



冠状病毒的糖基化研究进展*

惠子夜 于汉杰 舒 健 任夏萌 陈闻天**

(西北大学生命科学学院, 功能糖组学实验室, 西安 710069)

摘要 冠状病毒 (coronavirus) 为一类具有囊膜结构、单股正链 RNA 病毒, 可感染哺乳动物或禽类等。目前已知 7 种冠状病毒可造成人际间传播, 其中以新型冠状病毒 (Severe acute respiratory syndrome coronavirus-2, SARS-CoV-2) 为代表, 可引起严重的呼吸系统疾病, 并感染全球数亿人口, 从而对全球公共卫生安全造成严重危害。作为病毒表面蛋白最常见的糖基化修饰, 其在识别宿主、免疫逃逸、复制组装和裂解传播等病毒活动中起重要作用。本文以 SARS-CoV-2 等代表, 概括 N-与 O-糖基化研究在冠状病毒刺突蛋白 (spike protein)、膜蛋白 (membrane protein) 等的最新进展和方法。通过全局性总结冠状病毒科中糖蛋白的糖基化位点分布、糖链结构、生物学功能与研究技术等, 以推动冠状病毒相关的诊断、治疗与疫苗开发等。

关键词 冠状病毒, 新型冠状病毒, 糖基化修饰, 刺突蛋白, 膜蛋白

中图分类号 Q7

DOI: 10.16476/j.pibb.2021.0231

冠状病毒 (coronaviruses, CoVs) 对人类社会造成严重威胁, 在哺乳动物和禽类中可引起多种疾病^[1]。冠状病毒来源于网巢病毒目 (Nidovirales) 的冠状病毒科 (Coronaviridae), 是一类具有囊膜结构、单股正链的 RNA 病毒, 其形态较大, 呈圆或椭圆形, 直径约为 80~120 nm, 因其在电镜下呈日冕或皇冠状而得名^[2]。冠状病毒基因组为 27~32 kb, 是目前已知基因组最大的 RNA 病毒^[3]。根据血清学特性和分子进化特点, 可将近百种冠状病毒分为 4 个属, 即 α -、 β -、 γ -和 δ -CoV 属^[4]。其中造成小范围人群感染的 HCoV-229E (Human coronavirus-229E)、HCoV-NL63 (Human coronavirus-Netherlands 63)、HCoV-OC43 (Human coronavirus-organculture 43) 和 HCoV-HKU1 (Human coronavirus-Hong Kong University 1) 均属于 α -CoV 属; β -CoV 属成员众多, 可进一步划分为 5 个亚属^[5]。严重急性呼吸综合征冠状病毒, 或称非典病毒 (Severe acute respiratory syndrome coronavirus, SARS-CoV) 和中东呼吸综合征冠状病毒 (Middle East respiratory syndrome coronavirus, MERS-CoV) 及导致全球肆虐新型冠状病毒肺炎 (coronavirus disease 2019, COVID-19) 的新型冠状病毒 (Severe acute respiratory

syndrome coronavirus-2, SARS-CoV-2) 均来源于 β -CoV 属^[2, 6-9]; γ -与 δ -CoV 属包含较多以禽类为主宿主的病毒, 如鸡传染性支气管炎病毒 (Infectious bronchitis virus, IBV)、鹅口疮冠状病毒 (Thrush coronavirus HKU12) 和夜莺冠状病毒 (Bulbul coronavirus HKU11) 等^[10]。冠状病毒基因组编码多种结构蛋白, 包括刺突蛋白 (spike protein, S 蛋白)、膜蛋白 (membrane protein, M 蛋白)、核衣壳蛋白 (nucleocapsid protein, N 蛋白) 和囊膜蛋白 (envelope protein, E 蛋白) 等, 部分 β -CoV 属成员存在血凝素酯酶 (haemagglutinin-esterase, HE) 蛋白。此外还可编码 16 种非结构蛋白 (nonstructural proteins, nsp)。如 SARS-CoV-2 的 nsp1~nsp16, 但来源于 γ -或 δ -CoV 属部分冠状病毒含 15 种 nsp, 即 nsp2~nsp16^[11-12]。

糖基化 (glycosylation) 是最常见的生物大分子修饰方式, 包括 N-糖基化修饰、O-糖基化修饰和糖脂修饰等^[13-14]。糖基化修饰在人体中参与多

* 陕西省自然科学基金 (2021JQ-446, 2021JM-319), 中国博士后科学基金 (2020M673628XB) 和西北大学新型冠状病毒肺炎预防应急指导基金 (NWU002) 资助项目。

** 通讯联系人。

Tel: 029-88304104, E-mail: cwt@nwu.edu.cn

收稿日期: 2021-08-09, 接受日期: 2021-12-24

种生命活动, 例如蛋白质的折叠、运输和定位、细胞分化和发育、细胞识别、信号转导、免疫应答等^[15-20]。在含有囊膜结构的病毒中, 糖基化修饰对于识别宿主细胞受体与入侵宿主、免疫逃逸、适应温度酸碱等理化因素与子代病毒扩散等过程中发挥着重要的作用^[21]。

冠状病毒中存在极为复杂的糖基化修饰特性, 本文希望通过全局性总结冠状病毒糖蛋白的糖基化位点分布、糖链结构与生物学功能及糖生物学研究技术等, 有助于对冠状病毒相关的诊断、治疗和疫苗开发等工作。

1 冠状病毒S蛋白的糖基化修饰

冠状病毒同源三聚体S蛋白是I型膜蛋白, 其氨基酸数目从1 150到1 450个不等, 序列相似性在属间和种间差异较大, 但被描述为倒锥状或丁香状(clove)三维结构单元组成相同^[22-23]。S蛋白由S1和S2亚基组成^[24]。S1亚基包含N端结构域(N-terminal domain, NTD)和受体结合结构域(receptor binding domain, RBD)等, 其功能是结合宿主细胞受体^[25]。在SARS-CoV和SARS-CoV-2病毒中, RBD结合人体内糖蛋白血管紧张素转换酶2(angiotensin converting enzyme 2, ACE2)被认为是入侵宿主的第一步^[11]; 类似的结合机制也存在于鼠肝炎病毒(Mouse hepatitis virus, MHV)与鼠癌胚抗原相关细胞黏附分子1(mouse carcino embryonic antigen-related cell adhesion molecule 1, mCEACAM1a)^[26]和MERS-CoV与T细胞表面抗原CD26(DPP4)^[27]等之间。S2亚基富含 α 螺旋结构, 包含两段七肽重复域、跨膜区和C端胞内域等^[24]。当S蛋白被宿主细胞蛋白酶裂解引发S2亚基暴露, 两段三聚体的七肽重复域构象变化折叠成发夹状内螺旋的六聚体结构, 从而介导病毒和宿主细胞膜的融合及进一步的遗传物质释放^[25]。不同属或属间S1亚基结构差异较大, 例如MHV的NTD具有相对较大的空间结构域^[28]。在高度选择压力作用下, S1亚基是冠状病毒中最易变异区域, 而S2亚基相对保守^[29]。

1.1 S蛋白的N-糖基化

蛋白糖基化修饰现象非常普遍, 大多数以N-糖基化修饰为主。N-糖基化发生在内质网和高尔基体中, 其修饰位点以N-X-S/T(X不能为脯氨酸)为特征, 在所有物种中高度保守^[30]。囊膜病毒表面的糖蛋白布满丰富的N-糖基化修饰位点, 例如

甲型流感病毒囊膜上两种糖蛋白, 血凝素(hemagglutinin, HA)和神经氨酸苷酶(neuraminidase, NA), 通常含有6~12个N糖基化位点^[31]; 在人类免疫缺陷病毒表面的gp160含多达31个N-糖基化位点^[32]; 冠状病毒的N-糖基化修饰更加复杂多样。

通过对S蛋白进行生物信息学分析可见18~39个N-糖基化位点分布在4个属中, 其中HCoV-NL63中多达39个N-糖基化位点, 这意味其三聚体的N-糖基化位点数目可达117个(图1)^[33]; SARS-CoV-2和SARS-CoV的S蛋白相似性为80%, 均含22个N-糖基化位点, 但SARS-CoV-2的RBD区域缺失一个N-糖基化位点(对应SARS-CoV中的N370), 会暴露更多的受体结合区域^[2, 22]。值得注意的是, 在与SARS-CoV-2密切相关的蝙蝠来源(RaTG13)和穿山甲来源(PCoV_GX)的冠状病毒S蛋白序列更加相似(96.2%), 但多两个N-糖基化位点(N30和N370)^[34]。从三维结构角度看, 这些N-糖基化位点分散于S1亚基外表面上, 但S2的C端处6~8个保守N-糖基化位点均位于蛋白质与膜连接处, 可能涉及入侵宿主的膜融合行为^[35]。总之, 冠状病毒4个属之间糖基化位点数目和位置存在较大差异, 这种差异也见于亚属与种间(图1)^[36-39]。

由于含有至少54个N-糖基化位点, 故S蛋白分布着高度致密的N-糖链结构, 这些糖链对蛋白质折叠、宿主结合以及逃避宿主免疫系统均起到了重要的影响作用。但这些位点是否都会发生糖基化修饰? 由于病毒不编码糖链修饰相关酶, 因此其糖链结构来源于宿主系统: 绝大多数物种的新生肽链位于内质网时, 会在N-糖基化位点添加14糖前体, 随后伴随着肽链的折叠与运输, 受到宿主糖基转移酶和糖苷酶对糖链结构的加工与修饰^[40]。而病毒中滞留信号肽(retrieval signal)的存在会引导S蛋白等在内质网、内质网高尔基体中间体(endoplasmic reticulum-golgi intermediate compartment, ERGIC)或高尔基体中积累并形成出芽体, 这也导致了以高甘露糖型糖链为代表的未完全加工N-糖链结构在病毒膜蛋白中较为丰富^[41]。而Nal等^[42]利用C端标记蛋白分析SARS病毒S蛋白的细胞定位, 通过糖苷酶酶切与电泳分析等发现, S蛋白的高甘露糖糖链在进入高尔基体后成为复杂型的N-糖链。

因为病毒不表达参与糖链合成与修饰相关蛋白

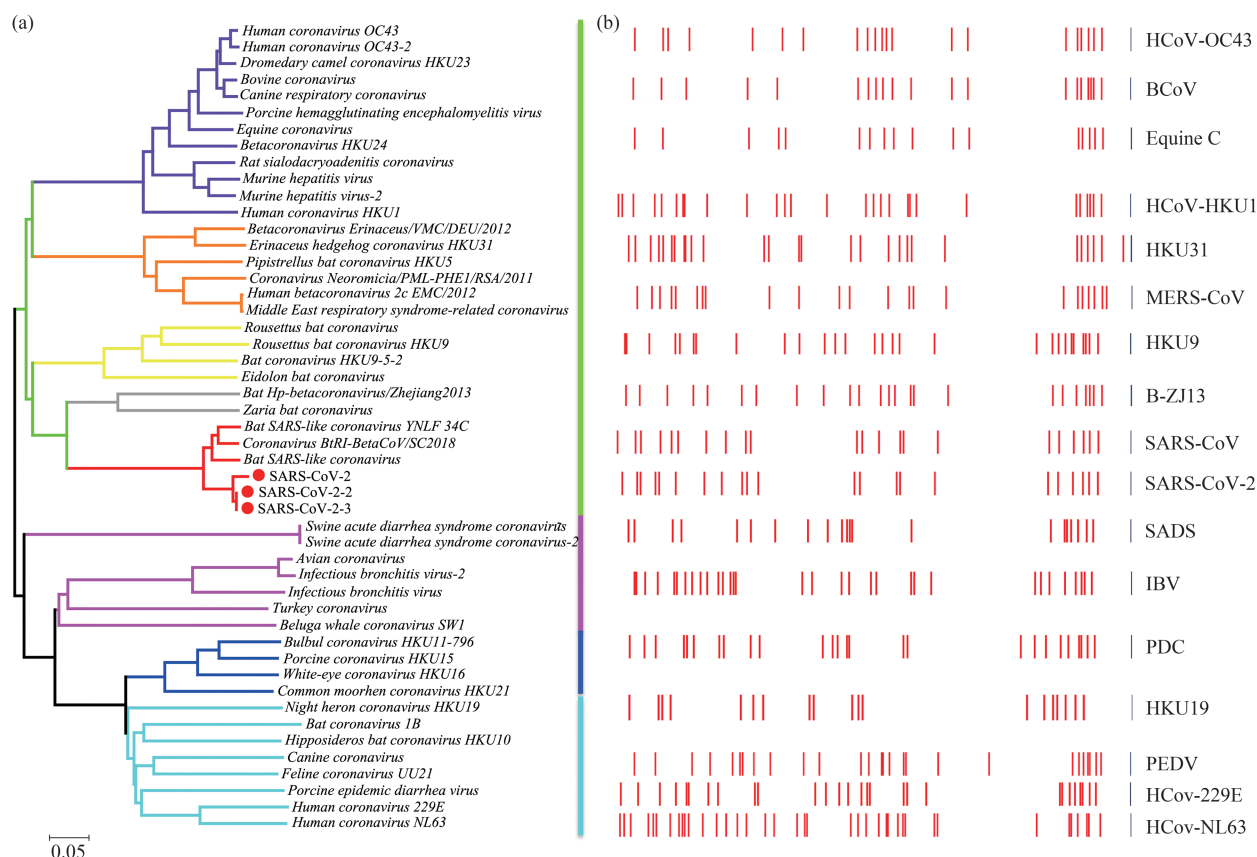


Fig. 1 The N-J phylogenetic tree and the distribution of N-glycosylation sites in spike proteins of CoVs

图1 冠状病毒S蛋白的N-J进化树与N-糖基化位点分布

(a) 冠状病毒可以划分为4个属： α （青色）、 β （绿色）、 γ （紫色）和 δ （蓝色）。其中 β 属冠状病毒还可以进一步划分为5个亚属：*Embecovirus*（靛蓝）、*Merbecovirus*（橙色）、*Nobecovirus*（黄色）、*Hibecovirus*（深灰）和*Sarbecovirus*（红色）。(b) 保守的N-糖基化位点在各冠状病毒的S蛋白上数目众多且分布方式多样，但C端通常含有保守的6~8个位点。

酶，所以冠状病毒培养或蛋白质表达体系不同会造成较大的糖链谱差异。Wang等^[43]通过特征离子触发电子转移/高能碰撞解离质谱（signature ions-triggered electron-transfer/higher-energy collisional dissociation (EThcD) mass spectrometry）发现，相比人源 HEK 细胞系表达 SARS-CoV-2、SARS-CoV、MERS-CoV 的 S 蛋白富含唾液酸等杂合型糖链结构，昆虫 SF9 细胞系表达的 SARS-CoV-2、SARS-CoV 的 S 蛋白含有大量寡甘露糖（paucimannosidic）或截断的糖链（truncated glycans）。Shajahan等^[44]通过高分辨率质谱（LC-MS/MS），检测表达于 HEK293 细胞系的 SARS-CoV-2 中 S 蛋白并鉴定出 17 个位点的 N-糖链结构，而其余 5 个位点未检测出（N17、N603、N1134、N1158 和 N1173）。这些 N-糖链包括高甘露糖型、杂合型和复杂型共 3 种类型，其中高甘露糖型和复杂型糖链占比较高；末端唾液酸基团见于大多数糖

链结构，尤其是位于 NTD 附近的 N234 和 N282 的糖链上，推测这些位点的糖链调节对受体蛋白或抗体的结合^[45-46]；相比 RBD 的 N331 和 N343 中以高甘露糖型和少量复杂型糖链，Antonopoulos 等^[47]对同样来源的这两个位点则鉴定到富含岩藻糖修饰和少见的乳糖二胺（LacdiNAc，即 GalNAc β (1-4) GlcNAc）结构的复杂型 N-糖链，这可能与仅表达 RBD 片段等因素有关；但 Tian 等^[48]通过 LC-MS 对 HEK293T 来源的 S 蛋白 N-糖链鉴定出 19 个位点的 N-糖链位点（除 N74、N1158 和 N1173），其中 N17 和 N1134 糖链末端存在唾液酸结构，而 N149、N165、N657、N709、N1134 与 N1194 含岩藻糖结构；Watanabe 等^[49]研究表达于 293F 细胞系 SARS-CoV-2 的 S 蛋白，他们发现 22 个 N-糖基化位点均发生 N-糖基化修饰。这些位点上 16% 的糖链含至少一个唾液酸残基，48% 的糖链发生了岩藻糖基化。

由于天然病毒的 N-糖链结构与重组表达系统

来源具有较大的差异^[50], 此外S蛋白的每个N-糖基化位点呈现出微不均一性的特点, 而非确定性的糖链结构^[51]。因此宿主环境与表达系统影响病毒糖链结构, 值得疫苗开发工作的注意。在世卫组织列出180余种应对COVID-19疫情的在研疫苗中, 有3种基于植物细胞表达^[52], 然而在植物中特有的N-糖链五糖核心木糖修饰可能会对人群产生免疫反应^[53]。在结构免疫原病毒蛋白疫苗、减毒或灭活病毒疫苗、重组载体疫苗和核酸疫苗等应对COVID-19的4类疫苗中^[54], 前两种疫苗中糖链来源于多种培养体系, 因此N-糖链不同形成的空间位阻差异会影响抗原位点的遮蔽效果^[55], 也会影响对细胞摄取、蛋白水解过程、MHC提呈以及随后的T细胞启动等^[56-57]。

1.2 S蛋白的O-糖基化

O-糖基化是另一种常见于病毒蛋白上的糖基化修饰, 其核心结构多样。O-糖基化修饰多发于丝氨酸或苏氨酸上, 但无固定基序。通过生物信息学方法可进行O-糖基化位点预测, 但准确率相比N-糖基化位点较低^[58]。例如, 有学者在COVID-19大流行初期针对SARS-CoV-2的S蛋白进行O-糖基化位点预测, 认为673S、678T和686S极具有修饰潜力, 而19T、22T、29T、250T和349S的潜力较高^[59]。但是Hajahan等^[44]仅鉴定出S1亚基的T323和S325发生了O-糖基化修饰, 这些糖链结构以Core-1和Core-2类型的O-糖链为主。这两个位点位于RBD, 因此该位置的O-糖基化可能在与ACE2受体结合中发挥作用。随后越来越多工作指出S蛋白存在丰富的O-糖基化修饰, 例如Dong等^[60]通过LC-MS/MS鉴定出26个O-糖基化位点和33种O-糖链结构。目前O-糖基化修饰对冠状病毒中生物学功能尚不明确, 有研究指出, 抑制O-糖链结构修饰对结合ACE2影响较小, 但可显著降低病毒入侵细胞能力^[61]。

O-糖基化修饰的发生也与宿主系统有关, Bagdonaite等^[62]通过建立的O-糖链分析技术比较了来源于昆虫细胞系(S2)与人源细胞系(HEK293T)表达SARS-CoV-2的S蛋白, 发现共25个O-糖基化位点。这些位点分布除在RBD上略有差异外基本相同, 例如仅人源细胞系来源的T478发生了O-糖基化修饰。有意思的是, 其中16个发生于已知N-糖基化位点中。他们解释这些O-糖基化发生率较低, 但可能会形成黏蛋白样结构域以保护抗原表位或关键位点。Tian等^[48]同样验证

HEK293T来源SARS-CoV-2的S蛋白中17个O-糖基化位点有11个发生在N-糖基化位点1至3个氨基酸附近(“N±1-3”), 其中的4个位点发生了唾液酸修饰(T236、S659、T1076和T1077)。他们通过定点突变验证N糖基化位点是这些O-糖基化位点出现的先决条件, 并将此现象称之为“O-Follow-N”^[48]。这些结论对拓展病毒的O-糖基化研究有重要意义。

2 冠状病毒M蛋白的糖基化修饰

M蛋白是形成冠状病毒蛋白外壳的重要成分, 其长度为200~250个氨基酸, 如SARS-CoV-2的M蛋白有222个氨基酸^[63]。M蛋白氨基酸相似性在属间极低, 属内相对较高。尽管在生物大分子蛋白质数据库(protein data bank, PDB)中缺少M蛋白相关完整的晶体结构, 但生物学信息学分析指出其包含N端胞外域、3次跨膜区及较长的C端胞内域^[64-65]。M蛋白招募其他结构蛋白进入ERGIC, 并进一步控制子代颗粒组装和出芽^[66]。M蛋白也可能参与结合宿主细胞, 例如HCoV-NL63的M蛋白协同S蛋白参与结合含硫酸类肝素的蛋白多糖实现宿主入侵^[67]。

2.1 M蛋白的N-糖基化

目前发现0~4个潜在N-糖基化位点存在于各种冠状病毒的M蛋白中。通常大多数M蛋白包含1个N-糖基化位点(图2); 但 δ 属的HKU15-CoV未发现相应位点^[68]; HKU4-2-CoV含有两个, N4和N20^[69]; γ 属中部分IBV也含有两个, N3和N6^[70]; 3个位点多见于 α 属, 例如PEDV(*Porcine epidemic diarrhea virus*)的N3、N19和N189^[71], HCoV-NL63的N3、N19和N188^[72], HCoV-229E的N5、N190和N207^[73]; 部分HKU2-CoV还含有4个N-糖基化位点, 如N6、N22、N192和N235^[74]。

SARS-CoV-2的M蛋白N5是唯一的N-糖基化位点^[59], SARS-CoV中类似位置糖基化位点是N4, 多项研究结果显示该位点出现了复杂型N-糖链类型。Voss等^[75]通过对该位点进行缺失突变, 发现M蛋白的N-糖基化对于病毒的组装和感染是必不可少的, 若该位点缺失则影响子代颗粒的组装; 然而, Liang等^[76]通过定点突变与反向遗传学技术发现, IBV中N3和N6糖基化位点缺失突变并不影响病毒复制与子代颗粒组装, 但会对病毒毒力产生一定影响。

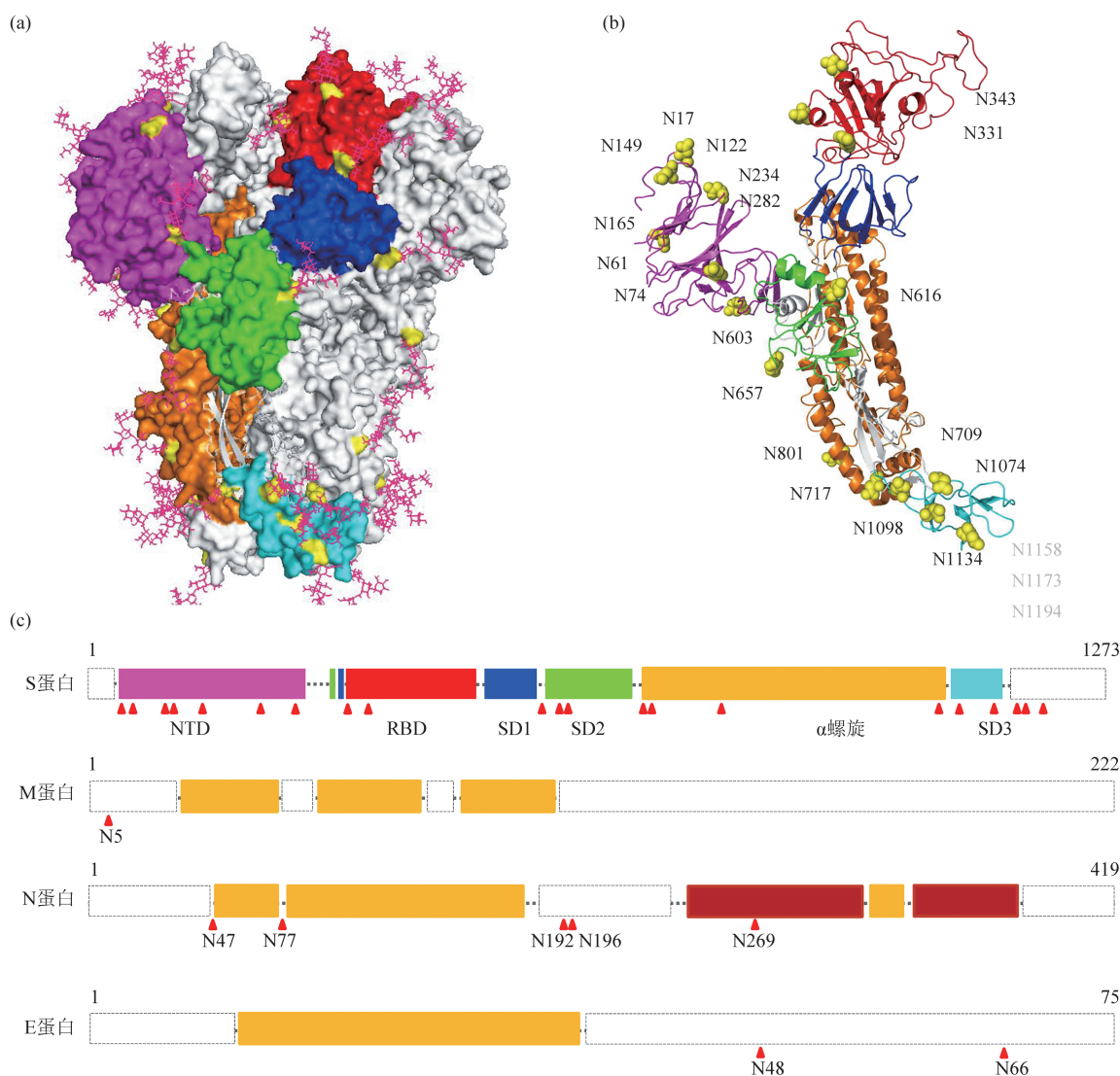


Fig. 2 The glycosylation sites in S-protein of SARS-CoV-2

图2 SARS-CoV-2结构蛋白中N-糖基化位点分布

(a) S蛋白是呈倒锥状的高度糖基化三聚体蛋白；(b) S蛋白亚基可分为NTD（紫色）、RBD（红色）、SD1（subdomain 1，次级结构1，蓝色）、SD2（绿色）、α螺旋束（橙色）、SD3（青色）等主要结构单元，N-糖基化位点用黄色球形标注；(c) SARS-CoV-2的结构蛋白潜在N-糖基化位点示意图。橙色与褐色区域分别代表二级结构中富含α螺旋和β片层结构的区域，虚线框为结构未知区域，N-糖基化位点用红色三角表示。

2.2 M蛋白的O-糖基化

除N-糖基化修饰，MHV为代表的部分冠状病毒也发现数目可变的O-糖基化修饰^[77]。De Haan等^[78-79]通过定点突变技术并观察表达产物的电泳条带，分析发现T5可能是M蛋白中发生O-糖基化修饰的位点，该糖链中包含N-乙酰半乳糖胺、半乳糖和唾液酸等；而如果MHV中O-糖基化位点被N-糖基化取代，会造成病毒颗粒重组的效率下降。

3 冠状病毒其他蛋白的糖基化修饰

冠状病毒N蛋白的结构保守且表达量丰富，通

过与RNA基因组形成螺旋状结构在病毒的感染、复制和包装等过程中发挥作用。早期对冠状病毒N蛋白研究表明，N蛋白是一种磷酸化蛋白^[80-82]。近来在SARS-CoV-2的N蛋白上发现5个潜在的N-糖基化位点（N47、N77、N192、N196和N269），但利用O¹⁸同位素标记实验证实只有N47和N269出现糖基化，而这两个位点分别以复杂型和高甘露糖型N-糖链占优势^[83-84]；也有学者利用LC-MS/MS分析并证明N蛋白上发生O-糖基化修饰的T148、T165、T166、T205、S206、T391和S404^[85]。作为参与组装和释放病毒颗粒的多功能E蛋白^[36]，

由于糖蛋白纯化结晶用于X射线晶体衍射的技术限制, PDB收录的生物大分子晶体结构中不含完整糖链结构。利用Glycosciences网站Glyprot在线程序能够选择将不同类型N-糖链添加到N-糖基化位点, 从而揭示糖链对相关功能区的影响^[95]; GlyCAM网站的Glycoprotein Builder在线程序也具有相同的功能, 并且进一步实现了O-糖链添加, 因此最早发布了糖基化修饰SARS-CoV-2的S蛋白三维结构并用于与其他分子相互作用研究^[96]。

4.2 质谱分析法

质谱分析方法具有高灵敏度和分辨率, 成为分析冠状病毒糖蛋白糖链结构的常规操作工具^[99-100]。针对病毒糖基化分析, 核心的问题就是糖基化位点鉴定与糖链结构解析。基于基质辅助激光解吸电离(matrix assisted laser desorption ionization, MALDI)和电喷雾电离(electrospray ionization, ESI)等质谱分析是现代糖组学中最常用工具之一。

4.2.1 质谱分析N-糖基化

糖苷酶PNGase F被用于释放糖蛋白/糖肽上的N-糖链, 并使肽段产生0.98 u分子质量的增加; 另外Endo D/F/H三种内切糖苷酶均可将N-五糖核心两个GlcNAc之间的糖苷键断开, 使肽段上增加了203 u或349 u的质量标签, 因此可作为质谱鉴定糖基化位点的标志^[101]。使用胍解法以释放N-糖链也是一种化学法去糖链策略^[102]。

针对冠状病毒S蛋白、M蛋白或HE蛋白N-糖链鉴定大多采用类似的操作流程: 蛋白质分离纯化、酶切成糖肽片段、糖苷酶切除N-糖链、糖链富集与纯化和MALDI-MS等质谱鉴定^[103]。这些操作策略早在多种冠状病毒N-糖基化中展开, 例如IBV^[104]、MERS-CoV^[35]、HCoV-HKU1^[105]和SARS-CoV-2等^[106-108]。

4.2.2 质谱分析O-糖基化

由于O-糖链的结构简单多变、缺乏保守序列模式且没有合适的酶切方法, 所以大多数病毒缺失O-糖基化报道。O-糖链的切除通常采用以高碘酸氧化-β消除法等为主的化学法^[109]。针对糖蛋白O-糖基化分析依然采取蛋白质分离、酶解、糖肽富集和质谱鉴定等操作思路。基于部分凝集素对O-糖链的亲和力进行分离是最新出现的O-糖蛋白分离/糖肽富集方法, 其中VVA、PNA和Jacalin等都是可选凝集素^[110-111]。对于O-糖链的富集与处理一直在改进之中, 例如Zhang等^[112-113]报道了基于PVDF膜在96孔板平台的细胞样本糖链处理方法,

释放的糖链无须染色或衍生便可使用多孔石墨化碳纳米液相色谱(porous graphitized carbon nano-liquid chromatography)串联质谱法分析糖链结构。他们针对CHO和HEK293细胞系表达SARS-CoV-2的RBD进行O-糖链分析, 同样发现T323和S325中以O-糖基化多以Core-1和Core-2型为主。还有学者将O-糖链末端通过ST3GAL1糖基转移酶标记带荧光的唾液酸, 随后采用胰蛋白酶水解等方法研究糖链指纹图谱。他们发现SF2和HEK293细胞系表达RBD分别以Core-1和Core-2糖链结构为主^[114]。

4.3 生物芯片法分析

以凝集素和糖芯片为代表的生物芯片被广泛应用于病毒糖生物学研究。利用凝集素(或称糖结合蛋白)对糖链结构具有不同特异性和亲和力结合差异的特性, 可联合数种凝集素推断SARS-CoV-2的S蛋白糖链基本特征^[115], 也可将几十种凝集素点制在一张片基, 以实现高通量快速地分析样本糖链谱(图3)^[116]。由于凝集素芯片可以分析糖苷键的连接方式, 具备质谱技术暂时无法达到优点, 因此经常用于联合质谱技术分析糖链结构^[117-118]。

糖链分子作为病毒膜蛋白的结合靶标已在多种病毒中获得报道, 而特异性糖链分布往往与病毒感染宿主的范围密切相关, 因此糖芯片技术通过将各种糖链点制在片基上用于快速筛选病毒蛋白潜在的糖链结合特异性。例如, 针对流感病毒HA的糖芯片分析, 通过比较对α-2,6唾液酸结构与α-2,3唾液酸结构的结合差异评估其宿主特异性^[119], 类似的工作也在BCoV的研究中进行^[120]。针对MERS-CoV的S蛋白进行糖芯片分析, 发现其具有较强唾液酸结合能力, 尤其是对α-2,3唾液酸糖链^[121]。Hao等^[122]则制备了包含有多种在糖胺聚糖中常见的糖链结构, 证实SARS-CoV-2的S蛋白也存在着对硫酸肝素的高度亲和能力。Heidepriem等^[123]则通过糖芯片分析愈后血清中SARS-CoV-2产生抗体的糖链结合能力, 反向推断出SARS-CoV-2中存在大量高甘露糖的N-糖链。此外糖组学研究的门户网站CFG(Consortium for Functional Glycomics)也提供了商品化的糖芯片并在γ属的火鸡冠状病毒、珍珠鸡冠状病毒及鹌鹑冠状病毒得到应用^[124]。

4.4 定点敲除糖基化位点分析

定点突变在病毒蛋白糖基化功能研究方面充当着重要角色, 将核苷酸序列进行突变, 从而造成糖基化位点的缺失或添加^[125]。Han等^[126]针对

SARS-CoV-2 的 S 蛋白 N-糖基化位点进行突变缺失, 发现其中 7 个位点 (N109、N118、N119、N158、N227、N589 和 N699) 对于与 ACE2 相互作用有重要影响, 其中 N227 和 N699 增强了病毒的传播能力。Bouwman 等^[127-128] 通过对 IBV 的 S 蛋白 RBD 区域附近 N-糖基化位点进行突变、反向遗传学操作和 ELISA 法测定, 最后发现 5 个 N-糖基化位点缺失会造成对 α -2,3 唾液酸结合能力的丧失, 部分位点变化还会导致结合其他糖链的能力。假病毒包装技术也应用于膜蛋白的定点突变研究, Li 等^[129] 利用假病毒技术构建了一系列 SARS-CoV-2 病毒 S 蛋白糖基化位点缺失突变体, 用以研究一组中和抗体和恢复期病人血清。他们发现 N331 和 N343 糖基化位点缺失显著降低了感染性, N234 缺失对显著提高病毒对中和抗体的拮抗能力而 N165 缺失增强了对抗体的敏感性。总之定点突变技术在病毒糖基化的应用较为成熟, 在疫苗和抗体的开发中已得到广泛应用。

5 总结与展望

当下肆虐的新冠病毒引起了全球对冠状病毒的关注, 针对病毒相关研究和探讨也在不断展开和深入。糖基化修饰在病毒入侵与离开宿主细胞、引发免疫反应以及针对病毒的检测与治疗等几乎所有环节均发挥不同作用。本文通过归纳冠状病毒中 S 蛋白和 M 蛋白等糖蛋白中糖基化修饰位点的分布、糖链结构和相关功能等, 促进了对糖基化修饰在病毒的认知; 对诸如糖生物学、生物芯片、质谱解析等糖生物学相关技术进行的总结提供了从糖生物学角度研究冠状病毒的方法。这些内容将提高糖基化在冠状病毒的病原学、流行病学和传染机制等领域的重视程度, 推动冠状病毒相关的感染、传播、诊断、治疗和预防等研究工作, 为尽早结束新冠大流行提供支持。

参 考 文 献

- [1] Ye Z W, Yuan S, Yuen K S, *et al.* Zoonotic origins of human coronaviruses. *Int J Biol Sci*, 2020, **16**(10): 1686-1697
- [2] Lu R, Zhao X, Li J, *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*, 2020, **395**(10224): 565-574
- [3] Ren L L, Wang Y M, Wu Z Q, *et al.* Identification of a novel coronavirus causing severe pneumonia in human: a descriptive study. *Chin Med J (Engl)*, 2020, **133**(9): 1015-1024
- [4] Hause B M, Collin E A, Liu R, *et al.* Characterization of a novel influenza virus in cattle and Swine: proposal for a new genus in the Orthomyxoviridae family. *mBio*, 2014, **5**(2): e00031-00014
- [5] Corman V M, Muth D, Niemeyer D, *et al.* Hosts and sources of endemic human coronaviruses. *Adv Virus Res*, 2018, **16**(100): 163-188
- [6] Ksiazek T G, Erdman D, Goldsmith C S, *et al.* A novel coronavirus associated with severe acute respiratory syndrome. *N Engl J Med*, 2003, **348**(20): 1953-1966
- [7] Peiris J S, Lai S T, Poon L L, *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet*, 2003, **361**(9366): 1319-1325
- [8] Zhou P, Yang X L, Wang X G, *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 2020, **579**(7798): 270-273
- [9] 刘文杰, 刘凯会, 张彦伟, 等. 糖基化在新型冠状病毒感染中的机制及药物研发中的应用. *化学进展*, 2021, **33**(4): 524-532
- [10] Liu W J, Liu K H, Zhang Y W, *et al.* *Prog Chem*. 2021, **33**(4): 524-532
- [10] Paim F C, Bowman A S, Miller L, *et al.* Epidemiology of Deltacoronaviruses (δ -CoV) and Gammaponaviruses (γ -CoV) in wild birds in the United States. *Viruses*, 2019, **11**(10): 897
- [11] Wrapp D, Wang N, Corbett K S, *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 2020, **367**(6483): 1260-1263
- [12] Kint J, Langereis M A, Maier H J, *et al.* Infectious bronchitis coronavirus limits interferon production by inducing a host shutoff that requires accessory protein 5b. *J Virol*, 2016, **90**(16): 7519-7528
- [13] Shrimall S, Cherepanova N A, Gilmore R. Cotranslational and posttranslational N-glycosylation of proteins in the endoplasmic reticulum. *Semin Cell Dev Biol*, 2015(41): 71-78
- [14] Van Den Steen P, Rudd P M, Dwek R A, *et al.* Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol*, 1998, **33**(3): 151-208
- [15] Stowell S R, Ju T, Cummings R D. Protein glycosylation in cancer. *Annu Rev Pathol*, 2015, **2015**(10): 473-510
- [16] Chen X L, Liu C, Tang B, *et al.* Quantitative proteomics analysis reveals important roles of N-glycosylation on ER quality control system for development and pathogenesis in *Magnaporthe oryzae*. *PLoS Pathog*, 2020, **16**(2): e1008355
- [17] Mammadova-Bach E, Jaeken J, Gudermann T, *et al.* Platelets and defective N-glycosylation. *Int J Mol Sci*, 2020, **21**(16): 5630
- [18] Zhang X. Alterations of golgi structural proteins and glycosylation defects in cancer. *Front Cell Dev Biol*, 2021, **12**(9): 665289
- [19] Hang Q, Isaji T, Hou S, *et al.* Integrin $\alpha 5$ Suppresses the phosphorylation of epidermal growth factor receptor and its cellular signaling of cell proliferation via N-glycosylation. *J Biol Chem*, 2015, **290**(49): 29345-29360
- [20] Martin T C, Ilieva K M, Visconti A, *et al.* Dysregulated antibody, natural killer cell and immune mediator profiles in autoimmune thyroid diseases. *Cells*, 2020, **9**(3): 665
- [21] 向田, 章晓联. 病毒与宿主细胞的糖基化修饰及相关功能. *生*

- 物化学与生物物理进展, 2017, **44**(10): 898-907
- Xiang T, Zhang X L. Prog Biochem Biophys. 2017, **44**(10): 898-907
- [22] Walls A C, Park Y J, Tortorici M A, *et al.* Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. Cell, 2020, **181**(2): 281-292
- [23] Guruprasad L. Human coronavirus spike protein-host receptor recognition. Prog Biophys Mol Biol, 2021, (161): 39-53
- [24] Li F. Structure, function, and evolution of coronavirus spike proteins. Annu Rev Virol, 2016, **3**(1): 237-261
- [25] Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell, 2020, **181**(2): 271-280
- [26] Miura T A, Travanty E A, Oko L, *et al.* The spike glycoprotein of murine coronavirus MHV-JHM mediates receptor-independent infection and spread in the central nervous systems of Ceacam1a-/- mice. J Virol, 2008, **82**(2): 755-763
- [27] Park Y J, Walls A C, Wang Z, *et al.* Structures of MERS-CoV spike glycoprotein in complex with sialoside attachment receptors. Nat Struct Mol Biol, 2019, **26**(12): 1151-1157
- [28] Shang J, Wan Y, Liu C, *et al.* Structure of mouse coronavirus spike protein complexed with receptor reveals mechanism for viral entry. PLoS Pathog, 2020, **16**(3): e1008392
- [29] Lacey J M, Minutti C Z, Magera M J, *et al.* Improved specificity of newborn screening for congenital adrenal hyperplasia by second-tier steroid profiling using tandem mass spectrometry. Clin Chem, 2004, **50**(3): 621-625
- [30] Hyakumura M, Walsh R, Thaysen-Andersen M, *et al.* Modification of asparagine-linked glycan density for the design of hepatitis B virus virus-like particles with enhanced immunogenicity. J Virol, 2015, **89**(22): 11312-11322
- [31] Chen W, Sun S, Li Z. Two glycosylation sites in H5N1 influenza virus hemagglutinin that affect binding preference by computer-based analysis. PLoS One, 2012, **7**(6): e38794
- [32] Bonomelli C, Doores K J, Dunlop D C, *et al.* The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. PLoS One, 2011, **6**(8): e23521
- [33] Kiyuka P K, Agoti C N, Munywoki P K, *et al.* Human coronavirus NL63 molecular epidemiology and evolutionary patterns in rural coastal kenya. J Infect Dis, 2018, **217**(11): 1728-1739
- [34] Zhang S, Qiao S, Yu J, *et al.* Bat and pangolin coronavirus spike glycoprotein structures provide insights into SARS-CoV-2 evolution. Nat Commun, 2021, **12**(1): 1607
- [35] Watanabe Y, Berendsen Z T, Raghvani J, *et al.* Vulnerabilities in coronavirus glycan shields despite extensive glycosylation. Nat Commun, 2020, **11**(1): 2688
- [36] Masters P S. The molecular biology of coronaviruses. Adv Virus Res, 2006, **66**: 193-292
- [37] Artika I M, Dewantari A K, Wiyatno A. Molecular biology of coronaviruses: current knowledge. Heliyon, 2020, **6**(8): e04743
- [38] Dawood A A. Glycosylation, ligand binding sites and antigenic variations between membrane glycoprotein of COVID-19 and related coronaviruses. Vacunas, 2021, **22**(1): 1-9
- [39] Hartenian E, Nandakumar D, Lari A, *et al.* The molecular virology of coronaviruses. J Biol Chem, 2020, **295**(37): 12910-12934
- [40] Moremen K W, Molinari M. N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. Curr Opin Struct Biol, 2006, **16**(5): 592-599
- [41] McBride C E, Li J, Machamer C E. The cytoplasmic tail of the severe acute respiratory syndrome coronavirus spike protein contains a novel endoplasmic reticulum retrieval signal that binds COPI and promotes interaction with membrane protein. J Virol, 2007, **81**(5): 2418-2428
- [42] Nal B, Chan C, Kien F, *et al.* Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. J Gen Virol, 2005, **86**(Pt 5): 1423-1434
- [43] Wang D, Baudys J, Bundy J L, *et al.* Comprehensive analysis of the glycan complement of SARS-CoV-2 spike proteins using signature ions-triggered electron-transfer/higher-energy collisional dissociation (EThcD) mass spectrometry. Anal Chem, 2020, **92**(21): 14730-14739
- [44] Shajahan A, Supekar N T, Gleimich A S, *et al.* Deducing the N- and O-glycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. Glycobiology, 2020, **30**(12): 981-988
- [45] Casalino L, Gaieb Z, Goldsmith J A, *et al.* Beyond Shielding: the roles of glycans in the SARS-CoV-2 spike protein. ACS Cent Sci, 2020, **6**(10): 1722-1734
- [46] Verkhrivker G M, Di Paola L. Integrated biophysical modeling of the SARS-CoV-2 spike protein binding and allosteric interactions with antibodies. J Phys Chem B, 2021, **125**(18): 4596-4619
- [47] Antonopoulos A, Broome S, Sharov V, *et al.* Site-specific characterization of SARS-CoV-2 spike glycoprotein receptor-binding domain. Glycobiology, 2021, **31**(3): 181-187
- [48] Tian W, Li D, Zhang N, *et al.* O-glycosylation pattern of the SARS-CoV-2 spike protein reveals an "O-Follow-N" rule. Cell Res, 2021, **31**(10): 1123-1125
- [49] Watanabe Y, Allen J D, Wrapp D, *et al.* Site-specific glycan analysis of the SARS-CoV-2 spike. Science, 2020, **369**(6501): 330-333
- [50] Yao H, Song Y, Chen Y, *et al.* Molecular architecture of the SARS-CoV-2 virus. Cell, 2020, **183**(3): 730-738
- [51] Gao C, Zeng J, Jia N, *et al.* SARS-CoV-2 spike protein interacts with multiple innate immune receptors. bioRxiv, 2020, 227462. doi: 10.1101/2020.07.29.227462
- [52] WHO. Draft landscape and tracker of COVID-19 candidate vaccines [M/OL]. Geneva: World Health Organization Press, 2021[2021-07-13]. <https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines>
- [53] Marusic C, Pioli C, Stelter S, *et al.* N-glycan engineering of a plant-produced anti-CD20-hIL-2 immunocytokine significantly enhances its effector functions. Biotechnol Bioeng, 2018, **115**(3): 565-576
- [54] Pandey S C, Pande V, Sati D, *et al.* Vaccination strategies to combat

- novel corona virus SARS-CoV-2. *Life Sci*, 2020, **256**:117956
- [55] Xu C, Ng D T. Glycosylation-directed quality control of protein folding. *Nat Rev Mol Cell Biol*, 2015, **16**(12): 742-752
- [56] Wolfert M A, Boons G J. Adaptive immune activation: glycosylation does matter. *Nat Chem Biol*, 2013, **9**(12): 776-784
- [57] Roth G A, Picece V, Ou B S, *et al.* Designing spatial and temporal control of vaccine responses. *Nat Rev Mater*, 2021:1-22
- [58] Bennett E P, Mandel U, Clausen H, *et al.* Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. *Glycobiology*, 2012, **22**(6): 736-756
- [59] Lo Presti A, Rezza G, Stefanelli P. Selective pressure on SARS-CoV-2 protein coding genes and glycosylation site prediction. *Heliyon*, 2020, **6**(9): e05001
- [60] Dong X, Chen C, Yan J, *et al.* Comprehensive O-glycosylation analysis of the SARS-CoV-2 spike protein with biomimetic Trp-Arg materials. *Anal Chem*, 2021, **93**(30): 10444-10452
- [61] Yang Q, Hughes T A, Kelkar A, *et al.* Inhibition of SARS-CoV-2 viral entry upon blocking N- and O-glycan elaboration. *Elife*, 2020, **26**(9): e61552
- [62] Bagdonaite I, Thompson A J, Wang X, *et al.* Site-specific O-glycosylation analysis of SARS-CoV-2 spike protein produced in insect and human cells. *Viruses*, 2021, **13**(4): 511
- [63] Wu F, Zhao S, Yu B, *et al.* A new coronavirus associated with human respiratory disease in China. *Nature*, 2020, **579**(7798): 265-269
- [64] Voss D, Pfefferle S, Drosten C, *et al.* Studies on membrane topology, N-glycosylation and functionality of SARS-CoV membrane protein. *Virol J*, 2009, **18**(6): 79
- [65] Thomas S. The structure of the membrane protein of SARS-CoV-2 resembles the sugar transporter SemiSWEET. *Pathog Immun*, 2020, **5**(1): 342-363
- [66] Armstrong J, Niemann H, Smeekens S, *et al.* Sequence and topology of a model intracellular membrane protein, E1 glycoprotein, from a coronavirus. *Nature*, 1984, **308**(5961): 751-752
- [67] Naskalska A, Dabrowska A, Szczepanski A, *et al.* Membrane protein of human coronavirus NL63 is responsible for interaction with the adhesion receptor. *J Virol*, 2019, **93**(19): e00355-00319
- [68] Wang L, Byrum B, Zhang Y. Detection and genetic characterization of deltacoronavirus in pigs, Ohio, USA, 2014. *Emerg Infect Dis*, 2014, **20**(7): 1227-1230
- [69] Woo P C, Wang M, Lau S K, *et al.* Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. *J Virol*, 2007, **81**(4): 1574-1585
- [70] Zhao Y, Liu X Y, Cheng J L, *et al.* Molecular characterization of an infectious bronchitis virus strain isolated from northern China in 2012. *Arch Virol*, 2014, **159**(12): 3457-3461
- [71] Park S J, Kim H K, Song D S, *et al.* Molecular characterization and phylogenetic analysis of porcine epidemic diarrhea virus (PEDV) field isolates in Korea. *Arch Virol*, 2011, **156**(4): 577-585
- [72] Dominguez S R, Sims G E, Wentworth D E, *et al.* Genomic analysis of 16 Colorado human NL63 coronaviruses identifies a new genotype, high sequence diversity in the N-terminal domain of the spike gene and evidence of recombination. *J Gen Virol*, 2012, **93**(Pt 11): 2387-2398
- [73] Farsani S M, Dijkman R, Jebbink M F, *et al.* The first complete genome sequences of clinical isolates of human coronavirus 229E. *Virus Genes*, 2012, **45**(3): 433-439
- [74] Lau S K, Woo P C, Li K S, *et al.* Complete genome sequence of bat coronavirus HKU2 from Chinese horseshoe bats revealed a much smaller spike gene with a different evolutionary lineage from the rest of the genome. *Virology*, 2007, **367**(2): 428-439
- [75] Voss D, Kern A, Traggiai E, *et al.* Characterization of severe acute respiratory syndrome coronavirus membrane protein. *FEBS Lett*, 2006, **580**(3): 968-973
- [76] Liang J Q, Fang S, Yuan Q, *et al.* N-Linked glycosylation of the membrane protein ectodomain regulates infectious bronchitis virus-induced ER stress response, apoptosis and pathogenesis. *Virology*, 2019, (531): 48-56
- [77] De Haan C A, De Wit M, Kuo L, *et al.* O-glycosylation of the mouse hepatitis coronavirus membrane protein. *Virus Res*, 2002, **82**(1-2): 77-81
- [78] De Haan C A, De Wit M, Kuo L, *et al.* The glycosylation status of the murine hepatitis coronavirus M protein affects the interferogenic capacity of the virus *in vitro* and its ability to replicate in the liver but not the brain. *Virology*, 2003, **312**(2): 395-406
- [79] Prentice E, Jerome W G, Yoshimori T, *et al.* Coronavirus replication complex formation utilizes components of cellular autophagy. *J Biol Chem*, 2004, **279**(11): 10136-10141
- [80] Fung T S, Liu D X. Post-translational modifications of coronavirus proteins: roles and function. *Future Virol*, 2018, **13**(6): 405-430
- [81] Delmas B, Kut E, Gelfi J, *et al.* Overexpression of TGEV cell receptor impairs the production of virus particles. *Adv Exp Med Biol*, 1995, **380**(380): 379-385
- [82] Parker M M, Masters P S. Sequence comparison of the N genes of five strains of the coronavirus mouse hepatitis virus suggests a three domain structure for the nucleocapsid protein. *Virology*, 1990, **179**(1): 463-468
- [83] Gouveia D, Miotello G, Gallais F, *et al.* Proteotyping SARS-CoV-2 virus from nasopharyngeal swabs: a proof-of-concept focused on a 3 min mass spectrometry window. *J Proteome Res*, 2020, **19**(11): 4407-4416
- [84] Sun Z, Zheng X, Ji F, *et al.* Mass spectrometry analysis of SARS-CoV-2 nucleocapsid protein reveals camouflaging glycans and unique post-translational modifications. *Infectious Microbes & Diseases*, 2021, **3**(3): 149-157
- [85] Ihling C, Tänzler D, Hagemann S, *et al.* Mass spectrometric identification of SARS-CoV-2 proteins from gargle solution samples of COVID-19 patients. *J Proteome Res*, 2020, **19**(11): 4389-4392
- [86] Huang X, Dong W, Milewska A, *et al.* Human coronavirus HKU1 spike protein uses O-acetylated sialic acid as an attachment

- receptor determinant and employs hemagglutinin-esterase protein as a receptor-destroying enzyme. *J Virology*, 2015, **89**(14): 7202-7213
- [87] Zeng Q, Langereis M A, Van Vliet A L, *et al.* Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc Natl Acad Sci USA*, 2008, **105**(26): 9065-9069
- [88] Bakkers M J, Zeng Q, Feitsma L J, *et al.* Coronavirus receptor switch explained from the stereochemistry of protein-carbohydrate interactions and a single mutation. *Proc Natl Acad Sci USA*, 2016, **113**(22): E3111-E3119
- [89] Gagneten S, Gout O, Dubois-Dalcq M, *et al.* Interaction of mouse hepatitis virus (MHV) spike glycoprotein with receptor glycoprotein MHVR is required for infection with an MHV strain that expresses the hemagglutinin-esterase glycoprotein. *J Virol*, 1995, **69**(2): 889-895
- [90] Soltani S. The hemagglutinin-esterase gene in human coronaviruses SARS-CoV-2, HKU1 and OC43. *Eur Rev Med Pharmacol Sci*, 2020, **24**(12): 6484-6485
- [91] 杨钊, 姜伟业, 陈闻天. 糖生物信息学数据库. *中国生物化学与分子生物学报*, 2015, **31**(3): 257-263
- Yang Z, Jiang W Y, Chen W T. *Chin J Biochem Mol Biol*, 2015, **31**(3): 257-263
- [92] Steentoft C, Vakhrushev S Y, Joshi H J, *et al.* Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J*, 2013, **32**(10): 1478-1488
- [93] Caragea C, Sinapov J, Silvescu A, *et al.* Glycosylation site prediction using ensembles of support vector machine classifiers. *BMC Bioinformatics*, 2007, **9**(8): 438
- [94] Rabbani N, Thornalley P J. Protein glycation - biomarkers of metabolic dysfunction and early-stage decline in health in the era of precision medicine. *Redox Biol*, 2021, (42): 101920
- [95] Bohne-Lang A, Von Der Lieth C W. GlyProt: in silico glycosylation of proteins. *Nucleic Acids Res*, 2005, **33**(Web Server issue): W214-W219
- [96] Grant O C, Montgomery D, Ito K, *et al.* Analysis of the SARS-CoV-2 spike protein glycan shield reveals implications for immune recognition. *Sci Rep*, 2020, **10**(1): 14991
- [97] Chen W, Zhong Y, Su R, *et al.* N-glycan profiles in H9N2 avian influenza viruses from chicken eggs and human embryonic lung fibroblast cells. *J Virol Methods*, 2017, **249**: 10-20
- [98] 李学甜, 孙宇, 杜亚蓉, 等. H7N2禽流感病毒HA的分离纯化及其糖链谱研究. *生物化学与生物物理进展*, 2015, **42**(6): 551-562
- Li X T, Sun Y, Du Y R, *et al.* *Prog Biochem Biophys*. 2015, **42**(6): 551-562
- [99] Xie H, Doneanu C, Chen W, *et al.* Characterization of a recombinant influenza vaccine candidate using complementary LC-MS methods. *Curr Pharm Biotechnol*, 2011, **12**(10): 1568-1579
- [100] Dubayle J, Vialle S, Schneider D, *et al.* Site-specific characterization of envelope protein N-glycosylation on Sanofi Pasteur's tetravalent CYD dengue vaccine. *Vaccine*, 2015, **33**(11): 1360-1368
- [101] O'Neill R A. Enzymatic release of oligosaccharides from glycoproteins for chromatographic and electrophoretic analysis. *J Chromatogr A*, 1996, **720**(1-2): 201-215
- [102] Patel T, Bruce J, Merry A, *et al.* Use of hydrazine to release in intact and unreduced form both N- and O-linked oligosaccharides from glycoproteins. *Biochemistry*, 1993, **32**(2): 679-693
- [103] Shirakawa A, Manabe Y, Fukase K. Recent advances in the chemical biology of N-glycans. *Molecules*, 2021, **26**(4): 1040
- [104] Zheng J, Yamada Y, Fung T S, *et al.* Identification of N-linked glycosylation sites in the spike protein and their functional impact on the replication and infectivity of coronavirus infectious bronchitis virus in cell culture. *Virology*, 2018, **1**(513): 65-74
- [105] Hurdiss D L, Drulyte I, Lang Y, *et al.* Cryo-EM structure of coronavirus-HKU1 haemagglutinin esterase reveals architectural changes arising from prolonged circulation in humans. *Nat Commun*, 2020, **11**(1): 4646
- [106] Zhang Y, Zhao W, Mao Y, *et al.* Site-specific N-glycosylation characterization of recombinant SARS-CoV-2 spike proteins. *Mol Cell Proteomics*, 2020, **19**(20): 100058
- [107] Barre A, Damme E, Simplicien M, *et al.* Man-specific, GalNAc/Tn-specific and Neu5Ac-specific seaweed lectins as glycan probes for the SARS-CoV-2 (COVID-19) coronavirus. *Mar Drugs*, 2020, **18**(11): 543
- [108] Shin Y J, König-Belhammer J, Vavra U, *et al.* N-glycosylation of the SARS-CoV-2 receptor binding domain is important for functional expression in plants. *Front Plant Sci*, 2021, **15**(12): 689104
- [109] Matsuno Y K, Yamada K, Tanabe A, *et al.* Development of an apparatus for rapid release of oligosaccharides at the glycosaminoglycan-protein linkage region in chondroitin sulfate-type proteoglycans. *Anal Biochem*, 2007, **362**(2): 245-257
- [110] Steentoft C, Vakhrushev S Y, Vester-Christensen M B, *et al.* Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. *Nat Methods*, 2011, **8**(11): 977-982
- [111] Yang Z, Cao J, He Y, *et al.* Macro-/micro-environment-sensitive chemosensing and biological imaging. *Chem Soc Rev*, 2014, **43**(13): 4563-4601
- [112] Zhang T, Madunić K, Holst S, *et al.* Development of a 96-well plate sample preparation method for integrated N- and O-glycomics using porous graphitized carbon liquid chromatography-mass spectrometry. *Mol Omics*, 2020, **16**(4): 355-363
- [113] Gstöttner C, Zhang T, Resemann A, *et al.* Structural and functional characterization of SARS-CoV-2 RBD domains produced in mammalian cells. *Anal Chem*, 2021, **93**(17): 6839-6847
- [114] Wu Z L, Ertelt J M. Fluorescent glycan fingerprinting of SARS2 spike proteins. *Sci Rep*, 2021, **11**(1): 20428
- [115] Hoffmann D, Mereiter S, Jin Oh Y, *et al.* Identification of lectin receptors for conserved SARS-CoV-2 glycosylation sites. *EMBO*

- J, 2021, **40**(19): e108375
- [116] Du H, Yu H, Ma T, *et al.* Analysis of glycosphingolipid glycans by lectin microarrays. *Anal Chem*, 2019, **91**(16): 10663-10671
- [117] Shu J, Ma J, Ren X, *et al.* The abnormal glycopatterns of salivary glycoproteins in esophageal squamous cell carcinoma patients. *Front Chem*, 2021, **4**(9): 637730
- [118] Yu H, Shu J, Li Z. Lectin microarrays for glycoproteomics: an overview of their use and potential. *Expert Rev Proteomics*, 2020, **17**(1): 27-39
- [119] Stevens J, Blixt O, Paulson J C, *et al.* Glycan microarray technologies: tools to survey host specificity of influenza viruses. *Nat Rev Microbiol*, 2006, **4**(11): 857-864
- [120] Peng G, Xu L, Lin Y L, *et al.* Crystal structure of bovine coronavirus spike protein lectin domain. *J Biol Chem*, 2012, **287**(50): 41931-41938
- [121] Li W, Hulswit R J G, Widjaja I, *et al.* Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *Proc Natl Acad Sci USA*, 2017, **114**(40): E8508-E8517
- [122] Hao W, Ma B, Li Z, *et al.* Binding of the SARS-CoV-2 spike protein to glycans. *Sci Bull (Beijing)*, 2021, **66**(12): 1205-1214
- [123] Heidepriem J, Dahlke C, Kobbe R, *et al.* Longitudinal development of antibody responses in COVID-19 patients of different severity with ELISA, peptide, and glycan arrays: an immunological case series. *Pathogens*, 2021, **10**(4): 438
- [124] Ambepitiya Wickramasinghe I N, De Vries R P, Weerts E A, *et al.* Novel receptor specificity of avian gammacoronaviruses that cause enteritis. *J Virol*, 2015, **89**(17): 8783-8792
- [125] Tsuchiya E, Sugawara K, Hongo S, *et al.* Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule. *J Gen Virol*, 2002, **83**(Pt 5): 1137-1146
- [126] Han D P, Lohani M, Cho M W. Specific asparagine-linked glycosylation sites are critical for DC-SIGN- and L-SIGN-mediated severe acute respiratory syndrome coronavirus entry. *J Virol*, 2007, **81**(21): 12029-12039
- [127] Bouwman K M, Habraeken N, Laconi A, *et al.* N-glycosylation of infectious bronchitis virus M41 spike determines receptor specificity. *J Gen Virol*, 2020, **101**(6): 599-608
- [128] Van Beurden S J, Berends A J, Krämer-Kühl A, *et al.* A reverse genetics system for avian coronavirus infectious bronchitis virus based on targeted RNA recombination. *Virol J*, 2017, **14**(1): 109
- [129] Li Q, Wu J, Nie J, *et al.* The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell*, 2020, **182**(5): 1284-1294

New Progress of Glycosylation in Coronaviruses*

HUI Zi-Ye, YU Han-Jie, SHU Jian, REN Xia-Meng, CHEN Wen-Tian**

(Laboratory for Functional Glycomics, College of Life Sciences, Northwest University, Xi'an 710069, China)

Abstract The coronaviruses (CoVs), which are a family of positive-strand RNA viruses, infect the mammals and birds. Seven CoVs are responsible for human-to-human transmission, especially the SARS-CoV-2, thereby posing a daunting challenge to global public health security. As the most common modification in viral glycoproteins, glycosylation plays the crucial role in host recognition, immunity avoidance, virus replication, assembly and transmission. In this review, we summarized and discussed the latest studies about glycosylation in coronaviridae members. Focused on the spike protein, nearly one hundred of N/O-glycosylation sites have been reported. The N-glycans from spike protein are dominated by the high-mannose and complex-type, while the O-glycosylation is rather complicated. Significantly, it is known that the viral glycosylation depend on host cells, thus the glycan pattern of the produced recombinant viral glycoproteins might be different from that of native viral proteins, which represent a crucial determinant for vaccine design. The latest results based on bioinformatics, biochip, mass spectrography and genetic technology facilitate the overall perspective for glycosylation researching in CoVs. By summarizing the distribution of glycosylation sites, the structure of glycans, the biological functions and the research technologies, this review will help promote diagnosis, treatment and vaccine development related to coronaviruses.

Key words coronavirus, SARS-CoV-2, glycosylation, spike protein, membrane protein

DOI: 10.16476/j.pibb.2021.0231

* This work was supported by grants from The Natural Science Foundation of Shaanxi Province (2021JQ-446, 2021JM-319), The China Postdoctoral Science Foundation (2020M673628XB), and The Emergency Guidance Fund for Prevention of Novel Coronavirus Pneumonia from Northwest University (NWU002).

** Corresponding author.

Tel: 86-29-88304104, E-mail: cwt@nwu.edu.cn

Received: August 9, 2021 Accepted: December 24, 2021