



ELSEVIER

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng



Research

Coronavirus Disease 2019—Article

新型冠状病毒 HCoV-19 S 蛋白与人 ACE2 蛋白表面糖链和独特翻译后修饰的质谱分析

孙泽宇^a, 任科焱^a, 张兴^b, 陈景华^b, 姜正一^a, 江静^a, 季飞洋^a, 欧阳晓希^a, 李兰娟^{a,*}

^a State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Disease, The First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou 310011, China

^b Department of Biophysics & Center of Cryo-Electron Microscopy, School of Medicine, Zhejiang University, Hangzhou 310011, China

ARTICLE INFO

Article history:

Received 12 May 2020

Revised 22 June 2020

Accepted 5 July 2020

Available online 30 August 2020

关键词

N-糖基化

COVID-19

棘突蛋白

hACE2

蛋白结构

摘要

当下新型冠状病毒肺炎(COVID-19)的国际大流行促使全球科学家努力尝试揭秘新冠病毒 HCoV-19 的生物学特性。质谱分析(MS)发现,在 HCoV-19 S 蛋白的 21 个潜在糖基化修饰位点中,有 20 个位点完全被 N-糖链占据,其糖型以低聚甘露糖为主。人血管紧张素转换酶 2(hACE2)的 7 个糖基化位点完全被复合型 N-糖链占据。然而,糖基化修饰并不能直接影响 HCoV-19 S 蛋白与 hACE2 之间的结合亲和力。另外,我们还发现了 S 蛋白和 hACE2 的多个甲基化修饰位点,以及 hACE2 的多个羟脯氨酸修饰位点。通过在最近发表的冷冻电镜(cryo-EM)结构的基础上加入 N-糖链和蛋白质翻译后修饰(PTM),我们构建了 HCoV-19 S 蛋白和 hACE2 的精细结构模型。本研究揭示的 HCoV-19 S 蛋白和 hACE2 的 PTM 及糖基化图谱为研究病毒的宿主黏附、免疫反应,以及相关药物和疫苗的研发提供了更多的蛋白质结构细节。

©2020 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. 引言

自 2019 年 12 月暴发的新型冠状病毒肺炎(COVID-19)已成为全球卫生领域的重大事件[1,3]。该疾病是由一种新发现的 β-冠状病毒 HCoV-19 (也称为 SARS-CoV-2)引起的。该病毒与严重急性呼吸综合征冠状病毒(SARS-CoV)密切相关。与 SARS 病毒一样, HCoV-19 病毒会引起下呼吸道感染(LRTI),最终可能发展为需要重症监护和机械通气的非典型肺炎[4,6]。与 SARS-CoV 相比, HCoV-19 具有更强的传染性,在全球范围内会有更多的感染人群,因此死亡人数更多[4-5]。虽然研究表明 HCoV-19 可能存在包括蝙蝠和穿山甲在内的多个中间宿主[3,7],

但仍需更多的研究来阐明其确切的病毒来源和传播途径[8,10]。迄今为止,还未出现针对冠状病毒感染(包括 HCoV-19)的特效疫苗或药物。

与 SARS-CoV 相似,由 HCoV-19 S 基因编码的病毒表面棘突糖蛋白(S 蛋白)对靶细胞的黏附和融合过程至关重要[11,13]。鉴于 S 蛋白在 COVID-19 发病机制中的重要性及其对疫苗开发的潜在免疫原性,在首个 HCoV-19 序列发表后不久,全球范围内多个研究组开始致力于阐明 S 蛋白的结构[1,13-14]。HCoV-19 S 糖蛋白是一种由 1273 个氨基酸构成的前体多肽,它可被宿主细胞蛋白酶(组织蛋白酶 L, TMPRSS2)切割成含连接宿主受体——人血管紧张素 I 转换酶 2(hACE2)——受体结合结构域(RBD)

* Corresponding author.

E-mail address: ljli@zju.edu.cn (L. Li).

的S1片段, 以及负责随后细胞膜融合的S2片段[1,12-13]。S蛋白被预测含有一个可切割的N端信号序列(1~15), 它可将S蛋白导向内质网(ER), 并在病毒颗粒包装之前进行广泛的聚糖修饰。

糖基化修饰是许多病毒棘突蛋白或包膜蛋白中最重要的翻译后修饰(PTM)类型之一, 且被证明可介导细胞附着、免疫反应以及病毒粒子包装和出芽过程[15,24]。各类冠状病毒与其各自受体的结合被证明是由S蛋白寡甘露糖N-糖链介导的[19-20,23]。此外, C型凝集素DC-SIGN和L-SIGN可通过与S蛋白糖链结合来促进病毒入侵[25-26]。因此, 通过干扰蛋白质糖基化或基于聚糖的相互作用可设计许多抗病毒策略[21,24,27-28]。糖基化也被认为是开发有效疫苗的关键因素[29-30]。

SARS-CoV冠状病毒的棘突蛋白含有23个N-连接的糖基化序列(N-X-S/T, X≠P)。与其类似, HCoV-19棘突蛋白预计每个单体包含22个糖基化序列, 也即每个三聚体含66个糖基化序列[13]。然而, HCoV-19 S蛋白的S1区的糖基化位点与SARS-CoV有所不同, 而S2区的糖基化位点则明显保守。这种S1糖基化模式的差异可能与HCoV-19和其他冠状病毒的生物学和临床特征的差异相关。在本研究中, 我们报道了HCoV-19 S蛋白和hACE2的N-糖基化及其他种类的PTM的高分辨率质谱分析(MS)结果, 并在此基础上重新描绘了目前HCoV-19 S蛋白的最新空间结构模型, 以突出与COVID-19发病机制相关的重要糖基化特征。但是, 研究证明, HCoV-19 S蛋白与hACE2的结合与糖基化状态无关。

2. 材料与方法

2.1. HCoV-19 S蛋白胞外结构域和hACE2蛋白的表达与纯化

利用昆虫细胞表达的优化密码子合成HCoV-19的S基因(病毒株: 武汉Hu-1; GenBank号QHD43416.1), 然后将其胞外区(Vall6-Pro1213)克隆到pFastBac载体(Life Technologies Inc., 美国)中, 并在N端和C端分别连接蜜蜂毒信号肽及His6和Flag标签。利用Bac-to-Bac系统(Life Technologies Inc., 美国), 在Sf9昆虫细胞中表达HCoV-19 S蛋白, 再从培养基中提取蛋白质, 并先后使用Ni-NTA柱和Superdex 200凝胶过滤柱(GE Healthcare, 英国)纯化S蛋白。含C端Fc标签的hACE2(Gln18-Ser740, NP_068576.1)胞外结构域则在HEK 293细胞中表达, 之后通过蛋白A sepharose层析柱(GE-Healthcare, 英国)进行纯化。

2.2. 聚丙烯酰胺凝胶电泳分析

为了检测HCoV-19 S蛋白和hACE2蛋白的糖基化状态, 在37 °C的磷酸盐平衡生理盐水(PBS)中用PNGase F(NEB, 1:50)对两者进行过夜去糖基酶解, 然后用15%的聚丙烯酰胺凝胶电泳(SDS-PAGE)和考马斯亮蓝染色分析糖基化和去糖基化形式的HCoV-19 S蛋白和hACE2蛋白。

2.3. 糖肽样品制备

参照之前文献[31]中的方法, hACE2和S蛋白先在含有10 mmol·L⁻¹二硫苏糖醇的50 mmol·L⁻¹碳酸氢铵(ABC)溶液中反应45 min, 之后在室温下用碘乙酰胺烷基化45 min。然后利用丙酮沉淀法收集蛋白质, 并在50 mmol·L⁻¹ ABC溶液中重悬。使用质谱级胰蛋白酶、糜蛋白酶或Lys-C蛋白酶(所有蛋白酶均购自美国普洛麦格公司), 按蛋白质与蛋白酶的比为1:50在37 °C下消化14 h。酶解后的多肽用固相萃取HLB柱(Waters, 美国)脱盐。

多肽样品被分成两份。第一份样品在纯H₂O¹⁸重水制备的50 mmol·L⁻¹ ABC中用PNGase F(NEB, 1:100)进行37 °C去糖基处理。第二份样品用亲水相互作用法(HIL-IC)富集完整的N-糖肽。简而言之, GlycoWorks柱(Waters, 美国)先用加样缓冲液[含有15 mmol·L⁻¹乙酸铵(AmA)和0.1%三氟乙酸(TFA)的80%乙腈(CAN)]进行预处理, 再将溶解在加样缓冲液中的多肽加载于GlycoWorks柱中, 并将未结合的流穿液收集起来。GlycoWorks柱用缓冲液洗涤两次后, 用0.1% TFA洗脱柱子上结合的糖肽。在进行液相串联质谱LC-MS/MS分析之前, 所有肽都用自制的C18柱进行脱盐处理。

2.4. LC-MS/MS实验

使用与Ultimate 3000液相色谱系统(Thermo Scientific, 美国)串连的Q-Exactive HFX质谱仪(Thermo Scientific, 美国)对多肽进行分析。质谱仪进行数据依赖性采集(DDA), 并只对强度排在前20的前体离子进行高能碰撞解离(HCD)和MS2扫描。使用Xcalibur 2.3版本软件(Thermo Scientific, 美国)记录MS和MS/MS谱图。液相色谱(LC)分离和MS采集的详细参数分别见Appendix A中的表S1和表S2。

2.5. 生物信息学

利用MaxQuant(1.6.10.43版本; Max-Planck-Institute of Biochemistry, 德国), 将从去糖基肽分析中获得的MS原始文件, 对照UniProtKB中的人ACE2序列和美国国家

生物技术信息中心 (NCBI) 中的 HCoV-19 S 序列 (YP_009724390.1_3) 进行搜索。将半胱氨酸氨甲酰化设置为固定修饰, 将蛋氨酸氧化和 ^{18}O 脱酰胺化设置为可变修饰。将胰蛋白酶的漏切位点数目设置为最高两个。预搜索和主搜索的质量误差分别设置为 15 ppm 和 4.5 ppm。对于全面的 PTM 分析, 将磷酸化 (S/T/Y)、乙酰化 (K)、甲基化 (K/R/E)、琥珀酰化 (K)、巴豆酰化 (K)、法尼酰化 (K/N_{term})、肉豆蔻酰化 (K/N_{term})、棕榈酰化或异戊烯化、糖基磷脂酰肌醇和脯氨酸氧化分别进行平行数据库检索。多肽水平的假阳性率 (FDR) 设置为 1%, 同时基于 99% 的定位概率对 PTM 进行了置信度筛选。利用修饰肽 (MP) 和相应的非修饰肽 (NP) 的峰值强度, 通过公式 $\text{MP}/(\text{MP}+\text{NP}) \times 100$ 计算每种 PTM 的位点占有率。

为了研究 *N*-糖基化形式, 我们用 pGlyco (2.2.2 版本; 中国科学院) [32] 根据相同的 hACE2 和 HCoV-19 S 序列搜索 HILIC 实验获得的 MS 原始文件。将半胱氨酸氨甲酰化设置为固定修饰, 将蛋氨酸氧化设置为可变修饰。将胰蛋白酶的漏切位点数目设置为最高两个。首次搜索和主搜索的质量误差分别设置为 15 ppm 和 4.5 ppm。MS2 图谱中潜在的糖链片段通过内置的糖链结构数据库 pGlyco.gdb 进行注释 [32]。利用 MaxQuant 特征检测算法提取每个糖肽的峰强度。本研究所有 LC-MS/MS 原始数据均存放在 iProX 数据库 (<https://www.iprox.cn/page/PSV023.html? url=1631583654953TDY8; password: th2J>)。

2.6. 蛋白结构模型的改进

通过 GLYCAM-Web (<https://dev.glycam.org/gp>) 下载或预测 PDB 格式的糖链结构模型。在前人构建的蛋白质模型 (PDB 编码 hACE2 蛋白的 6M18 及 S 蛋白的 6VXX) 基础上, 通过 Coot 软件在各个位点上添加质谱鉴定的糖基化修饰位点, 得到 hACE2 和三聚体 S 蛋白的 *N*-糖基化蛋白晶体结构 [33]。此外, 我们还用 HCoV-19 S 蛋白和 hACE2 复合物 (PDB 6M0J) 结构模型展示出 PTM 在两个蛋白结合区周围的分布。所有结构图均通过 PyMOL (Schrödinger Inc., 美国) 构建。

2.7. 生物膜干涉技术 (BLI) 评价 HCoV-19 S 蛋白与 hACE2 的结合力

采用 Octet Re96E 干涉仪 (ForteBio, 美国) 测定 HCoV-19 S 蛋白与 hACE2 的结合力。简而言之, 先使用氨基丙基硅烷生物传感器 (18-5045; ForteBio, 美国) 对 HCoV-19 S 蛋白进行固定化处理。为了评估 HCoV-19 S 的糖基化对结合力的影响, 我们预先在 37 °C 条件下将 S 蛋

白与 $1000 \text{ U} \cdot \text{mL}^{-1}$ 的 PNGase F 去糖基化酶或灭活的 PNGase F 共孵育 14 h。之后先用洗涤液 (PBS, pH 为 7.4, 0.01% 的牛血清白蛋白 BSA 和 0.002% 的吐温-20) 进行清洗, 再将 HCoV-19 S 蛋白与动力缓冲液中不同浓度 ($6.25 \text{ nmol} \cdot \text{L}^{-1}$ 、 $12.5 \text{ nmol} \cdot \text{L}^{-1}$ 、 $25 \text{ nmol} \cdot \text{L}^{-1}$ 、 $50 \text{ nmol} \cdot \text{L}^{-1}$ 、 $100 \text{ nmol} \cdot \text{L}^{-1}$ 、 $200 \text{ nmol} \cdot \text{L}^{-1}$) 的 hACE2 进行孵育, 通过传感器测量两种蛋白质之间的结合力, 其反应条件为在 30 °C 下停留 180 s。在结合阶段之后, 将传感器置于 30 °C 的 $1 \times$ 动力学缓冲液中停留 300 s 以解离蛋白质。在采用等摩尔结合模型进行数据拟合之前, 需要减去信号基线。结合速率常数 K_{on} 和解离速率常数 K_{off} 的平均值通过使用所有数据的全局拟合模型来确定。我们还通过使用不同浓度的 HCoV-19 S 蛋白溶液孵育载有固定化处理的 hACE2 传感器进行平行实验。同样, 固定化的 hACE2 用 PNGase F 或失活的 PNGase F 进行预处理, 以评估 hACE2 糖基化对蛋白质结合力的影响。在所有实验中, PNGase F 的失活处理条件是 75 °C 加热 10 min。

3. 结果

3.1. HCoV-19 S 蛋白和 hACE2 的 *N*-连接糖基化位点的测定

如图 1 (a) 所示, 在 SDS-PAGE 上, PNGase F 的去糖基化作用使 HCoV-19 S 蛋白和 hACE2 的分子量均降低。PNGase F 的消化作用使得 *N*-连接糖基化从天冬酰胺的 N-X-S/T 基序中释放, 这也引起了天冬酰胺的脱酰胺反应, 即加入来自溶剂的一个氧的同时损失了一个氢和一个氮 [34]。为了对糖基化位点进行彻底的排查, 我们将这两种蛋白的酶解多肽在 H_2O^{18} 重水中进行 PNGase F 去糖基化处理, 这引起 ^{18}O 的掺入, 并导致 2.98 Da 的质量增量, 可用于特异性标记糖基化位点。

通过 LC-MS/MS 对 S 蛋白的去糖基化肽进行分析, 确定了 20 个 *N*-糖基化位点, 包括 N61、N74、N122、N165、N234、N282、N331、N343、N603、N616、N657、N709、N717、N801、N1074、N1098、N1134、N1158、N1173 和 N1194 (表 1)。两个 N-X-S/T 位点——N17 和 N149 在任何去糖基化肽中均未被鉴定。然而, 如后文所述, 在未经 PNGase F 处理的样本中直接鉴定出了糖基化位点 N149。因此在 HCoV-19 S 蛋白中的 22 个 N-X-S/T 位点中, 共有 21 个糖基化位点被鉴定 [图 1 (b)]。此外, 对位点占有率的定量分析表明, 共有 18 个位点被完全糖基化, 而 N603、N657 的糖基化占有率分别为 43% 和 74% (表 1)。显示 N331 和 N343 糖基化修饰证据的示例质谱图见图 1 (c)。

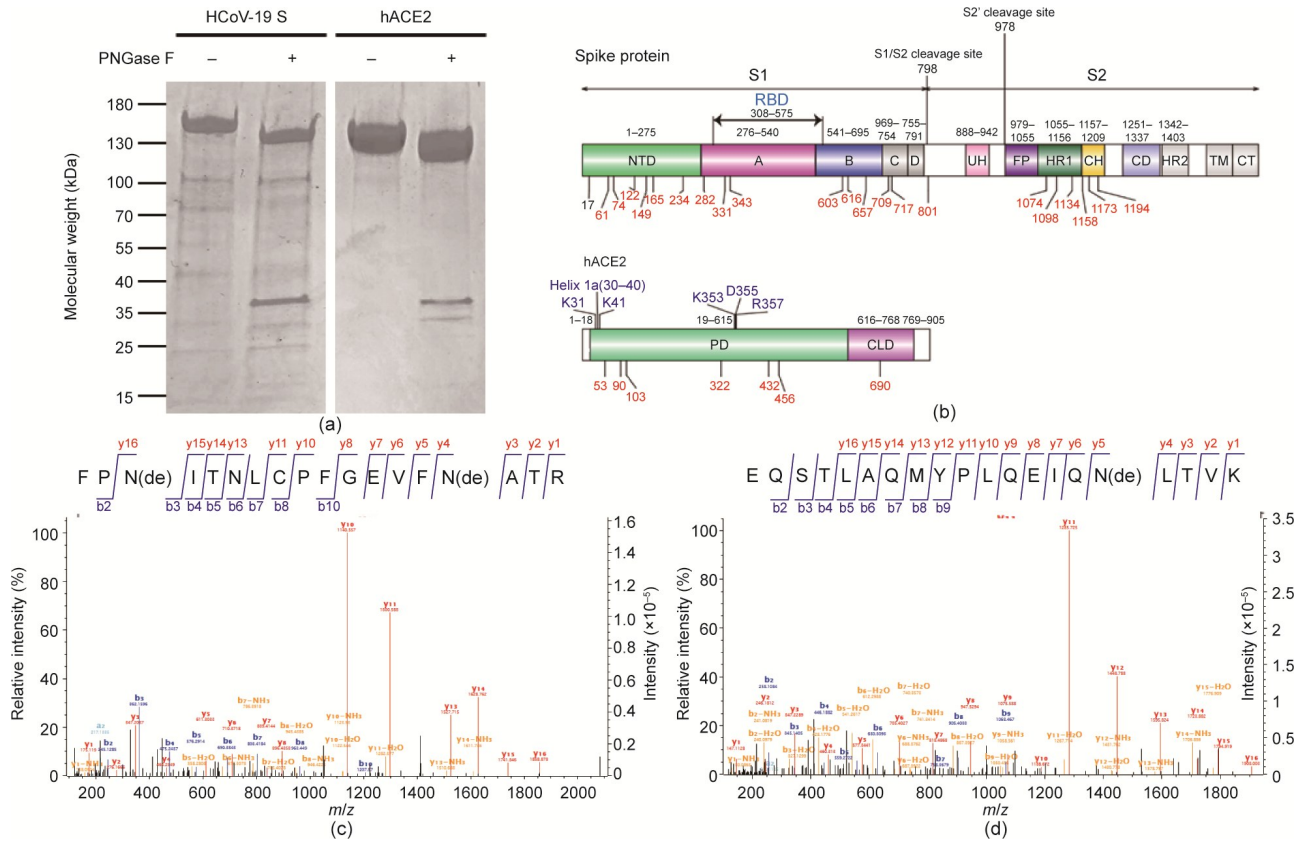


图1. HCoV-19 棘突蛋白和人hACE2潜在的糖基化位点。(a) 利用15% SDS-PAGE胶对糖基化与去糖基化的HCoV-19棘突蛋白及hACE2进行分析, 左侧为分子量标记。(b) HCoV-19棘突蛋白(上图)及hACE2(下图)的功能域示意图。CD: 连接区; CH: 中心螺旋区; CT: 细胞质尾部区; FP: 融合肽; TM: 穿膜区; UH: 上游螺旋区; HR1/2: 七肽重复区1/2。蓝色区域为可能的HCoV-19棘突蛋白与hACE2的结合区域。每个区域中可能的糖基化位点均已标出。红色标出的是本研究中证实存在糖基化的位点。HCoV-19棘突蛋白N331和N343位点(c)以及hACE2的N90位点(d)的去糖基质谱图。

hACE2蛋白中所有7个潜在的糖基化位点——N53、N90、N103、N322、N432、N546和N690, 在我们的实验中都证实是糖基化的(表1)。N90位点糖基化的质谱

图如图1(d)所示。位点占有率的定量分析显示7个位点都是完全糖基化(表1)。所有这些数据表明HCoV-19 S蛋白和hACE2都被N-糖基化高度修饰。

表1 HCoV-19 S蛋白和hACE2去糖基化肽中PTM的鉴定总结

Sites	PTM	Sequence	Occupancy (%)
HCoV-19 S Protein			
61, 74	N-glycosylation	SSVLHSTQDLFLPFFSN ^a VTWFHAIHVS ^a GTN ^a GTK	100
74	N-glycosylation	HAIHVS ^a GTN ^a GTKRF	100
78	Methylation	R ^a FDNPVLPFNDGVYFASTEK	100
122	N-glycosylation	TQSLIVNN ^a ATNVVIK IVNN ^a ATNVVIK/CEF	100
165	N-glycosylation	VYSSANN ^a CTFEYVSQPF ^a LM ^a DLE ^a GK	100
224	Methylation	DLPQGFSALE ^a PLVDLP ^a IGIN ^a ITR	100
234	N-glycosylation	V ^a DLPIGIN ^a ITRF	100
282	N-glycosylation	YNEN ^a G ^a TI ^a DAVDCALDPL ^a SETK NEN ^a G ^a TI ^a DAVDCALDPL ^a SETKCTL	100
331	N-glycosylation	RVQPTESIVRFPN ^a ITNL	100
331, 334	N-glycosylation	FPN ^a ITNLCPFG ^a EVFN ^a ATR	100
340	Methylation	FPNI ^a ITNLCPFG ^a EVFN ^a ATR	2.02
603	N-glycosylation	GGVSVIIPGTIN ^a TSNQVAVL	43.50

Sites	PTM	Sequence	Occupancy (%)
616	<i>N</i> -glycosylation	YQDVN ^a CTEVPVAIHADQL	100
654, 661	Methylation	AGCLIGAE ^a HVNN ^a SYE ^a CDIPIGAGICASYQTQTNSPR	100
657	<i>N</i> -glycosylation	AGCLIGAEHVN ^a NSYECDIPIGAGICASY	74.80
		IGAEHVN ^a NSYECDIPIGAGICASY	79.76
709, 717	<i>N</i> -glycosylation	SN ^a NSLAIPTN ^a F	100
801	<i>N</i> -glycosylation	TPPIKDFGGFN ^a FSQILPDPSKPSK	100
		N ^a FSQILPDPSKPSKRSF	100
1074	<i>N</i> -glycosylation	N ^a FTTAPAICHHDGK	100
1098	<i>N</i> -glycosylation	VSN ^a GTHWF	100
1134	<i>N</i> -glycosylation	VSGNCDVVIGIVN ^a NTVY	100
1158, 1173	<i>N</i> -glycosylation	N ^a HTSPDVLGDISGIN ^a ASVVNIQKEIDRLNEVAK	99.17
1194	<i>N</i> -glycosylation	NLN ^a ESLIDLQELGK	99.73
hACE2			
53	<i>N</i> -glycosylation	FNHEAEDLFYQSSLASWN ^a YNTN ^a ITEENVQNMN ^a NAGDK	100
		N ^a YNTN ^a ITEENVQNMN ^a NAGDKW	100
68	Methylation	N ^a YNTN ^a ITEENVQNMN ^a NAGDK ^a W	100
84	Hydroxyproline	EQSTLAQMYP ^a LQEIQN ^a LTVK	20.74
90	<i>N</i> -glycosylation	WSAFLKEQSTLAQMYP ^a LQEIQN ^a LTVK	100
103	<i>N</i> -glycosylation	LQLQALQQN ^a GSSVLSEDKSK	100
		QLQALQQN ^a GSSVL	100
253	Hydroxyproline	LMNAYP ^a SYISPIGCLPAHLLGDMWGR	98.35
263	Hydroxyproline	LMNAYPSYISPIGCLP ^a AHLLGDMWGR	88.57
284	Hydroxyproline	FWTNLYSLTVP ^a FGQKPNIDVTDAMVDQAWDAQR	4.06
289	Hydroxyproline	P ^a NIDVIDAMVDQAWDAQR	1.70
321	Hydroxyproline	FFVSVGLP ^a NMTQGFWENSMLTDPGNVQK	58.24
322	<i>N</i> -glycosylation	FFVSVGLPN ^a MTQGFWENSMLTDPGNVQK	100
		VSVGLPN ^a MTQGF	100
329	Methylation	FFVSVGLPNMTQGFWE ^a NSMLTDPGNVQK	100
346	Hydroxyproline	AVCHP ^a TAWDLGK	88.88
415	Hydroxyproline	NGANEGFHEAVGEMSLSAATP ^a K	0.68
432	<i>N</i> -glycosylation	SIGLLSPDFQEDN ^a ETEINFLK	100
		LSPDFQEDN ^a ETEINF	100
451	Hydroxyproline	QALTVGTLP ^a FTYMLEK	43.59
469	Hydroxyproline	GEIP ^a KDQWMK	4.09
546	<i>N</i> -glycosylation	CDISN ^a STEAGQK	100
		HKCDISN ^a STEAGQKLF	100
565	Hydroxyproline	SEP ^a WTLALENVVGA	0.37
583	Hydroxyproline	NMNVRP ^a LLNYFEP ^a LFTWLK	2.12
612	Hydroxyproline	NSFVGWSTDWSP ^a YADQSIK	0.21
690	<i>N</i> -glycosylation	ISFNFFVTAPKN ^a VSDIIPR	100
		VTAPKN ^a VSDIIPRTEVEKAIRM	100
		N ^a VSDIIPRTEVEK	100

^a Modified sites.

3.2. HCoV-19 S蛋白和hACE2的其他PTM鉴定

为了揭示糖基化以外的可能PTM，我们对HCoV-19 S蛋白和hACE2去糖肽的LC-MS/MS数据进行了多轮谱图

数据库匹配。在这两种蛋白质中，我们发现PTM以赖氨酸、精氨酸或谷氨酸的甲基化为主（表1）。以占有率大于50%为标准，我们发现HCoV-19 S蛋白中的78R、

224E、654E和661E，及hACE2中的57E、68K和329E高度甲基化。此外，hACE2中253、263、321和346位点的脯氨酸显著被氧化从而转化为羟脯氨酸。然而，HCoV-19 S蛋白和hACE2中均未发现磷酸化、乙酰化及其他形式的脂肪酸酰化。

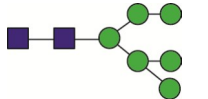
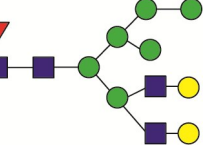
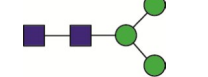
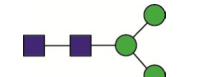

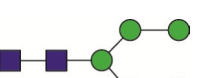
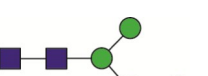
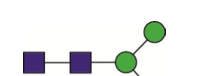
3.3. HCoV-19 S蛋白和hACE2的N-糖基化修饰图谱

为了揭示HCoV-19 S蛋白和hACE2表面覆盖的糖链，我们直接利用专用于检测N-糖基化修饰的高分子量肽的LC-MS/MS方法，对蛋白酶消化和HILIC SPE分馏后得到的完整糖肽进行分析。由于糖链的高度异质性，在我们的研究中，一种独特的糖肽被定义为一种独特的多肽序列和一种特定N-糖链形成的组合。根据该标准，我们分别从HCoV-19 S蛋白和hACE2中鉴定出了419个和467个独特的N-糖肽（表2、表3）。除N1134外，在PNGase-F去糖基化实验中，在鉴定出的20个HCoV-19 S糖基化位点中

有19个被含糖肽的质谱图证实。此外，N149位点的N-糖基化也被发现。尽管hACE2中的7个糖基化位点在带有脱酰胺标记的去糖基化肽中都被鉴定出来，但只在N90、N103、N432、N546和N690这5个位点上得到了N-糖基化图谱。HCoV-19 S蛋白和hACE2的N-糖基化图谱示例见Appendix A中的图S1。

HCoV-19 S蛋白中共发现了144种N-糖链，其中大多数含有共同的N-乙酰氨基葡萄糖核心（表2）。S蛋白中的所有N-糖基化位点均有多种N-糖链，其中N343位点含有种类最丰富的N-糖链形态。HCoV-19 S N-糖链成分主要是由寡糖型或高甘露糖型组成，但4个N-糖基化位点（N73、N343、N717和N1173）含有较高比例的复合型型和杂交型N-糖链[图2（a）、（b）和表2]。LC-MS分析表明，寡甘露糖Hex3HexNAc2是HCoV-19 S蛋白中最常见的N-糖链类型。有趣的是，在HCoV-19 S蛋白的N-糖链中没有发现唾液酸成分。

表2 HCoV-19S蛋白的糖肽和N-聚糖鉴定总结

Domain	Site	Sequences	Most abundant composition	Most likely structure
S1	61	FSN ^a VTWF FSN ^a VTW	Hex6HexNAc2	
	74	SSVLHSTQDLFLPFFSN ^a VTWFHAIHVSGTN ^a GTK SSVLHSTQDLFLPFFSN ^a VTWFHAIHVSGTN ^a GTKR HAIHVSGTN ^a GTKRF	Fuc1NeuGc2Hex6HexNAc4	
	122	TQSLIVNN ^a ATNVVIK IVNN ^a ATNVVIKVCEF LIVNN ^a ATNVVIKVCEF	Hex3HexNAc2	
	149	VCEFQFCNDPFLGVVYHKNN ^a K YHKNN ^a KSWMESEF HKNN ^a KSWMESEF	Hex3HexNAc2	
	165	VYSSANN ^a CTFEYVSQPFLMDLEGK SSANN ^a CTFEY SSANN ^a CTFEYVSQPF	Hex4HexNAc2	
	234	DLPQGFSALEPLVDLPIGIN ^a ITR VDLPIGIN ^a ITRF	Hex5HexNAc2	
	282	YNEN ^a GTITDAVDCALDPLSETK	Hex4HexNAc2	
	331	FPN ^a ITNLCPFGEVFN ^a ATR RVQPTESIVRFPN ^a ITNLCPF RVQPTESIVRFPN ^a ITNL	Hex3HexNAc2	

Domain	Site	Sequences	Most abundant composition	Most likely structure
	343	FPNITNLCPFGEVFN ^a ATRF CPFGEVFN ^a ATRF GEVFN ^a ATRF	Fuc1NeuGc2Hex5HexNAc4	
	603	GGVSVITPGTN ^a TSNQVAVLY GGVSVITPGTN ^a TSNQVAVL	Fuc1Hex3HexNAc2	
	616	YQDVN ^a CTEVPVAIHADQLTPTW QDVN ^a CTEVPVAIHADQLTPTW QDVN ^a CTEVPVAIHADQLTPTWRVY YQDVN ^a CTEVPVAIHADQL QDVN ^a CTEVPVAIHADQL	Hex3HexNAc2	
	657	AGCLIGAEHVN ^a NSYECDIPIGAGICASYQTQNSPR QTRAGCLIGAEHVN ^a NSYECDIPIGAGICASY IGAEHVN ^a NSYECDIPIGAGICASY QTRAGCLIGAEHVN ^a NSY	Fuc1Hex3HexNAc2	
S2	709	SN ^a NSIAIPTNF	Hex3HexNAc3	
	717	SNNSIAIPTN ^a F	Fuc1NeuGc2Hex4HexNAc4	
	801	DFGGFN ^a FSQILPDPSKPSK DFGGFN ^a FSQILPDPSKPSKR TPPIKDFGGFN ^a FSQILPDPSK TPPIKDFGGFN ^a FSQILPDPSKPSK DFGGFN ^a FSQILPDPSK N ^a FSQILPDPSKPSKRSF	Hex3HexNAc2	
S2(S2')	1074	N ^a FTTAPAICHGDK LHVTYVPAQEKN ^a F	Hex3HexNAc2	
	1098	EGVFVSN ^a GTHWFVTQR AHFPREGVFVSN ^a GTHWFVTQR VSN ^a GTHWF VSN ^a GTHW	Hex3HexNAc2	
	1134 ^b	—	—	—
	1158	N ^a HTSPDVDLGDISGINASVVNIQK YFKN ^a HTSPDVDLGDISGINASVVNIQK	Fuc1Hex3HexNAc2	
	1173	YFKNHTSPDVDLGDISGIN ^a ASVVNIQK NHTSPDVDLGDISGIN ^a ASVVNIQK	Fuc2NeuGc1Hex5HexNAc4	
	1194	NLN ^a ESLIDLQELGKYEQYIK NLN ^a ESLIDLQELGK	Fuc1Hex3HexNAc2	

^a Glycosylated *N* sites.

^b Sites identified only on deglycosylated peptides.

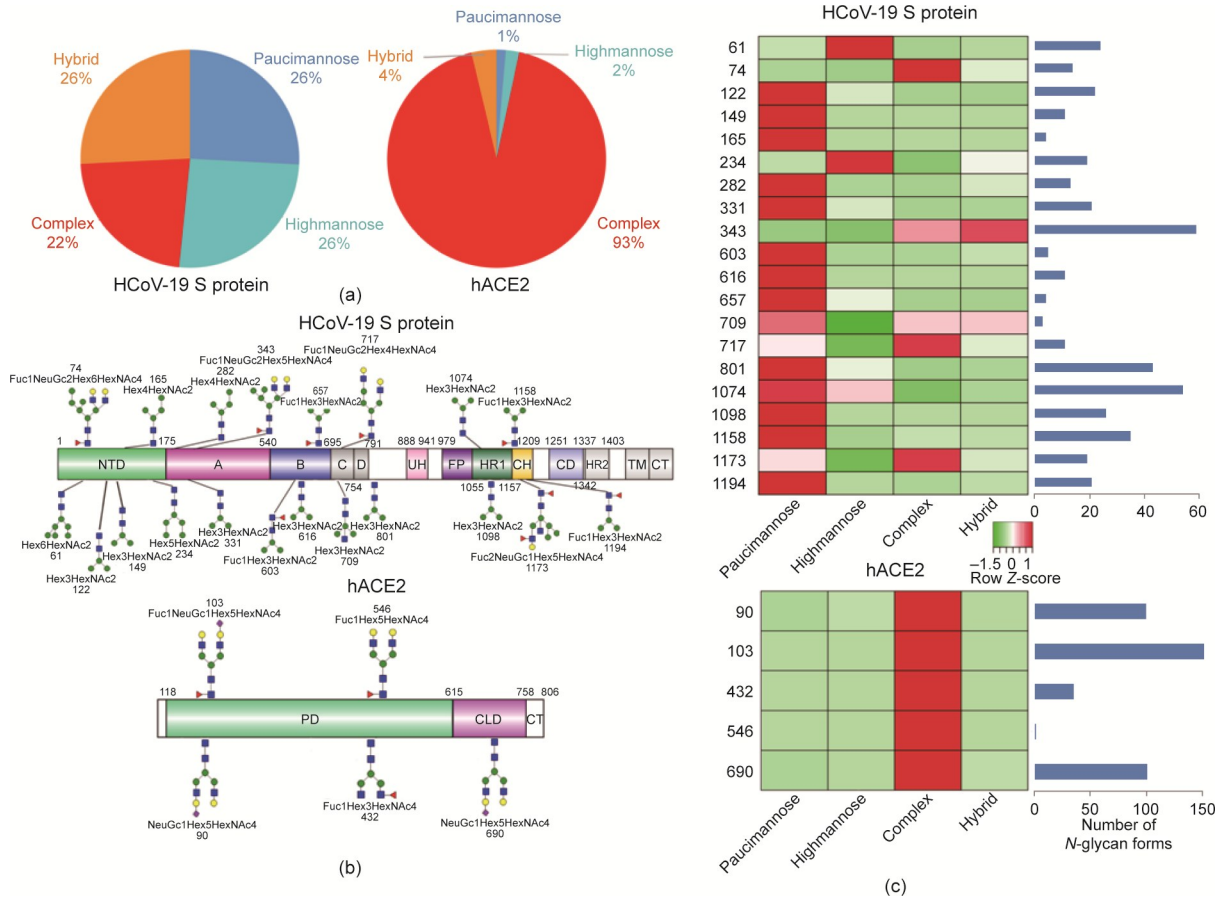


图2. HCoV-19 棘突蛋白和人 hACE2 的糖肽鉴定总结。(a) 四种主要糖型在两个蛋白中的整体比例；(b) 每个蛋白功能域中的糖基化位点上含量最高的糖基化类型图示；(c) 热图总结了每个位点上四种糖型的相对质谱信号的强度，右侧条状图显示了每个位点所有糖型种类总数。

hACE2 中共发现了 220 个 *N*-糖链，其中 93% 为复合型，而杂交型、高甘露糖型和寡甘露糖型糖链仅分别占 4%、2% 和 1% [图 2 (a) 和表 3]。LC-MS 分析表明，hACE2 多糖的主要成分为 NeuGc1Hex5HexNAc4，推测其为末端含有唾液酸的双支链复合糖。所有 hACE2 位点的 *N*-糖型主要为复合型。值得注意的是，位点 N90、N103 和 N690 都含有 100 多种 *N*-糖型，其中最主要的形式都含有唾液酸末端。

总之，我们的 LC-MS/MS 数据证实了 HCoV-19 S 蛋白及其受体 hACE2 在其预测的 N-X-S/T 序列中确实占有大量 *N*-糖链。HCoV-19 S 蛋白和 hACE2 的每个位点中最主要的 *N*-糖链成分和预测结构如图 2 (c) 所示。

3.4. 用糖基化和 PTM 改进的 HCoV-19 S-hACE2 复合物结构模型

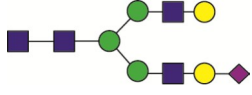
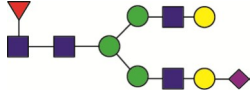
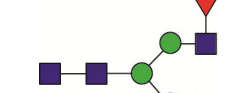
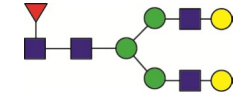
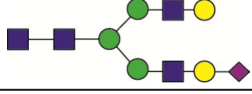
将 LC-MS/MS 结果中每个位点含量最丰富的 *N*-糖链结构添加到 HCoV-19 S 蛋白和 hACE2 的冷冻电镜模型中，生成了最新的原子模型，该模型代表了这两种蛋白质上 *N*-糖链最可能的空间分布 [图 3 (a)、(b)]。尽管 S 蛋白

的 4 个位点 (N74、N149、N1158 和 N1194) 在缺失残基模型 6VXX 中没有显示，但我们的模型表明，*N*-糖链屏蔽了 HCoV-19 S 蛋白表面 2/3 以上的部分，这可能与 S 蛋白的宿主黏附和免疫逃避相关。此外，HCoV-19 S-hACE2 结合模型显示，HCoV-19 S 蛋白的 N331、N343 位点糖链和 hACE2 的 N90 位点糖链在空间上靠近，尽管并不完全在它们的结合区内 [图 3 (c)、(d)]。HCoV-19 S-hACE2 结合模型还表明，hACE2 中的 3 个甲基化位点 (57E、68K 和 329E) 形成了一个三叉戟结构，包围了 hACE2 的 K353~R357 与 HCoV-19 S 的 N501 位点之间形成的接触区域 (见 Appendix A 中的图 S2)。

3.5. HCoV-19 S 蛋白与 hACE2 的结合不依赖 *N*-糖基化

为了解蛋白质糖基化对 HCoV-19 S 蛋白与 hACE2 之间相互作用的影响，我们在 BLI 实验中比较了纯化的 hACE2 胞外域与固定在生物传感器表面的糖基化和去糖基化的 HCoV-19 S-ECD 的结合动力学及亲和力。我们发现 hACE2 与活性或灭活的 PNGase F 预处理后的 HCoV-19 S 蛋白结合的平衡解离常数 K_d 分别为 $1.7 \text{ nmol} \cdot \text{L}^{-1}$ [图 4

表3 hACE2 蛋白的糖肽和 N-聚糖鉴定总结

Domain	Site	Sequences	Most abundant composition	Most likely structure
PD	53 ^a	–	–	–
	90	EQSTLAQMYPLQEIQN ^b LTVK	NeuGc1Hex5HexNAc4	
	103	LQLQALQQN ^b GSSVLSSEDK LQLQALQQN ^b GSSVLSSEDKSK LQLQALQQN ^b GSSVLSSEDKSKR	Fuc1NeuGc1Hex5HexNAc4	
	322 ^a	–	–	–
	423	SIGLLSPDFQEDN ^b ETEINFLK	Fuc1Hex3HexNAc4	
	546	CDISN ^b STEAGQK	Fuc1Hex5HexNAc4	
CLD	690	N ^b VSDIIPR	NeuGc1Hex5HexNAc4	
		ISFNFFVTAPKN ^b VSDIIPR		

^a Glycosylated N sites.

^b Sites identified only on deglycosylated peptides.

(a)]和 $1.5 \text{ nmol} \cdot \text{L}^{-1}$ [图4 (b)]。本研究中测定的 hACE2 与 HCoV-19 S 蛋白结合的亲和力与先前的报告一致[13]。当 HCoV-19 S ECD 与固定化的 hACE2 [$K_d = 16.7 \text{ nmol} \cdot \text{L}^{-1}$, 图4 (c)]或去糖基的 hACE2 [$K_d = 18.2 \text{ nmol} \cdot \text{L}^{-1}$, 图4 (d)]结合时, 也观察到类似的现象。因此, 我们的数据表明糖基化状态并不直接影响 HCoV-19 S 蛋白与 hACE2 的结合。BLI 结合动力学的详细总结见 Appendix A 中的表 S3。

4. 讨论

糖基化是一种普遍且复杂的蛋白质翻译后修饰类型, 它极大地丰富了蛋白质的结构和功能。然而, 表征蛋白质 N-糖基化在技术上具有挑战性。由于去糖基肽比其糖基化肽前体更容易被 LC-MS 分析, 为了确定 HCoV-19 S 蛋白和 hACE2 中的 N-糖基化位点, 本研究通过在 H_2O^{18} 的环境中, 由 PNGase F 介导的水解反应插入的 ^{18}O 同位素所造成的特殊修饰来确保糖基化位点鉴定的可信度。之后, 本研究通过进行完整的 N-糖肽分析进一步验证了糖基化修饰存在, 并提供了每个位点的 N-糖链组成和结构异质性的详细信息。我们的研究表明, 除 N17 位点外, HCoV-19 S 蛋白的所有位点都被糖基化高度修饰。根据 S 蛋白的 N-糖链模型, 蛋白质表面有很大一部分面积被糖链覆盖, 这与

以前各种 S 蛋白的报道一致[14,34,36]。当比较 HCoV-19 和 SARS-CoV 的 S 蛋白时, 我们注意到糖基化位点的大部分差异发生在 S1 亚单位的远端, 因此导致由棘突三聚体形成的病毒最外层的糖链谱有显著差异。糖基化的差异可能是宿主或环境压力选择的结果, 并可能暗示该病毒在感染性、发病机制和宿主反应方面的特性。在本研究进行过程中, 我们也注意到, 其他几个研究小组也在试图破解 HCoV-19 S 蛋白的糖链谱。所有这些研究都证实了 HCoV-19 S 蛋白是高度糖基化的[35–36]。

S 蛋白中的 N-糖链具有明显的异质性, 这极大地扩展了 S 蛋白的构象灵活性和表位多样性。虽然在 S 蛋白中发现了一小部分复合型 and 杂交型 N-糖链, 但大多数位点被低聚甘露糖型糖链占据, 这与先前关于 SARS-CoV 冠状病毒糖基化特征的描述是一致的[37,39]。许多病毒蛋白质都带有大量低聚甘露糖型糖链, 这可能是由高密度糖基化位点的空间位阻效应导致的糖链不完全成熟的标志[40]。有趣的是, 靠近结合面的 N343 位点具有最多多样性的 N-糖链类型, 且主要为杂交型和复合型的糖链。此外, 与 SARS-CoV 和 Marburg 病毒蛋白一样[41–42], 我们发现 SARS-CoV S 蛋白糖链中的唾液酸含量也是可以忽略的。鉴于目前全球正在努力研发以 S 蛋白为候选抗原的疫苗, 而糖链可以显著调节蛋白质的免疫原性, 因此疫苗开发人员应考虑在 S 蛋白三聚体表面占据大面积的不同糖链形式。有趣的是, 将我们的数据与其他糖基化分析研究进行比较时,

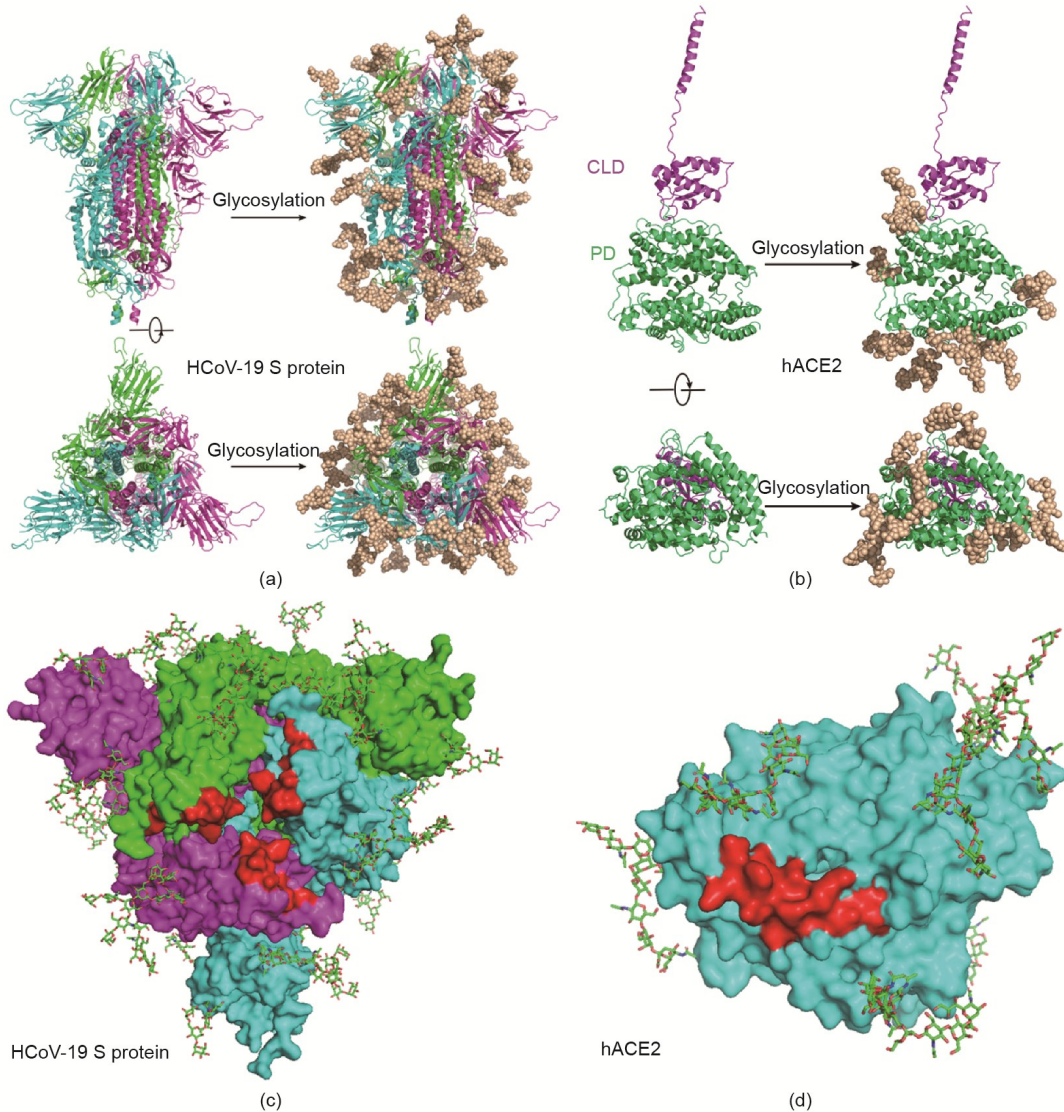


图3. 含糖链结构的HCoV-19 棘突蛋白三聚体和人hACE2的结构模型。(a) 每个单聚体由不同颜色表示的HCoV-19 棘突蛋白三聚体结构3D带状图；(b) 每个结构域由不同颜色表示的hACE2结构带状图。不含糖链（左侧图）与含糖链（右侧图）的蛋白质结构均提供了侧视图（上图）和病毒或细胞膜朝向的俯视图（下图）。HCoV-19 棘突蛋白三聚体（c）与hACE2（d）结合区域附近的糖链通过两个蛋白质的俯视图显示，其中红色为结合区域，枝条形状显示的是糖链。

我们发现不同的蛋白表达体系可以显著地影响糖基化类型[35–36]。与昆虫细胞相比，哺乳动物细胞更倾向于将复杂的糖链组装到S蛋白上。因此，我们也应对基于不同糖基化状态的S蛋白生产的疫苗的不同保护性保持警惕。

冠状病毒S蛋白与它们各自受体的结合被证明是由其寡甘露糖N-糖链介导的[19–20,23]。SARS-CoV冠状病毒S蛋白含有3个与RBD相关的糖基（N318、N330和N357），而HCoV-19只有两个（N331和N343）糖基化位点围绕在结合口袋区周围。我们假设这两个位点的N-糖链可能参与受体结合的极性相互作用。但事实与之相反，BLI结合实验证明，去糖基化不会改变S蛋白对hACE2的亲和力。然而，这种阴性的BLI结合结果并不能够排除糖

基化影响其他病毒入侵步骤的可能性，包括蛋白酶切割和聚糖介导DC/L-SIGN相互作用，这些步骤被证明可以增强SARS-CoV的病毒的入侵过程[25–26]。因此需要进一步的研究来确定糖基化在HCoV-19感染过程中的作用。

抗原糖基化在很大程度上决定了宿主的免疫反应。糖基化的一个重要功能是通过屏蔽病毒包膜或表面蛋白中的免疫原性表位来实现逃避免疫[17,43–44]。HCoV-19 S蛋白的大部分糖基化位点被糖基化完全修饰，表明该病毒能够以隐秘的方式入侵宿主细胞。HCoV-19在感染早期成功地逃避了先天性免疫，这可能解释了该病毒的较长的无症状潜伏期的由来[4,6]。HCoV-19短期内再感染的病例也已有报道，表明病毒可能会逃避抗体介导的中和作用[45]。

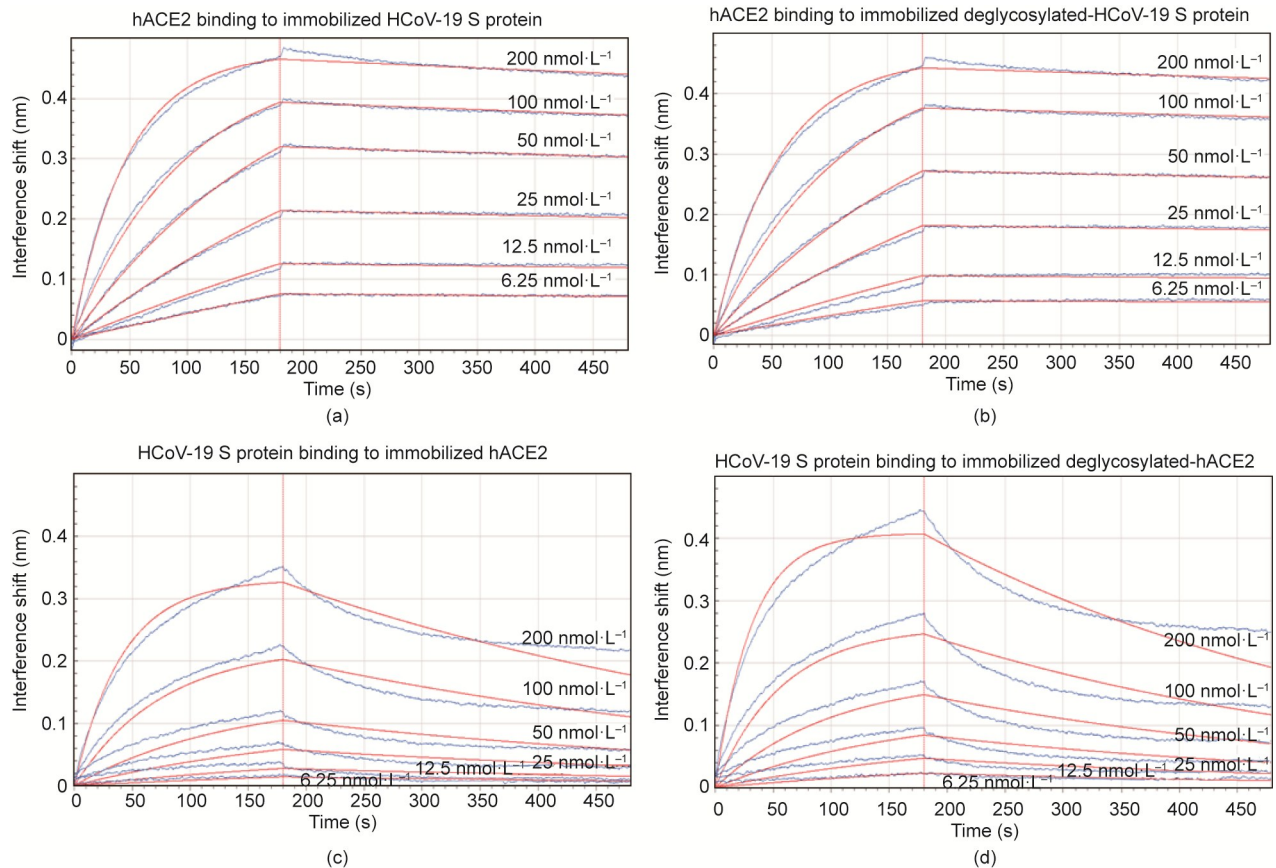


图4. 糖基化对HCoV-19 棘突蛋白和人hACE2结合的影响。生物膜层干涉仪检测HCoV-19 棘突蛋白和人hACE2的结合。在生物传感器上固定完整的HCoV-19 棘突蛋白(a)和去糖基化的棘突蛋白(b)并将其暴露在含6个不同浓度hACE2的溶液中。互换实验中则将完整的hACE2(c)和去糖基化的hACE2(d)固定后与含6个不同浓度HCoV-19 棘突蛋白进行孵育。红色标识的是通过等摩尔结合模型得出的数据整体拟合曲线。

另外, 抗体也可以通过糖基化反应促进对其他免疫分子的识别。*N*-糖链是半乳糖凝集素的已知配体, 可触发细胞因子释放和免疫细胞浸润等炎症事件[46–47], 这些事件可能参与了HCoV-19的发病机制, 并与疾病严重程度相关。此外, 流行病学研究表明, ABO血型多态性与HCoV-19和SARS-CoV在人群中的易感性差异有关[48–49]。由于ABH糖链表位和病毒蛋白糖链都是由内质网糖基化酶系统合成的, 抗-A或抗-B抗体可部分阻断病毒与宿主的相互作用[50], 因此O型血的个体对病毒感染的抵抗力更强。S蛋白的糖基化状态是否影响感染人群中病毒的反应和个体间临床结局的差异仍需进一步研究。

本研究还分析了hACE2的糖基化状况。hACE2中的7个糖基化位点均被糖链完全占据, 包括结合表面附近的N90位点。然而, 我们发现糖链并没有直接影响其与HCoV-19 S蛋白的结合。这与之前的发现一致, 即ACE2糖基化的抑制并不影响其与SARS-CoV冠状病毒S蛋白的结合[51]。但是去糖基化的ACE2确实可以抑制病毒入侵及传染性SARS-CoV病毒粒子的生成过程[51]。因此, 糖基化的ACE2仍然被认为是干预冠状病毒感染的重要靶

点。事实上, 氯喹对抗HCoV-19感染的能力被认为是其抑制ACE2糖基化以及提高核内体pH值的共同结果[28,52]。

除糖基化外, 我们还研究了其他翻译后修饰类型。甲基化修饰在S蛋白和hACE2的多个位点都被鉴定。尤其是hACE2与S蛋白的RBD结合位点周围的57E、68K和329E位点被完全甲基化。甲基化会导致电荷损失, 从而增加这些位点的疏水性。在hACE2的蛋白酶结构域中鉴定出4种羟脯氨酸修饰位点(即253、263、321和346位点)。额外的羟基可能会增加脯氨酸在hACE2细胞外区域的亲水性。未来还需要进一步研究以揭示这些PTM可能的生物学作用。在我们的数据集中没有发现磷酸化和乙酰化修饰, 这和先前对SARS-CoV的PTM研究相一致[38]。脂肪酸酰化在本研究中也考虑, 因为它们通常存在于表面病毒蛋白中, 可以促进膜融合和病毒入侵。然而, 本研究并没有发现HCoV-19 S蛋白中存在酰化修饰的证据。考虑到多种蛋白酶处理的蛋白质组学实验几乎可以实现对这两种蛋白质的全覆盖, 我们认为HCoV-19 S蛋白和hACE2的翻译后修饰组学主要由糖基化、甲基化和脯氨

酸氧化组成。

致谢

本研究得到了国家重点研发项目(2017YFC1200204、2017YFA0504803、2018YFA0507700)、浙江省科技厅应急项目(2020C03123-1)、中央高校基础研究经费(2018XZX01-13)和传染病诊治国家重点实验室自主项目基金的支持。感谢浙江大学传染病诊治国家重点实验室蛋白质组学和代谢组学平台进行的糖蛋白组学分析。感谢杭州双天生物技术有限公司协助进行了BLI分析。感谢中国科学院计算技术研究所pGlyco团队提供的技术支持。感谢马聘博士对蛋白质表达和纯化提供的建议。

Compliance with ethics guidelines

Zeyu Sun, Keyi Ren, Xing Zhang, Jinghua Chen, Zhengyi Jiang, Jing Jiang, Feiyang Ji, Xiaoxi Ouyang, and Lanjuan Li declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2020.07.014>.

References

- [1] Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020;579:265–9.
- [2] Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect Dis* 2020;20(5):533–4.
- [3] Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 2020;579:270–3.
- [4] Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708–20.
- [5] Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395(10229):1054–62.
- [6] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. *JAMA* 2020;323(11):1061–9.
- [7] Zhang T, Wu Q, Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol* 2020;30(7):1346–51.
- [8] Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, et al. Genome composition and divergence of the novel coronavirus (2019-nCoV). *Cell Host Microbe* 2020;27:325–8.
- [9] Xu Y. Unveiling the origin and transmission of 2019-nCoV. *Trends Microbiol* 2020;28(4):239–40.
- [10] Zhang YZ, Holmes EC. A genomic perspective on the origin and emergence of SARS-CoV-2. *Cell* 2020;181(2):223–7.
- [11] Wan Y, Shang J, Graham R, Baric RS, Li F. Receptor recognition by the novel coronavirus: an analysis based on decade-long structural studies of SARS coronavirus. *J Virol* 2020;94(7):e00127–220.
- [12] Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen 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):271–80.e8.
- [13] Walls AC, Park YI, Tortorici MA, Wall A, McGuire AT, StructureVeesler D., function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* 2020;181(2):281–92.e6.
- [14] Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* 2020;367(6483):1260–3.
- [15] Fukushi M, Yoshimaka Y, Matsuoka Y, Hatakeyama S, Ishizaka Y, Kirikae T, et al. Monitoring of S protein maturation in the endoplasmic reticulum by calnexin is important for the infectivity of severe acute respiratory syndrome coronavirus. *J Virol* 2012;86(21):11745–53.
- [16] de Groot RJ. Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj J* 2006;23(1–2):59–72.
- [17] Raman R, Tharakaraman K, Sasisekharan V, Sasisekharan R. Glycan-protein interactions in viral pathogenesis. *Curr Opin Struct Biol* 2016;40:153–62.
- [18] Chang D, Zaia J. Why glycosylation matters in building a better flu vaccine. *Mol Cell Proteomics* 2019;18(12):2348–58.
- [19] Li W, Hulswit RJG, Widjaja I, Raj VS, McBride R, Peng W, et al. Identification of sialic acid-binding function for the Middle East respiratory syndrome coronavirus spike glycoprotein. *PNAS* 2017;114(40):E8508–17.
- [20] Parsons LM, Bouwman KM, Azurmendi H, de Vries RP, Cipollo JF, Verheije MH. Glycosylation of the viral attachment protein of avian coronavirus is essential for host cell and receptor binding. *J Biol Chem* 2019;294:7797–809.
- [21] Shih YP, Chen CY, Liu SJ, Chen KH, Lee YM, Chao YC, et al. Identifying epitopes responsible for neutralizing antibody and DC-SIGN binding on the spike glycoprotein of the severe acute respiratory syndrome coronavirus. *J Virol* 2006;80(21):10315–24.
- [22] York IA, Stevens J, Alymova IV. Influenza virus N-linked glycosylation and innate immunity. *Biosci Rep* 2018;39(1):BSR20171505.
- [23] Zheng J, Yamada Y, Fung TS, Huang M, Chia R, Liu DX. 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;513:65–74.
- [24] Zhou Y, Lu K, Pfefferle S, Bertram S, Glowacka I, Drosten C, et al. A single asparagine-linked glycosylation site of the severe acute respiratory syndrome coronavirus spike glycoprotein facilitates inhibition by mannose-binding lectin through multiple mechanisms. *J Virol* 2010;84(17):8753–64.
- [25] Han DP, Lohani M, Cho MW. 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):12029–39.
- [26] Jeffers SA, Tüsell SM, Gillim-Ross L, Hemmila EM, Achenbach JE, Babcock GJ, et al. CD209L (L-SIGN) is a receptor for severe acute respiratory syndrome coronavirus. *PNAS* 2004;101(44):15748–53.
- [27] Zheng L, Li H, Fu L, Liu S, Yan Q, Leng SX. Blocking cellular N-glycosylation suppresses human cytomegalovirus entry in human fibroblasts. *Microb Pathog* 2020;138:103776.
- [28] Vincent MJ, Bergeron E, Benjannet S, Erickson BR, Rollin PE, Ksiazek TG, et al. Chloroquine is a potent inhibitor of SARS coronavirus infection and spread. *Virology* 2005;269:69.
- [29] Chen WH, Du L, Chag SM, Ma C, Tricoche N, Tao X, et al. Yeast-expressed recombinant protein of the receptor-binding domain in SARS-CoV spike protein with deglycosylated forms as a SARS vaccine candidate. *Hum Vaccines Immunother* 2014;10(3):648–58.
- [30] Kumar S, Maurya VK, Prasad AK, Bhatt MLB, Saxena SK. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *VirusDisease* 2020;31:13–21.
- [31] Sun Z, Liu X, Jiang J, Huang H, Wang J, Wu D, et al. Toward biomarker development in large clinical cohorts: an integrated high-throughput 96-wellplate-based sample preparation workflow for versatile downstream proteomic analyses. *Anal Chem* 2016;88(17):8518–25.
- [32] Liu MQ, Zeng WF, Fang P, Cao WQ, Liu C, Yan GQ, et al. pGlyco 2.0 enables precision N-glycoproteomics with comprehensive quality control and one-step mass spectrometry for intact glycopeptide identification.

- Nat Commun 2017;8:438.
- [33] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr* 2010;D66:486–501.
- [34] Bailey UM, Jamaluddin MF, Schulz BL. Analysis of congenital disorder of glycosylation-Iid in a yeast model system shows diverse site-specific underglycosylation of glycoproteins. *J Proteome Res* 2012;11:5376–83.
- [35] Shajahan A, Supekar NT, Gleinich AS, Azadi P. Deducing the N- and Oglycosylation profile of the spike protein of novel coronavirus SARS-CoV-2. *Glycobiology* 2020;30(12):981–8.
- [36] Watanabe Y, Allen JD, Wrapp D, McLellan JS, Crispin M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science* 2020;369(6501):330–3.
- [37] Krokhin O, Li Y, Andonov A, Feldmann H, Flick R, Jones S, et al. Mass spectrometric characterization of proteins from the SARS virus: a preliminary report. *Mol Cell Proteomics* 2003;2(5):346–56.
- [38] Ying W, Hao Y, Zhang Y, Peng W, Qin E, Cai Y, et al. Proteomic analysis on structural proteins of severe acute respiratory syndrome coronavirus. *Proteomics* 2004;4(2):492–504.
- [39] Song HC, Seo MY, Stadler K, Yoo BJ, Choo QL, Coates SR, et al. Synthesis and characterization of a native, oligomeric form of recombinant severe acute respiratory syndrome coronavirus spike glycoprotein. *J Virol* 2004;78(19):10328–35.
- [40] Watanabe Y, Bowden TA, Wilson IA, Crispin M. Exploitation of glycosylation in enveloped virus pathobiology. *Biochim Biophys Acta Gen Subj* 2019;1863(10):1480–97.
- [41] Feldmann H, Nichol ST, Klenk HD, Peters CJ, Sanchez A. Characterization of filoviruses based on differences in structure and antigenicity of the virion glycoprotein. *Virology* 1994;199(2):469–73.
- [42] Ritchie G, Harvey DJ, Feldmann F, Stroehner U, Feldmann H, Royle L, et al. Identification of N-linked carbohydrates from severe acute respiratory syndrome (SARS) spike glycoprotein. *Virology* 2010;399(2):257–69.
- [43] Yang TJ, Chang YC, Ko TP, Draczkowski P, Chien YC, Chang YC, et al. Cryo-EM analysis of a feline coronavirus spike protein reveals a unique structure and camouflaging glycans. *PNAS* 2020;117(3):1438–46.
- [44] Vigerust DJ, Shepherd VL. Virus glycosylation: role in virulence and immune interactions. *Trends Microbiol* 2007;15(5):211–8.
- [45] Biswas A, Bhattacharjee U, Chakrabarti AK, Tewari DN, Banu H, Dutta S. Emergence of novel coronavirus and COVID-19: whether to stay or die out? *Crit Rev Microbiol* 2020;46(2):182–93.
- [46] Robinson BS, Arthur CM, Evavold B, Roback E, Kamili NA, Stowell CS, et al. The sweet-side of leukocytes: galectins as master regulators of neutrophil function. *Front Immunol* 2019;53(6):925–35.
- [47] Wang WH, Lin CY, Chang MR, Urbina AN, Assavalapsakul W, Thitithanyanont A, et al. The role of galectins in virus infection—a systemic literature review. *J Microbiol Immunol Infect* 2019;53(6):925–35.
- [48] Cheng Y, Cheng G, Chui CH, Lau FY, Chan PK, Ng MH, et al. ABO blood group and susceptibility to severe acute respiratory syndrome. *JAMA* 2005;293(12):1450–1.
- [49] Zhao J, Yang Y, Huang H, Li D, Gu D, Lu X, et al. Relationship between the ABO blood group and the COVID-19 susceptibility. *Clin Infect Dis* 2020;73(2):328–31.
- [50] Guillon P, Clément M, Sébille V, Rivain JG, Chou CF, Ruvoën-Clouet N, et al. Inhibition of the interaction between the SARS-CoV spike protein and its cellular receptor by anti-histo-blood group antibodies. *Glycobiology* 2008;18(12):1085–93.
- [51] Zhao X, Guo F, Comunale MA, Mehta A, Sehgal M, Jain P, et al. Inhibition of endoplasmic reticulum-resident glucosidases impairs severe acute respiratory syndrome coronavirus and human coronavirus NL63 spike protein-mediated entry by altering the glycan processing of angiotensin I-converting enzyme 2. *Antimicrob Agents Chemother* 2015;59(1):206–16.
- [52] Liu J, Cao R, Xu M, Wang X, Zhang H, Hu H, et al. Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. *Cell Discovery* 2020;6:16.