

基于3种非天然糖代谢标记的分泌蛋白质组分析性能对比

毛源¹, 郑江南², 封顺^{1*}, 田瑞军^{2*}

(1. 西南交通大学生命科学与工程学院, 四川成都 610031; 2. 南方科技大学理学院化学系, 广东深圳 518055)

摘要:分泌蛋白质是调控细胞间信号转导的重要生物大分子。由于分泌蛋白的丰度相比于胞内蛋白以及培养基添加剂更低,因此分泌蛋白的高通量鉴定是目前蛋白质组学界研究的热点和难点。目前,基于生物质谱的分泌蛋白质组学分析一般均需要从无血清的条件培养基中获得分泌蛋白质,再对其进行富集和分析。该流程操作步骤繁琐,易造成分泌蛋白质的损失和降解。本工作采用基于生物正交化学生物学技术实现对分泌蛋白质的高选择性标记和高效富集。通过结合点击化学技术,综合评估了分泌蛋白质分析中用于代谢标记的不同糖类类似物。采用3种最常用的商品化糖类类似物,*N*-叠氮乙酰甘露糖胺(ManNAz)、*N*-叠氮乙酰半乳糖胺(GalNAz)和*N*-叠氮乙酰葡萄糖胺(GlcNAz)分别对HeLa细胞进行代谢标记,之后通过炔基生物素探针针对条件培养基中的分泌蛋白进行富集,结合质谱分析来对比3种糖类类似物对分泌蛋白的标记效率。最后通过无标定量蛋白质组学分析,系统评估了3种糖类类似物用于分泌蛋白质组分析的性能。结果表明,基于ManNAz的分泌蛋白标记方法鉴定到了282个分泌蛋白、224个细胞质膜蛋白以及846个*N*-糖基化位点;对分泌蛋白的富集效率分别较GalNAz和GlcNAz提高了130%和67.2%;对细胞质膜蛋白的富集效率较GalNAz和GlcNAz分别提高了273.3%和148.7%,体现出了明显的优势。本研究的实验结果为分泌蛋白高选择性富集和系统分析提供了有益的对比分析和新技术策略。

关键词:液相色谱;串联质谱;分泌蛋白;点击化学;代谢标记;糖类类似物

中图分类号:O658

文献标识码:A

文章编号:1000-8713(2021)10-1086-08

Comparison of the performance of secretome analysis based on metabolic labeling by three unnatural sugars

MAO Yuan¹, ZHENG Jiangnan², FENG Shun^{1*}, TIAN Ruijun^{2*}

(1. School of Life Science and Engineering, Southwest Jiaotong University, Chengdu 610031, China; 2. Department of Chemistry, School of Science, Southern University of Science and Technology, Shenzhen 518055, China)

Abstract: Many secreted proteins, including cytokines, growth factors and hormones, are crucial in processes like intercellular signaling. Dynamic changes in secreted proteins usually reflect the growth and pathological state of the cells. Many drug targets are secretory proteins. The proteins are also important biomarkers. Conditioned cell culture media are important samples for secretory proteomic studies. Biomass spectrometry-based proteomic analysis enables the systematic study of secretory proteins. The main problem in analyzing secretory proteins in conditioned culture media is the low concentration of these proteins and the presence of serum, amino acids, and additives in culture media that may interfere with the protein analysis. Conventional secretory proteome analysis uses serum-free cell culture to reduce sample complexity, and typically involves protein concentration, purification, and desalting using ultrafiltration, dialysis, lyophilization, and trichloroacetic acid (TCA) or acetone precipitation,

收稿日期:2021-04-18

* 通讯联系人.E-mail:fengshunxd@hotmail.com(封顺);E-mail:tianrj@sustech.edu.cn(田瑞军).

基金项目:国家重点基础研究计划基金(2016YFA0501403,2016YFA0501404,2020YFE0202200);国家自然科学基金(31700088);广东省杰出青年基金(2019B151502050);深圳市科委创新基金(JCYJ20170412154126026).

Foundation item: China State Key Basic Research Program Grants (Nos. 2016YFA0501403, 2016YFA0501404, 2020YFE0202200); National Natural Science Foundation of China (No. 31700088); Guangdong Provincial Fund for Distinguished Young Scholars (No. 2019B151502050); Shenzhen Innovation of Science and Technology Commission Grant (No. JCYJ20170412154126026).

followed by enzymatic digestion and mass spectrometry analysis. This analytical process does not allow specific enrichment of secreted proteins. Thus, only a few secreted proteins can be identified. In addition, prolonged serum-free incubation of cells also tends to lead to unexpected changes in their activity status. A bioorthogonal-based enrichment approach can effectively avoid this problem. In recent years, unnatural sugars containing bio-orthologous groups, such as azide groups, have been used to metabolically label glycosylated proteins, enabling cellular imaging or selective enrichment of glycoproteins and their use for proteomic analysis. The strategy is a two-step process. First, azide-based sugar analogues are added to the cell culture medium and introduced to glycoproteins via the intracellular glycan biosynthesis pathway. Second, they are specifically covalently labeled with imaging probes or affinity probes via click chemistry. Since secreted proteins are usually glycoproteins, this glycolytic labeling has been used to label and enrich secreted proteins. *N*-Azidoacetylgalactosamine (GalNAz), *N*-azidoacetylglucosamine (GlcNAz), and *N*-azidoacetylmannosamine (ManNAz) are classical azide-based sugar analogues. Their effects on cytoplasmic membrane proteins have been compared. However, only ManNAz has been used for metabolic labeling of secreted proteins. No other glyco-analogues that label secreted proteins have been reported. Here, the bio-orthogonal chemical biology technology achieved highly selective labeling and enriched secreted proteins. In combination with click chemistry, different sugar analogues were evaluated for metabolic labeling of secreted proteins. HeLa cells were metabolically labeled by ManNAz, GalNAz, and GlcNAz (the three most commonly used commercial sugar analogues). These glycolytic markers can selectively label specific types of glycosylation. For example, ManNAz, an analogue of the biosynthetic precursor of sialic acid, *N*-acetylmannosamine (ManNAc), can label sialylated *N*- or *O*-glycoproteins. GalNAz, an analogue of *N*-acetylgalactosamine (GalNAc), can replace GalNAc as a core residue of mucin-type *O*-glycans and thus label *O*-glycoproteins. In addition, the intracellular metabolic intermediate of GalNAz (pyrophosphate) UDP-GalNAz can be interconverted with UDP-GlcNAz catalyzed by UDP-galactose-4-differential isomerase (GALE) and thus can also label *N*-glycoproteins and *O*-GlcNAc glycoproteins instead of GlcNAc. The GlcNAz analogue is commonly used to label nuclear and cytoplasmic glycoproteins with β -*O*-GlcNAc residues, but can also label *N*-glycoproteins with mucin-type *O*-glycoproteins by converting GALE to GalNAz, followed by enrichment using a biotin-alkynyl probe. Label-free quantitative proteomic analysis was performed to evaluate their labeling efficiency. ManNAz-based secretory protein labeling identified 282 secretory proteins, 224 plasma membrane proteins, and 846 *N*-glycosites. Compared with GalNAz and GlcNAz, the enrichment of secreted proteins was increased 130% and 67.2%, respectively, and the enrichment of plasma membrane proteins was increased 273.3% and 148.7%, respectively. This study provides a useful comparative analysis and new strategies for highly selective enrichment and systematic secretome analysis.

Key words: liquid chromatography (LC); tandem mass spectrometry (MS/MS); secreted protein; click chemistry; metabolic labeling; sugar analogues

分泌蛋白质组 (secretome) 是指细胞、组织等分泌的全部蛋白质, 广义的分泌蛋白组还包括从细胞膜表面脱落 (ectodomain shedding) 的细胞质膜蛋白^[1]。许多分泌蛋白, 如细胞因子、生长因子和激素等, 在细胞间信号转导等过程中发挥着关键的功能。分泌蛋白的动态变化通常反映了细胞的生长情况和病理状态。分泌蛋白构成了很大一部分药物靶标, 同时也是重要的生物标志物^[2]。细胞的培养基是分泌蛋白质组研究的重要样本。目前, 基于物质谱的蛋白质组学分析可以实现对分泌蛋白的系统研究^[3]。分析培养基中的分泌蛋白的主要问题在于条件培养基中的分泌蛋白浓度较低^[4-7], 培养基中的血清、氨基酸和添加剂等的存在可能会干扰后续的蛋白质分析。常规的分泌蛋白质组分析使用无血清细胞培养以降低样品的复杂性, 一般先采用超滤、透析、冻干和三氯乙酸 (TCA) 或丙酮沉淀进行蛋白质的浓缩、纯化和除盐, 然后进行酶解和质谱分析。这一分析流程无法实现特异性的富集分泌蛋白, 从而导致分泌蛋白的鉴定数目较少。此外, 长时间无血清培养细胞也容易导致其活性状态的意外改变。而基于生物正交的富集方式则可以有效避免这个问题。

近年来, 含有生物正交基团 (如叠氮基) 的非天然糖已被用于代谢标记糖基化蛋白质^[8,9], 从而实现糖蛋白的细胞成像或选择性富集并用于蛋白质组学分析。该策略分为两个步骤: 叠氮基糖类物被添加到细胞培养基中, 通过细胞内的聚糖生物合成途径引入到糖蛋白上; 然后通过点击化学反应特异性地与成像探针或亲和探针进行共价标记。由于分泌蛋白通常是糖蛋白, 这种糖代谢标记已被用于分泌蛋白的标记和富集^[10]。*N*-叠氮乙酰半乳糖胺 (GalNAz)、*N*-叠氮乙酰葡萄糖胺 (GlcNAz) 和 *N*-叠氮乙酰甘露糖胺 (ManNAz) 是最经典的叠氮基糖类物^[8,11,12], 有研究对比了它们对细胞质膜蛋白的标记效果, 然而, 目前只有 ManNAz 被用来进行分泌蛋白的代谢标记, 尚未报道其他两种糖类物用于标记分泌蛋白。

本文通过结合糖代谢标记、点击化学和基于质谱的蛋白质组学, 系统地评估了 ManNAz、GalNAz 和 GlcNAz 3 种非天然糖对分泌蛋白的标记情况。通过无标定量 (LFQ) 蛋白质组学分析, 探索最适合的分泌蛋白的糖代谢标记方法。这项工作提供了一种对细胞分泌蛋白进行全局、位点特异性以及定量

研究的手段, 并为分泌蛋白的糖代谢标记方法的选择提供依据。

1 实验部分

1.1 仪器、试剂与材料

Easy-nLC 1000 色谱系统 (Thermo Fisher Scientific); Orbitrap Fusion 质谱仪 (Thermo Fisher Scientific)。

高糖 DMEM (达尔伯克改良伊格尔, dulbecco's modified eagle's medium) 培养基、磷酸盐缓冲液 (PBS) 购自 Corning 公司。胎牛血清 (FBS) 购自 GIBCO 公司。GalNAz、GlcNAz、ManNAz、三(3-羟丙基-三唑甲基)胺 (THPTA) 购自 Click Chemistry Tools 公司。0.22 μm 一次性无菌过滤器、Amicon Ultra 超滤管 (15 mL, 截止相对分子质量 3 kD) 购自 Millipore 公司。盐酸氨基胍、五水硫酸铜、三氟乙酸 (TFA) 购自阿拉丁试剂有限公司。链霉亲和素 (streptavidin) 琼脂糖微球购自 Thermo Scientific。乙二胺四乙酸二钠 (EDTA-Na₂) 购自 Amresco 公司。苯基甲磺酰氟 (PMSF) 购自 Auragene 公司。乙腈 (ACN, 色谱纯) 购自 BCR 国际贸易公司。甲醇 (色谱纯) 购自 Merck 公司。十二烷基硫酸钠 (SDS) 购自 Biosharp 公司。抗坏血酸钠、氯乙酰胺 (CAA)、三羟甲基氨基甲烷 (Tris)、碳酸氢铵、尿素、三(2-羧乙基)膦盐酸盐 (TCEP) 购自 Sigma-Aldrich 公司。胰蛋白酶 (测序级) 购自 Promega 公司。Lys-C (测序级) 购自 Wako 公司。炔基生物素探针由本实验室合成。

1.2 实验步骤

1.2.1 细胞培养与代谢标记

将 HeLa 细胞置于 37 °C、5% CO₂ 的加湿培养箱中, 在含 (10%, v/v) FBS 的高糖 DMEM 培养基中培养。待细胞生长至覆盖培养皿底部面积的 60% 时, 用 PBS 洗涤细胞 3 次, 然后分别向无血清 DMEM 培养基中添加 0.1 mmol/L ManNAz、GalNAz 和 GlcNAz 进行代谢标记, 细胞继续培养 24 h 后完成代谢标记过程。

1.2.2 条件培养基的收集

收集条件培养基, 以 300 g 的转速离心 5 min, 取上清液再以 3 000 g 的转速离心 10 min, 以去除细胞碎片和死细胞。之后收集上清液, 用 0.22 μm 滤膜过滤。在滤液中加入 1 mmol/L EDTA-Na₂、10 mmol/L CAA 和 0.1 mmol/L PMSF, 然后使用 3

kDa 超滤管于 4 °C、4 000 g 条件下将条件培养基浓缩至约 600 μ L。向超滤管中加入 12 mL PBS, 再将其超滤液浓缩至约 600 μ L, 重复此过程 3 次, 使分泌蛋白样品置换到 PBS 中。使用 Pierce 660 蛋白测定试剂测定蛋白质浓度。

1.2.3 利用点击化学探针进行铜催化的点击反应

铜 (I) 催化的叠氮化物-炔烃环加成反应 (CuAAC)^[13,14] 是应用最广泛的点击反应。将超滤后的样品通过 CuAAC 点击反应^[15] 与炔基化试剂进行反应。我们在每个样品中依次加入 100 μ mol/L 炔基生物探针、13.5 mmol/L 盐酸氨基胍和催化剂 (0.675 mmol/L 硫酸铜、3.38 mmol/L THPTA 和新制的 16.9 mmol/L 抗坏血酸钠的预混溶液) (均为加入后的浓度), 在常温下振荡反应 2 h。通过氯仿-甲醇蛋白沉淀法去除多余的探针。之后将蛋白沉淀溶解于尿素缓冲液 (8 mol/L 尿素, 200 mmol/L Tris, pH 7.4) 中。

1.2.4 生物素化蛋白的富集与酶解

将 1.2.3 中得到的蛋白溶液与 15 μ L Streptavidin 琼脂糖微球置于旋转混匀仪上于 4 °C 孵育 2 h。依次使用 0.4 mL 尿素缓冲液 (6 mol/L 尿素、0.1% SDS、10 mmol/L Tris, pH 7.4) 和 0.4 mL 1 mol/L 氯化钠溶液各洗涤 3 次。然后向样品管中加入 300 μ L 的炔基化反应缓冲液 (5 mmol/L TCEP、50 mmol/L CAA、0.2 mol/L 碳酸氢铵、0.5 mol/L 氯化钠), 于 37 °C 振荡孵育 30 min。之后向样品管中分 3 次加入 0.4 mL 的 20% (v/v) 乙醇水溶液洗涤。微球用胰蛋白酶 (酶与底物质量比为 1:200) 在 50 mmol/L 碳酸氢铵和 37 °C 条件下酶解 12 h。将微球悬浊液离心后, 收集上清液。之后用 PNGase F (酶与底物质量比为 1:200) 酶解 12 h; 将微球悬浊液离心后收集上清液。将上清液用三氟乙酸 (TFA) 酸化至 pH<3 后使用 StageTips 进行脱盐。

1.2.5 液相色谱-串联质谱分析

液相色谱柱为 C18 毛细管填充柱 (15 cm \times 100 μ m, 粒径 1.9 μ m, 孔径 12 nm, Dr. Maisch 公司)。流动相 A 相为 0.1% (v/v) 甲酸水溶液; B 相为 0.1% (v/v) 甲酸乙腈溶液。流速为 250 nL/min。色谱有效梯度为 7% B 在 100 min 提高至 22% B; 22% B 在 100~120 min 提高至 35% B。质谱使用数据依赖采集模式, 循环时间为 3 s, 一级质谱离子扫描的 m/z 范围为 350~1 550, 分辨率为 120 000, 自动增益控制 (AGC) 为 2×10^5 , 最大注入时间 (IT) 为

100 ms, 分离窗口 1.6 Da, 离子价态限定为 2~4, 动态排除时间为 30 s。二级质谱检测器类型为离子阱, 碎裂模式为 HCD, 最大 IT 为 35 ms, 归一化碰撞能 (NCE) 为 30。

1.2.6 数据分析

质谱数据使用 MaxQuant^[16] 和 UniProtKB 人类蛋白质组数据库 (2020 年 1 月发布, 74 823 个条目, 附加 188 个高丰度牛血清蛋白序列) 对质谱原始数据进行检索。酶的特异性设置为 Trypsin/P, 最多允许两个酶切遗漏点。固定修饰设置为半胱氨酸烷基化, 可变修饰设置为甲硫氨酸氧化和天冬酰胺脱酰胺化。LFQ 分析使用 “Match between runs” 功能, 蛋白质和多肽鉴定的假阳性率 (FDR) 控制在 1% 以内。

使用 Perseus 软件 (1.6.17.0 版)^[17] 进行统计分析。过滤掉反库序列、仅靠位点鉴定的蛋白和污染蛋白。鉴定的蛋白质至少在一组 (每组 3 次生物学独立重复) 中有 >1 个 unique 肽段并且拥有 3 个有效值的蛋白质数据被保留做定量分析。LFQ 强度进行 \log_2 变换, 缺失值依据正态分布替换 width (替换区间宽度): 0.3, downshift (替换位置下调度): 1.8。通过双侧 t 检验分析两组之间的差异性 (基于置换检验的 (permutation-based) FDR<0.05, $S_0=2$)。使用 DAVID (6.8 版)^[18,19] 进行基因本体 (GO) 富集分析, 分类的 p 值 <0.05。对于 N -糖基化位点分析, 仅保留符合 N-X-S/T/C ($X\neq P$) 序列的肽段, 并筛选出位点定位得分 ≥ 50 、定位概率 ≥ 0.8 的数据进行分析。

2 结果与讨论

2.1 实验设计与可行性讨论

糖代谢标记可以选择性地标记特定的糖基化类型 (见图 1a)^[20]。例如 ManNAz 是唾液酸的生物合成前体 N -乙酰甘露糖胺 (ManNAc) 的类似物, 可以标记唾液酸化的 N -或 O -糖蛋白^[20,21]。GalNAz 是 N -乙酰半乳糖胺 (GalNAc) 的类似物, 可以替代 GalNAc 成为黏蛋白型 O -聚糖的核心残基从而标记 O -糖蛋白^[20,21]。此外, GalNAz 在细胞内的代谢中间体尿苷二磷酸 (UDP)-GalNAz 可经过 UDP-半乳糖-4-差向异构酶 (GALE) 催化而与 UDP-GlcNAz 相互转化^[22], 因此也可替代 GlcNAc 对 N -糖蛋白和 O -GlcNAc 糖蛋白进行标记。 N -乙酰葡萄糖胺 (GlcNAc) 的类似物 GlcNAz 通常用于标记带有

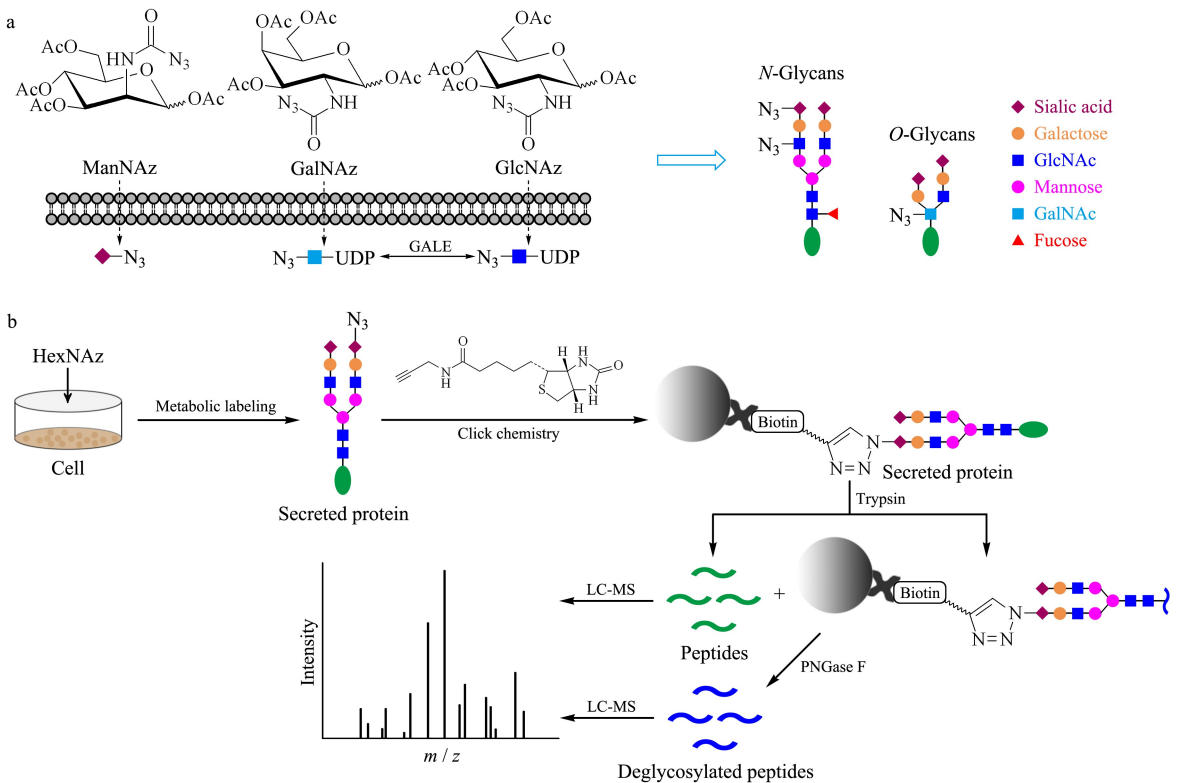


图 1 基于质谱分析的分泌蛋白富集

Fig. 1 Secreted protein enrichment based on MS analysis

a. structures of three sugar analogues and metabolic labeling of glycoproteins containing sialic acid, GlcNAc, GalNAc, mannose, galactose, and fucose; b. the experimental workflow.

UDP: uridine diphosphate; GALE: galactose-4-epimerase; PNGase F: peptide *N*-glycosidase F; GlcNAc: *N*-acetylglucosamine; GalNAc: *N*-acetylgalactosamine; HexNAz: six carbon carbohydrate analogues.

β -*O*-GlcNAc残基的核糖蛋白和细胞质糖蛋白,但是也能通过 GALE 转化为 GalNAz 从而标记 *N*-糖蛋白与黏蛋白型 *O*-糖蛋白^[20,21]。

因此,我们选择使用 ManNAz、GalNAz 和 GlcNAz 3 种最常用的叠氮基糖类类似物对 HeLa 细胞的分泌蛋白进行代谢标记。收集条件培养基中的叠氮标记的分泌蛋白,通过 CuAAC 与炔基生物素反应,再用 Streptavidin 微球富集生物素化的蛋白质。使用胰蛋白酶 (trypsin) 对富集到的蛋白质进行酶解,再将酶解下来的肽段进行 LC-MS 检测。之后通过 *N*-糖酰胺酶 F (PNGase F) 酶切释放微球上保留的 *N*-糖肽,将得到的去糖基化肽段进行质谱分析,以获得 *N*-糖基化位点信息 (见图 1b)。

2.2 不同糖类类似物标记的细胞中分泌蛋白的鉴定分析

将基于 3 种糖类类似物标记的方法在 3 次重复实验中富集到的蛋白质组和特异性肽段 (unique peptide) 筛选出来进行对比分析。GalNAz 组鉴定到了

1 487 个蛋白和 11 520 个特异性肽段;GlcNAz 组鉴定到了 1 462 个蛋白和 11 612 个特异性肽段;ManNAz 组鉴定到了 1 591 个蛋白以及 12 187 个特异性肽段 (见图 2a)。进一步对各平行实验进行定量重复性评估,各平行实验间的皮尔逊相关系数都大于 0.94,表现出了良好的生物学重复性 (见图 2b)。3 组实验鉴定到的共有蛋白超过了 80%。我们将 3 组实验中鉴定到的蛋白与 Uniprot 的蛋白数据库进行对比,结果表明 ManNAz 组检测到的 1 591 个蛋白中有 282 个蛋白被 Uniprot 注释为分泌蛋白,224 个蛋白被注释为细胞质膜蛋白;GalNAz 组鉴定到的 1 487 个蛋白中有 271 个蛋白被注释为分泌蛋白,204 个蛋白被注释为细胞质膜蛋白;GlcNAz 组鉴定到的 1 462 个蛋白中有 276 个蛋白被注释为分泌蛋白,192 个蛋白被注释为细胞质膜蛋白 (见图 2c 与图 2d)。我们也使用传统方法对无血清培养的 HeLa 细胞条件培养基进行了直接分析,3 次重复实验中共鉴定到了 1 473 个蛋白,其中 249 个为分泌

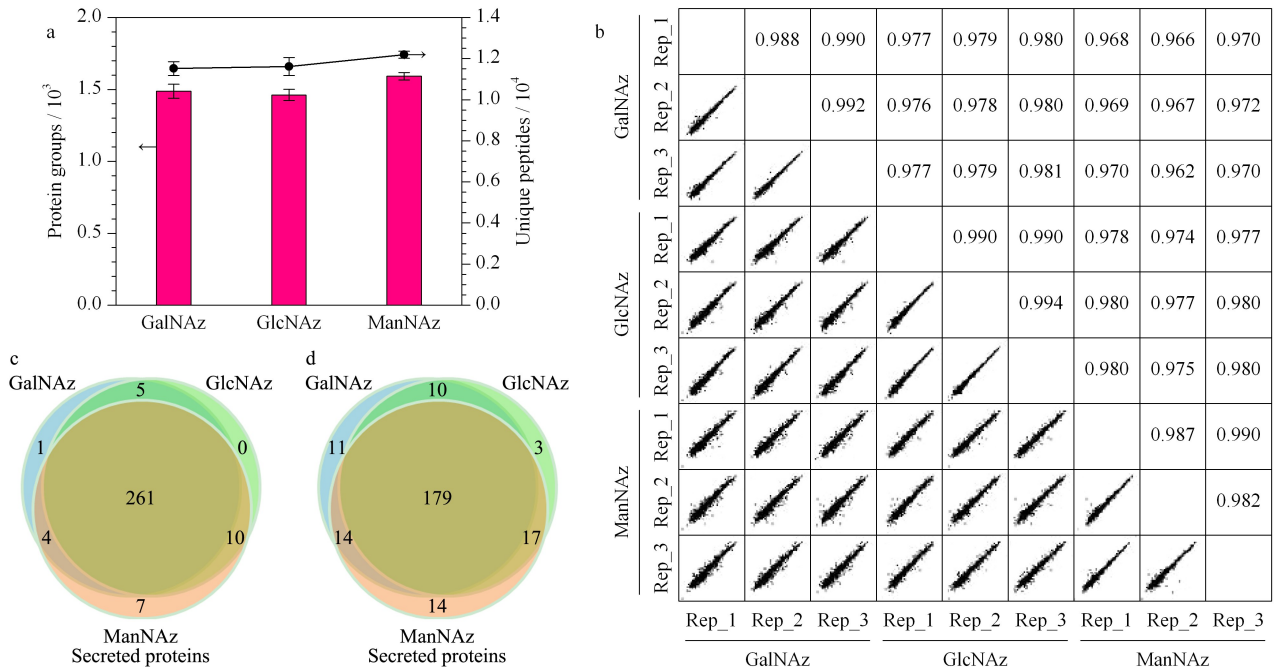


图 2 基于 3 种糖类类似物的分泌蛋白质分析结果

Fig. 2 Results of analysis of secreted proteins based on the three sugar analogues

a. numbers of protein groups and unique peptides ($n=3$); b. correlation analysis; c. numbers of identified secreted protein groups; d. numbers of identified plasma membrane protein groups.

蛋白。将传统方法与 3 种糖代谢标记方法鉴定到的分泌蛋白进行 LFQ intensity 的对比,可以看出,基于直接分析的传统方法,经过 3 次重复实验鉴定到的所有分泌蛋白的总 LFQ 强度(3 次重复实验的平均值)显著低于代谢标记方法(见表 1)。该结果表明,基于代谢标记的方式可以鉴定到更多的分泌蛋白。其中基于 ManNAz 的代谢标记在分泌蛋白质的富集中具有一定的优势,其富集到的分泌蛋白质数与细胞质膜蛋白质数多于其他两组,三者间的总体差异不大。共有的分泌蛋白质数为 261 个,在 3 组中的占比超过了 90%;共有的细胞质膜蛋白数为 179 个,在 3 组中的占比超过了 70%。

2.3 分泌蛋白的定量分析与糖基化位点鉴定分析

对 3 组糖类类似物鉴定到的蛋白质进行进一步的

表 1 直接鉴定与基于代谢标记方式的富集结果

Table 1 Direct enrichment and enrichment results based on metabolic labeling

Enrichment method	Number of proteins	Number of secreted proteins	LFQ intensity
Direct enrichment	1473	249	4.38×10^{10}
GalNAz	1487	271	1.04×10^{11}
GlcNAz	1462	276	1.44×10^{11}
ManNAz	1591	282	2.40×10^{11}

LFQ: label-free quantification.

无标定量分析。ManNAz 组在分泌蛋白的定量富集中鉴定到的 LFQ 强度值相比于 GalNAz 组与 GlcNAz 组分别提高了 130% 与 67.2%;在细胞质膜蛋白的定量富集中鉴定到的 LFQ 强度值相比于其他两组分别提高了 273.3% 与 148.7% (见图 3a)。三者分泌蛋白和细胞质膜蛋白的定量富集中存在着显著的差异。分析其原因可能是因为分泌蛋白在进行糖基化修饰时通常更多地利用唾液酸进行修饰,而 ManNAz 在进入细胞后会转化为 N-乙酰叠氮唾液酸,从而被有效地结合到分泌蛋白上。

图 3b 展示了 3 种糖类类似物分别鉴定到的糖蛋白、糖肽以及糖基化位点的结果。ManNAz 组鉴定到 352 个糖蛋白、754 个糖肽以及 846 个糖基化位点;GalNAz 组鉴定到 290 个糖蛋白、460 个糖肽以及 512 个糖基化位点;GlcNAz 组鉴定到 302 个糖蛋白、511 个糖肽以及 563 个糖基化位点。ManNAz 组的糖蛋白、糖肽与糖基化位点数目相比于 GalNAz 组分别提高了 21.4%、63.9% 以及 65.2%;相比于 GlcNAz 组则分别提高了 16.6%、47.6% 以及 50.3%。N-糖肽与 N-糖基化位点的显著差异结果进一步表明 ManNAz 组定量富集分泌蛋白的优势可能来源于 ManNAz 对糖基化位点的高效标记。

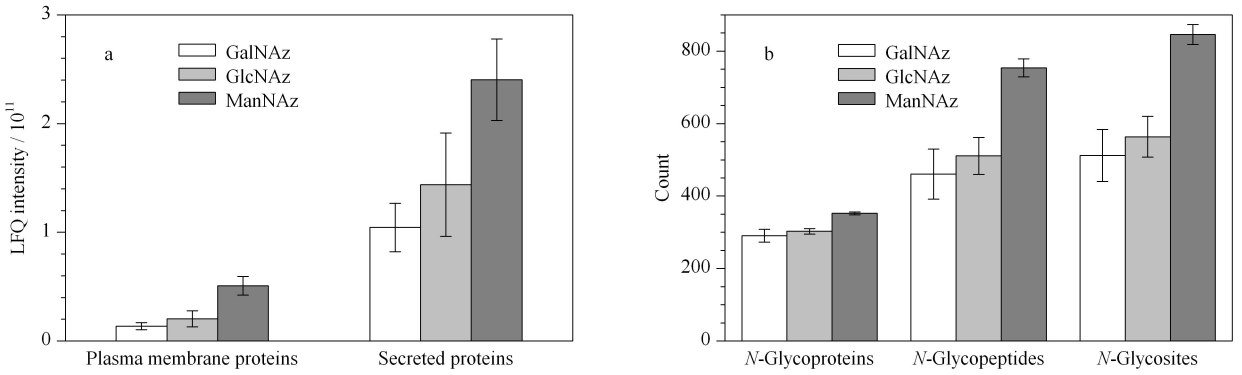


图 3 分泌蛋白的定量富集与糖基化鉴定结果 (n=3)

Fig. 3 Quantitative analysis and glycosylation identification of secreted proteins (n=3)

a. LFQ intensity of secreted proteins and plasma membrane proteins; b. numbers of identified glycoproteins, glycopeptides, and glycosylation sites.

2.4 分泌蛋白的差异化分析

对 GalNAz、GlcNAz 与 ManNAz 3 个实验组中鉴定到的分泌蛋白质进行差异化分析,其结果以火山图的形式表现。相比于 GalNAz 组,ManNAz 组

存在着 1 018 个显著增加的蛋白与 9 个显著减少的蛋白,其中有 141 个被鉴定为显著增加的分泌蛋白(红点表示)(见图 4a);相比于 GlcNAz 组,ManNAz 组存在着 121 个显著增加的蛋白,其中有 18 个被鉴

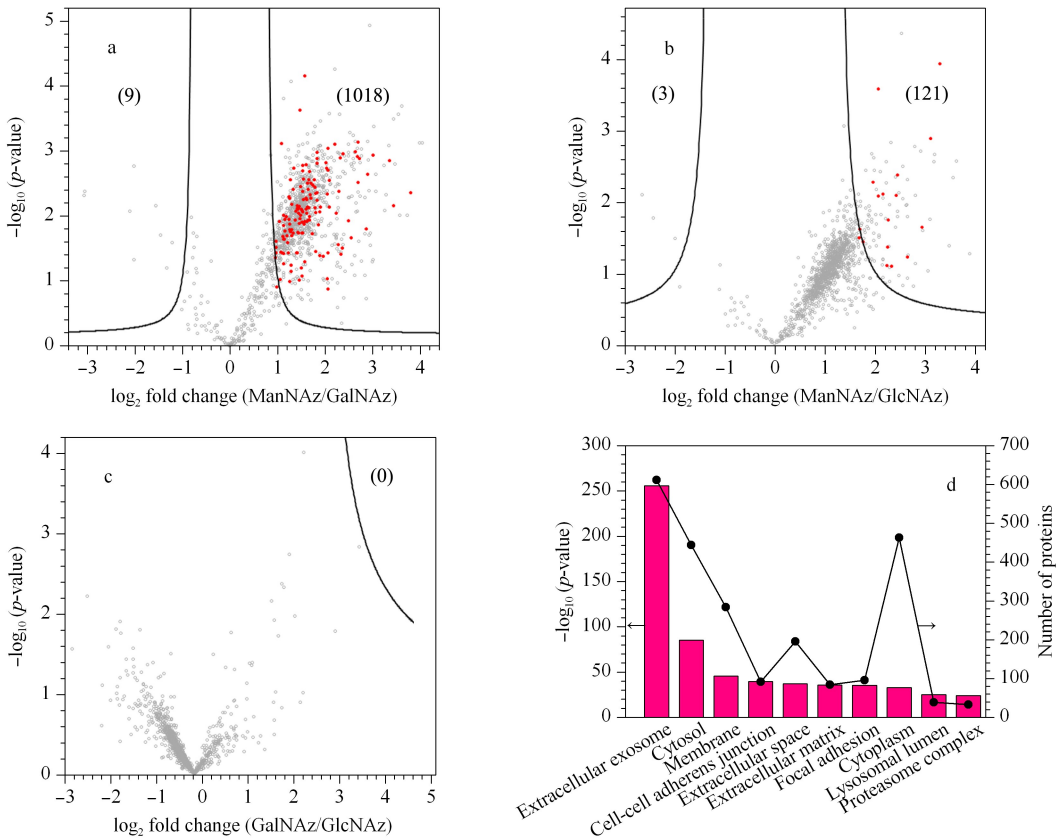


图 4 差异化蛋白分析

Fig. 4 Analysis of differential proteins

a. volcano plot analysis of ManNAz- and GalNAz-labeled proteins. The red dots denote secreted proteins and the numbers in parentheses represent significantly increased proteins. b. volcano plot analysis of ManNAz- and GlcNAz-labeled proteins. The red dots denote secreted proteins. c. volcano plot analysis of GalNAz- and GlcNAz-labeled proteins. d. cellular component analysis of significantly increased proteins in the ManNAz experiment.

定为分泌蛋白(见图 4b)。上述结果表明,ManNAz 相比于 GalNAz 与 GlcNAz 具有更好的标记分泌蛋白的能力。图 4c 展示了 GalNAz 与 GlcNAz 之间的差异蛋白火山图,可以看出两种糖之间并无显著性差异的分泌蛋白,这可能是由于二者可以在细胞内相互转化,能够以相同的形式标记到蛋白质上,从而导致了它们的标记效率没有明显差异^[20,21]。之后对图 4 中 ManNAz 组中显著增加的 1 018 个蛋白做了 GO 分析。结果表明,这些差异化的蛋白主要分布在细胞外区域,与分泌蛋白的分布情况相符(见图 4d)。上述结果表明 ManNAz 组相比于其他两组对分泌蛋白的定量富集具有高选择性。

3 结论

分泌蛋白是许多疾病发生过程中的重要标志物,对于分泌蛋白的系统研究是目前难点与热点。在这项工作中,我们将 3 种糖类似物(GalNAz、ManNAz 和 GlcNAz)进行了系统的比较,结合代谢标记、点击化学以及质谱检测对分泌蛋白进行了富集与分析。实验结果表明 ManNAz 对于分泌蛋白及相关蛋白的定量富集效果显著强于另外两种糖类似物。我们还发现基于 ManNAz 的代谢标记能鉴定到更多的糖基化位点,体现了 ManNAz 在分泌蛋白糖基化修饰过程中的巨大优势。本研究结果证明了 ManNAz 是研究基于糖代谢标记与点击化学相结合的分泌蛋白质组学的优秀工具,为分泌蛋白高选择性富集和系统分析提供了有益的对比分析和新技术策略。

参考文献:

- [1] Rabouille C. Trends Cell Biol, 2017, 27(3): 230
- [2] Stastna M, Van Eyk J E. Proteomics, 2012, 12(4/5): 722
- [3] Brown K J, Formolo C A, Seol H, et al. Exp Rev Proteomic, 2012, 9(3): 337
- [4] Colzani M, Waridel P, Laurent J, et al. J Proteome Res, 2009, 8(10): 4779
- [5] Weng Y, Sui Z, Shan Y, et al. Anal Chem, 2016, 88(9): 4971
- [6] Shin J, Rhim J, Kwon Y, et al. Sci Rep, 2019, 9(1): 3096
- [7] Pirkmajer S, Chibalin A V. Am J Physiol Cell Physiol, 2011, 301(2): C272
- [8] Ali Khan A, Hansson J, Weber P, et al. Mol Cell Proteomics, 2018, 17(12): 2358
- [9] Mahal L K, Yarema K F, Bertozzi C R. Science, 1997, 276: 1125
- [10] Xiao H, Tang G X. Anal Chem, 2016, 88(6): 3324
- [11] Hang H C, Yu C, Kato D L, et al. PNAS, 2003, 100(25): 14846
- [12] Mercier N, Ramakrishnan B, Boeggeman E, et al. Bioconjugate Chem, 2013, 24(1): 144
- [13] Rostovtsev V, Green L, Fokin V, et al. Angew Chem Int Ed Engl, 2002, 41: 2596
- [14] Tornøe C W, Christensen C, Meldal M. J Org Chem, 2002, 67(9): 3057
- [15] Hart C, Chase L G, Hajivandi M, et al. Methods Mol Biol, 2011, 698: 459
- [16] Cox J R, Mann M. Nat Biotechnol, 2008, 26(12): 1367
- [17] Tyanova S, Temu T, Sinitcyn P, et al. Nat Methods, 2016, 13(9): 731
- [18] Huang D W, Sherman B T, Lempicki R A. Nat Methods, 2009, 4(1): 44
- [19] Huang D W, Sherman B T, Lempicki R A. Nucleic Acids Res, 2009, 37(1): 1
- [20] Laughlin S T, Bertozzi C R. Nat Protoc, 2007, 2(11): 2930
- [21] Boyce M, Carrico I S, Ganguli A S, et al. PNAS, 2011, 108(8): 3141
- [22] Nandi A, Sprung R, Barma D K, et al. Anal Chem, 2006, 78(2): 452