

TCMSP 数据库中中药化学成分的 SARS-CoV-2 主蛋白酶抑制剂虚拟筛选

史海龙, 黄月, 程怡, 史永恒, 王川, 刘继平, 王斌(陕西中医药大学, 陕西 西咸新区 712046)

摘要: **目的** 运用计算机模拟技术从中药系统药理学数据库与分析平台(TCMSP)的中药化学成分中挖掘 SARS-CoV-2 主蛋白酶(Main protease, Mpro)抑制剂。**方法** 运用类药性评估、分子对接、ADME 预测、结合自由能计算、分子动力学模拟等多轮虚拟筛选的策略, 从 TCMSP 数据库中筛选 Mpro 抑制剂。**结果** 共筛选得到 10 个命中分子, 其中前 3 位最佳命中分子为 MOL003392、MOL011716、MOL002966, 表现出与靶标活性口袋关键氨基酸残基的强相互作用。**结论** 该研究可为发现 SARS-CoV-2 Mpro 小分子抑制剂提供新的思路。

关键词: SARS-CoV-2; 主蛋白酶(Mpro); 分子对接; 虚拟筛选; 分子动力学模拟

中图分类号: R285.5; R857.3 **文献标志码:** A **文章编号:** 1003-9783(2022)02-0219-09

doi: 10.19378/j.issn.1003-9783.2022.02.011

Virtual Screening of SARS-CoV-2 Mpro Protease Inhibitors Based on the Chemical Components of Traditional Chinese Medicine in TCMSP Database

SHI Hailong, HUANG Yue, CHENG Yi, SHI Yongheng, WANG Chuan, LIU Jiping, WANG Bin (Shaanxi University of Chinese Medicine, Xixian New Area 712046 Shaanxi, China)

Abstract: **Objective** Computer simulation technology was used to discover SARS-CoV-2 main protease (Mpro) inhibitors from chemical components of TCMSP database. **Methods** Multiple rounds of virtual screening strategies such as drug-like evaluation, molecular docking, ADME prediction, binding free energy calculation, molecular dynamics simulation, etc., were conducted to screen Mpro inhibitors from the TCMSP database. **Results** Ten hits were obtained, among which the top 3 best hits MOL003392, MOL011716, and MOL002966 showed strong interactions with key amino acid residues in the target active pocket. **Conclusion** This study provides a new idea for the discovery of small molecules as SARS-CoV-2 Mpro inhibitors.

Keywords: SARS-CoV-2; main protease (Mpro); molecular docking; virtual screening; molecular dynamic simulation

SARS-CoV-2 是一种属于冠状病毒科的新型冠状病毒, 可引起急性呼吸道疾病新型冠状病毒肺炎(COVID-19)。迄今为止, 由于病毒不断变异, 毒力和感染力不断增强, 现有的药物方案疗效有限, 尚未成功研发出有效控制 COVID-19 的临床药物^[1]。在 SARS-CoV-2 基因组中, 主蛋白酶(Main protease,

Mpro)也称为半胱氨酸蛋白酶, 负责切割病毒复制酶-转录酶复合物相关的一系列独立功能单元, 对于维持病毒的生命周期至关重要^[2]。因此, 靶向 SARS-CoV-2 Mpro 可能提供一种策略来遏制病毒复制并最终有效解决 COVID-19 危机^[3]。

目前在 PDB 蛋白质结构数据库(Protein data

收稿日期: 2021-03-15

作者简介: 史海龙, 男, 博士后, 副教授, 研究方向: 计算机辅助药物设计及生物大分子计算机模拟。Email: shl112@sntcm.edu.cn。通信作者: 王斌, 男, 教授, 硕士研究生导师, 研究方向: 中医药防治心脑血管疾病。Email: wangbing812@126.com。

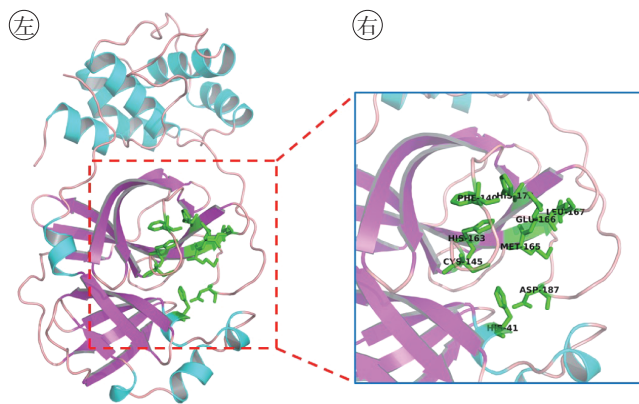
基金项目: 国家留学基金委资金资助项目(201908610138); 国家级大学生创新创业项目(201910716031); 陕西省中医管理局中医药科研课题(JCMS003); 陕西中医药大学创新团队项目(2019-YL13)。

bank, PDB; <https://www.rcsb.org/>)中已经可以得到高分辨率下测定的 Mpro 的晶体结构(PDB ID: 5R84), 基于此可以进行基于受体结构的药物设计, 高通量虚拟筛选天然产物化合物库, 成本与效率兼顾地挖掘潜在药效分子, 为新冠病毒药物研发提供了良好机遇^[4]。基于 SARS-CoV-2 Mpro 的活性位点结合口袋, 本研究拟从中药系统药理学数据库与分析平台(Traditional chinese medicine database and analysis platform, TCMSP^[5])筛选中药小分子抑制剂, 进行类药性评估、分子对接、吸收/分布/排泄/代谢(Absorption-Distribution-Metabolism-Excretion, ADME)预测、结合自由能计算、分子动力学模拟等多轮虚拟筛选。此外, 为了更加精准解析候选化合物与靶标 Mpro 的相互作用方式, 本研究拟借助分子动力学模拟技术, 预测在真实生理溶液系统中蛋白质-配体相互作用的稳定性, 并提供有关结合中涉及的氨基酸残基的信息。

1 材料与方法

1.1 软件 分子对接计算均采用美国 Schrodinger 公司的 Schrodinger-v2018.5 软件的 Glide 模块; ADME 参数计算采用 Schrodinger-v2018.5 软件的 QikProp 模块; MM/GBSA 结合自由能预测采用 Schrodinger-v2018.5 软件的 MM/GBSA 模块; 分子动力学模拟及 MM/PBSA 结合自由能预测均采用开源软件 Gromacs-v2019 程序包。

1.2 受体蛋白质的预处理及对接网格生成 从 PDB 数据库下载 COVID-19 Mpro 的晶体结构^[6](PDB ID: 5R84), 以其共结晶的配体 [2-cycloxy-*N*-(3-pyridyl)acetamide, Z31792168]作为虚拟筛选的阳性对照分子。使用 Schrodinger 软件的 Protein Preparation wizard 模块对 5R84 进行蛋白质结构预处理。以 Z31792168 的空间坐标为参考, 使用 Receptor Grid Generation 模块生成 Grid 文件, 用于基于受体结构的虚拟筛选。参考 Jin 等^[7]的报道, Mpro 形成催化底物结合袋的重要残基在所有冠状病毒中发现的主要蛋白酶中高度保守, 如 His41、Phe140、Cys145、His163、Met165、Glu166、Leu167、His172、Asp187, 然而 Mpro 的 Surface loop 和 Domain-III 包含最多的可变区, 因此选择催化结合袋的高度保守残基可作为关键氨基酸残基, 并使用 Pymol 进行可视化处理(见图 1)。



注: 左侧为活性口袋示意图; 右侧为活性部位的关键氨基酸示意图

图 1 SARS-CoV-2 主蛋白酶空间结构

Figure 1 The spatial structure of SARS-CoV-2 main protease

1.3 中药分子化学结构的预处理 以 TCMSP 数据库中的 13 144 个化合物作为配体库, 采用 Schrodinger 的 LigPrep 模块对各分子进行加氢、能量最小化、几何优化等处理。在力场 OPLS3 下, 计算 pH 值在 (7.0±2.0) 范围内的化合物离子状态, 搜索合理构象作为对接起始构象, 形成可能的互变异构体。以同样方式预处理 Z31792168。

1.4 基于类药性评估、分子对接模拟、ADME 预测、结合自由能计算的多轮虚拟筛选 首先, 基于类药性规则“Lipinski Ro5^[8]”及“Verber Ro3^[9]”对中药分子配体库进行首轮筛选。然后, 基于 Mpro 酶活性口袋结构 Glide 程序对命中分子进行第 2 轮筛选, 依次从高通量模式 HTVS-docking、标准模式 SP-docking、高精度模式 XP-docking 进行不同精度的分子对接。接着, 基于 ADME 药代动力学参数进行第 3 轮筛选, 从分子量、氢键供体、氢键受体、溶剂可及表面积、K⁺通道阻断相关 IC₅₀、水溶性、脂水分配系数、口服生物利用度等角度评估, 筛选具有良好类药性质的化合物。最后, 基于 Prime-MM/GBSA 结合自由能预测进行第 4 轮筛选, 由于高精度分子对接函数计算仍然无法估算配体-受体复合物结合的溶剂化效应因素, 故使用隐性溶剂模型下的 MM/GBSA 算法优化排序。

1.5 分子动力学模拟 为了从动力学和热力学角度动态分析 Mpro 与命中分子的互作模式, 使用 Gromacs 软件进行分子动力学模拟。蛋白拓扑参数取自 Amber 的 FF03 力场, 由 PDB2GMX 工具获得; 抑制剂的各原子电荷先用 Gaussian 09 程序包在 B3LYP/6-311G** 基组水平计算静电势, 再由 Amber 中的 RESP 电荷拟合程序计算。将复合物放入 10×10×10 Å³ 的立方体

TIP3P 水盒中，并加入 Na^+ 和 Cl^- ，中和该系统到 $0.15\text{ mol}\cdot\text{L}^{-1}$ NaCl 溶液浓度；随后进行能量最小化和预平衡，最后在 300 K 温度进行 100 ns 的分子动力学模拟，时间间隔为 2 fs，每 10 ps 保存 1 次能量和坐标文件。从 90~100 ns 稳定模拟轨迹中提取 100 帧轨迹文件，计算 MM/PBSA 结合自由能。

2 结果与分析

2.1 基于结构的虚拟筛选 基于靶标 Mpro 的三维空间结构，对 TCMSP 数据库的 13 144 个天然化合物进行了多轮虚拟筛选，最终确定了得分最高的 10 个候选化合物。与 Z31792168 相比较，命中分子具备有较强的靶标结合自由能、良好的药代动力学参数、较高的生物利用度等特性，筛选流程及结果见图 2、表 1、表 2。

2.2 结合模式分析及分子动力学模拟氢键数目解析 综合考虑了 Glide-XP 分子对接打分值、MM/GBSA 结合自由能预测值和 ADME 参数等各种因素，最终确定打分排前 3 位(MM/GBSA 结合自由能预测值 $\leq -55.10\text{ kcal}\cdot\text{mol}^{-1}$) 的化合物(MOL003392、

MOL011716、MOL002966)为最佳命中分子。采用 Discovery Studio Visualizer 软件绘制了命中分子、原

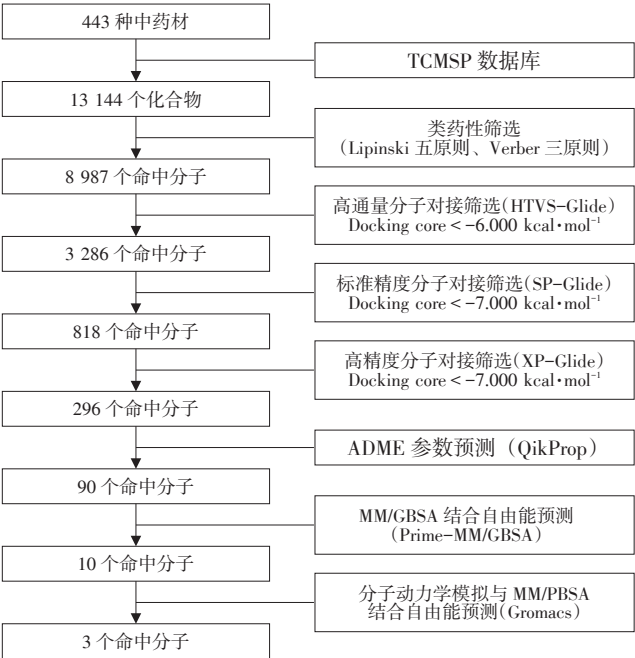


图 2 虚拟筛选流程及结果
Figure 2 Flow chart and results of virtual screening

表 1 排前 10 位的候选化合物与靶标的分子对接打分值及 MM/GBSA 结合能(kcal·mol⁻¹)
Table 1 The molecular docking scores and MM/GBSA binding energies of the top 10 hits(kcal·mol⁻¹)

TCMSP_ID	化合物(中文名)	MM/GBSA	Glide-XP	来源	化学结构式
MOL003392	Alpha-Terthienylmethanol (alpha-三联噻吩甲醇)	-56.65	-6.345	墨旱莲	
MOL011716	8-Hydroxy-9, 10-diisobutyryloxythymol (8-羟基-9, 10-二异丁酰氧基麝香草酚)	-56.09	-7.456	鹅不食草	
MOL002966	Dalbergin (黄檀素)	-55.10	-6.806	降香	
MOL005890	Pachypodol (藿香黄酮醇)	-54.00	-6.358	广藿香	
MOL004980	Inflacoumarin A (胀果香豆素甲)	-53.29	-7.068	甘草、糯稻根	
MOL007692	Dichotomide I (银柴胡氨 I)	-51.47	-7.457	银柴胡	
MOL010379	7'-n-butanol-N-trans-feruloyloctopamine	-51.33	-7.753	玉竹	
MOL012825	5-(3,4-dihydroxybutyn-1)-bithiophene	-50.21	-5.722	禹州漏芦	
MOL007701	Dichotomine D (银柴胡胺 D)	-49.61	-7.687	银柴胡	
MOL003956	Dihydrorutaecarpine (二氢吴茱萸次碱)	-48.88	-5.901	吴茱萸	
Z31792168	2-Cyclohexyl-N-(3-pyridyl) acetamide	-42.71	-5.701		

表 2 排前 10 位的命中化合物的 ADME 参数预测值

Table 2 The predicted values of ADME parameters of the top 10 hits

Entry	MW ^a	HB donor ^b	HB acceptor ^c	SASA ^d	LogHERG ^e	QlogS ^f	logPo/w ^g	PCaco-2 ^h	OB ⁱ /%
MOL003392	278	1	1.70	515.04	-5.482	-5.172	6.072	3 063.706	100.00
MOL011716	338	2	5.50	659.85	-5.011	-4.752	9.124	741.638	100.00
MOL002966	268	1	4.00	494.48	-5.082	-3.412	8.414	945.381	94.70
MOL005890	356	1	5.75	625.58	-5.143	-4.764	3.521	2 075.473	100.00
MOL004980	322	2	4.00	597.80	-5.414	-4.780	3.192	348.935	91.15
MOL007692	339	0	5.50	641.66	-5.816	-4.683	2.860	250.026	86.62
MOL010379	385	3	6.45	747.36	-6.349	-5.176	3.798	416.389	96.07
MOL012825	250	2	3.40	509.72	-5.429	-3.803	2.687	1 005.756	96.42
MOL007701	328	3	6.40	579.46	-4.836	-3.252	2.069	284.371	82.98
MOL003956	289	2	4.00	530.71	-5.545	-4.442	3.038	1 839.897	100.00
Z31792168	218	1	4.00	481.61	-4.375	-3.190	2.400	3 063.706	100.00

注: a. 分子量 (可接受范围: < 650); b. 氢键供体(可接受范围: ≤ 5); c. 氢键受体(可接受范围: ≤ 10); d. 溶剂可及表面积 (可接受范围: 300 ~ 1 000); e. IC₅₀值(推荐区间: < -5); f. 亲水性(可接受范围: -7.0 ~ 0.5); g. 脂水分配系数(可接受范围: -2 ~ 6.5); h. Caco-2 小肠细胞透过率(过差: < 25; 较好: > 500); i. 口服生物利用率(极差: < 50%; 较好: > 80%)

结晶配体分子与靶蛋白酶 Mpro 的结合模式图, 包括氢键、疏水相互作用、离子相互作用和水桥等不同形式的受体-配体相互作用, 结果见图 3~图 6。分子对接研究表明, 前 3 个命中化合物均具有相同的结合腔, 并显示出与 Mpro 的各种活性口袋关键残基(如 His41、Asn142、Gly143、Ser144、Cys145、Met165、Glu166)相互作用, 其中催化二元组由 His41 与 Cys145 组成。

由图 3 可知, 靶蛋白酶 Mpro 与 Z31792168 之间的主要作用力是常规氢键、盐桥、alkyl 相互作用。Z31792168 与 Mpro 氨基酸残基 His63、Glu166 均形成常规氢键, 与氨基酸残基 His164、Phe140 均形成碳氢键; 与受体蛋白空腔的氨基酸残基 His41、Met49、Met165 形成 alkyl 相互作用, 与氨基酸残基 Cys145 形成 π -alkyl 堆积作用, 与这些作用力促进复合物的形成, 抑制 Mpro 酶活性导致冠状病毒复制的阻断。

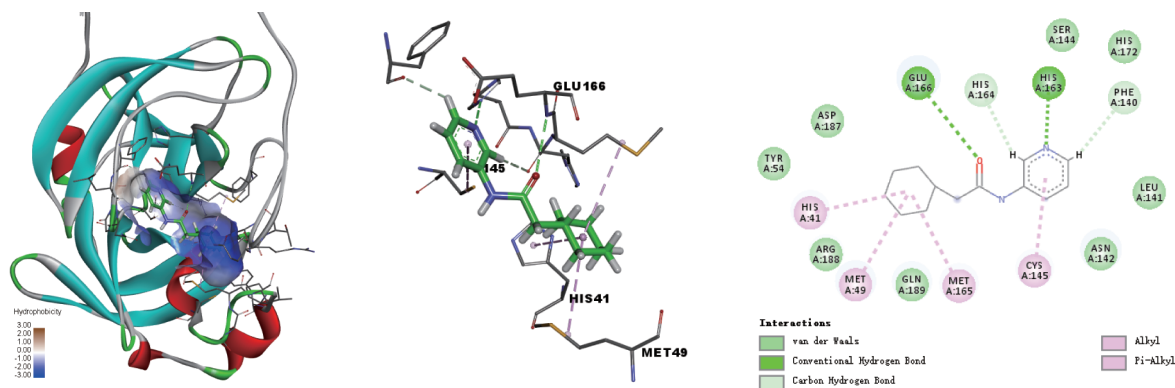


图 3 主蛋白酶与 Z31792168 的分子对接

Figure 3 Molecular docking of Mpro and Z31792168

由图 4 可知, 靶蛋白酶 Mpro 与 MOL003392 之间的主要作用力是氢键、范德华力、 π -sulfur 相互作用、 π -alkyl 相互作用、amide- π 堆积作用、 π - π 堆积作用。MOL003392 与靶标氨基酸残基 Tyr54、Glu189 均形成常规氢键, 与氨基酸残基 Glu166、Met49、Tyr54、Asp187 均形成碳氢键或 π -donor 氢

键。与氨基酸 Cys145 形成 π -sulfur 相互作用, 与 His41、Arg188 形成 amide- π 堆积作用或 π - π 堆积作用, 与氨基酸残基 Met165 形成 π -alkyl 相互作用, 与这些作用力促进复合物的形成。

由图 5 可见, 靶蛋白酶 Mpro 与 MOL011716 之间的相互作用不仅有氢键、范德华力, 还有 π - π 堆

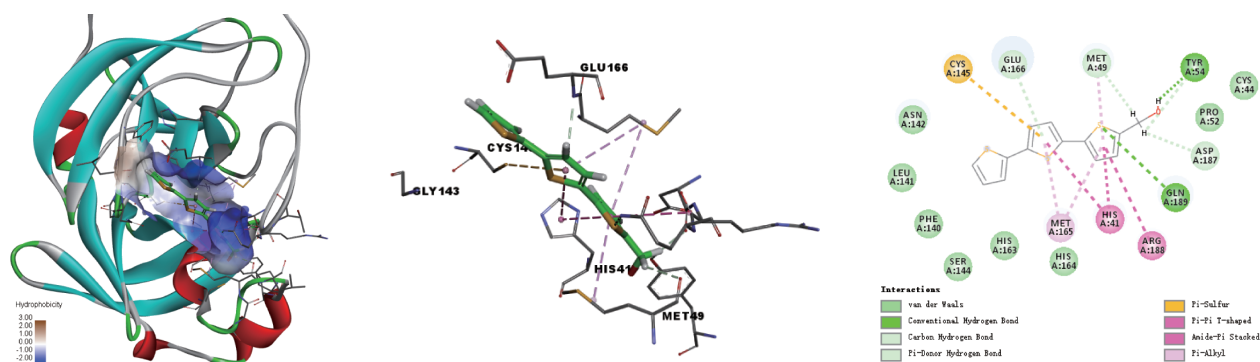


图 4 主蛋白酶与 MOL003392 的分子对接

Figure 4 Molecular docking of Mpro and MOL003392

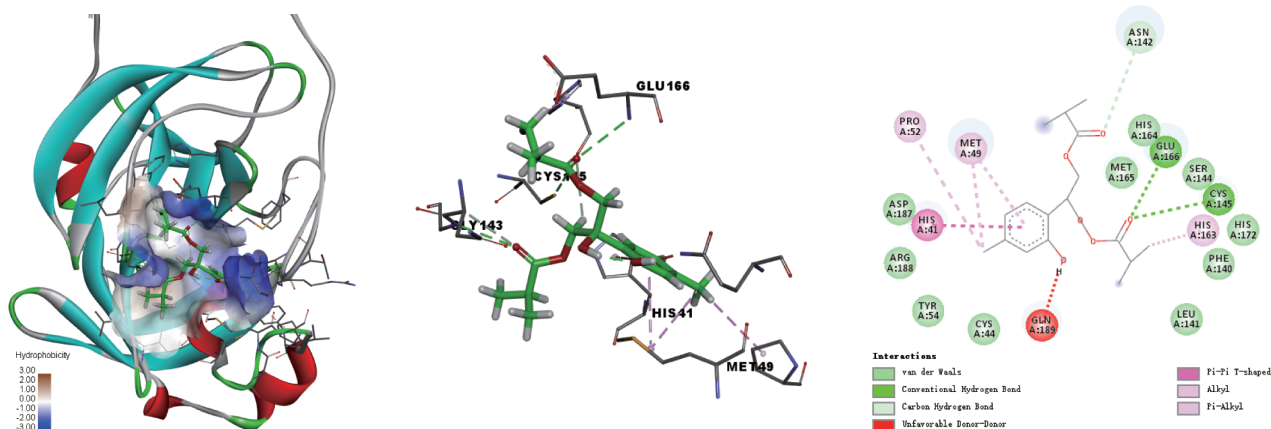


图 5 主蛋白酶与 MOL011716 的分子对接

Figure 5 Molecular docking of Mpro and MOL011716

积作用、 π -Alkyl 相互作用。MOL011716 与氨基酸残基 Cys145、Glu166 形成常规氢键，并且同样与 Asn142 形成碳氢键，还与 Met49、Pro52、His163 形成 π -Alkyl 或 Alkyl 相互作用，与氨基酸残基 His41 形成 π - π 堆积作用。这些作用力使结合能降低，亲和力增加，候选单体与蛋白酶 Mpro 活性空腔部位紧

密结合，形成稳定的复合物，以有效阻止外来病毒粒子附着并侵入宿主细胞。同时还观察到，Glu189 的侧链与药物分子之间可能形成了不利的 Donor-Donor 空间撞击。

由图 6 可见，靶蛋白酶 Mpro 与 MOL002966 之间的相互作用不仅有氢键，还有 π - π 堆积作用与

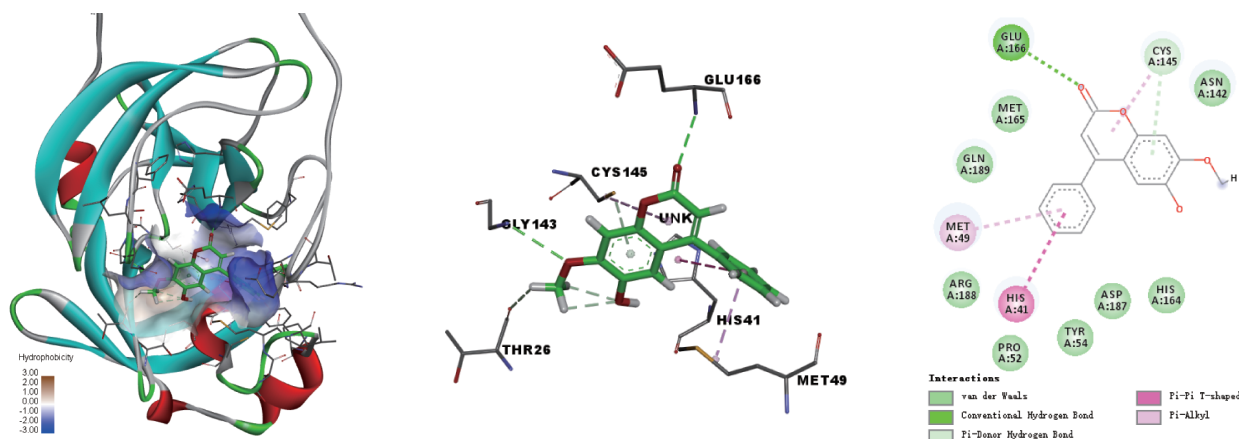


图 6 主蛋白酶与 MOL002966 的分子对接

Figure 6 Molecular docking of Mpro and MOL002966

π -Alkyl 相互作用。MOL002966 与 Mpro 形成氢键, 其中与 Glu166 形成常规氢键, 与 Cys145 形成 π -donor 氢键, 并且同样与 Met49、Cys145 均形成一个 π -Alkyl 相互作用, 与 His41 形成 π - π 堆积作用。这些作用力使结合能降低, 亲和力增加, 候选单体与靶蛋白酶的活性口袋紧密结合, 形成稳定的复合物, 可以有效阻止切割多聚蛋白体成为一系列独立功能单元。

2.3 分子动力学模拟

2.3.1 均方根偏差(RMSD)值计算 计算分子动力学模拟 100 ns 时长轨迹的 RMSD 值, 以此检验体系稳定性, 结果见图 7。在模拟运行 65 ns 时长之后, 所有体系 RMSD 平均值为 0.2~0.4 nm, 且涨落范围都稳定在 0.1 nm 以下, 均达到了平衡状态。

2.3.2 均方根波动(RMSF)值分析 Mpro 抑制剂复合物体系 Mpro-MOL003392、Mpro-MOL011716、Mpro-MOL002966 与原晶体分子复合物体系 Mpro-Z31792168 都有相似 RMSF 数值分布趋势, 结果见

图 8。Jin 等^[10]报道来自 residue 8~101、residue 102~184、residue 201~303 的残基分别形成 Domain-I、Domain-II、Domain-III。Domain-III 包含 5 个反向平行的 α -helices, 这些螺旋排列成球形簇, 并借助 residue 185~200 的 Long loop 与 Domain-II 相连, 其中底物结合位点位于 Domain-I 和 Domain-II 之间的裂隙内部。结合图 3 可以看出, Domain-I、Domain-II、Domain-III 以及底物结合位点内的残基 RMSF 都明显较原晶体 Z31792168 数值更小, 并且在活性中心催化二元组(His41 与 Cys145)具有更低 RMSF 数值, 促使关键催化氨基酸残基构象更加稳定, 有利于配体和受体的相互作用。

2.3.3 MM/PBSA 结合自由能计算 从 90~100 ns 分子动力学轨迹中提取 100 帧轨迹文件, 计算 MM/PBSA 结合自由能及各种能量贡献情况, 结果见表 3。结合自由能强弱关系为: MOL003392>MOL011716>MOL002966, 其中范德华力最有利于命中分子与靶标结合, 并起主导作用; 静电势能、表面溶剂化作

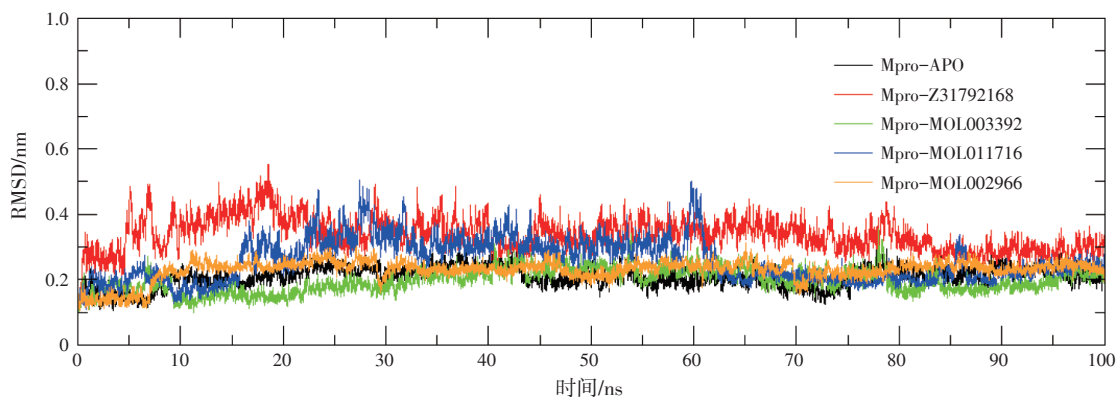


图 7 主蛋白酶蛋白骨架原子随时间变化的均方根偏差(RMSD)值

Figure 7 The RMSD graph of Mpro protein backbone

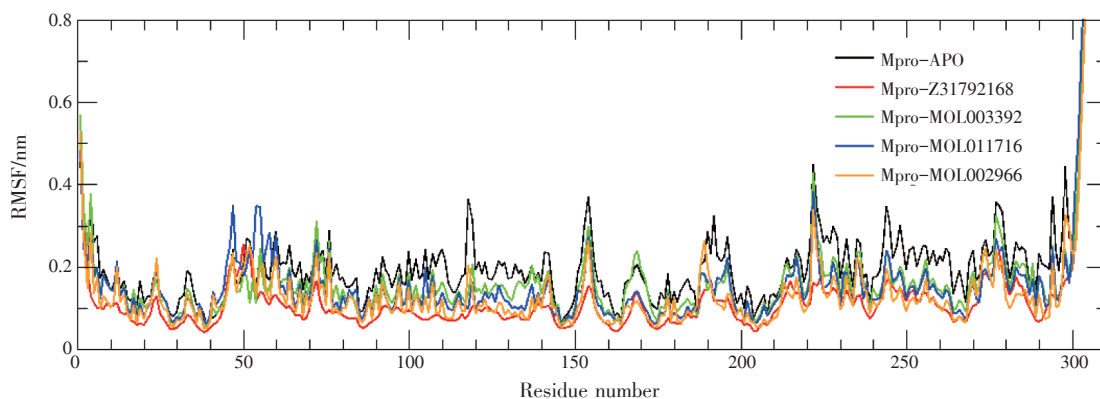


图 8 复合物中蛋白骨架原子的均方根波动(RMSF)值分布图

Figure 8 The RMSF graph of protein backbone in complex

用也有利于结合，但作用力较弱，而溶剂中极性溶剂化作用不利于体系结合。与 MM/GBSA 结合自由能的打分排序结果一致，均接近原结晶配体分子 Z31792168。

2.3.4 氢键数量的动态分析 借助 g-hbond 程序计算 Mpro 与命中分子之间的氢键数量随时间的变化，统计结果见图 9。总体而言，与 Z31792168 相比，

MOL003392、MOL011716、MOL002966 与 Mpro 结合氢键的数量近似；并且在分子动力学 65 ns 时长后，3 个命中分子与靶标至少存在 2 次氢键相互作用的接触，提示稳定后复合物均具有氢键的相互作用，推测 MOL003392、MOL011716、MOL002966 可能会产生较强的抑制效应。

表 3 前 3 位命中化合物与主蛋白酶的 MM/PBSA 结合自由能(kJ·mol⁻¹)

Table 3 MM/PBSA binding free energies of top 3 hits with Mpro protease(kJ·mol⁻¹)

化合物	范德华力	静电势能	极性溶剂化作用	表面溶剂化作用	结合自由能
MOL003392	-131.463 ± 11.373	-15.676 ± 6.085	80.403 ± 9.040	-13.082 ± 0.734	-79.818 ± 10.182
MOL011716	-119.386 ± 22.577	-17.063 ± 18.559	75.989 ± 26.082	-13.624 ± 1.719	-74.085 ± 17.782
MOL002966	-113.236 ± 9.708	-13.439 ± 7.553	69.701 ± 10.171	-10.907 ± 1.117	-67.882 ± 10.043
Z31792168	-122.333 ± 10.304	-26.633 ± 10.068	82.941 ± 13.603	-12.571 ± 0.771	-78.596 ± 9.642

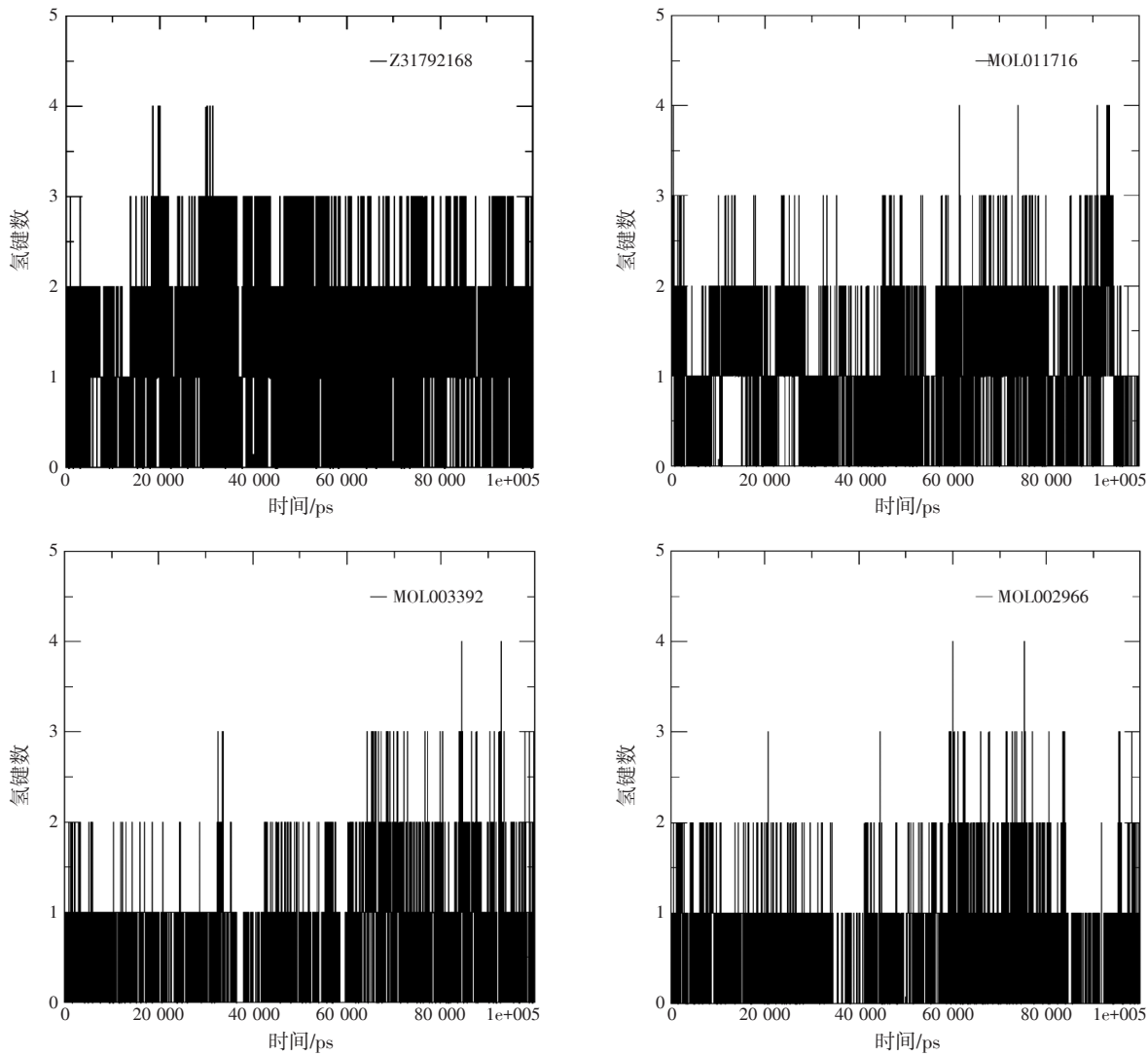


图 9 分子动力学模拟过程中命中分子与主蛋白酶形成的氢键数量

Figure 9 Hydrogen bonds formed between hits and Mpro during molecular dynamics simulation

4 讨论

由于抑制 Mpro 酶活性导致冠状病毒复制的阻断, 加上 SARS-CoV-2 Mpro 与目前已知的任何人类蛋白酶都不同源, 因此其抑制剂的毒性较小^[11]。基于此靶标结构的计算机虚拟筛选为前提, 然后进行药理学验证研究, 可能是研发抗 SARS-CoV-2 药物的有效策略。刘妍如等^[12]针对 SARS-CoV-2 3CL 水解酶的高分率晶体结构(PDB ID: 6LU7), 运用分子对接技术模拟预测了 TCMSP 数据库中药成分对 SARS-CoV-2 3CL 水解酶的作用, 共挖掘得到 60 个候选化合物。本研究针对相同问题, 重新设计了研究方案, 尽可能提高计算的精度, 主要从以下方面进行了改进或补充。

第一, 重新选择靶标。从 COVID-19 大流行爆发开始, 结构生物学家们已经前瞻性地进行了多种尝试, 解决不同药物靶标的三维结构问题, 以辅助药物发现过程。SARS-CoV-2 Mpro 的第 1 个三维晶体结构于 2020 年 2 月上旬发布(PDB ID: 6LU7)^[7], 其结晶原配体小分子为肽类分子。2020 年 10 月 Alice Douangamath 团队发布了新的 SARS-CoV-2 Mpro 晶体结构(PDB ID: 5R84)^[3], 其共结晶的配体[2-cycloxy-*N*-(3-pyridyl)acetamide, Z31792168]为有机小分子, 更加适合作为虚拟筛选 SARS-CoV-2 Mpro 抑制剂的阳性对照分子。

第二, 采用分子动力学模拟技术。任何基于分子对接算法的普遍局限性在于它将受体蛋白质视为刚性实体, 因此这种分子对接的结果仅仅提供了蛋白质-配体小分子之间相互作用的静态图像。然而, 在真实生理系统中, 这种相互作用本质上是动态的, 是不断变化的。因此, 为了深度解析候选化合物-靶标蛋白之间相互作用情况, 本研究采用分子动力学模拟程序 Gromacs-v2019 对命中分子相应的配体-蛋白质复合物对接构象进行了 100 ns 分子动力学模拟。这样做的基本原理是从分子动力学层面解析候选分子与 Mpro 催化活性位点结合的动力学状态。分子动力学模拟研究对于在模拟正常生理环境下检查配体-蛋白质复合物的稳定性和动态波动是有必要的。

第三, 基于 ADME 药代动力学参数的筛选。在临床前药物发现过程中纳入 ADME 预测可以帮助降低临床试验中候选药物失败的风险。因此, 本研究

使用 Schrodinger 2018 的 QikProp 模块对前 10 个命中分子及 Z31792168 进行了 ADME/T 参数预测, 选择了具有药理学重要意义的理化指标, 如溶剂可及表面积(SASA)、脂水分配系数(logPo/w)、水溶性(logS)、对 HERG K⁺通道阻塞的预测 IC₅₀ 值(logHERG)、预测表观肠血屏障的 Caco-2 细胞渗透性(PCaco-2)及口服生物利用率(OB)等。分析重要参数的预测值及其可接受范围发现, 化合物 MOL011716、MOL002966 的 logPo/w 参数略有偏差, 化合物 MOL003656、MOL006340、MOL008039 的 PCaco-2 略有偏差, 除此以外, 命中化合物和共结晶的配体 Z31792168 均具有良好的药代动力学参数, 可能通过临床药物测试。

第四, 分子模拟数据的深度分析。①选取排前 3 位的命中分子 Mpro-MOL003392、Mpro-MOL011716、Mpro-MOL002966 及 Mpro-Z31792168 进行 100 ns 的分子动力学模拟, 计算 Mpro 酶蛋白主链原子相对于初始优化结构的随时间变化的均方根偏差(Root mean square deviation, RMSD)函数, 用以判断分子动力学平衡的可靠性。②计算整个分子动力学模拟体系过程中 0~100 ns 时间内的 Mpro 氨基酸骨架的均方根波动值(Root mean square fluctuation, RMSF), 分析蛋白质中每个氨基酸残基的柔性, 评估不同小分子抑制剂对 Mpro 每个氨基酸残基波动的影响。结合口袋附近的氨基酸与配体相互作用会有较小的 RMSF, 因此比较不同小分子结合时蛋白的 RMSF, 并能用以考察蛋白与配体结合的强弱。③由于氢键在生物大分子中起重要作用, 氢键是蛋白质与配体相互作用的主要推动力之一, 在稳定蛋白质-配体复合物方面起着不可或缺的作用^[13]。借助 g-hbond 程序在整个 100 ns 的模拟时间内, 绘制了以氢键相互作用在时间轴上的变化, 计算 100 ns 模拟过程中 Mpro 与抑制剂分子间的氢键数量以及氢键距离或角度的分布。

综上所述, 采用分子对接、MM/GBSA 结合自由能计算、MM/PBSA 结合自由能、分子动力学模拟和 ADME 参数预测等一系列虚拟筛选的研究策略, 最终发现最佳命中分子为 MOL003392、MOL011716、MOL002966, 上述分子均紧密贴合在 Mpro 酶蛋白结合腔内, 并显示出与高度保守氨基酸残基(如 His41、Cys145、Glu166)的强相互作用。鉴于虚拟筛

选结果的局限性,后期还需进一步通过体内、体外实验来验证本研究结果,以期为抗 COVID-19 天然药物的研发提供实验依据。尽管尚未开展命中化合物的抗病毒生物学验证实验,本研究结果亦可为开发抗病毒药物的先导化合物提供参考。

参考文献:

- [1] SANDERS J M, MONOGUE M L, JODLOWSKI T Z, et al. Pharmacologic treatments for coronavirus disease 2019(COVID-19): a review[J]. JAMA, 2020, 323(18): 1824-1836.
- [2] CHEN Y, LIU Q, GUO D. Emerging coronaviruses: genome structure, replication, and pathogenesis[J]. Journal of Medical Virology, 2020, 92(4), 418-423.
- [3] ULLRICH S, NITSCHKE C. The SARS-CoV-2 main protease as drug target[J]. Bioorganic & Medicinal Chemistry Letters, 2020, 30(17): e127377.
- [4] REN Z, YAN L, NING Z, et al. The newly emerged SARS-Like coronavirus HCoV-EMC also has an "Achilles' heel": current effective inhibitor targeting a 3C-like protease[J]. Protein & Cell, 2013, 4(4): 248-250.
- [5] RU J, LI P, WANG J, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines[J]. Journal of Cheminformatics, 2014, 6(1): 1-6.
- [6] DOUANGAMATH A, FEARON D, GEHRTZ P, et al. Crystallographic and electrophilic fragment screening of the SARS-CoV-2 main protease[J]. Nature Communications, 2020, 11(1): 1-11.
- [7] JIN Z, DU X, XU Y, et al. Structure of Mpro from SARS-CoV-2 and discovery of its inhibitors[J]. Nature, 2020, 582: 289-293.
- [8] 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]. Advanced Drug Delivery Reviews, 1997, 23(1-3): 3-25.
- [9] VEBER D F, JOHNSON S R, CHENG H Y, et al. Molecular properties that influence the oral bioavailability of drug candidates[J]. Journal of Medicinal Chemistry, 2002, 45(12): 2615-2623.
- [10] GUO Y R, CAO Q D, HONG Z S, et al. The origin, transmission and clinical therapies on coronavirus disease 2019(COVID-19) outbreak-an update on the status[J]. Military Medical Research, 2020, 7(1): 1-10.
- [11] KIM Y, LIU H, GALASITI K A C, et al. Reversal of the progression of fatal coronavirus infection in cats by a broad-spectrum coronavirus protease inhibitor[J]. Plos Pathogens, 2016, 12(3): e1005531.
- [12] 刘妍如,唐志书,王梅,等.基于TCMSP数据库中药成分的新型冠状病毒3CL水解酶抑制剂虚拟筛选[J].中草药,2020,51(7): 1694-1703.
- [13] DU X, LI Y, XIA Y L, et al. Insights into protein-ligand interactions: mechanisms, models, and methods[J]. International Journal of Molecular Sciences, 2016, 17(2): 144-178.

(编辑:邹元平)