蛋白质的糖组学研究进展

王 山 李 钰*

(哈尔滨工业大学生命科学与工程系，哈尔滨 150001)

摘要 蛋白质的糖组学主要研究单个个体所包含的所有糖蛋白上聚糖的结构、功能等生物学作用。糖组学的出现使人类可以更深刻理解第三类生物信息大分子——聚糖在生命活动中的作用。目前，糖组学的发展仍处于初步阶段，基于DNA测序仪的荧光糖电泳、糖芯片等新技术的出现和质谱技术的应用对糖组学研究起到了很大的推动作用。

关键词 糖组学；基于DNA测序仪的荧光糖电泳；质谱；糖芯片；表层质子共振

基因组计划的提出到现在已有19年的时间了，截止到2005年1月已完成了包括人类基因组、秀丽隐杆线虫、小鼠和水稻等物种的基因组测序工作，同时猕猴、猪、牛、狗、家蚕等物种的基因组测序工作正在紧锣密鼓地的进行中[1,2]。基因序列的解析只是人类全面了解生物的一块奠基石，尚有许多工作需要研究。因此，在基因组计划提出8年后，澳大利亚科学家提出了全面解析生物所有蛋白质的蛋白质组学。随着研究的深入，科学家发现在复杂的细胞活动中，除了DNA与蛋白质扮演了绝对重要的作用外，糖类物质同样起着相当重要的作用，例如有8种人类遗传疾病是由N-糖链合成缺陷造成的[3]。于是有学者提出了以研究细胞中糖蛋白上糖链的结构和功能为主要内容的糖组学计划[4,5]，得到了科学家们广泛的响应。至此，人类在认识生命本质的道路上，形成了基因组-蛋白质组-糖组共同发展的研究模式。

1 糖链修饰的重要性

细胞中存在各种各样的糖复合物：糖蛋白、糖脂和蛋白聚糖等等，聚糖对于这些糖复合物的生物学功能起着至关重要的决定性作用。在细胞中，几乎所有的聚糖都是由甘露糖、果糖、N-乙酰葡萄糖胺等十几种单糖组成的。然而这十几种单糖包含的信息量是相当巨大的，3个己糖最多可以有27 648种可能的排列方式，而同样数目的核苷酸和氨基酸所组成的所有可能的序列数仅为6种。因此，寡糖链有可能靠多变的连接方式和分支结构而包含了相当量的重要的生物信息。

正因为聚糖包含了如此多的信息量，一旦在细胞的蛋白质后修饰过程中发生糖链修饰紊乱，则会产生相当严重的后果。研究发现糖基化修饰作用在肿瘤性钙质沉着症[6]和阿尔茨海默病[7,8]中均发生了改变。而且糖链对于维持蛋白质在体内的稳定性和活力方面也有重要的作用。Elliott等[9]通过对重组人促红细胞生成素(rHuEPO)的研究发现，N-糖链对于保持rHuEPO的体内活性和半衰期均有重要的作用。

由此可见，聚糖对于生物大分子的生物活性有着十分重要的作用。然而，在体内聚糖的作用远不只此，在病毒入侵、信号转导、炎症、细胞间相互作用、生殖和发育等过程中，都有各式各样的糖复合物的参与[10]。所以，对糖复合物尤其是所连接的糖链的研究具有十分重要的生物学意义。

2 糖链结构的测定

功能糖组学的主要研究内容之一就是对细胞糖蛋白上糖链结构的解析。解析糖链结构又可分为两步：第一步是高纯度糖链的获得；第二步是进行结构分析。真核生物中蛋白质的糖基化类型可分为3种：N-糖基化、O-糖基化和GPI糖基磷脂酰肌醇锚。大多数糖蛋白只含有一种糖基化类型，但是有些多肽同时连有糖氨聚糖和N-糖链或O-糖链[11]。

2.1 糖链的分离纯化

2.1.1 糖链与非糖物质的解离 糖链与非糖成份的分离是糖研究中的第一步，酶法和化学法是常用

收稿日期：2005-08-24 接受日期：2005-11-25

国家自然科学基金资助项目(No.30540066)

*通讯作者。Tel: 0451-86402052, Fax: 0451-86416944, E-mail: liyugene@hit.edu.cn

的两种裂解方法。酶法裂解的优点是可以使用特异性的内切酶(如PNGase F和内糖苷酶等)在特异部位切下糖链，且对非糖成份没有损伤，但不是所有的糖链都能找到合适的酶进行分离。对于有些酶无法切除的糖链则需要用化学法(如肼解和碱的β-消除反应等)来割裂糖链，化学法可以将糖链和非糖成份完全分离，但是化学法对非糖成份的损伤很大，一般非糖成份无法再用于分析[12]。

2.1.2 糖链的富集 植物凝集素是一类能选择性地结合糖链的糖结合蛋白。利用这种特性可以特异性地富集某种糖链。用于纯化糖链的植物凝集素主要有刀豆凝集素(concanavalin A, Con A)、麦胚凝集素(wheat germ agglutinin, WGA)、豌豆外源凝集素(pea lectin)、植物血凝素(phytohaemagglutinin-E4，E4-PHA)和网孢盘菌凝集素(*Aleuria aurantia* lectin, AAL)等。单一一种凝集素并不能够富集细胞中所有的糖链，因此科学家们在上世纪80年代发明了一种被称作连续凝集素亲和色谱(serial lectin affinity chromatography, SLAC)的方法来分离同一细胞中的各种糖链[13]。这种方法是将不同的凝集素亲和柱串联排列，样品混合物依次经过这些亲和柱，这样就可将不同的糖链富集在不同的柱中，提高了糖链的回收效率和混合糖链的分离效果。

氨基苯硼酸是一种可以和任何含有1,2-顺式-二醇结构的物质相互结合的化合物，将其固定在介质上可以对混合物中的糖链进行分离。Bossi等[14]将多聚氨基苯硼酸应用于毛细管电泳，成功地对含有糖基结构的血红蛋白进行了分离。

高效液相色谱(high performance liquid chromatography, HPLC)和毛细管电泳(capillary electrophoresis, CE)在糖链的分离中应用也是相当的广泛[15~17]。由于它们的自动化程度高、分辨率高，而且可以和质谱相连用，分离的样品直接可以进入质谱仪进行结构的解析，因此成为目前糖链分析领域的主流技术。

2.2 糖链结构的解析

分析糖复合物中的糖链结构是糖组学研究的核心内容之一，如何能快速准确测定糖链的结构成为糖生物学家最为关心的问题。

2.2.1 基于DNA测序仪的荧光糖电泳 荧光分子标记的单糖或多糖分子可以在20%~40%聚丙烯酰胺凝胶中得到分离，同时检测荧光并与标准品进行比较即可得到相应的糖分子信息。但是由于普通的电泳胶的大小限制，无法得到彻底的分析结果。基于以上的优缺点，Callewaert等[18]开发了一种基于DNA测序仪的荧光糖电泳(DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis, DSA-FACE)用于测定N-糖链的结构。基本程序是，先用酶切割糖链，然后用荧光物标记糖链，再用高分辨率的ABI-377A型DNA测序仪进行电泳，最后用Genescan软件对结果进行分析。实验结果表明，用内切酶消化后再进行电泳，可检测到质谱无法达到的飞摩尔水平的聚糖或皮摩尔水平的糖蛋白上的糖链，其中飞摩尔水平的壳四糖检测信号的信噪比超过4。而且在结构分析方面得到的信息与用基质辅助激光解析离子化飞行质谱所得到的结构信息完全一致。Callewaert等[19]将DSA-FACE技术用于肝硬化的无损伤性检测，通过测定血清中的蛋白质的N-糖链，找出了可用于诊断的生物标志物。在区分肝硬化和非硬化性慢性肝病方面，对于代谢失调型肝硬化可达到100%的敏感性和特异性，对于代偿性肝硬化可以达到79%的敏感性和86%的特异性。由此可见，此项技术有相当广阔的应用前景，必将会开创糖链结构分析的新篇章。

2.2.2 用于糖链结构解析的质谱分析法 到目前为止，质谱仍是糖生物学家应用最为广泛的测定糖链结构的工具。快原子轰击质谱(fast atom bombardment mass spectrometry, FAB-MS)、电喷雾质谱(electrospray mass spectrometry, ES-MS)、基质辅助激光解析离子化飞行质谱(matrix-assisted laser desorption ionization time-of-flight mass spectrometry, MALDI-TOF-MS)是目前在糖链结构测定方面应用广泛的3种质谱[20]，它们可以直接对样品进行离子化和解析作用以获得信息，而且可以对完整的糖复合物或者是片段化的样品进行分析。

一种称作气压基质辅助激光解析离子傅立叶转换离子回旋共振质谱(atmospheric pressure matrix-assisted laser desorption /ionization with Fourier transform ion cyclotron resonance mass spectrometry, AP-MALDI-FT-ICR-MS)被用于分析一些含有不稳定基团(如海藻糖和唾液酸)的多糖或糖基复合物。与传统的真空MALDI-MS相比，它的优点在于可以使糖链无论在正离子还是负离子条件下以更加完整的形式被分析，而较少或没有被片段化，因为AP-MALDI-FT-ICR-MS所产生的离子能量要比真空MALDI-MS少的多。在检测灵敏度方面，AP-

MALDI- FT-ICR-MS 可以检测到飞摩的水平，完全可以满足糖链结构分析的需要。Zhang 等[21]应用此种技术检测了一系列的 N- 连接寡糖和 O- 连接寡糖，结果显示AP-MALDI- FT-ICR-MS产生的糖链片段非常少，可以完整地分析其结构，而且可以达到 20 飞摩的检测敏感度。

质谱法测定糖链的另一个优点是可以与蛋白质分离设备，如 HPLC、CE 等进行联用，在 HPLC 和CE上分离的糖蛋白或寡糖可以直接进入质谱进行结构分析，这样就避免了在样品进行转运过程中的污染和损失，提高了分析效率和准确度。大量的糖链结构已由质谱法得到确定。质谱法不仅能够测定糖链的结构，而且可以测定糖链在糖复合物中的定位。Harazono 等[22]先用不同的蛋白酶水解人类载脂蛋白 B100 以得到不同大小的肽段，然后用液相色谱进行肽段分离，将分离后的片段直接进入质谱仪进行测定各个肽段中是否包含有糖链的信息。用这种方法，他们得到了载脂蛋白 B100 的糖基化位点信息。

3　与糖基转移相关的酶的基因

在复杂的细胞个体中，糖基转移酶对于蛋白质的糖基化后修饰过程起到了关键性的作用。因此，在细胞异常转化时这些糖基转移酶也会产生相应的变化，或是发生突变而失活，或是表达量发生变化。García-Vallejo 等[23]研究发现用肿瘤坏死因子 α（TNF-α）作用于人类上皮细胞，所研究的 74 种糖基化相关基因中有 17 种表达量明显上调，而有 2 种基因的表达量下降，上调和下降的基因改变了上皮细胞表面糖链的组成，从而使上皮细胞更容易发生炎症。因此将与糖基转移相关的酶基因点在一张芯片上，用于监测细胞变化前后相关基因表达量的变化对于糖组学研究也就具有非常重要的意义。美国功能糖组学协会的研究人员已将与糖基转移相关的 1 814 条基因制成芯片，可用于监测这些基因在细胞发生变化时转录的变化情况[24]。

4　糖组研究的新方法——糖芯片

随着糖组学研究的深入，基于芯片的原理人们发明了用于糖组研究的各类糖芯片。自 Wang 等[25]首次报道用糖芯片进行研究以来，这种方法得到了广泛的应用。由于其具有信息量大、可进行自动化操作等特点，就如基因芯片对于基因研究和蛋白质芯片对于蛋白质组研究一样，糖芯片在糖组学的研究中同样也将扮演重要的角色。

4.1　研究凝集素 - 糖相互作用的芯片

根据凝集素和寡糖之间的特异性相互作用，将寡糖或糖基复合物固定于芯片上，再用荧光素标记的凝集素进行杂交，检测阳性信号分析与已知凝集素相互作用的寡糖结构。这种方法的最大优点是可以同时检测多个样本，具有批量化、标准化和自动化的特点。

Houseman等[26]首先在有金涂层的玻片表面铺上烯烃硫醇单层并将苯醌基团暴露，通过 Diels-Alder 反应将 10 种单糖固定于一张芯片上。然后用 5 种已用罗丹明标记的凝集素进行检测，结果证实了凝集素和糖的特异性结合作用。

Nimrichter 等[27]采用含有 200 个点的糖芯片研究鸡肝细胞对糖的吸附情况。研究表明表面含有 C 型凝集素的鸡肝细胞可以特异结合到以各种链接方式固定的 N- 乙酰葡萄糖胺残基的非还原性末端，而不能与半乳糖或 N- 乙酰半乳糖胺的非还原性末端结合。他们同时研究了人类 $CD4^+$ T 细胞对糖的结合，结果发现 $CD4^+$ T 细胞对唾液酸化的 Lewis X 抗原 [Neu5Acα3Galβ4(Fucα3)-GlcNAcβ]具有很强的结合，而对没有果糖化的抗原[Neu5Acα3Galβ4GlcNAcβ]只有很少的结合，说明果糖修饰在抗原分子的结合中起到了至关重要的作用。

4.2　杂交糖类 / 糖蛋白芯片

Ratner 等[10]开发出一种杂交糖类 / 糖蛋白芯片，可以快速测定在结合过程中，蛋白质 - 蛋白质相互作用和糖类 - 蛋白质相互作用的主反应一方。通过将糖蛋白和它表面的寡糖都点到点阵上，就可以迅速证实结合的决定簇。将 GAPS II 型玻片用两种化学物质进行表面修饰，在一侧用马来酰亚胺修饰，另一侧用 N- 羟基丁二酰亚胺(N-hydroxysuccinimide, NHS)活化的酯进行修饰。然后分别将糖和糖蛋白点在同一张芯片的马来酰亚胺侧和 NHS 活化的酯一侧，用一种糖结合蛋白杂交来确定蛋白肽是否是结合时所必需的成分。Adams 等[28]应用这种技术证实了在HIV表面的糖蛋白gp120与人类蛋白相互作用过程中，gp120 的糖链起到了至关重要的作用。

4.3　用于研究蛋白质 - 糖相互作用的表面质子共振技术

为了监测凝集素和固定化糖在玻片表面的相互作用，糖生物学家于上世纪 90 年代将表面质子共振

(surface plasmon resonance, SPR)光谱应用于分析。该技术是利用了物理光学的原理，在研究两分子相互作用时，将一种分子固定在传感片表面，而含有另一种分子的溶液流过其表面，两种分子的结合会使传感片表面的折射率改变，因此检测两分子间的相互作用。这种技术不仅可以时时监测两种分子的结合情况，还可以测量结合亲和力的强弱。Smith等[29]将SPR技术用于分析刀豆凝集素Con A与甘露糖和木菠萝凝集素jacalin与半乳糖的相互作用。他们将硫醇修饰的糖固定于金膜的表面，然后测量蛋白质与糖修饰的膜表面结合的吸附系数和蛋白质-糖相互作用的平衡常数，结果表明甘露糖与Con A的结合要比半乳糖和jacalin的结合更为紧密。Ratner等[30]用这种技术研究了一系列高甘露糖结构和蓝藻抗病毒蛋白N(cyanovirin-N, CVN)之间的相互作用。研究结果与用荧光素标记的CVN和固定化聚糖相互作用结果完全一致。

SPR在糖芯片中的应用为糖组学研究提供了一种有力的研究工具。人们可以用它快速测定糖与凝集素或蛋白质之间的特异性作用，还可以对它们之间相互作用的增强剂或抑制剂进行高通量筛选分析。

5 展望

与基因组和蛋白质组学研究相比，糖组学的研究还处于起步阶段。阻碍糖组学迅速发展的主要是研究技术的限制和糖链本身结构的复杂性，目前尚没有一种技术可以快速、大量的测定细胞所有的糖链结构。在期望能有更便捷的测定糖链结构的技术的同时，如何利用现有技术开展糖组学研究则更应是糖生物学家需要考虑的问题。在糖链富集方面，虽然凝集素对聚糖有特异性的结合作用，但是并不是所有的聚糖都有与之相特异结合的凝集素。在结构测定方面质谱仪是最主要的结构测定工具，但是昂贵的价格限制了其更为广泛的应用。DAS-FACE技术的出现开辟了一条测定N-糖链结构的道路，但测定的结果还需要与质谱的结果进行比较确认。对于O-糖链和GPI糖基磷脂酰肌醇锚结构的测定目前仍是糖生物学家所面临的难题。糖芯片技术的出现和SPR技术的应用对于研究聚糖的结构和功能起到了相当大的推动作用，必将在糖组学研究中得到更为广泛的应用。

虽然目前还存在诸多的问题，但是糖组学研究这一崭新时代的到来是不可避免的，而且越来越多的科学家正投身到这一领域中。相信随着研究的深入，人们对生命本质的理解将会随之而达到新的高度。

参考文献 (References)

[1] International Human Genome Sequencing Consortium. *Nature*, 2004, **431**: 931
[2] http://www.genome.gov/
[3] Grünewald S *et al. Pediatr Res*, 2002, **52**: 618
[4] Hirabayashi J *et al. Proteomics*, 2001, **1**: 295
[5] Feizi T. *Glycoconj J*, 2000, **17**: 553
[6] Topaz O *et al. Nat Genet*, 2004, **36**: 579
[7] Sáez-Valero J *et al. Lancet*, 1997, **350**: 929
[8] Sáez-Valero J *et al. Brain Res*, 2001, **889**: 247
[9] Elliott S *et al. Exp Hematol*, 2004, **32**:1146
[10] Ratner DM *et al. Chembiochem*, 2004, **5**: 1375
[11] Varki A *et al. Essentials of Glycobiology*, New York: Cold Spring Harbor Laboratory Press, 1999, 7
[12] Kobata A. *Glycoconj J*, 2000, **17**: 443
[13] Satish PR *et al. J Biochem Biophys Methods*, 2001, **49**: 625
[14] Bossi A *et al. J Chromatogr A*, 2004, **1023**: 297
[15] Reinders J *et al. Proteomics*, 2004, **4**: 3686
[16] Lattova E *et al. J Chromatogr A*, 2003, **1016**: 71
[17] Zamfir A *et al. Electrophoresis*, 2004, **25**:1949
[18] Callewaert N *et al. Glycobiology*, 2001, **11**: 275
[19] Callewaert N *et al. Nat Med*, 2004, **10**: 429
[20] Dell A *et al. Science*, 2001, **291**: 2351
[21] Zhang J *et al. Anal Chem*, 2005, **77**: 4429
[22] Harazono A *et al. Glycobiology*, 2005, **15**: 447
[23] García-Vallejo JJ *et al. J Cell Physiol*, 2006, **206**: 203
[24] http://www.functionalglycomics.org/static/consortium/resources/resourcecoree.shtml
[25] Wang D *et al. Nat Biotechnol*, 2002, **20**:275
[26] Houseman BT *et al. Chem Biol*, 2002, **9**: 443
[27] Nimrichter L *et al. Glycobiology*, 2004, **14**: 197
[28] Adams EW *et al. Chem Biol*, 2004, **11**: 875
[29] Smith EA *et al. J Am Chem Soc*, 2003, **125**: 6140
[30] Ratner DM *et al. Chembiochem*, 2004, **5**: 379

Progress in Glycomics

Shan Wang, Yu Li*

(*Department of Life Science and Engineering, Harbin Institute of Technology, Harbin 150001, China*)

Abstract Glycomics is mainly to analyze the structure and function of the whole set of glycans linked to glycoproteins produced by individual organisms. Glycomics can make people understand more about the role of the third bioinformative macromolecules——glycan in biology. So far, the study of glycomics is on the first stage. Some new technology such as DNA sequencer-assisted, fluorophore-assisted carbohydrate electrophoresis (DSA-FACE), carbohydrate microarrays and the application of mass spectrum technology accelerate progress of the study on glycome.

Key words glycomics; DSA-FACE; mass spectrometry; carbohydrate microarrays; surface plasmon resonance

---

Received: August 24, 2005 Accepted: November 25, 2005

This work was supported by the National Natural Science Foundation of China (No.30540066)

*Corresponding author. Tel: 86-451-86402052, Fax: 86-451-86416944, E-mail: liyugene@hit.edu.cn

细胞生物学杂志