

基于埃博拉病毒糖蛋白单体结构从中药天然产物库中筛选新型抗病毒化合物

史海龙*, 黄月, 程怡, 王月雯, 冯雪松, 晁旭

陕西中医药大学, 西咸新区 712046

摘要: 基于埃博拉病毒跨膜单链糖蛋白(single glycoprotein of ebolavirus, EBOV-sGP)为靶标,运用类药性评估、分子对接、ADME 预测、结合自由能计算、分子动力学模拟等计算机虚拟筛选技术,从中药数据库 TCMS 中挖掘抗病毒天然产物活性分子。以药效分子托瑞米芬(toremifene)为阳性对照,最终筛选出 4 个类药性良好的先导化合物 MOL006834、MOL000174、MOL002192、MOL012524,它们与 EBOV-sGP 亲和力及相互作用基团均接近于 toremifene。该研究结果可促进从中药天然产物库中提取、设计以及实验合成抗埃博拉病毒的药效分子。

关键词: 埃博拉病毒; 中药天然产物库; 抗病毒活性分子; 虚拟筛选; 分子动力学模拟

中图分类号: R914.2

文献标识码: A

文章编号: 1001-6880(2022)5-0761-11

DOI: 10.16333/j.1001-6880.2022.5.004

Screening of new antiviral compounds from Chinese natural products library based on single glycoprotein of ebolavirus

SHI Hai-long*, HUANG Yue, CHENG Yi, WANG Yue-wen, FENG Xue-song, CHAO Xu

Shaanxi University of Chinese Medicine, Xixian New Area 712046, China

Abstract: In this study, ebolavirus single glycoprotein of ebolavirus (EBOV-sGP) was used as the target, using diverse virtual screening technologies, such as drug-like evaluation, molecular docking, ADME prediction, binding free energy calculation, molecular dynamics simulation, etc., in order to mine potential antiviral natural products from the database of traditional Chinese medicine. Taking toremifene as a positive control, four lead compounds, MOL006834, MOL000174, MOL002192, and MOL012524 with good drug-like properties were finally sorted. Their affinity values and interaction groups with EBOV-sGP are close to or similar to toremifene. The research results can promote the work of extracting, designing and experimentally synthesizing anti-ebolavirus compounds from natural products library of traditional Chinese medicine.

Key words: ebolavirus; Chinese natural products; antiviral active molecule; virtual screening; molecular dynamic simulation

埃博拉病毒(ebolavirus, EBOV)是一种包膜丝状病毒,其引起的急性出血热是多发于灵长类动物的人畜共患传染病,其病死率高达 50%~90%。近几年非洲西部多次暴发严重的埃博拉疫情,但是迄今为止尚无特效的抗埃博拉出血热治疗药物获准上市。EBOV 进入宿主细胞是感染的关键步骤,由病毒表面跨膜单链糖蛋白(single glycoprotein of ebolavirus, EBOV-sGP)介导,这种 I 类融合蛋白包含二

硫键连接的两个亚基,即表面暴露的 GP1 与嵌入病毒膜 GP2。在通过细胞表面附着因子与宿主细胞结合后,病毒被内吞,内体相应的半胱氨酸蛋白酶/组织蛋白酶切割掉 GP1 亚基的大部分肽链,裂解暴露受体结合位点^[1],这种 GP 构象的转变,有利于宿主-病毒膜融合和随后病毒基因组的释放。这种融合机制类似 HIV-1 病毒和流感病毒^[2]。因此,开发有效干扰 EBOV-sGP 构象转变的抑制剂分子,有助于阻断病毒入侵和复制。

2019 年英国牛津大学结构生物学系 Elizabeth E. Fry 及其团队通过使用蛋白表达纯化结晶技术已成功制备 EBOV-sGP 蛋白,并借助 X 射线衍射实验测定其三维结构^[3] (RCSB Protein Data Bank ID:

收稿日期: 2021-03-01 接受日期: 2022-03-22

基金项目: 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-YS05); 陕西省教育厅专项科研计划(19JK0237); 国家级大学生创新创业项目(201910716031)

* 通信作者 E-mail: shll12@sntcm.edu.cn

5JQ7), 提供了计算机虚拟筛选模型。除此之外, 由于天然产物, 尤其是传统中药成分库, 具有结构独特性与多样性, 已被用于治疗各种慢性病和传染病^[4,5], 同样可用于挖掘抗埃博拉病毒抑制剂。基于以上思考, 本研究采用计算机模拟技术, 从中药天然产物库中筛选 EBOV-sGP 抑制剂分子。

1 资料与计算方法

分子对接计算均采用 Schrodinger 2018 的 Glide 模块、ADME 参数计算采用 Schrodinger 2018 的 QikProp 模块, MM/GBSA 结合自由能预测^[6] 采用 Schrodinger 2018 的 Prime MM/GBSA 模块, 分子动力学模拟及 MM/PBSA 结合自由能预测^[7] 均采用 Gromacs 2019 程序包。采用 Pymol 与 Discovery Studio Visualizer 软件绘制受体配体结合模式图。

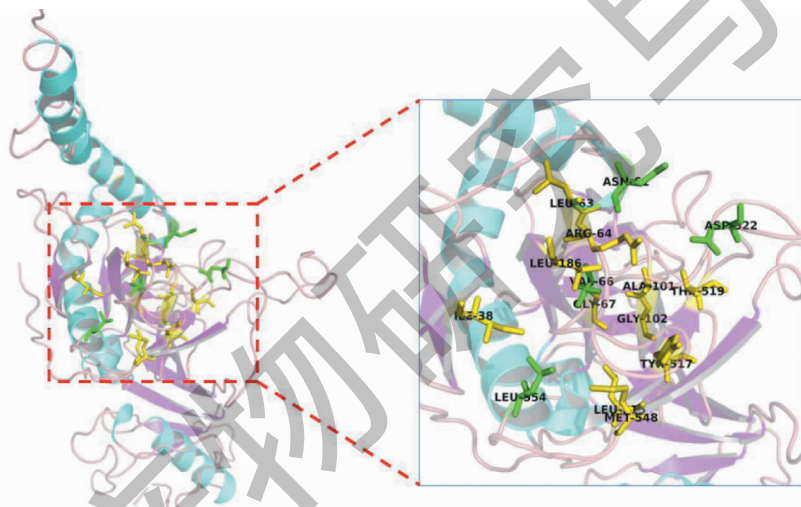


图1 EBOV-sGP 活性口袋的空间结构

Fig.1 Spatial structure of EBOV-sGP active pocket

1.2 中药分子化学结构的预处理

TCMSP 数据库^[9] 中 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 为 7.0 ± 2.0 范围的化合物离子状态, 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 toremifene。

1.3 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^[10]”, “Verber Ro3^[11]” 对中药分子配体库进行首轮筛选。然后, 基于 EBOV-sGP 活性口袋结构 Glide 程序对命中分子进行第二轮筛选, 依次从高通量模式 HTVS-

1.1 靶蛋白 EBOV-sGP 的预处理与活性口袋的确定

从 PDB 数据库下载 EBOV-sGP 结构 (PDB ID 为 5JQ7, 配体为 toremifene)。使用 Schrodinger2018 软件的 Protein Preparation wizard 模块对 5JQ7 结构进行预处理。以 toremifene 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 Elizabeth E. Fry 课题组发表的文章^[8], 选择 Asn61、Val66、Asp522、Leu554 可作为关键氨基酸残基, 并且关注 Ala101、Leu515、Tyr517、Gly67、Gly102、Arg64、Thr519、Met548、Ile38、Leu186、Leu63 与候选药效分子是否产生疏水相互作用、离子键等, 并使用 Pymol 进行可视化处理, 如图 1 所示。

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三轮筛选, 从分子量 (molecular weight, MW)、氢键供体 (hydrogen bond donor, HB donor)、氢键配体 (hydrogen bond acceptor, HB acceptor)、溶剂可及表面积 (solvent accessible surface area, SASA)、K⁺ 通道阻断相关 IC₅₀ (IC₅₀ value for blockage of HERG K⁺ channels, logHERG)、水溶性 (aqueous solubility, logS)、脂水分配系数 (octanol/water partition coefficient, logO/W)、口服生物利用度 (oral bioavailability, OB) 等评估, 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四轮筛选, 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 MM/GBSA 算法优化排序。

1.4 分子动力学模拟

为了从动力学和热力学角度动态分析 EBOV-sGP 与命中分子互作模式,使用 GROMACS 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 在 B3LYP/6-311G** 基组水平计算静电势,再用 Amber 中的 RESP 电荷拟合程序算出。将复合物放入 $10 \times 10 \times 10 \text{ \AA}^3$ 的立方体 TIP3P 水盒中,并加入 Na^+ 和 Cl^- 离子,中和该系统到 0.15 mol/L 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 50 ns 的分子动力学模拟,时间间隔为 2 fs,每 10 ps 保存一次能量和坐标文件。从 35 ~ 50 ns 稳定模拟轨迹中提取 150 帧轨迹文件,计算 MM/PBSA 结合自由能。

2 结果与讨论

2.1 虚拟筛选结果分析

基于靶标 EBOV-sGP 三维空间结构,从 TCMSP 数据库 13 144 个化合物进行多轮虚拟筛选,如图 2 所示,其具备有较强的靶标结合能、良好的成药性等特质。最终确定打分排名前 4 的化合物 MOL006834、MOL000174、MOL002192、MOL012524,具备有较强靶标结合自由能、良好药代动力学参数、

较高生物利用度等特质,结果见表 1 和表 2。随后进行文献检索,发现 MOL000174 属于中药材苍术的苍术醇类化合物,是动脉炎病毒的强抑制剂^[12]; MOL006834 属于中药材鸭跖草的阿魏酸酰胺类化合物,鸭跖草具有抑菌、抗氧化、抗病毒、降血脂、镇痛消炎的作用,其中阿魏酸酰胺类化合物具有抗炎活性^[13]; MOL002192 属于中药材川芎的阿魏酸酯,具有较好的血管扩张作用及抗菌抗癌活性^[14]; MOL012524 属于中药材牛膝的阿魏酰甲氧基酰胺类化合物,在体外药理实验中显示出有效的抗炎活性^[15]。从现有的文献中,4 种命中分子具有不同药理活性,但是还没有直接实验证据揭示与本研究靶点的相关性。

2.2 分子动力学模拟结果分析

计算分子动力学模拟 50 ns 时长轨迹的均方根偏差 (root mean square deviation, RMSD),以此检验体系稳定性,结果见图 3。在模拟运行 35 ns 时长之后,所有体系的 RMSD 的平均值为 0.5 ~ 0.7 nm,且涨落范围都稳定在 0.1 nm 以下,均达到了平衡状态。

通过计算化合物与受体蛋白复合物氨基酸骨架原子在分子动力学模拟过程中的均方根波动值 (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 EBOV-sGP 每个氨基酸残基波动的影响。发现 4 个

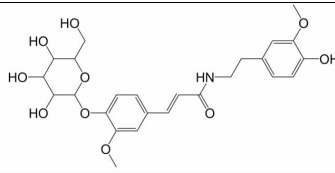
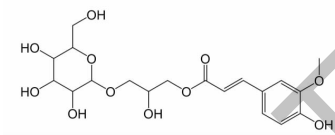
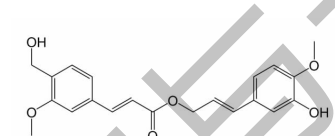
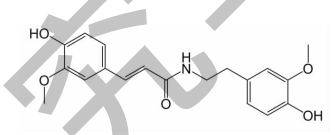
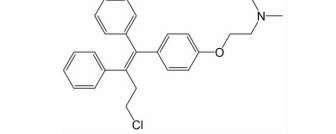


图 2 虚拟筛选流程

Fig. 2 The virtual screening workflow

表 1 候选化合物及阳性对照 Glide 分子对接打分值与 MM/GBSA 结合自由能

Table 1 Docking score calculated by Glide docking function and MM/GBSA binding energy (kcal/mol)

编号 Entry	Glide-HTVS score	Glide-SP score	Glide-XP score	MM/GBSA	化合物结构式 Chemical structure	来源 Resource
MOL006834	-7.532	-8.473	-8.044	-49.29		鸭跖草
MOL000174	-6.667	-7.016	-7.259	-51.5		苍术
MOL002192	-6.486	-8.259	-8.899	-58.05		川芎
MOL012524	-7.711	-7.952	-7.697	-49.41		牛膝
Toremifene	-9.111	-10.075	-10.296	-49.82		

注:埃博拉病毒糖蛋白单体结构在蛋白质结构数据库中的编号:5JQ7。

Note: EBOV-sGP's PDB-ID:5JQ7.

表 2 排名前 4 命中化合物的 ADME 参数预测值

Table 2 ADME predicted parameters of the top 4 hit compounds

分子编号 Entry	分子量 MW ^a	氢键供体 HB donor ^b	氢键受体 HB acceptor ^c	溶剂可及 表面积 SASA ^d	阻断 K ⁺ 通道的 IC ₅₀ 值 logHERG ^e	亲水性 logS ^f	脂水分配系数 logO/W ^g	口服生物利用度 OB ^h (%)
MOL006834	315	3	4.75	635.4	-4.6	-3.4	2.3	81.2
MOL000174	198	1	2.2	493.0	-5.8	-3.5	3.2	100.0
MOL002192	356	2	5	700.6	-6.4	-5.5	3.7	92.7
MOL012524	343	3	5.5	673.2	-6.1	-4.5	2.9	88.0
Toremifene	406	0	2.75	738.1	-7.5	-6.4	6.9	100.0

注:^a数值可接受范围小于 650;^b数值可接受范围不大于 5;^c数值可接受范围不大于 10;^d数值可接受范围为 300 ~ 1 000;^e数值推荐区间低于 -5;^f数值可接受范围为 -7.0 ~ 0.5;^g数值可接受范围为 -2 ~ 6.5;^h数值低于 50 代表较差,高于 80 代表较好。

Note:^aThe acceptable range is less than 650;^bThe acceptable range is less than or equal to 5;^cThe acceptable range is less than or equal to 10;^dThe acceptable range is from 300 to 1 000;^eThe recommended range is less than -5;^fThe acceptable range is from -7.0 to 0.5;^gThe acceptable range is from -2 to 6.5;^hThe value lower than 50 means poor, and the value higher than 80 means better.

抑制剂复合物体系与 toremifene 复合物体系都有相似 RMSF 数值分布趋势,结果见图 4。在铰链区(290 ~ 310)、催化 loop 区(186 ~ 201、99 ~ 104)、Activation β -sheet(514 ~ 520、62 ~ 68)这些结合位点内的残基 RMSF 都明显较小,说明 MOL006834、MOL000174、MOL002192、MOL012524 一样都可以与

EBOV-sGP 结合口袋形成稳定的相互作用,稳定蛋白结合口袋的残基。除此之外,复合物体系(MOL006834、MOL000174、MOL002192、MOL012524、toremifene)与靶蛋白空载体(in an inactive or unbound state, APO form)相比较,含配体分子的蛋白二元结构比空蛋白在活性中心关键氨基酸残基 520 ~ 550 区间具

有更小的 RMSF。

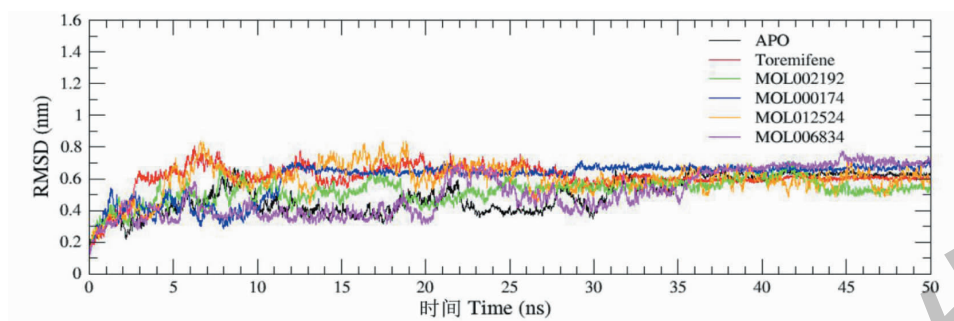
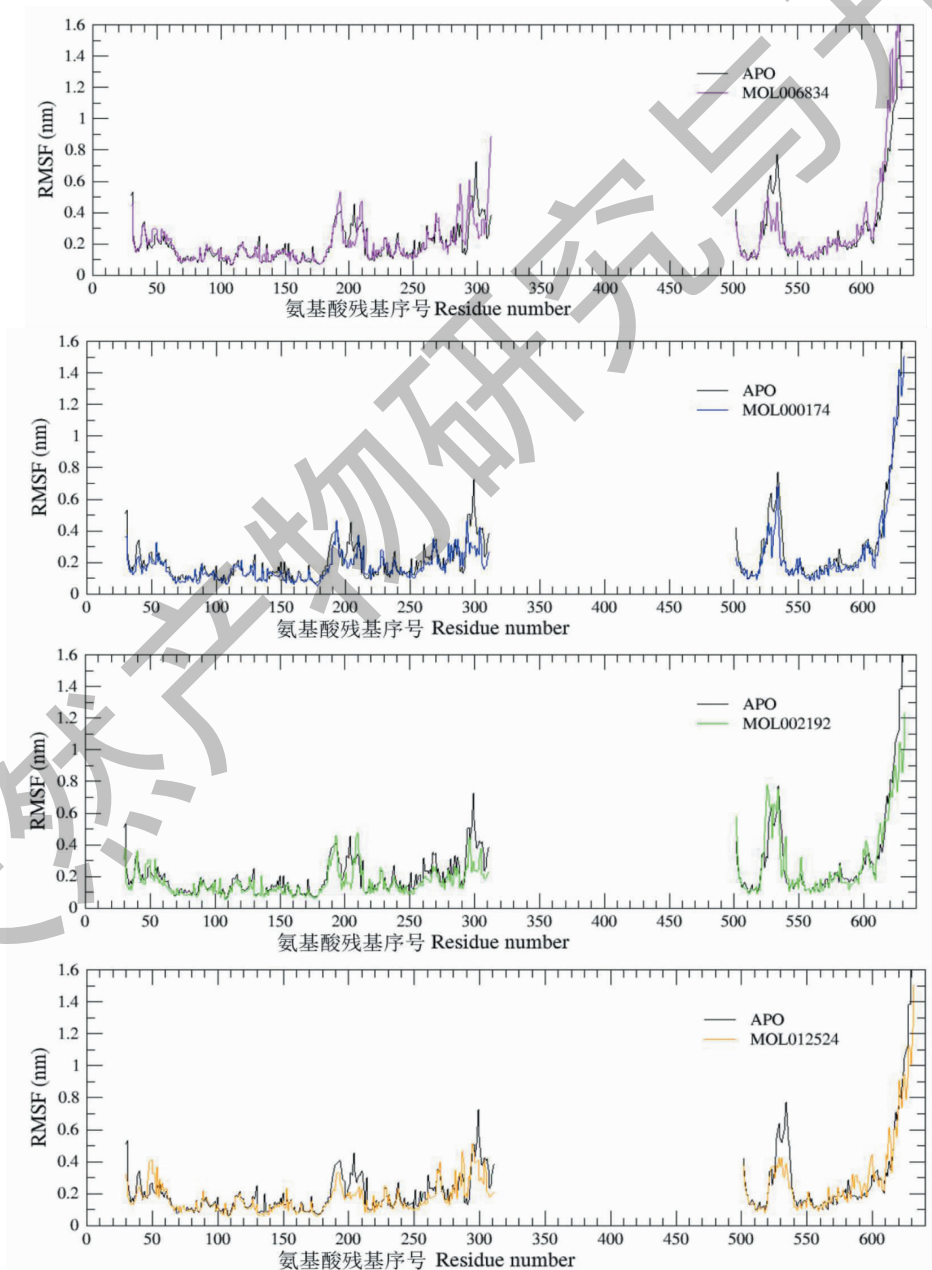


图3 靶蛋白 EBOV-sGP 的氨基酸骨架原子随时间变化的 RMSD 值
Fig.3 The RMSD value of backbone atoms of the target protein EBOV-sGP



续图 4(Continued Fig.4)

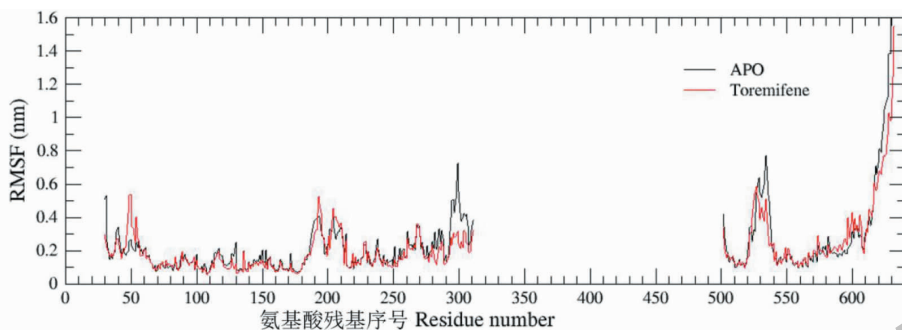


图4 配体-蛋白复合物氨基酸骨架原子的 RMSF 值波动图

Fig. 4 Evolution of the RMSF of the protein backbone during 50 ns simulation

注: EBOV-sGP 包括 A、B 两条链, A 链 30 ~ 311 位点, 共 281 个氨基酸残基; B 链 502 ~ 632 位点, 共 130 个氨基酸残基。Note: EBOV-sGP includes chain A and chain B, there are 281 residues in 30 to 311 Aa of chain A, and 130 residues in 502 to 632 Aa in chain B.

2.3 结合自由能计算结果分析

从 35 ~ 50 ns 分子动力学轨迹中提取 150 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。范德华作用能与静电势能最有利于 4 种命中化合物与 EBOV-sGP 结合, 表面溶剂化作用也有利于结合, 但与范德华作用能和静电势能的数值相比作用较弱, 而极性溶剂化作用不利于体系结合。表明范德华作用成分能在整个作用力占有主导地位, 范德华作用能成分极大地促进了抑制剂与靶标蛋白的结合, 决定了结合自由能的最终排名, 这也是 MOL006834、MOL000174 与靶蛋白 EBOV-sGP 结合自由能具有优势效应的关键因素。

MOL006834 与 MOL000174 来源于相同的化学结构母核, 空间结构相似, 因此范德华作用能、表面溶剂化作用、结合自由能均非常接近; 但是 MOL000174 具有较强的极性溶剂化负效应, 具体数值高达 205.799 kJ/mol, 非常不利于配体与受体结合, 因此与蛋白质结合过程中由于溶剂中的极性溶剂化相互作用很大程度上抵消了静电势能与范德华力之和, 不利于导致最终的结合自由能排名次于 MOL006834, 位居第二。

同样, MOL002192 与 MOL012524 也来源于相同的化学结构母核, 空间结构相似, 因此范德华作用能、表面溶剂化作用、结合自由能均非常接近。MOL012524 具有较强的极性溶剂化负效应, 具体数值为 156.173 kJ/mol, 不利于配体与受体结合, 虽然静电势能也是 -53.094 kJ/mol, 但是与蛋白质结合过程中由于溶剂中的极性溶剂化相互作用终究抵消了静电势能与范德华力之和。因此 MOL012524 具有最弱的范德华作用正效应与较强的极性溶剂化负效应, 导致最终的结合自由能不占优势, 排名第四。

结合自由能数值越小, 表示结合所需能量越低, 越有利于配体与蛋白的结合。比较 4 种化合物与 EBOV-sGP 的结合自由能, 可得出候选抑制剂分子结合靶标强弱关系如下: MOL006834 > MOL000174 > MOL002192 > MOL012524, 均优于阳性对照分子 toremifene。

2.4 候选化合物与靶标的相互作用位点分析

提取复合物分子动力学模拟 50ns 稳定后最后一帧构象, 观察命中化合物与 EBOV-sGP 的空间结合构象。由图 5 可见, 靶蛋白 EBOV-sGP 与 MOL006834 之间的相互作用不仅有范德华力、氢键, 还有

表 3 命中化合物与靶标 EBOV-sGP 的 MM/PBSA 结合自由能

Table 3 MM/PBSA binding energy of 4 selected compounds with target EBOV-sGP (kJ/mol)

化合物 Compound	范德华作用能 Van der Waals energy	静电势能 Electrostatic energy	极性溶剂化作用 Polar solvation energy	表面溶剂化作用 SASA energy	结合自由能 Binding energy
MOL006834	-232.97 ± 16.33	-45.04 ± 13.06	166.94 ± 16.74	-23.43 ± 1.20	-134.50 ± 17.12
MOL000174	-224.93 ± 13.34	-92.86 ± 10.63	205.80 ± 12.85	-21.38 ± 0.75	-133.37 ± 15.50
MOL002192	-200.79 ± 78.94	-27.97 ± 17.27	138.10 ± 68.42	-17.32 ± 6.88	-107.99 ± 36.31
MOL012524	-189.63 ± 13.90	-53.09 ± 9.65	156.17 ± 13.12	-19.40 ± 1.08	-105.95 ± 11.29
Toremifene	-162.39 ± 9.97	-11.13 ± 4.60	100.18 ± 9.94	-23.42 ± 0.77	-96.75 ± 12.89

π -Anion 相互作用、 π -Sulfur 相互作用、 π -Alkyl 相互作用、 π - π 堆积作用。MOL006834 与氨基酸残基 Met548、Arg200、Pro533、Pro537、Ala539、Asn204 形成常规氢键,并且同样与 Glu201、Asn204 均形成一个碳氢键。与 Ala101、Pro198、Pro202 分别形成一个 π -Alkyl 相互作用,与 Glu100 形成 π -Anion 相互作

用,与 Met548 形成 π -Sulfur 相互作用,与 Tyr517 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 EBOV-sGP 蛋白的活性部位紧密结合,形成稳定的复合物,就可以有效阻止外来病毒粒子附着并侵入宿主细胞。

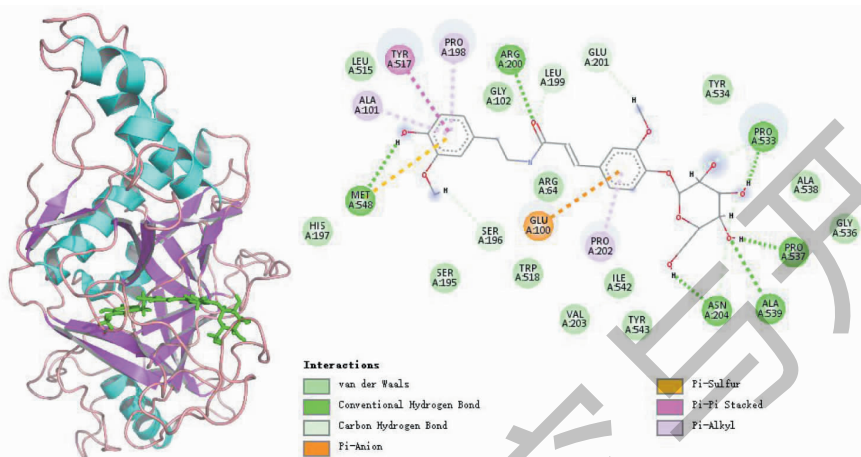


图5 靶蛋白 EBOV-sGP 与 MOL006834 相互作用

Fig. 5 The interaction between the target protein EBOV-sGP and MOL006834

由图 6 可见,靶蛋白 EBOV-sGP 与 MOL000174 之间的相互作用不仅有氢键、范德华力,还有 π - π 堆积作用、 π -Alkyl 相互作用。MOL000174 与氨基酸残基 Arg64、Arg164、Arg200、Glu201、Thr520 形成常规氢键,并且同样与 Gly67、Ala101、Arg200、Thr519、Glu201、Leu199 形成碳氢键,还与 Tyr517 形成 π - π 堆积作用,与 Ala101 形成 π -Alkyl 相互作用。这些作用力使结合能降低,亲和力增加,候选单体与 EBOV-sGP 活性空腔部位紧密结合,形成稳定

的复合物,就可以有效阻止外来病毒粒子附着并侵入宿主细胞。但是还可以观察到氨基酸残基 Arg64 与 Glu100 之间存在着不利相互作用,不利于构象稳定。

由图 7 可知,靶蛋白 EBOV-sGP 与 MOL002192 之间的主要作用力是范德华力、氢键、 π -cation 相互作用、amide- π 堆积作用、 π -Alkyl 相互作用。MOL002192 与靶标氨基酸残基 Ser195、Asp192、Arg200、Thr520、Asn61、His197 均形成氢键,其中

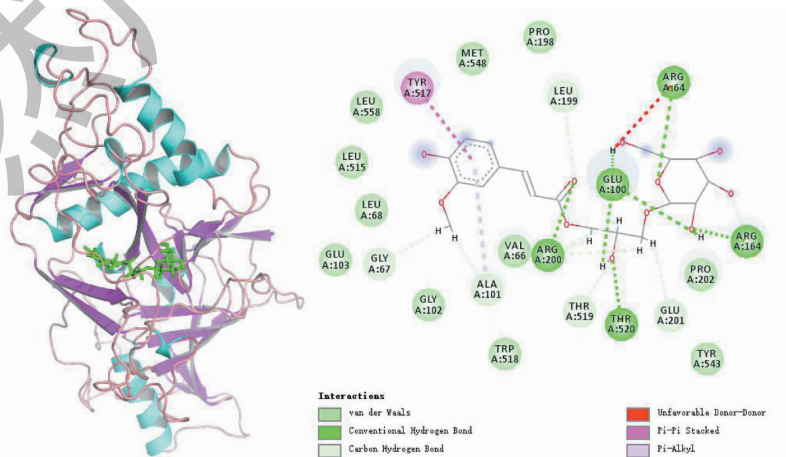


图6 靶蛋白 EBOV-sGP 与 MOL000174 相互作用

Fig. 6 The interaction between the target protein EBOV-sGP and MOL000174

Asn61 形成了 2 个常规氢键, Ser195 分别形成了常规氢键和碳氢键各 1 个。与碱性氨基酸残基 Arg200、Arg64 均形成 π -anion 相互作用, 与氨基酸

残基 Pro202 形成 π -Alkyl 相互作用, 并且 His197 形成 amide- π 堆积作用, 与这些作用力促进复合物的形成。

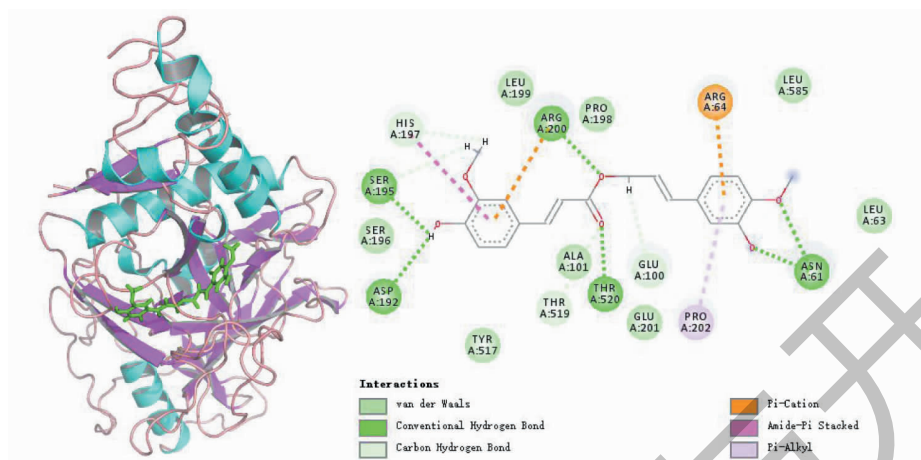


图 7 靶蛋白 EBOV-EBOV-sGP 与 MOL002192 相互作用

Fig. 7 The interaction between the target protein EBOV-sGP and MOL002192

由图 8 可见, 靶蛋白 EBOV-sGP 与 MOL012524 之间的相互作用不仅有范德华力、氢键, 还有 π -Alkyl 相互作用。MOL012524 与氨基酸残基 Glu100、Arg200、Glu201、His197 形成常规氢键, 同样并且与 Arg164、Asn98、Tyr99 均形成一个碳氢键, 与 Thr519 形成了一个 π -Donor 氢键, 其中 Glu100 形成了 2 个常规氢键。与 Arg64、Ala101 分别形成一个 π -Alkyl 相互作用。这些作用力使结合能降低, 亲和力增加, 候选单体与 EBOV-sGP 蛋白的活性部位紧密结合, 形成稳定的复合物, 就可以有效阻止外来病毒粒子附着并侵入宿主细胞。

氢键在生物大分子中起着重要作用, 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^[16]。Discovery Studio Visualizer 生成的结果显示, 阳性对照分子 toremifene 与氨基酸残基 Glu100、Thr520 形成氢键相互作用, 产生的氢键数目为 2 个; 然而 4 个候选单体均与 EBOV-sGP 氨基酸残基 Glu201 或 Asn61 形成氢键相互作用, 甚至形成双氢键, 这些残基均是候选单体的氢键受体, 并且常规氢键总数均大于 4 个 (其中氢键数 MOL006834 为 6 个, MOL00174 为 7 个, MOL002192 为 6 个, MOL012524 为 5 个)。借助 g-hbond 程序统计氢键数目, 计算 50 ns 模拟过程中 EBOV-sGP 与抑制剂分子间的氢键数目

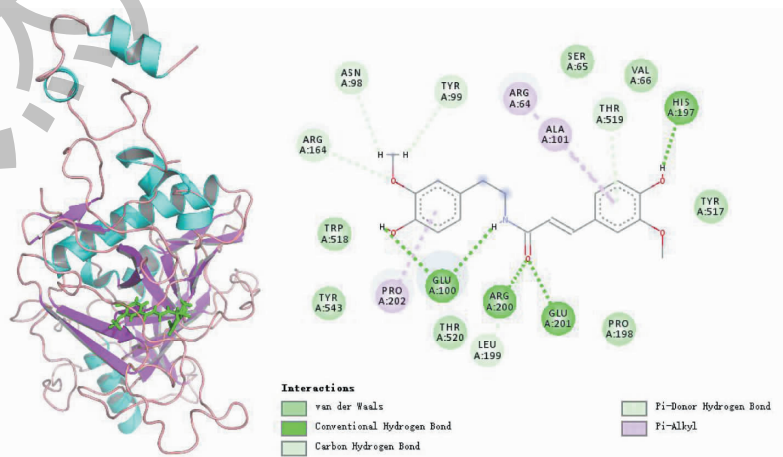


图 8 靶蛋白 EBOV-sGP 与 MOL012524 相互作用

Fig. 8 The interaction between the target protein EBOV-sGP and MOL012524

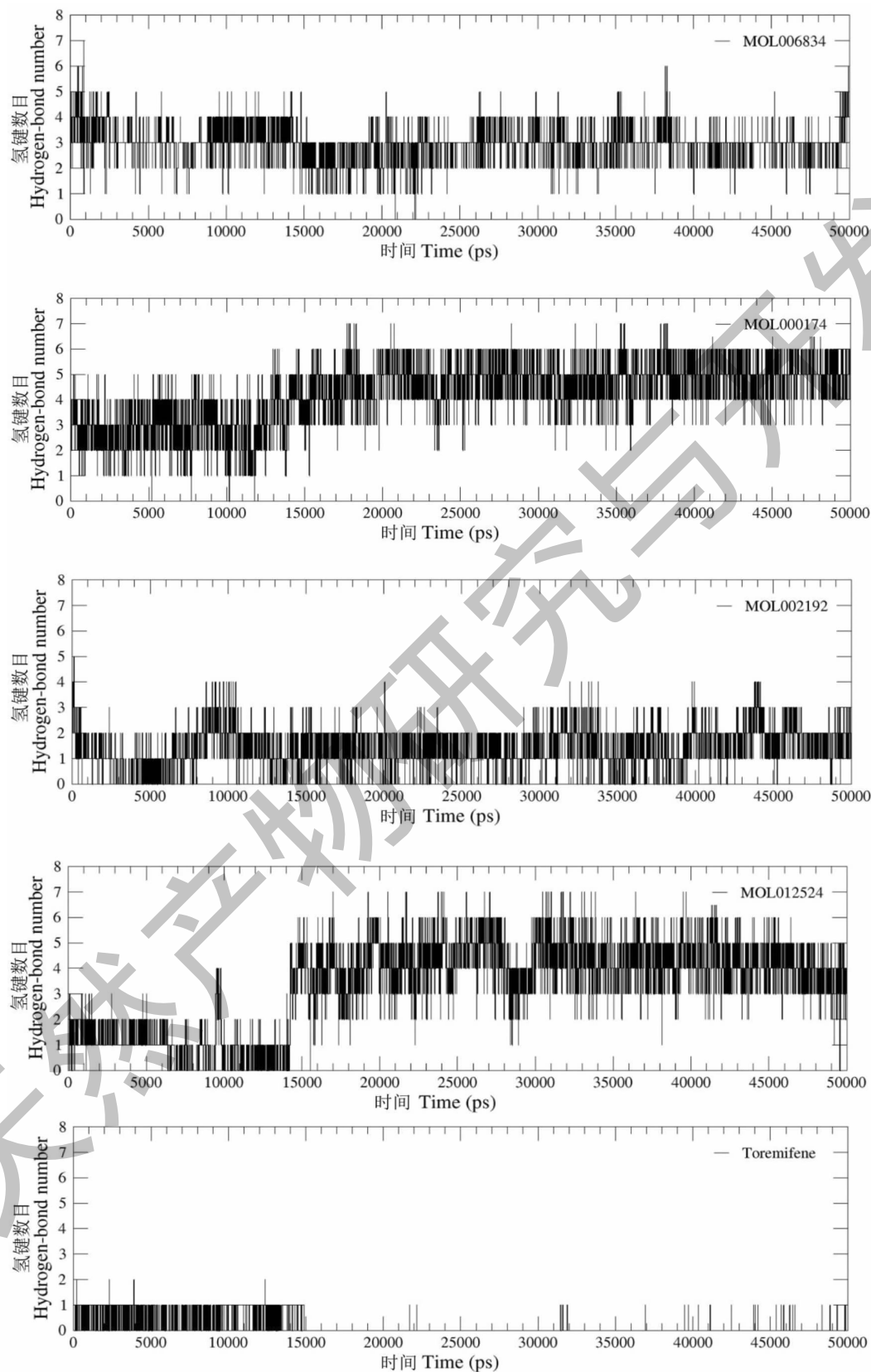


图9 50 ns 动力学模拟过程中命中分子与 EBOV-sGP 形成的氢键数目

Fig. 9 The number of hydrogen bonds formed between molecules and EBOV-sGP during 50 ns dynamic simulation

以及氢键距离或角度的分布。在 Gromacs 中判断氢键要满足两个条件,距离和角度,g-hbond 统计了满足距离限制的所有原子对,并且也同时满足角度限制,即可判定为氢键,统计结果如图 9 所示。总体而言,与阳性对照分子 toremifene 相比,命中的候选分子与 EBOV-sGP 结合氢键数目更加具有优势,均具有更强的相互作用,推测可能具有更高的抑制作用。

图 10 展示了阳性对照分子 toremifene 与靶标的结合构象,可以观察到 toremifene 与 EBOV-sGP 氨基酸残基 Val66、Glu515 均形成 π -sigma 相互作用,氨基酸残基 Ala101 同时与配体分子的两个苯环形成

π -alkyl 相互作用,与碱性氨基酸 Arg64 形成 π -Cation 堆积作用,3 个 Leu 氨基酸残基 184、186、558 同时与配体分子末端氧原子形成 alkyl 相互作用,这些作用力促进复合物的形成。同样在 4 个候选单体与靶蛋白的相互作用分析中,MOL006834、MOL002192 均与氨基酸残基 Pro202 形成 π -Alkyl 相互作用,MOL000174、MOL012524 与碱性氨基酸 Ala101 也同样均形成 π -Alkyl 相互作用,这些相互作用在候选配体分子与 EBOV-sGP 单体空腔的残基结合中也起了辅助有利作用。

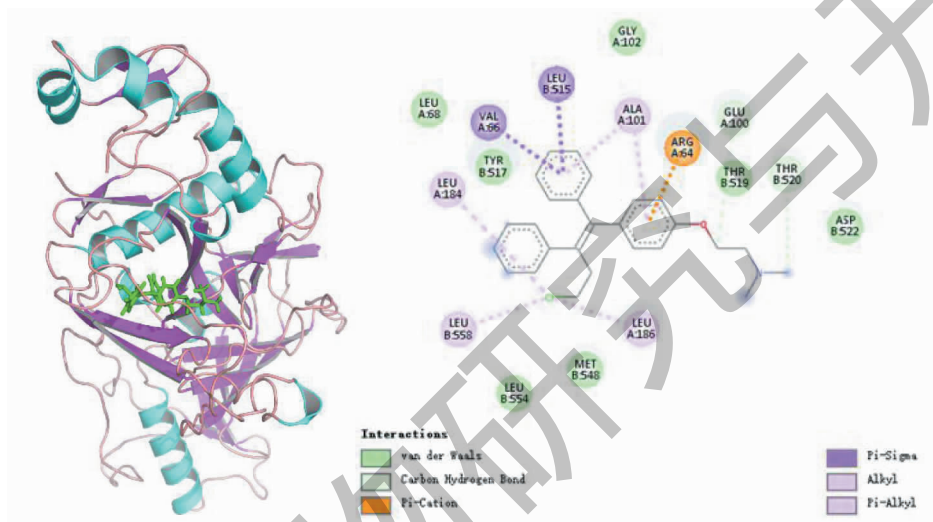


图 10 靶蛋白 EBOV-sGP 与 toremifene 相互作用

Fig. 10 The interaction between the target protein EBOV-sGP and toremifene

3 讨论与结论

Ahmad 课题组于 2017 年首次基于 EBOV-GP 跨膜糖蛋白三聚体空腔的三维结构为靶标,从 Molecule 化合物库 (<https://molecule.com/database/>) 进行大规模虚拟筛选抑制剂,采用 Auto Dock Vina 与 Flex-X 两种分子对接软件,并结合网上在线分析工具 admetSAR (<http://www.admetexp.org>) 进行 ADMET 过滤,找到了三个命中分子^[17]。从 2000 年以来,不断有埃博拉病毒疫情暴发,抗病毒药效分子的研发仍是全球医药研究热点。因此本研究为了从中药天然产物库中挖掘抗病毒活性分子,借鉴了此文章的研究思路,靶标从 EBOV-GP 三聚体改为单体,并重新设计了多轮虚拟筛选的实验方法,筛选策略使用成本更高的算法,尽可能提高预测的精度,降低虚拟筛选假阳性率,以及更好地预测结合构象。除此之外,分子动力学模拟和结合自由能计算也是研究蛋白质

与抑制剂相互作用强度的重要工具,采用 GROMACS 软件进行分子动力学模拟,并使用 MM/PBSA 算法预测体系稳定后的结合自由能,这类方法不仅能在原子和分子的层次上给出蛋白质和抑制剂的相互作用细节,而且还能便于阐明蛋白质和抑制剂复合体的结构亲和能关系。

为了有效干扰 EBOV-sGP 构象转变,本研究基于 EBOV-GP 单体活性空腔为靶标,从中药天然产物库中挖掘抑制剂分子,有助于阻断宿主-病毒膜融合。首先遵循 Lipinski 五原则、Verber 三原则,过滤掉违反类药性原则的化合物;然后采用 Glide 软件进行三种不同精度的分子对接计算,即高通量 HTVS、标准精度 SP、超精密 XP 等三种模式,逐步缩小候选化合物库的规模;随后采用 Schrodinger 的 QikProp 程序预测命中分子 ADMET 性质,较 admetSAR 进行过滤策略,新增了肝毒性 HERG_{IC₅₀}、口服

生物利用度等两项药代动力学与药物安全性分析的关键指标,并结合 Lipinski Ro5、Verber Ro3 预测值,做出相对全面的类药性评估;接着在分析靶标蛋白与命中分子结合方式过程中,使用高性能 GPU 加速技术模拟在真实生理盐水溶剂条件下生物大分子的运动状态,进行了 50 ns 时长的分子动力学模拟,并使用了 MM/GBSA 与 MM/PBSA 两种算法预测结合自由能,提升富集精准度。最后根据 XP-Glide 评分、RMSD、RMSF、MM/GBSA 结合能、MM/PBSA 结合能、ADME 预测,命中分子 MOL006834、MOL000174、MOL002192、MOL012524 表现出对 EBOV-sGP 活性位点显示出极好的结合亲和力和良好的类药性;从原子与分子层次上预测命中分子 MOL006834、MOL000174、MOL002192、MOL012524 与跨膜糖蛋白的真实结合构象,包括氢键、疏水相互作用、离子键、 π - π 堆积力等,与关键氨基酸残基显示出极好的相互作用。这四个命中分子均可在商业上购买获得,也可作为 EBOV-sGP 抑制剂先导化合物进行开发利用,以及药物活性实验验证。需要指出的是,埃博拉病毒具有高致死率和高传播性,相关生物安全要求别极其严格,抗病毒活性测试实验需要在生物安全 4 级实验室内进行,国内具备此级别的实验室并不多,制约了有关埃博拉病毒的基础研究。因此本研究成果仅局限于药物开发的早期阶段,缺乏抗病毒活性验证、ADMET 评估等实验数据支撑,后期需要联合相关的病毒学研究团队,借助专业的病毒学研究平台开展安全有效的生物学验证。

参考文献

- 1 Côté M, Misasi J, Ren T, et al. Small molecule inhibitors reveal Niemann-Pick C1 is essential for Ebola virus infection [J]. *Nature*, 2011, 477(7364):344-348.
- 2 Lee JE, Saphire EO. Ebolavirus glycoprotein structure and mechanism of entry[J]. *Future Virol*, 2009, 4:621-635.
- 3 Shaikh F, Zhao Y, Alvarez L, et al. Structure-based in silico screening identifies a potent ebolavirus inhibitor from a traditional Chinese medicine library [J]. *J Med Chem*, 2019, 62(6):2928-2937.
- 4 Newman DJ, Cragg GM. Natural products as sources of new drugs over the 30 years from 1981 to 2010 [J]. *J Nat Prod*, 2012, 75:311-335.
- 5 Li J, Wei BQ, Li KX, Shu XY, et al. Screening of RdRp inhibitors against SARS-CoV-2 based on prescription mining and pharmacophore [J]. *Nat Prod Res Dev* (天然产物研究与开发), 2020, 32:1981-1991.
- 6 Zhang X, Perez-Sanchez H, Lightstone FC. A comprehensive docking and MM/GBSA rescoring study of ligand recognition upon binding antithrombin [J]. *Curr Top Med Chem*, 2017, 17:1631-1639.
- 7 Kongsted J, Ryde U. An improved method to predict the entropy term with the MM/PBSA approach [J]. *J Comput Aided Mol Des*, 2009, 23(2):63-71.
- 8 Shaikh F, Zhao Y, Alvarez L, et al. Structure-based in silico screening identifies a potent ebolavirus inhibitor from a traditional Chinese medicine library [J]. *J Med Chem*, 2019, 62:2928-2937.
- 9 Ru J, Li P, Wang J, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines [J]. *J Cheminformatics*, 2014, 6:13.
- 10 Lipinski CA, Lombardo F, Dominy BW, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings [J]. *Adv Drug Deliv Rev*, 1997, 23(1-3):3-25.
- 11 Veber DF, Johnson SR, Cheng HY, et al. Molecular properties that influence the oral bioavailability of drug candidates [J]. *J Med Chem*, 2002, 45:2615-2623.
- 12 Kraus GA, Dong P, Qu Y, et al. A direct synthesis of atracytoldinol, a potent inhibitor of PRRSV, and its biological evaluation [J]. *Tetrahedron Lett*, 2016, 57:5185-5187.
- 13 Chae HS, Yoo H, Kim YM, et al. Anti-inflammatory effects of 6,8-diprenyl-7,4'-dihydroxyflavanone from *Sophora tonkinensis* on lipopolysaccharide-stimulated RAW 264.7 cells [J]. *Molecules*, 2016, 21(8):1049.
- 14 Li SJ, Zhang YC, Su PY, et al. Advances in Research of on Coniferyl Ferulate [J]. *Chin J Exp Tradit Med Form* (中国实验方剂学杂志), 2011, 17(3):229-231.
- 15 Lai YW, Wang SW, Hu YY, et al. Anti-inflammatory alkaloids from the root bark of *Hernandia nymphaeifolia* [J]. *Phytochemistry*, 2020, 173:112326.
- 16 Du X, Li Y, Xia YL, et al. Insights into protein-ligand interactions: mechanisms, models, and methods [J]. *Int J Mol Sci*, 2016, 17(2):144.
- 17 Onawole AT, Kolapo TU, Sulaiman KO, et al. Structure based virtual screening of the Ebola virus trimeric glycoprotein using consensus scoring [J]. *J Comput Chem*, 2018, 72:170-180.