentially in time which can be used with method of the invention are linear 2D or 3D Ion trap(IT) mass spectrometers or Fourier Transform (FT-MS) mass spectrometers (FT-ICR or Orbitrap®).

The MS spectrum production, the precursor selection, the dissociation of the precursor ions by CID or another dissociation technique, and the MS-MS spectrum production are produced successively in the IT or the FT-MS mass spectrometer used, as known by the skilled person in the art.

Other existing tandem mass spectrometers IT-MS are combining spatial separation and sequentially time operations, with a 3D ion trap as IT

and an axial or orthogonal injection RTOF as MS mass spectrometer, and with a linear 2D ion trap as IT and a FT mass spectrometer (FT-ICR or Orbitrap) as MS mass spectrometer.

The MS spectra are produced in the axial or orthogonal injection RTOF or in the FT mass spectrometers, the precursor ions selection and dissociation phases are successively produced in the 3D and 2D IT, and the MS-MS spectra are finally produced in the IT used, or in the MS mass spectrometer (axial or orthogonal injection RTOF, or FT-MS).

The existing single tandem mass spectrometers operating sequentially in time described above can produce successive multiplexed MS-MS spectra of successive selected MS-MS peaks in the MS^n mode as known by the person skilled in the art.

The method of the invention is well suited for applications using liquid chromatographic (LC) as separation technique 1 (LC-MS-MS). But the method of the invention is compatible with all existing methods of separation of the molecules studied before the introduction in tandem mass spectrometers such as 1D or 2D gel electrophoresis (PAGE) separation.

As non limited examples, LC is generally coupled with ESI (Electrospray Ionization) ion sources, and 1 D or 2D PAGE is generally used with MALDI (Matrix Assisted Laser Desorption Ionisation) ion sources.

The method of the invention can be used with all existing ion sources 2. The ion used source can be an ESI (Electro-Spray Ionization) ion source, a MALDI (Matrix Assisted Laser Desorption Ionization) pulsed laser ion source, a DESI (Desorption Electrospray Ionization) ion source, an APCI (Atmospheric Pressure Chemical Ionization) ion source, an APPI (Atmospheric Pressure Photo Ionisation) ion source, a DART (Direct Analysis in Real Time) ion source, a LDI (Laser Desorption Ionization) ion source, an ICP (Inductively Coupled Plasma) ion source, en EI (Electron Impact) ion source, a CI (Chemical Ionization) ion source, a FI (Field Ionization) ion source, a FAB (Fast Atom Bombardment) ion source, a LSIMS (Liquid Secondary Ion Mass Spectrometry) ion source, an API

(Atmospheric Pressure Ionization) ion source, a FD (Field Desorption) ion source, a DIOS (Desorption Ionization On Silicon) ion source, or any other type of ion source producing primary ions.

As known by the skilled person in the art, the most commonly precursor mass selectors 3 used in tandem mass spectrometers are: quadrupolar (Q), linear 2D or 3D ion trap (IT), Bradbury-Nielson temporal gate, Fourier Transform mass spectrometers (FT-ICR and Orbitrap).

The fragmentation in the dissociation device 3 for the production of the multiplexed MS-MS spectra by the tandem mass spectrometers using the method of the invention can be implemented with a collision chamber containing gas that allows dissociation by CID/CAD (Collision Induced Dissociation/Collision Activated Dissociation), a time–of–flight space allowing spontaneous dissociation (PSD or Post Source Decay) after increasing the internal energy of the primary molecule ionised in the ion source or over the time–of–flight path by photo ionisation, or with the SID (Surface Induced Dissociation) technique, the ECD (Electron Capture Dissociation) technique, the ETD (Electron Transfer Dissociation) technique, the IRMPD (Infra Red Multi Photon Dissociation) technique, the PD (Photo Dissociation) technique, the BIRD (Back Body Infra Red Dissociation) technique, or again any method of fragmentation of the primary ions.

Different techniques of production of multiplexed MS-MS spectra necessary to the method of the invention the can be used by the existing tandem mass spectrometers described above.

The first one is the In-Source-Dissociation (ISD) method where the primary ions of all the different type of precursors are fragmented in the ion source 2 before the injection into the mass spectrometer without any primary mass selection in the MS spectrum.

The ISD method can be used with MALDI ion sources for Top down (pure proteins) or Bottom up (peptides) analysis of protein samples by

producing prompt fragmentation in the MALDI ion source by increasing the laser power density on the MALDI target.

It can be used also with ESI ion sources for Top down (pure proteins) or Bottom up (peptides) analysis of protein samples by using collision fragmentation with a gas of the multi charged ions produced by the ESI ion sources, before the injection in the mass spectrometer.

The second technique of production of multiplexed MS-MS spectra consists in increasing the width of the precursor mass selection window of the mass spectrometer used to select more than one precursor in the primary MS spectrum instead of only one precursor.

All the existing tandem mass spectrometer described above can use this method of multiplexed MS-MS spectra production by using broader mass selection window for precursor MS peak selection.

Considering that the minimum width of the precursor mass selection windows used in the existing tandem mass spectrometer is typically about of 0.1- 0.2% of the selected precursor mass value, and that in practical applications it can be typically of 0.5-1% of the selected precursor mass value, a significant fraction of the MS-MS spectra produced in standard tandem mass spectrometry are generally multiplexed MS-MS spectra with more than one precursor selected.

Therefore, the method of the invention can be also used for the analysis of standard tandem mass spectrometry data.

The third technique of production of multiplexed MS-MS spectra is the successive dissociation of several different precursors individually selected, adjacent or not to the other selected precursors, by a primary mass selection window of the mass spectrometer used, before producing the single multiplexed MS-MS spectrum of the mixtures of the fragments of all the individually selected precursors.

The Q-q-MS tandem mass spectrometer described above where MS is a linear 2D IT (LIT) or a FT-ICR, can use the third method of multiplexed MS spectra production.

The Q-q-LIT spectrometer can select successively each precursor MS with the Q, fragment the selected precursor in the q, and stored successively the dissociated fragment ions of each selected precursor in the LIT, before to produce the corresponding single multiplexed MS-MS spectrum of the fragment mixture in the LIT.

The Q-q-FT-ICR spectrometer can select each precursor with the Q, fragment the selected precursor in the q, and store successively the dissociated fragment ions of each selected precursor in the q, before injecting the mixture of the dissociated fragments of all the selected precursors in the FT-ICR, to produce the corresponding single multiplexed MS-MS spectrum of the fragment mixture.

The IT-MS spectrometer, where IT is a linear ion trap and MS is a Fourier transform mass spectrometer (FT-ICR or Orbitrap®) 5 described above, can also use the third method of multiplexed MS-MS spectrum production.

Each precursor is successively selected by the IT before to be fragmented in the IT or in another external collision cell, the fragment ions of the plurality of the different selected precursors are finally stored in an intermediate cell, before to be injected altogether in the FT-MS (FT-ICR or Orbitrap) to produce the multiplexed MS-MS spectrum.

The MALDI-TOF-TOF mass spectrometer described above can also use the third method of multiplexed MS-MS spectrum production.

Instead of selecting only one precursor in the MS spectrum at each laser shot on the MALDI target, the primary ions of several different precursor can be selected successively at each laser shot with the Bradbury-Nielson temporal gate after their separation in the first linear TOF spectrometer, to produce the multiplexed MS-MS spectrum of the different selected precursors by the accumulation of the detected fragments of all the laser shots. The method of the invention is compatible with all the different types of fragment ions produced by using all the existing fragmentation

techniques known by the person skilled in the art, such as a, b, c, y, z, x, or w fragment ions.

A non-limiting application of the method of the invention is the analysis of complex samples of peptides (Bottom-up proteomic) and pure proteins (Top-down proteomic) by using LC-ESI, 2D PAGE-MALDI, or LC-MALDI with tandem mass spectrometers by using database searches with scoring methods using search tools such as Mascot or Sequest.

The method of the invention can be used also for small molecule applications such metabolomics, or the identification of impurities or pollutants.

First example

Now, a non–limiting first example of implementation of the method of the invention will be described with reference to Figure 4.

A protein sample of Escherichia Coli was prepared, as known by the skilled person in the art, for LC-MS-MS analysis by using LC-ESI-Q-q-TOF mass spectrometer.

100 ng of the protein sample was digested using trypsin to generate a mixture of peptides before the injection in the LC capillary column 1. Effluent from the LC column 1 was electrosprayed by the ESI ion source 2 into the used Q-q-TOF mass spectrometer to produce the MS and the multiplexed MS-MS spectra of the peptide mixture.

During the elution time, at each LC peak, MS spectra have been produced, each MS spectrum following by the corresponding MS-MS spectra, containing multiplexed MS-MS spectra, by using the Q-q-TOF mass spectrometer, as described above for the second technique of MS-MS production.

Each MS spectrum is produced in the RTOF mass spectrometer 5, after the selection of the precursors with the quadrupolar mass spectrometer 3. The selected primary ions are dissociated by CID in the collision cell q 4,

before to be injected in the RTOF mass spectrometer 5 to produce each multiplexed MS-MS spectrum.

The width of the mass selection window used for the precursor selection in the primary MS spectrum was about 0.5-1% of the mass-to-charge ratio (m/z) value of the selected precursor, and was similar to the ones used in standard LC-MS-MS.

The MS and MS-MS accuracy used in the analysis was 20 ppm.

Figure 2 shows an example of simplified MS spectrum corresponding to the list of MS mass-to-charge ratio (m/z) values and corresponding maximum intensity values presented in table 1 obtained from a primary MS spectrum containing the MS peaks of peptides from a LC peak produced by the LC-MS-MS acquisition of the Escherichia Coli protein sample and corresponding to step (a) of the method according to the invention.

In the particular case of multi-charged primary ions, the charge of the precursor ions, if determined, is added to the mass-to-charge ratio m/z and the corresponding maximum intensity value list, as shown in the example of table 1.

The person skilled in the art will be able to determine the charge of the primary ions corresponding to each primary mass peak selected in the MS spectrum as precursor with the identification techniques normally employed in mass spectrometry.

Figure 3 shows an example of a simplified multiplexed MS-MS spectrum obtained conformingly to steps (b) and (c). The corresponding list of MS-MS mass-to-charge ratio (m/z) values and the corresponding maximum intensity values is shown in table 2. Conformingly to steps (b) and (c), the simplified MS-MS spectrum is obtained from a multiplexed MS-MS spectrum produced by the dissociation of the primary ions of two MS peaks selected simultaneously in the MS spectrum of Figure 2. The corresponding mass-to-charge ratio (m/z) and maximum intensity values of the two selected MS peaks are written in bold in table 1.

According to the presentation graph conventionally used (though in no way limiting) by the person skilled in the art of mass spectrometry, the primary MS and the multiplexed MS-MS mass spectrum is generally shown, as in the examples of Figures 2 and 3, with two perpendicular axes, with the mass-to-charge ratio m/z values on the abscissa axis, and the corresponding intensity values on the ordinate axis.

In step (d) two individual MS-MS spectra are produced without using fragment filtering techniques by using the mass-to-charge ratio (m/z) and the corresponding charge values of each one of the two selected precursor listed in bold in table 1, and the simplified multiplexed MS-MS spectrum of table 2.

In step (e), the two individual MS-MS spectra and their corresponding precursor mass-to-charge ratio (m/z) and charge values, produced in step (d) have been submitted to real and corresponding decoy database searches by using Mascot without score identification threshold.

20 ppm MS accuracy and 0.05 Da MS-MS accuracy were used as parameters for the Mascot searches.

The mascot positive identification results of the real database search are shown in the second column of table 3a. The peptide precursors with mass-to-charge ratio m/z value of 652.3905 Da and 650.3741 Da obtained score identification of 63 and 15.

The Mascot positive identification results of decoy database searches are shown in the second column of table 3b, with an identification score of 4 and 3 for the peptide precursor with mass-to-charge ratio m/z value of 652.3905 Da, and 650.3741 Da.

All the possible theoretical fragment ion mass-to-charge ratio (m/z) values corresponding to the Mascot identifications of real database searches of step (e) of the two selected peptide precursors of the example of Figures 2 and 3, are shown in tables 4a and 4b. The amino acid sequences of the two corresponding identified peptides are shown in the first column of tables 4a and 4b.

All the possible theoretical ion fragment mass-to-charge ratio (m/z) values corresponding to the Mascot identification using decoy database searches of step (e) of the two selected peptide precursors of the example of Figures 2 and 3, are shown in tables 5a and 5b. The amino acid sequences of the two corresponding false identified peptides are shown in the first column of tables 5a and 5b.

The types of fragments listed in tables 4a, 4b, 5a and 5b are known to the person skilled in the art. These fragments comprises (b, y) fragments and the same fragments with neutral losses (H2O, NH3, CO) during the dissociation of the precursor ions.

The theoretical MS-MS mass-to-charge ratio (m/z) values corresponding to the identified experimental MS-MS mass-to-charge ratio (m/z) values for each one of the selected precursor of the real database searches of step (e) are listed in bold in tables 4a and 4b.

The theoretical MS-MS mass-to-charge ratio (m/z) values corresponding to the identified experimental MS-MS m/z values for each one of the selected precursor of the decoy database searches of step (e) are listed in bold in tables 5a and 5b.

In step (f), the two real individual MS-MS spectra of the two selected precursors corresponding to the results of the real data search of the step (e) are produced and are listed in the tables 6a and 6b.

The two real individual MS-MS spectra of tables 6a and 6b produced in step (f) are composed of the MS-MS mass-to-charge ratio (m/z) values and the corresponding maximum intensity values of ion fragments identified by the comparison within 20 ppm accuracy between the experimental MS-MS mass-to-charge ratio (m/z) values of the simplified MS-MS spectrum of table 2 and the theoretical mass-to-charge ratio m/z values of tables 4a and 4b.

In step (f), the two decoy individual MS-MS spectra of the two selected precursors corresponding to the results of the decoy data search of the step (e) are produced, and are listed in tables 7a and 7b.

The two decoy individual MS-MS spectra of tables 7a and 7b produced in step (f) are composed of the MS-MS mass-to-charge ratio (m/z) values and the corresponding maximum intensity values of ion fragments identified by the comparison within 20 ppm accuracy between the experimental MS-MS mass-to-charge ratio (m/z) values of the simplified MS-MS spectrum of table 2 and the theoretical mass-to-charge ratio m/z values of tables 5a and 5b. In step (g), the two real individual MS-MS spectra of tables 6a and 6b with the corresponding mass-to-charge ratio (m/z) values and charge values of the two selected peptide precursors have been submitted to real database searches by using Mascot with score identification threshold conditions.

The corresponding mascot positive identification results are shown in the third column of table 3a. The selected peptide precursor with mass-to-charge ratio (m/z) value of 652.3905 Da obtained an identification score of 107, and the selected peptide precursor with mass-to-charge ratio (m/z) value of 650.3741 Da obtained an identification score of 77.

In step (g), the two decoy individual MS-MS spectra of tables 7a and 7b with the corresponding mass-to-charge ratio (m/z) values and charge values of the two selected precursors have been submitted to decoy database searches by using Mascot with the same score identification threshold condition as used in the real database searches.

The corresponding Mascot false positive identification results are shown in the third column of table 3b. The selected peptide precursor with m/z value of 652.3905 Da obtained a false identification score of 51, and the selected peptide precursor with m/z value of 650.3741 Da obtained a false identification score of 31.

The identification scores of the real database searches of the third column of the example of table 3a are both significantly higher than the score identification threshold value of the Mascot analysis of all the LC-MS-MS data which is equal to 44, and corresponding to 0.5% FDR peptide value.

The two examples of peptides of table 3a (and their parent proteins) are positively identified in steps (h) and (i) of the method of the invention, by using the real database search results.

The higher identification score (which is 51) of the decoy database searches of the third column of the example of table 3b, corresponding to the selected precursor with mass-to-charge ratio (m/z) value equal to 652.3905 Da, is above the score identification threshold value of the Mascot analysis obtained with all the LC-MS-MS data, which equals to 44.

This positive identification of the decoy database search of step (g) will be used as false positive identification to estimate the number of false positive identifications of the real database search of step (g).

The lower identification score (which is 31) of the decoy database searches of the third column of the example of table 3b, corresponding to the selected precursor with mass-to-charge ratio (m/z) value equal to 650.3741 Da, is below the score identification threshold value of the Mascot analysis obtained with all the LC-MS-MS data, which equals to 44.

This negative identification resulting from the decoy database searches will not be used as false positive identification to statistically estimate the number of false positive identifications in the real database search.

The identification score threshold value of 44 used in the example of tables 3a and 3b, corresponding to an FDR value of 0.5%, has been obtained from the full LC-MS-MS data analysis by using the method of the invention as described further.

The Mascot results of standard database searches obtained without using the method of the invention, i.e. when only the selected precursor with higher intensity of the multiplexed spectrum is used in the analysis, will give only one positive precursor identification (with the mass-to-charge ratio m/z value of 652.3905 Da) above the threshold score value of 25 corresponding to a FDR value of 0.5% by using all the LC-MS-MS data in the standard analysis. The results of the third column of the table 3a corresponding to the

final result of the method of the invention shows that the method of the invention allows the identification of the two selected peptide precursors with the identification score threshold value of 44 corresponding to the same FDR value of 0.5%.

In standard analysis without using the method of the invention, only the most intense precursor is considered for each produced MS-MS spectrum. The Mascot results of the analysis of the complete LC-MS-MS acquisition of Escherichia Coli sample described above, without using the method of the invention, provide 3896 identified peptides and 674 corresponding identified proteins. These results were obtained with a score threshold value of 25 corresponding to a FDR value of about 0.5% for peptide identifications, used for the standard Mascot real and decoy database searches.

Steps (a) to (d) of the method of the invention described above for one example of multiplexed MS-MS spectrum were applied to all the multiplexed MS-MS spectra of the Escherichia Coli LC-MS-MS acquisition.

The total number of experimental multiplexed MS-MS spectra produced in the LC-MS-MS acquisition was 8690. The number of MS-MS spectra produced in the step (d) by using the steps (a) to (d) of the method of the invention was 33325, corresponding to an increase of the MS-MS throughput by a factor of about 3.8 by using the method of the invention.

The positive identification Mascot results obtained by using steps (e) to (i) of the method of the invention with real database searches were 6055 identified peptides and 828 corresponding identified proteins. These results were obtained with a score threshold value of 44 corresponding to an FDR value of about 0.5% for peptide identifications, used for the Mascot real and decoy database searches.

The use of the method of the invention to analyze the same Escherichia Coli LC-MS-MS data produced with a Q-q-TOF mass spectrometer increases the number of identified peptides by 55% and the number of identified proteins by 23% compared with standard analysis by

using the same Mascot parameters for the database searches and with the same FDR value of about 0.5%.

Table 1: example of simplified primary MS spectrum

m/z(Da) z	Relative Intensity %	m/z (Da) z	Relative intensity %	m/z(Da) z	Relative Intensity %	m/z(Da) z	Relative intensity%
299.2919 1+	0.73	602.6405 3+	0.68	710.0708 4+	2.38	872.4336 2+	2.25
400.5380 3+	2.27	607.7870 1+	2.60	710.3189 2+	3.75	873.9672 2+	0.65
405.8616 3+	4.44	608.2888 2+	3.44	712.8480 2+	5.04	881.0969 3+	1.86
427.6905 4+	1.15	608.2891 4+	6.14	718.6456 4+	8.59	883.1690 4+	2.93
435.2628 3+	6.46	627.3258 2+	0.89	729.6029 4+	1.19	903.4571 2+	1.54
455.2559 3+	0.62	631.3205 3+	2.72	740.4064 2+	5.14	913.7541 3+	0.56
473.8817 3+	0.20	639.3080 3+	9.01	747.3628 5+	6.72	927.7898 3+	1.17
475.5678 3+	0.56	641.2687 3+	4.03	750.9045 2+	2.68	928.7853 3+	7.48
503.7856 2+	5.34	643.6650 3+	4.52	764.4055 2+	9.86	933.9517 4+	1.95
513.6208 3+	0.42	645.0170 3+	1.21	767.7201 3+	4.68	946.4787 2+	1.33
518.2447 3+	0.64	647.3872 2+	1.34	769.7276 2+	1.05	946.7597 3+	0.80
521.9367 3+	0.94	**650.3741 2+**	**4.81**	776.0336 3+	4.9	947.9730 2+	1.44
533.2542 2+	1.7	**652.3905 2+**	**54.83**	776.3618 2+	3.2	957.8584 3+	1.39
543.6158 3+	1.39	655.6516 3+	1.82	782.4015 2+	2.13	958.9600 2+	0.64
556.2683 3+	0.47	661.3253 4+	0.29	810.7164 3+	3.21	961.3995 2+	4.97
557.3057 3+	1.62	662.9964 3+	1.86	814.9201 2+	1.98	965.7978 3+	4.34
560.8030 1+	1.08	672.3470 3+	5.10	823.3850 3+	3.30	972.4682 3+	0.94
563.9263 3+	0.63	682.3802 2+	4.63	827.9133 2+	30.03	982.9737 2+	3.45
569.9183 3+	3.41	685.5674 4+	2.06	833.8989 2+	2.87	1008.0169 2+	1.47
570.8379 2+	2.80	696.3444 2+	4.36	835.3838 3+	5.97	1047.0438 2+	0.29
574.9426 3+	1.64	696.5945 2+	2.89	835.4549 2+	6.83	1151.5782 2+	0.21
577.6211 3+	2.02	696.8408 4+	2.66	845.3859 2+	1.73	1177.2229 3+	2.00
581.3309 2+	100	698.3650 3+	3.43	847.4399 2+	1.70	1221.9809 1+	9.04
582.9805 3+	4.64	704.0228 3+	3.50	854.3738 2+	0.98	1392.6743 2+	0.34
599.3135 2+	1.41	705.0354 3+	5.35	865.9280 2+	3.05		

Table 2: Example of simplified multiplexed MS-MS spectrum

m/z(Da)	Relative Intensity %	m/z (Da)	Relative intensity %	m/z(Da)	Relative Intensity %	m/z(Da)	Relative intensity%
284.1593	2.94	387.2189	5.14	505.2732	1.98	723.4366	1.85
286.1735	2.93	389.2358	1.90	513.3019	2.25	735.4386	4.41
298.1740	41.75	397.2402	5.83	521.3057	3.02	743.4699	2.02
299.2116	2.67	399.2226	25.34	523.2842	8.54	745.4768	43.03
302.1676	3.12	410.1987	2.00	528.3314	4.21	753.4436	6.88
314.2963	2.95	411.2593	2.17	531.3459	16.10	771.4558	5.78
316.1838	21.41	413.2332	4.09	539.3141	12.50	783.4648	2.16
327.1648	3.04	415.2520	11.44	541.2955	22.53	800.4824	17.36
329.2115	2.02	417.2325	21.94	553.3339	3.07	816.5125	64.33
329.2146	5.01	422.2355	6.14	557.3242	9.98	838.5322	3.76
330.1631	4.89	425.2321	8.17	559.3050	11.31	839.5387	5.32
331.2315	27.38	428.2120	6.15	569.3297	2.41	854.5034	4.38
339.1646	4.43	430.2998	18.71	571.3453	2.36	862.5050	3.27
339.2001	5.89	434.2383	3.08	581.3602	2.40	872.4387	3.13
342.1990	9.68	440.2499	16.74	588.3317	10.86	887.5500	80.49
343.1598	6.25	443.2460	5.32	592.8623	2.13	896.4220	2.06
345.1733	4.22	450.2657	2.15	599.3738	5.11	913.5692	3.35
351.2008	5.14	452.2460	9.68	622.3526	5.03	951.6138	3.72
352.1571	1.95	456.2431	2.62	632.3928	39.19	955.5527	2.20
355.1958	3.68	456.2779	2.36	636.3648	4.10	969.4891	2.16
357.2106	23.58	458.2591	12.81	640.3635	17.55	970.5888	3.64
369.2106	17.35	468.2787	13.48	652.3948	3.18	976.4577	2.68
371.2232	4.32	470.2598	12.33	654.3769	7.60	988.5953	100
373.2068	6.10	472.2715	4.18	658.3730	8.86	998.5806	1.86
374.2378	2.78	485.3223	2.29	670.4085	8.20	1028.5927	2.17
375.1846	1.87	486.2888	11.94	672.3895	4.30	1082.6094	1.91
381.2117	4.62	488.2650	11.14	687.3981	11.42	1101.6792	37.09
387.1864	9.13	497.2673	1.94	712.4270	3.38	1129.6378	1.97

Table 3a: Real database search results

m/z values of peptide precursors	Mascot Identification scores of real database searches by using step (e) of the method of the invention	Mascot Identification scores of real database searches by using step (g) of the method of the invention
650.3741	15	77
652.3905	63	103

Table 3b: Decoy database search results

m/z values of peptide precursors	Mascot Identification scores of decoy database searches by using step (e) the method of the invention	Mascot Identification scores of decoy database searches by using step (g) of the method of the invention
650.3741	3	31
652.3905	4	51

Table 4a: m/z values of theoretical fragments for positive identification of precursor m/z = 652.3905 Da with real database searches (Da)

Amino Acid Sequence	b y	b^{++} y^{++}	b^{*} y^{*}	b^{*++} y^{*++}	b^{0} y^{0}	b^{0++} y^{0++}
T	102.0550 --	51.5311 --			84.0444 --	42.5258 --
T	203.1026 1202.7355	102.0550 601.8714	-- 1185.7089	-- 593.3581	185.0921 1184.7249	93.0497 **592.8661**
L	**316.1867** **1101.6878**	158.5970 551.3475	-- 1084.6612	-- 542.8343	**298.1761** 1083.6772	149.5917 542.3423
T	**417.2344** **988.6037**	209.1208 494.8055	-- 971.5772	-- **486.2922**	**399.2238** **970.5932**	200.1155 485.8002
A	**488.2715** **887.5560**	244.6394 444.2817	-- 870.5295	-- 435.7684	**470.2609** 869.5455	235.6341 435.2764
A	**559.3086** **816.5189**	280.1579 408.7631	-- 799.4924	-- 400.2498	**541.2980** 798.5084	271.1527 399.7578
I	**672.3927** **745.4818**	336.7000 373.2445	-- 728.4553	-- 364.7313	**654.3821** 727.4713	327.6747 364.2393
T	773.4403 **632.3978**	**387.2238** 316.7025	-- 615.3712	-- 308.1892	755.4298 614.3872	378.2185 307.6972
T	874.4880 **531.3501**	437.7477 266.1787	-- 514.3225	-- 257.6654	856.4775 **513.3395**	428.7424 257.1734
V	973.5564 **430.3024**	487.2819 215.6548	**--** **413.2758**	-- 207.1416	955.5459 --	478.2766 --
L	1086.6405 **331.2340**	543.8239 166.1206	**--** 314.2074	-- 157.6074	1068.6299 --	534.8186 --
A	1157.6776 218.1499	579.3424 109.5786	-- 201.1234	-- 101.0653	1139.6671 --	570.3372 --
K	-- 147.1128	-- 74.0600	-- 130.0863	-- 65.5468		

Table 4b: m/z values of theoretical fragments for positive identification of precursor m/z = 650.3741 with real database searches (Da)

Amino Acid Sequence	b y	b^{++} y^{++}	b^{*} y^{*}	b^{*++} y^{*++}	b^{0} y^{0}	b^{0++} y^{0++}
G	58.0287 --	29.5180 --				
I	171.1128 1242.7304	86.0600 621.8688	-- 1225.7038	-- 613.3556	-- 1224.7198	-- 612.8635
T	272.1605 **1129.6463**	136.5839 565.3268	-- 1112.6198	-- 556.8135	254.1499 1111.6458	127.5786 556.3215
D	**387.1874** **1028.5986**	194.0974 514.8030	-- 1011.5721	-- 506.2897	**369.1769** 1010.5881	185.0921 505.7977
I	500.2715 **913.5717**	250.6394 457.2895	-- 896.5451	-- 448.7762	482.2609 895.5611	241.6341 448.2842
L	613.3556 **800.4876**	307.1814 400.7475	**--** **783.4611**	-- 392.2342	595.3450 782.4771	**298.1761** 391.7422
V	**712.4240** **687.4036**	356.7156 344.2054	**--** **670.3770**	-- 335.6921	694.4134 669.3930	347.7103 335.2001
V	811.4924 **588.3352**	406.2498 294.6712	**--** **571.3086**	-- **286.1579**	793.4818 570.3246	**397.2445** 285.6659
D	926.5193 489.2667	463.7633 245.1370	**--** **472.2402**	-- 236.6237	908.5088 471.2562	454.7580 236.1317
N	1040.5623 374.2398	520.7848 187.6235	1023.5357 **357.2132**	512.2715 179.1103	1022.5517 --	511.7795 --
L	1153.6463 260.1969	577.3268 130.6021	1136.6198 243.1703	568.8135 122.0888	1135.6358 --	568.3215 --
K	-- 147.1128	-- 74.0600	-- 130.0863	-- 65.5468		

Table 5a: m/z values of theoretical fragments for positive identification of precursor m/z = 652.3905 with decoy database searches (Da)

Amino Acid Sequence	b y	b^{++} y^{++}	b^{*} y^{*}	b^{*++} y^{*++}	b^{0} y^{0}	b^{0++} y^{0++}
R	157.1084 --	79.0578 --	140.0818 --	70.5446 --		
I	270.1925 1147.6834	135.5999 574.3453	253.1659 1130.6568	127.0866 565.8320	-- **1129.6728**	-- 565.3400
S	**357.2245** 1034.5993	179.1159 517.8033	340.1979 1017.5728	170.6026 509.2900	339.2139 1016.5887	170.1106 508.7980
F	504.2929 947.5673	252.6501 474.2873	487.2663 930.5407	244.1368 465.7740	486.2823 929.5567	243.6448 465.2820
K	**632.3879** **800.4989**	316.6976 400.7531	615.3613 783.4723	308.1843 392.2398	614.3773 782.4883	307.6923 391.7478
L	**745.4719** 672.4039	373.2396 336.7056	728.4454 655.3774	364.7263 328.1923	727.4614 654.3933	364.2343 327.7003
S	832.5039 **559.3198**	416.7556 280.1636	815.4774 542.2933	408.2423 271.6503	814.4934 **541.3093**	407.7503 271.1583
P	929.5567 472.2878	465.2820 236.6475	912.5302 455.2613	456.7687 228.1343	911.5461 454.2772	456.2767 227.6423
S	1016.5887 375.2350	508.7980 188.1212	999.5622 358.2085	500.2847 179.6079	998.5782 **357.2245**	499.7927 179.1159
L	**1129.6728** 288.2030	565.3400 144.6051	1112.6463 271.1765	556.8268 136.0919	1111.6622 --	556.3348 --
R	-- 175.1190	-- 88.0631	-- 158.0924	-- 79.5498		

Table 5b: m/z values of theoretical fragments for positive identification of precursor m/z = 650.3741 with decoy database searches (Da)

Amino Acid Sequence	b y	b^{++} y^{++}	b^{*} y^{*}	b^{*++} y^{*++}	b^{0} y^{0}	b^{0++} y^{0++}
A	72.0444 --	36.5258 --				
L	185.1285 1228.7008	93.0679 614.8540	-- 1211.6743	-- 606.3408	-- 1210.6902	-- 605.8488
I	**298.2125** 1115.6167	149.6099 558.3120	-- 1098.5902	-- 549.7987	-- 1097.6002	-- 549.3067
D	413.2395 1002.5327	207.1234 501.7700	-- 985.5061	-- 493.2567	395.2289 984.5221	198.1181 492.7647
A	484.2766 **887.5057**	242.6419 444.2565	-- 870.4792	-- 435.7432	466.2660 869.4952	233.6366 435.2512
L	597.3606 **816.4686**	299.1840 408.7380	-- 799.4421	-- 400.2247	579.3501 798.4581	290.1787 399.7327
S	684.3927 703.3846	342.7000 352.1959	-- 686.3580	-- 343.6826	666.3821 685.3740	333.6947 343.1906
R	840.4938 616.3525	420.7505 308.6799	823.4672 599.4260	412.2373 300.1666	822.4832 598.3420	411.7452 299.6746
T	941.5415 460.2514	471.2744 230.6293	924.5149 443.2249	462.7611 222.1161	923.5309 442.2409	462.2691 221.6241
S	**1028.5735** 359.2037	514.7904 180.1055	1011.5469 342.1772	506.2771 171.5922	1010.5629 341.1932	505.7851 171.1002
P	1125.6262 272.1717	563.3168 136.5895	1108.5997 255.1452	554.8035 128.0762	1107.6157 --	554.3115 --
R	-- 175.1190	-- 88.0631	-- 158.0924	-- 79.5498		

Table 6a: corrected individual MS-MS spectrum of step (f) of the method of the invention for real database search Precursor m/z = 652.3095 Da

m/z (Da)	Relative Intensity %
298.1740	41.75
316.1838	21.41
331.2315	27.38
387.2189	5.14
399.2226	25.34
417.2325	21.94
430.2998	18.71
470.2598	12.33
486.2888	11.94
488.2650	11.14
531.3459	16.10
541.2955	22.53
559.3050	11.31
632.3928	39.19
654.3769	7.60
672.3895	4.30
745.4768	43.03
816.5125	64.33
887.5500	80.49
970.5888	3.64
988.5953	100
1101.6792	37.09

Table 6b: corrected individual MS-MS spectrum of step (f) of the method of the invention for real database search

Precursor m/z = 650.3741 Da

m/z (Da)	Relative Intensity %
298.1740	41.75
357.2106	23.58
374.2378	2.78
387.1864	9.13
397.2402	5.83
588.3317	10.86
687.3981	11.42
712.4270	3.38
783.4648	2.16
800.4824	17.36
913.5692	3.35
1028.5927	2.17
1129.6378	1.97

Table 7a: corrected individual MS-MS spectrum of step (f) of the method of the invention for decoy database search

Precursor m/z = 652.3095 Da

m/z (Da)	Relative Intensity %
486.2888	11.94
745.4768	43.03
783.4648	2.16
998.5806	1.86

Table 7b: corrected individual MS-MS spectrum of step (f) of the method of the invention for decoy database search

Precursor m/z = 650.3741 Da

m/z (Da)	Relative Intensity %
413.2332	4.09
1028.5927	2.17

Second example

A non–limiting second example of implementation of the method of the invention will now be described with reference to Figure 4.

A protein sample of Human cell was prepared, as known by the skilled person in the art, for LC-MS-MS analysis by using an LC-ESI-IT(LTQ)-FT-MS (Orbitrap) mass spectrometer.

1 µg of the protein sample was digested using trypsin to generate a mixture of peptides before the injection in the LC capillary column 1. Effluent from the LC column 1 was electrosprayed by the ESI ion source 2 into the used IT-FT-MS mass spectrometer 5 to produce the MS and the multiplexed MS-MS spectra of the peptide mixture.

During the elution time, at each LC peak, MS spectra have been produced using the FT-MS mass spectrometer, following by the multiplexed MS-MS spectra production corresponding to the second method of multiplexed MS-MS production described above.

Each MS spectrum is produced in the FT-MS mass spectrometer. After each selection of the precursors with the IT 3, the selected primary ions are injected in the collision cell (HCD) 4 in order to be dissociated by CID, before to be injected in the FT-MS mass spectrometer 5 to produce each multiplexed MS-MS spectrum.

The width of the mass selection windows used for the precursor selection in the MS spectrum was about 6 Da, instead of the one of 3 Da normally used in standard LC-MS-MS with the used IT-FT-MS mass spectrometer.

The MS resolution used to produce the MS spectrum was 30000, and the MS-MS resolution was 7500. The corresponding MS and MS-MS accuracies used in the analysis were 4 ppm and 10 ppm.

The Mascot results of the analysis of the complete LC-MS-MS acquisition of Human cell sample described above, without using the

method of the invention, provide 2838 identified peptides and 761 corresponding identified proteins. These results were obtained with a score threshold value of 37 corresponding to a FDR value of about 0.85% for peptide identifications, used for the standard Mascot real and decoy database searches.

Steps (a) to (d) of the method of the invention described above were applied to all the multiplexed MS-MS spectra of the LC-MS-MS acquisition.

The total number of experimental multiplexed MS-MS spectra produced in the LC-MS-MS acquisition was 15242. The number of MS-MS spectra produced in the step (d) by using the steps (a) to (d) of the method of the invention was 49605, corresponding to an increase of the MS-MS throughput by a factor of about 3.25 by using the method of the invention.

The positive identification Mascot results obtained by using steps (e) to (i) of the method of the invention with real database searches provided 9742 identified peptides and 1318 corresponding identified proteins. These results were obtained with a score threshold value of 66 corresponding to a FDR value of about 0.86% for peptide identifications, used for the Mascot real and decoy database searches.

4 ppm MS accuracy and 0.01 Da MS-MS accuracy were used as parameters for the Mascot searches.

The use of the method of the invention to analyze the same Human cell LC-MS-MS data produced with an LTQ-Orbitrap increases the number of identified peptides by 243% and the number of identified proteins by 73% compared with standard analysis by using the same Mascot parameters for the database searches and with the same FDR value of about 0.85%.

BIBLIOGRAPHICAL REFERENCES

[1] J.D. Pinston et al, Rev. Sci.Instrum., 57 (4), (1983), p.583.

C. G. Enke et al, US Patent 4,472,631 (1984).

[2] C. G. Enke, patent PCT/US 2004/008424.

[3] Alderdice et al, US patent 5,206,508 (1993).

[4] R.H. Bateman, J. M. Brown, D. J. Kenny, US patent 2005/0098721 A1 (2005).

[5] R.W. Vachet and J. Wilson, US patent 7,141, 784 (2006).

[6] D.Scigocki, patent PCT/EP2008/056428.

[7] R.H. Bateman et al, EP 1 385 194 A2 (2004).

[8] C. Masselon et al, Proteomics, 3, (2003), p. 1279.

[9] D.Scigocki, patent application PCT/EP2007/05059.

[10] D.Scigocki, patent application PCT/EP2008/068271.

CLAIMS

1. A method for multiplexed tandem mass spectrometry of a sample to be analysed containing at least two precursors, wherein at least two simplified multiplexed MS-MS spectra are obtained each from at least two selected precursors of the sample, the method comprising:

(d) for each selected precursor generating an individual MS-MS spectrum from the simplified multiplexed MS-MS spectrum by selecting maximum intensity values and corresponding mass-to-charge ratio m/z values of fragment ions of the simplified multiplexed MS-MS spectrum, the fragment ions are potential fragment ions obtained from the precursor;

(e) submitting each individual MS-MS spectrum of step (d) to a real and a decoy database searches using a scoring process without score threshold condition or low score threshold condition for identifying candidate precursors and their fragment ions;

(f) producing real individual MS-MS spectra from identified candidate precursors resulting from the real database search of step (e), one real individual MS-MS spectrum being produced for one identified candidate precursor; and

producing decoy individual MS-MS spectra from identified candidate precursors resulting from the decoy database search of step (e), one decoy individual MS-MS spectrum being produced for one identified candidate precursor;

(g) submitting the real and decoy individual MS-MS spectra to a further scoring process with a score threshold condition for determining a score for each real and decoy individual MS-MS spectra.

2. The method of claim 1, wherein the simplified multiplexed MS-MS spectrum is obtained using a mass spectrometer, and wherein producing a real, respectively decoy, individual MS-MS spectrum of step (f) comprises:

computing from a candidate precursor identified in step (e) using the real, respectively decoy, database search a list of mass-to-charge ratio m/z values corresponding to theoretical fragment ions of the candidate precursor;

selecting all fragment ions of the simplified multiplexed MS-MS spectrum, of which the mass-to-charge ratio m/z values match with a mass-to-charge ratio m/z value of the list, within MS-MS accuracy of the mass spectrometer.

3. The method of claim 1, wherein producing a real, respectively decoy, individual MS-MS spectrum of step (f) comprises:

selecting fragment ions in the simplified multiplexed MS-MS spectrum, which match the fragment ions of the candidate precursor, the fragment ions of the candidate precursor being identified in step (e) using the real, respectively decoy, database search.

4. The method of one of claims 1 to 3, wherein step (g) comprises submitting the real, respectively decoy, individual MS-MS spectra to a real, respectively decoy, database search using scoring process with score threshold condition.

5. The method of claim 4, wherein the real, respectively decoy, databases used in step (e) and step (g) are identical.

6. The method of claim 4, wherein the real, respectively decoy databases used in step (e) and step (g) are different.

7. The method of one of claims 1 to 6, wherein the scoring processes used in step (e) and step (g) are identical processes, respectively without and with a threshold condition.

8. The method of one of claims 1 to 6, wherein the scoring processes used in step (e) and step (g) are different.

9. The method of one of claims 1 to 3, wherein determining a score for a real, respectively decoy, individual MS-MS spectrum in step (g) comprises dividing the number of fragment ions of the real, respectively decoy, individual MS-MS spectrum by the number of all theoretically possible fragment ions of the candidate precursor identified in step (e).

10. The method of one of claims 1 to 9, wherein, for each selected precursor, the individual MS-MS spectrum of step (d) comprises the simplified multiplexed MS-MS spectrum and mass or mass-to-charge ratio (m/z) value of the selected precursor.

11. The method of one of claims 1 to 9, further comprising, prior to step (d):

forming fragment ion pairs or multiplets from masses of the fragment ions of the simplified multiplexed MS-MS spectrum; when the sum of the masses of at least two fragment ions equals the mass of one given selected precursor, the at least two fragment ions form a fragment ion pair or multiplet and are assigned to the given selected precursor; and wherein

in step (d), the individual MS-MS spectrum of the given selected precursor comprises the assigned fragment ion pairs and/or multiplets and the mass or mass-to-charge ratio (m/z) value of the given selected precursor.

12. A computer program designed to be implemented in a tandem mass spectrometry system, including a set of instructions adapted to control said mass spectrometry system so that it performs the method of one of claims 1 to 11 when the computer program is run in the tandem mass spectrometry system.

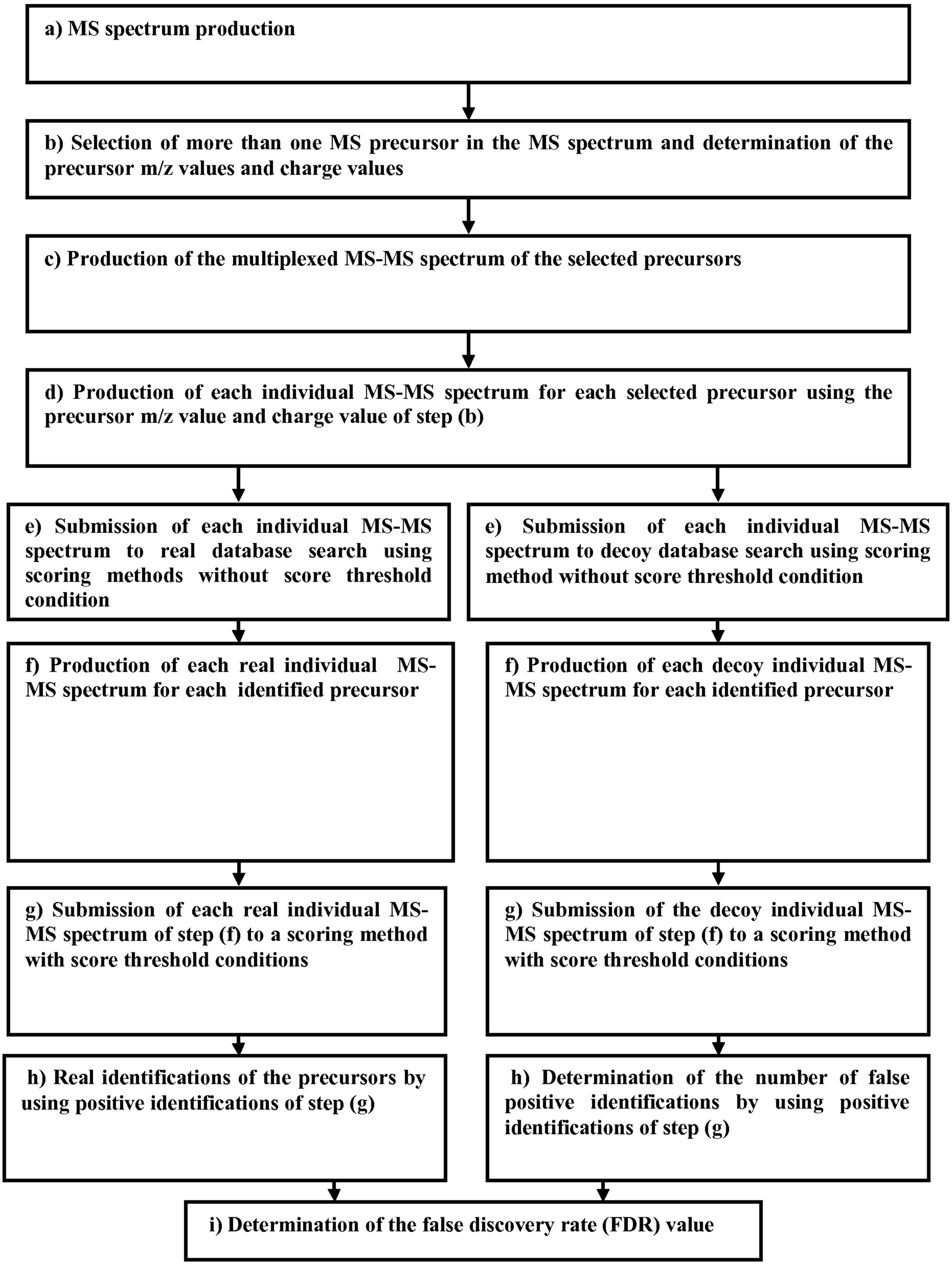

a) MS spectrum production
b) Selection of more than one MS precursor in the MS spectrum and determination of the precursor m/z values and charge values
c) Production of the multiplexed MS-MS spectrum of the selected precursors
d) Production of each individual MS-MS spectrum for each selected precursor using the precursor m/z value and charge value of step (b)
e) Submission of each individual MS-MS spectrum to real database search using scoring methods without score threshold condition
e) Submission of each individual MS-MS spectrum to decoy database search using scoring method without score threshold condition
f) Production of each real individual MS-MS spectrum for each identified precursor
f) Production of each decoy individual MS-MS spectrum for each identified precursor
g) Submission of each real individual MS-MS spectrum of step (f) to a scoring method with score threshold conditions
g) Submission of the decoy individual MS-MS spectrum of step (f) to a scoring method with score threshold conditions
h) Real identifications of the precursors by using positive identifications of step (g)
h) Determination of the number of false positive identifications by using positive identifications of step (g)
i) Determination of the false discovery rate (FDR) value

FIG.1

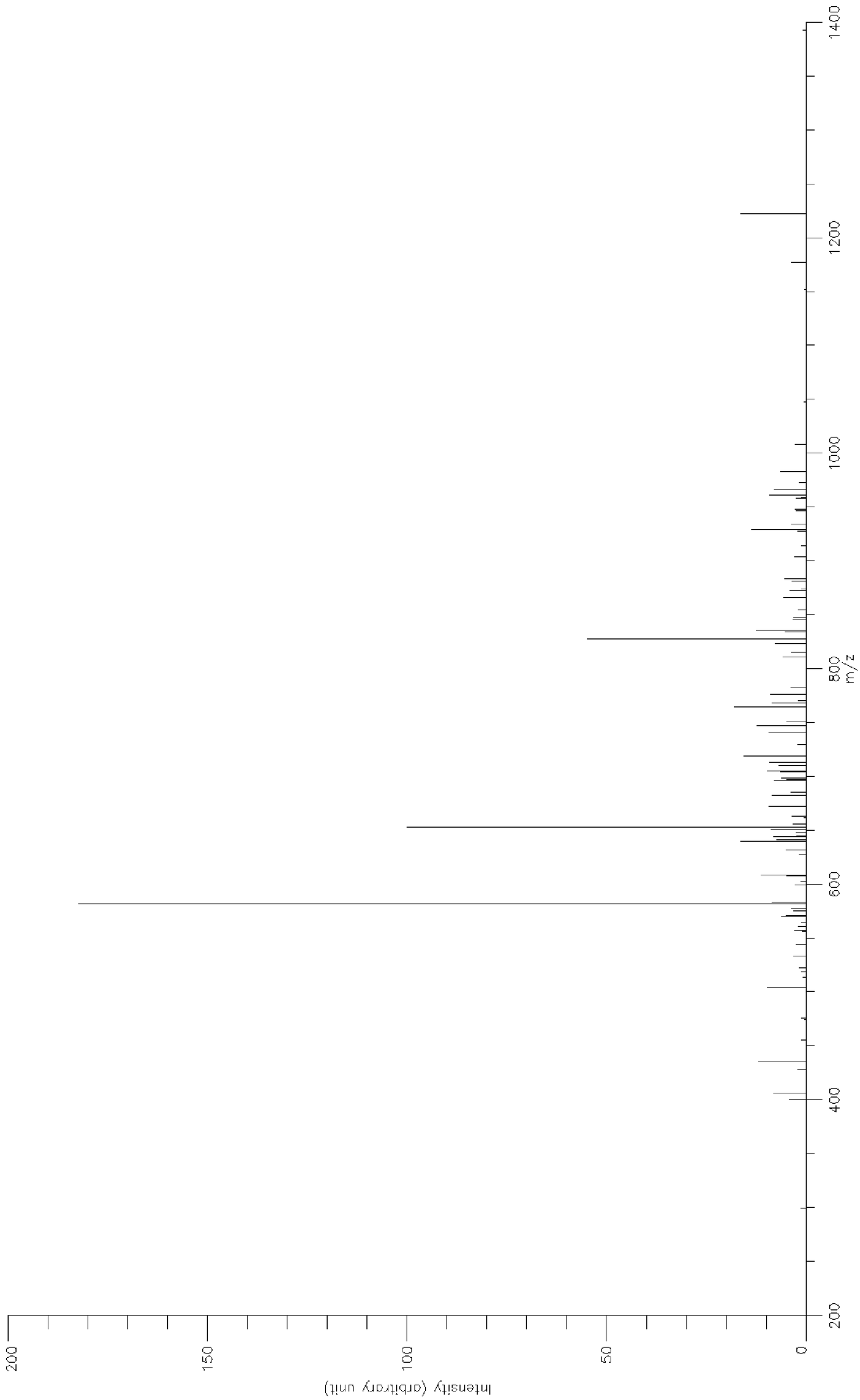

1400
1200
1000
800
600
400
200
m/z
200
150
100
50
0
Intensity (arbitrary unit)

FIG.2

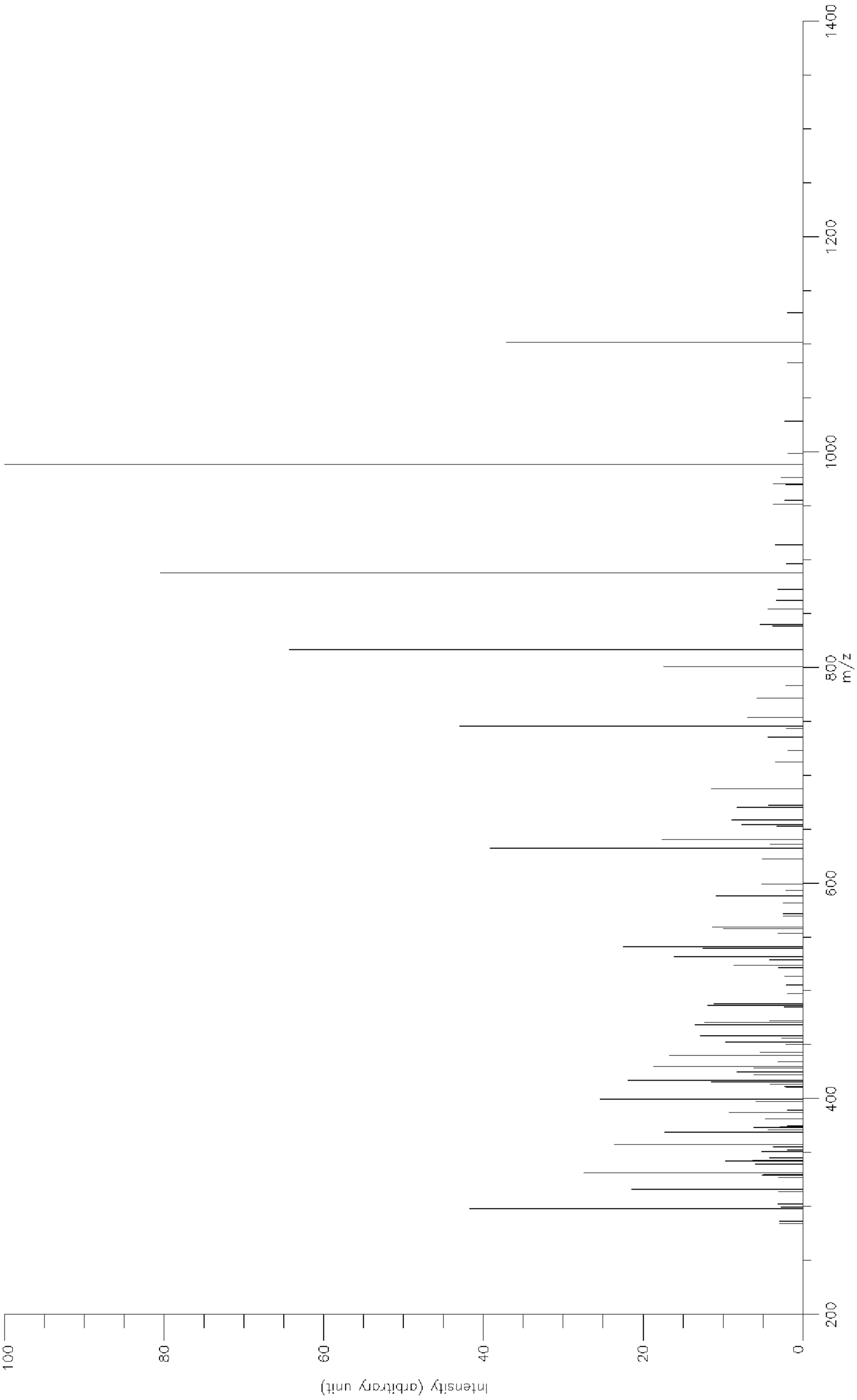

1400
1200
1000
800
600
400
200
m/z
100
80
60
40
20
0
Intensity (arbitrary unit)

FIG.3

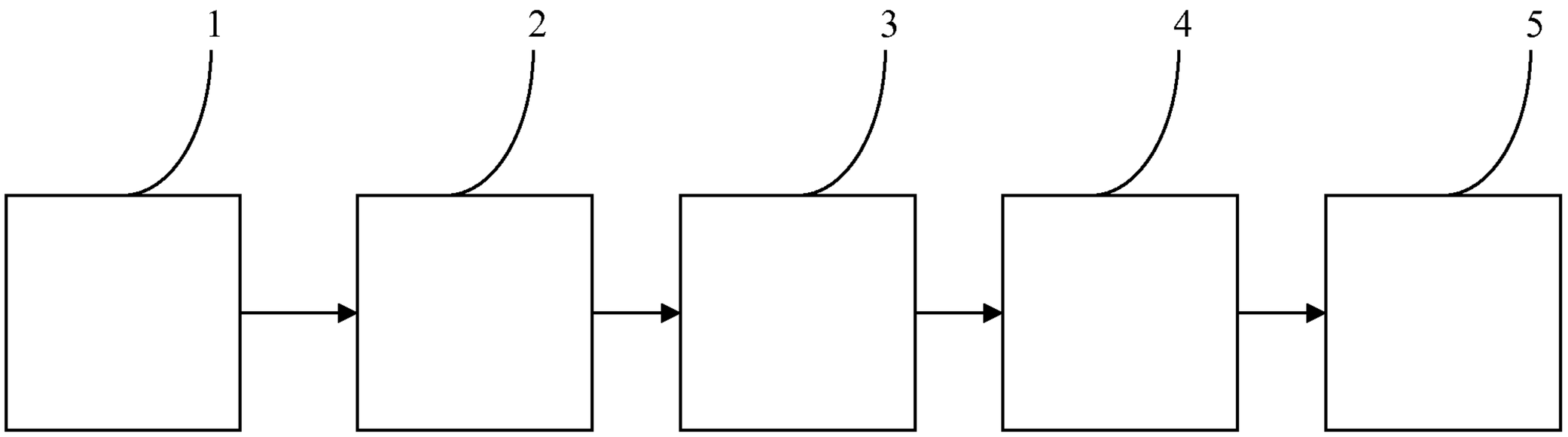


FIG.4