

基于 HERB 数据库草药成分的新型冠状病毒多靶点新型抑制剂虚拟筛选

沈文静, 饶青青, 王 星, 戚懿予*, 杨胜祥*

浙江农林大学化学与材料工程学院, 浙江 杭州 311300

摘要: 目的 通过计算机虚拟筛选从 HERB 数据库草药成分中筛选出靶向新型冠状病毒主蛋白酶 (main protease, Mpro)、木瓜样蛋白酶 (papain-like protease, PLpro) 和受体结合域 (receptor binding domain, RBD) 的新型潜在抑制剂并阐述其相互作用机制。方法 以从 HERB 数据库中获得天然化合物为筛选对象, 综合运用药效团模型、分子对接、ADME-T 预测、分子动力学模拟和分子力学泊松-玻尔兹曼表面积 (molecular mechanics Poisson-Boltzmann surface area, MM-PBSA) 等计算方法, 获取靶向新型冠状病毒中 Mpro、PLpro 和 RBD 等多个靶点的潜在抑制剂, 并采用酶活性抑制实验评价其植物提取物对 Mpro 的抑制效果。结果 化合物 HBIN006454、HBIN032286 和 HBIN031053 对 SARS-CoV-2 的 Mpro、PLpro 和 RBD 靶点均具有较高的结合亲和力。其中, 巴戟天中的 HBIN031053 具有更好的类药性和药动学特性, 且与各靶点所形成的复合物在分子动力学模拟中均表现出良好的稳定性。MM-PBSA 分析结果进一步揭示了在 HBIN031053-Mpro、HBIN031053-PLpro 以及 HBIN031053-RBD 的结合中发挥关键作用的氨基酸残基, 并深入阐述了其抑制病毒蛋白酶活性的机制。酶活性抑制实验表明, 巴戟天提取物对 Mpro 有抑制效果。结论 天然化合物 HBIN031053 可作为靶向 SARS-CoV-2 中 Mpro、PLpro 和 RBD 等靶点的潜在新型抑制剂。

关键词: HERB 数据库; 新型冠状病毒; 多靶点; 药效团; 分子对接; 分子动力学模拟; 巴戟天

中图分类号: R284.1 **文献标志码:** A **文章编号:** 0253-2670(2024)11-3609-11

DOI: 10.7501/j.issn.0253-2670.2024.11.005

Virtual screening of novel multi-target inhibitors against SARS-CoV-2 based on herbal ingredients of HERB database

SHEN Wenjing, RAO Qingqing, WANG Xing, QI Yiyu, YANG Shengxiang

College of Chemical and Materials Engineering, Zhejiang A&F University, Hangzhou 311300, China

Abstract: Objective To screen novel potential inhibitors against SARS-CoV-2 main protease (Mpro), papain-like protease (PLpro), and receptor binding domain (RBD), and elucidate their interaction mechanisms through computer virtual screening from herbal ingredients of the HERB database. **Methods** Natural compounds obtained from the HERB database were prepared for screening objects. Computational methods involving pharmacophore modeling, molecular docking, ADME-T prediction, molecular dynamics (MD) simulation and molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) were comprehensively used to obtain potent inhibitors against SARS-CoV-2 Mpro, PLpro, and RBD. The enzyme activity inhibition assay was used to evaluate the inhibitory effect of the plant extracts on Mpro. **Results** Compounds HBIN006454, HBIN032286, and HBIN031053 all showed high binding affinities against SARS-CoV-2 Mpro, PLpro, and RBD. Among them, HBIN031053 from *Morindae Officinalis Radix* had better drug-likeness and pharmacokinetic properties, and the complexes formed with each target demonstrated favorable stability in MD simulations. The results of MM-PBSA analysis revealed the amino acid residues that played a key role in the binding of HBIN031053-Mpro, HBIN031053-PLpro and HBIN031053-RBD, and further elaborated on their mechanism of inhibiting viral protease activity. Enzyme activity inhibition assay showed that the plant extract of *Morindae Officinalis Radix* had an inhibitory effect on Mpro. **Conclusion** The natural compound HBIN031053 can be used as a novel potential inhibitor targeting SARS-CoV-2 Mpro, PLpro, and RBD.

收稿日期: 2024-02-26

基金项目: 国家自然科学基金项目 (32371536); 国家自然科学基金项目 (32301259); 国家自然科学基金项目 (32271527)

作者简介: 沈文静 (1997—), 硕士研究生。E-mail: isshehwj@163.com

*通信作者: 戚懿予, 讲师, 从事药物化学与药理学研究。E-mail: qyy@zafu.edu.cn

杨胜祥, 教授, 博士生导师, 从事天然药物化学研究。E-mail: shengxiangyang2000@163.com

Key words: HERB database; SARS-CoV-2; multi-target; pharmacophore; molecular docking; molecular dynamics simulation; *Morindae Officilis Radix*

由严重急性呼吸综合征冠状病毒 2 型 (severe acute respiratory syndrome coronavirus 2, SARS-CoV-2) 引起的新型冠状病毒肺炎 (coronavirus disease 2019, COVID-19) 的爆发给人类的生命健康安全造成了严重的危害。截至 2024 年 2 月, 世界卫生组织公布的 COVID-19 全球确诊病例已超过 7.7 亿例, 并导致超过 700 万例死亡。随着 SARS-CoV-2 的不断演变, 数据仍在持续增长。然而, 受到广泛认可的治疗 SARS-CoV-2 感染的药物还未出现。

SARS-CoV-2 是一种有包膜的正向单链 RNA 病毒, 其基因组可编码 4 种结构蛋白和 16 种非结构蛋白 (non-structural proteins, NSPs)^[1]。前者包括棘突蛋白 (spike protein, S)、包膜蛋白 (envelope protein, E)、膜蛋白 (membrane protein, M) 和核衣壳蛋白 (nucleocapsid protein, N), 在冠状病毒中较为常见; 后者 (NSP1-NSP16) 则经 SARS-CoV-2 主蛋白酶 (main protease, Mpro) 和木瓜样蛋白酶 (papain-like protease, PLpro) 裂解前体蛋白得到。Mpro 和 PLpro 在病毒复制中发挥关键作用, 已成为抗 SARS-CoV-2 药物研发的热点靶标。Vuong 等^[2]发现 GC376 通过结合 Mpro 抑制其活性, 进而有效抑制 SARS-CoV-2 的复制。Fu 等^[3]发现, 使用 GRL0617 抑制 PLpro 同样具有阻断 SARS-CoV-2 病毒复制的能力。Mpro 的结构在冠状病毒中高度保守, 在已知的 SARS-CoV-2 变体中突变频率较低, 将其作为靶标更有利于应对病毒变体层出不穷的现状^[4]。此外, SARS-CoV-2 通过其表面 S 蛋白的受体结合域 (receptor binding domain, RBD) 识别细胞表面的血管紧张素转化酶 2 (angiotensin-converting enzyme, ACE2), 从而侵入宿主的正常细胞^[5]。因此, 阻断 RBD 与 ACE2 的结合也为 SARS-CoV-2 抑制剂的研发提供了一个重要思路。综上所述, 同时靶向 Mpro、PLpro 和 RBD, 进而阻止病毒入侵和抑制病毒复制将为抗 SARS-CoV-2 感染药物的研发提供新的策略。

中药是中国传统医学 (traditional Chinese medicine, TCM) 的重要组成部分, 包含大量可利用的天然化合物。迄今为止, 一部分来自中草药的活性成分已被应用于 COVID-19 的治疗当中, 如穿心莲内酯、苦参碱、甘草酸等^[6], 表明中草药在开

发新型 SARS-CoV-2 抑制剂方面具有广阔的前景。HERB 本草组鉴数据库 (<http://herb.ac.cn/>) 汇集了来自 7 000 余种中草药的数万个天然化合物, 对现代药物研发具有重要意义^[7]。本研究旨在借助计算机虚拟筛选方法, 从大量天然化合物中获取针对 SARS-CoV-2 多靶点的新型潜在抑制剂, 为推进治疗 SARS-CoV-2 感染药物的研发提供依据。

1 材料

新型冠状病毒 Mpro/3CLpro 抑制剂筛选试剂盒 (增强型, 货号 P0315S) 购自上海碧云天生物技术有限公司, 巴戟天提取物 (分析纯, 批号 2401168) 购自南京道斯夫生物科技有限公司。

2 方法

2.1 靶标蛋白准备与药效团构建

从 RCSB PDB 数据库 (<https://www.rcsb.org/>) 下载 SARS-CoV-2 Mpro (PDB ID: 7VH8, 0.159 nm)、PLpro (PDB ID: 7JRN, 0.248 nm) 和 RBD (PDB ID: 6LZG, 0.25 nm) 的蛋白质结构^[4,8-9], 并使用 PyMol 对蛋白质结构进行预处理。使用 AncPhore 提取蛋白质结构的药效团信息并将其上传至 Pharmit 平台 (<https://pharmit.csb.pitt.edu/>) 以构建药效团模型^[10-11]。

2.2 药效团验证与筛选

参照 Halimi 等^[12]的方法对药效团进行验证, 评价方式借助如下指标: (1) 灵敏度 (sensitivity); (2) 特异性 (specificity); (3) 活性化合物命中率 (yield of actives, YA); (4) 富集因子 (enrichment factor, EF); (5) 命中优度 (goodness of hit, GH)。上述指标按照公式的方法进行计算。

$$\text{sensitivity} = \text{Ha}/A \times 100 \quad (1)$$

$$\text{specificity} = \text{ture negative}/D \times 100 \quad (2)$$

$$\text{YA} = \text{Ha}/\text{Ht} \times 100 \quad (3)$$

$$\text{EF} = \text{YA}/(A/D) \times 100 \quad (4)$$

$$\text{GH} = [\text{Ha}(3A + \text{Ht})/4\text{HtA}] [1 - (\text{Ht} - \text{Ha})/(D - A)] \quad (5)$$

Ha 指活性化合物命中数 (active hits), ture negative 指未命中的诱饵化合物数, Ht 指总命中数 (total hits), A 指活性化合物数 (actives), D 指诱饵化合物数 (decoys)

通过文献调研获得经过活性实验验证对靶标蛋白具有抑制效果的小分子, 并使用 DUD.E 平台 (<http://dude.docking.org/>) 生成一系列诱饵化合物^[13]。

将准备好的活性化合物和诱饵化合物作为 2 个独立的库上传至 Pharmit 网站, 并使用构建的药效团模型进行筛选, 以确定能够产生最佳结果的药效团模型。将从 HERB 数据库获得的共计 33 495 个天然化合物的 SMILES 式上传到 Pharmit 网站, 并使用验证过的药效团进行初步筛选。

2.3 分子对接

使用 AutoDock Vina 对经过药效团筛选的化合物与 SARS-CoV-2 Mpro、PLpro 和 RBD 靶点分别进行分子对接^[14]。Mpro 和 PLpro 的活性位点盒子参照了所使用的共晶结构中原配体的位置, RBD 的活性位点盒子参考了所使用的共晶结构中 RBD 与 ACE2 的结合区域。其中, Mpro 和 PLpro 的对接盒子大小均设置为 2.6 nm×3.0 nm×2.6 nm, RBD 的对接盒子大小设置为 2.6 nm×3.8 nm×2.6 nm。3 个靶标蛋白(Mpro、PLpro 和 RBD)的对接盒子中心位点分别为(-19.952, 17.263, -23.280)、(16.191, -10.650, 29.451)和(-33.550, 28.300, 5.080)。为了更好地显示和分析对接构象, 使用 PyMol 和 LigPlot⁺ 分别以 3D 和 2D 模式对分子对接结果进行可视化^[15]。

2.4 类药性评估和 ADME-T 预测

使用 Lipinski 类药性五规则对分子对接中综合表现较好的天然化合物进行类药性评估。使用吸收、分布、代谢、排泄和毒性(absorption, distribution, metabolism, excretion, and toxicity, ADME-T)参数评价化合物的药动学特性和毒性^[16]。利用 SwissADME (<http://www.swissadme.ch/>) 和 Protox II (https://tox-new.charite.de/protox_II/) 预测所选化合物的类药性和 ADME-T 参数^[17-18]。

2.5 分子动力学模拟

使用 Gromacs 对筛选出的化合物与 SARS-CoV-2 Mpro、PLpro 和 RBD 分别进行 100 ns 的分子动力学模拟^[19]。每个分子动力学模拟体系均采用 Charmm36-jul2021 力场和 CHARMM-modified TIP3P 水模型^[20]。采用 pdb2gmx 模块生成蛋白质拓扑文件。采用 SwissParam (<https://www.swissparam.ch/>) 生成化合物拓扑文件^[21]。每个蛋白质被放置在立方体水盒子的中心并确保其距离边界不小于 1 nm。为了中和每个模拟体系的电荷, 在 Mpro 体系中加入 4 个 Na⁺, 在 PLpro 体系中加入 2 个 Cl⁻, 在 RBD 体系中加入 2 个 Cl⁻。接着, 对体系进行能量最小化并依次进行 100 ps 的 NVT 和 NPT 平衡, 使体系处于 300 K 的温度和 100 kPa 的压强下。为了进一步

阐明复合物形成过程中的能量贡献, 使用 gmx_MM-PBSA 程序对分子动力学模拟得到的轨迹进行 MM-PBSA 计算^[22]。

2.6 酶活性抑制实验

基于 SARS-CoV-2 中 Mpro 在病毒生命周期中发挥的关键作用, 以及其结构的保守性和较低的突变频率, 选择 Mpro 进行酶活性抑制实验以评估筛选结果的生物活性。

参照试剂盒使用说明进行实验。将巴戟天提取物溶解在 DMSO 溶液中以配制不同浓度的样品溶液, 并根据样品数量使用分析缓冲液和 2019-nCoV Mpro/3CLpro 配制适量检定试剂(分析缓冲液与 2019-nCoV Mpro/3CLpro 的比例为 90:1)。在 96 孔黑板中依次加入 91 μ L 的检测试剂和 5 μ L 的样品溶液, 混匀, 再将 4 μ L 的底物快速加入各孔, 混匀。以加入 91 μ L 的检定试剂、5 μ L 的 DMSO 溶液和 4 μ L 的底物的孔为 100%酶活性对照, 以 91 μ L 的分析缓冲液、5 μ L 的 DMSO 溶液和 4 μ L 的底物的孔为空白对照, 每个样品进行 2 个重复孔的检测及 3 次重复实验。37 $^{\circ}$ C 避光孵育 5 min 后, 使用多功能酶标仪进行荧光测定(激发波长为 325 nm, 发射波长为 393 nm)。将每个空白对照孔和样品孔测得的相对荧光单位(relative fluorescence unit, RFU)的平均值分别记录为 RFU_{空白对照}、RFU_{100%酶活性对照}和 RFU_{样品}。为排除植物提取物本身荧光对实验结果的影响, 增设了相应的不同浓度的样品溶液孔以测定其荧光值排除影响(RFU_{植物提取物})。每个样品的抑制率计算方法如下。

$$\text{抑制率} = \frac{\text{RFU}_{100\% \text{酶活性对照}} - (\text{RFU}_{\text{样品}} - \text{RFU}_{\text{植物提取物}})}{(\text{RFU}_{100\% \text{酶活性对照}} - \text{RFU}_{\text{空白对照}})}$$

3 结果

3.1 药效团的构建与筛选

构建药效团模型的评价指标及验证结果(表 1)显示, 靶向 SARS-CoV-2 Mpro、PLpro 和 RBD 所构建的药效团模型的 YA 值均不低于 0.8, 说明其具有较高的活性化合物命中率。EF 将命中率与验证数据库中活性化合物数、诱饵化合物数相关联, 其值越高, 表明药效团模型越优。其中, 靶向 RBD 的药效团模型的 EF 值最高, 其次是 PLpro 和 RBD。GH 介于 0~1, 药效团模型越好, 其 GH 值越接近于 1。如表 1 所示, PLpro 的 GH 值略低(0.787), 而 Mpro 和 RBD 的 GH 值相对较高, 分别为 0.849 和 0.874。经过验证的药效团模型参数信息如表 2 所示。其中,

表 1 SARS-CoV-2 Mpro、PLpro 和 RBD 的药效团验证结果

Table 1 Results of pharmacophore validation for SARS-CoV-2 Mpro, PLpro, and RBD

评价指标	Mpro	PLpro	RBD
总化合物数	790	719	785
活性化合物数	32	26	30
命中活性化合物数	26	20	25
未命中活性化合物数	6	6	5
诱饵化合物数	758	693	755
未命中诱饵化合物数	754	688	752
命中诱饵化合物数	4	5	3
总命中数	30	25	28
灵敏度	0.813	0.769	0.833
特异性	0.995	0.993	0.996
活性化合物命中率	0.867	0.800	0.893
富集因子	21.396	22.123	23.363
命中优度	0.849	0.787	0.874

表 2 经过验证的药效团模型参数

Table 2 Validated pharmacophore model parameters

靶标	药效团	坐标		
		x	y	z
Mpro	HD	-21.018 0	13.341 0	-19.814 9
	HA	-18.451 7	12.332 9	-22.435 1
	HY	-21.035 7	19.229 7	-30.361 8
PLpro	HD	10.480 4	-5.455 9	33.985 6
	HD	5.596 8	-2.902 7	32.488 5
	HA	7.784 1	-3.031 7	33.955 4
RBD	HD	-38.040 8	21.260 2	2.526 5
	HA	-36.165 8	23.224 4	4.758 6
	HA	-26.963 9	23.455 2	1.220 1

HA 表示氢键受体 (hydrogen bond acceptors, H-bond acceptors), HD 表示氢键供体 (H-bond donors), HY 表示疏水作用 (hydrophobic interaction)。使用经过验证的药效团模型对 33 495 个来自 HERB 数据库的天然化合物进行初步筛选, 得到了 899 个满足表 2 中 SARS-CoV-2 Mpro、PLpro 和 RBD 药效团模型参数的天然化合物。这些通过药效团模型筛选的化合物将进行进一步的分子对接研究。

3.2 基于 Mpro、PLpro 和 RBD 的分子对接

首先, 为验证对接参数的可靠性, 选择文献报道过的 3 个 SARS-CoV-2 小分子抑制剂 (PF-07321332^[23]、GRL0617^[8] 和 glycyrrhetic acid^[24]) 分别与其对应靶标 (Mpro、PLpro 和 RBD) 进行对接。结果显示, PF-07321332 与 Mpro、GRL0617 与 PLpro、glycyrrhetic acid 与 RBD 的

接结合能分别为-39.748、-41.003、-33.890 kJ/mol, 均<-29.288 kJ/mol, 表明对接参数可靠。药效团模型筛选得到的 899 个天然化合物分别与 SARS-CoV-2 的 3 个靶标进行分子对接的总体结果见图 1。靶向 Mpro、PLpro 和 RBD 靶点的对接能量均值分别为-31.380、-30.543 和-29.288 kJ/mol, 表明使用药效团模型进行筛选已筛选一部分与 SARS-CoV-2 多靶点结合较弱的化合物。

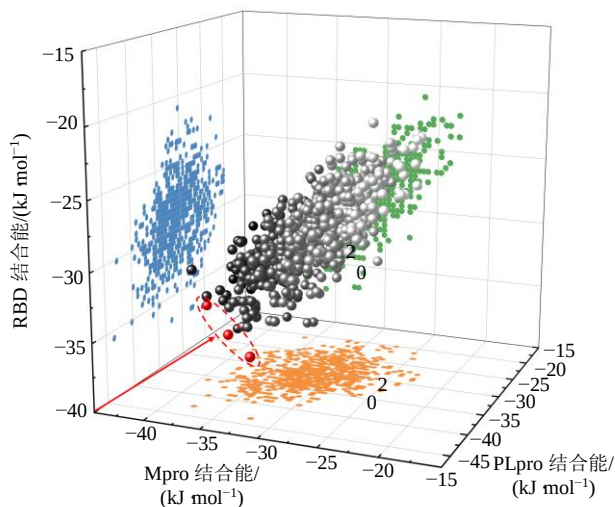


图 1 899 个天然化合物与 SARS-CoV-2 中 Mpro、PLpro 和 RBD 的对接结果

Fig. 1 Docking energy between 899 natural compounds and SARS-CoV-2 Mpro, PLpro and RBD

对图 1 进行分析可知, 越靠近左下角的化合物与 3 个靶点的结合亲和力越强。根据对 SARS-CoV-2 Mpro、PLpro 和 RBD 靶点综合表现较好的 3 个天然化合物 (HBIN006454、HBIN032286 和 HBIN031053, 图 2) 的具体的对接能量数据可知 (表 3), 3 个化合物均能与至少 2 个靶标有较好结合,

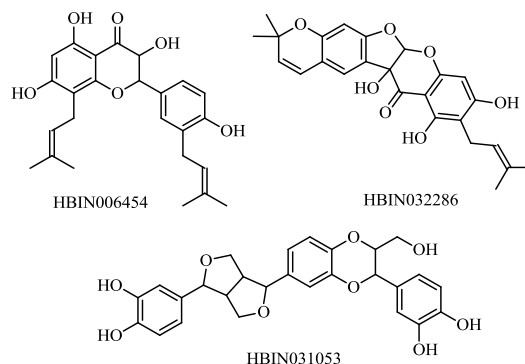
图 2 HBIN006454、HBIN032286 和 HBIN031053 的结构
Fig. 2 Structures of HBIN006454, HBIN032286, and HBIN031053

表 3 针对 SARS-CoV-2 Mpro、PLpro 和 RBD 靶点综合表现较好的化合物的对接能量

Table 3 Docking energy of compounds with good comprehensive performance against SARS-CoV-2 Mpro, PLpro and RBD

化合物	对接能量/(kJ mol ⁻¹)		
	Mpro	PLpro	RBD
HBIN006454	-32.217	-43.514	-35.564
HBIN032286	-35.564	-40.166	-35.146
HBIN031053	-39.330	-35.982	-34.309

其对接能量接近或优于文献报道过的 SARS-CoV-2 小分子抑制剂。

3.3 类药性和 ADME-T 研究

采用 Lipinski 类药性五规则 (molecular weight ≤ 500, H-bond acceptors ≤ 10, H-bond donors ≤ 5, MlogP ≤ 4.15) 评价分子对接综合表现最好的 3 个化合物的类药性^[25]。如表 4 所示, HBIN006454、HBIN032286 和 HBIN031053 均符合 Lipinski 规则。接着, 进一步评估了 HBIN006454、HBIN032286 和 HBIN031053 的药动学特性。如表 5 所示, 化合物 HBIN031053 具有最低的 iLOGP 值和最高的 ESOL log S 值, 表明其在人体内更容易被吸收。此外, 该化合物具有最少的细胞色素 P450 酶 (cytochrome P450, CYP450) 抑制和毒性, 表明其生物安全性更高。因此, 选择 HBIN031053 作为研究对象开展进一步的研究。

表 4 所筛选化合物的类药性评估

Table 4 Drug-likeness assessment of selected compounds

化合物	相对分子质量	氢键受体	氢键供体	脂水分配系数	Lipinski 违背项数
HBIN006454	424.49	6	4	2.01	0
HBIN032286	436.45	7	3	1.88	0
HBIN031053	494.49	9	5	0.58	0

3.4 HBIN031053 与 Mpro、PLpro 和 RBD 的分子对接结合构象分析

如图 3-A 所示, 在靶标蛋白 Mpro 包含催化二元组组氨酸(histidine, His)41 和半胱氨酸(cysteine, Cys) 145 在内的活性位点处, HBIN031053 与苏氨酸(threonine, Thr) 26 上的 N 原子之间形成了氢键。此外, 该化合物与 Thr25、亮氨酸(leucine, Leu) 27、His41、甘氨酸(glycine, Gly) 143、Cys145、蛋氨酸(methionine, Met) 165、Leu167、精氨酸(arginine, Arg) 188、谷氨酰胺(glutamine, Gln) 189 及 Gln192 等多个氨基酸残基之间存在疏水相互

作用, 这可能是其与 Mpro 结合能的主要来源。在与 PLpro 的对接中, HBIN031053 分别与 Gln269 和 Thr301 产生氢键(图 3-B)。其中, Gln269 属于 PLpro 催化位点附近的封闭环 2(blocking loop 2, BL2), 可参与 PLpro 底物识别并控制底物进入活性位点^[26]。因此, 当 HBIN031053 与 Gln269 形成氢键后, 可通过限制 PLpro 与其底物的结合来抑制该蛋白的酶活性。此外, HBIN031053 与结合位点附近的天冬氨酸(aspartic acid, Asp) 164、脯氨酸(proline, Pro) 247、Pro248、酪氨酸(tyrosine, Tyr) 264、Gly266、天冬酰胺(asparagine, Asn) 267 和 Tyr268 产生疏水作用, 有利于提升其与 PLpro 的结合亲和力。在与 RBD 的结合中, HBIN031053 分别与 Gln409、Gly496 和 Gln498 产生氢键, 且在结合位点处与氨基酸残基 Arg403、谷氨酸(glutamic acid, Glu) 406、赖氨酸(lysine, Lys) 417、Tyr449、Tyr453、Leu455、丝氨酸(serine, Ser) 494、Tyr495 和 Tyr505 产生疏水作用(图 3-C)。其中, Lys417、Tyr449、Leu455、Gln498 和 Tyr505 等在 RBD 与 ACE2 的紧密结合中发挥重要作用^[27], 而 HBIN031053 通过占据 RBD 与 ACE2 的结合区域, 抑制病毒 RBD 与人体 ACE2 的结合(图 4)。

3.5 HBIN031053 与 Mpro、PLpro 和 RBD 的分子动力学模拟

为了评价 HBIN031053 与 SARS-CoV-2 Mpro、PLpro 和 RBD 靶点结合的稳定性, 对复合物 HBIN031053-Mpro、HBIN031053-PLpro 和 HBIN031053-RBD 结合前后的蛋白分别进行分子动力学模拟, 并通过均方根偏差(root mean square deviation, RMSD)、均方根波动(root mean square fluctuation, RMSF)、回转半径(radius of gyration, Rg)以及氢键(H-bond)综合评判模拟过程中体系结构的稳定性。RMSD 的分析结果如图 5 所示。在 100 ns 的分子动力学模拟过程中, 3 个复合物体系中 Mpro、PLpro 和 RBD 的 RMSD 均值处于 0.16~0.20 nm, 且无明显波动。其中, HBIN031053 与 Mpro 和 RBD 的复合体的 RMSD 值在 50~100 ns 呈上升趋势, 但分别在 0.23 nm 和 0.20 nm 处趋于稳定, 并维持在一个较低的水平(< 0.25 nm)(图 5-A、C)。在 HBIN031053-PLpro 复合物的模拟过程中, 后 50 ns 与前 50 ns 的 RMSD 均值仅相差 0.000 3 nm, 表明其结合稳定性较好, 且与单独进行模拟的 PLpro 的 RMSD 基本保持一致(图 5-B)。

表 5 化合物的 ADME-T 特性预测
Table 5 ADME-T properties prediction of compounds

ADME-T 参数	HBIN006454	HBIN032286	HBIN031053
iLOGP	3.62	3.54	2.66
ESOL Log S	-6.01	-5.78	-4.44
GI absorption	High	High	High
BBB permeant	No	No	No
P-糖蛋白底物	No	No	No
CYP1A2 抑制剂	No	No	No
CYP2C19 抑制剂	No	Yes	No
CYP2C9 抑制剂	Yes	Yes	No
CYP2D6 抑制剂	No	No	No
CYP3A4 抑制剂	Yes	Yes	Yes
生物利用度评分	0.55	0.55	0.55
肝毒性	No	No	No
致癌性	No	No	No
免疫毒性	Yes	Yes	Yes
致突变性	No	Yes	No
细胞毒性	No	No	No

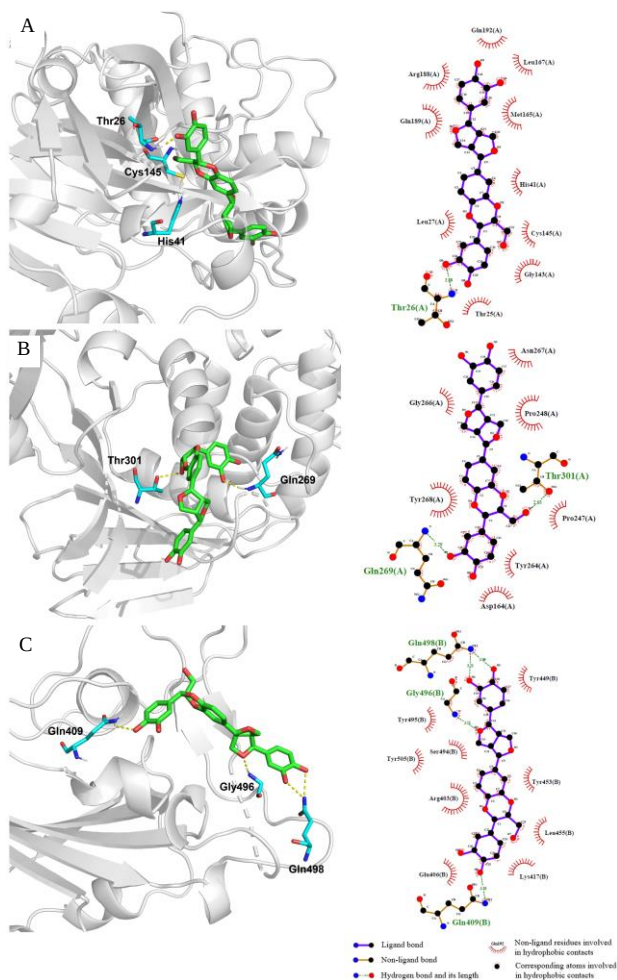
High 表示良好的胃肠道吸收表现, No 表示不能透过血脑屏障、非 P-糖蛋白底物、非细胞色素 P450 酶抑制剂或者没有毒性, Yes 表示能透过血脑屏障、是 P-糖蛋白底物、是细胞色素 P450 酶抑制剂或者没有毒性。
High indicates favorable gastrointestinal absorption (GI absorption), No indicates that it cannot penetrate the blood-brain barrier (BBB permeant), non-P-glycoprotein substrate (Pgp substrate) or non-cytochrome P450 (CYP450) enzyme inhibitor, Yes indicates that it can penetrate the blood-brain barrier, is a P-glycoprotein substrate, a CYP450 enzyme inhibitor or non-toxicity.

RMSF 的分析结果如图 6 所示, 3 个复合物体系的 RMSF 均值均低于 0.11 nm, 整体波动幅度较小。其中, HBIN031053-Mpro 体系中, 氨基酸残基 7~21、26~40、98~123、143~152、156~167 段的 RMSF 值较低 (<0.1 nm) (图 6-A)。在 HBIN031053-RBD 体系中, 氨基酸残基 347~361、374~381、391~443、448~474、488~495 段的 RMSF 值较低 (≤0.1 nm) (图 6-C)。上述结果与分子对接结果中产生氢键和疏水作用的主要区域相吻合。此外, 与单独进行模拟的 Mpro 相比, 复合物中的 Mpro 在氨基酸残基 47、50 和 190 处的 RMSF 达到峰值, 而在其他区域二者基本保持一致。值得注意的是, HBIN031053-PLpro 体系在氨基酸残基 202~219、233~253、295~307 处波动较小 (<0.1 nm), 而在存在氢键和疏水相互作用的 266~270 区域却波动较大 (>0.12 nm) (图 6-B), 这可能是由于氨基酸残基 266~270 属于 PLpro 上的由 6 氨基酸 (266~271) 组成的柔性环 BL2, 其较高的灵活性使得与 HBIN031053 相互作用时发生结构的变化。

最后, 3 个蛋白-配体复合物的 Rg 值的分析结果显示, 各个体系的 Rg 值整体呈平稳态势 (图 7), 表

明蛋白-配体复合物结构在模拟过程中始终保持稳定紧凑。其中, 在 HBIN031053-Mpro 体系和 HBIN031053-PLpro 体系中, 复合物的蛋白的 Rg 值与单独进行模拟的蛋白的 Rg 值基本一致。特别地, 对于复合物 HBIN031053-PLpro, 与 HBIN031053 形成复合物的 PLpro 的 Rg 均值 (2.35 nm) 低于单独模拟的 PLpro 的 Rg 均值 (2.36 nm), 表明与 HBIN031053 的结合使 PLpro 的结构更为紧凑。在 HBIN031053-RBD 体系中, 复合物中的 RBD 的 Rg 均值高于单独进行模拟的 RBD 的 Rg 值, 但整体处于较低水平 (1.86 nm)。

图 8 显示了在 100 ns 的分子动力学模拟过程中 HBIN031053 分别与靶标蛋白 Mpro、PLpro 和 RBD 产生的氢键数量。HBIN031053-Mpro 体系和 HBIN031053-PLpro 体系在 100 ns 的模拟期间均形成最少 0 个、最多 7 个氢键, 且在整个模拟过程中的氢键平均数量分别为 1.50 和 2.35 (图 8-A、B), 与分子对接结果保持一致。HBIN031053-RBD 体系中 (图 8-C), 形成最少 0 个、最多 4 个氢键, 整体氢键数量少于前两者, 这可能是由于其氢键强度较低。



A-HBIN031053-Mpro, B-HBIN031053-PLpro, C-HBIN031053-RBD, 下图同。

A-HBIN031053-Mpro, B-HBIN031053-PLpro, C-HBIN031053-RBD, same as Figs below.

图 3 HBIN031053 与 SARS-CoV-2 中 3 个靶点的对接构象及 2D 相互作用图

Fig. 3 Docked conformations of HBIN031053 and SARS-CoV-2 three targets along with 2D interaction plots

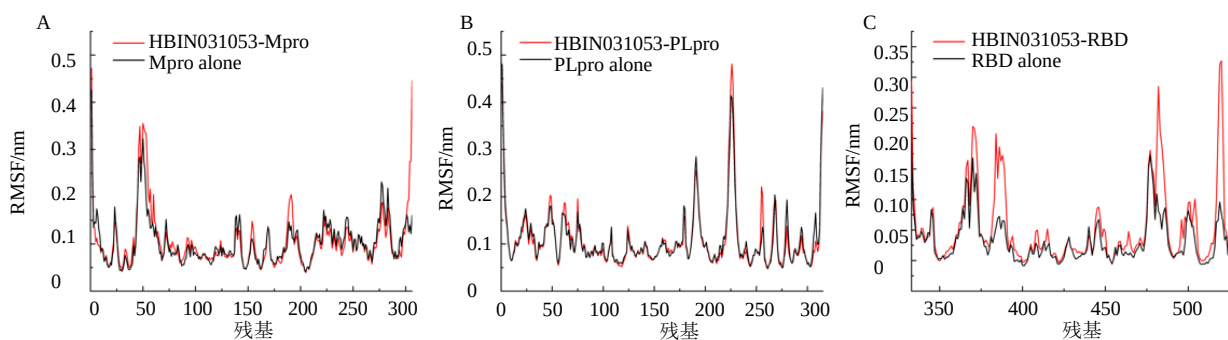


图 4 HBIN031053 占据 RBD 与 ACE2 的结合区域的示意图

Fig. 4 Schematic illustration of HBIN031053 occupying binding region of RBD and ACE2

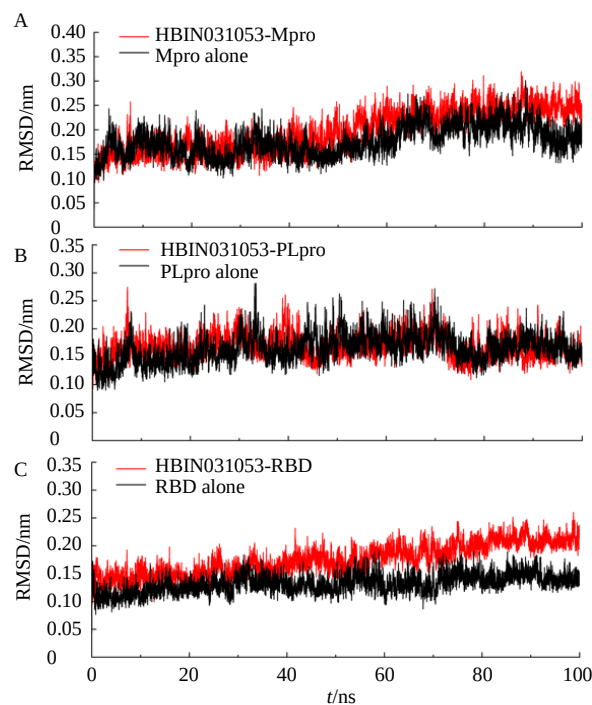


图 5 分子动力学模拟的 RMSD 分析

Fig. 5 RMSD analysis of MD simulation

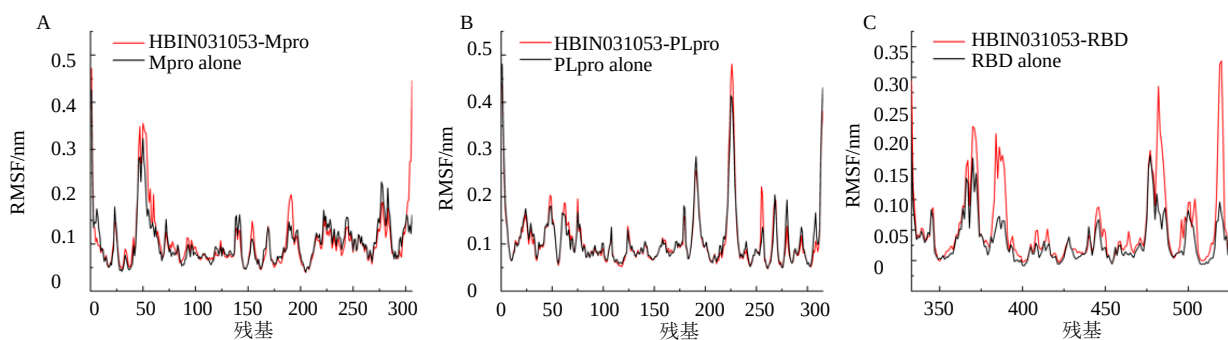


图 6 分子动力学模拟的 RMSF 分析

Fig. 6 RMSF analysis of MD simulation

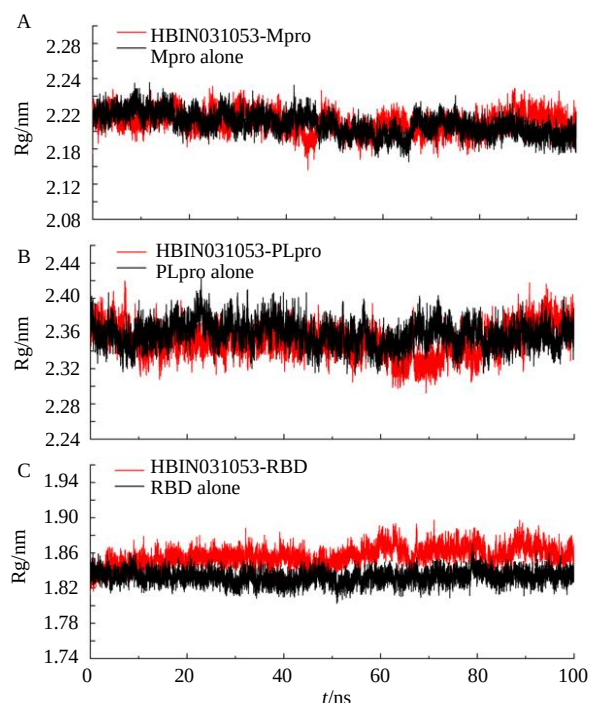


图 7 分子动力学模拟的 Rg 分析

Fig. 7 Rg analysis of MD simulation

3.6 HBIN031053 与 Mpro、PLpro 和 RBD 的 MM-PBSA 分析

采用 MM-PBSA 方法计算 HBIN031053 与 SARS-CoV-2 3 个靶标蛋白的结合能,以进一步分析其结合亲和力,并获得结合的关键氨基酸残基。由表 6 可知,HBIN031053-Mpro、HBIN031053-PLpro 和 HBIN031053-RBD 复合物均具有较低的结合自由能 (binding energy < -41.840 kJ/mol), 表明结合亲和力强,与分子对接结果一致。通过分析结合作用力可知,范德华力 (van der Waal energy) 是促进复合物形成的主要作用力,其次是静电势能 (electrostatic energy),最后是表面溶剂化能 (SASA energy)。相反,极性溶剂化能 (polar solvation energy) 则不同程度地阻碍了复合物的形成。尤其是复合物 HBIN031053-PLpro,其极性溶剂化能为 169.452 kJ/mol,严重地限制了 HBIN031053 与 PLpro 更紧密的结合。然而,由于非极性溶剂化能占主导作用,特别是范德华力和静电势能,使得三者的结合自由能产生差异。

根据复合物结合过程中氨基酸残基的能量贡献

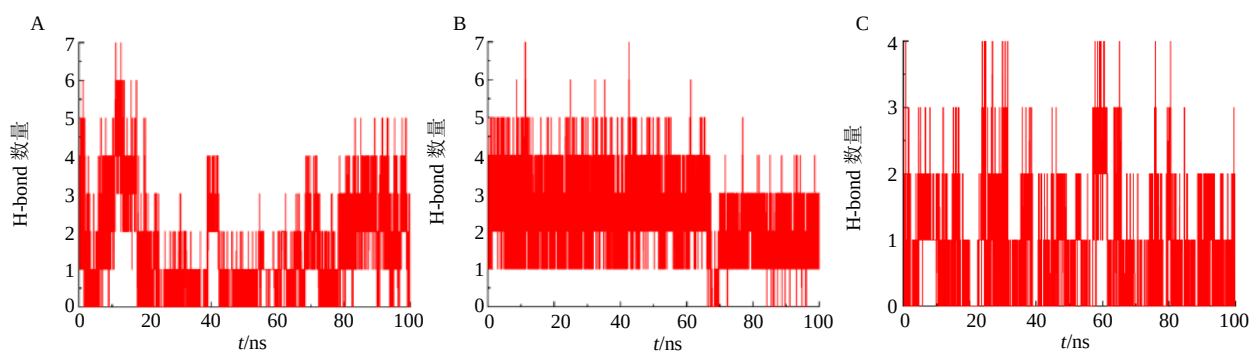


图 8 分子动力学模拟的 H-bond 分析

Fig. 8 H-bond analysis of MD simulation

表 6 SARS-CoV-2 Mpro、PLpro、RBD 与 HBIN031053 结合的能量贡献

Table 6 Energy contribution of SARS-CoV-2 Mpro, PLpro, RBD in binding with HBIN031053, respectively

复合物	范德华力/ (kJ mol ⁻¹)	静电势能/ (kJ mol ⁻¹)	极性溶剂化作用/ (kJ mol ⁻¹)	表面溶剂化作用/ (kJ mol ⁻¹)	结合能/ (kJ mol ⁻¹)
HBIN031053-Mpro	-71.756	-64.057	101.337	-9.372	-43.848
HBIN031053-PLpro	-120.708	-114.600	169.452	-14.016	-79.831
HBIN031053-RBD	-82.006	-75.187	115.939	-11.004	-52.258

分析结果可知,在 HBIN031053-Mpro 复合物中,Arg4(-1.088 kJ/mol)、Glu47(-2.259 kJ/mol)、Leu50(-1.423 kJ/mol)、Met165(-1.255 kJ/mol)、Gln189(-0.920 kJ/mol)、Thr190(-0.753 kJ/mol)、丙氨酸

(alanine, Ala) 191(-1.004 kJ/mol)等残基的能量值最低,有助于配体和蛋白质的结合(图 9-A)。其中,Glu47、Leu50 和 Thr190 与分子动力学模拟过程中 RMSF 产生波动峰值的区域对应。在 HBIN031053-

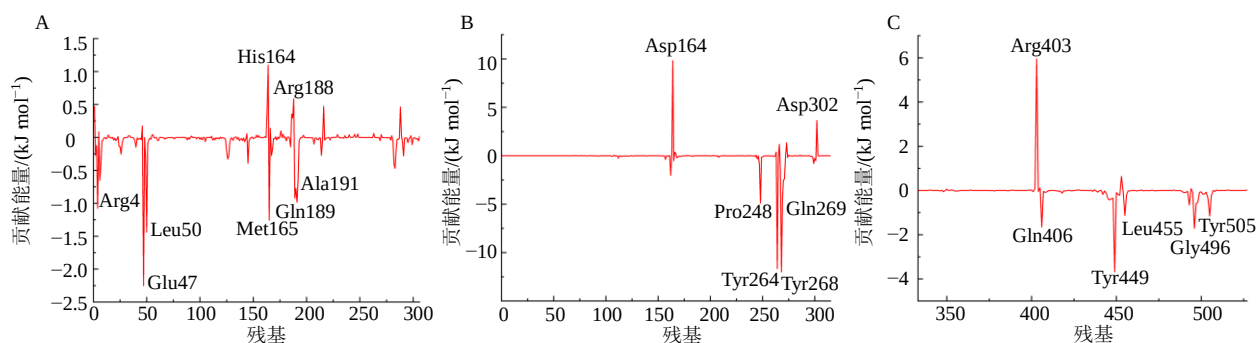


图 9 SARS-CoV-2 Mpro、PLpro、RBD 与 HBIN031053 结合中氨基酸残基的能量贡献

Fig. 9 Energy contribution of amino acid residues in binding of SARS-CoV-2 Mpro, PLpro, RBD to HBIN031053, respectively

PLpro 结合过程中, Pro248 (−4.895 kJ/mol)、Tyr264 (−11.673 kJ/mol)、Tyr268 (−12.008 kJ/mol) 和 Gln269 (−4.393 kJ/mol) 对复合物的稳定结合具有显著贡献(图 9-B)。其中, Tyr268 和 Gln269 同属于柔性环 BL2, 在分子动力学模拟过程中其 RMSF 值产生了较大的波动(图 6-B), 这可能有助于提升结合亲和力。此外, Gln269 在分子对接结合模式中与 HBIN031053 形成氢键, 与其对形成稳定复合物所做出的有利能量贡献 (−4.393 kJ/mol) 匹配。在 HBIN031053-RBD 结合过程中, Arg403 的贡献能量值为 5.941 kJ/mol, 但 Glu406 (−1.674 kJ/mol)、Tyr449 (−3.682 kJ/mol)、Leu455 (−1.130 kJ/mol)、Gly496 (−1.715 kJ/mol) 和 Tyr505 (−1.172 kJ/mol) 等抵消了前者的不利影响并促进了 HBIN031053 与 RBD 的稳定结合(图 9-C)。其中, Gly496 在分子对接结合模式中与 HBIN031053 形成氢键, Glu406、Tyr449、Leu455 和 Tyr505 与 HBIN031053 之间存在疏水相互作用, 进一步验证了其结合的促进作用。

3.7 酶活性抑制结果

鉴于天然化合物 HBIN031053 获取难度高, 根据 HERB 数据库中的信息, 采用其源自的中草药巴戟天的提取物对 SARS-CoV-2 中关键蛋白酶 Mpro 进行酶活性抑制实验, 用以评价其抑制效果。根据巴戟天提取物对 Mpro 的抑制曲线可知, 巴戟天提取物对 Mpro 酶活性的 IC_{50} 值为 5.36 mg/mL, 且随着浓度增加, 抑制效果逐渐增强, 并在 12.5 mg/mL 处达到 76.75% 的抑制率(图 10)。

4 讨论

SARS-CoV-2 感染已经成为困扰人类生命健康安全的一个长期问题。以 SARS-CoV-2 Mpro、PLpro 和 RBD 为靶点探索新型潜在 SARS-CoV-2 抑制剂, 可从阻断病毒入侵人体和病毒复制 2 方面抑制病毒

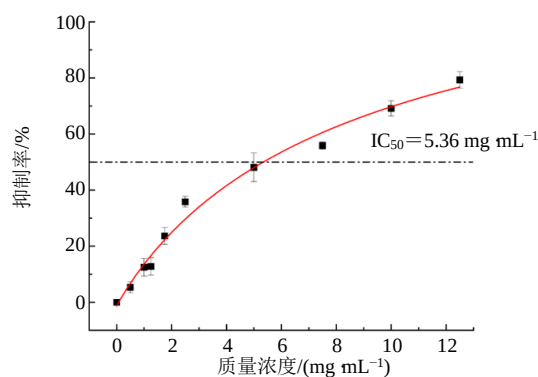


图 10 巴戟天提取物对 SARS-CoV-2 Mpro 的抑制曲线

Fig. 10 Inhibition curve of *Morinda officinalis Radix* extract against SARS-CoV-2 Mpro

感染, 以达到治疗 COVID-19 的目的。

本研究首先构建并验证了靶向 SARS-CoV-2 Mpro、PLpro 和 RBD 的药效团模型, 并使用该药效团模型对 HERB 数据库的草药成分进行初步筛选, 最终得到 899 个满足要求的天然化合物。随后, 通过分子对接评价上述化合物与 Mpro、PLpro 和 RBD 的对接分数和结合模式, 得到了对 3 个靶点均有较高结合亲和力的化合物 HBIN006454、HBIN032286 和 HBIN031053。Lipinski 类药性评估和 ADME-T 预测表明, HBIN031053 具有更好的类药性和药动学特性。进一步的分子动力学模拟结果显示, HBIN031053 与 Mpro、PLpro 和 RBD 的结合均非常稳定。

与此同时, MM-PBSA 分析进一步验证了各复合物的结合亲和力, 获得了发挥主导作用的能量组成, 并揭示了结合过程中发挥重要作用的氨基酸残基。在 HBIN031053-Mpro 中, Met165 主要产生疏水作用, 在分子动力学模拟过程中可稳定周围结构, 对 MM-PBSA 结合自由能做出了重要贡献, 可促进 HBIN031053 和 Mpro 的结合。此外, 在分子动力学

模拟过程中 RMSF 产生波动峰值的氨基酸残基 Glu47、Leu50 和 Thr190 在 MM-PBSA 结合能中同样做出了有利贡献。在 HBIN031053 与 PLpro 的相互作用中, 参与 PLpro 底物识别并控制底物进入活性位点的柔性环 BL2 是关键的结构片段。HBIN031053 与 BL2 中的 Gln269 和 266~268 分别形成氢键和疏水作用, 其中, Tyr268 和 Gln269 为 HBIN031053-PLpro 结合自由能的重要贡献者。因此, HBIN031053 通过与 BL2 形成多种相互作用, 阻碍其识别和控制底物, 从而抑制病毒蛋白酶活性。在与 RBD 形成的复合物中, HBIN031053 通过与 Tyr449、Leu455、Tyr505 等促进病毒 RBD 与宿主 ACE2 结合的关键氨基酸残基发生相互作用, 在一定程度上可实现占位抑制的效果。综上所述, 当前的研究表明, 天然化合物 HBIN031053 具有基于 Mpro、PLpro 和 RBD 多靶点的抑制 SARS-CoV-2 病毒的潜在活性, 可作为结构优化和抗 SARS-CoV-2 感染的新型潜在抑制剂的先导化合物。酶活性抑制结果表明, 含 HBIN031053 的草药巴戟天的提取物对 SARS-CoV-2 关键蛋白 Mpro 的酶活性具有抑制效果。本研究可为 COVID-19 治疗药物的开发提供参考。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] Troyano-Hernández P, Reinosa R, Holguín Á. Evolution of SARS-CoV-2 envelope, membrane, nucleocapsid, and spike structural proteins from the beginning of the pandemic to September 2020: A global and regional approach by epidemiological week [J]. *Viruses*, 2021, 13(2): 243.
- [2] Vuong W, Khan M B, Fischer C, et al. Feline coronavirus drug inhibits the main protease of SARS-CoV-2 and blocks virus replication [J]. *Nat Commun*, 2020, 11(1): 4282.
- [3] Fu Z Y, Huang B, Tang J L, et al. The complex structure of GRL0617 and SARS-CoV-2 PLpro reveals a hot spot for antiviral drug discovery [J]. *Nat Commun*, 2021, 12(1): 488.
- [4] Zhao Y, Fang C, Zhang Q, et al. Crystal structure of SARS-CoV-2 main protease in complex with protease inhibitor PF-07321332 [J]. *Protein Cell*, 2022, 13(9): 689-693.
- [5] Jackson C B, Farzan M, Chen B, et al. Mechanisms of SARS-CoV-2 entry into cells [J]. *Nat Rev Mol Cell Biol*, 2022, 23(1): 3-20.
- [6] Wang Z L, Yang L Y. Chinese herbal medicine: Fighting

- SARS-CoV-2 infection on all fronts [J]. *J Ethnopharmacol*, 2021, 270: 113869.
- [7] Fang S S, Dong L, Liu L, et al. HERB: A high-throughput experiment- and reference-guided database of traditional Chinese medicine [J]. *Nucleic Acids Res*, 2021, 49(D1): D1197-D1206.
- [8] Ma C L, Sacco M D, Xia Z L, et al. Discovery of SARS-CoV-2 papain-like protease inhibitors through a combination of high-throughput screening and a FlipGFP-based reporter assay [J]. *ACS Cent Sci*, 2021, 7(7): 1245-1260.
- [9] Wang Q H, Zhang Y F, Wu L L, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2 [J]. *Cell*, 2020, 181(4): 894-904.e9.
- [10] Dai Q Q, Yan Y H, Ning X L, et al. AncPhore: A versatile tool for anchor pharmacophore steered drug discovery with applications in discovery of new inhibitors targeting metallo- β -lactamases and indoleamine/tryptophan 2, 3-dioxygenases [J]. *Acta Pharm Sin B*, 2021, 11(7): 1931-1946.
- [11] Sunseri J, Koes D R. Pharmit: Interactive exploration of chemical space [J]. *Nucleic Acids Res*, 2016, 44(W1): W442-W448.
- [12] Halimi M, Bararpour P. Natural inhibitors of SARS-CoV-2 main protease: Structure based pharmacophore modeling, molecular docking and molecular dynamic simulation studies [J]. *J Mol Model*, 2022, 28(9): 279.
- [13] Mysinger M M, Carchia M, Irwin J J, et al. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking [J]. *J Med Chem*, 2012, 55(14): 6582-6594.
- [14] Trott O, Olson A J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading [J]. *J Comput Chem*, 2010, 31(2): 455-461.
- [15] Laskowski R A, Swindells M B. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery [J]. *J Chem Inf Model*, 2011, 51(10): 2778-2786.
- [16] Yuda G P W C, Hanif N, Hermawan A. Computational screening using a combination of ligand-based machine learning and molecular docking methods for the repurposing of antivirals targeting the SARS-CoV-2 main protease [J]. *Daru*, 2024, 32: 47-65.
- [17] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules [J]. *Sci Rep*, 2017, 7: 42717.

- [18] Banerjee P, Eckert A O, Schrey A K, *et al.* ProTox-II: A webserver for the prediction of toxicity of chemicals [J]. *Nucleic Acids Res*, 2018, 46(W1): W257-W263.
- [19] Abraham M J, Murtola T, Schulz R, *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers [J]. *SoftwareX*, 2015, 1/2: 19-25.
- [20] Yu Y L, Krämer A, Venable R M, *et al.* Semi-automated optimization of the CHARMM36 lipid force field to include explicit treatment of long-range dispersion [J]. *J Chem Theory Comput*, 2021, 17(3): 1562-1580.
- [21] Zoete V, Cuendet M A, Grosdidier A, *et al.* SwissParam: A fast force field generation tool for small organic molecules [J]. *J Comput Chem*, 2011, 32(11): 2359-2368.
- [22] Valdés-Tresanco M S, Valdés-Tresanco M E, Valiente P A, *et al.* gmx_MMPBSA: A new tool to perform end-state free energy calculations with GROMACS [J]. *J Chem Theory Comput*, 2021, 17(10): 6281-6291.
- [23] Owen D R, Allerton C M N, Anderson A S, *et al.* An oral SARS-CoV-2 M^{pro} inhibitor clinical candidate for the treatment of COVID-19 [J]. *Science*, 2021, 374(6575): 1586-1593.
- [24] Yi Y, Li J H, Lai X Y, *et al.* Natural triterpenoids from licorice potently inhibit SARS-CoV-2 infection [J]. *J Adv Res*, 2022, 36: 201-210.
- [25] Lipinski C A, Lombardo F, Dominy B W, *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings [J]. *Adv Drug Deliv Rev*, 2001, 46(1/2/3): 3-26.
- [26] Gao X P, Qin B, Chen P, *et al.* Crystal structure of SARS-CoV-2 papain-like protease [J]. *Acta Pharm Sin B*, 2021, 11(1): 237-245.
- [27] Jawad B, Adhikari P, Podgornik R, *et al.* Key interacting residues between RBD of SARS-CoV-2 and ACE2 receptor: Combination of molecular dynamics simulation and density functional calculation [J]. *J Chem Inf Model*, 2021, 61(9): 4425-41.

[责任编辑 王文倩]