



(12)发明专利

(10)授权公告号 CN 106461630 B

(45)授权公告日 2018.09.28

(21)申请号 201580029290.5

R · F · 邦纳

(22)申请日 2015.05.07

(74)专利代理机构 北京律盟知识产权代理有限公司 11287

(65)同一申请的已公布的文献号

代理人 关旭颖

申请公布号 CN 106461630 A

(51)Int.Cl.

G01N 30/86(2006.01)

(43)申请公布日 2017.02.22

G01N 30/88(2006.01)

(30)优先权数据

G01N 30/72(2006.01)

62/006,805 2014.06.02 US

(85)PCT国际申请进入国家阶段日

(56)对比文件

2016.11.30

US 2012/0108448 A1, 2012.05.03,

(86)PCT国际申请的申请数据

US 2004/0031918 A1, 2004.02.19,

PCT/IB2015/053351 2015.05.07

US 2007/0023633 A1, 2007.02.01,

(87)PCT国际申请的公布数据

CN 103109345 A, 2013.05.15,

W02015/186012 EN 2015.12.10

US 2013/0087701 A1, 2013.04.11,

(73)专利权人 DH科技发展私人贸易有限公司

US 2009/0173878 A1, 2009.07.09,

地址 新加坡新加坡

US 2011/0295521 A1, 2011.12.01,

(72)发明人 E · 杜霍斯拉夫 L · L · 伯顿

审查员 王方

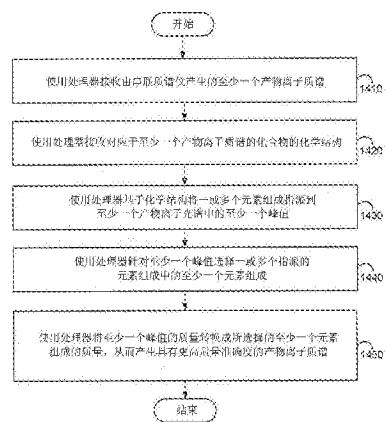
权利要求书3页 说明书12页 附图15页

(54)发明名称

用于将质谱库转换成准确质谱库的方法

(57)摘要

接收由串联质谱仪产生的至少一个产物离子质谱。接收对应于所述至少一个产物离子质谱的化合物的化学结构。使用处理器基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值。使用所述处理器针对所述至少一个峰值选择所述一或多个指派的元素组成中的至少一个元素组成。使用所述处理器将所述至少一个峰值的质量转换成所述所选择的至少一个元素组成的质量，从而产生具有更高质量准确度的产物离子质谱。



1. 一种用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的系统，其包括：

处理器，所述处理器

接收由串联质谱仪产生的至少一个产物离子质谱，

接收对应于所述至少一个产物离子质谱的化合物的化学结构，

基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值，所述指派通过以下步骤完成：根据所述化学结构中具有在所述至少一个峰值的质量的质量容限内的质量的元素计算所述一或多个元素组成，及将所述一或多个元素组成指派到所述至少一个峰值，

对一或多个指派的元素组成进行计分，

模拟所述化学结构的一或多个碎片以产生所述化学结构的多个子结构，

将所述多个子结构与所述一或多个指派的元素组成进行比较以找到一或多个匹配的子结构，

将所述一或多个匹配的子结构指派到它们相应的指派的元素组成，

对所述一或多个匹配的子结构进行计分，

组合所述一或多个指派的元素组成和它们相应的一或多个匹配的子结构的分值以产生所述一或多个指派的元素组成中的每一者的总分值，

选择所述一或多个指派的元素组成中的具有最高总分值的至少一个元素组成，及

将所述至少一个峰值的质量转换成所选择的至少一个元素组成质量，从而产生具有更高质量准确度的产物离子质谱。

2. 根据权利要求1所述的系统，其中基于包括以下规则中的一或多者的碎片化规则对所述一或多个匹配的子结构进行计分：碎片离子的所述元素组成与已知前体离子的组成一致、丢失与其前体一致、较高级的化学键比较低级的键更难断裂、及碳C与杂原子氮N、氧O、及硫S之间的化学键比C-C键更易断裂。

3. 根据权利要求1所述的系统，其中基于至少一个元素组成与所述至少一个峰值的所述质量之间的质量差值对所述一或多个指派的元素组成进行计分。

4. 根据权利要求2所述的系统，其中所述处理器进一步接收对应于所述至少一个产物离子质谱的至少一个数据收集条件，且所述碎片化规则包括使用所述至少一个数据收集条件的规则。

5. 根据权利要求4所述的系统，其中所述至少一个数据收集条件包括以下各者中的一或多者：极性、第一四极Q1分辨率、前体质荷比m/z、m/z误差分布、目标产物离子光谱Q1宽度、及碰撞能量。

6. 根据权利要求1所述的系统，其中在所述处理器将所述至少一个峰值的所述质量转换成所述所选择的至少一个元素组成的所述质量之后，所述处理器进一步将所述至少一个峰值的一或多个同位素峰值加入到具有更高质量准确度的所述产物离子质谱。

7. 一种用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法，其包括：

使用处理器接收由串联质谱仪产生的至少一个产物离子质谱；

使用所述处理器接收对应于所述至少一个产物离子质谱的化合物的化学结构；

使用所述处理器基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值，所述指派通过以下步骤完成：根据所述化学结构中具有在所述至少一个峰值的质量的质量容限内的质量的元素计算所述一或多个元素组成，及将所述一或多个元素组成指派到所述至少一个峰值；

使用所述处理器对一或多个指派的元素组成进行计分，

使用所述处理器模拟所述化学结构的一或多个碎片以产生所述化学结构的多个子结构，

使用所述处理器将所述多个子结构与所述一或多个指派的元素组成进行比较以找到一或多个匹配的子结构，

使用所述处理器将所述一或多个匹配的子结构指派到它们相应的指派的元素组成，

使用所述处理器对所述一或多个匹配的子结构进行计分，

使用所述处理器组合所述一或多个指派的元素组成和它们相应的一或多个匹配的子结构的分值以产生所述一或多个指派的元素组成中的每一者的总分值，

使用所述处理器选择所述一或多个指派的元素组成中的具有最高总分值的至少一个元素组成，及

使用所述处理器将所述至少一个峰值的质量转换成所选择的至少一个元素组成的质量，从而产生具有更高质量准确度的产物离子质谱。

8.一种存储有指令的非暂时性且有形计算机可读存储媒体，所述指令在由处理器执行时致使所述处理器通过以下步骤执行用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法：

提供系统，其中所述系统包括一或多个相异软件模块，且其中所述相异软件模块包括输入模块及分析模块；

使用所述输入模块接收由串联质谱仪产生的至少一个产物离子质谱；

使用所述输入模块接收对应于所述至少一个产物离子质谱的化合物的化学结构；

使用所述分析模块基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值，所述指派通过以下步骤完成：通过根据所述化学结构中具有在所述至少一个峰值的质量的质量容限内的质量的元素计算所述一或多个元素组成，及将所述一或多个元素组成指派到所述至少一个峰值；

使用所述分析模块对一或多个指派的元素组成进行计分，

使用所述分析模块模拟所述化学结构的一或多个碎片以产生所述化学结构的多个子结构，

使用所述分析模块将所述多个子结构与所述一或多个指派的元素组成进行比较以找到一或多个匹配的子结构，

使用所述分析模块将所述一或多个匹配的子结构指派到它们相应的指派的元素组成，

使用所述分析模块对所述一或多个匹配的子结构进行计分，

使用所述分析模块组合所述一或多个指派的元素组成和它们相应的一或多个匹配的子结构的分值以产生所述一或多个指派的元素组成中的每一者的总分值，

使用所述分析模块选择所述一或多个指派的元素组成中的具有最高总分值的至少一个元素组成；及

使用所述分析模块将所述至少一个峰值的质量转换成所选择的至少一个元素组成质量,从而产生具有更高质量准确度的产物离子质谱。

## 用于将质谱库转换成准确质谱库的方法

[0001] 相关申请案的交叉参考

[0002] 本申请案主张2014年6月2日申请的序列号为62/006,805的美国临时专利申请案的权益，所述临时专利申请案的全部内容以引用的方式并入本文中。

### 背景技术

[0003] 准确的质谱法/质谱法 (MS/MS) 光谱库匹配具有提高未知筛选工作流的效率的巨大潜力。然而，大规模的准确质谱库目前并不像标称或非准确质量库那样覆盖广泛的化学空间。建立准确的质谱库比较耗时 (例如，使用仪器的时间、化学品的可用性等等) 且并非总是可行的。

[0004] 已对准确质量数据的自动重新校准做了一些工作来改进准确质谱库的质量。此外，近期已开发用于仅从化合物结构 (在不存在任何MS数据的情况下) 自动产生理论MS<sub>n</sub>光谱的工具。然而，这些工具倾向于过度估计碎片。

### 发明内容

[0005] 本发明揭示一种用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的系统。所述系统包含处理器。所述处理器接收由串联质谱仪产生的至少一个产物离子质谱、接收对应于所述至少一个产物离子质谱的化合物的化学结构、及基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值。所述处理器进一步针对所述至少一个峰值选择所述一或多个指派的元素组成中的至少一个元素组成、及将所述至少一个峰值的质量转换成所述所选择的至少一个元素组成质量，从而产生具有高质量准确度的产物离子质谱。

[0006] 本发明揭示一种将产物离子质谱转换成具有高质量准确度的产物离子质谱的方法。使用处理器接收由串联质谱仪产生的至少一个产物离子质谱。使用所述处理器接收对应于所述至少一个产物离子质谱的化合物的化学结构。使用处理器基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值。使用所述处理器针对所述至少一个峰值选择所述一或多个指派的元素组成中的至少一个元素组成。使用所述处理器将所述至少一个峰值的质量转换成所述所选择的至少一个元素组成质量，从而产生具有高质量准确度的产物离子质谱。

[0007] 本发明揭示一种计算机程序产品，其包含非暂时性且有形计算机可读存储媒体，所述计算机可读存储媒体的内容包含具有指令的程序，所述指令在处理器上执行以便执行用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的方法。

[0008] 所述方法包含提供系统，其中所述系统包括一或多个相异软件模块，且其中所述相异软件模块包括输入模块及分析模块。所述输入模块接收由串联质谱仪产生的至少一个产物离子质谱。所述输入模块接收对应于所述至少一个产物离子质谱的化合物的化学结构。所述分析模块基于所述化学结构将一或多个元素组成指派到所述至少一个产物离子光谱中的至少一个峰值。所述分析模块针对所述至少一个峰值选择所述一或多个指派的元素

组成中的至少一个元素组成。所述分析模块将所述至少一个峰值的质量转换成所述所选择的至少一个元素组成质量,从而产生具有高质量准确度的产物离子质谱。

[0009] 本文陈述申请人的教示的这些特征及其它特征。

## 附图说明

[0010] 技术人员将理解下文描述的图式仅用于说明目的。不希望图式以任何方式限制本教示的范围。

[0011] 图1是说明可在其上实施本教示的实施例的计算机系统的方框图。

[0012] 图2是根据各种实施例的用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的示范性系统的示意图。

[0013] 图3是根据各种实施例的来自展示同量异位碎片的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0014] 图4是根据各种实施例的来自展示级联中性丢失的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0015] 图5是根据各种实施例的来自展示由两种不同类型的断裂键产生的碎片的碎片化评估工具的两个叠加的显示窗的信息的示范性屏幕捕获。

[0016] 图6是根据各种实施例的来自展示文拉法辛(Venlafaxine)的碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0017] 图7是根据各种实施例的来自展示同样具有132.0570的质量的文拉法辛的另一碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0018] 图8是根据各种实施例的来自展示具有193的标称质量且具有最高分值的7-氨基氯硝西泮(7-aminoclazepam)的碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0019] 图9是根据各种实施例的来自展示具有与图8中的最高分值碎片类似的结构的地西泮(diazepam)的碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获。

[0020] 图10是根据各种实施例的来自展示同样具有193的质量的替马西泮(temazepam)的两个碎片的化学结构的碎片化评估工具的两个叠加的显示窗的信息的示范性屏幕捕获。

[0021] 图11是根据各种实施例的展示如何通过将每一注释的碎片的理论同位素模式注入到光谱中来将光谱转换成准确的非指定碎片化光谱的一系列示范性的质谱标绘图。

[0022] 图12是根据各种实施例的针对肾上腺素的示范性标称或低准确度产物离子质谱。

[0023] 图13是根据各种实施例的针对肾上腺素的示范性转换的准确产物离子质谱。

[0024] 图14是展示根据各种实施例的用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的方法的流程图。

[0025] 图15是根据各种实施例的包含一或多个相异软件模块的执行用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的方法的系统的示意图。

[0026] 在详细描述本教示的一或多个实施例之前,所属领域的技术人员应了解,本教示在其应用上不受限于下文详细描述中阐释或图式中说明的构造细节、组件布置及步骤布置。此外,应理解,本文使用的短语及术语用于描述目的且不应被视为具限制性。

## 具体实施方式

[0027] 图1是说明计算机系统100的框图,本教示的实施例可实施于计算机系统100上。计算机系统100包含:总线102或用于传达信息的其它通信机构;及与总线102耦合以用于处理信息的处理器104。计算机系统100还包含存储器106,其可为随机存取存储器(RAM)或其它动态存储装置,存储器106耦合到总线102以用于存储待由处理器104执行的指令。存储器106还可用于在执行待由处理器104执行的指令期间存储临时变量或其它中间信息。计算机系统100进一步包含只读存储器(ROM)108或耦合到总线102以用于存储用于处理器104的静态信息及指令的其它静态存储装置。存储装置110(例如磁盘或光盘)经提供且耦合到总线102以用于存储信息及指令。

[0028] 计算机系统100可经由总线102耦合到显示器112(例如阴极射线管(CRT)或液晶显示器(LCD))以将信息显示给计算机用户。包含字母数字键及其它键的输入装置114耦合到总线102以将信息及命令选择传达到处理器104。另一类型的用户输入装置是光标控制件116,例如用于将方向信息及命令选择传达到处理器104及用于控制光标在显示器112上移动的鼠标、轨迹球或光标方向键。此输入装置通常在两个轴(第一轴(即,x)及第二轴(即,y))上具有两个自由度,这允许所述装置在平面中指定位置。

[0029] 计算机系统100可执行本教示。与本教示的某些实施方案一致,由计算机系统100响应于处理器104执行包含于存储器106中的一或多个指令的一或多个序列而提供结果。此类指令可从另一计算机可读媒体(例如存储装置110)读取到存储器106中。包含于存储器106中的指令序列的执行致使处理器104执行本文描述的过程。替代地,可使用硬接线电路替代软件指令或与软件指令组合以实施本教示。因此,本教示的实施方案不受限于硬件电路及软件的任何特定组合。

[0030] 在各种实施例中,计算机系统100可跨越网络连接到一或多个其它计算机系统(如计算机系统100)以形成网络化系统。网络可包含私有网络或公共网络,例如因特网。在网络化系统中,一或多个计算机系统可存储数据且将所述数据发送到其它计算机系统。可将存储及发送数据的一或多个计算机系统称为服务器,或在云计算场景中称为云。举例来说,一或多个计算机系统可包含一或多个网页服务器。举例来说,可将数据发送到服务器或云及从服务器或云接收数据的其它计算机系统称为客户机或云装置。

[0031] 如本文使用的术语“计算机可读媒体”是指参与将指令提供到处理器104以供执行的任何媒体。此类媒体可呈许多形式,包含(但不限于)非易失性媒体、易失性媒体及传输媒体。非易失性媒体包含(例如)光盘或磁盘,例如存储装置110。易失性媒体包含动态存储器,例如存储器106。传输媒体包含同轴电缆、铜线及光纤,其包含包括总线102的导线。

[0032] 计算机可读媒体或计算机程序产品的常见形式包含(例如)软磁盘、软盘、硬盘、磁带、或任何其它磁性媒体、CD-ROM、数字视频光盘(DVD)、蓝光光盘、任何其它光学媒体、指状驱动器、存储卡、RAM、PROM、及EPROM、FLASH-EPROM、任何其它存储器片或存储器匣、或计算机可从其读取的任何其它有形媒体。

[0033] 在将一或多个指令的一或多个序列载送到处理器104以供执行时可涉及到各种形式的计算机可读媒体。举例来说,指令最初可携载在远程计算机的磁盘上。远程计算机可将指令加载到其动态存储器中并使用调制解调器通过电话线发送指令。计算机系统100本地

的调制解调器可接收电话线上的数据并使用红外线发射器来将数据转换为红外线信号。耦合到总线102的红外线检测器可接收红外线信号中载送的数据并将数据置于总线102上。总线102将数据载送到存储器106，处理器104从存储器106检索指令并执行指令。由存储器106接收到的指令可任选地在由处理器104执行之前或之后存储在存储装置110上。

[0034] 根据各种实施例，被配置为由处理器执行以执行方法的指令存储在计算机可读媒体上。计算机可读媒体可为存储数字信息的装置。举例来说，计算机可读媒体包含所属领域已知的用于存储软件的光盘只读存储器(CD-ROM)。计算机可读媒体由适于执行经配置以被执行的指令的处理器来存取。

[0035] 已出于说明及描述的目的呈现本教示的各种实施方案的以下描述。所述描述并非为详尽的且不将本教示限于所揭示的精确形式。鉴于上文教示修改及变化为可能的，或可从实践本教示来获得修改及变化。另外，虽然所描述的实施方案包含软件，但本教示可被实施为硬件及软件的组合或单独以硬件实施。可使用面向对象编程系统及非面向对象编程系统两者来实施本教示。

#### [0036] 用于转换质量的系统及方法

[0037] 串联质谱法或质谱法/质谱法(MS/MS)用来通过使从实验获得的产物离子光谱与从已知化合物的可信标准样本获得的参考产物离子光谱匹配来识别未知化合物。测量的质量反映通过将标准样本的前体离子碎片化产生的产物离子的元素组成及结构。通常仅选择单一同位素形式的前体离子，使得产物离子光谱不含有关于同位素峰值的信息。参考产物离子光谱的集合称为库，且通常将产物离子光谱的库的质量及强度存储于数据库中。

[0038] 通过比较不同光谱中的峰值的质量及强度执行匹配。因为始终存在与质量测量相关联的误差，所以质量比较使用容限窗以决定两个质量是否相同。然而，如果容限窗较大，那么可能出现的情况是：具有不同但类似的质量的离子可能被不正确地假设为匹配。因此，匹配过程的准确度取决于质量在实验光谱及参考光谱两者中的测量准确度。

[0039] 已使用具有大约0.1质量单位(amu或道尔顿(Da))的质量测量准确度的串联质谱仪(例如，三重四极仪器)及在类似仪器上获得的实验光谱产生许多库。

[0040] 近来，使用具有高得多的准确度(例如，0.01或0.001amu)的仪器来测量质谱已变得越来越常见。虽然来自这些质谱仪的实验光谱可与在较低准确度装置上获得的库光谱匹配，但如果库也含有高准确度光谱，那么极大地改进光谱比较的执行。通常，产生较高准确度库需要使用不同的高准确度仪器重新分析标准样本，但这比较耗时，且(例如)如果标准样本不再可用那么这可能是不可行的。

[0041] 在各种实施例中，使用系统及方法来改进现有库的准确度，即，以高准确度等效物或计算出的精确质量来取代标称或较低准确度质量值，同时保持参考光谱的强度比。

[0042] 这可在可确定碎片离子的结构的情况下实现，因为这提供可根据其计算精确质量的元素组成且精确质量用于取代参考库中的较低准确度值。有可能通过将结构指派到观测到的碎片离子手动地注释光谱，但这比较耗时，且需要渊博的知识及经验。

[0043] 在各种实施例中，算法基于已知化合物的化学结构的计算机可读表示产生硅中碎片的全部潜在元素组成，但此类算法通常预测太多碎片。化学结构的计算机可读表示是(例如)MOL文件，或可以SMILES注释表示。此外，因为库光谱具有较低准确度，两个或两个以上预测的碎片通常具有在参考数据的容限内匹配观测到的碎片的质量。

[0044] 在各种实施例中,算法进一步基于碎片化规则对预测的碎片进行计分。碎片化规则可包含(但不限于)以下规则:例如碎片离子的元素组成必须与(已知)前体离子的组成一致、丢失(通常易于指派)也必须与其前体一致、碳(C)与杂原子(例如,氮(N)、氧(O)及硫(S))之间的化学键比C-C键更易断裂、环比线性结构更难断裂(尤其是,如果环也为芳香族的)。另外,如果存在化合物族(例如,作为常见架体结构的衍生物的药物),指派到族的一或多个成员的碎片的结构很可能与族的第二成员中具有相同质量的碎片相同。

[0045] 基于这些分值,使用那个产物离子的最可能的结构注释每一产物离子光谱的每一产物离子,确定那个结构的元素组成,且使用计算出的那个元素组成的精确质量更新参考库质量。

[0046] 此外,如果碎片具有不能分辨的两个或两个以上可能结构,可对两者进行计分。换句话说,如果应用硅中碎片化及碎片化规则导致产物离子的两个或两个以上可能的精确质量,那么可使参考库对产物离子的两个或两个以上质量值进行计分。

[0047] 在各种实施例中,随后光谱匹配算法经修改以选择最接近于实验质量的质量。举例来说,参考库可接收化合物的产物离子光谱的峰值的99.9及100.1的质量且对其进行计分。通过上文描述的硅中碎片化及碎片化规则发现这些值。接着,如果稍后发现99.91的实验值,那么匹配算法使用最接近于实验质量的精确质量,且忽略任何替代精确质量值。在此情况下,最接近于实验质量的精确质量是99.9。在各种实施例中,从参考库移除替代精确质量值而非修改匹配算法。

[0048] 在各种实施例中,根据产物离子的元素组成,可计算预期的同位素模式(质量及强度),且可将这些额外峰值加入到希望与有意使用较宽的前体离子窗使得同位素被包含在内的技术一起使用的库。有意使用较宽的前体离子窗的此类技术包含(但不限于)数据独立获取(DIA)及循序窗口化获取(SWATH)。

[0049] 图2是根据各种实施例的用于将产物离子质谱转换成具有高质量准确度的产物离子质谱的示范性系统200的示意图。系统200包含处理器220。处理器220接收产物离子光谱211及对应于产物离子光谱211的已知化合物的化学结构212。举例来说,产物离子光谱211来自具有较低准确度质量值的现有库。

[0050] 也可从现有库获得化学结构212。通常,所属领域的一般技术人员将现有库想象为含有光谱及(通常)化学结构的单个数据库。然而,也可从存储化学结构的某一计算机目录或其中响应于化合物识别符(名称,等等)获得结构的化学结构的可搜索数据库获得化学结构212。

[0051] 除产物离子光谱211及化学结构212外,算法200还可从现有库接收补充信息213作为输入。补充信息213可包含(但不限于)数据收集条件,例如极性、Q1分辨率、前体m/z、m/z误差分布、目标产物离子光谱Q1宽度、及碰撞能量。

[0052] 处理器220通过将一或多个元素组成指派到质量峰值及针对质量峰值选择至少一个元素组成将产物离子光谱211的质量峰值转换成更准确或精确的质量峰值。处理器220可以至少两种不同方式执行此转换。

[0053] 首先,处理器220可执行化学结构211的硅中碎片化,且比较每一模拟碎片的元素组成质量与产物离子光谱211的每一质量峰值。将质量在产物离子光谱211的质量峰值的质量的质量容限内的碎片的元素组成指派到质量峰值。还可给予指派的元素组成某一分

值。分值可为基于考虑(例如)断裂的键的数目、断裂的键的类型(鉴于用于碎片化的CE)、内部键的类型、级联碎片化的证据、氢迁移、重新布置及来自具有类似结构的化合物的产物离子光谱中的碎片的证据的碎片化规则。

[0054] 接着,举例来说,针对产物离子光谱211的质量峰值基于最高分值选择至少一个指派的元素组成。接着,将质量峰值的质量转换成所选择的元素组成的较高准确度或精确质量。在转换产物离子光谱211的全部质量峰值之后,处理器220输出较高准确度的产物离子光谱230。

[0055] 处理器220也可通过首先基于化学结构212中的元素的数目及类型针对产物离子光谱211中的每一质量峰值计算全部可能元素组成来转换产物离子光谱211的质量。因此,处理器220将一或多个元素组成指派到产物离子光谱211的质量峰值。还可给予这些指派的元素组成某一分值。分值是(例如)基于质量或m/z误差的组分值。在各种实施例中,可确定产物离子光谱211的一些质量峰值为基于Q1宽度的同位素峰值,且可给予其适当的元素组成。

[0056] 然而,仅基于质量针对每一质量峰值指派元素组成可导致对一些峰值来说太多的元素组成,从而难以选择一个元素组成。因此,在各种实施例中,处理器220另外执行如上文描述的化学结构211的硅中碎片化。然而,在此情况中,替代比较碎片与质量峰值,比较碎片与已指派到每一质量峰值的元素组成。如上文描述,还可基于碎片化规则给予碎片某一碎片化分值。因此,为了选择元素组成,组合组分值及碎片分值以提供每一指派的元素组成的总分值。

[0057] 如在上文描述的另一转换方法中,接着,举例来说,针对产物离子光谱211的质量峰值,基于最高分值选择至少一个指派的元素组成。接着,将质量峰值的质量转换成所选择的元素组成的较高准确度或精确质量,且在转换产物离子光谱211的全部质量峰值之后,处理器220输出较高准确度的产物离子光谱230。

[0058] 将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法可改进对用于较小分子定性工作(例如,筛选及识别工作流)中的内部及公共光谱库的可访问性及其质量。其可提供用于使现有质谱库的兼容性限制与准确质量数据桥接的容易选项。其还可解决使用单位Q1分辨率收集到的光谱库与使用非指定前体离子选择收集到的数据之间的差异。

[0059] 众所周知的蛋白质组学技术也使用碎片预测方法。在此蛋白质组学技术中,一组蛋白质是硅中酶,其经消化以形成肽,且预测其碎片,且将其存储于理论光谱数据库中。比较实验产物离子光谱与此数据库,以便确定哪些肽存在于样本中。肽碎片化是简单且好理解的,所以碎片质量预测非常准确,尽管不能保证特定碎片将形成及不能预测碎片强度。相比之下,从较小分子的碎片化产生的碎片难以预测,且强度比率是匹配算法的重要部分,因此有利的是,更新可信的产物离子光谱库的质量。同样相比之下,理论肽光谱不用于增加先前存储的实验光谱的质量准确度。

[0060] 近来,已关注建立可信肽的参考光谱库,其可能已以实验形式合成或观测,且因此确定观测到的碎片的强度比率。如果以低准确度产生这些光谱,那么本文描述的技术还可用于改进所测量的蛋白质组学数据的质量的准确度。

[0061] 因此,在各种实施例中,使用计算机将标称或非准确质量产物离子光谱库自动转

换成准确质量产物离子光谱库,使得在比较未知准确质谱与此类库时,最小公分母是准确质量值。可将此类转换称为硅中,这意味着在计算机或处理器中或由计算机或处理器执行转换。

[0062] 在各种实施例中,至少三条信息用于转换。第一条信息是化合物结构本身。第二条信息是现有光谱库中的数据,其包含数据收集条件(例如,极性、Q1分辨率及碰撞能量)。第三条信息是一或多个实验量身定制的母体化合物结构的硅中碎片化规则。举例来说,调整光谱库中的m/z值(同时保持相对碎片强度)以将其转换成在相同离子化及CID条件下使用准确质量仪器将收集到对应的准确质量。

[0063] 在各种实施例中,为了将质谱成功转换成准确质谱,充分利用光谱库中的补充信息(如果可用的话)。这可通过存储推定的光谱碎片及具有支持信息(例如,m/z误差、分值、环加双键(RDB)、氢迁移或断裂的键的类型)的中性丢失注释来完成。

[0064] 在各种实施例中,为了分辨是同量异位的或具有在低准确度数据或库光谱的准确度内的相同质量的预测碎片,使用对碎片的可能性进行计分来筛选出较不可能的条目。以至少四种不同方式扩大现有碎片计分(基于m/z误差、奇/偶电子、用于产生碎片的断裂的键的数目及类型、氢迁移)。

[0065] (i) 使用来自级联碎片的证据以对同量异位碎片进行计分且筛选同量异位碎片(即,可通过就前一碎片而言的常见中性丢失的差异解释碎片存在)。

[0066] 图3是根据各种实施例的来自展示同量异位碎片的碎片化评估工具的显示窗的信息的示范性屏幕捕获300。屏幕捕获300展示丁螺环酮的两个较高分值的同量异位预测碎片310及320。屏幕捕获300还展示丁螺环酮的产物离子光谱340、突出碎片310的子结构350的丁螺环酮的化学结构、及突出碎片320的子结构360的丁螺环酮的化学结构。碎片310及320的精确质量对应于在0.002amu内所测量的质量222.1476。

[0067] 图4是根据各种实施例的来自展示级联中性丢失的碎片化评估工具的显示窗的信息的示范性屏幕捕获400。屏幕捕获400展示可通过常见中性丢失的差异解释碎片的存在。换句话说,通过跟踪来自潜在级联中性丢失的信息及其对碎片分值的贡献,可筛选出不太可能的同量异位碎片指派。

[0068] 举例来说,屏幕捕获400展示突出图3的子结构350的丁螺环酮的产物离子光谱440及化学结构。中性丢失410在产物离子光谱440中具有从峰值460到峰值470的54.0447Da的丢失。因为图3的子结构350可包含中性丢失410。图3的碎片310比图式的碎片320更可能。因此,图3及4展示可如何使用中性丢失信息对同量异位碎片进行计分及筛选。

[0069] (ii) 使用碎片稳定性以对同量异位碎片进行计分及筛选(用于产生碎片的某些断裂的键的类型比其它键类型需要更少的能量)。

[0070] 图5是根据各种实施例的来自展示由两种不同类型的断裂键产生的碎片的碎片化评估工具的两个叠加的显示窗的信息的示范性屏幕捕获500。显示窗510及显示窗520两者分别显示比索洛尔511及522的化学结构。然而,显示窗510展示由3个C杂原子键被断裂产生的比索洛尔515的碎片的化学结构。相比之下,显示窗520展示由2个芳香族键被断裂产生的比索洛尔525的碎片的化学结构。尽管与结构515相比,需要断裂更少的键来产生结构525,但结构525实际上更不可能。这是因为芳香族环键比C杂原子键稳定得多。

[0071] 图6是根据各种实施例的来自展示文拉法辛的碎片的化学结构的碎片化评估工具

的显示窗的信息的示范性屏幕捕获600。屏幕捕获600展示具有132.0570的质量的文拉法辛610的碎片的化学结构。屏幕捕获600还展示文拉法辛的产物离子光谱640及化学结构650。

[0072] 图7是根据各种实施例的来自展示同样具有132.0570的质量的文拉法辛的另一碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获700。屏幕捕获700展示具有132.0570的质量的文拉法辛的碎片720的化学结构。

[0073] 图6的碎片结构610及图7的碎片结构720是文拉法辛的同量异位碎片。图6的结构610更稳定,且因此更可能,这是因为与C-C键相比,更易于使C-N键断裂。换句话说,图7的结构720包含含有未断裂的C-N键的缺陷且因此是较不可能的。

[0074] (iii) 基于CID条件使用碎片化规则以对同量异位碎片进行计分及筛选(例如,在较低CE下不可能断裂芳香族键)。也参见图5到7。

[0075] (iv) 在可能的情况下,使用来自针对光谱库中的给定子结构所研究及搜索的化学空间及对应实验数据及其指派的碎片证据以对同量异位碎片进行计分及筛选。

[0076] 举例来说,西洋(zepam)化合物的化学空间包含7-氨基氯硝西洋、地西洋及替马西洋。全部西洋化合物具有193Da下的碎片。为了确定193Da下的碎片,比较193Da下或大约193Da的7-氨基氯硝西洋、地西洋及替马西洋的碎片。

[0077] 图8是根据各种实施例的来自展示具有193的标称质量且具有最高分值的7-氨基氯硝西洋的碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获800。屏幕捕获800展示具有193Da的m/z的7-氨基氯硝西洋的最高分值碎片810 ( $C_{13}H_9N_2$ )。屏幕捕获800还展示7-氨基氯硝西洋的产物离子光谱820、7-氨基氯硝西洋的化学结构830、及7-氨基氯硝西洋的碎片810的碎片子结构840。

[0078] 图9是根据各种实施例的来自展示具有与图8中的最高分值碎片类似的结构的地西洋的碎片的化学结构的碎片化评估工具的显示窗的信息的示范性屏幕捕获900。屏幕捕获900展示具有193Da的m/z的地西洋的碎片910 ( $C_{13}H_9N_2$ )。屏幕捕获900展示地西洋的产物离子光谱920、地西洋的化学结构930及地西洋的碎片910的碎片子结构940。屏幕捕获900还展示碎片950 ( $C_{14}H_{11}N$ ) ,其具有比碎片910 ( $C_{13}H_9N_2$ ) 更高的分值。

[0079] 图10是根据各种实施例的来自展示同样具有193的标称质量的替马西洋的两个碎片的化学结构的碎片化评估工具的两个叠加的显示窗的信息的示范性屏幕捕获1000。显示窗1010展示替马西洋的碎片1011 ( $C_{13}H_9N_2$ ) 的化学结构。显示窗1020展示替马西洋的不同碎片1021 ( $C_{10}H_{13}N_2O_2$ ) 的化学结构。

[0080] 图8到10展示彼此非常类似的7-氨基氯硝西洋、地西洋及替马西洋的化学结构。另外,全部三个化合物的光谱都具有在m/z 193.08下的碎片。为了一致性,此碎片的结构应在全部四种情况中相同或至少十分类似是有意义的。参看图8,7-氨基氯硝西洋的综合最高分值碎片是碎片810 ( $C_{13}H_9N_2$ )。所以如果假设这是此分子的正确指派,那么即使碎片950 ( $C_{14}H_{11}N$ ) 具有更高分值,图8中的地西洋的正确指派也应为类同的指派,即,碎片910 ( $C_{13}H_9N_2$ )。

[0081] 图10更清楚地展示正确碎片的选择。碎片1021有可能不正确,这是因为其化学结构并不分别与图8的化学结构840及图9的化学结构940(推定的正确碎片)十分类似。举例来说,图10的碎片1021在结构的“底部”处不具有苯环,但图8的化学结构840及图9的化学结构940具有苯环。然而,图10的碎片1011类似于图8的化学结构840及图9的化学结构940,且因

此,更有可能是正确的。

[0082] 在各种实施例中,当不能将一个唯一的碎片或组成指派到给定碎片(即,两种可能性具有类似分值)时,可使用多种可能性注释碎片。

[0083] 在各种实施例中,使用元素组成及任何潜在子结构块注释碎片。留存具有较高分值但不具有子结构的元素组成以允许未预料的碎片化,例如重新布置。

[0084] 在各种实施例中,一旦将正确的元素组成指派到MS/MS或产物离子光谱中的峰值,就通过将每一注释的碎片的理论同位素模式注入到光谱中来将使用近似单位Q1分辨率收集到的光谱转换成非指定碎片光谱。

[0085] 图11是根据各种实施例的展示如何通过将每一注释的碎片的理论同位素模式注入到光谱中来将光谱转换成准确的非指定碎片化光谱的一系列示范性质谱标绘图1100。标绘图1110展示从准确的实验非指定碎片化方法产生的示范性非指定碎片化光谱的部分。因为使用了非指定的碎片化方法,所以碎片或产物离子具有全同位素模式。如果使用较窄的前体离子窗或较窄的Q1窗,那么情况将不会是如此。标绘图1120中的化合物具有溴原子,其具有相异的同位素模式。模式包含具有由2Da分离的大致相等强度的两个同位素。在图11中将两个同位素展示为峰值1111及1112。使用较窄的前体窗获取的典型的产物离子光谱将仅展示这些同位素中的一者。

[0086] 标绘图1120展示示范性库光谱的部分。光谱展示已知化合物的碎片的低准确度质量383。使用较窄的前体离子窗获取标绘图1120的光谱。因此,标绘图1120的光谱仅展展示已知化合物的一个同位素峰值1121。

[0087] 在各种实施例中,在将标称或低准确度质量库光谱转换到准确质量库光谱期间,将理论同位素质量或处理器产生的同位素质量加入回到高准确度质量库光谱。

[0088] 标绘图1130展示示范性转换的准确质量库光谱的部分。比较标绘图1120及1130展示将标绘图1120中的质量库光谱的质量峰值1121转换成标绘图1130中的准确质量库光谱中的准确质量峰值1131。另外,将理论同位素质量峰值1132加入到标绘图1130中的高准确度质量库光谱。

[0089] 接着,举例来说,可使用标绘图1130中的准确的质量库光谱确定已知化合物是否在例如标绘图1110中所展示的任何非指定的碎片化光谱中。举例来说,针对标绘图1130中的高准确度质量库光谱搜索标绘图1110的非指定碎片化光谱。

#### [0090] 用于转换产物离子质谱的系统

[0091] 各种实施例包含根据各种实施例的用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的系统。此系统包含经配置以处理获取后的串联质谱数据的处理器。处理器可为(但不限于)图1的计算机、微处理器、计算机系统、图2的处理器或能够处理数据及发送及接收数据的任何装置。

[0092] 处理器接收由串联质谱仪产生的至少一个产物离子质谱。举例来说,串联质谱仪是低准确度串联质谱仪。举例来说,处理器从低准确度光谱库接收至少一个产物离子质谱。

[0093] 图12是根据各种实施例的针对肾上腺素的示范性标称或低准确度产物离子质谱1200。产物离子质谱1200包含分别具有91Da及120Da的质量的峰值1210及1220。

[0094] 除至少一个产物离子质谱外,处理器接收对应于至少一个产物离子质谱的化合物的化学结构。在各种实施例中,也从光谱库接收对应于至少一个产物离子质谱的化合物的

化学结构。在其它实施例中,可从另一库或数据库接收化学结构。

[0095] 处理器基于化学结构将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值。处理器针对至少一个峰值选择一或多个指派的元素组成的至少一个元素组成。最后,处理器将至少一个峰值的质量转换成所选择的至少一个元素组成质量,从而产生具有更高质量准确度的产物离子质谱。

[0096] 图13是根据各种实施例的针对肾上腺素的示范性转换的准确产物离子质谱1300。转换的准确产物离子质谱1300包含分别具有91.0369 Da及120.1252 Da的准确质量的峰值1310及1320。

[0097] 如上文描述,处理器可以多种方式指派及选择元素组成。在各种实施例中,处理器通过以下操作将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值:模拟产生化学结构的一或多个子结构的化学结构的一或多个碎片化、及将具有在至少一个峰值的质量的质量容限内的质量的一或多个子结构的至少一个元素组成指派到至少一个峰值。接着,举例来说,处理器通过以下操作选择至少一个元素组成:对一或多个指派的元素组成进行计分、及选择具有最高分值的至少一个元素组成。可基于碎片化规则进行计分。

[0098] 在替代实施例中,处理器通过以下操作将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值:根据化学结构中具有在至少一个峰值的质量的质量容限内的质量的元素计算一或多个元素组成、及将一或多个元素组成指派到至少一个峰值。可以数种方式选择至少一个元素组成。

[0099] 在各种实施例中,处理器通过以下操作选择至少一个元素组成:对一或多个指派的元素组成进行计分、及选择具有最高分值的至少一个元素组成。举例来说,基于至少一个元素组成与至少一个峰值的质量之间的质量差值对一或多个指派的元素组成进行计分。

[0100] 在各种实施例中,处理器基于两个分值选择至少一个元素组成。如上文提及,处理器通过对一或多个指派的元素组成进行计分来选择至少一个元素组成。另外,处理器模拟产生化学结构的一或多个子结构的化学结构的一或多个碎片化、将具有在至少一个峰值的质量的质量容限内的质量的一或多个子结构指派到至少一个峰值、及对一或多个子结构进行计分。最后,处理器组合指派的子结构与其对应元素组成的分值、及选择具有最高组合分值的指派的子结构的对应元素组成。

[0101] 在各种实施例中,基于至少一个元素组成与至少一个峰值的质量之间的质量差值对一或多个指派的元素组成进行计分,且基于碎片化规则对一或多个子结构进行计分。

[0102] 在各种实施例中,处理器进一步接收对应于至少一个产物离子质谱的至少一个数据收集条件。举例来说,收集条件由碎片化规则在计分期间使用。收集条件可包含(但不限于)极性、第一四极Q1分辨率、前体质荷比( $m/z$ )、 $m/z$ 误差分布、目标产物离子光谱Q1宽度、及碰撞能量中的一或者者。

[0103] 在各种实施例中,在处理器将至少一个峰值的质量转换成所选择的至少一个元素组成质量之后,所述处理器进一步将至少一个峰值的一或多个同位素峰值加入到具有更高质量准确度的产物离子质谱。

[0104] 在各种替代实施例中,处理器将至少一个峰值的一或多个同位素峰值加入到具有更高质量准确度的产物离子质谱而不将至少一个峰值的质量转换成所选择的至少一个元素组成质量。换句话来说,处理器接收由串联质谱仪产生的至少一个产物离子质谱。处理

器接收对应于至少一个产物离子质谱的化合物的化学结构。处理器基于化学结构将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值。处理器针对至少一个峰值选择一或多个指派的元素组成中的至少一个元素组成。然而，替代将至少一个峰值的质量转换成所选择的至少一个元素组成的质量，处理器将至少一个峰值的一或多个同位素峰值加入到产物离子质谱，从而产生适于与非指定前体离子选择方法一起使用的产物离子质谱。

[0105] 在各种实施例中，质量容限是由串联质谱仪测量的较低准确度质量的已知误差范围。

[0106] 用于转换产物离子质谱的方法

[0107] 图14是展示根据各种实施例的用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法1400的流程图。

[0108] 在方法1400的步骤1410中，使用处理器接收由串联质谱仪产生的至少一个产物离子质谱。

[0109] 在步骤1420中，使用处理器接收对应于至少一个产物离子质谱的化合物的化学结构。

[0110] 在步骤1430中，使用处理器基于化学结构将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值。

[0111] 在步骤1440中，使用处理器针对至少一个峰值选择一或多个指派的元素组成中的至少一个元素组成。

[0112] 在步骤1450中，使用处理器将至少一个峰值的质量转换成所选择的至少一个元素组成的质量，从而产生具有更高质量准确度的产物离子质谱。

[0113] 用于转换产物离子质谱的计算机程序产品

[0114] 在各种实施例中，计算机程序产品包含有形计算机可读媒体，其内容包含具有指令的程序，在处理器上执行所述指令以便执行用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法。此方法由包含一或多个相异软件模块的系统执行。

[0115] 图15是根据各种实施例的包含一或多个相异软件模块的执行用于将产物离子质谱转换成具有更高质量准确度的产物离子质谱的方法的系统1500的示意图。系统1500包含输入模块1510及分析模块1520。

[0116] 输入模块1510模块接收由串联质谱仪产生的至少一个产物离子质谱。输入模块1510接收对应于至少一个产物离子质谱的化合物的化学结构。

[0117] 分析模块1520基于化学结构将一或多个元素组成指派到至少一个产物离子光谱中的至少一个峰值。分析模块1520使用分析模块针对至少一个峰值选择一或多个指派的元素组成中的至少一个元素组成。最后，分析模块1520使用分析模块将至少一个峰值的质量转换成所选择的至少一个元素组成的质量，从而产生具有更高质量准确度的产物离子质谱。

[0118] 所属领域的一般技术人员应了解，本文关于质谱数据使用的术语“质量”可与术语“质荷比 ( $m/z$ )”互换使用。

[0119] 尽管结合各种实施例描述本教示，但不希望将本教示限于此类实施例。相反，本教示涵盖各种替代、修改及等效物，如由所属领域的技术人员将了解。

[0120] 此外，在描述各种实施例时，本说明书可能已将方法及/或过程呈现为特定步骤序

列。然而,在某种程度上,所述方法或过程不依赖于本文所阐述的特定步骤顺序,不应将所述方法或过程限于所描述的特定步骤序列。如所属领域的一般技术人员应了解,其它步骤序列可为可能的。因此,本说明书中所阐述的特定步骤顺序不应被解释为对权利要求书的限制。另外,不应将涉及方法及/或过程的权利要求限于按所书写的顺序执行其步骤,且所属领域的技术人员可容易地了解,序列可变化且仍保持在各种实施例的精神及范围内。

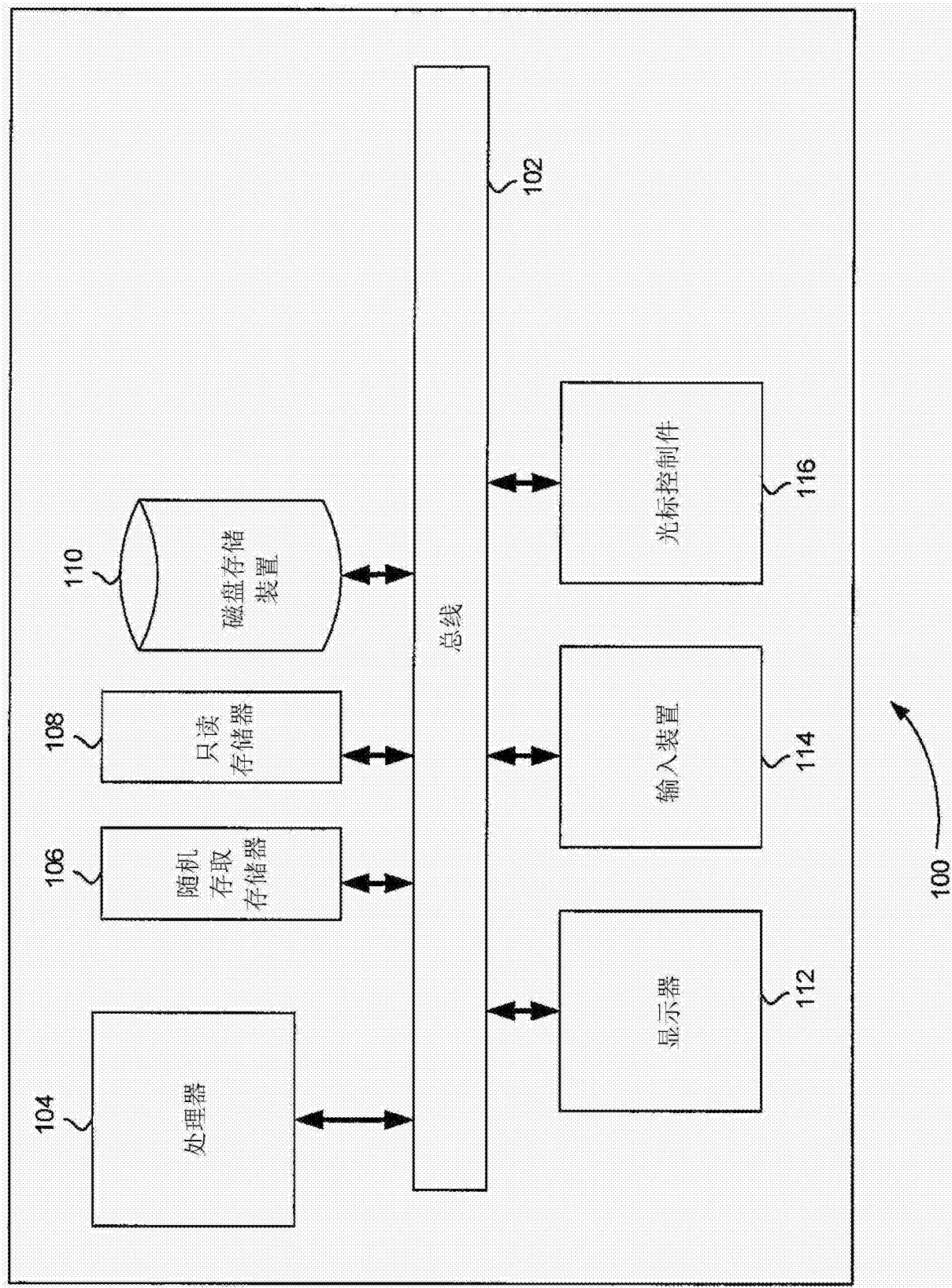


图1

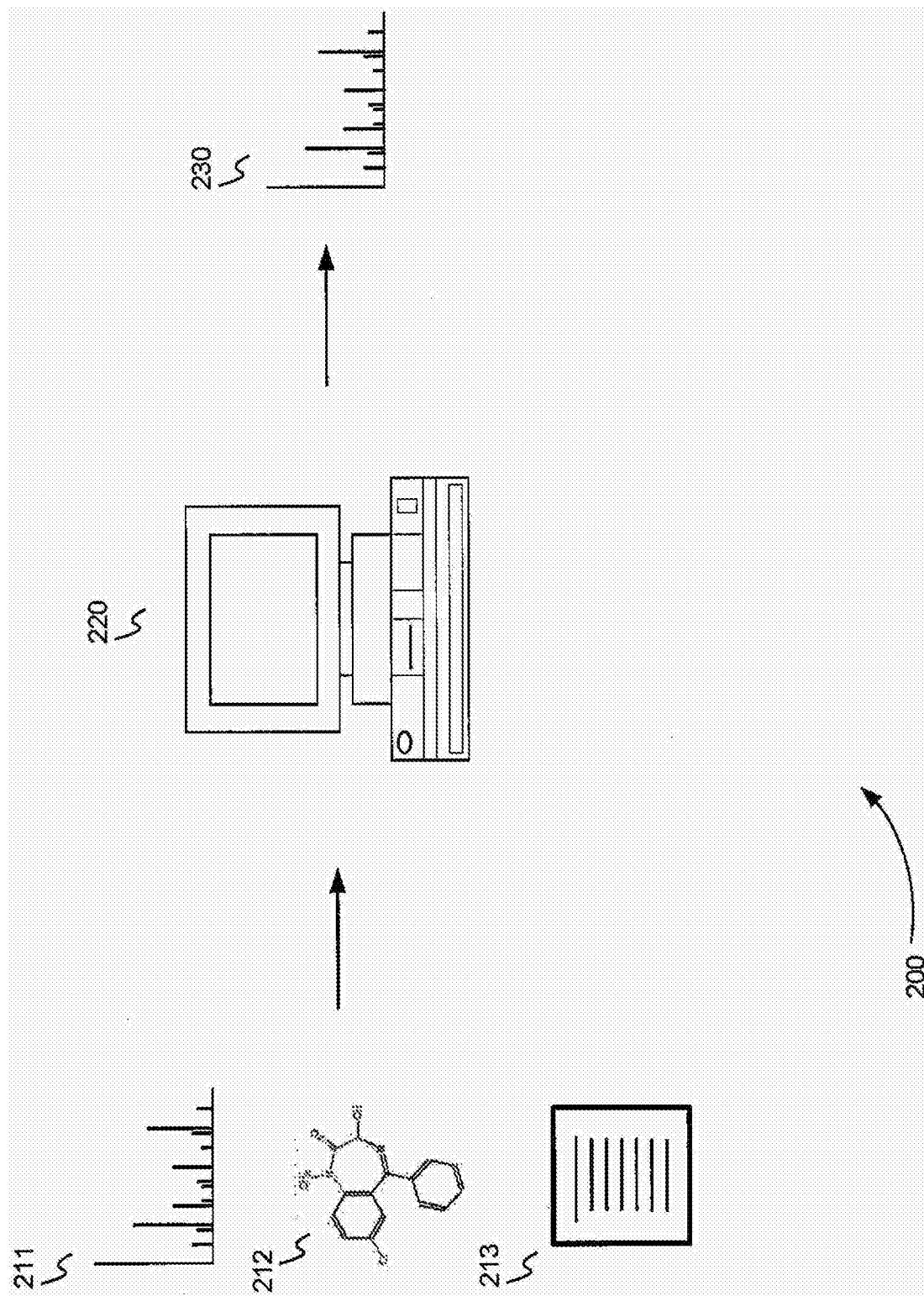


图2

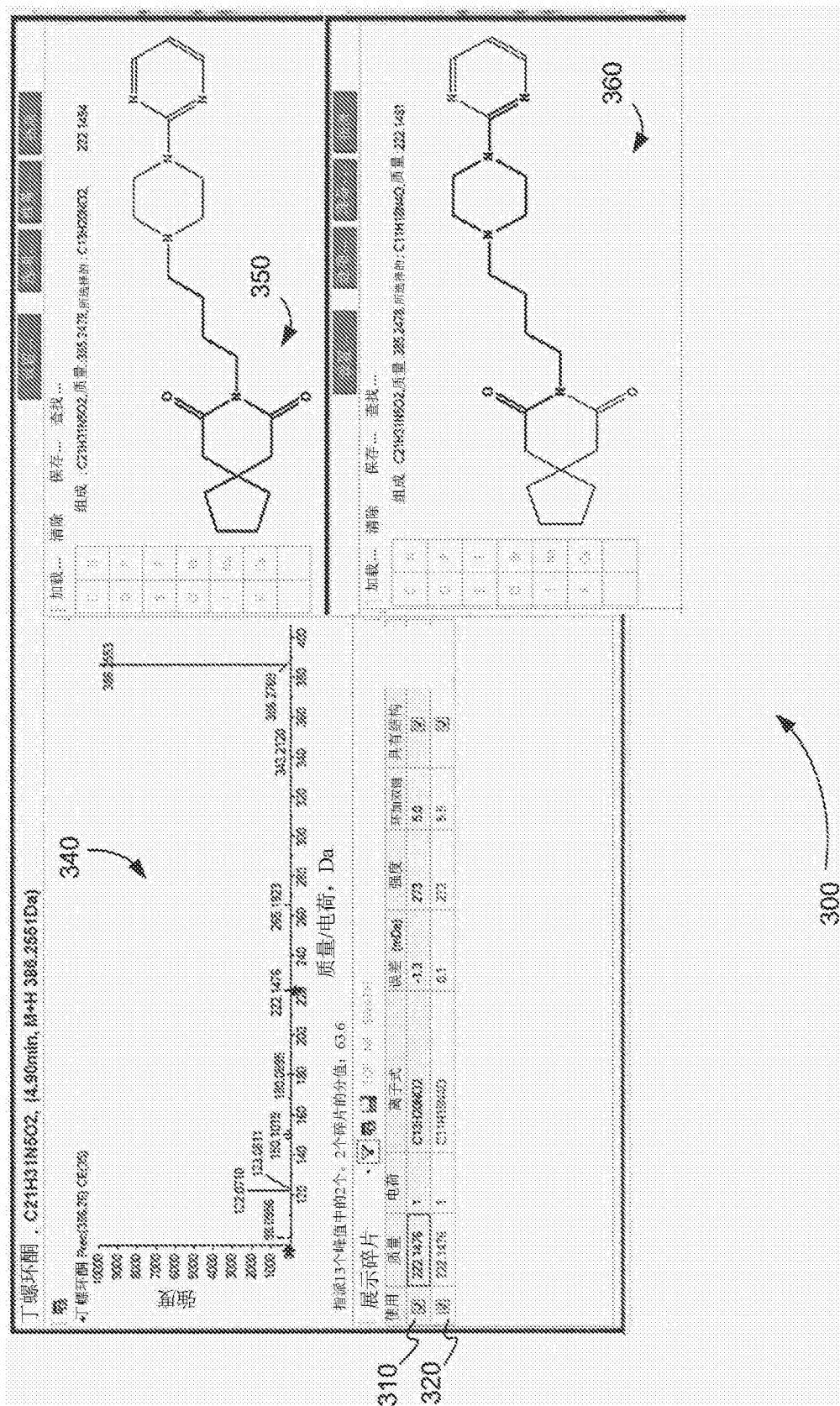


图3

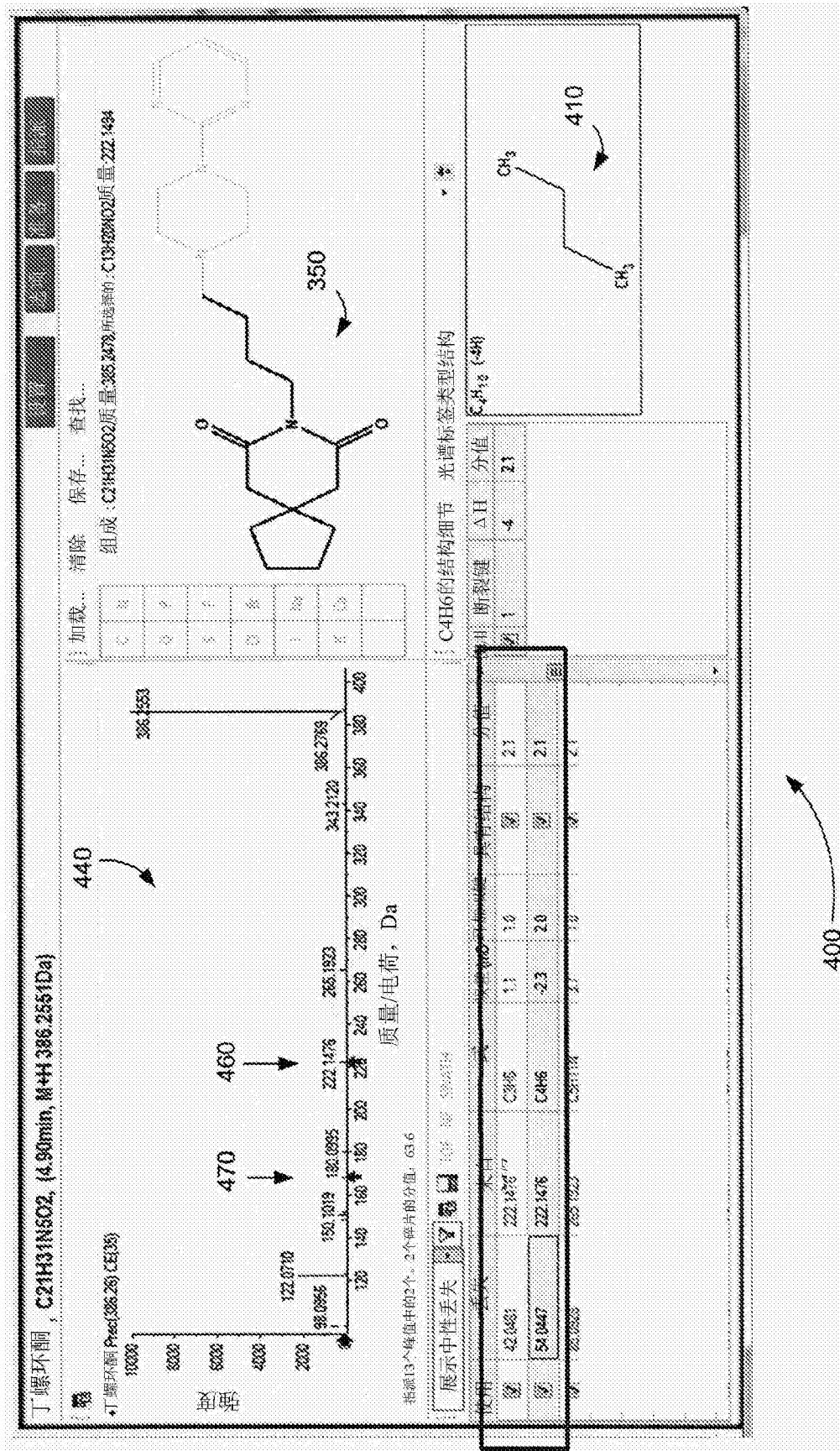


图4

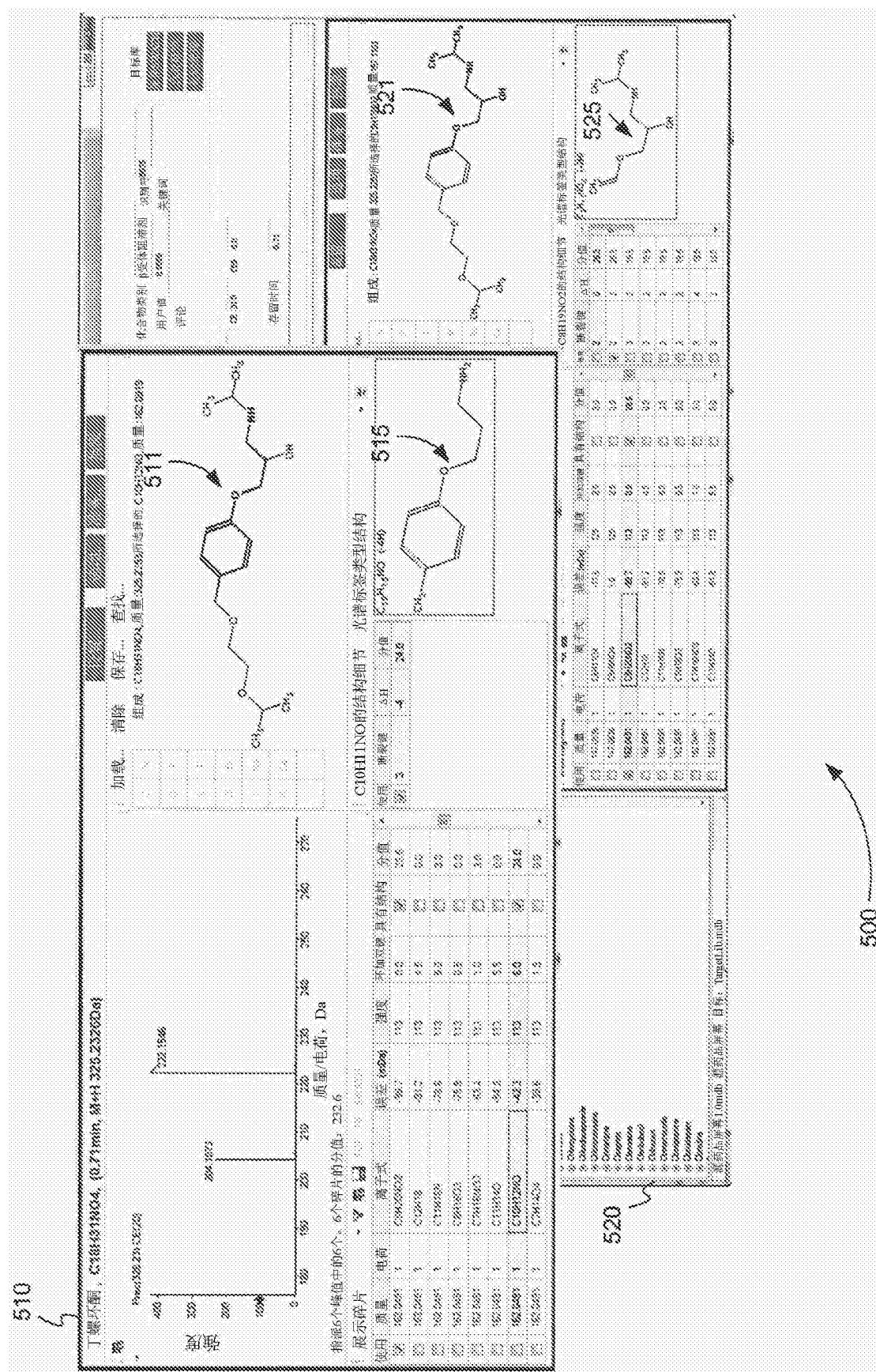


图 5

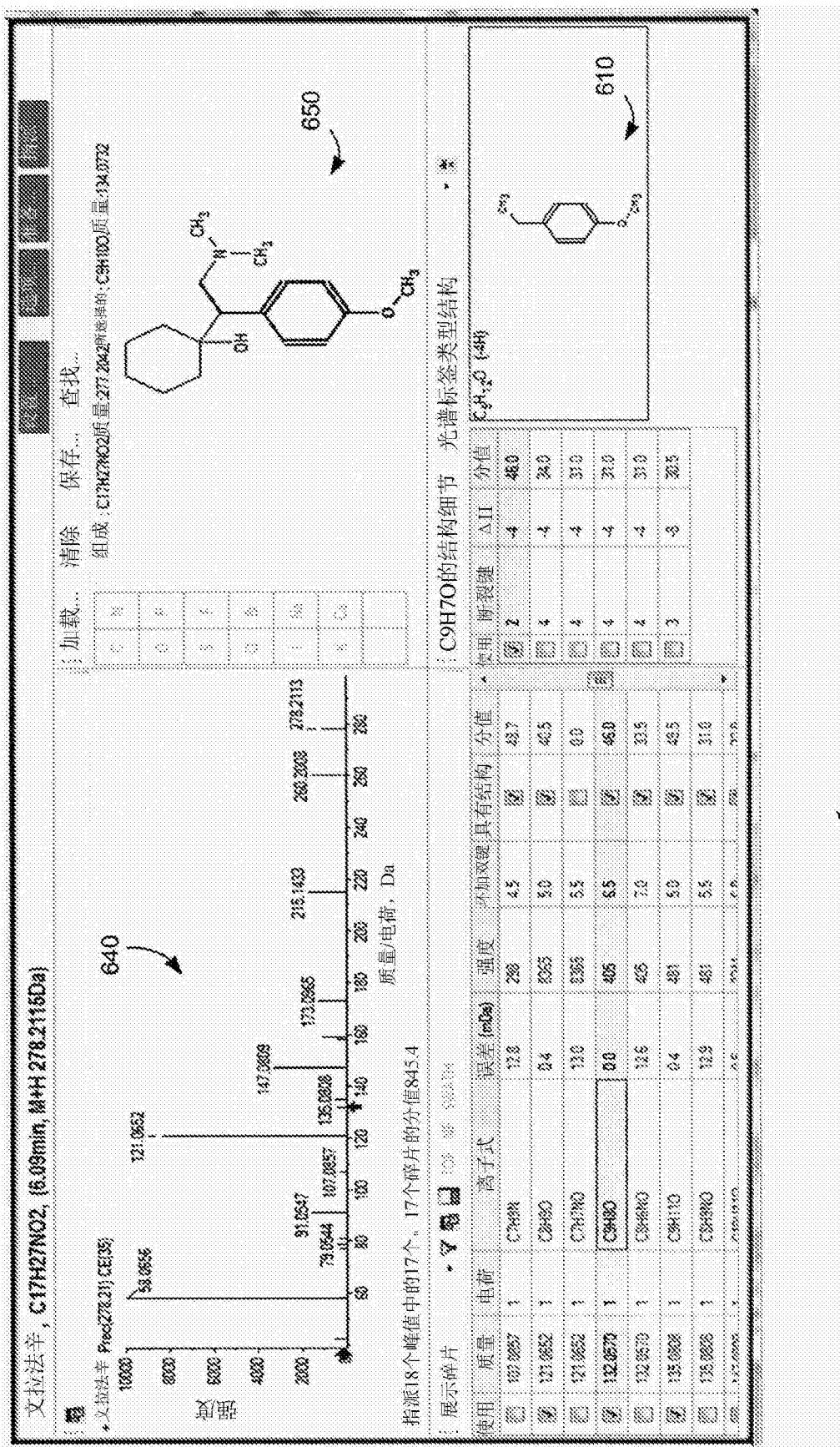


图6

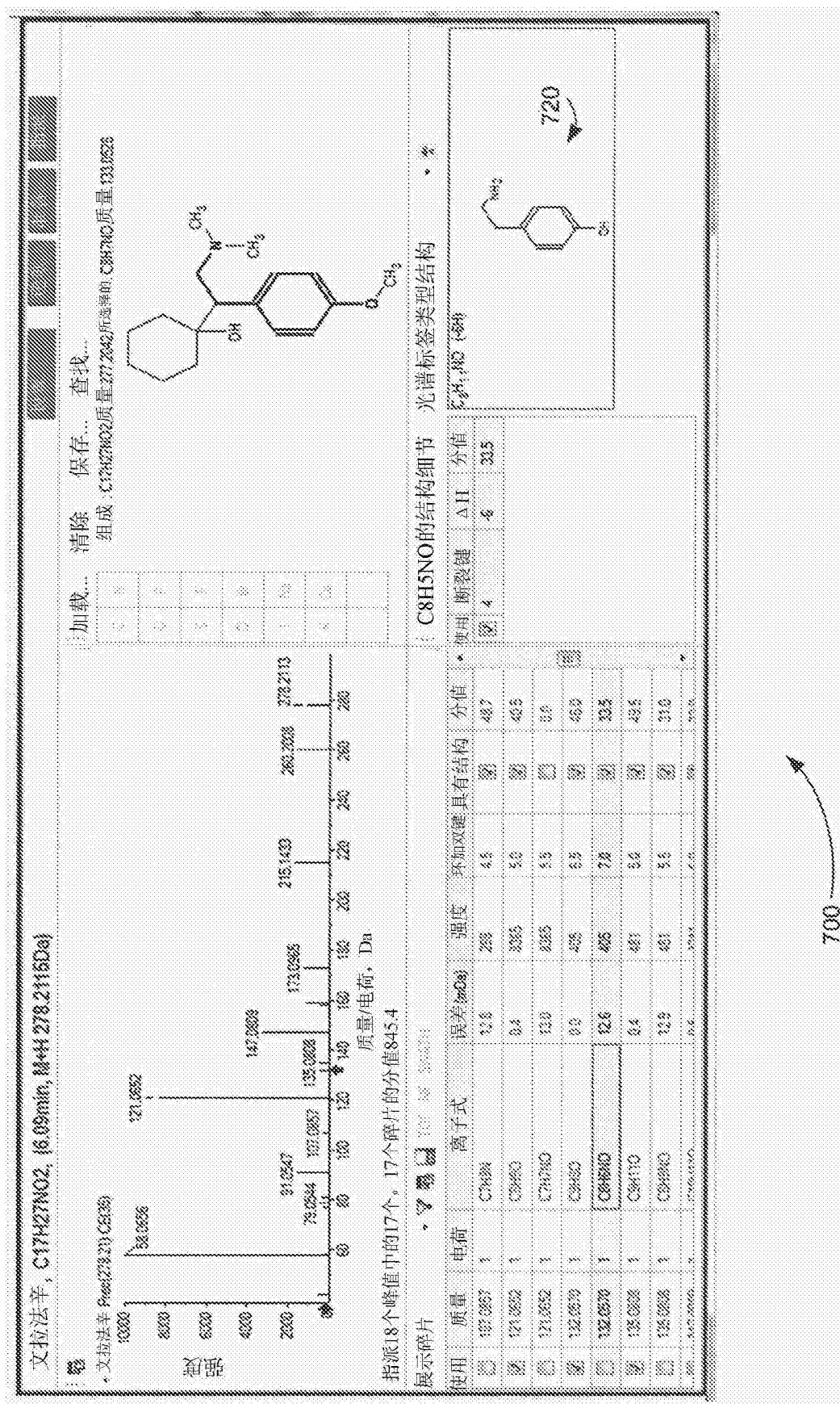


图7

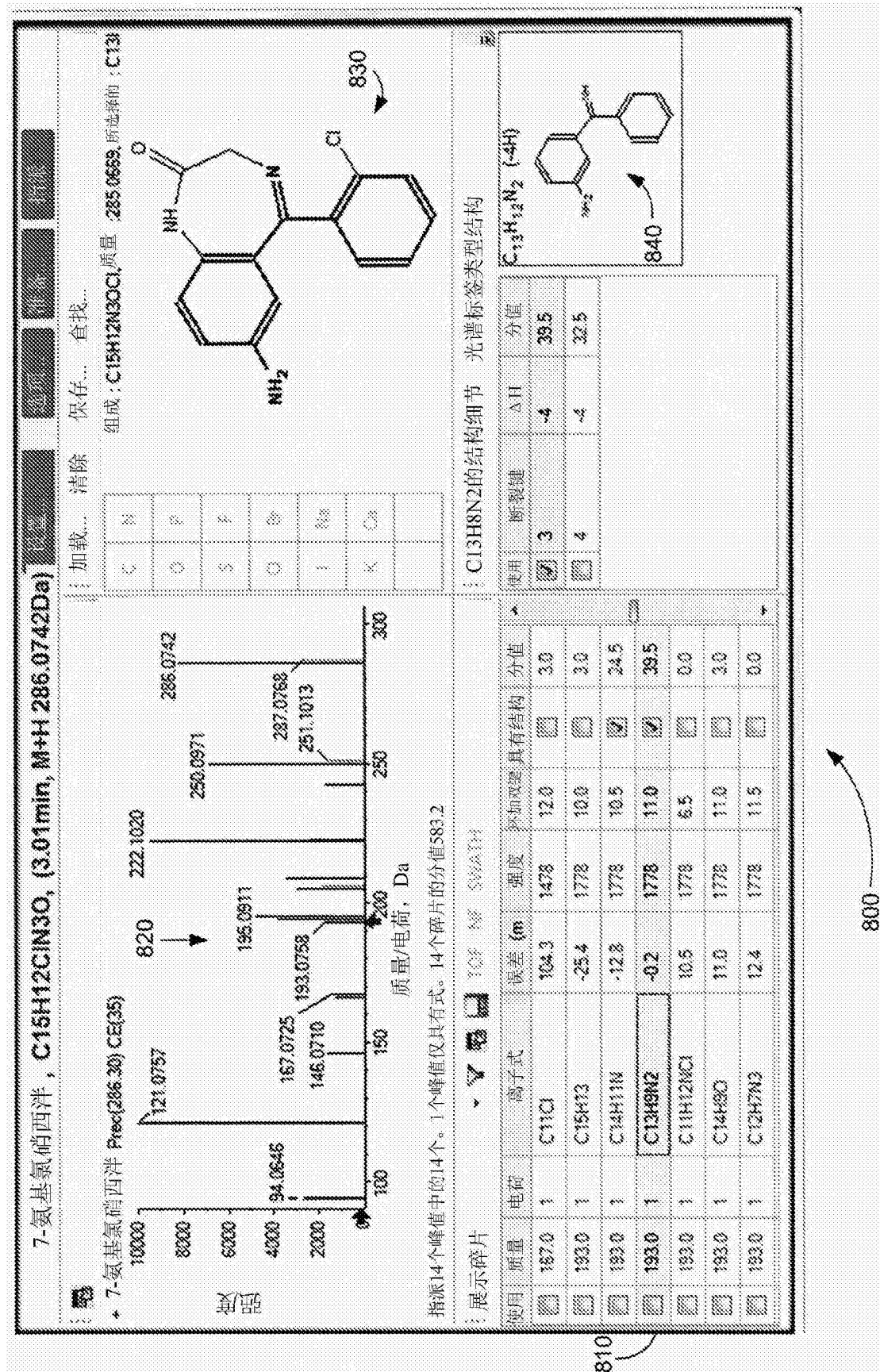


图8

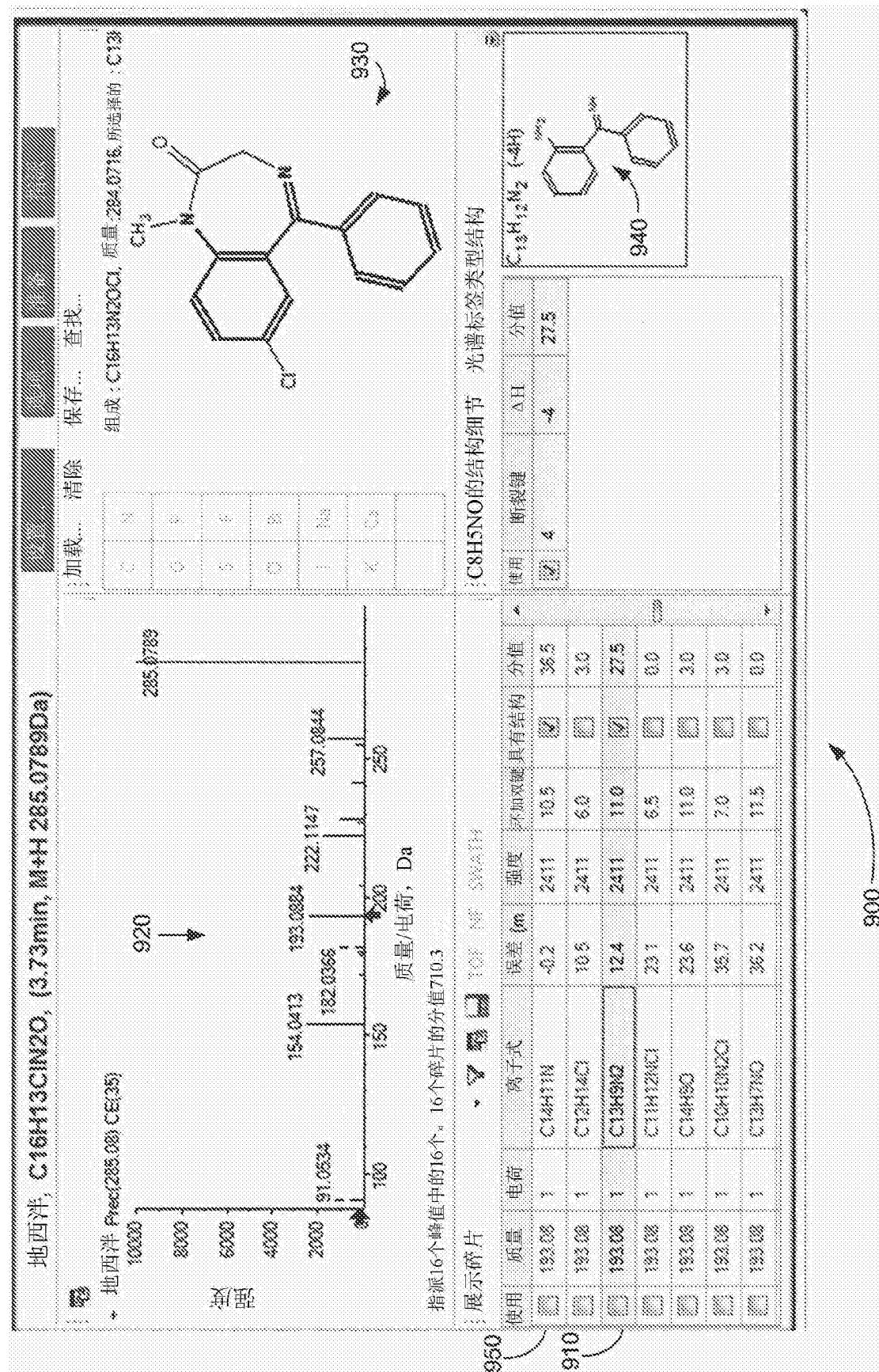


图9

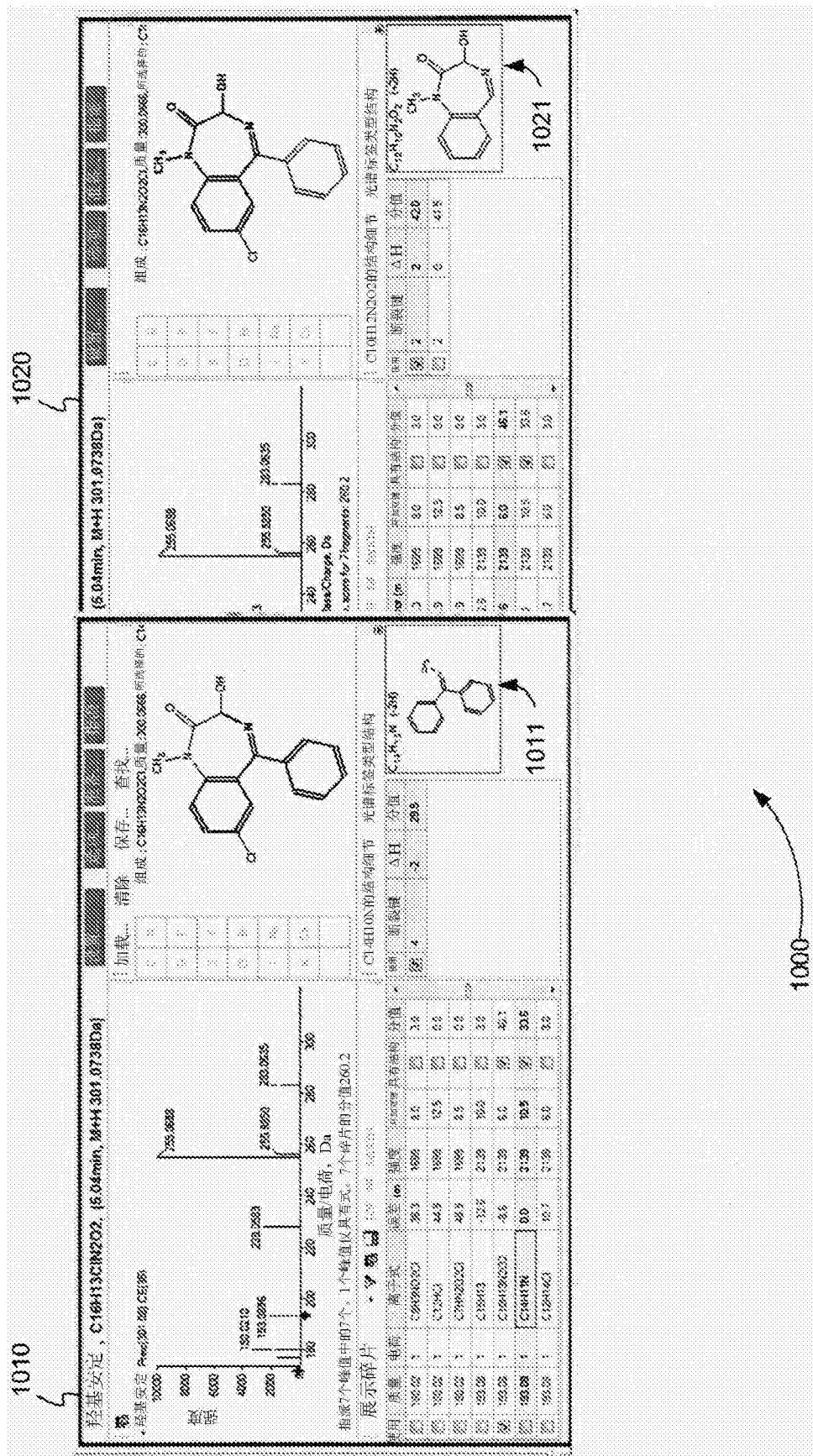


图 10

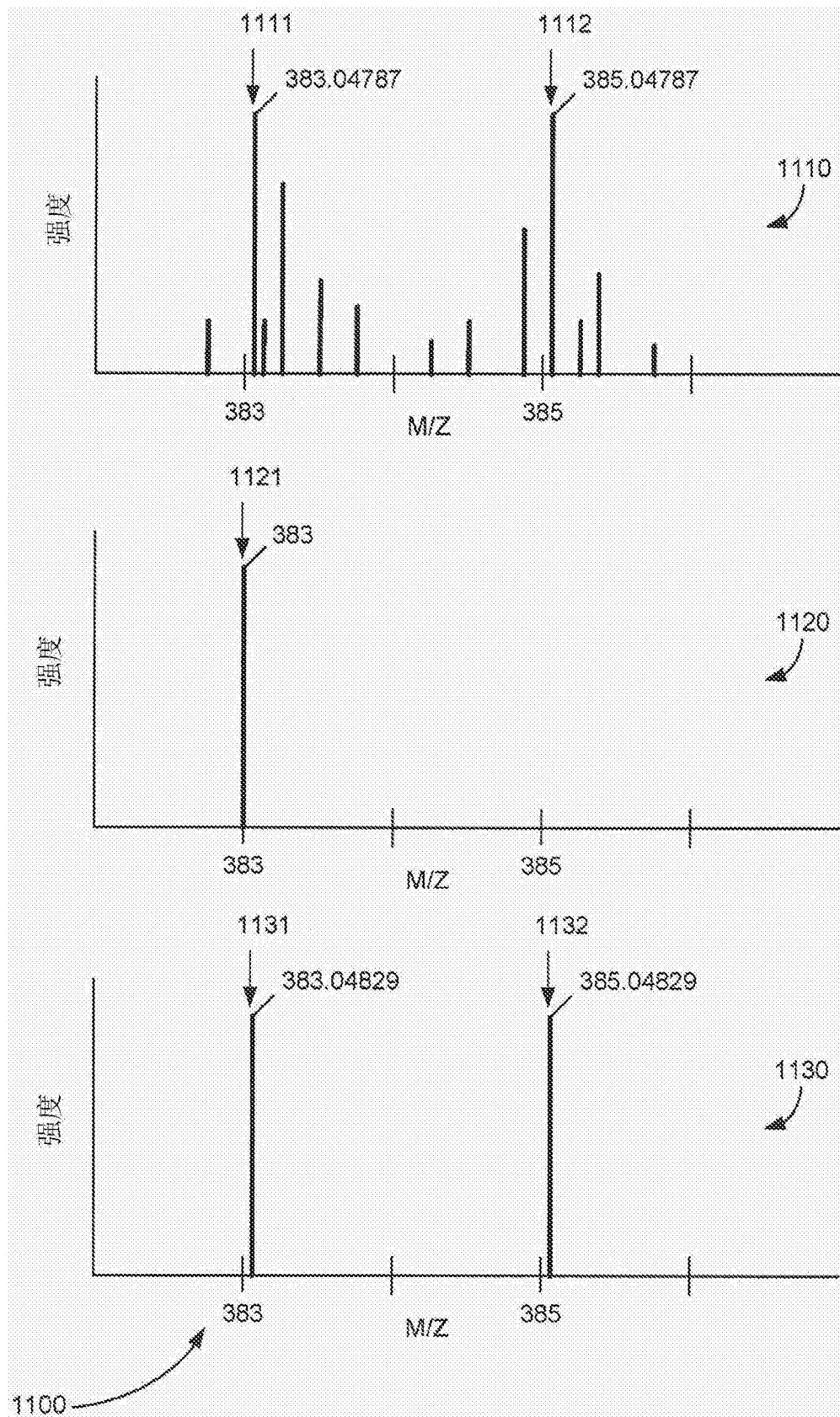


图11

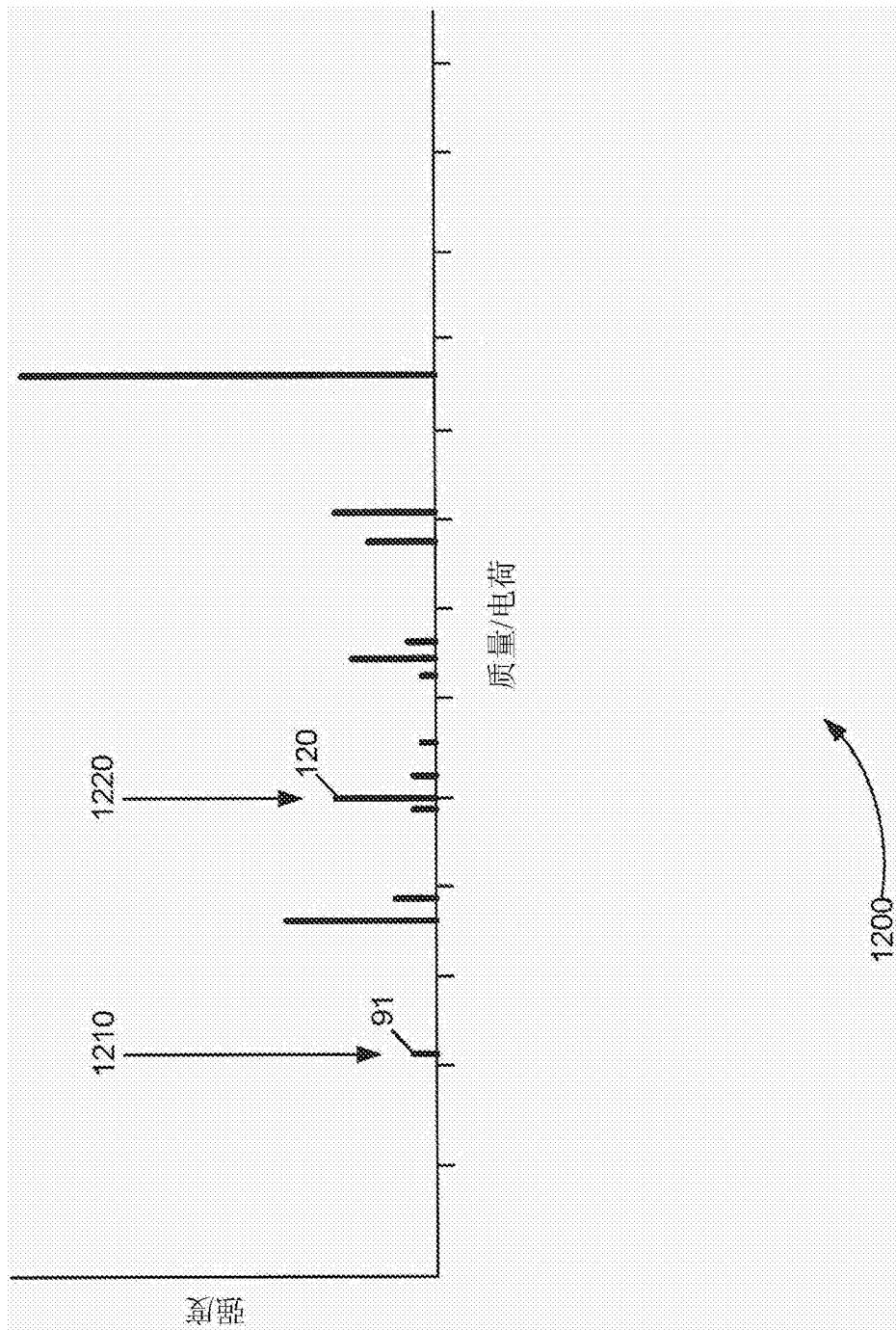


图12

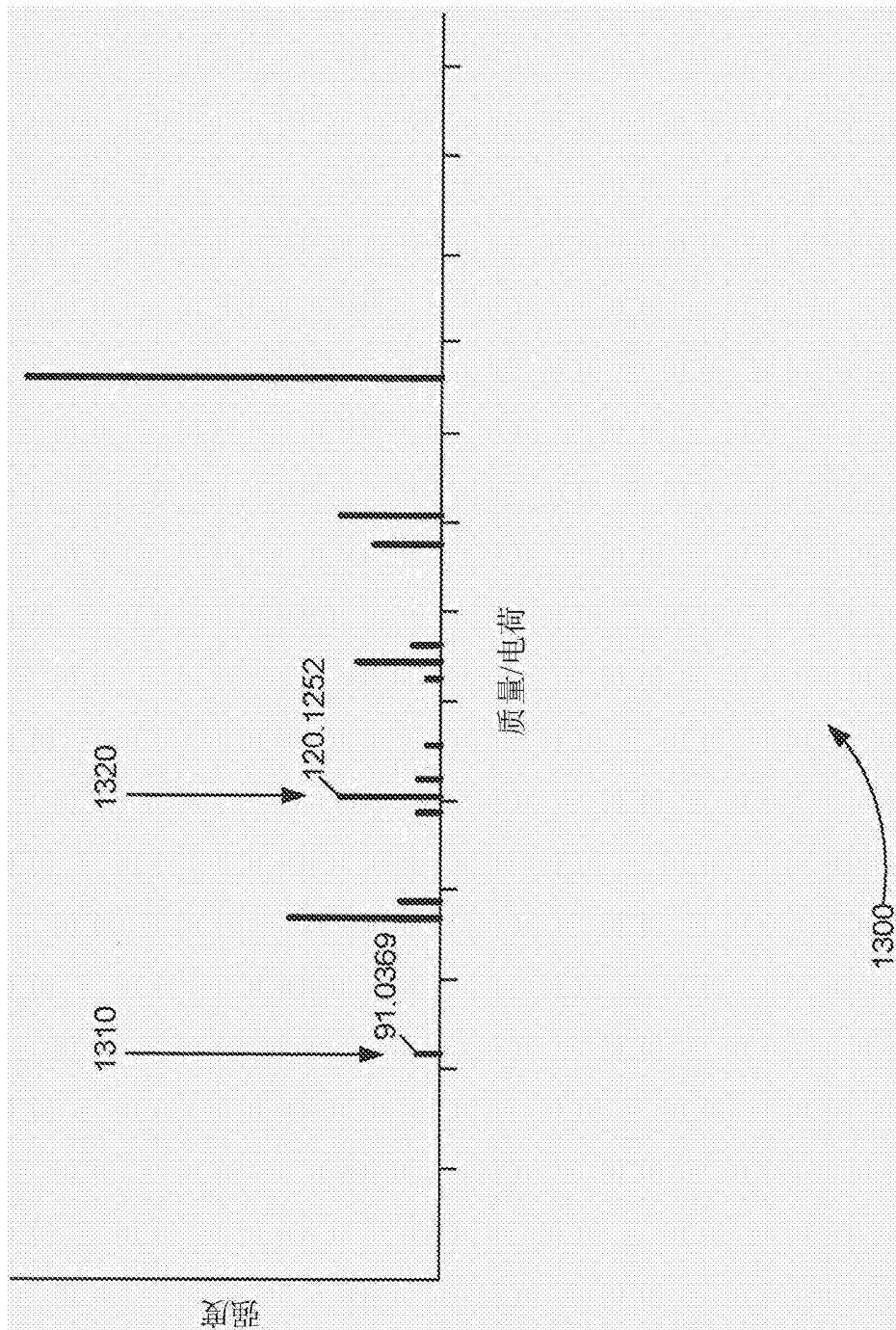


图13

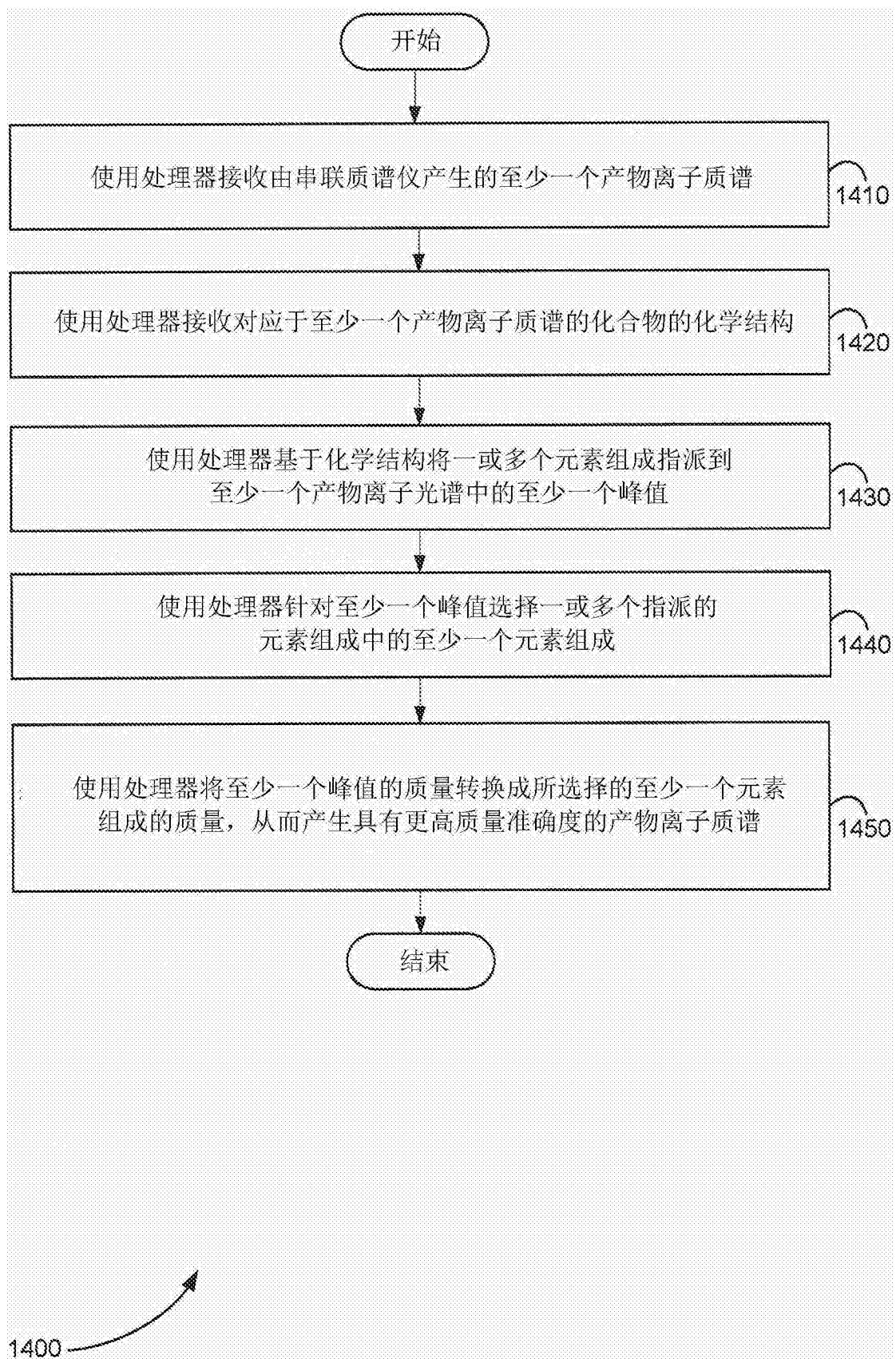


图14

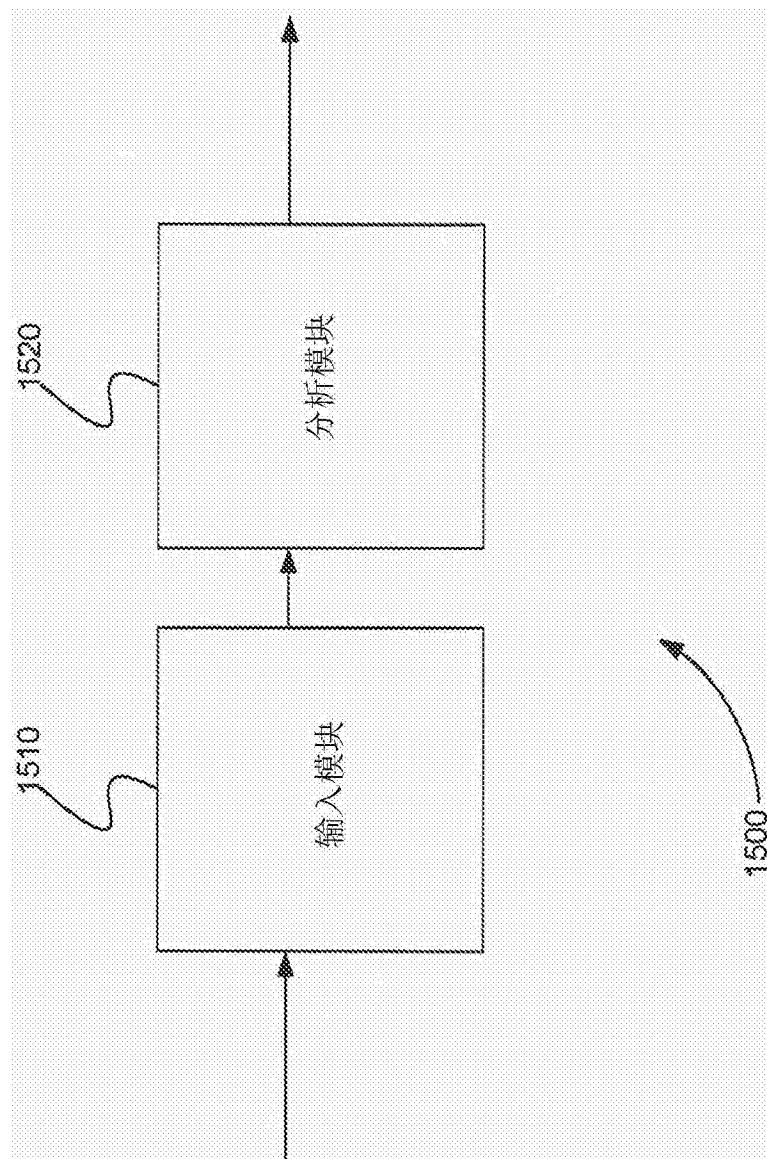


图15