



甲型流感病毒神经氨酸酶抗原蛋白 表位与抗体的结构空间识别*

朱 拯^{1,2)} 陈郑珊¹⁾ 张冠英¹⁾ 房 婷¹⁾ 樊 璞¹⁾ 毕 磊¹⁾ 崔 越¹⁾ 李泽亚¹⁾

苏纯懿²⁾ 迟象阳^{1)**} 于长明^{1)**}

(¹⁾ 军事科学院军事医学研究院, 北京 100071; ²⁾ 联勤保障部队第九六一医院卫勤处, 齐齐哈尔 161000)

摘要 **目的** 基于甲型流感病毒神经氨酸酶 (neuraminidase, NA) 蛋白抗原抗体复合物结构数据, 探究抗原蛋白表位与抗体的空间结构识别关系。**方法** 完整收集 SAbDab 数据库中 NA 蛋白抗原抗体复合物结构数据, 对蛋白质结构进行处理分析, 获取抗原表位和相应抗体配位氨基酸序列和空间分布信息, 统计分析结合抗体序列、基因使用频率、氨基酸使用偏好性和互补决定区 (CDR) 氨基酸长度; 分析抗体结合的热点区域, 进一步对抗体配位空间结构相似性进行计算和层次聚类, 深入探究抗原表位与抗体的空间识别关系, 通过生物膜干涉 (BLI) 实验验证靶向不同抗原表位的抗体的识别特异性。**结果** SAbDab 数据库中甲型流感病毒 NA 蛋白的抗原抗体复合物结构数据以 H3N2、H7N9、H1N1 亚型为主, 抗原表位的热点位置主要集中在催化活性位点区域, 抗体物种来源以人源和鼠源为主, 鼠源抗体 VJ 基因组合使用频率最高的是 IGHV1-12*01/IGHJ2*01, 人源抗体 VJ 基因组合使用频率最高的是 IGHV1-69*01/IGHJ6*01。结合在催化活性位点区域内与结合在催化活性位点区域以外其他区域的抗体, 在重链 CDR 氨基酸长度与氨基酸使用偏好性方面存在显著差异。抗原表位与抗体配位具有特定的识别关系, 部分相似的抗体配位结构特异性地识别相同的表位, 可在结构上观察到结合区域的重叠, BLI 实验也验证了这类抗体对表位的竞争结合。**结论** NA 蛋白抗原表位点分布以催化活性位点及周围环状区域为主, 空间构象互补与静电相互作用在抗体与抗原的催化活性区域识别和结合过程中发挥重要作用。抗原与抗体存在不依赖于抗体唯一序列的空间识别关系, 不同序列的抗体也可能形成识别同一抗原表位的局部空间结构。

关键词 甲型流感病毒, 神经氨酸酶, 抗原表位, 抗体, 抗体配位相似性

中图分类号 R373.1, Q511, Q71

DOI: 10.16476/j.pibb.2024.0309

CSTR: 32369.14.pibb.20240309

甲型流感病毒 (influenza A virus) 是一种单链 RNA 病毒, 属于正黏液病毒科。具有高度传染性和变异性, 分类复杂、亚型众多、容易突变, 根据世界卫生组织报道, 流感季节性流行每年可导致 300 万~500 万重症病例, 30 万~65 万呼吸道疾病相关死亡^[1]。20 世纪以来, 全球范围内共出现了 4 次流感大流行: 1918 年西班牙 H1N1 流感、1957 年亚洲 H2N2 流感、1968 年香港 H3N2 流感和 2009 年甲型 H1N1 流感, 均造成了大量的人员死亡。不可预测的大流行以及不断爆发的季节性流感、禽流感, 对人类健康及社会稳定造成巨大威胁^[1]。

甲型流感病毒的众多编码蛋白中, 血凝素 (hemagglutinin, HA) 和神经酰胺酶 (neuraminidase, NA) 蛋白是当前疫苗及药物研发

的主要靶标蛋白, HA 蛋白主要参与病毒对宿主细胞的吸附和侵入过程^[2], NA 蛋白主要在病毒的释放和传播的过程中发挥作用^[3]。NA 蛋白能够特异性识别、催化切割唾液酸, 阻止病毒颗粒在细胞表面相互聚集, 便于病毒颗粒脱离宿主细胞并继续扩散感染^[4]。NA 蛋白上的催化活性位点在其功能发挥中起着核心作用, 是奥司他韦、扎那米韦等抗甲流药物的主要靶点^[5]。NA 蛋白的催化活性位点位于蛋白质表面的口袋区域, 主要由带电荷的氨基酸

* 北京市科技新星计划 (20240484733) 资助项目。

** 通讯联系人。

于长明 Tel: 010-66948876, E-mail: yuchangming@126.com

迟象阳 Tel: 010-66948563, E-mail: xiangyangchi@163.com

收稿日期: 2024-07-10, 接受日期: 2024-12-03

残基组成^[6], 研究表明, 催化活性位点在所有流感病毒株的NA蛋白中均高度保守^[5, 7]。与HA蛋白相比, NA蛋白的整体结构在甲型流感病毒的不同亚型中非常保守, 且抗原变异发生速度明显缓慢、抗原漂移发生率低^[8]。

疫苗和抗体是控制疫情大规模传播的重要武器。当前, 流感疫苗的开发主要集中在HA蛋白上, 疫苗激发机体产生HA特异性保护抗体^[9-10]。近年研究表明, NA抗体反应具有与HA抗体反应不同的免疫保护机制。鉴于目前的季节性流感疫苗提供的保护不够理想, 作为潜在的候选抗原, NA蛋白正受到越来越多的关注。针对甲型流感病毒, 靶向NA的抗体研究也取得了很多重要进展, 已有研究表明, 靶向NA的抗体能够抑制流感病毒复制和传播, 降低病毒对哺乳动物的感染力, 并且能够缓解流感临床症状和缩短病毒复制时间^[11]。此外, Johansson等^[2, 12]研究表明, 人体内诱导产生的NA特异性抗体具有广泛交叉反应性和广谱保护性, 与保护机体免受同源或异源病毒感染、降低病毒滴度以及减轻临床症状相关。Liu等^[13]研究发现, 许多NA特异性抗体可以结合到同一NA亚型中的抗原保守表位上, 保护机体免受异源病毒侵袭, 表明基于NA设计的NA结合抗体和NA抑制剂具有保护作用 and 巨大的应用潜力^[14-16]。HA的抗原变异性和免疫原性已得到充分证明, 而NA的免疫原性以及与其保护性抗体的相互作用有待更加充分的研究^[15, 17]。

抗原表位是指抗原分子中能被抗体或免疫细胞识别并与其结合的区域, 抗体配位是指抗体与抗原表位结合的区域, 抗体通过轻链和重链的可变区与抗原表位相互作用, 形成抗原-抗体复合物。随着抗原抗体复合物实验解析结构的积累, 在空间结构上理解NA抗原蛋白与抗体识别和相互作用成为可能^[18-20]。但目前尚未有系统性的利用基于复合物三维结构研究NA抗原表位与抗体空间识别关系的报道。因此, 本研究旨在通过对NA抗原蛋白与抗体复合物结构的全面分析, 揭示NA抗原蛋白表位分布及不同表位与抗体结合的特征, 探究甲型流感病毒NA抗原蛋白表位与抗体配位的空间识别关系。

1 材料与方法

1.1 数据收集

截止2024年3月, 从SAbDab数据库^[21-22] (<http://opig.stats.ox.ac.uk/webapps/newsabdab/sabdab/>) 下载

抗原抗体复合物结构数据, 按照以下条件筛选:

a. 抗原类型为 Influenza A virus; b. 抗原名称为 Neuraminidase; c. 复合物结构数据同时包含抗原氨基酸链和抗体氨基酸链; d. 抗体结构同时包含重链和轻链。对于缺失流感病毒亚型标注的数据, 通过 RCSB PDB 数据库 (<https://www.rcsb.org/>) 检索 PDB accession code (PDB identification code) 补充 NA 蛋白所属病毒亚型。

1.2 抗原抗体复合物蛋白质结构处理

在 Python 3.8.13 中调用 BioPython v.1.83^[23-24] 提取蛋白质序列, 调用 pymol v.2.6.0 库的 cmd 包^[25], 去除复合物结构中的抗原抗体重复的氨基酸链、金属离子、小分子化合物以及溶剂分子, 获得表位和配位氨基酸残基位点编号和氨基酸类型。在本研究中, 定义表位 (epitope) 为抗原结构表面与抗体结构原子距离 ≤ 3.5 Å 的区域, 配位 (paratope) 为抗体结构表面与相抗原结构原子距离 ≤ 3.5 Å 的区域^[20, 26-28]。

1.3 抗体特征计算

使用 ANARCI 工具^[29] (<http://opig.stats.ox.ac.uk/webapps/newsabdab/sabpred/anarci/>) 下载的抗原抗体复合物中的抗体结构, 根据 IMGT 标准化编码规则^[30] 进行编号 (CDR1: 27~38; CDR2: 56~65; CDR3: 105~117)。对抗体序列, 使用 Abalign 软件^[31] 的 “Align” 功能对抗体序列进行多重序比对和对齐, 用 “Gene Usage” 工具注释抗体基因并统计基因使用频率, 用 “Length Distribution” 工具统计抗体互补决定区 (complementarity determining region, CDR) 氨基酸长度, 用 “Mutation Profile” 工具统计抗体各位点上 20 种氨基酸类型的使用频数。

1.4 抗体配位空间结构相似性计算、序列相似性计算和层次聚类

使用 Ab-Ligity 工具^[32] (<https://opig.stats.ox.ac.uk/resources>) 计算抗体配位之间的相似性, 得到抗体配位相似性矩阵。使用 Abalign 命令行工具^[31] (<http://cao.labshare.cn/abalign/>) 计算抗体序列之间的相似性, 得到抗体序列相似性矩阵。利用 R 中的 dist (method= “Euclidean”) 和 hclust (method= “average”) 函数^[33] 分别对抗体配位相似性矩阵与抗体序列相似性矩阵进行层次聚类, 得到相似性聚类树。

1.5 甲型流感病毒NA蛋白与抗体的制备

甲型流感病毒 H3N2 NA 蛋白 (A/Darwin/9/

2021, 货号: 40860-V08B) 和 H7N9 NA 蛋白 (A/Shanghai/1/2013, 货号: 40109-V07H) 购自北京义翘神州科技有限公司。

甲型流感病毒 NA 抗体轻、重链可变区序列从相应 PDB 结构文件中提取, 由江苏赛索飞生物科技有限公司合成表达质粒。使用 ExpiFectamine™293 转染试剂将抗体轻、重链质粒转染至 Expi293F 细胞中, 培养后收取细胞上清, 经过 rProtein A 亲和层析柱纯化, 制备得到 Ig 抗体。

1.6 生物膜干涉 (bio-layer interferometry, BLI) 法测定抗体的竞争结合

对 NA 蛋白使用 EZ-Link™ Sulfo-NHS-LC-LC-生物素 (Thermo Scientific) 进行标记, 并且在含有 0.2% 牛血清白蛋白 (bovine serum albumin, BSA) 和 0.02% Tween-20 的 PBS 缓冲液中稀释至适当浓度 (如 30 mg/L、60 mg/L)。将稀释的标记抗原按每孔 250 μl 加入 96 孔板中, 在 Gator™ 非标记生物分子分析仪上, 按 400 r/min, 60 s 的条件固定在链霉亲和素生物传感器 (SA 探针) 上。固定抗原后的生物传感器用缓冲溶液洗涤 60 s, 浸入含有初始抗体的分析孔并孵育 180 s, 使生物传感器上的抗原与初始抗体结合, 而后将生物传感器移至竞争抗体的分析孔并孵育 180 s, 以进行竞争抗体的结合。抗体的结合曲线数据由 Gator 集成软件 (Version 2.10.4.0713) 导出。

1.7 数据统计学分析及可视化

在 Python 3.8.13 中利用 Pandas v. 1.4.2、Numpy v.1.22.3 库进行数值计算和统计分析; 蛋白质三维结构数据可视化使用 ChimeraX 1.2.4 软件; 聚类树可视化和图形标注使用 ChiPlot 网站 (<https://www.chiplot.online/>) 完成; 抗体重链 CDR 区长度的组间比较使用 Graphpad Pism 9.5 完成, 统计学分析使用双因素方差分析 (two-way analysis of variance); 序列 Weblogo 图的绘制使用 WEBLOGO 网站 (<http://weblogo.berkeley.edu/logo.cgi>) 完成; 序列多重比对结果的可视化使用 ESPript 3.0 网站 (<https://esprict.ibcp.fr/ESPript/ESPript/>) 完成。

2 结 果

2.1 甲型流感病毒 NA 抗原蛋白与相应抗体的统计

总共获得完整有效的甲型流感病毒 NA 抗原抗体复合物 37 个。复合物结构中 NA 抗原蛋白所属流感病毒亚型以 H3N2、H7N9、H1N1 为主 (图 1a), N1 型 NA 蛋白占比 16.22%、N2 型 NA 蛋白占比 40.54%、N9 型 NA 蛋白占比 43.24% (图 1b)。37 个抗原抗体复合物中的抗体物种来源以人源和鼠源为主, 人源抗体 25 个, 占比 67.57%, 鼠源抗体 12 个, 占比 32.43% (图 1c)。完整的抗原抗体复合物信息在 Github 提供: https://github.com/shanshangg/IAV_neuraminidase_Ab_epitope_paratope_mapping。

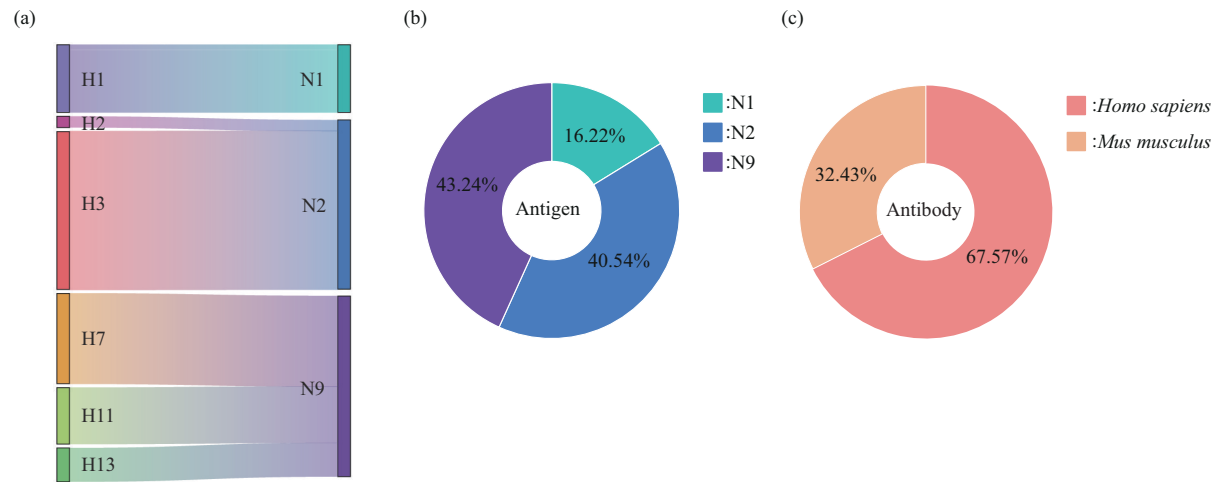


Fig. 1 Influenza A virus NA protein and bound antibodies in complex structures

(a) Influenza A virus subtypes; (b) frequency of N1, N2 and N9 NA proteins; (c) frequency of human antibodies and murine antibodies.

2.2 甲型流感病毒 NA 抗原蛋白表位位点分布分析

抗原表位是免疫识别的关键, 为了探究不同病

毒亚型 NA 蛋白上的表位分布与表位氨基酸类型的异同。本研究统计了全部 37 个抗原抗体复合物上

的表位位点分布,定位出NA抗原蛋白上激发机体产生中和抗体的关键区域。图2列出高频表位氨基酸分布,其中黄色标记的氨基酸位置(residue ID)为催化活性位点区域、蓝色标记催化活性位点以外的其他区域。在经过多重比对和对齐的序列中,使用红色标记出抗原表位,从左往右依次是N1(绿)、N2(蓝)、N9(紫)亚型的NA抗原表位。在NA蛋白上R118、R119、D151、R152、R156、W179、S180、D/N198、I222、R224、E228、D244、H274、E276、E277、R292、N294、D331、K351、R372、Y407、E425等氨基酸直接参与唾液酸结合,形成催化活性位点区域^[6, 8, 17]。统计结果

表明,表位分布的前10位热点氨基酸位置为152、118、151、348、372、407、345、179、200、228位氨基酸,其中7个(R152、R118、D151、R372、Y407、W179、E228)分布于催化活性位点区域,并且7个位点在N1、N2、N9型NA蛋白中氨基酸类型均相同,是保守的表位区域,也提示了结合该区域的NA抗体具有广谱结合的潜力。考虑到NA上的表位主要是构象表位,将表位分布频率在NA蛋白三维结构上进行可视化,高频出现的表位位点呈现更深的颜色,抗体结合位点集中分布于催化活性区域的口袋中,或分散分布于周围环状区(图3)。

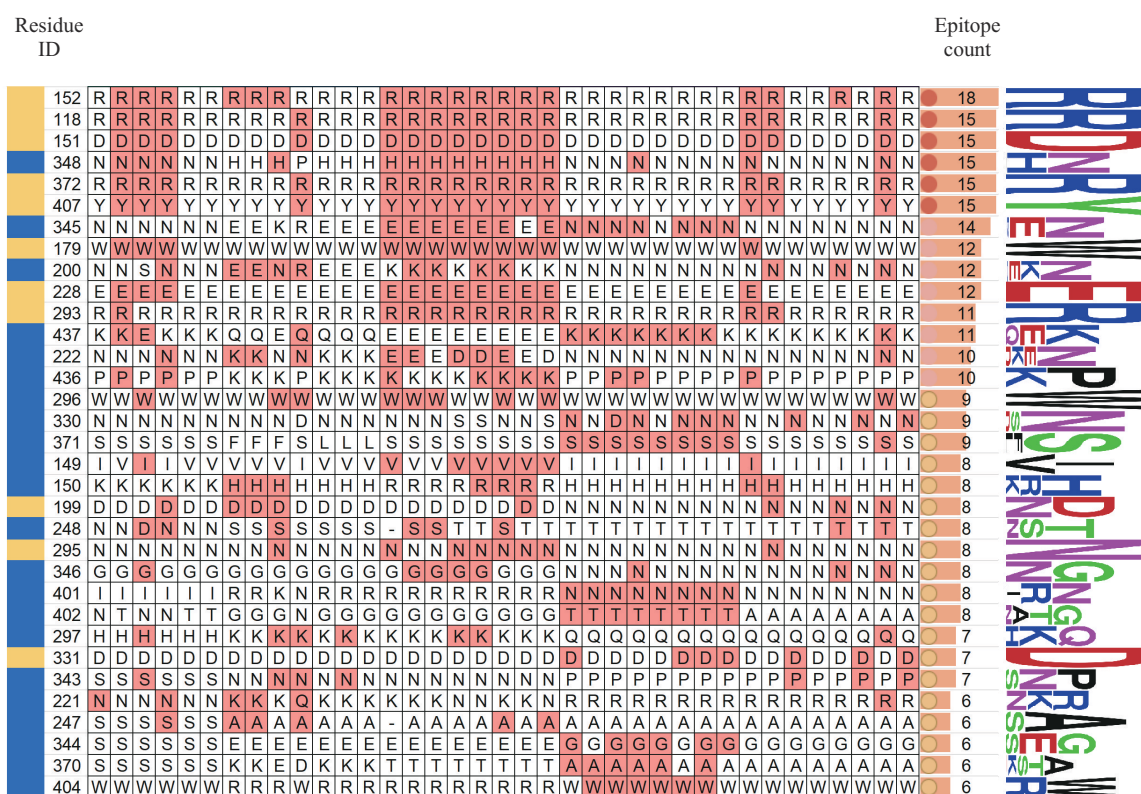


Fig. 2 Epitope distribution on the sequence of influenza A virus NA protein

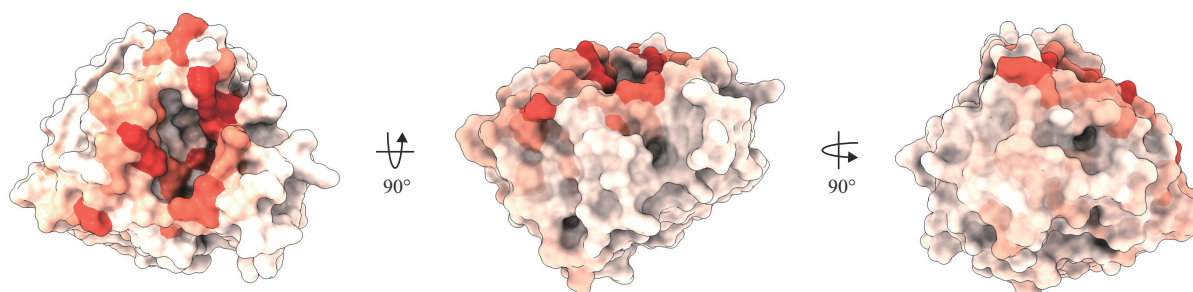


Fig. 3 Epitope distribution on the three-dimensional structure of influenza A virus NA protein

根据对每个抗原抗体复合物结构中表位的提取结果, 43.24% (16/37) 复合物的抗原表位主要分布在催化活性位点区域, 56.76% (21/37) 复合物的抗原表位主要分布在催化活性位点以外的其他区域。

2.3 甲型流感病毒NA抗体特征分析

抗体的 V (D) J 基因的多样性和基因重排是抗体多样性形成的重要过程之一, 本研究对结合 NA 蛋白不同表位区域的抗体进行了基因注释, 探

究不同表位的抗体基因使用和基因组合的偏好性。在 37 个复合物结构中具有 30 个可变区序列不同的抗体 (但是在复合物结构中序列相同的抗体所结合的抗原不同, 所以仍保留 37 个复合物结构), 考虑到鼠源和人源抗体基因的差异, 区分鼠源和人源抗体分别统计基因使用情况 (图 4)。在 10 个鼠源抗体中, 重链 V 基因使用频率最高的是 IGHV1-12*01, J 基因使用频率最高的是 IGHJ2*01, VJ 基因组合主要是 IGHV1-12*01/ IGHJ2*01 (图 4a)。在 20 个

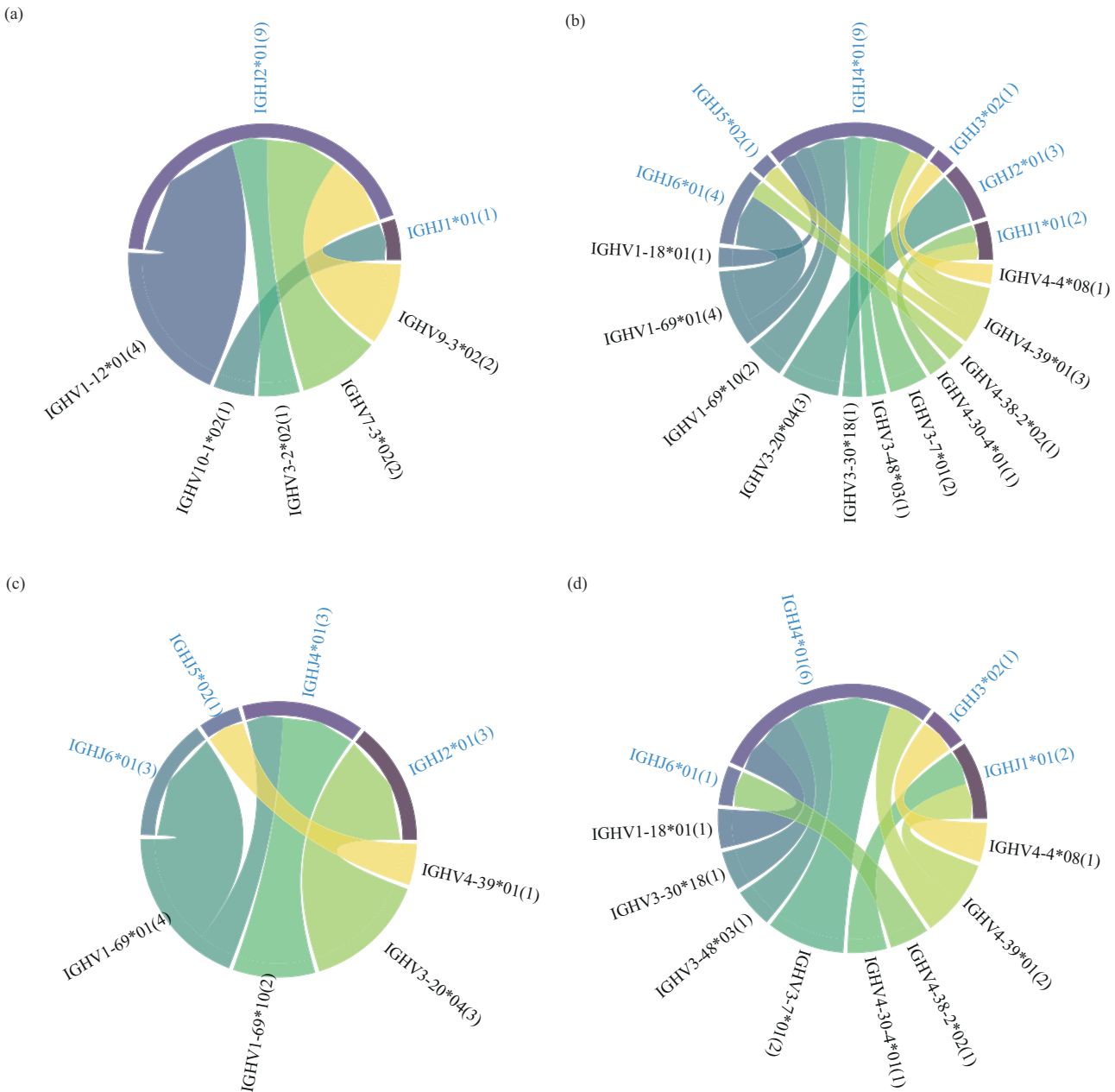


Fig. 4 V-J gene combinations in the heavy chain of antibodies against influenza A virus NA protein
(a) V-J gene combinations in murine antibodies; (b) V-J gene combinations in human antibodies; (c) V-J gene combinations in human antibodies binding to the catalytic active site region of NA protein; (d) V-J gene combinations in human antibodies binding to other regions outside the catalytic active site of NA protein.

人源抗体中, 重链 V 基因使用频率最高的是 IGHV1-69*01, J 基因使用频率最高的是 IGHJ4*01, IGHV1-69*01/IGHJ6*01 是出现频率最高的 V 基因和 J 基因组合 (图 4b)。

20 个人源抗体中, 含 10 个人源抗体其对应的抗原表位分布在 NA 蛋白的催化活性位点区域, 使用的 V 基因主要是 IGHV1-69*01, J 基因主要是 IGHJ4*01、IGHJ2*01、IGHJ6*01, VJ 基因组合主要是 IGHV1-69*01/IGHJ6*01 (图 4c), 分布在 NA 蛋白催化活性位点以外的其他区域的人源抗体 10 个, J 基因使用偏好性与催化活性位点区域相似、V 基因主要是 IGHV3-7*01、IGHV4-39*01, VJ 基因组合主要是 IGHV3-7*01/IGHJ4*01 (图 4d)。

抗体重链 CDR 结构在抗原抗体结合过程中通常起着重要作用, 尤其是 CDR3 区的三维结构呈现出一个环状或类似“V”型的构象, 有利于与抗原表位的空间匹配。本研究对重链 CDR 氨基酸序列长度的统计分析, 结果显示, 结合在催化活性位点区域的抗体与结合催化活性位点以外的其他区域的抗体, 在重链 CDR1、CDR2 的长度上均无显著性差异。然而结合在催化活性位点区域的抗体的重链 CDR3 区长度, 显著大于结合在催化活性位点以外的其他区域的抗体 CDR3 区长度 ($P < 0.001$) (图 5a)。提示催化活性位点区域为口袋形状, 较长的 CDR3 区有利于抗体结合区域形成“突出”的结构, 与抗原适配互补, 使得抗体能够特异识别和结合 NA 蛋白。

为进一步探究抗原与抗体识别和结合的特点, 本研究对结合 NA 蛋白不同区域的抗体氨基酸组成进行分类统计, 结果显示, 结合催化活性位点区域的抗体在 CDR3 区的第 106 位高频使用精氨酸 (Arg, R), 在第 107、110、116 位高频使用天冬氨酸 (Asp, D), 在第 112、114 位高频使用酪氨酸 (Tyr, Y) (图 5c)。结合在催化活性位点以外的其他区域的抗体在 CDR3 区末端的第 107 位高频使用丝氨酸 (Ser, S), 第 115 位高频使用苯丙氨酸 (Phe, F), 第 117 位高频使用酪氨酸 (Tyr, Y) (图 5d)。更具体的, 对抗体相互作用的抗原配位结构上的氨基酸使用情况进行统计, 结果表明, 结合催化活性位点区域的抗体主要由精氨酸、谷氨酸、天冬氨酸等极性带电荷氨基酸组成, 结合催化活性位点区域以外的其他区域的抗体主要由酪氨酸、丝氨酸、天冬酰胺、苏氨酸等极性不带电荷氨

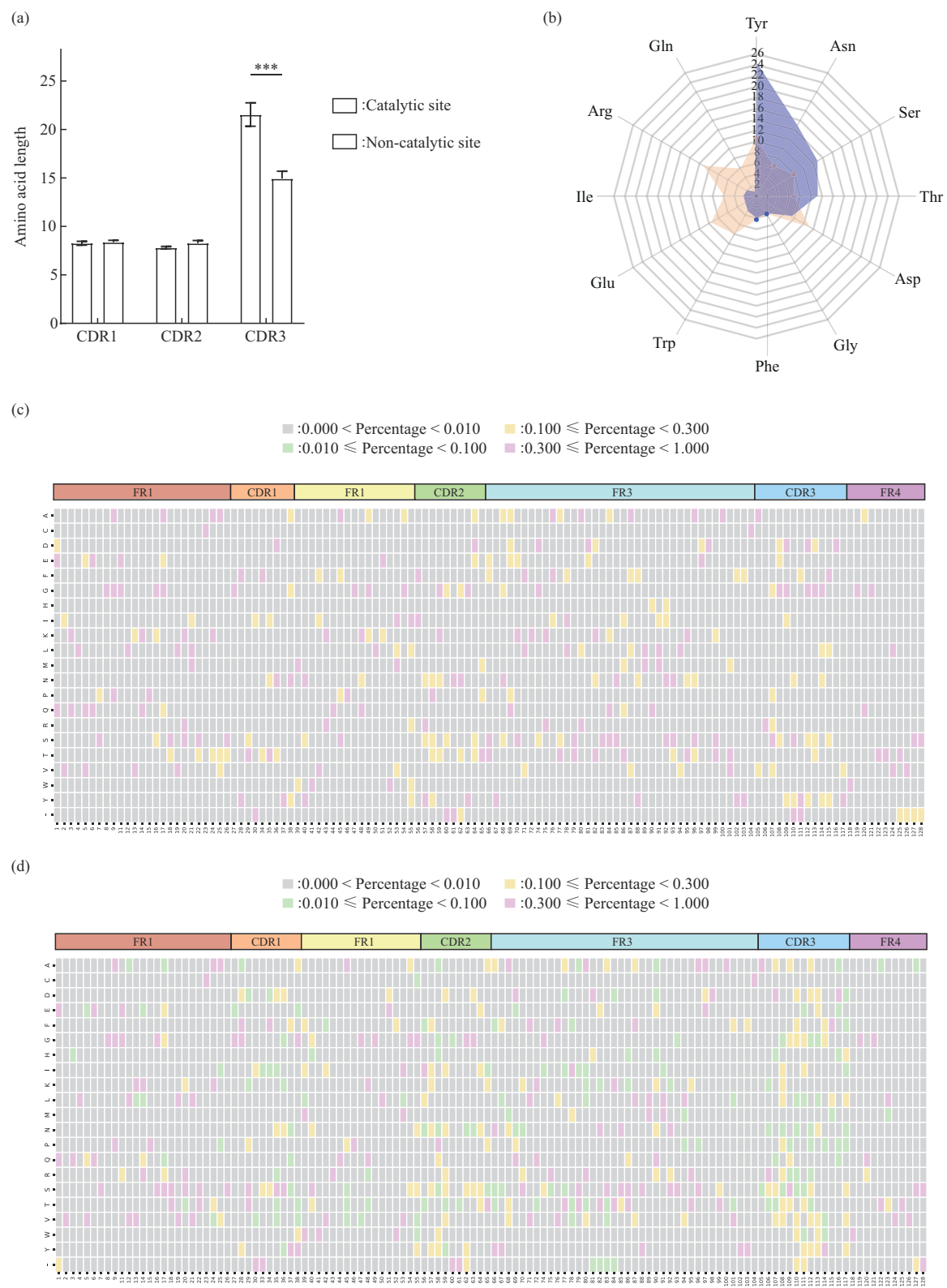
基酸组成 (图 5b)。

NA 蛋白催化活性位点区域主要由精氨酸、谷氨酸、天冬氨酸等极性带电荷氨基酸组成^[6, 8, 17], 结合催化活性位点区域的抗体在 CDR3 区结构上高频使用极性带电荷的精氨酸和天冬氨酸。结果提示了抗体结合位点氨基酸的电性与催化活性位点区域的电荷分布存在潜在的对应关系, 静电相互作用可能在抗体与抗原的催化活性区域识别和结合过程中发挥重要作用。

2.4 甲型流感病毒 NA 抗体配位空间结构相似性识别关系分析

为进一步比较结合在催化活性位点区域和以外的其他区域的抗体配位结构特点、抗体对抗原的空间识别关系, 本研究对提取的抗原抗体复合物中抗体上的配位结构, 利用 Ab-Ligity 工具对抗体配位结构进行表征和表征的相似性计算, 得到相似性矩阵并作层次聚类, 分析抗体配位相似性与表位分布区域存在的对应关系 (图 6a)。同时利用 Abalign 命令行工具 (MSA) 计算得到抗体序列之间的相似性 (图 6b)。相似序列的抗体在结合不同亚型 NA 蛋白的复合物结构中, 具有相近的配位结构, 如 8g3m、8g3n、8g3o、8g3p、8g3q、8g3r、8g3v、8g30 位于相同的抗体配位和序列聚类簇。提示广谱抗体是以相近的配位结构结合不同的 NA 抗原蛋白的保守表位。值得注意的是, 配位结构相似性聚类还揭示了一些序列完全不同但配位结构相近的抗体 (图 6b, c), 如 6qzy 和 6lxx、6pzz 和 6pzf、8ez8 和 8ez7。分别将上述 3 组复合物结构中的抗原结构相重叠 (图 6d~f), 可见虽然抗体序列不同, 但形成了相似的空间构象与对应的抗原表位结构相适配。如结合 NA 蛋白的催化活性位点区域的 2 株抗体, 具有“长而突出”的 loop 结构 (图 6d), 结合催化活性位点以外的其他区域的 4 株抗体, 则形成相对“平坦”的接触面 (图 6e, f)。

为了通过抗体竞争结合实验验证以上序列不同但配位结构相近的抗体具有相同的表位结合特异性, 对复合物结构中包含的 6 株抗体进行合成和表达: 抗体 1F04 (提取自 8ez7 结构)、3C08 (8ez8)、NA63 (6pzf)、NA80 (6pzz)、Z2B3 (6lxx)、NA73 (6pzy)。通过 BLI 法测定出抗体 1F04 和 3C08 对 A/Darwin/9/2021 H3N2 NA 蛋白、抗体 NA63 和 NA80 对 A/Shanghai/1/2013 H7N9 NA 蛋白的结合竞争曲线 (抗体 Z2B3 和 NA73 对以上 NA 蛋白无结合活性)。BLI 的实验结果验证了 1F04 和



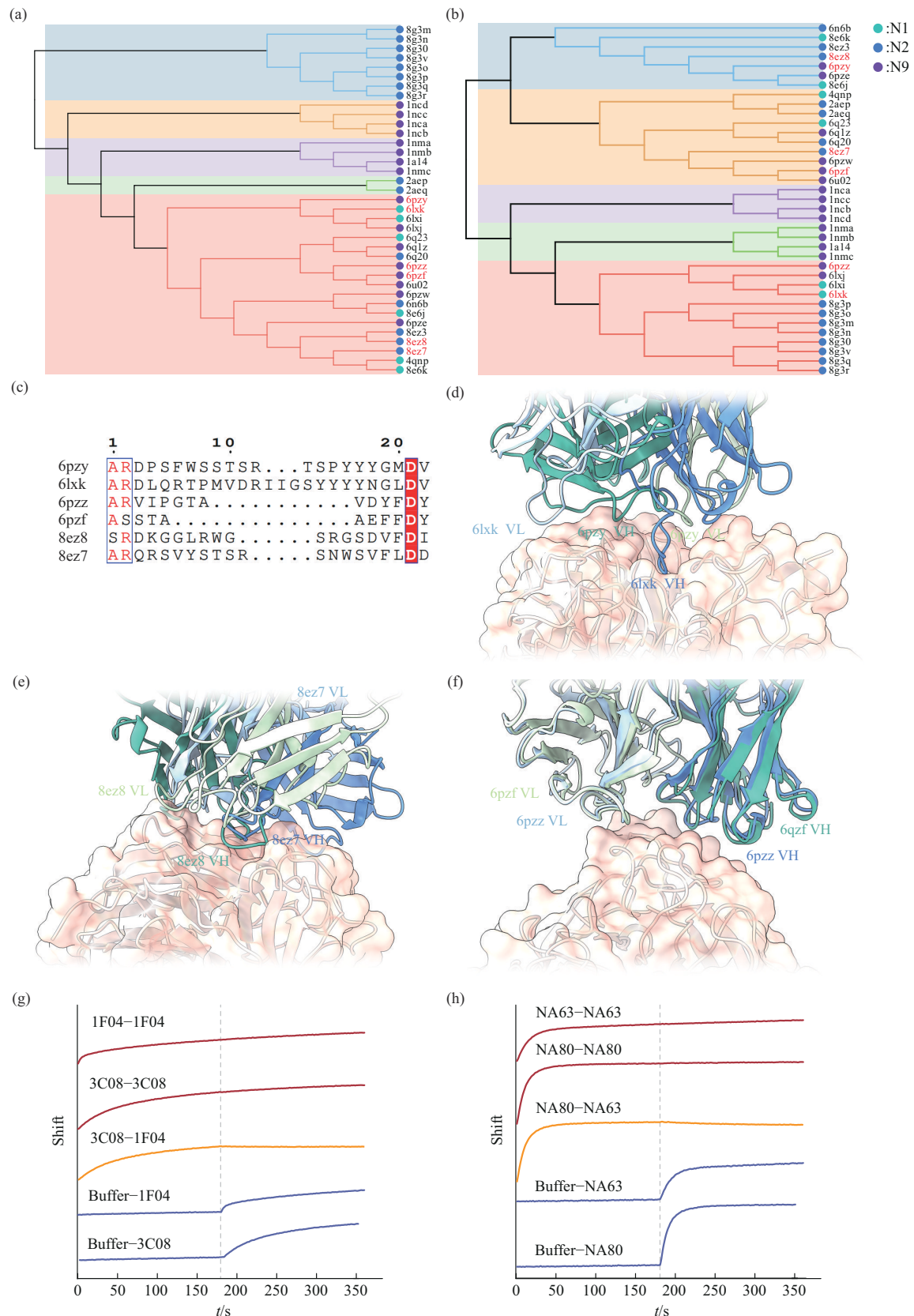


Fig. 6 Spatial recognition of antibody paratope structures targeting the influenza A virus NA protein

(a) Similarity of the antibody paratope structures. (b) Similarity of the antibody sequences. (c) Sequences of CDR3 region in the representative antibodies. Superposition of antigen-antibody complex structures: (d) 6lzk.pdb vs. 6lzk.pdb; (e) 8ez7.pdb vs. 8ez8.pdb; (f) 6qzz.pdb vs. 6pzf.pdb; (g) competitive binding curves of antibodies 1F04 (8ez7) and 3C08 (8ez8) against the antigen (A/Darwin/9/2021 H3N2 NA). (h) Competitive binding curves of antibodies NA63 (6pzf) and NA80 (6pzz) against the antigen (A/Shanghai/1/2013 H7N9 NA).

3C08、NA63和NA80结合相同或相近的表位, 在结合抗原时存在竞争效应(图6g, h), 表明抗体配位结构相似性与其特定表位的选择性存在密切联系。结果提示, 抗原表位与抗体配位具有特定的识别关系, 且与两者的空间结构具有密切的相关性, 不同序列的抗体也有可能形成与抗原表位相识别的局部空间结构, 从而识别相同的抗原表位。此类序列不同但配位空间结构相近的抗体无法通过序列相似性聚类的方法发现, 体现出结构相似性聚类方法的优势。同时, 不同抗体在配位结构上的相似性一定程度上反映了其结合表位的特异性。

3 讨 论

甲型流感病毒是一种具有高度传染性的呼吸道病原体, 对人类生命健康和社会经济都构成了严重威胁。神经氨酸酶对病毒从受感染细胞释放起到关键作用, 是抗甲型流感病毒药物和疫苗开发的主要靶标之一。抗原表位是免疫系统发挥作用的关键靶点, 是病毒上被免疫系统识别并引发免疫反应的特定区域。体液免疫中, 抗体对这些表位特异性地识别和结合是机体免疫发挥抗病毒感染的重要环节。基于抗原抗体复合物三维结构, 系统研究NA抗原表位结构和抗体配位结构的特征以及两者在空间结构上的识别关系, 有助于理解NA抗原与抗体的相互作用的关键因素和结构基础。

本研究基于甲型流感病毒NA蛋白抗原抗体复合物数据集, 分析NA蛋白表位分布特点、结合不同表位抗体的特点, 并对抗原表位与抗体配位的空间识别关系进行探究。甲型流感病毒NA蛋白抗原抗体复合物数据集由37个复合物结构组成, 甲型流感病毒NA抗原蛋白来源以H3N2、H7N9、H1N1流感病毒亚型为主, 抗体物种来源以人源和鼠源为主。表位氨基酸提取及统计结果表明, NA抗原表位主要分布在催化活性位点及其周围环状区域, NA抗原中的R118、D151、R152等催化活性位点区域呈现出高度保守性, 保守的表位区域对广谱抗体的筛选和表位疫苗的设计具有重要的提示意义。Momont等^[34-35]的研究也证实了靶向NA催化活性位点区域的抗体(如FNI9、1G01)可以广谱性中和所有甲型流感病毒NA亚型。

甲型流感病毒NA抗体的胚系基因使用具有偏好性, 如IGHV1-69*01/IGHJ6*01的VJ基因组合在人源抗体中出现频率最高, Jackson等^[36]针对三价灭活季节性流感疫苗(TIV)接种后人群的外周

血B细胞VDJ测序, 结果显示, IGHV3-7/IGHJ6、IGHV1-18/IGHJ4、IGHV1-69/IGHJ6的VJ基因组合在疫苗诱导增殖的B细胞中高频出现。这表明机体在感染流感病毒或接种流感疫苗后, 激发的抗体倾向于使用特定的V基因、J基因及其组合。结合NA催化活性口袋区域的抗体具有较长的CDR3区, 抗体配位上更多地出现精氨酸、谷氨酸、天冬氨酸等极性带电荷氨基酸, 这些特点与催化活性口袋的空间结构和电荷分布相适应, 提示了空间上的“锁钥式”的形状互补与静电相互作用在抗体与表位的识别和结合中发挥重要作用。

通过对抗体配位空间结构的表征和相似性分析, 发现结合在相同或相近表位上的抗体具有相似的配位结构, 即使是CDR序列不相同的抗体也存在着配位的相似性, 在蛋白质三维结构叠合比较和抗体竞争结合实验中均得到验证, 提示了潜在的基于空间结构的抗原表位与抗体配位的识别关系和结合规律。类似的, 在binder或抗体的从头设计工作中, RFdiffusion通过逐渐“去噪”蛋白质结构, 生成能够与目标蛋白质表面互补的结合物三维结构(backbone)^[37-38], 再由ProteinMPNN^[39]设计出相应的序列, 也是基于蛋白质结构上的空间识别(而不是序列配对识别)关系的应用。由于抗体的结构预测^[40]和配位预测方法^[41]已经取得较大的进展, 在实验室初步筛选抗体或B细胞受体(B cell receptor, BCR)的高通量测序之后, 通过将功能未知抗体的预测配位与已知表位中和抗体的配位进行相似性比较, 有望初步判断抗体的结合特异性或大致表位区域。基于甲型流感病毒NA蛋白抗原抗体复合物结构数据探究抗原蛋白表位与抗体的结构空间识别关系, 有助于理解特定抗原与抗体相互作用的机制, 也显示了抗原表位、抗体配位等三维结构的表征与相似性度量在免疫学基础研究中的重要作用与价值。数量有限的抗原抗体复合物仍然是制约相关研究开展的瓶颈, 虽然现有的AlphaFold2-Multimer和AlphaFold3等对抗原抗体复合物结构预测的准确性尚不足以替代实验解析结构^[42-43], 但随着蛋白质结构预测工具不断发展, 未来有可能使用预测的抗原抗体复合物结构作为相关研究的起点, 以弥补实验解析蛋白质复合物结构成本高、周期长、通量低等制约性因素。

4 结 论

甲型流感病毒每年在全球范围内引发大规模流

行病, 给公共卫生带来巨大挑战。神经氨酸酶是抗甲型流感病毒药物和疫苗开发的主要靶标之一。本研究系统地收集了甲型流感病毒神经氨酸酶蛋白抗原抗体复合物结构数据, 统计分析了抗原抗体相互作用的关键氨基酸残基, 深入解析了抗体与神经氨酸酶蛋白结合过程中的关键因素(氨基酸使用偏好、HCDR3长度和电荷匹配等), 并且通过对抗体配位三维结构的表征与相似性分析, 发现并验证了抗体对抗原表位特异性识别的结构机制。本研究揭示NA抗原表位与抗体的结构空间识别和相互作用过程中的关键因素, 为抗体与抗原特异性识别提供结构相关证据。

参 考 文 献

- [1] Iuliano A D, Roguski K M, Chang H H, *et al.* Estimates of global seasonal influenza-associated respiratory mortality: a modelling study. *Lancet*, 2018, **391**(10127): 1285-1300
- [2] Chen Y Q, Wohlbold T J, Zheng N Y, *et al.* Influenza infection in humans induces broadly cross-reactive and protective neuraminidase-reactive antibodies. *Cell*, 2018, **173**(2): 417-429. e10
- [3] Krammer F, Fouchier R A M, Eichelberger M C, *et al.* NAction! how can neuraminidase-based immunity contribute to better influenza virus vaccines?. *mBio*, 2018, **9**(2): e02332-17
- [4] McAuley J L, Gilbertson B P, Trifkovic S, *et al.* Influenza virus neuraminidase structure and functions. *Front Microbiol*, 2019, **10**: 39
- [5] Madsen A, Dai Y N, McMahon M, *et al.* Human antibodies targeting influenza B virus neuraminidase active site are broadly protective. *Immunity*, 2020, **53**(4): 852-863. e7
- [6] Xu X, Zhu X, Dwek R A, *et al.* Structural characterization of the 1918 influenza virus H1N1 neuraminidase. *J Virol*, 2008, **82**(21): 10493-10501
- [7] Yasuhara A, Yamayoshi S, Kiso M, *et al.* A broadly protective human monoclonal antibody targeting the sialidase activity of influenza A and B virus neuraminidases. *Nat Commun*, 2022, **13**(1): 6602
- [8] Sun X, Ma H, Wang X, *et al.* Broadly neutralizing antibodies to combat influenza virus infection. *Antiviral Res*, 2024, **221**: 105785
- [9] Wei C J, Crank M C, Shiver J, *et al.* Next-generation influenza vaccines: opportunities and challenges. *Nat Rev Drug Discov*, 2020, **19**(4): 239-252
- [10] McMillan C L D, Young P R, Watterson D, *et al.* The next generation of influenza vaccines: towards a universal solution. *Vaccines*, 2021, **9**(1): 26
- [11] Eichelberger M C, Morens D M, Taubenberger J K. Neuraminidase as an influenza vaccine antigen: a low hanging fruit, ready for picking to improve vaccine effectiveness. *Curr Opin Immunol*, 2018, **53**: 38-44
- [12] Johansson B E, Cox M M J. Influenza viral neuraminidase: the forgotten antigen. *Expert Rev Vaccines*, 2011, **10**(12): 1683-1695
- [13] Liu W C, Lin C Y, Tsou Y T, *et al.* Cross-reactive neuraminidase-inhibiting antibodies elicited by immunization with recombinant neuraminidase proteins of H5N1 and pandemic H1N1 influenza A viruses. *J Virol*, 2015, **89**(14): 7224-7234
- [14] Lederhofer J, Tsybovsky Y, Nguyen L, *et al.* Protective human monoclonal antibodies target conserved sites of vulnerability on the underside of influenza virus neuraminidase. *Immunity*, 2024, **57**(3): 574-586. e7
- [15] Ng S, Nachbagauer R, Balmaseda A, *et al.* Novel correlates of protection against pandemic H1N1 influenza A virus infection. *Nat Med*, 2019, **25**(6): 962-967
- [16] Memoli M J, Shaw P A, Han A, *et al.* Evaluation of anti-hemagglutinin and anti-neuraminidase antibodies as correlates of protection in an influenza A/H1N1 virus healthy human challenge model. *mBio*, 2016, **7**(2): e00417-16
- [17] Jagadeesh A, Salam A A A, Mudgal P P, *et al.* Influenza virus neuraminidase (NA): a target for antivirals and vaccines. *Arch Virol*, 2016, **161**(8): 2087-2094
- [18] Hummer A M, Abanades B, Deane C M. Advances in computational structure-based antibody design. *Curr Opin Struct Biol*, 2022, **74**: 102379
- [19] Ramaraj T, Angel T, Dratz E A, *et al.* Antigen-antibody interface properties: composition, residue interactions, and features of 53 non-redundant structures. *Biochim Biophys Acta*, 2012, **1824**(3): 520-532
- [20] Wang M, Zhu D, Zhu J, *et al.* Local and global anatomy of antibody-protein antigen recognition. *J Mol Recognit*, 2018, **31**(5): e2693
- [21] Schneider C, Raybould M I J, Deane C M. SAbDab in the age of biotherapeutics: updates including SAbDab-nano, the nanobody structure tracker. *Nucleic Acids Res*, 2022, **50**(D1): D1368-D1372
- [22] Dunbar J, Krawczyk K, Leem J, *et al.* SAbDab: the structural antibody database. *Nucleic Acids Res*, 2014, **42**(Database issue): D1140-D1146
- [23] Cock P J A, Antao T, Chang J T, *et al.* Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 2009, **25**(11): 1422-1423
- [24] Hamelryck T, Manderick B. PDB file parser and structure class implemented in Python. *Bioinformatics*, 2003, **19**(17): 2308-2310
- [25] Langenstein M, Hermjakob H, Bernal Llinares M. A decoupled, modular and scriptable architecture for tools to curate data platforms. *Bioinformatics*, 2021, **37**(20): 3693-3694
- [26] Gordon G L, Capel H L, Guloglu B, *et al.* A comparison of the binding sites of antibodies and single-domain antibodies. *Front Immunol*, 2023, **14**: 1231623
- [27] Akbar R, Robert P A, Pavlović M, *et al.* A compact vocabulary of paratope-epitope interactions enables predictability of antibody-antigen binding. *Cell Rep*, 2021, **34**(11): 108856
- [28] Madsen A V, Mejias-Gomez O, Pedersen L E, *et al.* Structural trends in antibody-antigen binding interfaces: a computational

- analysis of 1833 experimentally determined 3D structures. *Comput Struct Biotechnol J*, 2023, **23**: 199-211
- [29] Dunbar J, Deane C M. ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics*, 2016, **32**(2): 298-300
- [30] Lefranc M P, Pommié C, Ruiz M, *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev Comp Immunol*, 2003, **27**(1): 55-77
- [31] Zong F, Long C, Hu W, *et al.* Abalign: a comprehensive multiple sequence alignment platform for B-cell receptor immune repertoires. *Nucleic Acids Res*, 2023, **51**(W1): W17-W24
- [32] Wong W K, Robinson S A, Bujotzek A, *et al.* Ab-Ligity: identifying sequence-dissimilar antibodies that bind to the same epitope. *MAbs*, 2021, **13**(1): 1873478
- [33] Paradis E, Schliep K. Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 2019, **35**(3): 526-528
- [34] Stadlbauer D, Zhu X, McMahon M, *et al.* Broadly protective human antibodies that target the active site of influenza virus neuraminidase. *Science*, 2019, **366**(6464): 499-504
- [35] Momont C, Dang H V, Zatta F, *et al.* A pan-influenza antibody inhibiting neuraminidase *via* receptor mimicry. *Nature*, 2023, **618**(7965): 590-597
- [36] Jackson K J L, Liu Y, Roskin K M, *et al.* Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe*, 2014, **16**(1): 105-114
- [37] Watson J L, Juergens D, Bennett N R, *et al.* *De novo* design of protein structure and function with RFdiffusion. *Nature*, 2023, **620**(7976): 1089-1100
- [38] Bennett N R, Watson J L, Ragotte R J, *et al.* Atomically accurate *de novo* design of single-domain antibodies. *bioRxiv*, 2024. DOI: 10.1101/2024.03.14.585103
- [39] Dauparas J, Anishchenko I, Bennett N, *et al.* Robust deep learning-based protein sequence design using ProteinMPNN. *Science*, 2022, **378**(6615): 49-56
- [40] Ruffolo J A, Chu L S, Mahajan S P, *et al.* Fast, accurate antibody structure prediction from deep learning on massive set of natural antibodies. *Nat Commun*, 2023, **14**(1): 2389
- [41] Chinery L, Wahome N, Moal I, *et al.* Paragraph-antibody paratope prediction using graph neural networks with minimal feature vectors. *Bioinformatics*, 2023, **39**(1): btac732
- [42] Abramson J, Adler J, Dunger J, *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 2024, **630**(8016): 493-500
- [43] Eshak F, Goupil-Lamy A. AI driven approaches in nanobody epitope prediction: are we there yet?. *bioRxiv*, 2024. DOI: 10.1101/2024.10.07.616899

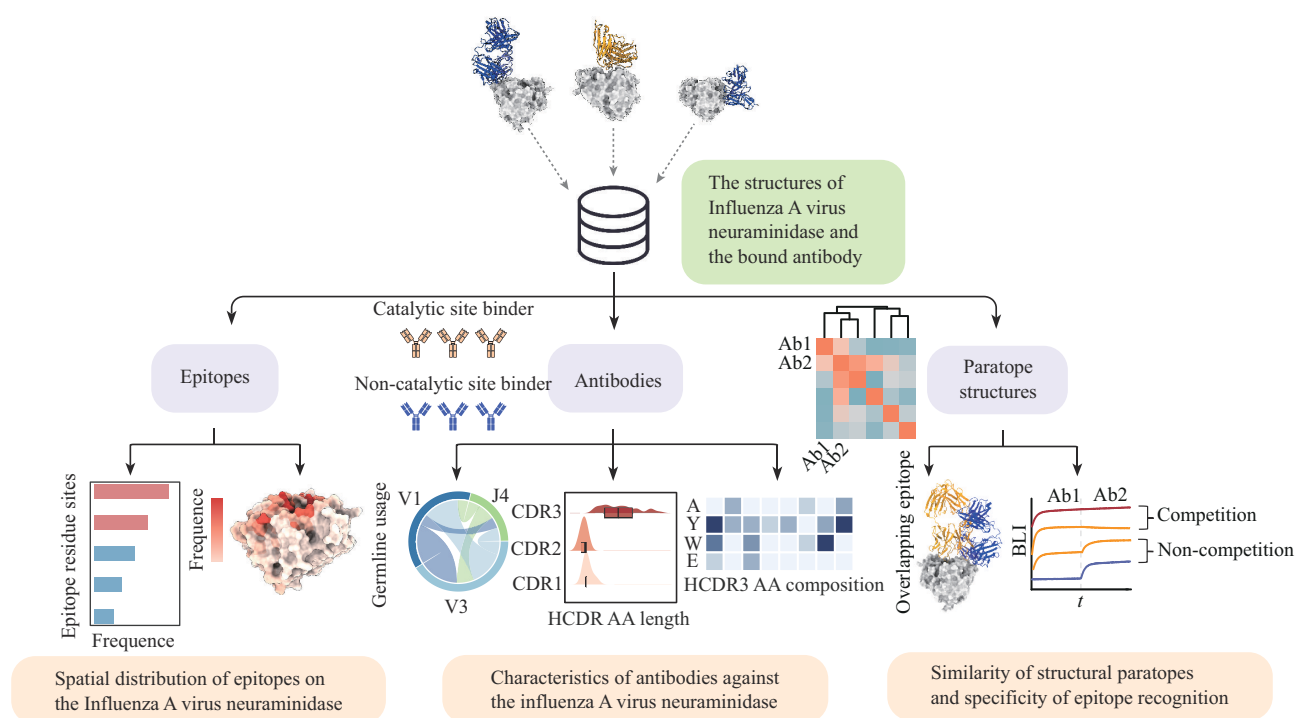
Structural and Spatial Analysis of The Recognition Relationship Between Influenza A Virus Neuraminidase Antigenic Epitopes and Antibodies*

ZHU Zheng^{1,2)}, CHEN Zheng-Shan¹⁾, ZHANG Guan-Ying¹⁾, FANG Ting¹⁾, FAN Pu¹⁾, BI Lei¹⁾, CUI Yue¹⁾, LI Ze-Ya¹⁾, SU Chun-Yi²⁾, CHI Xiang-Yang^{1)**}, YU Chang-Ming^{1)**}

⁽¹⁾Academy of Military Medical Sciences, Academy of Military Sciences, Beijing 100071, China;

⁽²⁾Department of Medical Service, The 961st Hospital of Joint Logistics Support Force, Qiqihar 161000, China)

Graphical abstract



Abstract Objective This study leverages structural data from antigen-antibody complexes of the influenza A virus neuraminidase (NA) protein to investigate the spatial recognition relationship between the antigenic epitopes and antibody paratopes. **Methods** Structural data on NA protein antigen-antibody complexes were comprehensively collected from the SAbDab database, and processed to obtain the amino acid sequences and spatial distribution information on antigenic epitopes and corresponding antibody paratopes. Statistical analysis was conducted on the antibody sequences, frequency of use of genes, amino acid preferences, and the lengths of complementarity determining regions (CDR). Epitope hotspots for antibody binding were analyzed, and the

* This work was supported by a grant from Beijing Nova Program Funding Project (20240484733).

** Corresponding author.

YU Chang-Ming. Tel: 86-10-66948876, E-mail: yuchangming@126.com

CHI Xiang-Yang. Tel: 86-10-66948563, E-mail: xiangyangchi@163.com

Received: July 10, 2024 Accepted: December 3, 2024

spatial structural similarity of antibody paratopes was calculated and subjected to clustering, which allowed for a comprehensive exploration of the spatial recognition relationship between antigenic epitopes and antibodies. The specificity of antibodies targeting different antigenic epitope clusters was further validated through bio-layer interferometry (BLI) experiments. **Results** The collected data revealed that the antigen-antibody complex structure data of influenza A virus NA protein in SAbDab database were mainly from H3N2, H7N9 and H1N1 subtypes. The hotspot regions of antigen epitopes were primarily located around the catalytic active site. The antibodies used for structural analysis were primarily derived from human and murine sources. Among murine antibodies, the most frequently used V-J gene combination was IGHV1-12*01/IGHJ2*01, while for human antibodies, the most common combination was IGHV1-69*01/IGHJ6*01. There were significant differences in the lengths and usage preferences of heavy chain CDR amino acids between antibodies that bind within the catalytic active site and those that bind to regions outside the catalytic active site. The results revealed that structurally similar antibodies could recognize the same epitopes, indicating a specific spatial recognition between antibody and antigen epitopes. Structural overlap in the binding regions was observed for antibodies with similar paratope structures, and the competitive binding of these antibodies to the epitope was confirmed through BLI experiments. **Conclusion** The antigen epitopes of NA protein mainly distributed around the catalytic active site and its surrounding loops. Spatial complementarity and electrostatic interactions play crucial roles in the recognition and binding of antibodies to antigenic epitopes in the catalytic region. There existed a spatial recognition relationship between antigens and antibodies that was independent of the uniqueness of antibody sequences, which means that antibodies with different sequences could potentially form similar local spatial structures and recognize the same epitopes.

Key words influenza A virus, neuraminidase, epitopes, antibody, similarity of paratope

DOI: 10.16476/j.pibb.2024.0309

CSTR: 32369.14.pibb.20240309