

## 基于质谱技术的糖链结构解析研究进展\*

赵克力<sup>1,2)</sup> 黄纯翠<sup>1)</sup> 武红梅<sup>1,3)</sup> 李岩<sup>1,2)\*\*</sup>

(<sup>1)</sup>中国科学院生物物理研究所, 北京 100101; (<sup>2)</sup>中国科学院大学, 北京 100049; (<sup>3)</sup>广东中科康仪生物技术有限公司, 佛山 528000)

### 通讯作者简介

李岩, 1998年获得北京大学医学部药学学士学位, 2004年获得美国马里兰大学分析化学博士学位, 2005至2011年在美国约翰霍普金斯大学医学院病理系进行博士后培训, 并留任助理研究员, 2011年7月作为“中国科学院引进杰出技术人才”引进, 现任中国科学院生物物理研究所研究员, 博士生导师, 从事临床检验新方法新技术的开发应用. 主要从事物质谱与临床质谱技术的研究与转化. 研究方向主要包括以下三个方向: 发展适用于生物样品的高通量糖组学分析技术方法, 满足临床检验与生物制药等科研需要; 发展临床质谱检验技术用于疾病精准检测与辅助治疗, 进行产品化研发; 承担中国科学院中国人群前瞻性生物样本库建设任务, 进行中国人群基线数据采集等健康相关研究.

**摘要** 糖组学是继基因组学、蛋白质组学之后, 又一门新兴的学科, 其主要是研究糖分子的结构与功能. 糖是一类比核酸、蛋白质更加独特的生物分子, 它们不仅是生物体储存能量和释放能量的主要物质, 更是生物体内的信息传递分子, 并且在生理和病理过程中扮演着重要的角色, 如细胞间的识别作用、炎症以及自身免疫疾病等. 在结构上, 糖类物质更为复杂, 具有宏观不均一性(蛋白质上有多个糖基化位点)和微观不均一性(同一结合位点上可以连接不同的多糖), 所以糖链的结构解析一直是糖组学研究的难题. 相较于传统的分析方法, 质谱法具有高灵敏度、高精度、高通量等优势, 被认为是在糖链结构解析过程中重要的分析方法. 本文综述了质谱、多级质谱、液相色谱-质谱、毛细管电泳-质谱等方法在糖组学中糖链结构解析的研究进展.

**关键词** 糖组学, 聚糖, 质谱, 多级质谱, 结构分析  
**学科分类号** Q53, O65

**DOI:** 10.16476/j.pibb.2017.0008

过去人们一直认为糖是生物体内的储存能量和释放能量的主要物质, 但是随着研究的不断深入, 研究人员发现糖类分子是继核酸和蛋白质之后的第三类生物信息分子, 广泛地参与了生理病理活动, 如信息传递、细胞生长发育、免疫细胞调节、肿瘤发生发展等过程. 糖类分子可以根据其组成单元数量分为单糖、寡糖、多糖. 在生命体中, 糖类物质通常以糖复合物的形式出现, 即糖蛋白、蛋白聚糖、糖肽、糖脂等, 从而实现其生物学功能<sup>[1-4]</sup>.

糖组学(glycomics)与基因组学、蛋白质组学有着密切的关系, 被认为是基因组学和蛋白质组学的延续, 其主要研究目标是在大规模水平上研究糖类物质, 包括了对糖类物质的富集与分离, 糖链结构的鉴定和糖分子在生命系统中功能的研究. 根据基

因组和蛋白质组的定义, 将一个细胞或某个组织表达的全部糖链或是某一类特定的糖链称为糖组(glycome)<sup>[5-7]</sup>. 与核酸和蛋白质相比, 糖类分子的一级结构更为复杂, 不仅需要确定糖链的序列信息和单糖种类, 还需要确定其手性、连接方式和异头物构型等立体结构. 以相同数目单体组成的寡糖链通过不同的连接位点、不同的拓扑结构以及不同的

\* 国家重点基础研究发展计划(973)(2013CB910100), 国家自然科学基金(31270909), 国家高技术研究发展计划(863)(2014AA021101)和二维微流控糖芯片分析仪的研制(YZ201249)资助项目.

\*\* 通讯联系人.

Tel: 010-64887212, E-mail: yanli001@gmail.com

收稿日期: 2017-01-04, 接受日期: 2017-06-19

空间构型( $\alpha$ 、 $\beta$ )等形成的异构体数目就远远大于前两者。此外, 由于大多数糖类化合物是呈中性或酸性, 很容易形成聚合物( $n=1\sim 103$ )以及多种修饰结构, 如乙酰化、磷酸化、硫酸化等, 而且单糖的种类数目繁多<sup>[8-10]</sup>。所以, 糖链结构鉴定与解析是糖组学研究中最具有挑战性的难题之一。与蛋白质组学和代谢组学一样, 传统的分析技术如核磁共振、色谱、凝集素芯片等也应用到了糖类物质的分析, 由于糖类物质结构复杂, 具有宏观和微观不均一性, 所以在糖链结构研究过程中遇到了瓶颈<sup>[11]</sup>。而基于质谱的糖链解析技术, 由于其可以直接对糖链进行定性分析, 提供糖链的结构信息, 并且具有高灵敏度、高通量、高精度等优势, 特别是软电离技术的发展, 使得生物质谱成为糖组学研究的主要工具<sup>[12]</sup>。

近年来, 糖链结构解析和糖类物质生物学功能等方面的研究取得了不少进展。伴随着以质谱分析技术为平台, 糖链的释放与制备技术的不断提升, 以及各种分析方法的相继建立, 使得糖结构信息逐渐被解析出来。本文主要对生物质谱技术在糖结构分析中的应用及进展进行分析与总结。

## 1 糖链的制备与获取

绝大多数的糖类分子是以糖复合物的形式出现, 如糖蛋白、糖脂、糖肽、蛋白质聚糖等。在哺乳动物中, 糖类分子主要以糖蛋白的形式出现, 根据糖类分子与氨基酸连接方式的不同大致上可以分为3类: N-糖、O-糖以及葡萄糖胺聚糖(GAGs)。因此, 对每种糖链进行分析之前, 需要将其从蛋白质、脂质上释放出来。在糖链的定性分析中, 每种糖链的释放都有其独特或者首选的方法, 释放糖链的质量高低和效率都会影响后续实验的效果。所以, 从糖复合物中获取糖链是对其功能研究和结构解析至关重要的一步。目前, 糖链的释放主要有酶解法和化学方法两大类<sup>[5, 13-14]</sup>。

N-糖链一般是指糖链的还原性末端以N-糖肽键方式连接到肽链的天冬酰胺(Asn)上, 其核心结构是 $\text{Man}_3\text{GlcNAc}_2$ 。一般情况下, 对于N-糖链的获取采用酶解法, 通常使用肽-N-糖苷酶F(PNGase F), 特异性切断N-乙酰葡萄糖胺与天冬酰胺形成的糖肽键, 得到带有还原性末端的完整N-糖链。除了PNGase F外, 还有PNGase A和内切糖苷酶ENDO F1、F2、F3、ENDO H等, 这些酶也经常配合使用。此外, 非特异性的蛋白水解酶如链霉菌蛋白酶等也可以用来释放N-糖链, 通过控制

使用量和酶切时间, 可以将糖蛋白水解为低分子质量的糖肽甚至是只含有一个氨基酸的糖链, 从而解决长链蛋白对糖基化位点掩盖<sup>[15]</sup>。2017年, Chen等<sup>[16]</sup>利用人工改造后的Fbs1可以识别并结合N-糖基化糖肽中N-糖链的核心结构( $\text{M}_3\text{N}_2$ )的特性, 从而实现了N-糖的高度富集。与凝集素相比, 此方法大大地增加了N-糖链信息的覆盖性和完整性。

O-糖链是指糖链中某一糖基的羟基通过糖肽键与肽链的丝氨酸(Ser)或苏氨酸(Thr)相连。与N-糖相比, 其结构上更为复杂, 存在多种连接形式和多种核心结构, 例如通过N-乙酰葡萄糖胺与丝氨酸、苏氨酸形成的糖肽键, 除此之外半乳糖、甘露糖、木糖也可以与丝氨酸、苏氨酸共价结合形成O-糖肽键, 所以没有特征修饰序列作为其位点的预测。另外, 尚未发现可以将O-糖链整体从蛋白质或多肽上水解的内切酶。因此, 对于O-糖链的释放一般采用化学法, 即在碱性条件下通过 $\beta$ 消除反应完成, O-糖链的还原末端被还原成糖醇<sup>[17]</sup>。最近, Wu等<sup>[18-19]</sup>基于实验之前发明使用超滤膜富集细胞和血清中N-糖链的方法, 此次又使用10 ku超滤膜辅助O-糖链的富集, 利用分子质量的差异, 将 $\beta$ 消除后的O-糖进行分离富集, 再使用MALDI-MS对甲基化后的O-糖链进行分析。这种方法实现了快速和高通量地解析O-糖链, 并将此方法应用到了临床样本的分析中, 鉴定了血清、尿液中的O-糖。

糖脂是糖分子与脂质分子通过共价作用形成的糖复合物。根据脂质物质的不同, 又可以分为以糖链的还原末端与神经酰胺连接形成的鞘糖脂(glycosphingolipid-glycans, GSL), 以及与丙三醇或二酰基甘油连接形成的甘油糖脂(glycoglycerolipids)<sup>[20-21]</sup>。对于糖脂分子上糖链的释放一般会使用酶解法和化学法两种。Anugraham等<sup>[21]</sup>建立了将中性和酸性糖脂固定到PVDF膜上, 再使用神经酰胺糖内切酶II释放糖链的方法。并利用这种方法分析了子宫内膜癌血浆、组织、细胞系(107)中的鞘糖脂糖链结构。

有时候, 为了提高质谱检测的灵敏度, 常常需要对糖分子进行衍生化, 主要包括甲基化法、乙酰化法和还原氨化法。Desantos-Garcia等<sup>[22]</sup>就是对由胎球蛋白等糖蛋白上获取的N-糖进行了固相全甲基化的尝试, 这种方法不仅提高了衍生化效率、降低氢氧化钠的使用量、增强了信号强度, 而且此方法已经在人血清的糖组分析中得到应用。目前, 阻碍糖链的获取主要有两方面因素: 一方面是多糖不

能像核酸、蛋白质拥有可以编码的模板，从而进行扩增；另外，对不同连接方式的糖链难以分离，影响了糖链的获取，造成了糖链信息的丢失，进而阻碍了糖链结构的解析工作。Song 等<sup>[23]</sup>提出了一种氧化法获取糖链的策略，通过使用次氯酸钠 (NaClO) 以及控制反应的时间与条件，可以将不同连接方式的糖链从糖蛋白质、糖脂上游离下来，因此可以获得更为完整的糖组信息，从而便于对整个糖组进行分析。尽管此方法在低样品量下的重现性不是很好以及在实际样本(血清、组织)中的应用存在限制，但是也提供了一种新的获取糖链的方法，同时降低了糖链获取的成本。

### 2 糖链碎片在质谱中的命名方法

在质谱解析糖链结构方面，一般采用的是 Domon 和 Costello<sup>[24]</sup>提出的命名方法，这种方法准确地总结了糖链由串级质谱、多级质谱产生离子碎

片的规律和特点。糖链的断裂方式一般有两种，即糖苷键断裂和跨环断裂。如图 1 所示，其命名方法如：**a.** 在糖苷键断裂过程中，糖苷键的氧左侧键断裂后产生的非还原性端的糖链用字母 B 表示，还原端的糖链用字母 Y 表示，糖苷键的氧右侧键断裂后产生的非还原性端的糖链用字母 C 表示，还原性端糖链用字母 Z 表示。字母下方的角标表示糖苷键的断裂位置， $Y_0$  和  $Z_0$  代表还原端半缩醛羟基的氧左键或右键断裂产生的碎片。下标数字代表化学键断裂的位置， $\alpha$  表示含有单糖数量最多的糖链分支。**b.** 在跨环断裂中，非还原端部分用字母 A 表示，还原端部分用字母 X 表示，其角标表示发生跨环断裂的糖环位置，异头碳的位置为 1，顺时针依次标记。**c.** 若糖链结构中存在分支结构，使用字母 D 表示从含有分支结构的两端发生糖苷键断裂后产生的离子碎片，其角标表示两个糖苷键的位置。

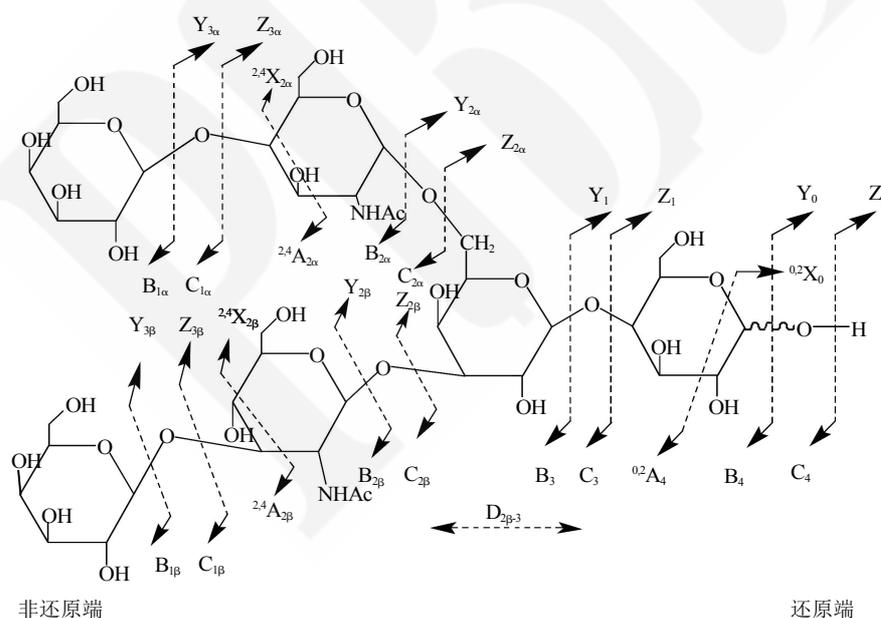


Fig. 1 The nomenclature for glycan fragments generated by multistage mass spectrometry

图 1 在多级质谱中寡糖链的断裂方式和命名

### 3 质谱技术在糖链分析中的应用

质谱(mass spectrometry, MS)的基本原理是基于分子或原子等粒子在离子源中发生电离，在加速电场的作用下使产生的离子进入质量分子器，再利用电场和磁场使离子发生色散和聚焦，得到质谱图

从而确定其质量。基质辅助激光解析(matrix-assisted laser desorption ionization, MALDI)和电喷雾离子源(electrospray ionization, ESI)是目前用于糖链电离的两种主要方法，并且在糖链离子化方面具有很大的互补性。此外，通过与多种高性能质量分析器(离子阱、时间飞行、三重四极杆、

傅里叶变换离子回旋共振等)串联后, 可以实现串联质谱(MS/MS)或多级质谱(MS<sup>n</sup>)分析, 从而获得更多的糖结构信息。

### 3.1 基于基质解析飞行时间质谱和电喷雾质谱的糖谱获取

基质辅助激光解析质谱(MALDI-MS)是一种广泛地用于糖组学分析的技术, 具有样品易于制备、数据采集速度快以及数据信息简单(观察到单电荷离子)等优点<sup>[25]</sup>。当对糖链结构进行分析时, 酸性末端是限制其准确分析的主要因素, 如末端的唾液酸, 常常在高能量的碰撞中丢失。MALDI-MS产生的离子常用飞行时间检测器(time of flight, TOF)进行检测, 这种组合成的质谱被称为基质辅助激光解析飞行时间质谱(MALDI-TOF-MS), 主要用于完全甲基化的糖链分析<sup>[26]</sup>。MALDI-TOF-MS具有灵敏度高、分子质量检测精确度高、检测速度快、成本低等特点。基于这些特征, Canis等<sup>[27]</sup>先用凝胶过滤和亲和色谱对血清中的血管性血友病因子(von willebrand factor, vWF)进行纯化, 并对酶解游离出的N-糖进行了正负离子两种模式的MALDI-TOF-MS分析。分析了300个人血管性血友病因子携带的N-糖谱, 同时也证明了H抗原不受N-糖基化位点的影响。此外, MALDI-TOF-MS也是用于大规模体外诊断和生物标志物研究的有力工具, 目前已用于分析癌症患者血清中N-糖组和O-糖组的组成和丰度的变化。例如, Park等<sup>[28-29]</sup>采用2, 5-二羟基苯甲酸(DHB)作为基质, 利用MALDI-TOF-MS和MALDI-QIT-TOF-MS分析正常胰腺导管上皮细胞和3种胰腺癌细胞系细胞膜表面的N-糖组, 通过对质谱数据的研究, 发现在肿瘤细胞中会出现高甘露糖化、高唾液酸化和高分支化等特征, 从而区分了正常细胞和胰腺癌细胞。Holst等<sup>[30]</sup>应用MALDI-TOF-TOF-MS对25种不同的结肠直肠癌细胞系的N-糖谱进行了大规模研究, 发现了 $\alpha$ -2, 3和 $\alpha$ -2, 6连接的唾液酸糖肽在不同细胞膜蛋白的改变, 并证明了高甘露糖型N-糖不仅存在于细胞内前体上, 也存在于细胞膜上。

电喷雾电离质谱(ESI-MS)不仅可以产生多电荷离子和更稳定的母离子, 获得比单电荷母离子更多的结构信息, 还可以补充MALDI-MS所获得的信息。与MALDI-MS相比, ESI-MS可以用于天然或衍生化多糖的分析。此外, 由于ESI产生的分子离子具有多电荷现象, 使得其质量检测范围可以成倍增加。但是对于ESI-MS而言, 样品的制备是十分

繁杂的, 主要是因为其容易受到样品添加剂、缓冲盐的影响<sup>[31]</sup>。因此, 常常需要与色谱、毛细管电泳等技术进行联用, 对组分为复杂的样品进行线上线下分离分析。Sethi等<sup>[32]</sup>通过使用PNGase F对3种不同表型的结肠癌细胞系中糖蛋白进行水解, 得到的N-糖采用LC-ESI-MS/MS在负离子模式下进行分析, 结果表明3种表型均有甘露糖高表达, 同时揭示了二天线 $\alpha$ -2, 3-唾液酸N-糖结构对于肿瘤细胞分型的影响。

糖链的糖谱分析可以从完整的分子离子峰中获取, 但是想要阐明单糖组成、序列以及连接方式等更多、更为详细糖链结构的信息, 就需要使用碰撞诱导解离技术(CID)、串联质谱或多级质谱技术了。

### 3.2 串级质谱/多级质谱

由于糖链结构的复杂性, 一级质谱数据可以检测到糖分子或糖复合物的质量数, 常规的二级质谱数据可以简单地判断确定糖链的线性结构。如Zhang等<sup>[33]</sup>使用ESI-CID-MS/MS分析不同血型抗原上的中性寡糖链结构, 为研究寡糖链对血型系统分型的影响提供了重要依据, 并将此方法推广到了主要血型抗原的鉴定, 如ABO、Lea、Leb等。而阐明某一种糖链结构需要知道其分子质量、单糖的种类、连接顺序、糖苷键的位置和构型以及环状结构的类型等, 而常规的二级质谱无法测定这些糖链的结构信息。因此, 对于糖链结构的完全解析, 则需要多级质谱的数据确定单糖组成、序列、糖链的连接方式、糖链拓扑结构等信息。在多级质谱模式下, 可以选择特定的母离子, 使用气体进行碰撞, 产生子离子碎片, 子离子碎片越小, 其包含的结构信息越少, 更有利于确定出其准确的结构。通过分析单糖与单糖之间的裂解信息, 可以确定糖链序列, 分析单糖的环内裂解信息, 可以确定糖链的分支情况。在确定碎片结构后, 将所有的碎片信息重组, 可以获得完整的糖链结构信息。Reinhold实验室<sup>[34-36]</sup>就应用多级质谱对糖链进行测序, 并提供了三种分析的策略和算法, 并对糖链的连接位点、单体、空间构型等信息进行了推测与判断。在多级质谱分析鉴定过程中, 其主要的困难在于如何整合质谱数据, 以及如何选择特定的母离子进行分析。因此需要借助计算机软件的开发, 就目前已有一些商业化、公开化的软件, 如Cartoonist、Glycoworkbench、STAT、StrOligo和GlycoSeq等<sup>[37-39]</sup>, 可以很好地分析由多级质谱产生的离子碎片信息, 降低了质谱实验时间和人工解析糖链结构的时间, 使得糖链结

构可以朝着串联质谱、多级质谱全面解析的方向发展。利用基于碰撞诱导解离技术(CID)的串联质谱分析方法和 GlycoSeq 软件, Yu 等<sup>[40]</sup>对从乳腺癌细胞系中获取的 N-糖链进行了分析和结构验证, 并且实现了此种策略的自动化分析。

虽然碰撞诱导解离技术是经典和已经广泛地应用在串联质谱中的解离方法, 但仍存在信息单一、缺少环内裂解碎片信息等问题。随着解离技术的发展, 在进行糖多级质谱实验时已经开始尝试使用新的解离技术, 如电子转移解离技术(ETD)和高能量的碰撞解离技术(HCD)等<sup>[41-42]</sup>。目前, 已有实验室将 ETD 应用到解析牛乳寡糖链结构中, 并获得一些新的关于糖链结构的信息, 同时也阐明了 ETD 产生离子碎片的机制以及金属离子依赖性。此外, 在对样品分析过程中, 发现某些离子的碎片信息也存在 N-糖和 O-糖中, 如乙酰葡萄糖胺和唾液酸残基等碎片信息<sup>[43]</sup>。在 CID 谱图中, 被检测到分子离子峰主要是不同位点的糖苷键断裂的信息, ETD 可以获得更多的环内裂解碎片信息, 这些补充信息可以更好地阐释单糖间的连接方式和拓扑结构。还有实验室尝试使用 ETD 对 O-糖组进行了解析, 尽管 ETD-MS<sup>2</sup>、ETD-MS<sup>3</sup> 对低丰度的母离子峰也具有着良好的响应, 但是此技术仍受制于 O-糖链长度或结合蛋白的大小。尽管如此, ETD 仍是一种潜在的、强有力的糖链分析技术。如果能够得到较为完整的离子碎片信息, 就可以直接从质谱图中推断出糖链结构和糖基化位点。

#### 4 基于质谱联用技术在糖链分析中的应用

由于单一质谱不能够区分相同分子质量的糖类化合物, 也不能提供准确的糖苷键信息和原子数量, 使得单一质谱法面临着巨大的挑战。相反地, 液相色谱和毛细管电泳可以解决相同分子质量的糖类化合物区分的问题, 但不能直接提供结构上的信息。因此, 质谱与液相色谱、毛细管电泳的联用, 不仅可以获得更高质量的信息, 而且多种技术的联用可以更迅速更准确地对糖链进行分析与鉴定。

##### 4.1 液相色谱-质谱联用(LC-MS)

液相色谱(liquid chromatography, LC)是最常见和有效的分离分析技术。分离维度的增加可以提高分析的准确度, 根据保留时间和质量位置可以获得更多的糖链信息。目前常使用的色谱包括亲水相互作用色谱(HILIC)、反相液相色谱(PR)、多孔石墨化碳色谱(PGC)等。与电喷雾质谱的结合, 可以实

现复杂样本的线上分离分析。相较于其他色谱柱 PGC 是一种应用比较广泛的技术, 不仅仅可以作为固相萃取材料, 而且也可以作为液相色谱的固定相, 用于天然多糖和还原性多糖的分析。PGC 具有较高的区分聚糖同分异构体的能力, 例如对于  $\alpha$ -2, 3/6 位唾液酸化聚糖的分离以及二天线复合 N-连接寡糖中乙酰葡萄糖胺的识别<sup>[44]</sup>。通过与质谱技术(如 ESI)相结合对某些特殊糖类分子的分析具有更高的灵敏性和比较稳定的实验重现性。有实验室采用 PGC-LC-ESI-MS/MS 的方法测定了来自于溃疡性结肠炎患者的糖组进行分析, 通过对质谱色谱图的分析, 确定了三种相对分子质量为 1827.78 u 的寡糖链结构, 并发现了平分型 N-乙酰葡萄糖胺末端对于保留时间的影响<sup>[45]</sup>。目前, PGC-LC-ESI-MS/MS 已经成为临床质谱实验中主要方法之一, 如肿瘤标志物的发现和治疗性重组免疫球蛋白表征分析。最近, Zhu 所在实验室<sup>[46]</sup>开发了一种基于 96 色谱柱的方法, 分析了肝硬化与肝癌患者血清中结合珠蛋白异常糖基化修饰的方法, 通过 HPLC 将结合珠蛋白分离纯化, 并且在 96 孔板中完成了获取 N-糖链的过程, 在已负载多孔石墨化碳的 96 孔板中进行了甲基化反应, 结合 MALDI-QIT-TOF-MS 确定了 N-糖链中岩藻糖的变化, 此过程仅仅需要 10  $\mu$ l 的血清样品量, 这不仅降低了样品的消耗量和实验时间, 还扩大了样本分析的数量。随着液相色谱技术和质谱技术的革新和应用, 为糖链分析和结构解析提供了技术上的支持, 使其可以朝着分析自动化、一体化、高通量化和临床化应用的方向发展<sup>[47]</sup>。

##### 4.2 毛细管电泳-质谱联用(CE-MS)

毛细管电泳(capillary electrophoresis, CE)具有高通量以及高分辨率等特点, 通过与其联用提高了质谱检测速度和准确性。CE-MS 可以在一次分析中获得迁移时间、分子质量和母离子-子离子碎片信息。因此, CE-MS 技术可以用于生命体内的低丰度糖蛋白、蛋白聚糖中糖链的结构分析。但是, 并不是所有的毛细管电泳分离模式可以直接与质谱联用, 需要进一步研究毛细管电泳与质谱的接口技术。近些年, 一些技术上的进步, 保证了系统对糖链结构的快速且稳定的分析<sup>[48]</sup>。Mechref 等<sup>[49]</sup>系统地介绍了与糖链分析和结构解析相关的毛细管电泳-质谱接口技术研究进展。Sun 等<sup>[50]</sup>采用一种新型电动压力的接口技术用于 CE-ESI-MS 中, 通过使用正向分离模式的毛细管电泳和正离子模式的

电喷雾质谱, 分析了肝素寡糖和低分子量的肝素, 并鉴定了其二糖结构, 增加了实验的重现性. 由于糖类物质既不带电, 也不含有紫外吸收基团和荧光基团, 所以需要糖进行衍生化, 使其可以进行毛细管电泳实验. Tie 等<sup>[51]</sup>提供了一类以三嗪为基础结构的新型衍生化试剂, 从而提高了毛细管电泳的分离效率和质谱的检测性能. 尽管如此, 毛细管电泳仍存在处理样品量低、溶剂难挥发以及与质谱兼容性等问题, 使得该技术在糖链结构解析中受到一定的限制.

## 5 糖链的定量分析

糖链的研究除了要获取不同生物样本中糖链的种类和结构信息外, 对于糖链的定量分析在解释和揭示糖链的生物学功能和与生理病理的关系等方面具有重要意义, 是糖链分析与实际应用的重要手段和桥梁. 生物质谱检测技术不仅为糖链结构的研究提供帮助, 同时也为糖链定量分析提供了支持. 目前, 糖链的定量研究主要包括了绝对定量和相对定量两种. 绝对定量需要获得待测组内某种糖链或是所有糖链的绝对含量, 需要待测糖链的标准品才能进行分析. 然而因为糖链结构的多样性以及微观结构的不均一性, 生物样品中糖链组成复杂, 标准品难获得等问题, 使得对糖链的绝对定量变得十分困难<sup>[52]</sup>. 而 Echeverria 等<sup>[53]</sup>首先使用化学酶法合成了  $^{13}\text{C}$  标记的 N-糖链, 随之建立了标记后的 N-糖数据库. 将标记后的 N-糖作为内标加入到样品中, 并与 MALDI-TOF 质谱结合, 成功地对 IgG 1 中二天线结构的 N-糖链进行了定性和绝对定量分析. 相反, 相对定量则不需要知道某种糖链或是所有糖链的绝对含量, 只需要知道某种糖链或所有糖链在两个或多个样本中的相对值. 相对定量通常会在样品中加入内标. 在各类内标中, 一般会选用同位素标记( $^{13}\text{C}$ 、 $^{15}\text{N}$ )进行实验. 利用稳定同位素标记进行糖链的相对定量, 是将两组或多组样品分别标记, 再按比例混合进行质谱分析, 利用同位素标记的糖链具有相似的化学特性, 从而形成一组特征峰, 通过分析不同来源的同种糖链的丰度从而达到相对定量. Váradi 等<sup>[54]</sup>建立了使用( $^{12/13}\text{C}$ ) 2-AA 标记对 N-糖链定量的方法. 首先使用( $^{12/13}\text{C}$ ) 2-AA 对从样品中获取的 N-糖链进行衍生化反应, HILIC 色谱进行纯化, 然后样品按照比例混合, 再利用毛细管电泳-质谱联用的方式分析混合物中  $^{12}\text{C}$ 、 $^{13}\text{C}$  的比例, 实现了对 N-糖链的相对定量分析, 并完成了

对西妥昔单抗 Fc 端糖链的定量分析. 然而在糖组研究的过程中, 如何对不同连接类型的糖链进行定量分析也是一个难点. Zhang 实验室<sup>[55-56]</sup>开发了一种 GIG (glycoprotein immobilization for glycan extraction) 富集糖链的技术, 可以同时 N-糖和 O-糖进行定性和定量分析. 首先通过胺化反应将糖蛋白固定到含有醛基活性物质的固相介质上, 然后通过 PNGase F 释放 N-糖链, 并采用 QUANTITY (quaternary amine containing isobaric tag for glycan) 进行标记,  $\beta$ -消除释放的 O-糖链被 1-苯基-3-甲基-5-吡啶啉酮(PMP)标记, 最后通过串联质谱完成了糖链的定性和定量分析.

## 6 总结和展望

蛋白质的糖基化作为调控蛋白质功能的重要因素之一, 糖基化位点的改变或者糖链结构的改变都会引起蛋白质功能的改变. 已有研究表明, 糖复合物中糖类分子的异常表达常常出现于肿瘤的发生、发展、转移等过程中. 而目前临床上常用的肿瘤标志物如 CA125、CA242 等都是异常糖基化修饰的糖蛋白, 或者是如 CA19-9、CA15-3 等异常的糖链抗原<sup>[57-58]</sup>. 但是, 这些糖复合物中糖链的准确结构还不是很清楚, 并且也不能表现出肿瘤特异性和器官特异性. 所以, 需要更好地发展糖组学, 将糖组学上升到与基因组学、蛋白质组学、代谢组学同样的高度, 从糖组学的角度解决生物学问题、医学问题, 如细胞间的相互作用、肿瘤标志物的发现等, 相信这些未知的领域将会是研究的重点与热点. 在糖链结构的解析过程, 质谱既可以获得糖链的结构信息, 又可以与其他技术的联用, 如色谱技术对糖链进行定性定量分析. 随着生物质谱的进一步普及与不断发展, 质谱数据解析的进一步深入挖掘和研究, 质谱技术也在糖组学研究中呈现出了自动化、高通量化的发展趋势, 也越来越多地用于与疾病相关的糖链结构分析中, 成为寻找糖类分子作为新型标志物的重要工具, 从而为疾病的诊断、治疗提供了帮助.

## 参 考 文 献

- [1] Bond M R, Hanover J A. A little sugar goes a long way: the cell biology of O-GlcNAc. *J Cell Biol*, 2015, **208**(7): 869-880
- [2] Swamy M, Pathak S, Grzes K M, *et al.* Glucose and glutamine fuel protein O-GlcNAcylation to control T cell self-renewal and malignancy. *Nat Immunol*, 2016, **17**(6): 712-720
- [3] Antonopoulos A, North S J, Haslam S M, *et al.* Glycosylation of

- mouse and human immune cells: insights emerging from N-glycomics analyses. *Biochem Soc Trans*, 2011, **39** (5): 1334–1340
- [4] Hennet T, Cabalzar J. Congenital disorders of glycosylation: a concise chart of glycoalyx dysfunction. *Trends in Biochemical Sciences*, 2015, **40**(7): 377–384
- [5] Cummings R D, Pierce J M. The challenge and promise of glycomics. *Chemistry & Biology*, 2014, **21**(1): 1–15
- [6] Thaysen-Andersen M, Packer N H. Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. *Biochim Biophys Acta*, 2014, **1844**(9): 1437–1452
- [7] Aoki-Kinoshita K F. An introduction to bioinformatics for glycomics research. *PLoS Comput Biol*, 2008, **4**(5): e1000075
- [8] Robinson L N, Artpradit C, Raman R, *et al.* Harnessing glycomics technologies: integrating structure with function for glycan characterization. *Electrophoresis*, 2012, **33**(5): 797–814
- [9] Alley W R, Jr., Novotny M V. Structural glycomics analyses at high sensitivity: a decade of progress. *Annu Rev Anal Chem (Palo Alto Calif)*, 2013, **6**: 237–265
- [10] Corfield A P, Berry M. Glycan variation and evolution in the eukaryotes. *Trends in Biochemical Sciences*, 2015, **40**(7): 351–359
- [11] Irague R, Massou S, Moulis C, *et al.* NMR-based structural glycomics for high-throughput screening of carbohydrate-active enzyme specificity. *Analytical Chemistry*, 2011, **83**(4): 1202–1206
- [12] Lössl P, Van De Waterbeemd M, Heck A J R. The diverse and expanding role of mass spectrometry in structural and molecular biology. *The EMBO Journal*, 2016, **35**(24): 2634–2657
- [13] Christopher A. Waddling T H P, Jr., Anthony L. Tarentino, and patrick van roey structural basis for the substrate specificity of endo- $\beta$ -N-acetylglucosaminidase F3. *Biochemistry*, 2000, **39** (27): 7878–7885
- [14] Yang S, Rubin A, Eshghi S T, *et al.* Chemoenzymatic method for glycomics: Isolation, identification, and quantitation. *Proteomics*, 2016, **16**(2): 241–256
- [15] Stavenhagen K, Plomp R, Wuhrer M. Site-specific protein N- and O-glycosylation analysis by a C18-porous graphitized carbon-liquid chromatography-electrospray ionization mass spectrometry approach using pronase treated glycopeptides. *Anal Chem*, 2015, **87**(23): 11691–11699
- [16] Chen M, Shi X, Duke R M, *et al.* An engineered high affinity Fbs1 carbohydrate binding protein for selective capture of N-glycans and N-glycopeptides. *Nat Commun*, 2017, **8**: 15487
- [17] Levery S B, Steentoft C, Halim A, *et al.* Advances in mass spectrometry driven O-glycoproteomics. *Biochim Biophys Acta*, 2015, **1850**(1): 33–42
- [18] 武艳丽, 杨刚龙, 缪明永, 等. 基于超滤膜辅助的糖蛋白全 O-连接糖链的富集和 MALDI-TOF/TOF 质谱结构解析. *生物化学与生物物理进展*, 2017, **44**(1): 70–79
- Wu Y L, Yang G L, Miao M Y, *et al.* *Prog Biochem Biophys*, 2017, **44**(1): 70–79
- [19] Yang G L, Ma T R, Li Z. Enrichment and characterization of total N-linked glycans from glycoproteins by ultrafiltration units and mass spectrometry. *Progress in Biochemistry and Biophysics*, 2014, **41**(4): 403–408
- [20] Du H Q, Yu H J, Jia L Y, *et al.* Advances in analysis of glycosphingolipids and related diseases. *Progress in Biochemistry & Biophysics*, 2016, **43**(1): 14–27
- [21] Anugraham M, Everest-Dass A V, Jacob F, *et al.* A platform for the structural characterization of glycans enzymatically released from glycosphingolipids extracted from tissue and cells. *Rapid Commun Mass Spectrom*, 2015, **29**(7): 545–561
- [22] Desantos-Garcia J L, Khalil S I, Hussein A, *et al.* Enhanced sensitivity of LC-MS analysis of permethylated N-glycans through online purification. *Electrophoresis*, 2011, **32**(24): 3516–3525
- [23] Song X, Ju H, Lasanajak Y, *et al.* Oxidative release of natural glycans for functional glycomics. *Nat Methods*, 2016, **13**(6): 528–534
- [24] Domon B, Costello C E. Asystematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal*, 1988, **5**(4): 397–409
- [25] Harvey D J. Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: an update for 2009–2010. *Mass Spectrom Rev*, 2015, **34**(3): 268–422
- [26] Wada Y, Azadi P, Costello C E, *et al.* Comparison of the methods for profiling glycoprotein glycans—HUPO human disease glycomics/proteome initiative multi-institutional study. *Glycobiology*, 2007, **17**(4): 411–422
- [27] Canis K, Mckinnon T A, Nowak A, *et al.* Mapping the N-glycome of human von Willebrand factor. *Biochem J*, 2012, **447** (2): 217–228
- [28] Park H M, Hwang M P, Kim Y W, *et al.* Mass spectrometry-based N-linked glycomics profiling as a means for tracking pancreatic cancer metastasis. *Carbohydr Res*, 2015, **413**: 5–11
- [29] Mechref Y, Hu Y, Garcia A, *et al.* Identifying cancer biomarkers by mass spectrometry-based glycomics. *Electrophoresis*, 2012, **33**(12): 1755–1767
- [30] Holst S, Deuss A J, Van Pelt G W, *et al.* N-glycosylation profiling of colorectal cancer cell lines reveals association of fucosylation with differentiation and caudal type homeobox 1 (CDX1)/villin mRNA expression. *Mol Cell Proteomics*, 2016, **15**(1): 124–140
- [31] Zaia J. Mass spectrometry and the emerging field of glycomics. *Chem Biol*, 2008, **15**(9): 881–892
- [32] Sethi M K, Thaysen-Andersen M, Smith J T, *et al.* Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and alpha-2,3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines. *J Proteome Res*, 2014, **13**(1): 277–288
- [33] Zhang H, Zhang S, Tao G, *et al.* Typing of blood-group antigens on neutral oligosaccharides by negative-ion electrospray ionization tandem mass spectrometry. *Anal Chem*, 2013, **85**(12): 5940–5949
- [34] Ashline D, Singh S, Hanneman A, *et al.* Congruent strategies for carbohydrate sequencing. 1. mining structural details by mSn. *Anal Chem*, 2005, **77**(19): 6250–6262

- [35] Lapadula A J, Hatcher P J, Hanneman A J, *et al.* Congruent strategies for carbohydrate sequencing. 3. OSCAR: an algorithm for assigning oligosaccharide topology from MSn data. *Anal Chem*, 2005, **77**(19): 6271–6279
- [36] Zhang H, Singh S, Reinhold V N. Congruent strategies for carbohydrate sequencing. 2. FragLib: an MSn spectral library. *Anal Chem*, 2005, **77**(19): 6263–6270
- [37] An H J, Lebrilla C B. Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial). *Mass Spectrom Rev*, 2011, **30**(4): 560–578
- [38] Ceroni A, Maass K, Geyer H, *et al.* GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *Journal of Proteome Research*, 2008, **7**(4): 1650–1659
- [39] Ethier M, Saba J A, Spearman M, *et al.* Application of the StrOligo algorithm for the automated structure assignment of complex N-linked glycans from glycoproteins using tandem mass spectrometry. *Rapid Commun Mass Spectrom*, 2003, **17** (24): 2713–2720
- [40] Yu C Y, Mayampurath A, Zhu R, *et al.* Automated glycan sequencing from tandem mass spectra of N-linked glycopeptides. *Anal Chem*, 2016, **88**(11): 5725–5732
- [41] Segu Z M, Mechref Y. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Commun Mass Spectrom*, 2010, **24**(9): 1217–1225
- [42] Wen-Feng Z, Yang Z, Ming-Qi L, *et al.* Trends in mass spectrometry-based large-scale N-glycopeptides analysis. *Progress in Biochemistry and Biophysics*, 2016, **43**(6): 550–562
- [43] Han L, Costello C E. Electron transfer dissociation of milk oligosaccharides. *J Am Soc Mass Spectrom*, 2011, **22**(6): 997–1013
- [44] Everest-Dass A V, Abrahams J L, Kolarich D, *et al.* Structural feature ions for distinguishing N- and O-linked glycan isomers by LC-ESI-IT MS/MS. *J Am Soc Mass Spectrom*, 2013, **24** (6): 895–906
- [45] Stavenhagen K, Kolarich D, Wuhrer M. Clinical glycomics employing graphitized carbon liquid chromatography-mass spectrometry. *Chromatographia*, 2015, **78**(5–6): 307–320
- [46] Zhu J, Wu J, Yin H, *et al.* Mass spectrometric N-glycan analysis of haptoglobin from patient serum samples using a 96-well plate format. *J Proteome Res*, 2015, **14**(11): 4932–4939
- [47] Shubhakar A, Reiding K R, Gardner R A, *et al.* High-throughput analysis and automation for glycomics studies. *Chromatographia*, 2015, **78**(5–6): 321–333
- [48] Maxwell E J, Ratnayake C, Jayo R, *et al.* A promising capillary electrophoresis-electrospray ionization-mass spectrometry method for carbohydrate analysis. *Electrophoresis*, 2011, **32** (16): 2161–2166
- [49] Mechref Y. Analysis of glycans derived from glycoconjugates by capillary electrophoresis-mass spectrometry. *Electrophoresis*, 2011, **32**(24): 3467–3481
- [50] Sun X, Lin L, Liu X, *et al.* Capillary electrophoresis-mass spectrometry for the analysis of heparin oligosaccharides and low molecular weight heparin. *Anal Chem*, 2016, **88**(3): 1937–1943
- [51] Tie C, Zhang X-X. A new labelling reagent for glycans analysis by capillary electrophoresis-mass spectrometry. *Analytical Methods*, 2012, **4**(2): 357
- [52] Etxebarria J, Reichardt N C. Methods for the absolute quantification of N-glycan biomarkers. *Biochim Biophys Acta*, 2016, **1860** (8): 1676–1687
- [53] Echeverria B, Etxebarria J, Ruiz N, *et al.* Chemo-enzymatic synthesis of (13)C labeled complex N-glycans as internal standards for the absolute glycan quantification by mass spectrometry. *Anal Chem*, 2015, **87**(22): 11460–11467
- [54] Váradi C, Mittermayr S, Millánmartín S, *et al.* Quantitative twoplex glycan analysis using (12)C6 and (13)C6 stable isotope 2-aminobenzoic acid labelling and capillary electrophoresis mass spectrometry. *Analytical & Bioanalytical Chemistry*, 2016, 1–10
- [55] Yang S, Hu Y, Sokoll L, *et al.* Simultaneous quantification of N- and O-glycans using a solid-phase method. *Nat Protoc*, 2017, **12**(6): 1229–1244
- [56] Yang S, Hoti N, Yang W, *et al.* Simultaneous analyses of N-linked and O-linked glycans of ovarian cancer cells using solid-phase chemoenzymatic method. *Clin Proteomics*, 2017, **14**: 3
- [57] Stowell S R, Ju T, Cummings R D. Protein glycosylation in cancer. *Annual Review of Pathology*, 2015, **10**: 473–510
- [58] Kirwan A, Utratna M, O'dwyer M E, *et al.* Glycosylation-based serum biomarkers for cancer diagnostics and prognostics. *Biomed Res Int*, 2015, **2015**: 490531

## Progress in Understanding of Glycan Structure Analysis Based on Mass Spectrometry\*

ZHAO Ke-Li<sup>1,2)</sup>, HUANG Chun-Cui<sup>1)</sup>, WU Hong-Mei<sup>1,3)</sup>, LI Yan<sup>1,2)\*\*</sup>

<sup>1)</sup> Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China;

<sup>2)</sup> University of Chinese Academy of Sciences, Beijing 100049, China; <sup>3)</sup> Guangdong Bio-heattech Advanced Co., Ltd., Foshan 528000, China)

**Abstract** Glycomics represents a novel omics technology developed after the genomics and proteomics, which focus on the glycan structure and function. As a group of biomolecule, glycans provide and store energy within the cell, and play many significant roles in physiological and pathological processes, such as cell-cell recognition and inflammatory and autoimmune diseases. Structural analysis of glycan is the bottleneck of glycomics research, due to their macro-heterogeneity (site occupancy) and micro-heterogeneity (different glycoforms attached to one glycosylation site). Comparing with traditional analysis approaches, mass spectrometry has become the method of choice for structural analysis methods in glycan, thanks to its ability to require minute amounts of sample, gain more different information and to be used in combination with high-throughput analysis methods. This paper summarizes the advance in our understanding of glycan structures using mass spectrometry (MS), including multistage mass spectrometry (MS<sup>n</sup>), LC-MS and CE-MS.

**Key words** glycomics, glycan, mass spectrometry, multistage mass spectrometry, structure analysis

**DOI:** 10.16476/j.pibb.2017.0008

---

\* This work was supported by grants from National Basic Research Program of China (2013CB910100), The National Natural Science Foundation of China (31270909), Hi-Tech Research and Development Project of China(2014AA021101) and Scientific Equipment Development Project of Chinese Academy of Sciences (YZ201249).

\*\*Corresponding author.

Tel: 86-10-64887212, E-mail: yanli001@gmail.com

Received: January 4, 2017 Accepted: June 19, 2017