宇际互影
Mutiin Scientific Service

GM5 分子与蛋白对接分析报告

分子对接分析报告

项目编号： SAMPLE-02

报告日期： 2026-11-08

版本号： v1.0

上海宇际互影科技服务有限公司

www.mutiin.com

摘要

本报告针对 GM5 与 Prot-B 的分子识别开展 HADDOCK3 对接分析。多糖的共价连接与立体化学依据经审计的结构拓扑严格保留，蛋白结合区域经残基映射后用于构建模糊相互作用约束。排名首位构象的 HADDOCK score 为 −87.61，用于同一计算设置下的相对排序，不代表实验结合自由能。在代表构象中识别到 6 项通过化学合理性审核的可视化相互作用，主要包括氢键 5 项、疏水作用 1 项。综合候选排序、界面几何和关键残基分布，本结果为多糖-蛋白识别模式及后续分子动力学与实验验证提供结构假设。

关键词：GM5； Prot-B ；多糖-蛋白分子对接；HADDOCK3；糖苷键拓扑；界面相互作用

目录

1 引言 **2**

2 模型与方法 **2**

2.1 研究体系与结构来源 2

2.2 结构预处理与多糖拓扑保持 2

2.3 HADDOCK3 对接策略与参数 3

2.4 相互作用识别与可视化 3

3 结果与分析 **4**

3.1 模型生成与筛选质量 4

3.2 代表性构象与评分 4

3.3 三维结合模式 5

3.4 二维相互作用与关键残基 6

3.5 结果解释 8

4 结论 **8**

A 附录：附件文件说明 **8**

1 引言

多糖-蛋白分子识别广泛参与细胞黏附、免疫调控、受体激活、病原体侵染和细胞外基质组装等生物学过程。与典型小分子相比，多糖通常具有更多可旋转糖苷键、密集分布的氢键供受体以及明确的单糖立体化学，其识别模式往往由多个局部接触协同维持，并对糖链构象与蛋白表面形貌高度敏感。因此，多糖-蛋白对接既要评估界面几何互补和非共价作用网络，也要在结构准备、采样和结果恢复过程中保持糖苷键连接、单糖构型及多链蛋白身份的一致性 [1]。

本研究以 GM5 与 Prot-B 为对象，采用 HADDOCK3 在经审计的候选结合区域约束下开展刚体取向采样、界面聚类、候选筛选与柔性精修，并以 HADDOCK score、能量分项、约束满足度和界面接触评价候选构象 [2]。本任务的结合区域来源为 peptide_-envelope_blind（低置信度）。对柔性精修模型给出 Top 10 单模型排名，并将簇级群体支持度作为独立证据报告，避免把低群体小簇等同于稳定构象共识。

2 模型与方法

2.1 研究体系与结构来源

本案例以 GM5 作为多糖分子，以 Prot-B 作为蛋白受体。 Prot-B 采用客户提供的结构文件（ Prot-B.cif ），结构中包含 1 条链（A）；GM5 采用客户提供的结构文件（GM5.mol2）。结构审计将该多糖归类为线性寡糖、5 个单糖单元、1 个共价连通组分、4 条糖苷连接。上述体系组成与拓扑信息用于后续约束定义、构象评价和结果解释。

2.2 结构预处理与多糖拓扑保持

对接前对蛋白原子记录、链标识、残基编号及约束位点的可映射性进行核验。多链蛋白作为一个完整受体伙伴参与刚体取向采样；为适配 HADDOCK3 的伙伴标识规则，计算阶段对链与残基编号建立可逆映射，最终代表构象恢复为输入结构中的原始链和残基身份，避免多链信息在计算与报告之间发生混淆。

对于 GM5，其原子连接、糖苷键、原子类型、电荷及立体化学以经审计的 MOL2 拓扑为依据。输入共价图经连通性、残基映射、糖苷键和立体化学一致性审核后转换为 HADDOCK3 可识别的糖类残基与连接定义；坐标文件不被用于重新猜测糖苷键。该处理采用 GLYCAM 06j-1 糖类参数 [3]。

2.3 HADDOCK3 对接策略与参数

经结构映射确定的蛋白活性(active)残基为 A:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20，位点来源为 peptide_envelope_blind（低置信度）；依据其空间邻接关系扩展 0 个邻近（passive）残基。两类位点共同用于构建模糊相互作用约束（ambiguous interaction restraints, AIRs），引导多糖在候选界面附近进行取向采样 [4]。该结合区域来自自动预测而非实验位点；因此结果只评价该候选区域内的相对构象，不能据此排除蛋白表面的其他潜在多糖结合区域。HADDOCK3 依次执行拓扑准备、刚体采样、CAPRI 评分、界面-配体 RMSD 聚类、跨簇候选筛选、柔性精修和最终全原子评价；该组合策略与经蛋白-多糖基准验证的 HADDOCK 建模协议一致 [5, 2]。柔性精修阶段将 Z:1–5 设为连续柔性片段，使糖链在保持共价拓扑的前提下调整局部构象。

表 1: 多糖-蛋白分子对接的主要计算参数

参数	设置
对接程序	HADDOCK3
多糖拓扑与参数	GLYCAM 06j-1；MOL2 拓扑审计
采样档位	标准（standard）
并行核心数	32
刚体取向采样数	1000 个候选模型
刚体采样容差	5
初始界面聚类	基于界面 RMSD 的 maxclust 聚类，目标簇数 50
每簇进入柔性精修的模型数	5
多糖柔性片段	Z:1–5
柔性精修容差	20
最终界面聚类	界面-配体 RMSD 距离聚类，average linkage，截断值 2.5 Å，最小有效簇规模 4
客户排名	柔性精修后按单模型 HADDOCK score 输出 Top 10；簇级支持度单独报告
蛋白 active 残基	A:1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20
邻近 passive 残基	0 个

2.4 相互作用识别与可视化

代表性复合物中的氢键、疏水接触及其他非共价作用依据原子级几何与化学规则识别，并对蛋白残基身份、作用端点、距离和多糖原子映射进行一致性审核。三维图用于呈

现多糖相对蛋白整体结构的位置、结合区域和局部界面空间关系；二维图采用 LigPlot+ 排版并保持经校验的多糖共价图，用于比较关键蛋白残基、作用类型和几何距离 [6, 7]。

3 结果与分析

3.1 模型生成与筛选质量

刚体阶段成功生成 1000/1000（100.0%）的预期模型；进入柔性精修的候选中，175/175（100.0%）完成精修并进入后续评价。最终 2.5 Å 界面-配体 RMSD 聚类未形成至少含 4 个模型的有效簇；仅 28/175 个精修模型获得临时小簇归属。因此，本报告采用柔性精修后的单模型 HADDOCK score 给出 Top 10 排名，并将小簇信息仅作为构象异质性的审计证据，不解释为稳定构象共识。模型生成数量与精修完成率说明本次计算获得了可用于相对排序的候选集合；簇级证据强弱则单独决定是否能够讨论构象群体共识。

3.2 代表性构象与评分

在本次约束与采样设置下，柔性精修后单模型排序首位构象的 HADDOCK score 为 −87.61。其范德华能、静电能、去溶剂化能和 AIR 约束项分别为 −28.04、−61.93、1.50 和 78.96，埋藏表面积为 703.58 Å^2。该模型不依赖未达最小群体要求的小簇进行推荐。HADDOCK score 是由多个能量和约束分项构成的相对排序指标，其数值不等同于实验结合自由能，也不应直接换算为 kcal·mol^{-}1。AIR 约束项为 78.96，提示代表构象仍存在一定约束偏离；因此，结合位点匹配应结合约束来源与实验位点信息审慎解释。

表 2: 代表性对接构象的 HADDOCK3 评分与界面参数

参数	数值
HADDOCK score	−87.61
范德华能	−28.04
静电能	−61.93
去溶剂化能	1.50
AIR 约束项	78.96
埋藏表面积（Å^2）	703.58
单模型交付排序	1
有效簇支持	无

表 3: 柔性精修后 Top 10 单模型 HADDOCK3 排名。各能量项仅用于相同设置下的相对比较。

排名	HADDOCK score	AIR	E_{vdW}	E_{elec}	E_{desolv}	BSA（Å^2）	簇证据
1	−87.61	78.96	−28.04	−61.93	1.50	703.58	–
2	−83.55	31.84	−39.16	−38.12	−1.22	823.07	–
3	−81.21	78.44	−25.65	−57.91	1.66	715.30	–
4	−78.17	70.77	−12.54	−65.39	−1.28	604.02	–
5	−74.46	114.45	−27.46	−50.82	−0.81	681.46	–
6	−72.99	106.85	−33.21	−44.20	1.22	747.63	–
7	−72.11	45.76	−23.11	−45.99	−0.64	694.64	–
8	−72.10	96.05	−36.27	−31.77	−6.00	766.25	–
9	−68.42	33.82	−28.97	−35.99	0.23	705.63	11（n=2，不足）
10	−67.77	94.40	−34.47	−34.10	−0.11	854.25	–

3.3 三维结合模式

图 1 展示了单模型排名第 1 位的多糖构象在蛋白表面的总体位置及局部界面。整体视图用于判断结合区域与蛋白结构域的空间对应关系，局部视图则突出直接参与作用的残基与多糖原子。该图用于解释当前优先候选，不代表 Top 10 构象已经收敛为唯一结合模式。

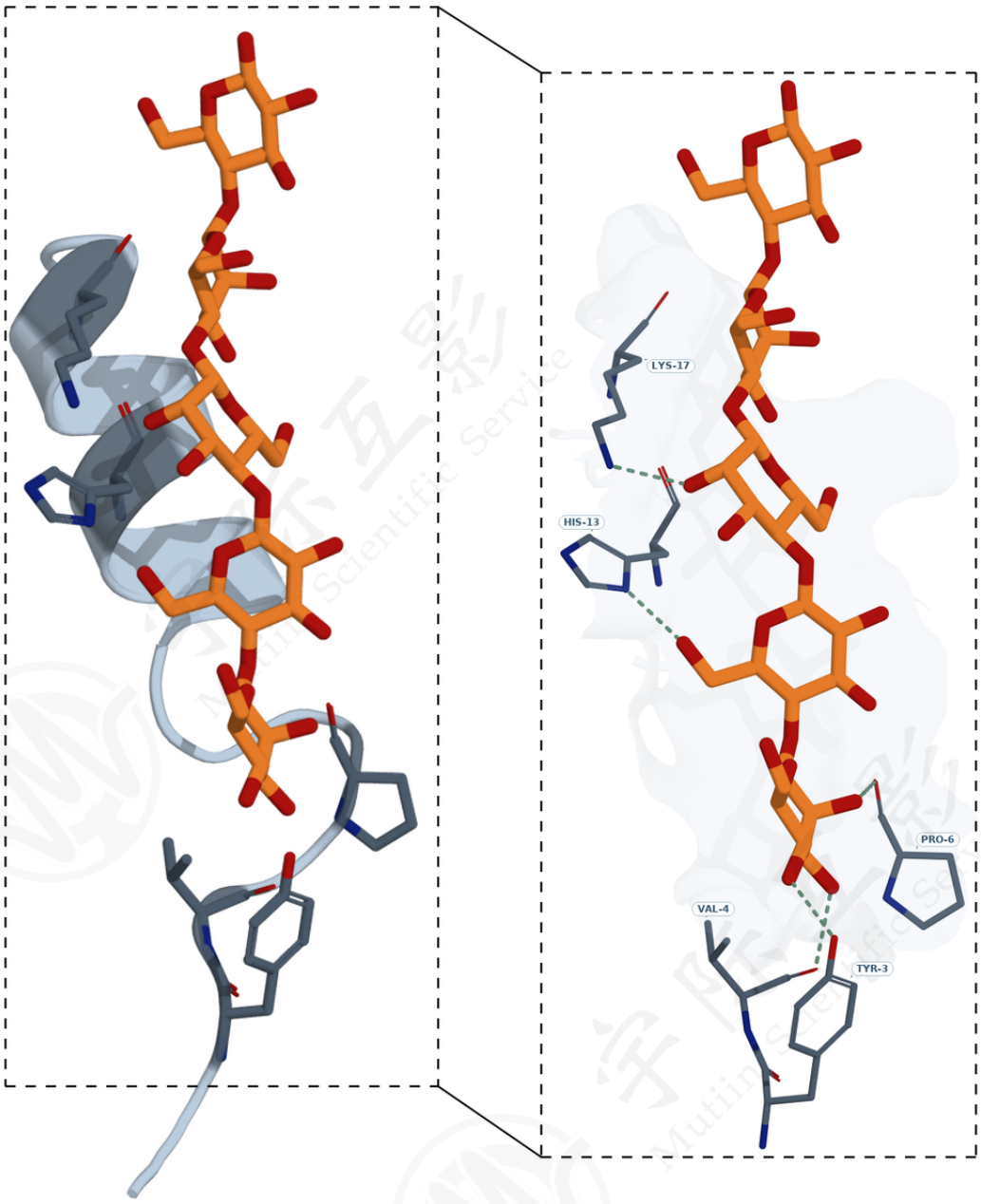


图 1: GM5 与 Prot-B 代表性对接构象的三维结合模式。左侧为整体结构与结合区域定位，右侧为局部界面及关键相互作用。

3.4 二维相互作用与关键残基

代表性构象中共识别到 6 项满足几何与化学审核标准的可视化相互作用，主要包括氢键 5 项、疏水作用 1 项。相关蛋白残基包括 A:LYS-17、A:TYR-3、A:HIS-13、A:PRO-6、A:VAL-4。这些距离描述单一代表构象中的局部几何关系，不代表动力学占有率、结合自由能或实验作用强度。

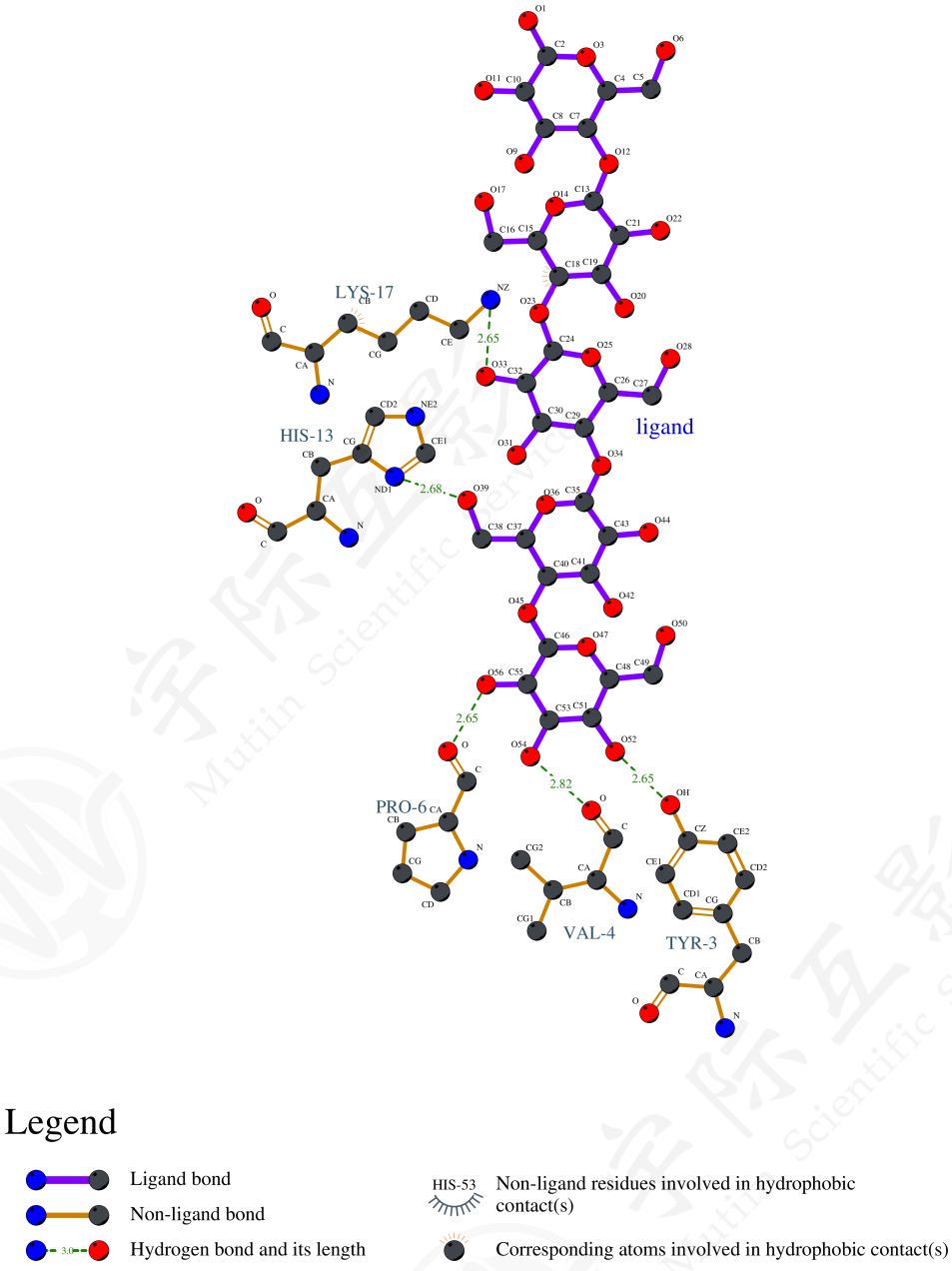


图 2: GM5 与 Prot-B 代表性构象的二维相互作用图。

表 4: 代表性构象中的主要多糖-蛋白相互作用

序号	类型	蛋白残基	端点原子	距离（Å）
1	氢键	A:LYS-17	NZ–O01R	2.65
2	疏水作用	A:LYS-17	CB–C00Y	3.88
3	氢键	A:TYR-3	OH–O02R	2.65
4	氢键	A:HIS-13	ND1–O022	2.68
5	氢键	A:PRO-6	O–O02Z	2.65
6	氢键	A:VAL-4	O–O02V	2.82

3.5 结果解释

排名第 1 位构象显示 GM5 在 Prot-B 的候选区域形成可解析界面；经审核的相互作用主要包括氢键 5 项、疏水作用 1 项，涉及 A:LYS-17、A:TYR-3、A:HIS-13、A:PRO-6、A:VAL-4。这些接触为局部极性互补和界面几何匹配提供结构依据。鉴于多糖具有较高构象自由度、AIR 约束项非零，界面约束尚未被完全满足，且最终聚类未形成满足最小群体要求的构象共识，本结果应被视为位点约束与采样条件下的候选结合假设，而非稳定结合或高亲和力的直接证据。该结合区域来自自动预测而非实验位点；因此结果只评价该候选区域内的相对构象，不能据此排除蛋白表面的其他潜在多糖结合区域。显式溶剂分子动力学、独立重复采样及定点突变实验可进一步检验关键接触的持续性与功能相关性。

4 结论

HADDOCK3 对接获得了 GM5 与 Prot-B 的 Top 10 柔性精修候选构象；排名第 1 位构象中识别到 6 项通过审核的可视化相互作用。综合模型生成质量、单模型相对排序、簇级支持度、界面几何及关键残基分布，该构象可作为后续验证的优先结构假设。

本结果反映本次输入结构、peptide_envelope_blind（低置信度）和采样范围内的相对优先构象。最终 2.5 Å 界面-配体 RMSD 聚类未形成至少含 4 个模型的有效簇；仅 28/175 个精修模型获得临时小簇归属。因此，本报告采用柔性精修后的单模型 HADDOCK score 给出 Top 10 排名，并将小簇信息仅作为构象异质性的审计证据，不解释为稳定构象共识。由于单一静态构象不能反映溶剂竞争、界面重排和接触占有率，建议将排名构象用于后续显式溶剂分子动力学模拟，并结合独立重复轨迹、自由能评估或关键残基实验验证其稳定性与生物学相关性。

A 附录：附件文件说明

本报告随任务结果提供下列多糖-蛋白分子对接核心附件，涵盖第 1 名代表结构、前 10 名候选构象、评分与能量分项以及相互作用图件，便于结果复核、展示及后续建模。

表 5: 附件文件说明

附件文件	内容说明	建议用途
protein_glycan_docking_report.pdf	本次多糖-蛋白分子对接分析的正式 PDF 报告。	阅读、归档、提交
protein_glycan_docking_report_editable.docx	与 PDF 内容一致的可编辑 Word 文档。	二次编辑、补充实验信息

附件文件	内容说明	建议用途
docked_complex_best.pdb	按 HADDOCK3 单构象评分排序得到的第 1 名多糖-蛋白复合物结构。	结构复查、后续模拟
glycan_pose_best.sdf	与第 1 名复合物坐标对应的多糖 SDF 文件，保留糖苷键连接、键级、电荷和立体化学。	多糖结构复核、后续建模
glycan_poses_ranked.sdf	评分前 10 名多糖构象的 SDF 合集，构象顺序与 docking_scores.xlsx 中的排名对应。	多构象比较、二次筛选
docking_scores.xlsx	评分前 10 名精修构象的 HADDOCK3 总分与能量分项；HADDOCK 分数不等同于结合自由能。	排序复核、数据归档
interaction_3d.png	代表性构象的三维综合图，包含结合区域总览和局部相互作用细节。	报告配图、展示汇报
interaction_3d_partA.png	多糖-蛋白结合区域的整体视图。	结构定位、界面说明
interaction_3d_partB.png	多糖-蛋白局部相互作用细节图。	关键残基解读
interaction_2d.png	基于 LigPlot+/相互作用识别结果生成的二维相互作用图。	快速查看关键相互作用
pymol_session.pse	保留推荐观察角度、颜色和显示状态的 PyMOL 会话。	继续调整视角或导出图片

参考文献

[1] Zhang, S., Chen, K. Y. & Zou, X. Carbohydrate–protein interactions: Advances and challenges. *Communications in Information and Systems* **21**, 147–163 (2021).

[2] Giulini, M. *et al.* HADDOCK3: A modular and versatile platform for integrative modeling of biomolecular complexes. *Journal of Chemical Information and Modeling* **65**, 7315–7324 (2025).

[3] Kirschner, K. N. *et al.* GLYCAM06: A generalizable biomolecular force field. carbohydrates. *Journal of Computational Chemistry* **29**, 622–655 (2008).

[4] Dominguez, C., Boelens, R. & Bonvin, A. M. J. J. HADDOCK: A protein–protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society* **125**, 1731–1737 (2003).

[5] Ranaudo, A., Giulini, M., Pelissou Ayuso, A. & Bonvin, A. M. J. J. Modeling protein–glycan interactions with HADDOCK. *Journal of Chemical Information and Modeling* **64**, 7816–7825 (2024).

[6] Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Research* **43**, W443–W447 (2015).

[7] Laskowski, R. A. & Swindells, M. B. LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. *Journal of Chemical Information and Modeling* **51**, 2778–2786 (2011).