



(51) International Patent Classification:

G01N 30/72 (2006.01) G01N 30/02 (2006.01)  
G06K 9/00 (2006.01)

(21) International Application Number:

PCT/US2019/019354

(22) International Filing Date:

25 February 2019 (25.02.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/635,379 26 February 2018 (26.02.2018) US

(71) Applicant: LECO CORPORATION [US/US]; 3000 Lakeview Avenue, St. Joseph, MI 49085 (US).

(72) Inventors: ARTAEV, Viatcheslav; 3866 Derry Drive, St. Joseph, MI 49085 (US). RICHARDS, Todd; 931 Western Avenue, Watervliet, MI 49098 (US). FELL, Lorne; 4454 Preserve Place, Stevensville, MI 49127 (US). MCNITT, Kevin, Lee; 4703 Vince Court, Saint Joseph, MI 49085 (US).

(74) Agent: APPLIEDORN, Thomas, J.; Honigman LLP, 39400 Woodward Avenue, Suite 101, Bloomfield Hills, MI 48304-5151 (US).

(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, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) Title: METHOD FOR RANKING LIBRARY HITS IN MASS SPECTROMETRY

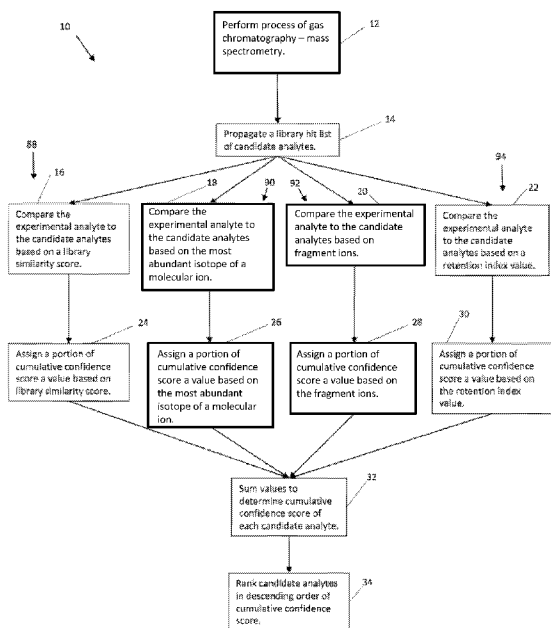


FIG. 1

(57) Abstract: A method for ranking analytes includes the steps of analyzing an experimental analyte in a mass spectrometer. The method includes comparing the experimental analyte to a plurality of candidate analytes in a library hit list, and assigning a cumulative confidence score to each candidate analyte based on the steps of comparing the experimental analyte to the candidate analytes based on a library similarity score, comparing the experimental analyte to the candidate analytes based on a presence of the most abundant isotope of a molecular ion and its mass, comparing the experimental analyte to the candidate analytes based on an abundance of fragment ions and a mass of the fragment ions, and, in some implementations, comparing the experimental analyte to the candidate analytes based on a retention index value. The method includes ranking the candidate analytes based on the cumulative confidence score of each candidate analyte.



## Method for Ranking Library Hits in Mass Spectrometry

### TECHNICAL FIELD

**[0001]** This disclosure relates to the field of processing data obtained from analyses of samples in a mass spectrometer, including by means of Gas Chromatography – High Resolution Mass Spectrometry (GC-HRMS), Comprehensive Two-Dimensional Gas Chromatography – High Resolution Mass Spectrometry (GC×GC-HRMS), and the like.

### BACKGROUND

**[0002]** It may be desirable to improve reliability of analytes assignments resulting from analysis of complex mixtures of volatiles and semi-volatiles compounds. It may be desirable to reduce an analyst's efforts in reviewing the results of analytes assignment by sorting the possible library hits in a more meaningful order and reducing the probability of erroneous results in the library hit lists.

**[0003]** In Gas Chromatography – High Resolution Mass Spectrometry (GC-HRMS), as well as in Comprehensive Two-Dimensional Gas Chromatography – High Resolution Mass Spectrometry (GC×GC-HRMS), a mixture of unknown or suspected analytes is introduced into the GC injector and the analytes are separated according to their physicochemical properties. The separated analytes consequently elute from a GC column and are introduced into an ion source of the mass spectrometer for ionization. Analyte molecules are ionized and fragment ions are created. The mass spectrometer then detects the ions from analyte molecules and fragment ions, and the data is presented as mass spectra, which is a list of ion mass and intensity combinations corresponding to each moment of detection.

**[0004]** The mass spectra may be recorded, saved, and searched against mass spectral databases. Mass spectral databases include commercially available databases, such as, for example, NIST, Willey, etc. As another example, the mass spectral databases may be custom-created by users. The results of the search may be presented as a hit list for each of the analytes in question. The hit list contains a list of possible candidates for assignment suggested for the analyte in question. Each of the hit list entries contains information describing the candidate, e.g., name, molecular formula, molecular structure,

mass spectrum, ring double bond equivalents (RDBE), etc., and information describing a matching probability of the experimental mass spectrum to the library mass spectrum of the candidate, information such as, for example, similarity score, reverse search score, probability score, etc.

5 **[0005]** The hit list may contain many entries (depending on a set of matching parameters set by user), and the hit list may be sorted in order of similarity scores from high to low, with the assumption that candidates with higher similarity scores are more likely to be a correct assignment. An analyst may perform a review of the analyte assignments. During the review, the analyst may select the highest ranked hit as the peak  
10 assignment. Alternatively, the analyst may select the peak assignment from a range of the highest ranked hits, e.g., from a top five, a top ten, or any other suitable range.

**[0006]** The review may include making sense of the possibility of the presence of the suggested candidate in the analyzed sample based on sample type and origin, looking for characteristic fragment ions, correspondence to chromatographic conditions (i.e., how  
15 volatile the candidates are and how it corresponds to injector and column temperature), elution order, etc. The review may require substantial expertise in GC-MS and related topics, and may be trusted to experts in the GC-MS field.

**[0007]** In some cases, the analyst may obtain a candidate analyte (as a purchased standard or synthesized), conduct an analysis on the known analyte with identical, or  
20 similar, conditions as the candidate analyte, and compare the results from the known analyte with the candidate analyte. If the mass spectra and the chromatographic retention time match, then typically the assignment will be confirmed. The process of analytes assignment by manual curation of chromatograms and mass spectra, as well as confirmation by running standards, may be long, expensive, and may not guarantee 100%  
25 reliability of the results. Therefore, improvements in reliability and simplification of the process of analytes assignments may be desirable for analytical chemists and researchers using the GC-MS results.

## SUMMARY

**[0008]** One aspect of the disclosure provides a method for ranking analytes includes  
30 the steps of analyzing an experimental analyte in a mass spectrometer, for example, in

connection with gas chromatography - mass spectrometry. The method includes comparing the experimental analyte to a plurality of candidate analytes in a library hit list, and assigning a cumulative confidence score to each candidate analyte based on the steps of comparing the experimental analyte to the candidate analytes based on a library similarity score, comparing the experimental analyte to the candidate analytes based on a presence and mass accuracy of a most abundant isotope of the molecular ion, comparing the experimental analyte to the candidate analytes based on a presence and mass accuracy of fragment ions, and, in some implementations, comparing the experimental analyte to the candidate analytes based on a retention index value. The method includes ranking the candidate analytes based on the cumulative confidence score of each candidate analyte.

**[0009]** Implementations of the disclosure may include one or more of the following features.

**[0010]** The library similarity score may have a failing range of values equal to and below X, a passing range of values equal to and above Y, and a neutral range of values between X and Y.

**[0011]** A portion of the cumulative confidence score of a candidate analyte in the failing range may be assigned a value of -1.5, a portion of the cumulative confidence score of a candidate analyte in the neutral range may be assigned a value of 0, and a portion of the cumulative confidence score of a candidate in the passing range may be assigned a value of 1.

**[0012]** The mass accuracy, relative to the candidate, of the most abundant isotope of the molecular ion may have a failing criteria, a passing criteria, and a neutral criteria based on a minimum acceptable abundance and a mass accuracy window. The failing criteria may require the abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and a mass of the most abundant isotope of the molecular ion to be outside of the mass accuracy window. The passing criteria may require the abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and the mass of the most abundant isotope of the molecular ion to be within the mass accuracy window. The neutral criteria may be satisfied if the molecular ion is not expected to be in the mass spectrum, its abundance in

the library is below the established minimum abundance, or the number of the ions in the ion peak is below minimum number.

**[0013]** A portion of the cumulative confidence score of a candidate analyte satisfying the failing criteria may be assigned a value of -1.5, a portion of the cumulative  
5 confidence score of a candidate analyte satisfying the passing criteria may be assigned a value of 1, and a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria may be assigned a value of 0.

**[0014]** The fragment ions may have a passing criteria, neutral criteria, and a failing criteria based on a minimum acceptable abundance and a mass accuracy window. The  
10 passing criteria may be satisfied if the abundance of the fragment ions exceeds the minimum acceptable abundance and the mass of the fragment ions, in which chemical formulae are generated only from the subsets of the chemical elements of the molecular formula of the candidate analyte, is within the mass accuracy window, while fragment ions chemical formulae may only be a subset of the elements from the molecular formula  
15 of the candidate analyte. The failing criteria may be satisfied if the abundance of the fragment ions exceed the minimum acceptable abundance and the mass of the fragment ions, in which chemical formulae are generated only from the subsets of the chemical elements of the molecular formula of the candidate analyte, is outside of the mass accuracy window. The neutral criteria may be satisfied if there are no fragment ions to  
20 exceed the minimum acceptable abundance or the number of ions in the peaks are below the minimum criteria.

**[0015]** A portion of the cumulative confidence score of a candidate analyte satisfying the passing criteria may be assigned a value of 1, and a portion of the cumulative  
25 confidence score of a candidate analyte satisfying the failing criteria may be assigned a value of -1.5, and a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria may be assigned a value of 0.

**[0016]** The minimum acceptable abundance of the fragment ions may be Z%.

**[0017]** The retention index value may have a passing range, a neutral range, and a failing range based on an acceptable retention index window. The passing range may be  
30 within the acceptable retention index window. The neutral criteria may require the

retention index value of the candidate analyte to be unavailable. The failing range may be outside the acceptable retention index window.

[0018] A portion of the cumulative confidence score of a candidate analyte in the passing range may be assigned a value of 1, a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria may be assigned a value of 0, and a portion of the cumulative confidence score of a candidate analyte in the failing range may be assigned a value of 0.

[0019] The acceptable retention index window may be  $W$  units.

[0020] The cumulative confidence score may range from -6 to 4.

[0021] The candidate analytes may be ranked in descending order of the cumulative confidence score of each candidate analyte.

[0022] The details of one or more implementations of the disclosure are set forth in the accompanying drawings and the description below. Other aspects, features, and advantages will be apparent from the description and drawings, and from the claims.

#### DESCRIPTION OF DRAWINGS

[0023] FIG. 1 is a flowchart of a method for ranking analytes.

[0024] FIG. 2 is a flowchart of a portion of the method for ranking analytes.

[0025] FIG. 3 is a flowchart of a portion of the method for ranking analytes.

[0026] FIG. 4 is a flowchart of a portion of the method for ranking analytes.

[0027] FIG. 5 is a flowchart of a portion of the method for ranking analytes.

[0028] Like reference symbols in the various drawings indicate like elements.

#### DETAILED DESCRIPTION

[0029] With reference to FIG. 1, a method for ranking analytes 10 is generally shown. The method for ranking analytes 10 includes the steps of analyzing an experimental analyte using gas chromatography – mass spectrometry 12. While the method is disclosed in connection with the chromatography – mass spectrometry arrangement shown at 12, other configurations are suitable for the ranking method in connection with other mass spectrometer configurations. With continued reference to Figure 1, the method 10 includes comparing the experimental analyte to a plurality of

candidate analytes in a library hit list 16–22, and assigning a cumulative confidence score to each candidate analyte 24–30 based on the steps of comparing the experimental analyte to the candidate analytes based on a library similarity score 16, comparing the experimental analyte to the candidate analytes based on a most abundant isotope of a molecular ion 18, comparing the experimental analyte to the candidate analytes based on an abundance and mass accuracy of the fragment ions 20, and, in some implementations, comparing the experimental analyte to the candidate analytes based on a retention index value 22. The method includes ranking the candidate analytes based on the cumulative confidence score of each candidate analyte 34.

**[0030]** In an implementation and as described herein, the experimental analyte may be analyzed using gas chromatography – mass spectrometry (GC-MS) or gas chromatography – high resolution mass spectrometry (GC-HRMS). As another example, the analysis of the experimental analyte 12 may be performed by any suitable process. The use of GC-MS, GC×GC-MS, GC-HRMS or GC×GC-HRMS may be such as that known in the art.

**[0031]** The library hit list of candidate analytes may be propagated 14 by any suitable means, such as, for example, a software application, a manual process, etc. The library hit list may contain data about each of the candidate analytes such that the candidate analytes may be compared to the experimental analyte. The comparison of the candidate analytes to the experimental analyte, and the corresponding data, may be presented in a hit table.

**[0032]** With reference to FIG. 2, a method for comparing library similarity scores 88 is generally shown. Library similarity score, and related criteria, may also be referred to as Factor 1. The method 88 may include determining the library similarity score of each candidate analyte 36. For example, the library similarity score may have a failing range of values equal to and below X, a passing range of values equal to and above Y, and a neutral range of values between X and Y. X may be, for example, 600, and Y may be, for example, 800. Alternatively, X and Y may be any suitable value. The failing range, neutral range, and passing range may be defined by a user, and may be any suitable ranges of values. If the library similarity score is within the failing range 38, e.g., equal to or below 600, then a portion of the cumulative confidence score may be assigned a

value of -1.5 at step 44. If the library similarity score is within the passing range 40, e.g., equal to or above 800, then a portion of the cumulative confidence score may be assigned a value of 1 at step 46. If the library similarity score is within the neutral range 42, e.g., between 601 and 799, then a portion of the cumulative confidence score may be assigned a value of 0 at step 48.

5 [0033] With reference to FIG. 3, a method for comparing the most abundant isotope of a molecular ion 90 is generally shown. The most abundant isotope of a molecular ion, and related criteria, may also be referred to as Factor 2. The user may assign a mass accuracy window, for example, in ppm or mDa units, and the user may assign a minimum acceptable abundance of molecular ions most abundant isotope in a library spectrum. 10 The method 90 may include determining the presence of a most abundant isotope of a molecular ion and its m/z. The most abundant isotope of the molecular ion may have a failing criteria, a passing criteria, and a neutral criteria based on the minimum acceptable abundance and the mass accuracy window. The failing criteria may require the 15 abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and a mass of the most abundant isotope of the molecular ion to be outside of the mass accuracy window. The passing criteria may require the abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and the mass of the most abundant isotope of the molecular ion to be within 20 the mass accuracy window. The neutral criteria may be satisfied if the failing criteria and the passing criteria are not satisfied. If the passing criteria is satisfied 52, e.g., the abundance of a most abundant isotope of a molecular ion is above the minimum acceptable abundance and the mass is within the mass accuracy window, a portion of the cumulative confidence score may be assigned a value of 1 in step 58. If the failing 25 criteria is satisfied 54, e.g., the abundance of a most abundant isotope of a molecular ion is above the minimum acceptable abundance and the mass is outside of the mass accuracy window, a portion of the cumulative confidence score may be assigned a value of -1.5 in step 60. If the neutral criteria is satisfied 56, e.g., neither the passing criteria nor the failing criteria are satisfied, a portion of the cumulative confidence score may be assigned 30 a value of 0 in step 62.

**[0034]** A minimum number of ions in an ion peak to be considered for mass accuracy criteria may be defined as, for example, 25 ions. If an ion peak in a mass spectrum contains less than the minimum number of ions, e.g., 25 ions, and those ions are used for defining the corresponding cumulative confidence scores, then the neutral criteria may be satisfied and a portion of the cumulative confidence score may be assigned a value of 0. As another example, a more elaborate definition for the minimum number of ions and mass accuracy window may be implemented based on, for example, resolving power of the system, m/z of the ion in consideration, spectra acquisition rate, chromatographic peak width, etc. Ion peaks which have intensity above the system capability to detect them correctly, e.g., saturated ions, may be ignored.

**[0035]** With reference to FIG. 4, a method for comparing the fragment ions 92 is generally shown. The fragment ions, and related criteria, may also be referred to as Factor 3. The user may define a minimum acceptable abundance of the fragment ions. For example, the minimum acceptable abundance of the fragment ions may be Z%. For example, Z may be equal to 300 out of 1000, i.e., 30%, or any other suitable value. The user may define a mass accuracy window, for example, in ppm or mDa units. The fragment ions may have a passing criteria, a neutral criteria and a failing criteria based on the minimum acceptable abundance and the mass accuracy window. The passing criteria may be satisfied if the abundance of the fragment ions exceeds the minimum acceptable abundance and the mass of the fragment ions, in which chemical formulae are generated only from the subsets of the chemical elements of the molecular formula of the candidate analyte, is within the mass accuracy window, while fragment ions chemical formulae may only be a subset of the elements from the molecular formula of the candidate analyte. The failing criteria may be satisfied if the abundance of the fragment ions exceed the minimum acceptable abundance and the mass of the fragment ions, in which chemical formulae are generated only from the subsets of the chemical elements of the molecular formula of the candidate analyte, is outside of the mass accuracy window. The neutral criteria may be satisfied if there are no fragment ions to exceed the minimum acceptable abundance or the number of ions in the peaks are below the minimum criteria. If the passing criteria is satisfied 66, e.g., the abundance of the fragment ions is above the minimum acceptable abundance and the mass is within the mass accuracy window, a

portion of the cumulative confidence score may be assigned a value of 1 in step 70. If the failing criteria is satisfied 68, e.g., the passing criteria is not satisfied, a portion of the cumulative confidence score may be assigned a value of -1.5 in step 72. If the neutral criteria is satisfied 67, e.g., the abundance of the fragment ions is below the minimum acceptable abundance or none of the fragment ion peaks are above the minimum criteria, a portion of the cumulative confidence score may be assigned a value of 0 in step 71.

**[0036]** A minimum number of ions in an ion peak to be considered for mass accuracy criteria may be defined as, for example, 25 ions. If an ion peak in a mass spectrum contains less than the minimum number of ions, e.g., 25 ions, and those ions are used for defining the corresponding cumulative confidence scores, then a neutral criteria may be satisfied and a portion of the cumulative confidence score may be assigned a value of 0. As another example, a more elaborate definition for the minimum number of ions and mass accuracy window may be implemented based on, for example, resolving power of the system, m/z of the ion in consideration, spectra acquisition rate, chromatographic peak width, etc. Ion peaks which have intensity above the system capability to detect them correctly, e.g., saturated ions, may be ignored.

**[0037]** With reference to FIG. 5, a method for comparing the retention index values 94 is generally shown. Retention index values may also be referred to as Factor 4. The user may define an acceptable retention index window. For example, the acceptable retention index window may be W units. For example, W may have a value of 50, or any other suitable value. The retention index value may have a passing range, a neutral criteria, and a failing range based on the acceptable retention index window. The passing range may be within the acceptable retention index window. The neutral criteria may require the retention index value of the candidate analyte to be unavailable. The failing range may be outside the acceptable retention index window. If the passing criteria is satisfied 76, e.g., the retention index value of the candidate analyte is within the acceptable retention index window, a portion of the cumulative confidence score may be assigned a value of 1 in step 82. If the failing criteria is satisfied 78, e.g., the retention index value of the candidate analyte is outside of the acceptable retention index window, a portion of the cumulative confidence score may be assigned a value of -1.5 in step 84. If the neutral criteria is satisfied 80, e.g., the retention index value of the candidate

analyte is unavailable, a portion of the cumulative confidence score may be assigned a value of 0 in step 86.

**[0038]** With reference to FIG. 1, the portions of the cumulative confidence score may be summed to determine the cumulative confidence score for each candidate analyte 32.

5 The cumulative confidence score may range from -6 to 4. The candidate analytes may be ranked in descending order of the cumulative confidence score of each candidate analyte 34. For example, the candidate analytes may be ranked in descending order of the cumulative confidence score of each candidate analyte by default, and a secondary sort may be by spectral similarity score in case of equal values of cumulative confidence  
10 scores for two or more candidate analytes. If there is still a tie, the first found candidate analyte may be ranked higher than the later found candidate analytes.

**[0039]** The method for ranking analytes 10 may improve reliability of the candidate analyte suggestions by excluding erroneous hits, and moving to the top of the hit list candidate analytes which are more probably from gas chromatography, e.g., Factor 4, and  
15 mass spectrometry, e.g., Factors 1–3.

**[0040]** To assist the user with understanding how the cumulative confidence scores were calculated, the hit table may include a concerns column with notes indicated which factors were assigned a neutral value, e.g., 0, or a failed value, e.g., -1.5.

**[0041]** As another example, the value assigned to a portion of the cumulative  
20 confidence scores, e.g., the passing value, failing value, and neutral value, may be modified for each factor to better reflect importance of the various factors on reliability of the assignment. For example, if Factor 1 was determined to be more important than Factor 2, the values assigned to the failing, neutral, and passing ranges in Factor 1 may be changed from -1.5, 0, and 1 to -2.5, 0, and 2, respectively. As another example,  
25 additional factors may be added based on other criteria, such as, for example, isotope abundance.

**[0042]** A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made without departing from the spirit and scope of the disclosure. Accordingly, other implementations are within the scope of the  
30 following claims.

## WHAT IS CLAIMED IS:

1. A method for ranking analytes, comprising the steps of:
  - analyzing an experimental analyte in a mass spectrometer;
  - comparing the experimental analyte to a plurality of candidate analytes in a
  - 5 library hit list, and assigning a cumulative confidence score to each candidate analyte based on the steps of:
    - comparing the experimental analyte to the candidate analytes based on a library similarity score;
    - comparing the experimental analyte to the candidate analytes based on a
    - 10 presence of the most abundant isotope of a molecular ion and its mass; and
    - comparing the experimental analyte to the candidate analytes based on an abundance of fragment ions, number of ions in the peaks, and a mass of the fragment ions; and
    - ranking the candidate analytes based on the cumulative confidence score of each
    - 15 candidate analyte.
2. The method for ranking analytes of claim 1, wherein the library similarity score has a failing range of values equal to and below X, a passing range of values equal to and above Y, and a neutral range of values between X and Y.
- 20 3. The method for ranking analytes of claim 2, wherein a portion of the cumulative confidence score of a candidate analyte in the failing range is assigned a value of -1.5, a portion of the cumulative confidence score of a candidate analyte in the neutral range is assigned a value of 0, and a portion of the cumulative confidence score of a candidate in
- 25 the passing range is assigned a value of 1.
4. The method for ranking analytes of claim 1, wherein the most abundant isotope of the molecular ion has a failing criteria, a passing criteria, and a neutral criteria based on a minimum acceptable abundance and a mass accuracy window, the failing criteria
- 30 requiring the abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and a mass of the most abundant isotope of the

molecular ion to be outside of the mass accuracy window, the passing criteria requiring the abundance of the most abundant isotope of the molecular ion to exceed the minimum acceptable abundance and the mass of the most abundant isotope of the molecular ion to be within the mass accuracy window, and the neutral criteria being satisfied if the failing  
5 criteria and the passing criteria are not satisfied.

5. The method for ranking analytes of claim 4, wherein a portion of the cumulative confidence score of a candidate analyte satisfying the failing criteria is assigned a value of -1.5, a portion of the cumulative confidence score of a candidate analyte satisfying the  
10 passing criteria is assigned a value of 1, and a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria is assigned a value of 0.

6. The method for ranking analytes of claim 1, wherein the fragment ions has a passing criteria, a neutral criteria and a failing criteria based on a minimum acceptable  
15 abundance and a mass accuracy window, the passing criteria requiring the abundance of the fragment ions to exceed the minimum acceptable abundance and the mass of the fragment ions to be within the mass accuracy window, and the neutral criteria being satisfied if there are no fragment ions exceeding the minimum abundance criteria or if none of the fragment ion peaks contain at least a minimum number of ions.

7. The method for ranking analytes of claim 6, wherein a portion of the cumulative confidence score of a candidate analyte satisfying the passing criteria is assigned a value of 1, a portion of the cumulative confidence score of a candidate analyte satisfying the  
25 failing criteria is assigned a value of -1.5, and a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria is assigned a value of 0.

8. The method for ranking analytes of claim 6, wherein the minimum acceptable abundance of the fragment ions is Z%.

9. The method for ranking analytes of claim 1, wherein the step for comparing the experimental analyte to a plurality of candidate analytes in a library hit list, and assigning a cumulative confidence score to each candidate analyte further comprises the step of:

5                    comparing the experimental analyte to the candidate analytes based on a retention index value.

10. . The method for ranking analytes of claim 9, wherein the retention index value is selected from retention index values from one or more dimensions.

10            11. The method for ranking analytes of claim 9, wherein the retention index value has passing range, a neutral criteria, and a failing range based on an acceptable retention index window, the passing range being within the acceptable retention index window, the neutral criteria requiring the retention index value of the candidate analyte to be unavailable, and the failing range being outside the acceptable retention index window.

15            12. The method for ranking analytes of claim 11, wherein a portion of the cumulative confidence score of a candidate analyte in the passing range is assigned a value of 1, a portion of the cumulative confidence score of a candidate analyte satisfying the neutral criteria is assigned a value of 0, and a portion of the cumulative confidence score of a candidate analyte in the failing range is assigned a value of 0.

20            13. The method for ranking analytes of claim 11, wherein the acceptable retention index window is W units.

25            14. The method for ranking analytes of claim 1, wherein the cumulative confidence score ranges from -6 to 4.

30            15. The method for ranking analytes of claim 1, wherein the candidate analytes are ranked in descending order of the cumulative confidence score of each candidate analyte.

16. The method for ranking analytes of claim 1, wherein the step of analyzing an experimental analyte occurs using a gas chromatography - mass spectrometer;

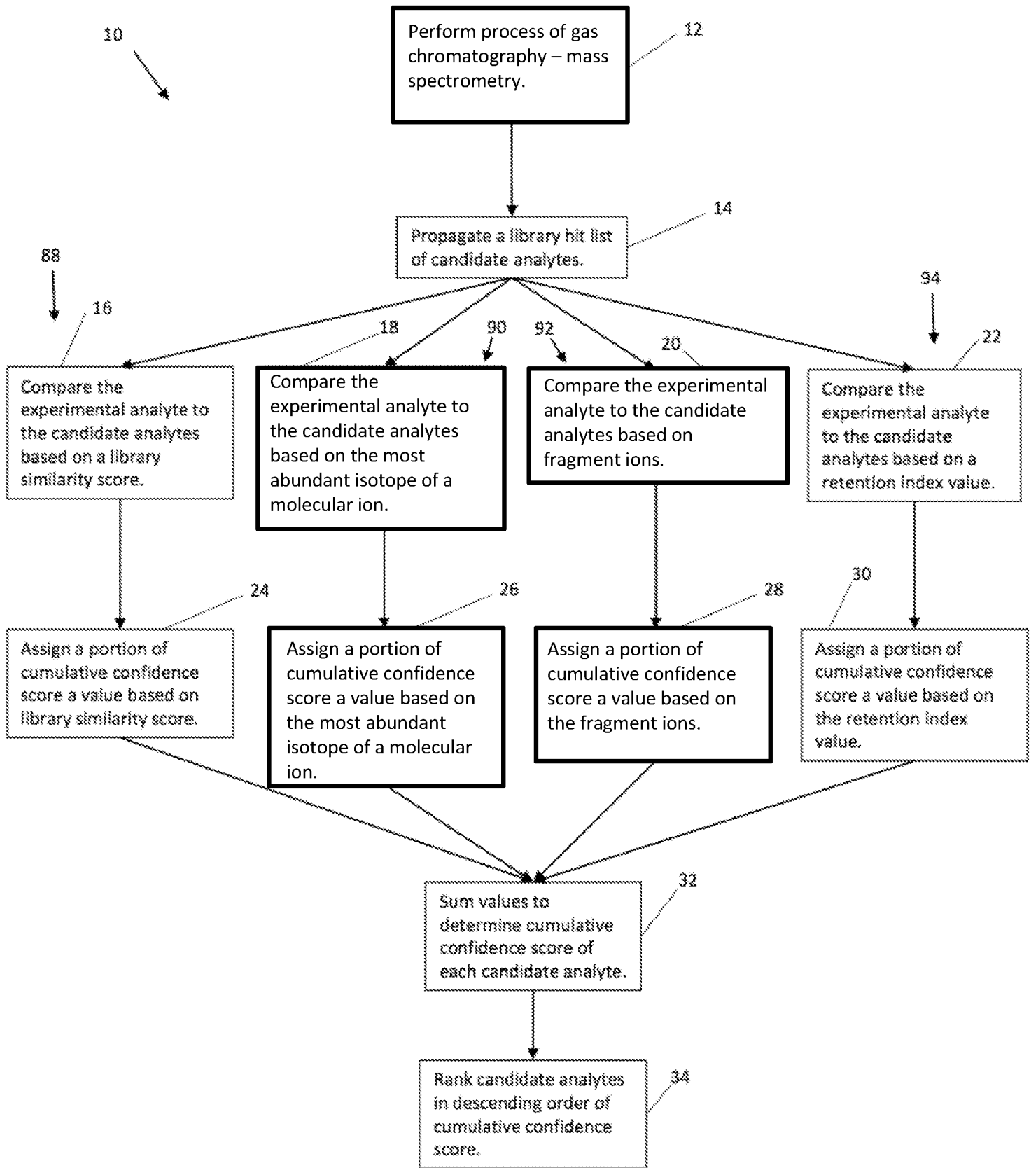


FIG. 1

88

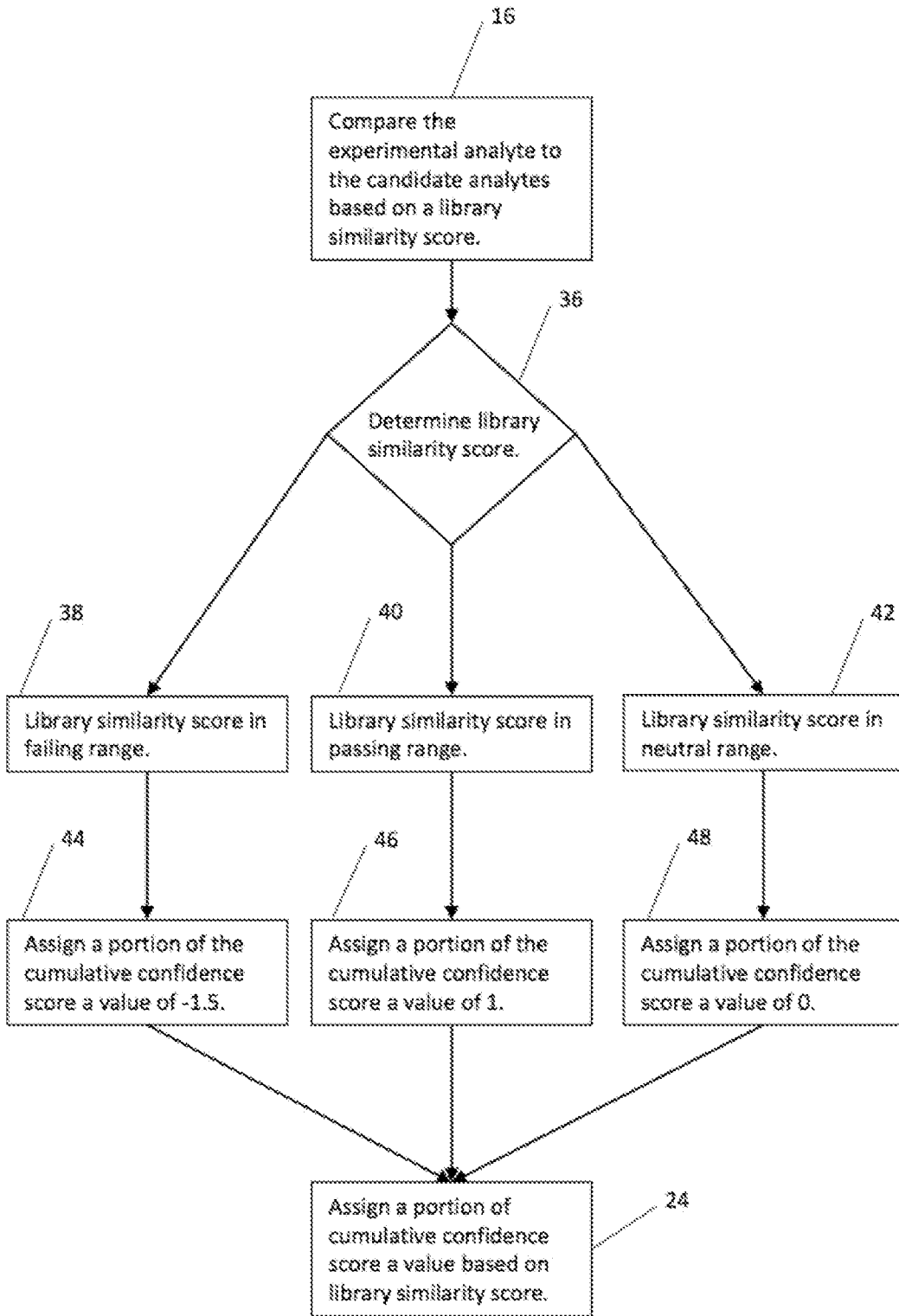


FIG. 2

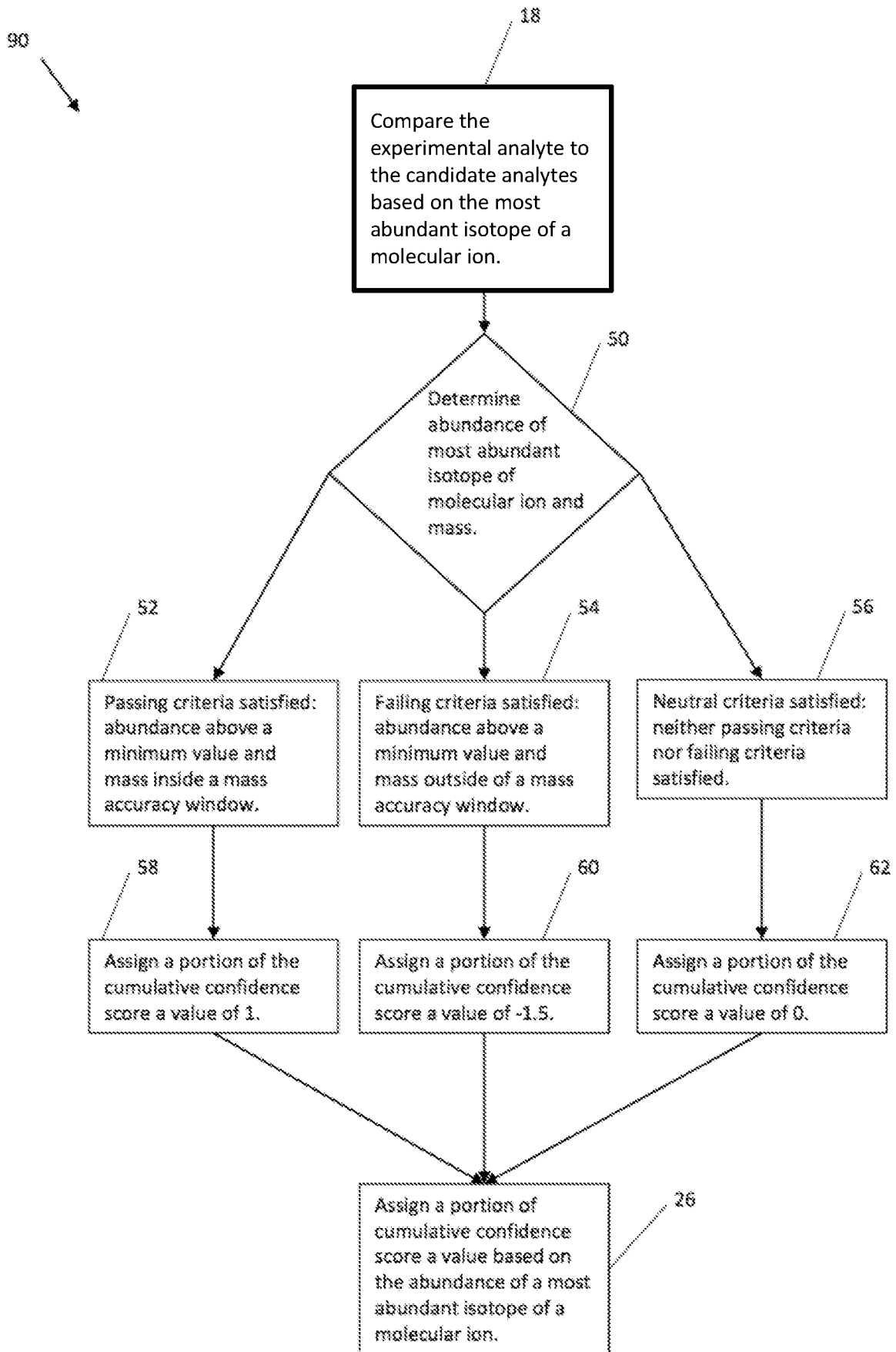


FIG. 3

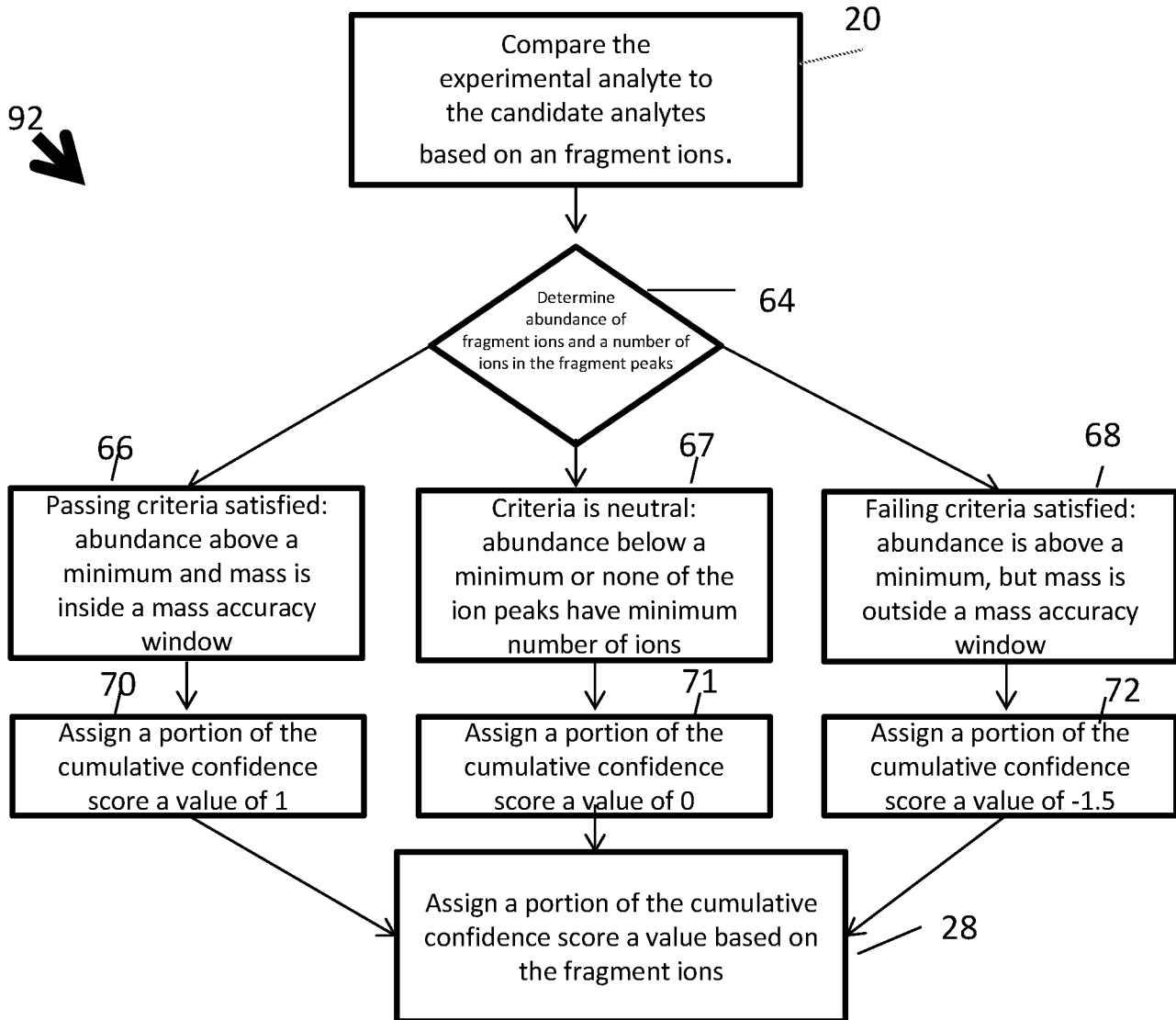


Fig. 4

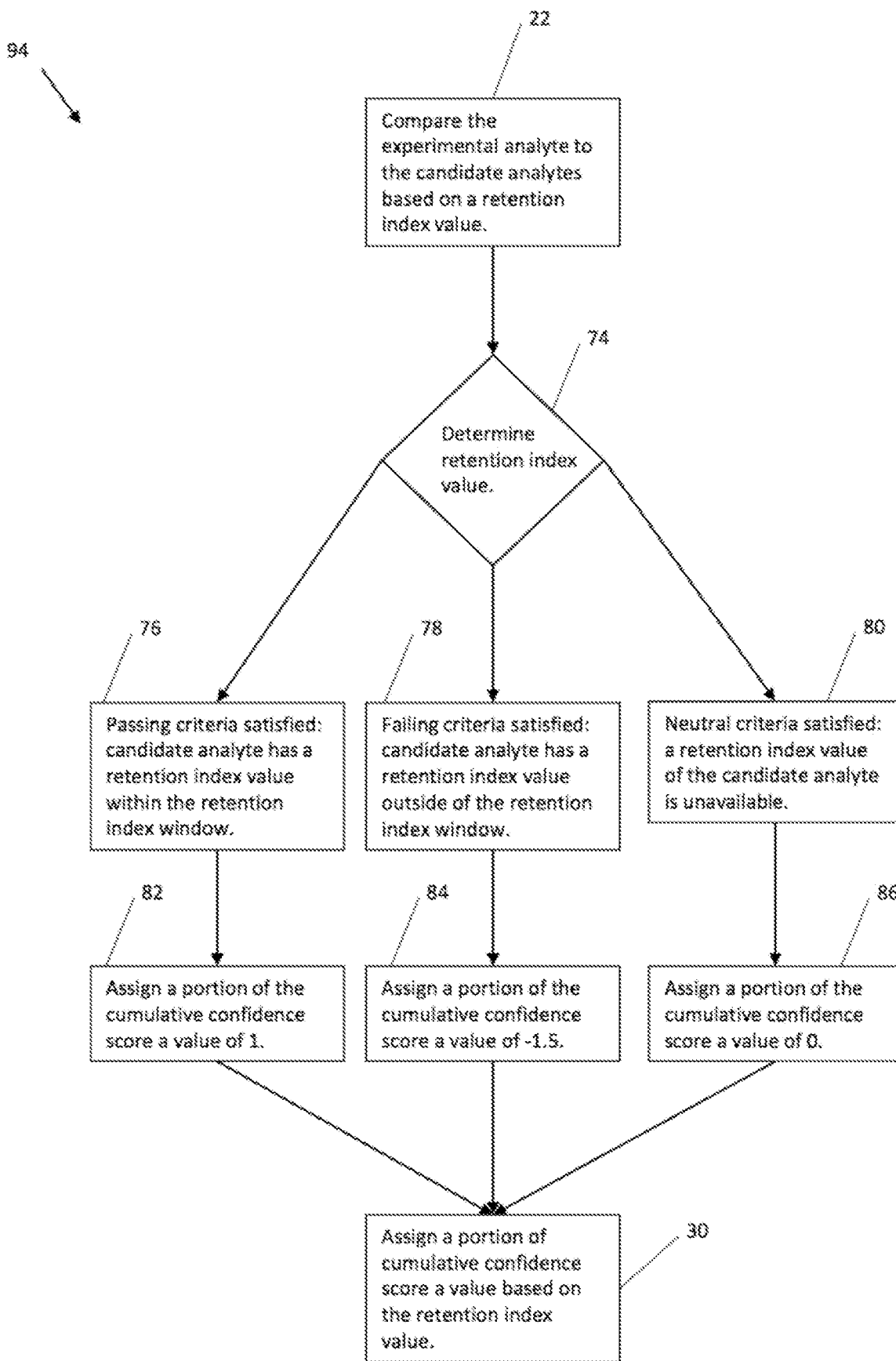


FIG. 5

**A. CLASSIFICATION OF SUBJECT MATTER**

G01N 30/72(2006.01)i, G06K 9/00(2006.01)i, G01N 30/02(2006.01)i

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

**B. FIELDS SEARCHED**Minimum documentation searched (classification system followed by classification symbols)  
G01N 30/72; G01N 27/62; H01J 49/00; G06K 9/00; G01N 30/02Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Korean utility models and applications for utility models  
Japanese utility models and applications for utility modelsElectronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
eKOMPASS(KIPO internal) & Keywords:experimental analyte, mass spectrometer, comparing, candidate analytes, assigning, cumulative confidence score, library similarity score, abundant isotope, molecular ion, fragment ion, retention index value, ranking**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2014-0297201 A1 (KNORR, A. et al.) 02 October 2014 See abstract; paragraphs [0032], [0033], [0134]-[0150]; claims 1-20; figures 1-18.	1-16
A	US 2015-0340216 A1 (WISCONSIN ALUMINI RESEARCH FOUNDATION) 26 November 2015 See the whole document.	1-16
A	KERBAR, A. et al., "CASE via MS: Ranking structure candidates by mass spectra", CROATICA Chemica Acta, 2006, Vol, 79, No. 3, pp.449-464 See the whole document.	1-16
A	FELL, L. et al., "Accurate Mass Spectral Database: Harnessing the Power of High Performance Mass Spectrometry at Long Last", Poster, In: Analysis & Separations from TECHNOLOGY NETWORKS, 16 June 2016, < <a href="https://www.technology-networks.com/analysis/posters/accurate-mass-spectral-database-harnessing-the-power-of-high-performance-mass-spectrometry-at-long-last-229369">https://www.technology-networks.com/analysis/posters/accurate-mass-spectral-database-harnessing-the-power-of-high-performance-mass-spectrometry-at-long-last-229369</a> > See the whole document.	1-16
A	JP 2005-083952 A (SHIMADZU CORP.) 31 March 2005 See the whole document.	1-16

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

\* Special categories of cited documents:

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

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

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

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

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

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

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

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

"&amp;" document member of the same patent family

Date of the actual completion of the international search

12 June 2019 (12.06.2019)

Date of mailing of the international search report

12 June 2019 (12.06.2019)

Name and mailing address of the ISA/KR

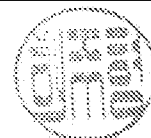
International Application Division  
Korean Intellectual Property Office  
189 Cheongsa-ro, Seo-gu, Daejeon, 35208, Republic of Korea

Facsimile No. +82-42-481-8578

Authorized officer

HEO, Joo Hyung

Telephone No. +82-42-481-8150



**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International application No.

**PCT/US2019/019354**

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2014-0297201 A1	02/10/2014	CN 103650100 A EP 2541585 A1 EP 2710621 A1 WO 2012-146787 A1	19/03/2014 02/01/2013 26/03/2014 01/11/2012
US 2015-0340216 A1	26/11/2015	CN 107077592 A EP 3123495 A1 US 10153146 B2 WO 2015-148941 A1	18/08/2017 01/02/2017 11/12/2018 01/10/2015
JP 2005-083952 A	31/03/2005	None	