



(12) 发明专利

(10) 授权公告号 CN 103109345 B

(45) 授权公告日 2016.06.22

(21) 申请号 201180044570.5

(22) 申请日 2011.09.14

(30) 优先权数据

100095959 2010.09.15 EP

61/383,137 2010.09.15 US

(85) PCT国际申请进入国家阶段日

2013.03.15

(86) PCT国际申请的申请数据

PCT/IB2011/002139 2011.09.14

(87) PCT国际申请的公布数据

W02012/035412 EN 2012.03.22

(73) 专利权人 DH 科技发展私人贸易有限公司

地址 新加坡新加坡

专利权人 苏黎世联邦理工学院

(72) 发明人 罗纳德·邦纳 斯蒂芬·泰特

鲁道夫·埃伯索尔德

佩德罗·何塞·纳瓦罗阿尔瓦雷斯

奥利弗·林纳 卢卡斯·赖特尔

卢多维克·吉莱

(74) 专利代理机构 北京律盟知识产权代理有限公司
11287

代理人 容春霞

(51) Int. Cl.

H01J 49/00(2006.01)

(56) 对比文件

GB 2438488 A, 2007.11.28,

US 2005/0230611 A1, 2005.10.20,

CN 1689134 A, 2005.10.26,

PAPAYANNOPOULOS I A 等. The

interpretation of collision-induced dissociation tandem mass spectra of peptidea. 《Mass Spectrometry Reviews》. 1995, 第14卷(第1期), 全文.

审查员 李晨雄

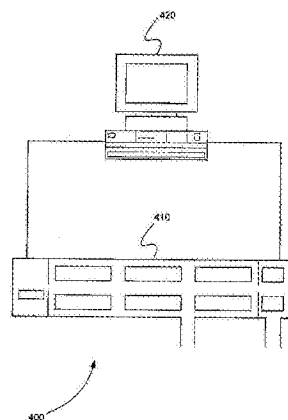
权利要求书2页 说明书13页 附图6页

(54) 发明名称

产物离子光谱的数据独立获取及参考光谱库匹配

(57) 摘要

本发明揭示用来存储样本的所有可检测化合物的全部产物离子光谱的电子记录的系统及方法。在串联质谱仪上在单个样本分析中使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描。产生每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱。使用处理器从所述串联质谱仪接收每一质量选择窗口的全部样本产物离子光谱。使用所述处理器将每一质量选择窗口的全部样本产物离子光谱存储为所述样本的所有可检测化合物的电子记录。所述电子记录用于对在存储所述电子记录时已知的化合物进行表征或对在存储所述电子记录之后变为已知的化合物进行表征。



1. 一种用于定量样本的可检测化合物的系统,其包括:

处理器,其

针对多个时间步骤中的每个时间步骤接收多个前驱物质量选择窗口中的每一质量选择窗口的样本产物离子光谱,其中通过随着时间的推进而分离样本中的可检测的化合物来产生所述样本产物离子光谱,以及在每个时间步骤处通过使用串联光谱仪对在所述样本的整个质量范围内选择的多个前驱物质量选择窗口执行多次产物离子扫描来分析所述可检测的化合物;及

通过以下步骤来利用已知产物离子光谱搜索所接收的产物离子光谱中的一种或多种已知化合物的存在:

a)从库中检索所述已知产物离子光谱,

b)检索与预期包含与所述已知产物离子光谱相对应的前驱物离子的所述前驱物质量选择窗口相对应的所述样本产物离子光谱,

c)及时生成一个或多个所述已知产物离子光谱的所述样本产物离子光谱的产物离子迹线,

d)计算所述产物离子迹线和产物离子光谱的得分,所述得分表示所述已知产物离子与样本产物离子的匹配程度,以及

e)当所述得分超过阈值时,计算来自所述产物离子迹线的已知化合物的定量值。

2. 根据权利要求1所述的系统,其中所述多个质量选择窗口中的每一质量选择窗口具有大于15amu的宽度。

3. 根据权利要求1所述的系统,其中所述已知化合物包括肽、蛋白质、完整蛋白质组、内源性代谢物、脂质、碳水化合物或其组合中的一者或多者。

4. 根据权利要求1所述的系统,其中所述已知化合物包括具有药品的、环境的、法医学的或工业重要性的一种或多种化合物。

5. 根据权利要求1所述的系统,其中所述已知产物离子光谱是从所述已知化合物的可信标准、从对包含所述化合物的样本的分析、从现存光谱库确定,或者通过将经验的或先验的片段化或改质规则应用到所述已知化合物而以计算的方式产生。

6. 根据权利要求1所述的系统,其中所述处理器进一步使用来自步骤d)的一种或多种匹配的样本产物离子来识别所述已知化合物中的一种或多种化合物的经改质形式。

7. 根据权利要求6所述的系统,其中所述处理器将所述匹配的样本离子与从已知产物离子及所述改质所预测的质量进行比较,以对所述改质的类型及位置进行表征。

8. 根据权利要求6所述的系统,其中所述处理器进一步从所述样本产物离子光谱提取被识别的经改质形式的光谱,以便对所述被识别的经改质形式中的所述改质的类型及位置进行表征。

9. 一种用于定量样本的可检测化合物的方法,其包括:

针对多个时间步骤中的每个时间步骤,使用处理器接收多个前驱物质量选择窗口中的每一质量选择窗口的样本产物离子光谱,其中通过随着时间的推进而分离样本中的可检测的化合物来产生所述样本产物离子光谱,以及在每个时间步骤处通过使用串联光谱仪对在所述样本的整个质量范围内选择的多个前驱物质量选择窗口执行多次产物离子扫描来分析所述可检测的化合物;及

使用处理器通过以下步骤来利用已知产物离子光谱搜索所接收的样本产物离子光谱中的感兴趣的一种或多种已知化合物的存在：

- a) 从库中检索所述已知产物离子光谱，
- b) 检索与预期包含与所述已知产物离子光谱相对应的前驱物离子的所述前驱物质量选择窗口相对应的所述样本产物离子光谱，
- c) 及时生成一个或多个所述已知产物离子光谱的所述样本产物离子光谱的产物离子迹线，
- d) 计算所述产物离子迹线和产物离子光谱的得分，所述得分表示所述已知产物离子与样本产物离子的匹配程度，以及
- e) 当所述得分超过阈值时，计算来自所述产物离子迹线的所述已知化合物的定量值。

产物离子光谱的数据独立获取及参考光谱库匹配

[0001] 相关申请案的交叉参考

[0002] 本申请案主张2010年9月15日申请的第61/383,137号美国临时专利申请案及2010年9月15日申请的第10009595.9号欧洲专利申请案的权益,其全文以引用的方式并入本文中。

技术领域

[0003] 本申请案涉及质谱分析领域。

背景技术

[0004] 多年来,已使用质谱分析对复杂混合物中的化合物进行识别及定量。一般的化合物可包括(但不限于)蛋白质、肽、药品化合物及衍生物(例如代谢物、麻醉剂、杀虫剂等)。常用的质谱分析技术为串联质谱分析。在串联质谱分析中,通过质量分析器来选择前驱物离子,所述前驱物离子以某种方式经片段化且所述片段在第二质量分析器内或在第一分析器的第二扫描内进行分析。所产生的片段可用于识别或定量。

[0005] 常用的定量技术为选择反应监视(SRM)。SRM已长期用于对小型分子进行定量,且最近已应用于肽、蛋白质及其它生物化合物(例如脂质及碳水化合物)。一般在三级四极仪器上执行SRM,其中第一及第二质量分析器具有约0.7的质量离析峰值宽度,且在液相色谱-质谱联用仪(LC-MS)分析期间,对前驱物与片段质量的一个或一个以上结合物(已知为过渡物)进行监视。

[0006] 尽管SRM具有灵敏性及稳健性,但其至少具有限制其应用的以下若干问题:

[0007] 1.在数据获取之前必须界定待测量的化合物,且必须从感兴趣的化合物的经验片段光谱或从此光谱的库来确定待监视的过渡物。

[0008] 2.由于可在一个分析期间进行测量的过渡物且因此化合物的数目由于若干原因受到限制,所以所获得的信息是不完整的。必须很好地界定色谱峰值,原因在于定量是基于过渡物的色谱中的峰值的高度或面积(即过渡物对时间的响应的曲线)进行的。因此,有必要尽可能的降低测量一组过渡物所花费的时间。灵敏性(可检测到的材料的最少量)取决于监视过渡物所花费的时间长度(驻留时间),因此较好的灵敏性花费更多的时间,这意味着可分析较少的化合物。人们认为,已识别的正确化合物通常要求测量若干过渡物并将响应与从标准光谱所期望的那些响应进行对比。如果质量过渡窗口相对较宽以使在复杂化合物中可同时选择一个或一个以上前驱物或片段且干扰目标片段离子的测量,则以上看法就特别真实。精确且准确的定量要求包含可靠的标准材料(一般为目标化合物的同位素标记形式),其产生可与目标化合物区分开的不同过渡物;也必须对这些进行监视,从而减少可进行分析的化合物的总数目。

[0009] 3.此外,可用数据受到在执行分析之前所界定的数据的限制。因此,如果必须监视不同的或额外的过渡物以提高定量的置信度的准确性,或如果需要额外数据来检测不同化合物或者目标化合物的经改质形式,那么就经常有必要对样本进行再分析以产生额外的数

据。

[0010] 4.由于一次仅可分析有限数量的化合物,所以获得呈现在样本中的所有化合物的数据就需要许多单独分析。

[0011] 一个替代获取方法使用高的及低的片段化交替进行扫描,接着对所述片段化进行处理以确定合为一体的前驱物(低能量)及片段(高能量)。由于在复杂混合物中,可同时对许多离子进行片段化,所以基于从此数据所提取的离子迹线的定量(类似于SRM)易于受到干扰。

[0012] 其它替代获取方法选择跨越感兴趣的质量范围的小型质量窗口,但完全覆盖整个质量范围就需要大量的分析且花费相当多的时间。

[0013] 因此,在同一分析中,在可进行分析的化合物的数目、且因此样本生产量(如果需要完全覆盖)、灵敏性与检测到使定量行为降级的干扰的可能性之间存在折衷。

发明内容

[0014] 本申请涉及一种用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的系统,其包括:串联质谱仪,所述串联质谱仪:在单一样本分析中使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描,从而产生每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱;及处理器,其与所述串联质谱仪通信,所述处理器:从所述串联质谱仪接收每一质量选择窗口的所述全部样本产物离子光谱,且将每一质量选择窗口的所述全部样本产物离子光谱存储为所述样本的所有可检测化合物的电子记录,其中所述电子记录用于对在存储所述电子记录时已知的化合物进行表征或对在存储所述电子记录之后变为已知的化合物进行表征。

[0015] 本申请还涉及一种用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法,其包括:使用串联质谱仪在单一样本分析中使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描,从而产生每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱;使用处理器从所述串联质谱仪接收每一质量选择窗口的所述全部样本产物离子光谱;及使用所述处理器将每一质量选择窗口的所述全部样本产物离子光谱存储为所述样本的所有可检测化合物的电子记录,其中所述电子记录用于对在存储所述电子记录时已知的化合物进行表征或对在存储所述电子记录之后变为已知的化合物进行表征。

[0016] 本申请还涉及一种计算机程序产品,其包括有形的计算机可读存储媒体,所述有形的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存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法,所述方法包括:提供系统,其中所述系统包括一个或一个以上截然不同的软件模块,且其中所述截然不同的软件模块包括测量模块及存储模块;及使用所述测量模块从串联质谱仪接收每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱,所述串联质谱仪在单一样本分析中使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描;及使用所述存储模块将每一质量选择窗口的所述全部样本产物离子光谱存储为所述样本的所有可检测化合物的电子记录,其中所述电子记录用于对在存储所述电子记录时已知的化合物进行表征或对在存储所述电子记录之后变为已知的化合物进行表征。

附图说明

[0017] 所属领域的技术人员将了解,以下所描述的附图仅出于说明的目的。所述附图无意以任何方式限制本发明教示的范围。

[0018] 图1为阐述可在其上实施本发明教示的实施例的计算机系统的框图。

[0019] 图2为根据各种实施例展示如何使用质量分析器的步进式前驱物离子选择窗口获取整个质量范围的数据的示意图。

[0020] 图3为根据各种实施例展示针对从参考光谱库所确定的片段从图2的离子迹线所抽取的质量迹线的示范性图表。

[0021] 图4为根据各种实施例展示用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的系统的示意图。

[0022] 图5为根据各种实施例展示用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法的示范性流程图。

[0023] 图6为根据各种实施例包括一个或一个以上截然不同的软件模块的系统的示意图,所述软件模块执行用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法。

[0024] 在详细描述本发明的一个或一个以上实施例之前,所属领域的技术人员将了解,本发明教示的应用并不受限于构造细节、组件布置以及在以下详细描述中所阐述的或在附图中所说明的步骤布置。而且,应了解,本文所使用的措辞及术语是出于描述的目的且不应被视为限制。

具体实施方式

[0025] 由计算机实施的系统

[0026] 图1为阐述可在其上实施本发明教示的实施例的计算机系统100的框图。计算机系统100包括总线102或用于传递信息的其它通信机构及与总线102耦合的用于处理信息的处理器104。计算机系统100还包括存储器106,其可为随机存取存储器(RAM)或其它动态存储装置,耦合到总线102用于存储待由处理器104执行的指令。存储器106还可用于在待由处理器104执行的指令的执行期间存储临时变量或其它中间信息。计算机系统100进一步包括只读存储器(ROM)108或其它静态存储装置,其耦合到总线102用于存储用于处理器104的静态信息及指令。提供存储装置110(例如磁盘或光盘)且将其耦合到总线102用于存储信息及指令。

[0027] 计算机系统100可经由总线102耦合到用于将信息显示给计算机用户的显示器112(例如阴极射线管(CRT)或液晶显示器(LCD))。将输入装置114(包括字母数字及其它键)耦合到总线102用于将信息及命令选择传递给处理器104。另一类型的用户输入装置为光标控制116,例如鼠标、追踪球或光标方向键,其用于将方向信息及命令选择传递给处理器104且用于控制显示器112上的光标移动。此输入装置一般在两个轴线(第一轴线(即,x)和第二轴线(即,y))上具有两个自由度,这允许所述装置指定平面中的位置。

[0028] 计算机系统100可执行本发明教示。与本发明教示的某些实施方案一致,响应于处理器104执行包含在存储器106中的一个或一个以上指令的一个或一个以上序列,计算机系

统100提供结果。可从另一计算机可读媒体(例如存储装置110)将此类指令读入到存储器106内。执行包含在存储器106中的指令序列引起处理器104执行本文所描述的过程。替代地,可使用硬连线电路来替代或结合软件指令而实施本发明教导。因此,本发明教导的实施方式并不限于硬件电路与软件的任何特定结合。

[0029] 本文所使用的术语“计算机可读媒体”意指参与将指令提供给处理器104用以执行的任何媒体。此一媒体可采取许多形式,包括但不限于非易失性媒体、易失性媒体及传输媒体。非易失性媒体包括(例如)光盘或磁盘(例如存储装置110)。易失性媒体包括动态存储器(例如存储器106)。传输媒体包括同轴电缆、铜线及光纤,包括构成总线102的电线。

[0030] 计算机可读媒体的共同形式包括(例如)软盘(floppy disk)、软盘(flexible disk)、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、存储器卡、RAM、PROM及EPROM、FLASH-EPROM、任何其它存储器芯片或盒带、或任何其它计算机可读取的有形媒体。

[0031] 在将一个或一个以上指令的一个或一个以上序列载运到处理器104用以执行时可涉及到多种形式的计算机可读媒体。举例来说,开始可在远程计算机的磁盘上载运指令。远程计算机可将指令装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的本地调制解调器可在电话线上接收数据且使用红外发射器将数据转换为红外信号。耦合到总线102的红外检测器可接收在红外信号中所载运的数据且将所述数据放置在总线102上。总线102将数据载运到存储器106,处理器104从存储器106检索并执行指令。可任选地在被处理器104执行之前或之后将由存储器106接收的指令存储在存储装置110上。

[0032] 根据各种实施例,经配置以被处理器执行而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,计算机可读媒体包括在本行业中已知用于存储软件的压缩光盘只读存储器(CD-ROM)。由适于执行经配置待执行的指令的处理器存取计算机可读媒体。

[0033] 已出于说明和描述的目的呈现本发明教导的多个实施方案的以下描述。其并不详尽且并不将本发明教导限制于所揭示的精确形式。根据以上教导的修改和变更是可能的,或可从本发明教导的实践获得修改和变更。此外,所描述的实施方案包括软件,但本发明教导可作为硬件与软件的结合或仅以硬件的形式而进行实施。可使用面向对象的及非面向对象的编程系统来实施本发明教导。

[0034] 数据处理的系统及方法

[0035] 如以上所描述,在传统质谱分析方法中,在同一分析中,在可分析的化合物的数目、灵敏性与检测到使定量行为降级的干扰的可能性之间存在折衷。

[0036] 因此,需要一种方法,其提供在复杂混合物中存在的所有可检测化合物的完整记录,所述记录可用于对已知化合物进行定量、对已知化合物的经改质形式进行确定及定量或者确定未知的改质的类型及位置。此外,应存储所述记录,以便可在获取数据的时候或之后的某个时候执行这些操作。此一记录使得能够对可用分离装置及所使用的质谱仪系统检测到的所有化合物进行完整的分析。

[0037] 此方法允许在不需要在样本上重新获取数据的情况下确定动态定量目标过渡物及目标化合物的经改质形式(例如,代谢物或者转译后的改质)。

[0038] 在各种实施例中,系统及方法提供用于产生所有可检测化合物的记录的方法,且

包括产生所有化合物的片段光谱的数据获取方法与目标数据分析方法的新颖结合。

[0039] 当前的选择反应监视(SRM)方法需要漫长的准备工作来在样本注入(包括参考光谱库产生)之前设计方法;其允许监视每一注入的有限数目的肽的过渡物;其需要通过参考库匹配进行大量数据分析以确认所监视的肽的身份;且其需要新的数据获取以改善定量精确性(通过用新的过渡物取代被污染的过渡物)及/或扩张对新的过渡物及/或在原始数据集中没有进行监视的分析物的定量。

[0040] 在各种实施例中,系统及方法解决了SRM方法的当前限制且使得能够迭代地且递归地从单一样本注入探测整个蛋白质组。虽然这些系统及方法解决了SRM方法的当前限制,但其并不以任何方式受限于SRM或应用SRM的试验类型。这些系统及方法包括新颖的LC-MS获取装置与用于数据分析的生物信息管线的结合。以下呈现各种实施例的细节:

[0041] 方法设置

[0042] 与SRM相反,各种实施例在样本注入之前并不需要任何初步的方法设计。由于LC-MS获取可在所记录的质量及保留时间范围内涵盖样本的全部分析内容(参看下文),所以可后验地提取任何感兴趣的化合物的数据。可设定保留时间及质量范围以产生特别感兴趣的范围的信息。

[0043] LC-MS获取方法

[0044] 在各种实施例中,获取方法涵盖所有可用所使用的色谱及质量范围检测到的化合物;这些可为广泛的且一般的以检测尽可能多的化合物或可经调节以聚焦于特别感兴趣的化合物或化合物类型。前驱物离子的宽的窗口(例如 >10 , >15 , >20 amu)经选择且经片段化以产生呈现在所述窗口中的所有前驱物的片段光谱。以步进方式移动窗口以涵盖其余的前驱物空间(例如,对于25 amu的窗口宽度,第一窗口可涵盖100-125,第二窗口可涵盖125-150,第三窗口可涵盖150-175等等)。

[0045] 窗口可重叠以确保化合物的所有同位素形式一起呈现在至少一个窗口中。窗口具有相对四方的形状以保持小的重叠且最小化所需窗口数目,其为有利的。

[0046] 因此,获取整个质量范围的数据所需的时间并不取决于待片段化的前驱物的数目,而是取决于窗口的数目及窗口的累积时间。所述时间足够短以维持分离所产生的峰值的保真度。

[0047] 取决于是否使用分离系统,将产生整个质量范围的产物离子光谱重复一次或一次以上。所有未经片段化的前驱物离子的质谱可作为循环的部分而并入。存储所获取的所有数据用于随后的提取。窗口的宽度可恒定或可改变。

[0048] 图2为根据各种实施例展示如何使用质量分析器的步进式前驱物离子选择窗口获取全部质量范围的数据的示意图200。图形200描绘LC-MS方法,其中通过在整个色谱分析230期间在整个质量范围220内重复地逐步平移质量分析器的隔离窗口而获得对全部片段离子光谱的数据独立获取。举例来说,图形200为LC-MS图谱。注意,在图形200中的每一循环开始之前的虚线描绘对高分辨率的精确质量调查(MS1)扫描的任选获取,如果分析过程需要,所述质量调查扫描也可用于将片段离子与片段离子所起源于的前驱物进行再关联。

[0049] 通过将针对每一隔离窗口所获取的产物离子光谱组合为单独的MS2图谱可解释图形200的数据。MS2图谱240为隔离窗口210的所有产物离子光谱的示范性组合。MS2图谱240包括离子迹线250,其绘示为质量对电荷(m/z)、保留时间及信号强度的函数。符号260识别

属于对应分析物的片段的离子迹线250。符号270识别从参考光谱库所确定的片段的离子迹线250。

[0050] 图3为根据各种实施例展示针对从参考光谱库所确定的片段从图2的离子迹线所提取的质量迹线的示范性图表300。

[0051] 在各种实施例中,LC-MS获取方法包括对从离子前驱物的片段化所产生的产物离子的监视,如下所述:

[0052] 样本的所有分析物的全部内容涵盖:

[0053] 1)基于质量分析器的隔离窗口的步进的数据独立获取:

[0054] 不考虑样本的内容,以数据独立的方式在整个质量范围内且通过整个色谱分析获取MS信息,而不监视每一运转的少许离散前驱物/过渡物。在各种实施例中,并非聚焦于预定或目标前驱物,而是通过在整个质量范围内逐步步进质量分析器的前驱物离子选择窗口而可实现这一点(参看图2)。因此,并非通过在SRM中待监视的过渡物的数目,而是通过涵盖整个质量范围(图2)所必要的步骤数目来确定这些测量的循环次数(或驻留时间)。此类步进式数据独立的片段化测量允许在一个单个运转内获取有关包含在样本中的分析物/前驱物的全部信息。实际上,此数据获取方法可产生存在于样本中的所有分析物的全部片段离子图谱,且使片段离子光谱涉及在其中获取片段离子光谱的前驱物离子选择窗口。

[0055] 2)第一质量分析器的选择窗口的变宽:

[0056] 2a)即使在质量分析器可实现的最窄选择窗口处,也几乎不可能确保仅选择不含污染物的感兴趣的前驱物用于片段化。因此,在许多实施例中,采用相反的方法:使质量分析器的前驱物隔离窗口变宽且因此包括多个前驱物共洗脱,且有助于分析期间所记录的片段化模式。下文在数据分析区段内描述了由多个前驱物的片段化所导致的复杂产物离子光谱的解译。

[0057] 2b)如在各种实施例中所实践的,使质量分析器的选择窗口变宽的附带积极效应为(i)在点(1)中所提到的循环时间的缩短,且因此获取所监视的前驱物的经较好界定且分辨率较高的色谱洗脱外形;及(ii)片段的信号强度增加,因为前驱物的整个同位素模式现在参与片段化,而不仅是SRM中的单个同位素峰值参与片段化。

[0058] 肽识别的置信度增益:

[0059] 3)获取全部的产物离子光谱,而非如经典SRM试验中监视给定前驱物的若干产物离子。在前驱物的洗脱内获得一系列全部产物离子(MS2)光谱,从而允许通过将全部MS2光谱的完整片段化模式(而非若干离子片段)与来自库的参考光谱进行匹配而实现前驱物识别中的更好的置信度。

[0060] 总之,在许多实施例中,LC-MS方法导致获取一系列MS2图谱而非每一前驱物的若干过渡物,所述MS2图谱可涵盖完整质量范围及色谱分析外形且可作为在宽的隔离窗口(图2-3)上所获取的伪SRM迹线进行处理。

[0061] 使用某种前驱物隔离窗口重叠可获取连续的MS2图谱,以确保在至少一个隔离窗口内任何给定前驱物离子的完整同位素模式的传送,且借此在任何LC时间点处维持母体与片段同位素峰值之间的最佳相关。此重叠可减小到可通过实验而确定的最小值,以最佳匹配在质谱仪中所使用的离子选择装置上可实现的片段离子传送外形。减少连续的隔离窗口之间的重叠允许维持最小的窗口尺寸、涵盖给定m/z范围的最小的窗口数目以及循环的隔

离窗口获取之间的最小驻留时间。

[0062] 在各种组合中,各种实施例包括以下MS获取方法;且单独及结合数据分析原理的各种组合在下文进一步描述:

[0063] 1)基于以内容独立的方式步进质量分析器的前驱物隔离窗口循环获取前驱物离子的全部片段化(MS2)光谱(参看图2)。获取并不包括(例如)将前驱物隔离窗口聚焦于预定(例如,数据依赖获取/胁迫)或目标(例如内含物列表或SRM)前驱物离子的质量上。

[0064] 2)通过在数据分析部分中所描述的本发明的各种实施例,对多个母体前驱物的这些MS2光谱进行精细搜索,所述母体前驱物是在母体离子前驱物隔离窗口内伴随选择且伴随参与所观察到的片段模式。换句话说,在各种实施例中,并不针对前驱物离子进行搜索。在预期包含前驱物离子的窗口内针对片段进行搜索。

[0065] 3)针对前驱物选择的重叠窗口的用途。

[0066] 可在各种实施例中使用重叠窗口以确保(i)适当地选择所有的前驱物离子,甚至在并不理想的质量分析器的情形下,及(ii)在质量分析器/选择窗口的边界处的前驱物离子获得其全部或大体上全部的同位素模式,所述同位素模式被选择用于在同一隔离窗口内的片段化。

[0067] 4)在同一获取运转期间前驱物隔离窗口的固定及/或可变宽度的用途。在各种实施例中,固定及/或可变宽度在同一循环期间(即,整个质量范围上的一组扫描)用于前驱物隔离窗口。使用较大窗口允许在质量/色谱分析空间的较不拥挤的部分中(即,在其中预期最小的分析物数目)缩短获取的循环时间。使用较窄的窗口可允许在质量/色谱分析空间的最复杂部分中增加分析的动态范围。实际上,较窄的窗口包含较少的将片段化的前驱物离子且因此具有较少的机会来包含大量的具有较大差异的前驱物。

[0068] 5)在同一获取运转期间,每一前驱物选择窗口的单个及/或多个(可变的或离散的)碰撞能量的用途。在此类多个碰撞能量实验期间所获取的增加的或减少的片段离子强度可经检查以用于共洗脱且与来自光谱库的参考片段离子强度相关联,且可加强对源自同一母体离子的片段离子峰值群组的识别(参看数据分析部分)。

[0069] 6)在同一获取运转期间,每一前驱物隔离窗口的固定及/或可变时间对MS2光谱的获取的用途。由于在获取时间期间可将信号报告为计数(例如,cts/msec),所以可变获取时间仍可用于定量的目的(如在“动态的”或“按时间表的”SRM中)。较长的获取时间可允许更灵敏地监视低含量的前驱物。

[0070] 关于生物信息学/数据分析管线的各种实施例

[0071] 数据分析包括片段离子洗脱信息的使用及参考光谱库的数据提取。针对使用合成肽池(Picotti等人,Nat Methods2010)及/或从在那些有机体上执行的大量先前的MS定序蛋白质学分析可产生整个有机体的原型肽(唯独在一个蛋白质中发现的MS可观察到的肽,因此所述肽定性且定量地对所述蛋白质进行明确地特征化)的参考光谱库。类似地,从合成分析物参考及/或从先前的分析物MS分析可产生其它分析物的参考光谱库。重要地,一旦已产生参考片段离子库,其就可被永久使用。

[0072] 由于LC-MS数据包括从宽的前驱物选择窗口所获取的全部产物离子光谱,所以数据处理经修改以考虑在所记录的MS2光谱中所观察到的潜在地参与片段化模式的多个前驱物且考虑所有的片段离子的存在。

[0073] 后验地搜索前驱物

[0074] 与必须在样本注入之前选择感兴趣的前驱物的SRM方法相反,“完整内容涵盖”获取方法使得能够在LC-MS/MS数据集中后验地且在各种实施例中搜索且定量存在于光谱库中的任何分析物。数据分析包括从在感兴趣的前驱物的预期选择窗口(m/z)中所获取的一系列全部产物离子光谱提取所述前驱物的片段质量迹线(从参考光谱库及/或从计算机模拟预测而确定)(参看图2-3)。

[0075] 举例来说,基于将所获取的产物离子片段的质量准确性及/或相对强度与参考(或预测)片段化光谱进行对比,基于匹配片段的数目,基于这些片段的所提取的离子迹线的类似色谱分析特征(共洗脱、峰值形状等),可对前驱物识别的置信度进行打分。举例来说,通过从同一LC-MS数据集类似地搜索(及打分)假目标前驱物片段离子,可确定识别的概率。可通过在前驱物的整个色谱分析洗脱上整合产物离子迹线来执行相对定量。在各种实施例中,使用经不同同位素标记的参考分析物(类似地识别、定量及打分)来实现感兴趣的对应前驱物的绝对定量。

[0076] 肽识别的置信度增益:

[0077] 通过将全部MS2光谱的整个片段化模式与来自库的参考光谱(而非若干离子片段)进行匹配,可围绕得分最佳的片段离子峰值群组候选者的洗脱而提取一系列全部产物离子(MS2)光谱以实现前驱物识别的更好置信度。

[0078] 以上所描述的数据提取策略使用无偏提取从LC-MS/MS数据集提取参考片段离子迹线(来自光谱库)。由于全部产物离子图谱是使用来自库的独立匹配的片段离子迹线提取的,因此提取全部产物离子图谱用于多个前驱物的识别。因此这并不限制在质量分析器的选择窗口内的前驱物共洗脱的搜索数目,且可允许在同一产物离子光谱内识别多个前驱物。

[0079] 通过MS2特征提取进行替代数据处理

[0080] 另一数据处理实施例包括来自经重建的MS2图谱的所有或大体上所有的片段离子信号的重新特征提取(图2)。接着可由进行参考光谱库匹配(或最后针对分析物的预计算理论片段离子的数据库)以确定其原始前驱物来分组并搜索共洗脱片段离子信号。在各种实施例中,通过在整个其洗脱中重复前驱物识别且减去所述前驱物的产物离子信号而继续进行所述方法,从而增加分析的灵敏性且揭示含量较低的前驱物的离子片段。

[0081] 各种实施例在各种组合中包括下述数据分析原理;且单独及与MS获取原理的各种组合一起在下文进一步描述:

[0082] 1)从在所述前驱物的预期选择窗口(或多个窗口,用于经改质的或多个电荷状态的肽)中所获取的全部产物离子光谱系列提取感兴趣的前驱物的片段质量迹线(从参考光谱库或从计算机模拟预测中确定)(参看图2到3)。

[0083] 2)通过基于以下参数对所提取的片段离子迹线进行打分而识别分析物,举例来说,例如:(i)所提取的片段离子迹线的共洗脱,(ii)其峰值形状的相关,(iii)其相对强度与来自参考光谱库(或来自计算机模拟预测)的相对强度的相关,(iv)接近预期的参考色谱分析的保留时间,(v)同一前驱物的多个电荷状态的片段离子迹线的共洗脱及峰值形状相关,(vi)与一个或一个以上不同同位素标记参考(例如,轻的或重的内生样本的相应的重的或轻的参考分析物)的片段离子迹线的共洗脱、峰值形状及相对强度相关,(vii)从在各种

碰撞能量下所获取的窗口得到的片段离子迹线的共洗脱及峰值形状相关,(viii)从在各种碰撞能量下所获取的窗口得到的片段离子的相对强度与来自参考光谱库(或来自计算机模拟预测)的那些相对强度的相关,及(ix)以上各者中的两者或两者以上的组合。

[0084] 3)通过基于在同一LC-MS/MS数据集中类似地搜索假目标前驱物片段离子而进行的错误发现率评估来区分真与假肯定识别。假目标命中可大体上用于使用机器学习技术(例如半监督学习)最优化以上所提及的一个或一个以上得分的组合,且通过假设其类似于识别的零分布而估计错误发现率。

[0085] 4)使用共洗脱片段离子强度来定量包含在样本中的经识别的分析物。

[0086] 5)通过在其整个色谱分析洗脱上大体上移除被污染的片段离子迹线或已被识别的分析物的那些离子迹线来“精化”及再搜索(例如,多次反复)所获取的数据。

[0087] 6)从在所述前驱物的预期选择窗口(或多个窗口,用于经改质的或多个电荷状态肽)中所获取的全部产物离子光谱的系列提取任何感兴趣的前驱物的预计算理论片段质量迹线(例如,用于对天然的或合成的化合物的光谱库的获取及精化)。

[0088] 7)对来自经重组MS2图谱的片段离子信号的重新“特征提取”(图2);对(例如)那些在上文所描述的片段离子信号进行分组及打分(点2);通过进行参考光谱库匹配(或最后针对分析物的预计算理论片段离子的数据库)以确定其原始前驱物而对那些片段离子信号进行搜索;基于其共洗脱片段离子强度对被识别的分析物进行定量。

[0089] 总之,各种实施例可允许(i)在单一LC-MS注入或分析中,对存在于样本中的所有分析物的产物离子光谱进行详尽的获取,(ii)通过特定的数据提取策略而对那些产物离子光谱进行完全的识别与定量分析,及(iii)通过反复数据提取而对那些分析进行精化及/或补充。因此,在数据一致性、识别速率及定量速度方面,此组合的LC-MS获取及数据处理方法相对于传统方法获得了显著改善。这些发明使得对完整蛋白质组图谱的获取及对那些蛋白质组图谱的定性及定量数据提取的方法成为可能。

[0090] 这些发明的潜在应用本质上与SRM定量蛋白质组的那些应用相同且包括依赖于定性及定量LC-MS分析的任何生物技术、生物医学、药品及生物应用。举例来说,在各种实施例中,所述方法特别适合对可能仅在有限数量中有效的复杂样本(例如,完全有机物、细胞、器官、体液等)中的大量感兴趣候选前驱物(例如,肽)执行分析。

[0091] 各种实施例包括以下应用,其中:

[0092] -对天然的或合成的化合物(例如肽)的光谱库进行快速获取及精化。

[0093] -对感兴趣的天然的或合成的化合物进行定性及定量分析(例如,在特定分析物或生物标志测量的背景下,或分析蛋白质络合物的成分)。

[0094] -对经天然或人工改质的分析物进行定性及定量分析,所述分析物与其未经改质的对等物(例如,与活性型探针反应的具有转译后的改质的蛋白质/肽,或化学交联的蛋白质/肽)共享片段离子,或其改质共享片段离子(例如,泛素或类泛素的分子)或共同报导离子,或通过使用此改质对那些分析物的片段离子所带来的(正的或负的)质量差异,对经天然或人工改质的分析物进行定性及定量分析。

[0095] -对存在于光谱库中或重新识别的所有可检测分析物进行定性及定量分析(参考数据分析部分)(例如,在部分或完全蛋白质组分析的背景下)。

[0096] -通过对所获取的数据集进行反复数据提取而对那些定性及定量分析进行精化

及/或补充的能力。

[0097] 这些各种实施例可为采取高生产力方式对复杂样本的整个代谢物组/蛋白质组进行完全的定性及定量分析做准备。

[0098] 定量数据处理

[0099] 在各种实施例中,从单一前驱物质量窗口所获取的所有片段数据可一起进行处理。虽然所述数据可含有来自一个或一个以上前驱物离子(化合物)的片段,但其可经处理以定量感兴趣的化合物或搜索此类化合物的经改质的形式。

[0100] 高分辨率及高质量准确性的感兴趣的化合物的前驱物质量及预期的一组片段是从库获得,或通过分析化合物的可信标准形式而获得,或者从以前的分析(不论化合物是否已知)获得,或者通过使用已知片段化规则的预报而获得。可基于其预期强度、其对感兴趣的化合物而言为唯一的可能性或其它特征来选择片段组。对于含有预期的前驱物质量的窗口来说,片段质量组用于产生包括一个或一个以上峰值的离子迹线(例如色谱)。

[0101] 对迹线打分以确定正确的或最可能的峰值。可基于来自质谱的以下信息进行打分,例如:样本片段离子的经检测质量在多大程度上匹配预定产物离子的预期质量;样本片段离子的相对强度在多大程度上匹配预定产物离子的相对强度;所测量的样本离子处于正确的同位素形式,通常其为单一同位素;前驱物及片段离子具有预期的电荷状态。

[0102] 如果包含分离步骤,就可基于以下额外信息进行打分,例如:在形状及位置方面,所检测的离子迹线在多大程度上彼此匹配。如果对样本的不同同位素形式(例如,经标记的与天然的形式)的组合)进行分析,则可使用来自不同形式的数据对打分进行进一步精化。如果所述组中的一个或一个以上片段由于存在干扰而得分较低,则可将其从所述组排除,且如果需要的话,就用来自预定光谱的另一片段取代所述片段。

[0103] 得到可接受的得分的离子可用于产生目标化合物的定量值,可将所述值与来自其它样本(例如,用于时间进程研究的成员、已进行不同处理或制备的样本群组、来自健康的或有病的对象的样本群组等)的类似值进行比较。由于来自所有被检测前驱物的所有片段离子均存在于数据中,所以在各种实施例中可通过使用减少测量误差的替代片段离子来执行最佳定量。

[0104] 由于所获取的数据包括来自所有可检测化合物的片段,所以其可针对任何数量的化合物进行提取且打分可产生定量值。

[0105] 定性数据处理

[0106] 在各种实施例中,可对数据进行进一步提取以提取有关存在于样本中的化合物的定性信息。举例来说,通过未预期的保留时间处将同一组片段离子定位于同一前驱物窗口内或不同窗口内可检测到经改质的形式。可基于由所述改质所引起的预期的质量差异来确定窗口。在各种实施例中,可通过定位具有改质的特性的离子来检测经改质的形式。

[0107] 一旦检测到经改质的形式,就可以多种方式确定所述改质的类型及位置。举例来说,可通过对依赖于所述改质的位置或类型的离子进行预测且通过产生从那些经预测质量的数据所提取的迹线并对其打分,可确定所述改质的类型及位置。在各种实施例中,可通过从所述数据产生光谱且对所述光谱进行解译而确定所述改质的类型及位置。

[0108] 此外,可处理来自每一窗口的数据以确定或识别相关离子且借此提取已知或未知化合物的光谱,可对所述光谱进行解译以确定化合物的身份。

[0109] 串联质谱系统

[0110] 图4为根据各种实施例展示用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的系统400的示意图。系统400包括串联质谱仪410及处理器420。处理器420可为(但不限于)计算机、微处理器或任何能够发送及接收来自质谱仪410的控制信号及数据且对数据进行处理装置。

[0111] 串联质谱仪410可包括一个或一个以上执行两个或两个以上质量分析的物理质量分析器。串联质谱仪的质量分析器可包括(但不限于)飞行时间(TOF)、四极子、离子阱、线性离子阱、轨道阱或者傅里叶变换质量分析器。串联质谱仪410还可包括分离装置(未图示)。分离装置可执行分离技术,其包括(但不限于)液相色谱分析、气相色谱分析、毛细管电泳或离子迁移。串联质谱仪410可包括分别在空间或时间内的分离质谱分析阶段或步骤。

[0112] 串联质谱仪410使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描。在单一样本分析中执行多个产物离子扫描。举例来说,单一样本分析为单一样本注入。串联质谱仪410从多个产物离子扫描产生每一质量选择窗口的所有可检测化合物的所有样本产物离子光谱。

[0113] 处理器420与串联质谱仪410通信。处理器420从串联质谱仪410接收每一质量选择窗口的所有样本产物离子光谱。接着处理器420将每一质量选择窗口的所有样本产物离子光谱存储为样本的所有可检测化合物的电子记录。电子记录用于对在存储电子记录时已知的化合物进行表征或用于对在存储电子记录之后变为已知的化合物进行表征。

[0114] 在各种实施例中,多个质量选择窗口中的每一质量选择窗口具有大于10原子质量单位(amu)的宽度或大于15amu的宽度。

[0115] 在各种实施例中,多个质量选择窗口中的至少两个质量选择窗口具有不同宽度。

[0116] 在各种实施例中,搜索来自电子记录的一个或一个以上质量选择窗口的所有样本产物离子光谱是否有来自参考库的预定产物离子光谱。举例来说,处理器420接收与已知化合物对应的预定产物离子光谱。处理器420从电子记录接收一个或一个以上质量选择窗口的所有样本产物离子光谱。接着处理器420将预定产物离子光谱的预定产物离子与一个或一个以上质量选择窗口的所述所有样本产物离子光谱的样本产物离子进行比较。从所述比较,一种或一种以上匹配的样本产物离子对在样本中可检测到的已知化合物进行表征。举例来说,已知化合物包括产生可再生产物离子光谱的任何化合物。

[0117] 在各种实施例中,已知化合物包括肽、蛋白质、完整蛋白质组、内源性代谢物、脂质或碳水化合物中的一者或一者以上。

[0118] 在各种实施例中,已知化合物包括具有生物的、药品的、环境的、法医学的或工业重要性的一种或一种以上化合物。所述具有生物的、药品的、环境的、法医学的或工业重要性的一种或一种以上化合物可包括(但不限于)杀虫剂、除草剂、杀真菌剂、工业化学品、麻醉剂、掺杂剂或炸药中的一者或一者以上。

[0119] 在各种实施例中,通过将经验的或先验的片段化或改质规则应用到已知化合物而通过计算产生预定产物离子光谱。

[0120] 在各种实施例中,从对与天然发生源隔离或化学合成的可信的标准化化合物的分析获得预定产物离子光谱。

[0121] 在各种实施例中,从公共的或私有的光谱库获得预定产物离子光谱。

[0122] 在各种实施例中,在对与预定产物离子光谱相对应的化合物进行识别或不进行识别的情况下,从对代表性样本的先前分析获得预定产物离子光谱。

[0123] 在各种实施例中,从参考库搜索预定产物离子光谱包括计算得分。举例来说,处理器420通过计算得分而将预定产物离子光谱的预定产物离子与一个或一个以上质量选择窗口的所有样本产物离子光谱的样本产物离子进行比较,所述得分表示预定产物离子与样本产物离子的匹配程度。举例来说,所述得分可包括比较离子质量及相对强度。在各种实施例中,得分包括确定样本前驱物离子具有预期的同位素形式。在各种实施例中,得分可包括有关前驱物及片段离子的预期电荷状态的信息。

[0124] 在各种实施例中,分离装置随着时间的推进而分离单个样本分析的样本化合物。在正对样本化合物进行分离时,串联质谱仪410在单个样本分析上执行多个产物离子扫描。在各种实施例中,基于对离子迹线的峰值形状或匹配的样本产物离子的检测时间类似性进行比较,处理器420进一步计算所述匹配的得分。

[0125] 在各种实施例中,处理器420进一步使用来自对搜索的比较的一种或一种以上匹配样本产物离子来计算样本的化合物的定量值。举例来说,使用样本产物离子光谱中的一种或一种以上匹配样本产物离子的强度来计算定量值。在各种实施例中,使用一种或一种以上匹配离子迹线峰值的强度或面积来计算定量值。

[0126] 在各种实施例中,处理器420进一步使用来自对搜索的比较的一种或一种以上匹配样本产物离子来识别已知化合物中的一种或一种以上化合物的经改质形式。举例来说,通过在同一质量选择窗口内或在不同质量选择窗口内发现一种或一种以上匹配样本产物离子而识别经改质形式。举例来说,从预期改质的质量确定不同质量选择窗口。

[0127] 在各种实施例中,通过发现与通过改质的质量进行调节的一种或一种以上匹配样本产物离子相对应的质量来识别经改质形式。举例来说,所述改质为已知改质、由已知反应引起或由其它实验提出。

[0128] 在各种实施例中,通过发现已知化合物的经改质部分的质量特性来识别经改质形式。

[0129] 在各种实施例中,通过从经结合分子发现络合物合成光谱来识别经改质形式。

[0130] 在各种实施例中,处理器420进一步从电子记录提取被识别的经改质形式的光谱,以便对被识别的经改质形式中的改质的类型及位置进行表征。

[0131] 在各种形式中,处理器420进一步使用预定产物离子及改质来预测将指示改质的位点的质量,且产生每一所述质量的得分以确定所述改质的位置。

[0132] 串联质谱分析方法

[0133] 图5为根据各种实施例展示用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法500的示范性流程图。

[0134] 在方法500的步骤510中,在单一样本分析中,在串联质谱仪上使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描。产生每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱。

[0135] 在步骤520中,使用处理器从串联质谱仪接收每一质量选择窗口的全部样本产物离子光谱。

[0136] 在步骤530中,使用处理器将每一质量选择窗口的全部样本产物离子光谱存储为

样本的所有可检测化合物的电子记录。所述电子记录用于对在存储电子记录时已知的化合物进行表征或对在存储电子记录之后变为已知的化合物进行表征。

[0137] 串联质谱分析计算机程序产品

[0138] 在各种实施例中,计算机程序产品包括有形的计算机可读存储媒体,其内容包括具有在处理器上执行的指令的程序,以便执行用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法。由包括一个或一个以上截然不同的软件模块的系统执行此方法。

[0139] 图6为根据各种实施例的包括执行用于存储样本的所有可检测化合物的全部产物离子光谱的电子记录的方法的一个或一个以上截然不同的软件模块的系统600的示意图。系统600包括测量模块610及存储模块620。

[0140] 测量模块610从串联质谱仪接收质量范围的每一质量选择窗口的所有可检测化合物的全部样本产物离子光谱。通过在单一样本分析中使用多个质量选择窗口在整个质量范围内执行一次或一次以上多个产物离子扫描,串联质谱仪产生样本产物离子光谱。

[0141] 存储模块620将每一质量选择窗口的全部样本产物离子光谱存储为样本的所有可检测化合物的电子记录。电子记录用于对在存储电子记录时已知的化合物进行表征或用于对在存储电子记录之后变为已知的化合物进行表征。

[0142] 虽然结合各种实施例描述了本发明教导,但本发明教导并无意受限于此类实施例。相反,如所属领域的技术人员将了解,本发明教导包括各种替代物、修改及等效物。

[0143] 此外,在描述各种实施例中,说明书可能已将方法及/或工艺呈现为特定的步骤顺序。然而,在所述方法或工艺并不依赖本文所阐述的特定的步骤次序的程度上,所述方法或工艺不应受限于所描述的特定的步骤顺序。如所属领域的技术人员将了解,其它的步骤顺序也是可能的。因此,本说明书中所阐述的特定的步骤次序不应解释为对权利要求书的限制。此外,针对方法及/或工艺的权利要求书不应受限于以所写出的次序执行其步骤,且所属领域的技术人员可易于了解,所述顺序可改变且仍保持在各种实施例的精神和范围内。

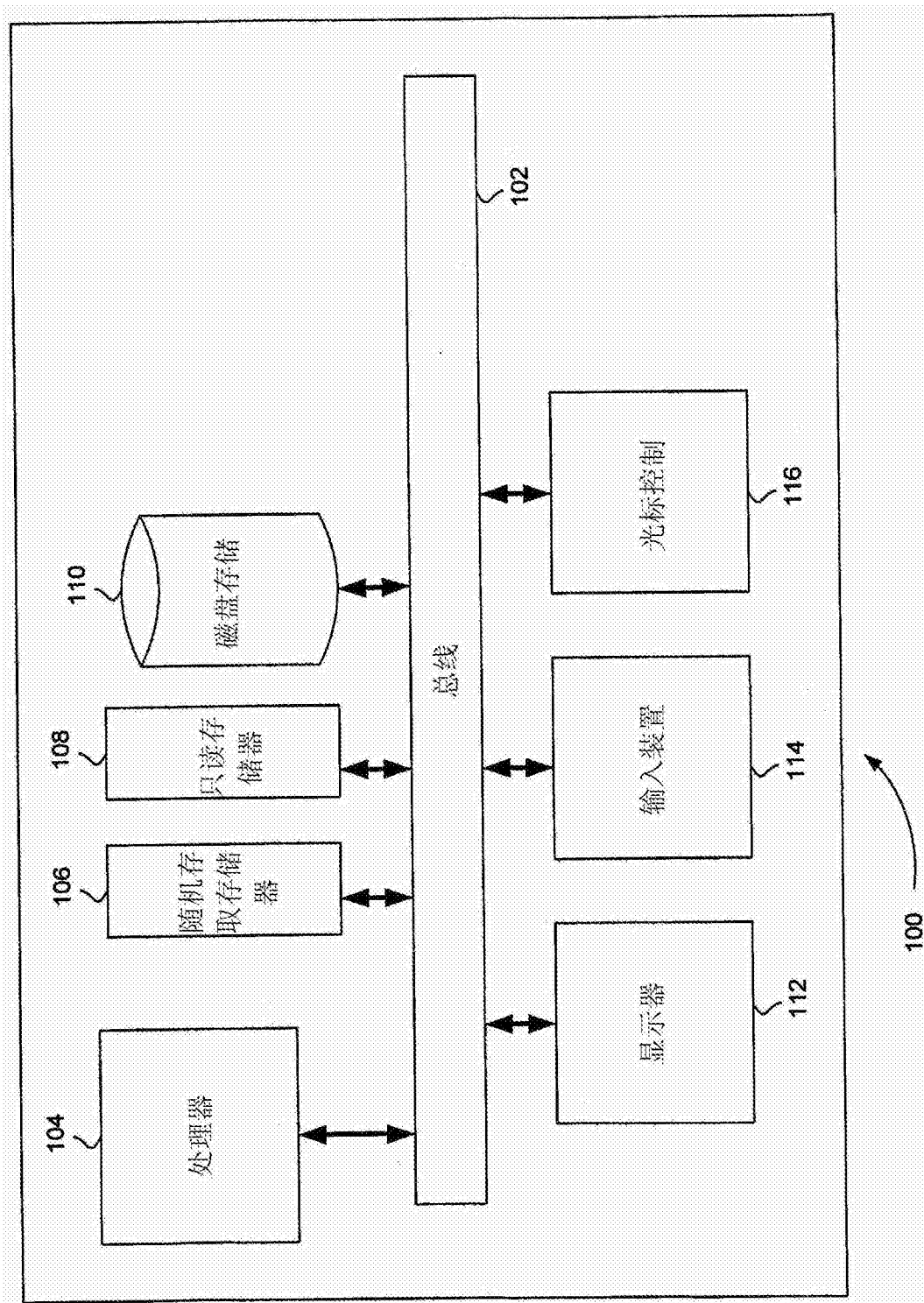


图1

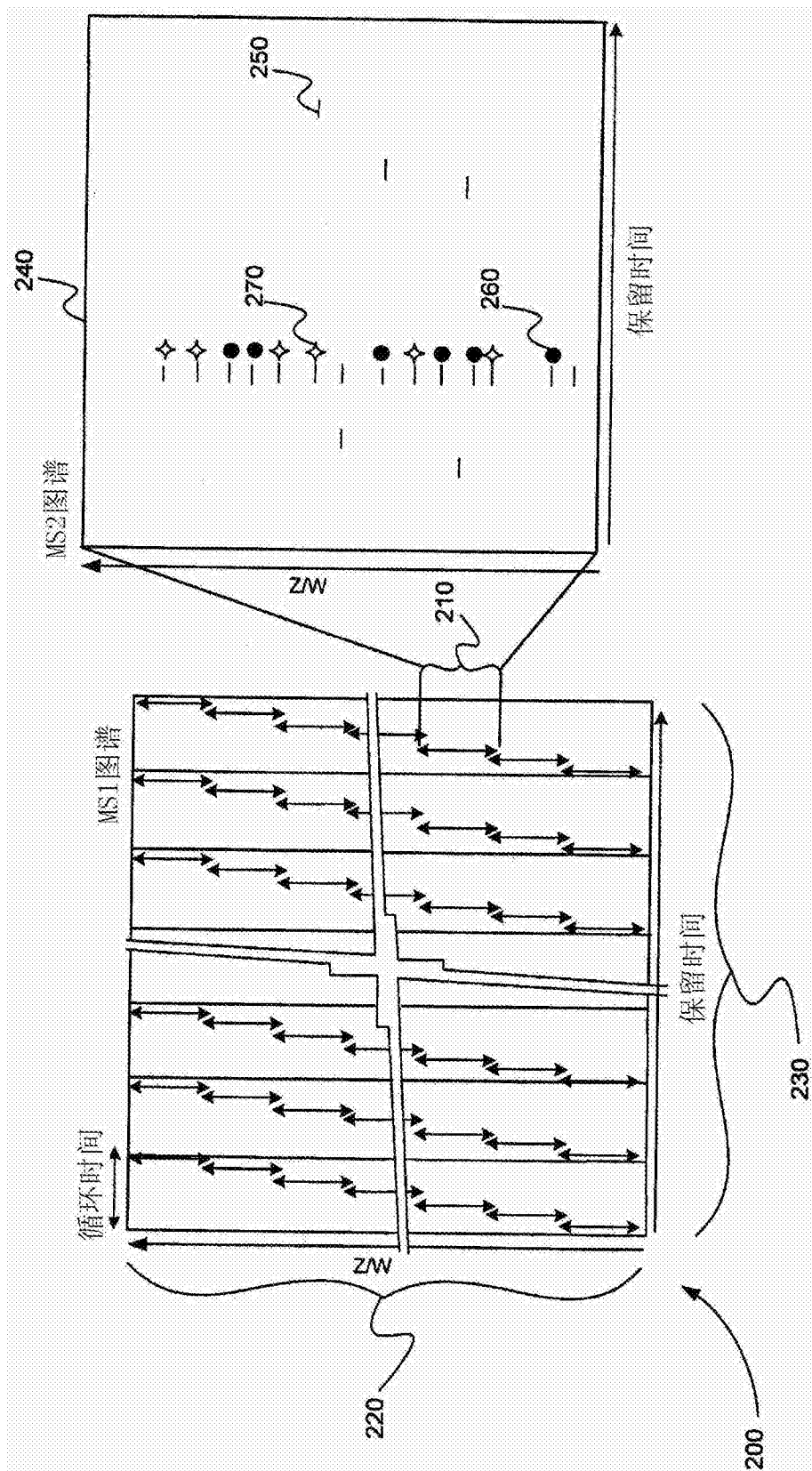


图2

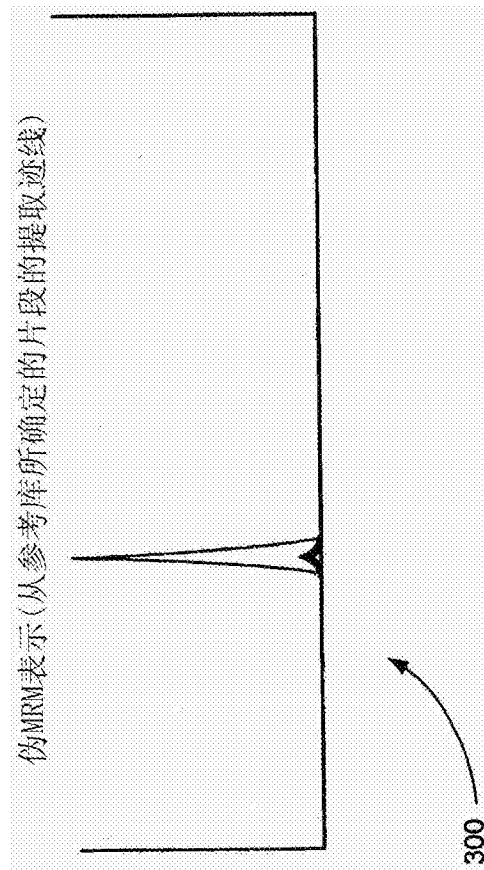


图3

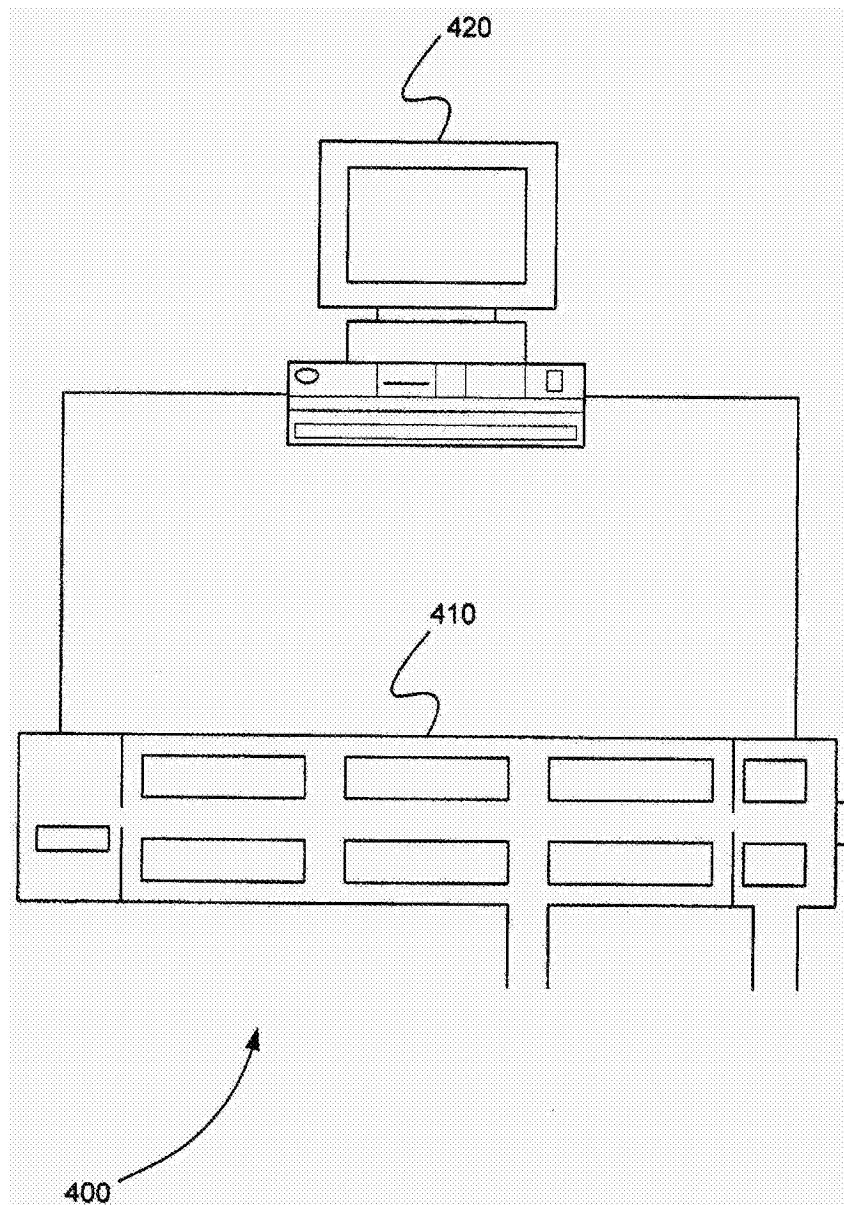


图4

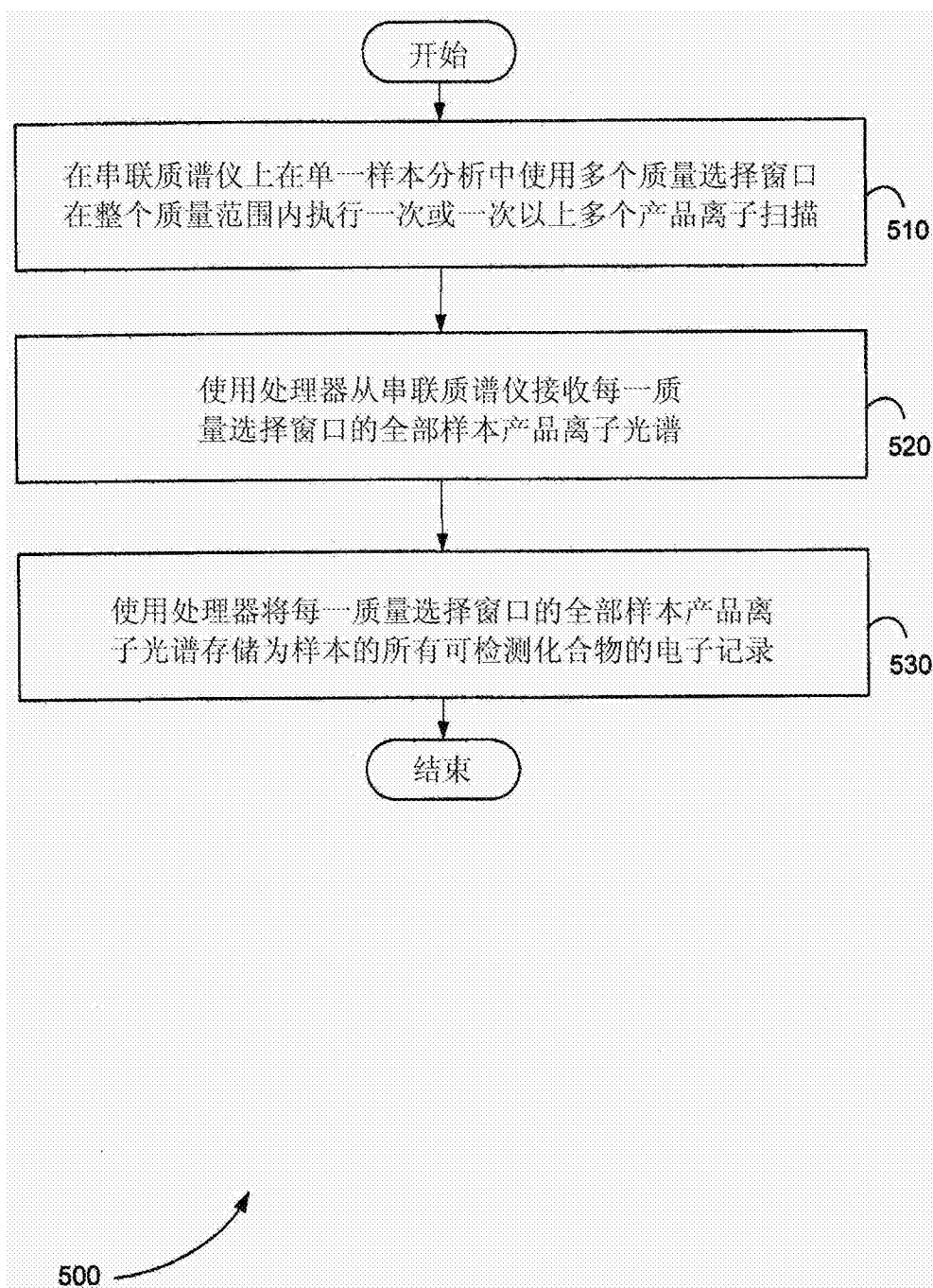


图5

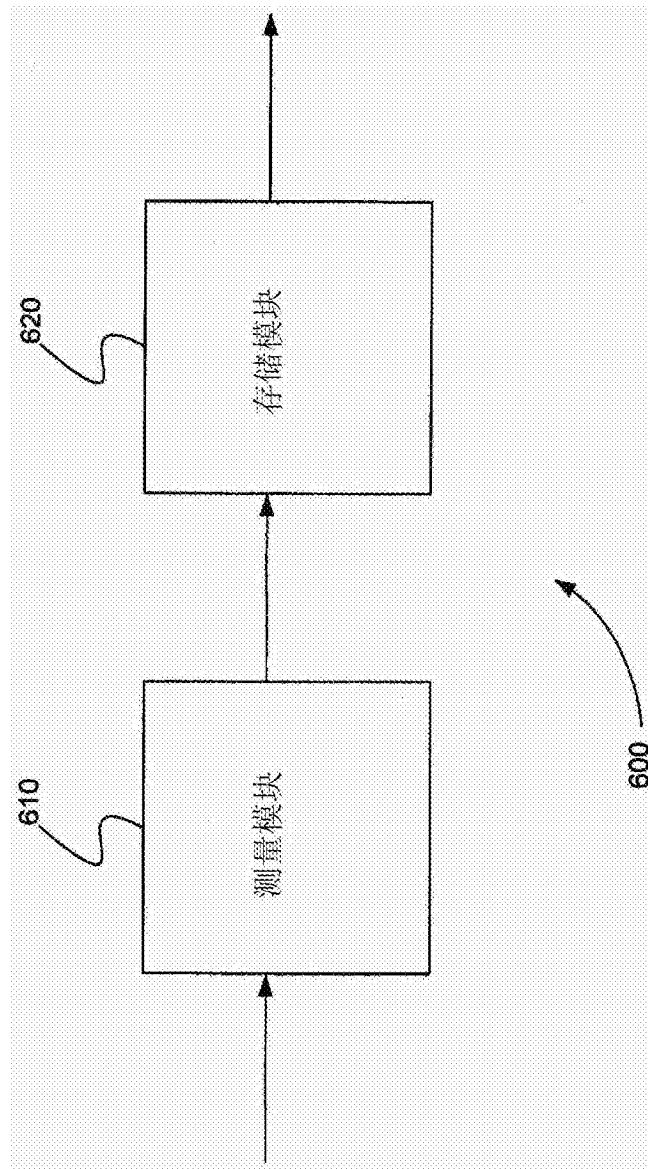


图6