宇际互影
Mutiin Scientific Service

citric_acid 与 protein_2Zn 分子动力学模拟分析报告

分子动力学模拟分析报告

项目编号： SAMPLE-03

报告日期： 2026 年 11 月 06 日

上海宇际互影科技服务有限公司

www.mutiin.com

摘要

本报告以 citric_acid 与 protein_2Zn 复合物为对象，开展 100 ns 显式溶剂分子动力学模拟。体系包含 2 个 Zn 配位中心，采用 MCPB.py 成键模型描述其配位相互作用。模拟保持预设成键拓扑，不描述配位键断裂、形成或配体交换。在 90–100 ns 内，RMSD 为 0.259 ± 0.017 nm，Rg 为 1.677 ± 0.005 nm，SASA 为 108.004 ± 1.382 nm^2。在所选时间窗口并采用 0.15 nm 截断值时，1001 帧归入同一构象簇，支持末段局部构象同质性，但不单独证明全程收敛或唯一全局构象态。末段代表构象中识别到 3 项几何接触，包括氢键 1 项、金属配位 2 项。单对氢键最高占有率为 30.8%，最长连续寿命为 1.40 ns，连续寿命受采样间隔影响，两个极值不一定来自同一氢键。综合结果提示，本次轨迹中的代表构象保留了受体–配体接触；该结论描述的是本次轨迹中的结构保持，不等同于已证实的高亲和力结合。上述判断限定于本次单条轨迹与所选末段窗口，亲和力排序或机制推断仍需独立重复模拟和实验数据验证。

关键词：citric_acid；protein_2Zn；分子动力学；构象稳定性；构象聚类；离子配位；金属感知质子化；MCPB.py

目录

1 引言 3

2 模型与方法 3

2.1 初始复合物与力场参数化 3

2.2 分子动力学模拟方案 4

2.3 轨迹处理、统计分析与稳定构象提取 4

3 结果与分析 5

3.1 轨迹稳定性与构象变化概览 5

3.2 整体构象偏移与局部柔性 6

3.3 整体紧致性、界面氢键与溶剂暴露 6

3.4 构象分布与末段聚类 6

3.5 代表构象的三维与二维相互作用 7

3.6 补充结构动力学分析 8

3.6.1 二级结构时间演化 8

3.6.2 残基协同运动与主要构象空间 9

3.6.3 主链二面角分布 11

3.6.4 氢键寿命与界面接触持续性 12

4 结论 15

A 附录：附件文件说明 16

1 引言

分子动力学模拟（molecular dynamics, MD）可用于考察候选复合物在显式溶剂、离子环境与热涨落条件下的结构响应，但单条轨迹本身不能直接证明实验亲和力或作用机制。本报告以分子对接获得的 citric_acid 与 protein_2Zn 复合物为起始模型。本研究开展 100 ns 显式溶剂模拟，并从整体构象偏移、局部柔性、蛋白紧致性和界面接触等方面评估该复合物在本次采样时间尺度内的表现。

报告基于末段分析窗口进行构象聚类并提取代表帧，随后对同一构象开展三维和二维相互作用分析 [1, 2]。代表帧用于描述该时间窗口内的典型结构，不被解释为唯一结合构象。除常规轨迹指标外，报告还结合二级结构、协同运动、主成分空间、主链构象和界面接触持续性，用于区分蛋白整体稳定、局部构象变化与配体接触保持这三类证据。

2 模型与方法

2.1 初始复合物与力场参数化

本研究以 protein_2Zn 为受体蛋白、citric_acid 为小分子配体建立复合物模型。受体经结构准备与质子化处理，保留配位金属及所属辅因子。小分子的键级与质子化状态沿用对接前准备的结构，采用 AutoDock4 metal maps 在指定活性位点内进行姿态搜索 [3]，再结合对接评分、金属配位几何和口袋相互作用合理性选取后续分子动力学初始构象。

配体的键级与质子化状态沿用对接前准备的质子化结构，重原子位置来自对接姿态。所用配体形式电荷为 citric_acid：−3。局部量子簇的几何优化和 Hessian 计算采用 TPSS-D3(0)/def2-SVP with CPCM water。优化中每 20 步重新计算 Hessian。金属位点的 8 项键参数和 25 项角参数由 EasyPARM 从量子 Hessian 提取，并通过非封端原子映射转移到完整体系。金属位点原子电荷采用 MCPB 大模型的两阶段 RESP 拟合结果，小模型电荷不用于完整体系 [4][5]。配体其余有机参数采用 GAFF2。金属中心属于对接位点，并采用固定成键拓扑。该模型可描述给定配位态附近的结构涨落，不能用于验证配位键自发形成、断裂或配体交换。对 2 个金属中心采用 MCPB.py 成键模型建立局部参数 [4]。局部拓扑包含 8 条金属–供体配位键；依据配位原子类型确定供体残基的质子化状态，包括组氨酸的互变异构状态以及羧酸盐和硫醇盐状态。

蛋白标准残基采用 Amber ff14SB 力场描述 [6]。体系采用 TIP3P 三点水模型进行显式溶剂化 [7]。建盒时设置溶质与盒边缘的缓冲距离为 1.0 nm。加入 0.15 M 的 Na/Cl 盐，并中和体系净电荷。上述参数共同构成生产期动力学使用的受体、配体、离子与溶剂模型。

2.2 分子动力学模拟方案

分子动力学计算由 GROMACS 2026.3 执行 [8, 9]。首先对复合物体系进行能量最小化，以消除局部不合理接触；随后依次进行 NVT 与 NPT 平衡，使体系在 300 K 温度和常压条件下达到较合理的密度与热力学状态；最后开展生产期 MD 模拟并输出轨迹。生产期轨迹用于计算 RMSD、RMSF、Rg、蛋白-配体氢键数量、SASA 和 RMSD-Rg 构象分布。RMSD-Rg 图由轨迹的二维占据概率换算为相对自由能，用于描述本次采样的构象分布，不是配体结合自由能。

结构动力学分析包括 DSSP、DCCM、PCA、Ramachandran 分布、氢键寿命及界面接触占有率。结构动力学分析基于全程轨迹，与末段构象聚类分别统计。分析前进行周期性边界修正并按受体主链对齐；接触指标使用受体和配体各自的原子分组。

2.3 轨迹处理、统计分析与稳定构象提取

轨迹先进行周期性边界修正与受体主链对齐，再计算 RMSD、RMSF、Rg、蛋白-配体氢键、SASA 及 RMSD-Rg 构象分布。数值表汇总于 md_analysis_data.xlsx，标准动力学、结构动力学分别置于独立工作表。

构象分析采用 90–100 ns 轨迹区间，并以 GROMOS 方法完成聚类，截断值为 0.15 nm。最大构象簇的中心构象作为代表帧，用于三维口袋定位、二维相互作用识别和关键残基归属。聚类结果仅适用于所选时间窗口和截断值，不作为整条轨迹已充分收敛的单独证据。

3 结果与分析

3.1 轨迹稳定性与构象变化概览

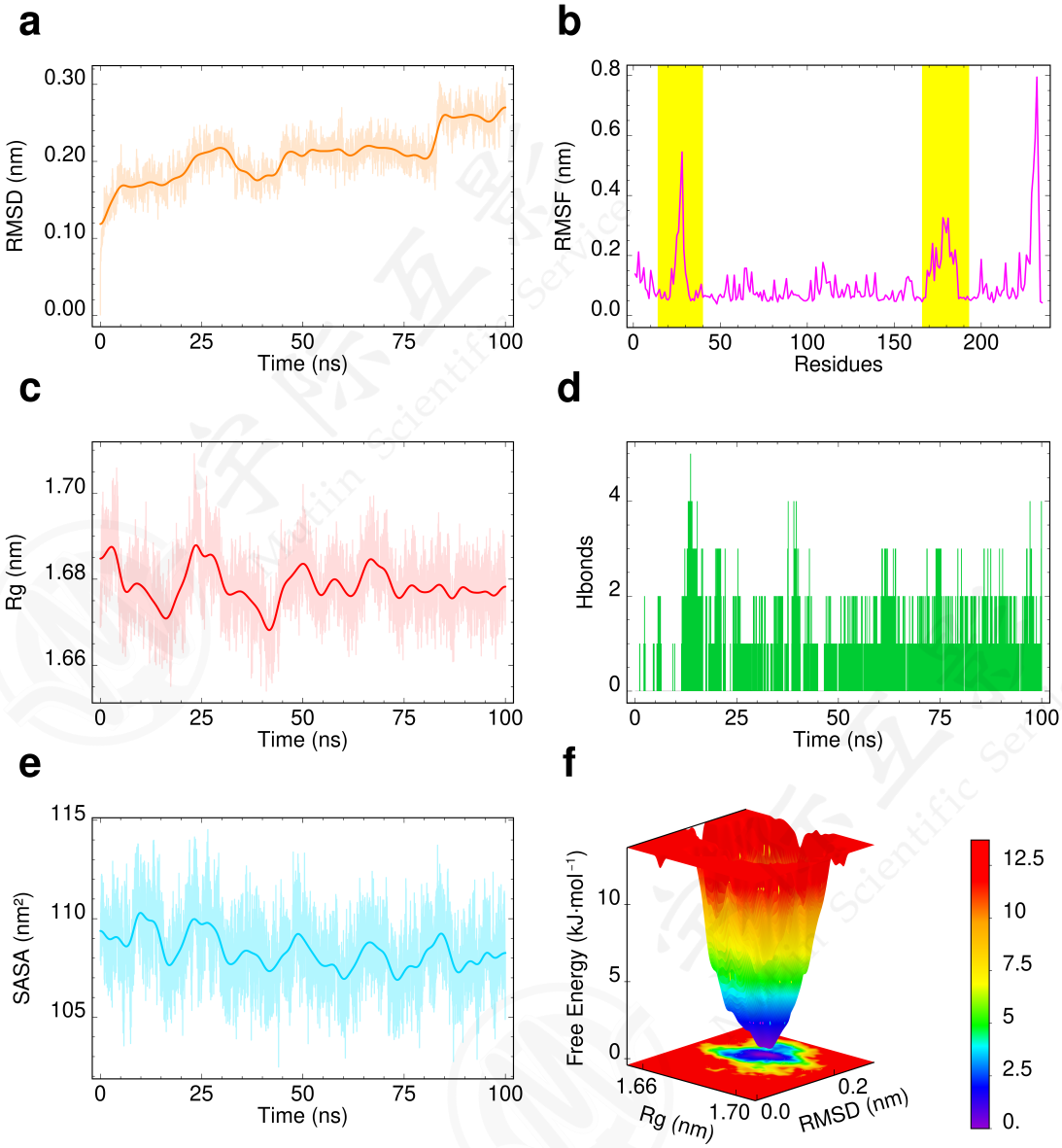


图 1: 分子动力学轨迹综合分析：RMSD、RMSF、Rg、蛋白-配体氢键数量、SASA 及 RMSD-Rg 构象分布。

图1 与表1 汇总了本次轨迹的主要结构指标。以下分析重点区分“末段波动收窄”与“相对初始构象偏移较小”这两个不同概念，并将蛋白整体稳定性与配体结合证据分别讨论。

表 1: 分子动力学轨迹主要指标及分析窗口统计。

指标	单位	全程均值	全程范围	分析窗口均值 ± 标准差
RMSD	nm	0.206	0.001–0.309	0.259 ± 0.017
Rg	nm	1.678	1.654–1.709	1.677 ± 0.005
蛋白-配体氢键	数量	0.718	0.000–5.000	0.898 ± 0.699
SASA	nm²	108.362	102.472–114.522	108.004 ± 1.382

3.2 整体构象偏移与局部柔性

RMSD 全程均值为 0.206 nm，范围为 0.001–0.309 nm。在选定的末段分析窗口（90–100 ns）内，RMSD 为 0.259 ± 0.017 nm。RMSD 在分析窗口内的波动较全程明显收窄，在该窗口内未见强持续单向漂移。分析窗口内 RMSD 进入相对集中的取值区间。该窗口约占整段轨迹的 10%，因此它支持末段局部稳定，但不足以单独证明全程收敛。

相对波动较大的残基包括：PRO-232（0.794 nm）、VAL-231（0.589 nm）、PHE-28（0.545 nm）、LEU-230（0.481 nm）、LYS-229（0.408 nm）、GLY-27（0.403 nm）。这些残基的柔性涉及的结构区域应结合三维位置判断。

3.3 整体紧致性、界面氢键与溶剂暴露

Rg 全程均值为 1.678 nm，分析窗口均值为 1.677 ± 0.005 nm。Rg 在分析窗口内变异很小，在该窗口内未见强持续单向漂移，未显示分析窗口内发生明显整体塌缩或展开。蛋白-配体氢键数量分析窗口均值为 0.898 ± 0.699，提示氢键具有一定间歇性，需要结合占有率或关键帧进一步确认；该数值表示同一时刻的平均氢键数量，不能替代单个氢键的占有率与寿命。SASA 分析窗口均值为 108.004 ± 1.382 nm²，SASA 在分析窗口内存在有限波动，在该窗口内未见强持续单向漂移。上述指标描述整体紧致度、溶剂暴露和氢键数量；氢键的持续性还需结合逐对占有率和连续寿命判断。

3.4 构象分布与末段聚类

在 90–100 ns 内采用 GROMOS 方法聚类，主簇包含 1,001 帧，代表帧时间约为 93.47 ns，导出的代表复合物包含 3,439 个原子。主簇占已报告聚类帧的 100.0%，表明在当前窗口和截断值下构象较为集中。聚类结果受分析窗口与 RMSD 截断值影响；即使全部帧归入同一簇，也只能说明该末段窗口内的局部构象同质性，不能据此认定整条轨迹仅存在一个全局构象态。

表 2: 分析窗口内 GROMOS 聚类结果摘要。

构象簇	帧数	平均 RMSD (nm)	代表帧时间 (ns)
1	1,001	0.101	93.47

3.5 代表构象的三维与二维相互作用

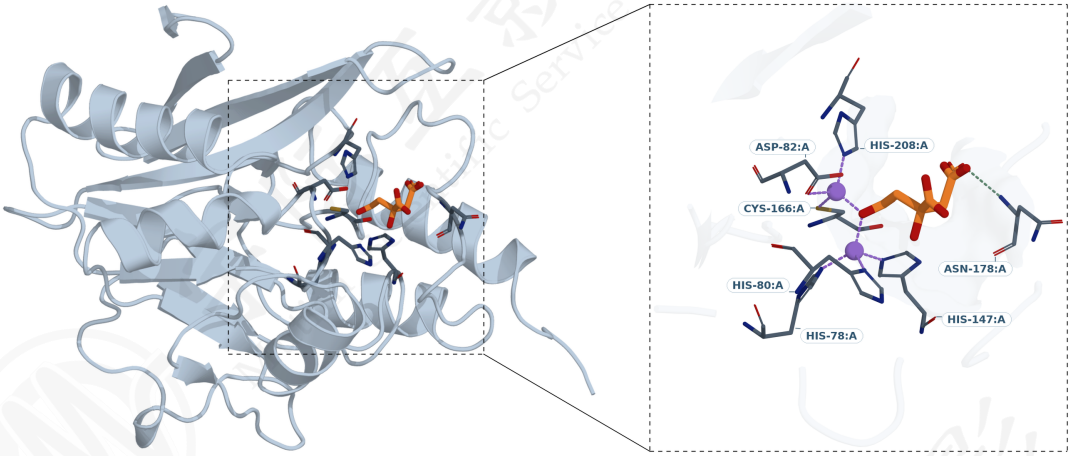


图 2: 分析窗口代表构象的三维相互作用图。

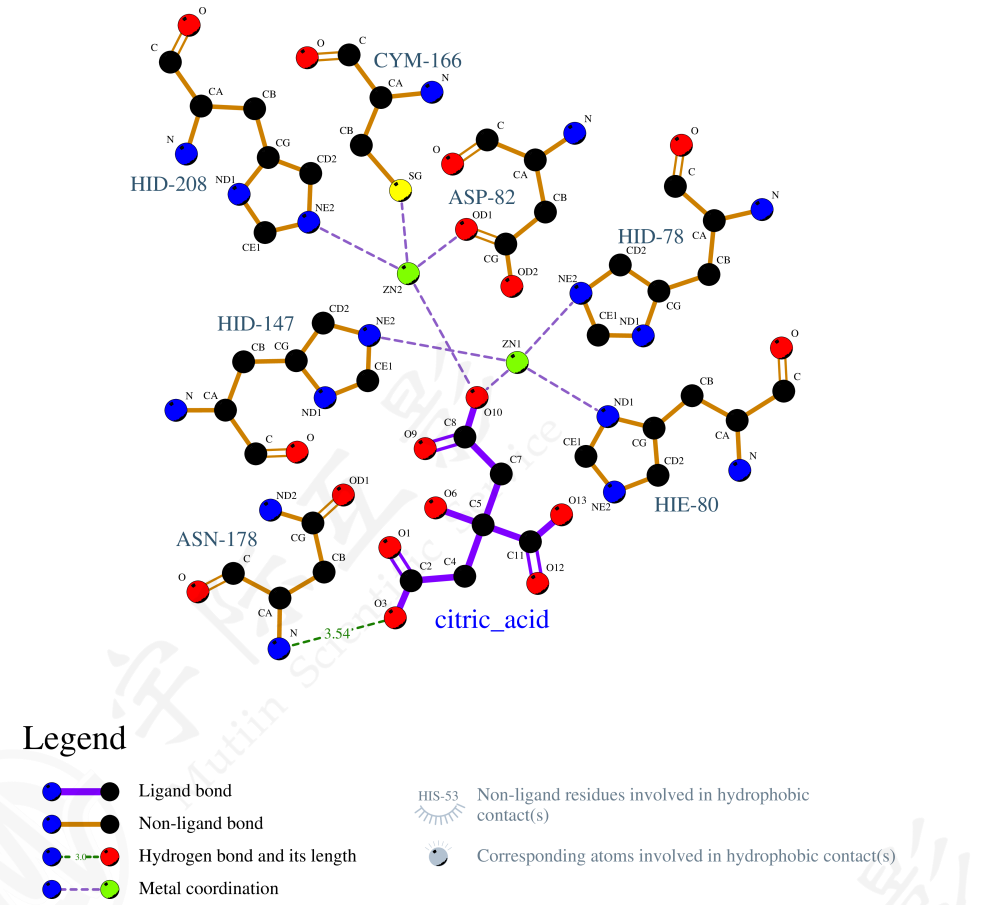


图 3: 与三维图对应的二维相互作用图。

分析窗口代表构象中共识别到 3 项几何接触，主要包括氢键 1 项、金属配位 2 项，涉及残基包括 ASN-178、ZN-234、ZN-235。氢键可作为极性锚定线索。相互作用类型按代表帧的原子几何识别。三维图与二维图对应同一代表帧，用于核对口袋位置、作用方向、距离及残基归属；这些单帧结果需与轨迹占有率和氢键寿命联合解释。

表 3: 分析窗口代表构象中的主要蛋白-配体接触。

接触类型	残基	距离 (Å)
氢键	ASN-178	3.54
金属配位	ZN-234	2.02
金属配位	ZN-235	1.87

3.6 补充结构动力学分析

3.6.1 二级结构时间演化

DSSP 分析覆盖 1,001 个轨迹帧和 232 个蛋白残基。按每帧每残基归并统计，螺旋、折叠和转角/无规卷曲状态分别占 23.8%、28.6% 和 47.6%。这些比例反映所采样残基与

时间点的平均组成；局部构象变化见热图。

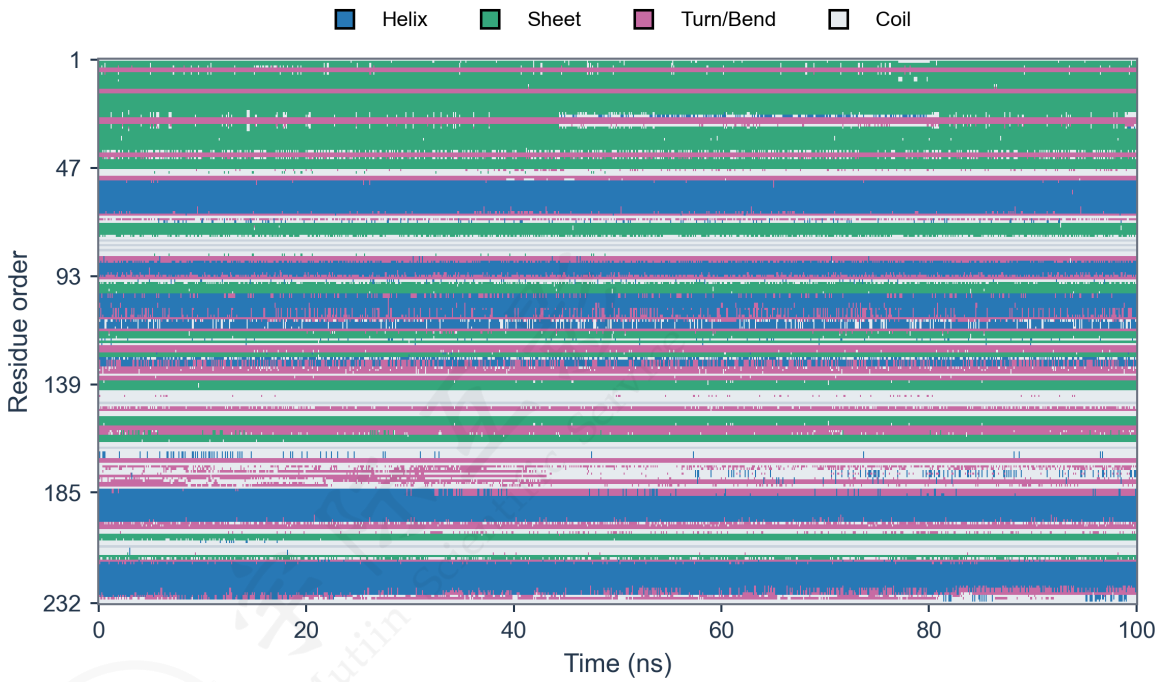


图 4: DSSP 二级结构随时间演化热图。横轴为模拟时间，纵轴为蛋白残基；颜色表示各帧的二级结构类型。

3.6.2 残基协同运动与主要构象空间

DCCM 基于 1,001 个轨迹帧、232 个 Cα 原子计算。非对角区域的相关系数用于描述不同残基涨落的相关程度，最低值为 -0.529。

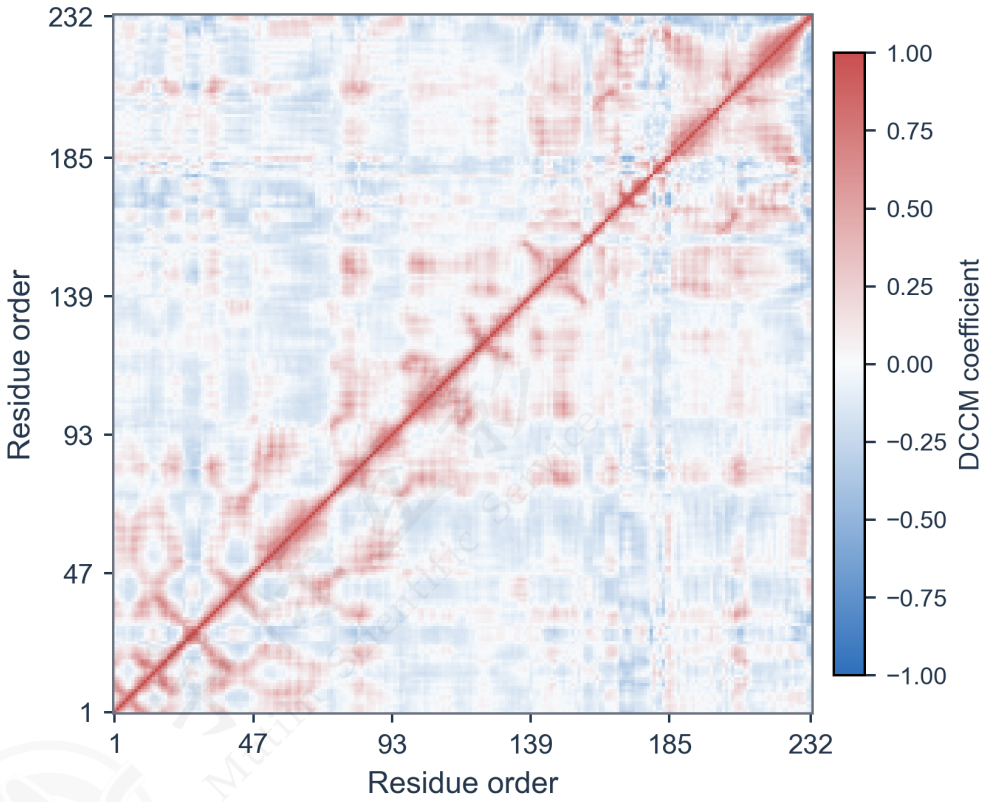


图 5: 蛋白 Cα 原子动态互相关矩阵（DCCM）。正、负相关分别表示同向和反向协同运动。

PCA 使用 1,001 个轨迹帧和 232 个 Cα 原子建立协方差模型。PC1 与 PC2 分别解释 25.20% 和 16.42% 的坐标方差，累计解释率为 41.62%。

PC1 与 PC2 表示对齐后 Cα 坐标沿前两个协方差特征向量的投影，以 nm 表示。图中对 PC1–PC2 平面内的轨迹帧进行二维分箱，并经轻度高斯平滑后以对数尺度显示局部采样频数。高频区域表示本次轨迹中重复访问较多的构象盆地；起始构象所在区域标记为 Initial region，局部采样频数最高的区域标记为 Dominant basin。上述名称仅描述采样过程和占据特征；在缺少开放度结构指标或已知参考构象时，不将其直接解释为 Open 或 Closed 功能状态。

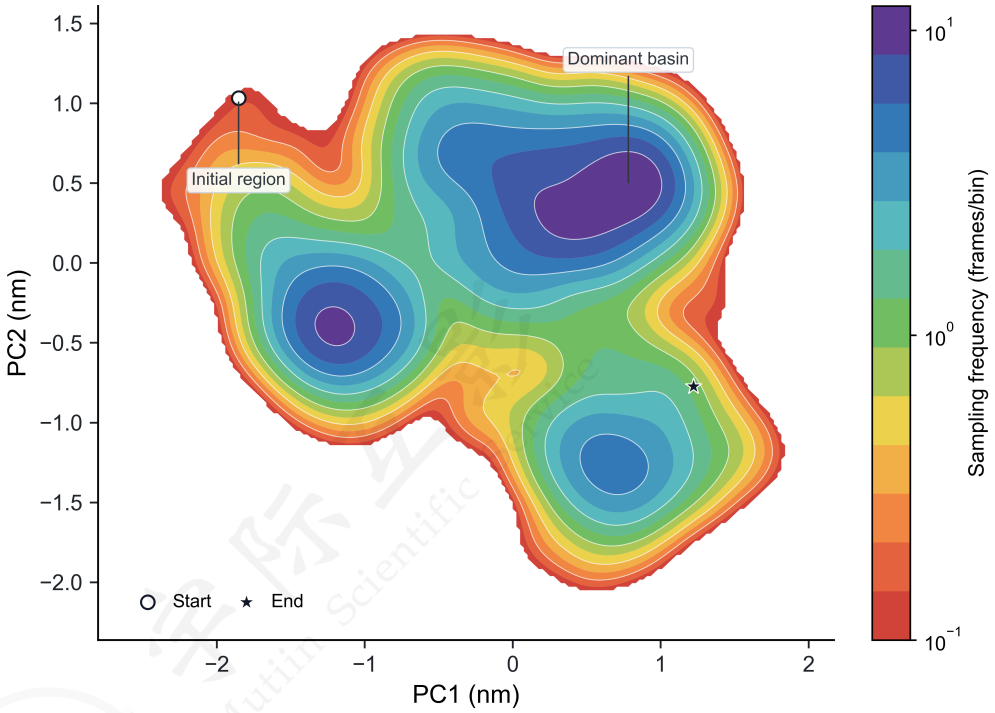


图 6: 主成分空间对数采样频率分布。横、纵坐标分别为 PC1 和 PC2 投影（nm）；颜色表示经轻度平滑的局部采样频数（frames/bin，对数尺度），空心圆与星形分别标记起始帧和结束帧。

3.6.3 主链二面角分布

Ramachandran 分析汇总 230,230 个 ϕ/ψ 二面角样本。

按残基类别分为 General 189,189 点、Glycine 21,021 点、Pre-Proline 9,009 点、Proline 11,011 点。Gly、Pro、Pre-Proline 与一般残基分别采用独立的对数密度着色，并叠加稀疏采样点以兼顾高频构象盆和低频采样点。各面板的 n 表示残基与轨迹帧组合形成的二面角样本数。密度反映本次轨迹的采样分布，不代表外部参考有利区；未计算有利区或离群区比例。

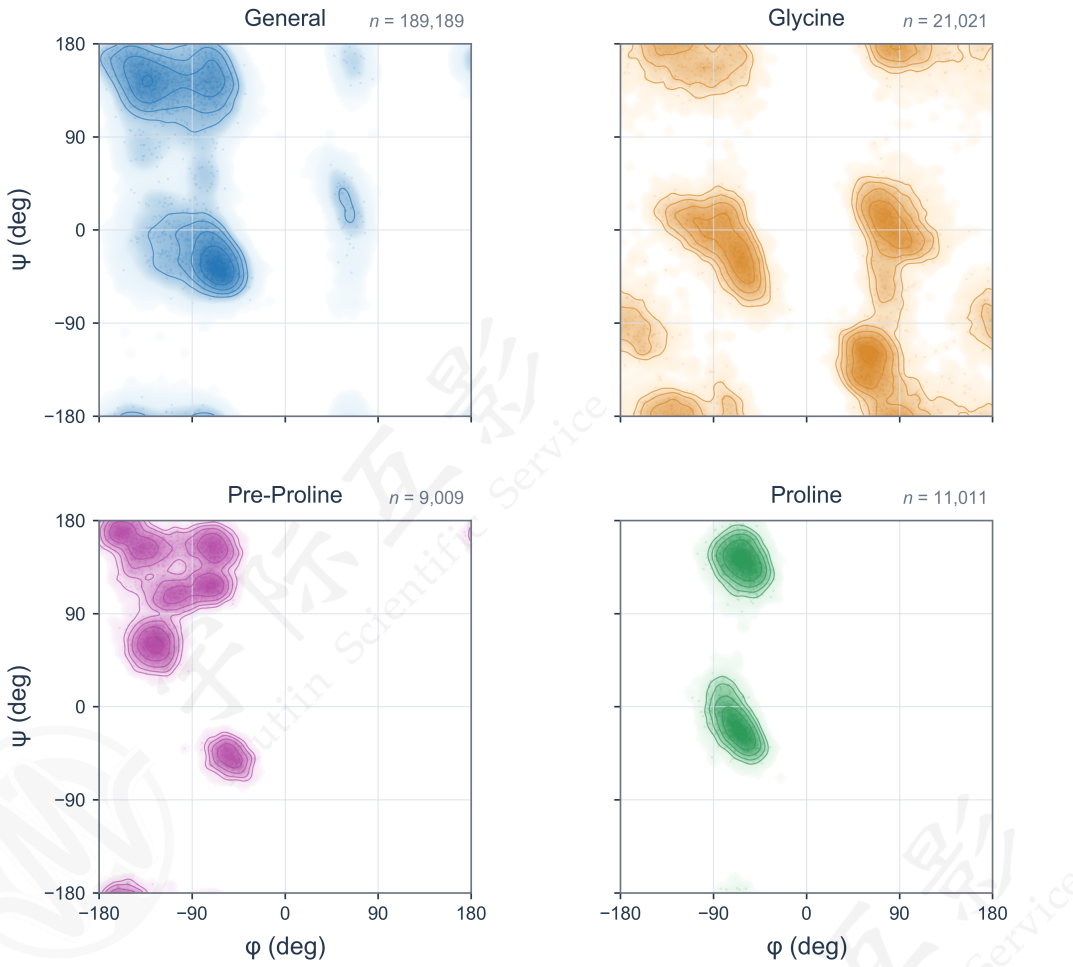


图 7: 按 General、Glycine、Pre-Proline 和 Proline 分类的 Ramachandran 二面角分布图；对数密度着色呈现高频构象盆，稀疏采样点保留低频采样信息。

3.6.4 氢键寿命与界面接触持续性

氢键寿命分析识别 16 对候选氢键。最高占有率为 30.8%，最长连续保持时间为 1.40 ns。各氢键均未在全部采样帧中持续存在。占有率和连续寿命描述各氢键的持续性；两个极值可能来自不同原子对，连续寿命受轨迹采样间隔影响。

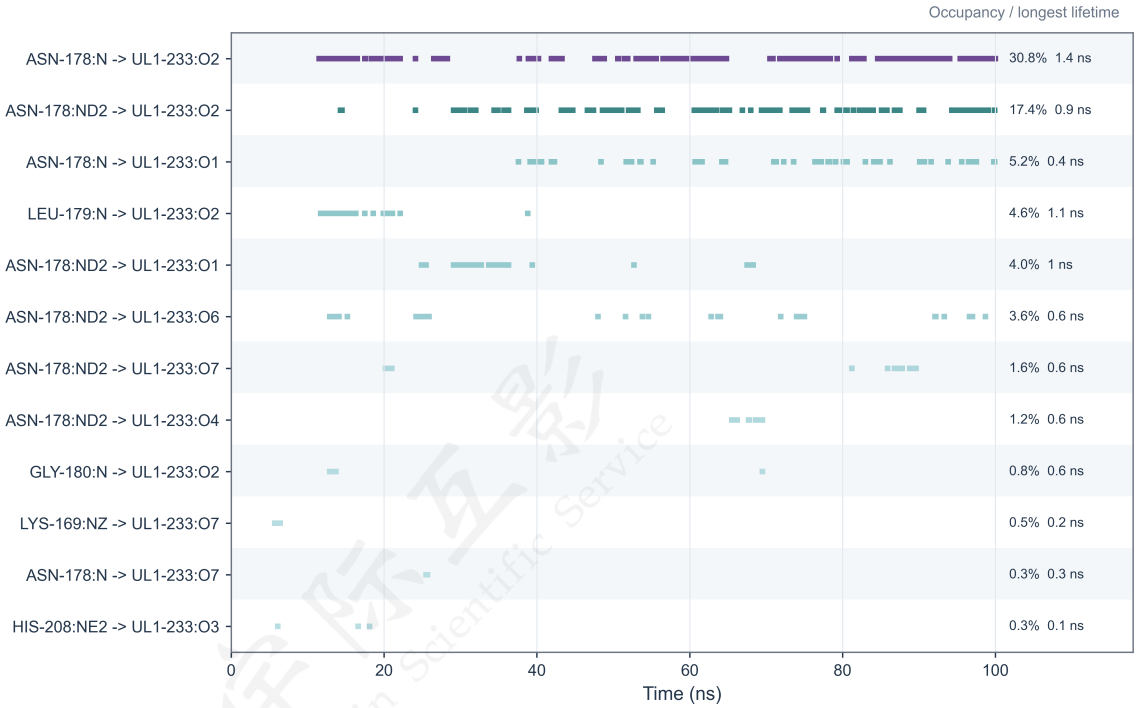


图 8: 蛋白-配体关键氢键的轨迹占有率与最长连续寿命。两项指标共同反映极性接触的持续性，标记颜色由浅青至紫表示占有率由低到高。

界面接触分析以 0.45 nm 为距离阈值，在 1,001 个轨迹帧中识别到 21 对非零占有接触。占有率最高的接触包括 ASP-82–UL1-233（100.0%）、CYS-166–UL1-233（100.0%）、HIS-147–UL1-233（100.0%）、HIS-208–UL1-233（100.0%）、HIS-78–UL1-233（100.0%）。高占有率说明相应残基与配体在所用距离判据下长期邻近，但不自动等同于特定氢键、盐桥或疏水作用；化学类型仍以代表构象 2D/3D 相互作用和原子几何为准。

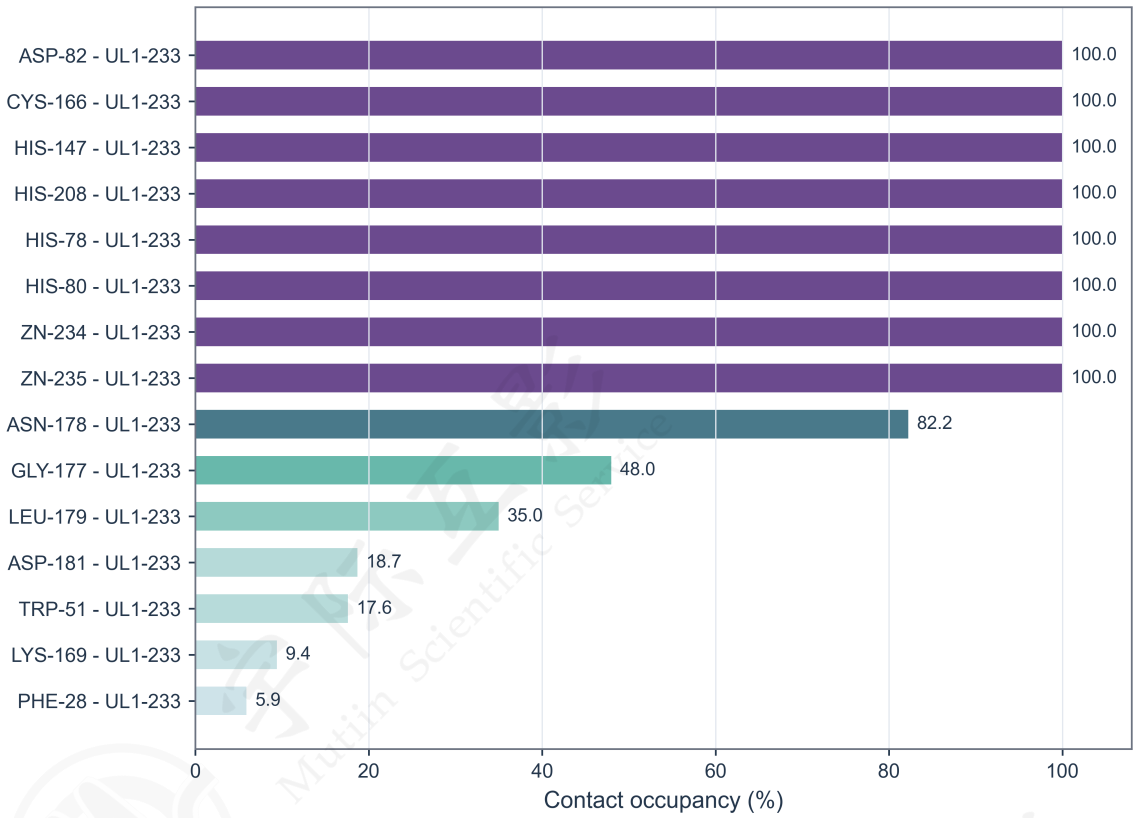


图 9: 关键蛋白-配体接触对的轨迹占有率。图中最多展示占有率最高的 15 项，条形颜色由浅青至紫表示占有率由低到高。

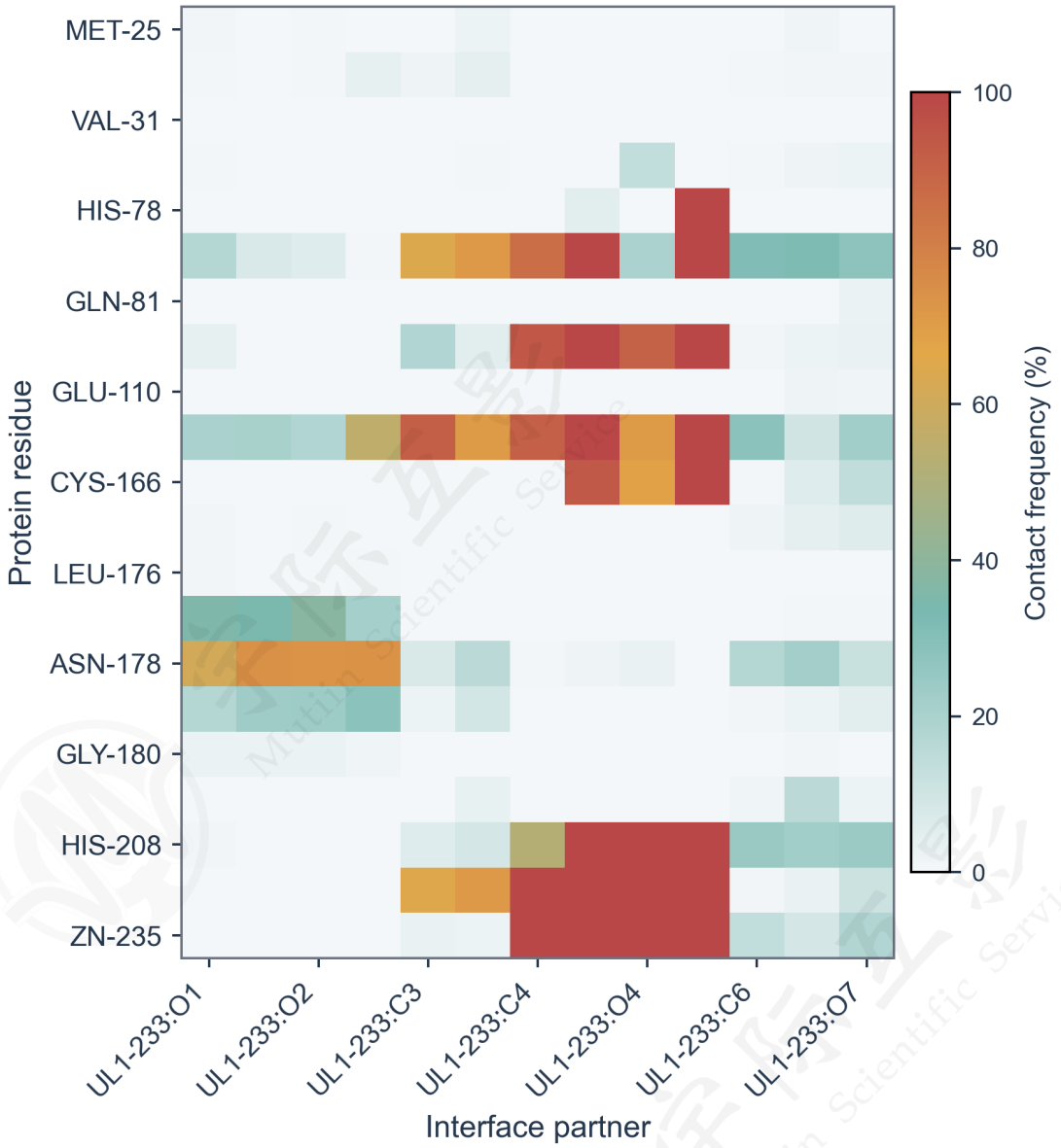


图 10: 蛋白-配体界面接触频率热图。颜色强度用于比较不同残基接触在轨迹中的持续程度。

补充分析显示，二级结构组成、集体运动、主链几何和界面接触提供的是互补而非相互替代的证据。其中，接触占有率用于描述配体在所分析轨迹中与受体残基的邻近关系；氢键寿命则反映极性作用的持续性。这些观察用于描述本次轨迹，不外推为确定的功能机制。

4 结论

本研究完成了 100 ns 显式溶剂分子动力学模拟，并以 90–100 ns 作为末段稳定性、构象聚类和代表姿态分析窗口。该窗口内受体主链 RMSD 为 0.259 ± 0.017 nm，Rg 为 1.677 ± 0.005 nm，SASA 为 108.004 ± 1.382 nm²。Rg 与 SASA 的相对波动均较小，说明

蛋白整体紧致程度和溶剂暴露在该窗口内保持稳定。分析窗口内蛋白-配体氢键数为 0.898 ± 0.699，单对氢键最高占有率为 30.8%，表明各氢键并非在全部采样帧中持续存在。代表构象共识别到 3 项相互作用，主要为氢键 1 项、金属配位 2 项，涉及 ASN-178、ZN-234、ZN-235。在当前时间窗口和聚类截断值下，主簇包含 1001 帧，占 100.0%；这支持末段窗口内的局部构象同质性，但不能单独证明全程收敛或唯一全局构象态。受体与配体通过金属配位键连接，常规分离态 MM/PBSA 不适用于该成键模型，因此未计算结合能。

综合轨迹稳定性、聚类和代表构象结果，本次模拟给出了该结合姿态在所研究时间尺度内的结构证据，这些结构指标不能直接定量给出结合强度或实验亲和力。

上述结论限定于本次单条 100 ns 轨迹、90–100 ns 分析窗口；聚类结论依赖所用时间窗口和截断值，若用于亲和力排序或机制判断，仍需独立重复轨迹及实验数据验证。

补充结构动力学结果与常规轨迹指标共同描述本次单条轨迹中的构象行为；其用途是完善结构解释，而不是把相关性图或单次采样结果等同于功能机制或实验因果关系。

金属中心位于所研究的配体结合口袋；MCPB.py 成键模型保持预设金属–供体连接，轨迹中的配位几何反映该拓扑下的局部涨落，不能据此验证其他配位态或配体交换。

A 附录：附件文件说明

本报告随任务结果提供下列小分子-蛋白动力学核心附件。附件名称及内容如下。

表 4: 附件文件说明

附件文件	内容说明	建议用途
protein_ligand_md_report.pdf	本次分子动力学模拟分析的正式 PDF 报告。	阅读、归档、提交
protein_ligand_md_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
md_representative_complex_structure.pdb	末段分析窗口聚类后提取的代表性复合物结构文件。	结构复查、PyMOL 或 ChimeraX 可视化
md_run_input.tpr	GROMACS 运行输入文件，记录生产期模拟所需的拓扑、参数和坐标信息。	复算、参数复核
md_system_structure.gro	生产期体系结构文件。	体系结构检查、GROMACS 可视化

附件文件	内容说明	建议用途
ion_retention_audit.xlsx	客户所选离子的保留、删除及结构角色审计表。	离子处理与结构角色复核
md_interaction_3d.png	代表构象的三维综合图，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
md_interaction_3d_partA.png	结合区域整体视图，用于观察配体相对受体结构的位置。	结构定位、口袋说明
md_interaction_3d_partB.png	代表构象的三维相互作用局部图，展示结合口袋中的关键残基和作用细节。	局部相互作用解读
md_interaction_2d.png	基于 LigPlot+/相互作用识别结果生成的二维相互作用图，用于快速查看氢键、疏水作用和 π 相关作用等。	快速查看关键相互作用
pymol_session.pse	PyMOL 会话文件，保留推荐观察角度、颜色和显示状态。	继续调整视角或导出图片
md_analysis_data.xlsx	分子动力学标准指标与结构动力学数据的汇总工作簿。	数据复核、重新制图
md_dynamics_summary.png	RMSD、RMSF、Rg、氢键数量、SASA 和自由能面的标准六联图。	动力学概览、汇报展示
md_dssp_heatmap.png	DSSP 二级结构热图。	二级结构稳定性复核
md_dccm.png	C-α 动态互相关矩阵图。	协同运动分析
md_hbond_lifetime.png	关键氢键寿命图。	持续氢键筛选
md_pca.png	主成分分析投影图。	构象采样评估
md_ramachandran.png	Ramachandran 主链二面角密度图。	主链构象分布分析
md_contact_occupancy.png	关键界面接触占有率图。	稳定接触热点筛选
md_contact_frequency_heatmap.png	界面接触频率热图。	残基-配体接触网络复核

参考文献

[1] 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).

[2] 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).

[3] Morris, G. M. *et al.* Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry* **30**, 2785–2791 (2009).

[4] Li, P. & Merz, K. M. MCPB.py: A python based metal center parameter builder. *Journal of Chemical Information and Modeling* **56**, 599–604 (2016).

[5] Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *The Journal of Physical Chemistry* **97**, 10269–10280 (1993).

[6] Maier, J. A. *et al.* ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99sb. *Journal of Chemical Theory and Computation* **11**, 3696–3713 (2015).

[7] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **79**, 926–935 (1983).

[8] Berendsen, H. J. C., van der Spoel, D. & van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* **91**, 43–56 (1995).

[9] Abraham, M. J. *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).