

【实验研究】

以跨膜丝氨酸蛋白酶2(TMPS2)为受体挖掘治疗新型冠状病毒肺炎(COVID-19)潜在单体化合物

徐昕怡^{1,2}, 曹学帅², 龙茜², 刘乐平³, 唐吉伟¹, 顾晖^{1*}

1. 湖南省脑科医院 麻醉手术科, 湖南 长沙 410006

2. 湖南中医药大学 药理实验室, 湖南 长沙 410208

3. 中南大学湘雅三医院 输血科, 湖南 长沙 410000

摘要: 目的 探讨以跨膜丝氨酸蛋白酶2(TMPS2)为受体的单体化合物治疗新型冠状病毒肺炎(COVID-19)的作用机制。方法 借助BindingDB数据库检索作用于TMPS2受体的单体化合物。通过UniProt、GeneCards等数据库查询单体化合物作用靶点对应的基因名,进而运用Cytoscape 3.2.1构建化合物-靶点(基因)网络,通过WebGestalt数据库进行GO功能注释和Pathway通路分析,预测其作用机制。运用SWISS-MODEL进行TMPS2蛋白质三级结构预测,利用AutoDock软件进行化合物与TMPS2受体的分子对接。结果 通过BindingDB数据库发现3个作用于TMPS2的单体化合物:(2S)-1-[(2R)-2-(苄基磺酰基氨基)-5-(二氨基亚甲基氨基)戊酰基]-N-[(4-氨基甲酰基苯基)甲基]吡咯烷-2-羧酰胺(CID-46899577,分子式 $C_{26}H_{36}N_8O_4S$)、4-[[[(5R)-5-(苄基磺酰基氨基)-8-(二氨基亚甲基氨基)氨基-4-氧代辛烷-3-基]氨基]甲基]苯磺酰亚胺(CID-56677007,分子式 $C_{24}H_{35}N_7O_3S$)、4-[[[(4S,6R)-6-(苄基磺酰基氨基)-9-(二氨基亚甲基氨基)-5-氧肟酮-4-基]氨基]甲基]苯磺酰亚胺(CID-56663319,分子式 $C_{25}H_{37}N_7O_3S$);单体化合物-靶点网络中靶点58个,GO功能富集分析得到GO条目380个($P<0.05$),其中生物过程(BP)条目310个,细胞组成(CC)条目14个,分子功能(MF)条目56个。KEGG通路富集筛选得到26条信号通路,涉及干扰素- γ 信号传导途径、转化生长因子- β (TGF- β)信号通路、白介素信号通路等。分子对接结果显示,TMPS2与 $C_{26}H_{36}N_8O_4S$ 、 $C_{24}H_{35}N_7O_3S$ 及甲磺酸卡莫司他都能自发结合,且2种化合物与甲磺酸卡莫司他结合能相近。结论 单体化合物 $C_{26}H_{36}N_8O_4S$ 、 $C_{24}H_{35}N_7O_3S$ 与TMPS2能自发结合,作用于F2、PLG、PLAU、PLAT、HSP90AA1、XIAP、AKT1、AKT2、AKT3等靶点调节多条信号通路,有可能对COVID-19有治疗作用。

关键词: 跨膜丝氨酸蛋白酶2; 新型冠状病毒肺炎; 丝氨酸蛋白酶抑制剂; 网络药理学; 分子对接; $C_{26}H_{36}N_8O_4S$; $C_{24}H_{35}N_7O_3S$; 甲磺酸卡莫司他

中图分类号: R962.2

文献标志码: A

文章编号: 1674-6376(2020)05-0813-09

DOI: 10.7501/j.issn.1674-6376.2020.05.004

Potential monomer compounds for treatment of corona virus disease 2019 (COVID-19) by transmembrane serine proteinase 2 (TMPS2)

XU Xinyi^{1,2}, CAO Xueshuai², LONG Xi², LIU Leping³, TANG Jiwei¹, GU Hui¹

1. Department of Anesthesiology, Hunan Brain Hospital, Changsha 410006, China

2. Pharmacology Laboratory of Hunan University of traditional Chinese medicine, Changsha 410208, China

3. Department of blood transfusion, Xiangya Third Hospital, Central South University, Changsha 410000, China

Abstract: Objective To explore the mechanism of monomeric treatment of corona virus disease 2019 (COVID-19) by using transmembrane serine protease 2 (TMPS2) as a receptor. **Methods** The monomer compounds acting on the TMPS2 receptor were searched by the BindingDB database. Use UniProt, GeneCards and other databases to query the gene name corresponding to the target of the monomer compound, and then use Cytoscape 3.2.1 to build a compound-target (gene) network. Use the WebGestalt database to perform GO function annotation and Pathway analysis to predict its mechanism. The third-order structure of TMPS2

收稿日期: 2020-03-20

基金项目: 湖南省卫计委科研计划课题(B20180102、B20180714)

第一作者: 徐昕怡,女,湖南中医药大学2017级中医学5+3专业学生。E-mail: 840494992@qq.com

*通信作者: 顾晖,男,副主任医师。E-mail: 13975130528@139.com

protein was predicted by SWISS-MODEL, and the molecular docking between the compound and TMPRSS2 receptor was carried out by auto dock software. **Results** Three monomer compounds acting on TMPRSS2 were found through the BindingDB database: (2S)-1-[(2R)-2-(Benzylsulfonylamino)-5-Guanidino-Pentanoyl]-N-[(4-Carbamimidoylphenyl)methyl]pyrrolidine-2-Carboxamide (CID-46899577, Molecular formula: $C_{26}H_{36}N_8O_4S$), 4-[[[(5R)-5-(Benzylsulfonylamino)-8-(diaminomethylideneamino)-4-oxooctan-3-yl]amino]methyl]benzenecarboximidamide (CID-56677007, Molecular formula: $C_{24}H_{35}N_7O_3S$), 4-[[[(4S, 6R)-6-(Benzylsulfonylamino)-9-(diaminomethylideneamino)-5-oxononan-4-yl]amino]methyl]benzenecarboximidamide (CID-56663319, Molecular formula: $C_{25}H_{37}N_7O_3S$). There were 58 targets in the monomer compound-target network, and 380 GO entries were obtained by GO functional enrichment analysis ($P < 0.05$), including 310 biological process (BP) entries, 14 cell composition (CC) entries, and molecules. There are 56 function (MF) entries. The KEGG pathway was enriched and screened to obtain 26 signaling pathways, which involved the interferon- γ signaling pathway, the TGF-beta signaling pathway, and the interleukin signaling pathway. The molecular docking results show that TMPRSS2 can spontaneously bind to the three compounds carnoselita mesylate $C_{26}H_{36}N_8O_4S$, and $C_{24}H_{35}N_7O_3S$. **Conclusion** Monomeric compounds $C_{26}H_{36}N_8O_4S$ and $C_{24}H_{35}N_7O_3S$ combined with TMPRSS2 act on targets such as F2, PLG, PLAU, PLAT, HSP90AA1, XIAP, AKT1, AKT2, and AKT3 to regulate multiple signaling pathways, which may have therapeutic effects on COVID-19.

Key words: transmembrane serine protease 2; novel coronavirus pneumonia; serine protease inhibitors; network pharmacology; molecular docking; $C_{26}H_{36}N_8O_4S$; $C_{24}H_{35}N_7O_3S$; carnoselita mesylate

自2019年以来,世界上一些国家爆发了一种新型冠状病毒感染的肺炎,由于病原体传播速度快,感染性极强,导致该肺炎迅速传播。这种病毒是继SARS、MERS冠状病毒之后发现的一种新型冠状病毒。2020年2月11日,世界卫生组织将新型冠状病毒肺炎命名为COVID-19,同时将新型冠状病毒命名为SARS-CoV-2。COVID-19的临床表现主要为发热、干咳、乏力,严重者可进展为急性呼吸窘迫综合征、脓毒血症、多器官衰竭,甚至死亡^[1-2]。

该病毒尚无特效药治疗,目前国家颁布第7版诊疗方案主要为西医的对症支持结合中医辨证治疗,国家也报道了相关疫苗开展工作。SARS-CoV-2的研究对疫情的控制具有重大意义,目前上海科技大学饶子和课题组已对病毒的晶体结构进行了解析,西湖大学周强课题组对病毒S蛋白作用于血管紧张素转换酶2(ACE2)靶点的结构机制进行了解析^[3]。近日,德国科学家Markus Hoffmann等^[4]发现SARS-CoV-2进入细胞需要蛋白跨膜丝氨酸蛋白酶2(TMPRSS2)的辅助,且这一过程能被蛋白酶抑制剂甲磺酸卡莫司他阻断。另有研究发现,除了甲磺酸卡莫司他是TMPRSS2的抑制剂,还有3种存在抑制作用的化合物^[5]。TMPRSS2作为S蛋白的启动剂,对其进行深入研究,尤其是挖掘其抑制剂,对抗病毒干预治疗有关键意义,对疫情的控制也有深远影响。

本研究对TMPRSS2进行了蛋白质性质的分析以及三级结构的预测,收集3种对TMPRSS2具有抑制作用的化合物的靶点,将化合物与TMPRSS2进

行了分子对接,旨在发现有效的可能治疗COVID-19的潜在单体化合物。

1 方法

1.1 单体化合物收集

通过BindingDB数据库(<http://www.bindingdb.org/bind/index.jsp>)输入TMPRSS2全称“Transmembrane protease serine 2”为关键词,检索作用于TMPRSS2的化合物。

1.2 作用靶点的获取

在PubChem数据库(<https://pubchem.ncbi.nlm.nih.gov/>)检索在BindingDB中得到的化合物以收集3个化合物的靶点,同时将TMPRSS2的抑制剂化合物的SMLIES格式结构上传到在线预测单化合物作用靶点数据库SwissTargetPrediction(<http://www.swisstargetprediction.ch/>),将物种设定为homo sapiens,进行反向药效团匹配预测化合物靶点。

1.3 靶点-网络的构建

为观察化合物作用靶点间的相互作用关系,利用在线蛋白互作关系分析数据库String 11.0(<https://string-db.org/>, 11.0)^[6]分析化合物的作用靶点之间的相互作用关系,将结果导入网络可视化软件Cytoscape 3.7.2(<http://www.cytoscape.org/>, 版本:3.6.1)^[7]进行可视化处理,进一步分析蛋白质与蛋白质之间相互作用关系的拓扑学性质。

1.4 通路和GO分析

为了进一步了解上述筛选出的核心靶点的功能以及在信号通路中的作用,将筛选得到的单体化合物作用靶点导入WebGestalt数据库(<http://www.webgestalt.org/>)。

org/), 设定阈值 $P < 0.05$, 进行GO生物学过程富集分析和KEGG信号通路富集分析, 并用在线绘图网站Omicshare Tools (<http://www.omicshare.com/tools/index.php/>) 将其结果以高级气泡图的形式呈现。

1.5 TMPRSS2 三级结构预测

从 Uniprot (<https://www.uniprot.org/>) 获取 TMPRSS2 的氨基酸序列 FANSTA 文件导入 SWISS-MODEL (<https://swissmodel.expasy.org/interactive#sequence>) 进行蛋白质三级结构预测, 根据 GMQE 和 QMEAN 选择合适的模型, 下载对应的 PDB 文件^[8-9]。将 TMPRSS2 的蛋白质序列输入 ProtParam 在线工具 (<http://expasy.org/tools/protparam.html>) 进行分析得到 TMPRSS2 的基本理化性质。

1.6 成分-靶点分子对接

利用 ZINC (<http://zinc.docking.org/>) 得到收集的 3 个单体化合物以及甲磺酸卡莫司他结构的 *.mol2 格式, 由于 4-[[[(4*S*, 6*R*)-6-(苄基磺酰基氨基)-9-(二氨基亚甲基氨基)-5-氧肟酮-4-基]氨基]甲基]苯磺酰亚胺分子性质不稳定, 所以只对剩下 3

个化合物进行对接。运用 PyMOL 软件对 TMPRSS2 蛋白质进行去水、加氢等操作, 利用 Auto Dock 软件将化合物及蛋白质格式转换为 *.pdbqt 格式后进行对接。将对接结果导入 PyMOL 进行可视化处理。

2 结果

2.1 单体化合物的发现

通过 BindingDB 数据库发现 3 个作用于 TMPRSS2 的单体化合物: (2*S*)-1-[(2*R*)-2-(苄基磺酰基氨基)-5-(二氨基亚甲基氨基)戊酰基]-*N*-[(4-氨基甲酰基苯基)甲基]吡咯烷-2-羧酰胺 (CID-46899577, 分子式 $C_{26}H_{36}N_8O_4S$)、4-[[[(4*S*, 6*R*)-6-(苄基磺酰基氨基)-9-(二氨基亚甲基氨基)氨基-4-氧代辛烷-3-基]氨基]甲基]苯磺酰亚胺 (CID-56677007, 分子式 $C_{24}H_{35}N_7O_3S$)、4-[[[(4*S*, 6*R*)-6-(苄基磺酰基氨基)-9-(二氨基亚甲基氨基)-5-氧肟酮-4-基]氨基]甲基]苯磺酰亚胺 (CID-56663319, 分子式 $C_{25}H_{37}N_7O_3S$), 二维结构式见图 1。

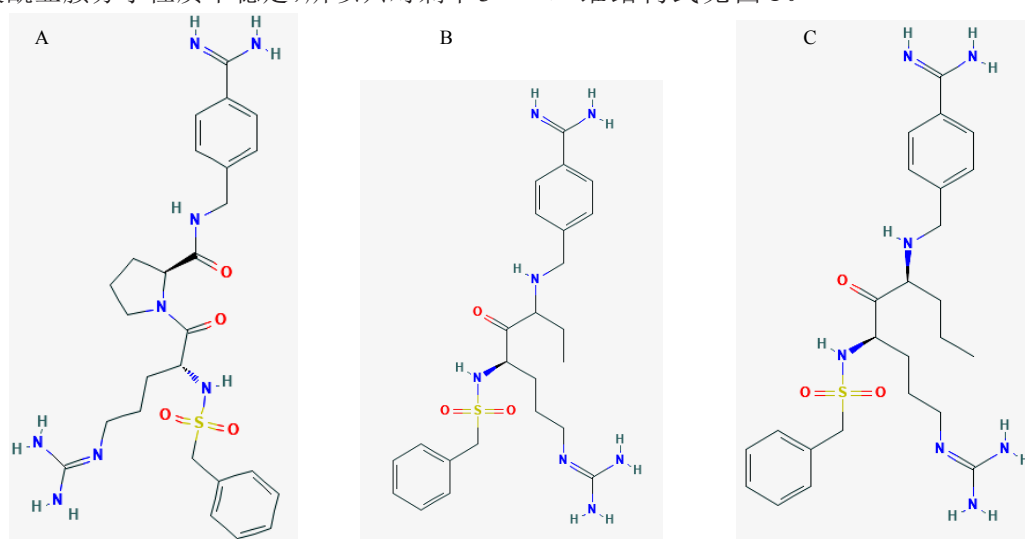


图1 $C_{26}H_{36}N_8O_4S$ (A)、 $C_{24}H_{35}N_7O_3S$ (B)、 $C_{25}H_{37}N_7O_3S$ (C)二维结构式

Fig. 1 Two dimensional structure of $C_{26}H_{36}N_8O_4S$ (A), $C_{24}H_{35}N_7O_3S$ (B), and $C_{25}H_{37}N_7O_3S$ (C)

2.2 单体化合物潜在靶点信息

在 PubChem 数据库和 Swiss Target Prediction 数据库收集整理后得到 CID-46899577、CID-56677007、CID-56663319 的靶点分别为 38 个、30 个、37 个, 共 105 个, 除重后得到 58 个相关靶点。

2.3 单体化合物作用靶点网络分析

将查询得到的 58 个靶点导入 STRING 数据库通过 Cytoscape 软件中可视化, 见图 2。利用 Network Analyzer^[10] 分析 PPI 网络拓扑学属性, 蛋白互作网络包含 58 个节点, 455 条边, 其中节点表示蛋白, 每条边则表示蛋白与蛋白之间的相互作用关

系, 靶点的直径越大表示与其相互作用的蛋白越多、关联度越强, 其中平均节点度为 15.690, 网络密度为 0.275。利用 cytoscape 软件中的 cytoHubba 功能中的 MCC 算法分析出网络中的 Hub 基因^[11], 见表 1(排名前 20)。

2.4 单体化合物作用靶点的生物信息学分析

运用 WebGestalt 数据库进行 GO 功能注释和 Pathway 通路分析。GO 功能注释是通过生物过程(BP)、细胞组分(CC)和分子功能(MF)对基因进行注释和分类, 见图 3。每个生物过程、细胞成分和分子功能类别分别由红色、蓝色和绿色的条形表

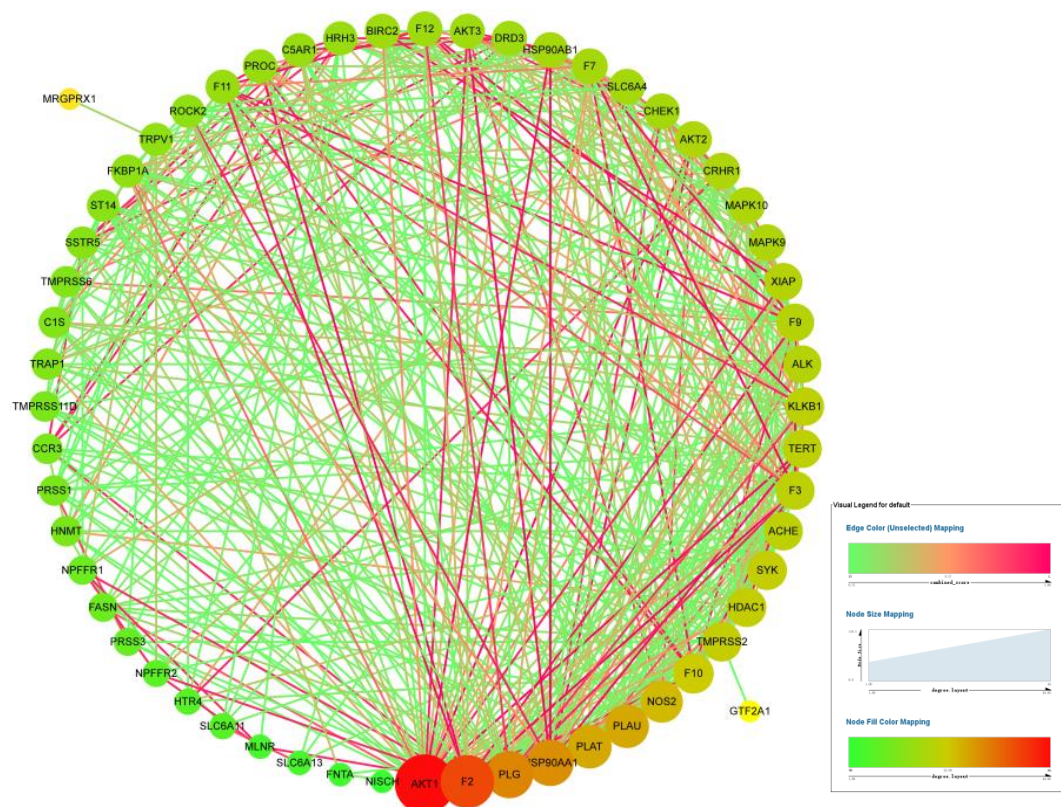


图2 单体化合物作用靶点网络

Fig. 2 Network of targets of monomer compounds

表1 按MCC方法排名的PPI网络前20名

Table 1 Top 20 in PPI network ranked by MCC method

靶点名称	正式基因 符号	得分
Prothrombin	F2	8 450 374
Tissue-type plasminogen activator	PLAT	8 308 712
Coagulation factor X (human)	F10	8 277 222
Plasma kallikrein	KLKB1	8 268 408
Coagulation factor VII	F7	8 241 288
Coagulation factor IX	F9	8 211 270
Vitamin K-dependent protein C	PROC	8 150 400
Coagulation factor XI	F11	8 070 480
Plasminogen	PLG	7 525 872
Coagulation factor VII/tissue factor	F3	7 468 008
Serine/threonine-protein kinase AKT	AKT1	6 091 730
Heat shock protein HSP 90- α	HSP90AA1	5 944 926
Histone deacetylase 1	HDAC1	5 832 968
Serine/threonine-protein kinase AKT2	AKT2	5 657 640
Serine/threonine-protein kinase AKT	AKT3	5 645 520
Inhibitor of apoptosis protein 3	XIAP	5 441 424
Serine/threonine-protein kinase Chk1	CHEK1	5 268 960
ALK tyrosine kinase receptor	ALK	5 209 560
Baculoviral IAP repeat-containing protein 2	BIRC2	4 603 200

示。条形图的高度代表类别中注释基因重叠的基因数量。在生物过程中生物调节、应激反应、代谢过程等相关性较高,其中包括一氧化氮的生物合成过程、调节突触传递、抑制酶原、蛋白质水解的负调控、防御响应、抑制半胱氨酸型内肽酶活性、单加氧酶活性的调节、细胞对铵离子的反应、调节蛋白激酶活性、蛋白质定位到细胞核的调控、免疫系统过程、内皮细胞增殖的调控、调节凋亡过程中的半胱氨酸型内肽酶活性、水解酶活性的正调控、通过细胞调节分泌、巨噬细胞激活、纤维蛋白溶解的负调控、细胞对有机氮化合物的反应等;在细胞组分中细胞膜、细胞外间隙、囊泡所占比例大;在分子功能中蛋白结合、离子结合、水解酶活性影响相对靠前。

KEGG 通路富集筛选得到 26 条信号通路($P < 0.05$),见图 4,按伪发现率(FDR)值降序排列前 20 名,分别为胰岛素/胰岛素样生长因子(IGF)途径-蛋白激酶 B 信号级联反应、5-羟色胺(5-HT)₄型受体介导的信号通路、干扰素- γ 信号传导途径、转化生长因子- β (TGF- β)信号通路、白介素信号通路、内皮素信号传导途径、异三聚体 G 蛋白信号传导途径-Gi α 和 Gs α 介导的途径、B 细胞活化、趋化因子和细胞因子信号传导途径介导的炎症、亨廷顿病、

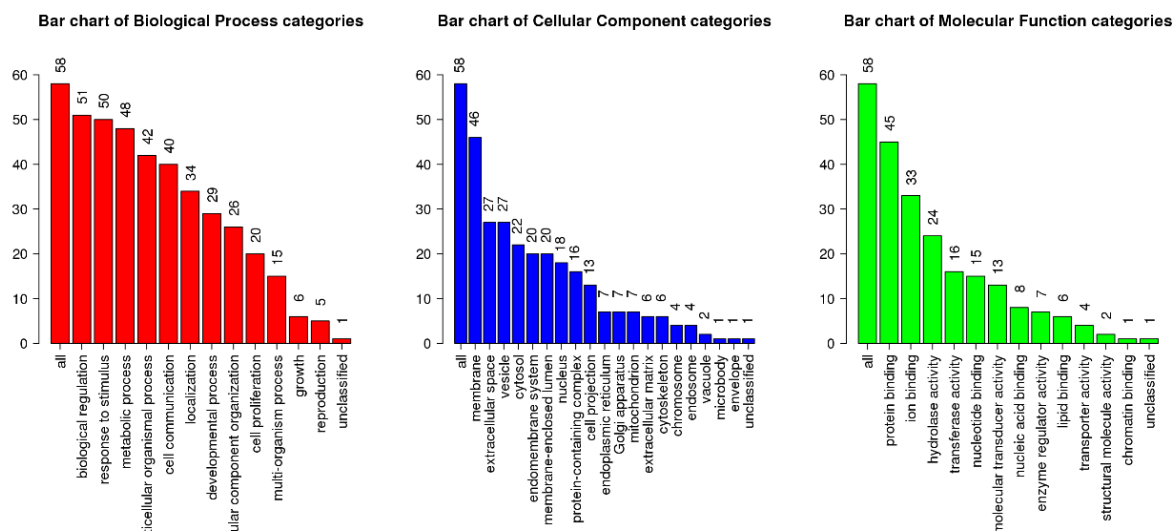


图3 化合物目标靶点GO富集分析

Fig. 3 GO enrichment analysis of targets of compounds

p53 途径反馈回路2、CCKR 信号通路、通过缺氧诱导因子(HIF)激活引起的缺氧反应、p53 途径、Ras 通路、表皮生长因子(EGF)受体信号传导途径、葡萄糖剥夺引起的 p53 途径、成纤维细胞生长因子(FGF)信号通路、纤溶酶原激活级联、磷脂酰肌醇激酶(PI3K)途径。

2.5 TMPRSS2 三级结构及基本理化性质

通过 SWISS-MODEL 采用同源建模法,目标序列与模板序列(SMTL ID: 5ce1.1)一致度为 33.82%(>30%),符合同源建模法条件^[12],将预测到的蛋白质三级结构根据 GMQE 和 QMEAN 选择合适的模型^[13],GMQE 表示模型的准确程度,取值范围在 0 到 1 之间,而 QMEAN 表示模型的质量^[13]。对

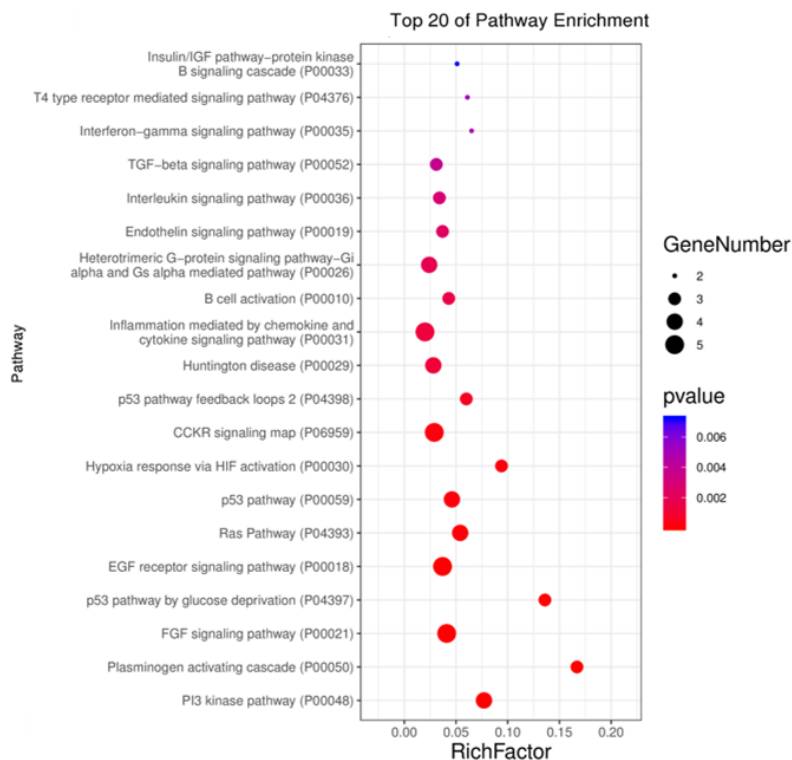


图4 化合物作用靶点的KEGG富集分析的前20条通路

Fig. 4 Top20 pathways of target of compounds by KEGG enrichment analysis

其选择的标准是GMQE越高越好,而QMEAN最好大于-4。选择GMQE为0.53,QMEAN为-1.62的蛋白三级结构,见图5。

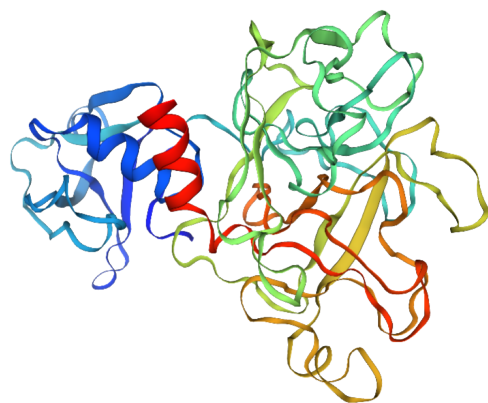


图5 氨基酸序列预测的TMPRSS2蛋白三级结构

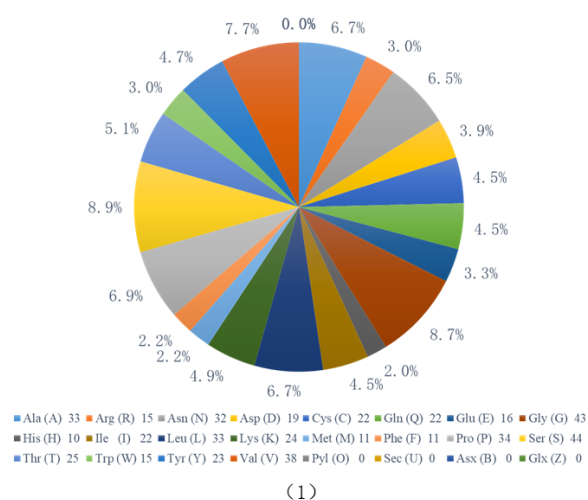
Fig. 5 Tertiary structure of TMPRSS2 protein predicted by amino acid sequence

通过Protparam在线工具进行分析,TMPRSS2的总原子数为7 433个,氨基酸数目为492个,相对分子质量为53 859.18,理论等电点为8.12,负电荷

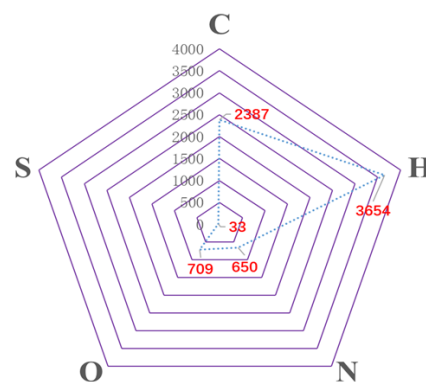
总残基数(Asp+Glu)为35,正电荷总残基数(Arg+Lys)为39。具体氨基酸和原子组成见图6,其中丝氨酸(Ser)占氨基酸总数的8.9%,含量最多;H原子有3 654个,为原子中最多。不稳定系数为41.94,脂肪系数为72.70。根据判别标准^[14-15],将不稳定系数大于40的蛋白判定为在体外较不稳定。脂肪系数是由丙氨酸、缬氨酸、亮氨酸、异亮氨酸这4个疏水氨基酸的出现频次于相关权重相乘并加和得到,也可以用来衡量目的蛋白的稳定性。脂肪系数越高,蛋白相对越稳定^[14-15]。在哺乳动物网织红细胞、酵母和大肠杆菌这3种模式细胞中的半衰期分别为30、20和10 h,总平均亲水性(GRAVY)为-0.248。数值越大,亲水性越强^[14-15]。

2.6 单体化合物与TMPRSS2的分子对接结果

一般认为能量越低,配体与受体结合的构象越稳定,发生的作用可能性越大。分子对接结果显示,甲磺酸卡莫司他、2种单体化合物与TMPRSS2蛋白的分子对接亲和力均小于零,由此这3种小分子与TMPRSS2可自发结合,对接结果见图7。且2种单体化合物和TMPRSS2的结合能与甲磺酸卡莫



(1)



分子式: $C_{2387}H_{3654}N_{650}O_{709}S_{33}$

(2)

图6 TMPRSS2蛋白的氨基酸和原子组成

Fig. 6 Amino acid composition and atomic composition of TMPRSS2 protein

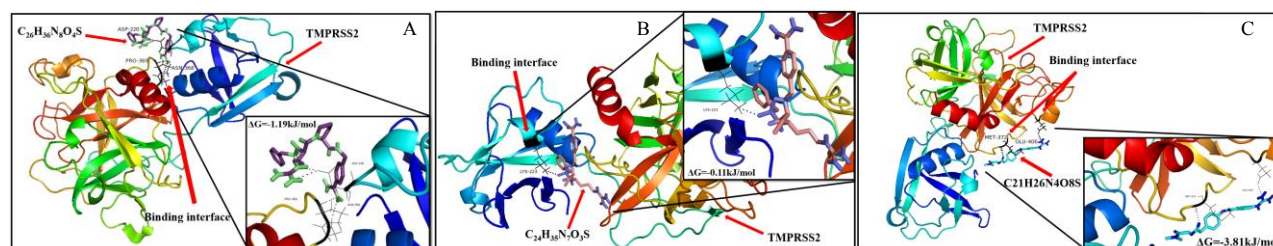


图7 TMPRSS2- $C_{26}H_{36}N_8O_4S$ (A)、 $C_{24}H_{35}N_7O_3S$ (B)、甲磺酸卡莫司他(C)对接图

Fig. 7 TMPRSS2- $C_{26}H_{36}N_8O_4S$ (A), $C_{24}H_{35}N_7O_3S$ (B), and Carmostat mesylate (C) docking diagram

司他相近,说明2种单体化合物对TMPRSS2有较好的抑制作用,结果见表2。 $C_{26}H_{36}N_8O_4S$ 与TMPRSS2相互作用的残基包括ASP220、PRO369、PHE194、PHE357、GLU376、GLY370、ARG240、ASN368、LYS350、LYS353、GLN352、MET478、GLU329、GLU289、GLU406、ASP175; $C_{24}H_{35}N_7O_3S$ 与TMPRSS2相互作用的残基包括LYS223、SER163、ASN218、THR407、GLU289、ASP359、GLN408、ASN336、PRO369、ASN418、ASP338、ALA466、GLN431、GLN289、LYS224、ASP220、PRO422、GLU406、LEU355、GLY370。

3 讨论

关于SARS-CoV-2的蛋白质结构生物学研究仍处于早期阶段,目前关于病毒的晶体结构只有3CL水解酶一种(PDB: 6LU7)。在SARS-CoV-2与ACE2的结合这一过程中,S蛋白进入人体细胞的启动剂TMPRSS2发挥着重要作用。早期研究表明,SARS-CoV的S蛋白可以利用体内半胱氨酸蛋白酶(CatB/L)和TMPRSS2进入靶细胞^[16-17];近日,德国科学家Markus Hoffmann等^[4]的研究表示,SARS-CoV-2的传播也依赖于TMPRSS2,他们应用了TMPRSS2抑制剂甲磺酸卡莫他司减轻了SARS-CoV-2的感染。TMPRSS2在多种病毒传播过程中至关重要,并且它在体内的表达可受调控^[18],所以TMPRSS2的抑制物,很可能具有抑制病毒入侵细胞的作用。Frank Sielaff等^[5]制备了一系列以4-胍基苄胺作为P1残基的丝氨酸蛋白酶(HAT)底物类似物抑制剂,其中最有效的化合物以D构型的碱性氨基酸作为P3残基的特征,这些化合物在表达HAT的细胞模型中抑制了流感病毒的复制。本课题组认为TMPRSS2抑制剂可以考虑应用于COVID-19患者的靶向治疗。

靶点分析结果可知,单体化合物作用的靶点有F2、PLG、PLAU、PLAT、HSP90AA1、XIAP、AKT1、AKT2、AKT3等,主要与激活凝血系统、抗凝、抗刺

激、抑制细胞凋亡、抗炎等相关^[19-23]。AKT1、AKT2和AKT3是AKT激酶的3种紧密相关的丝氨酸/苏氨酸蛋白激酶,能够调节许多过程^[24],而AKT还是细胞内外信号传导的重要节点,在PI3K、PDK1、ILK的作用下发生磷酸化,并与相应的下游底物蛋白发生作用,调控着细胞的存活、增殖及凋亡^[25]。细胞因子风暴在肺炎的进展过程中发挥重要作用,能导致非上皮细胞和肺血管内皮细胞的损伤引起急性肺损伤,从而引起急性呼吸窘迫症^[26-27]。有实验揭示了血管内皮细胞是细胞因子风暴的重要参与者,也是中心调控者^[28]。还有研究发现,细胞凋亡是引起炎症反应,进而发展成为炎症因子风暴的重要因素^[29]。 γ -干扰素(IFN- γ)又称II型干扰素,主要由活化的T细胞和NK细胞产生,其主要作用是能诱导巨噬细胞的激活,促使多种细胞MHC因子的表达增强,进一步促进Th1细胞分化、内皮细胞黏附分子及白细胞的表达,具有干扰病毒复制的活性,参与机体炎症反应^[30]。结合靶点分析、GO功能分析和KEGG富集通路分析结果,单体化合物可能是通过作用于AKT调控胰岛素/IGF途径-蛋白激酶B信号级联反应增加内皮细胞的增殖、抑制细胞凋亡从而对抗或抑制细胞因子风暴,减少肺组织损伤,干预病毒感染后的细胞信号级联反应和干扰素- γ 信号通路抑制病毒的复制。

本研究对TMPRSS2的三级结构进行了预测,并对其蛋白质特征进行了分析,结果显示,TMPRSS2相对较稳定。将预测的TMPRSS2和2种单体化合物及已上市抑制剂甲磺酸卡莫他司进行对接,通过分子对接结果可以看出,TMPRSS2与两种单体化合物都能自发结合,且结合能与TMPRSS2和甲磺酸卡莫他司的结合能相近,说明两种单体化合物对TMPRSS2有较好的抑制作用。 $C_{26}H_{36}N_8O_4S$ 与TMPRSS2的之间的相互作用,以疏水作用为主, $C_{24}H_{35}N_7O_3S$ 与TMPRSS2的作用以亲水作用为主,二者共同作用的残基有PRO369、

表2 3种小分子与TMPRSS2的结合能

Table 2 Binding energies of three small molecules to TMPRSS2

受体	配体	化学式	相对分子质量	结合能/(kJ·mol ⁻¹)
TMPRSS2	(2S)-1-[(2R)-2-(苄基磺酰基氨基)-5-(二氨基亚甲基氨基)戊酰基]-N-[(4-氨基甲酰基苯基)甲基]吡咯烷-2-羧酰胺	$C_{26}H_{36}N_8O_4S$	556.7	-1.19
TMPRSS2	4-[[[(5R)-5-(苄基磺酰基氨基)-8-(二氨基甲基亚氨基)氨基]-4-氧代辛烷-3-基]氨基]甲基]苯羧酰亚胺	$C_{24}H_{35}N_7O_3S$	501.6	-0.11
TMPRSS2	甲磺酸卡莫他司	$C_{21}H_{26}N_4O_8S$	494.5	-3.81

GLY370、GLU406。此外,与 $C_{26}H_{36}N_8O_4S$ 作用的疏水残基包括 PRO369、PHE194、PHE357、MET478 等。疏水性氨基酸在蛋白质内部,由于其疏水的相互作用,而在酶和基质等各种非共价键的分子结合、保持蛋白质的三级结构、维持生物膜的结构方面起作用^[31]。抗体的与抗原结合的部位上也有很多疏水性氨基酸,它参与与半抗原的结合^[31]。 $C_{24}H_{35}N_7O_3S$ 主要与 TMPRSS2 上的 SER163、ASN218、THR407、GLN408、ASN336、ASN418、GLN431、GLN289、GLY370 等天冬酰胺、谷氨酰胺等残基之间具有亲水作用。在构成受体或酶的蛋白质和多肽结构中含有大量的酰胺键,因此酰胺类药物易与生物大分子形成氢键,增强与受体的结合能力^[32]。此外,还有研究表明,这2种化合物都具有一定的抗凝血作用^[33-35]。故推测3种化合物可以通过竞争性拮抗作用、抗凝血作用对新型冠状病毒肺炎发挥较好的治疗作用。

本研究对治疗 COVID-19 的潜在化合物进行了探索性研究,对 TMPRSS2 进行了三级结构和靶点的预测,发现了单体化合物 $C_{26}H_{36}N_8O_4S$ 、 $C_{24}H_{35}N_7O_3S$ 、甲磺酸卡莫司他可能是通过 F2、PLG、PLAU、PLAT、HSP90AA1、XIAP、AKT1、AKT2、AKT3 等靶点和细胞信号级联反应、 γ -IFN 信号通路等抑制 TMPRSS2 发挥治疗 COVID-19 的作用。鉴于网络药理学及分子对接技术的局限性,本课题组后期将在此基础上,对化合物作用的主要靶点、通路进行实验验证,为治疗 COVID-19 提供坚实的理论依据。

参考文献

- [1] Du Z W, Wang L, Cauchemez S, et al. Risk for transportation of 2019 novel coronavirus disease from Wuhan to other cities in China [J]. *Emerging Infect Dis*, 2020, 26(5). DOI:10.3201/eid2605.200146.
- [2] Lai C C, Shih T P, Ko W C, et al. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): the epidemic and the challenges [J]. *Int J Antimicrob Agents*, 2020, 55(3): 105924.
- [3] Yan R, Zhang Y, Guo Y, et al. Structural basis for the recognition of the 2019-nCoV by human ACE2 [J]. *bioRxiv*, 2020. doi: <https://doi.org/10.1101/2020.02.19.956946>
- [4] Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor [J]. *Cell*, 2020: S0092-S8674. Doi:10.1016/j.cell.2020.02.052
- [5] Sielaff F, Böttcher-Friebertshäuser E, Meyer D, et al. Development of substrate analogue inhibitors for the human airway trypsin-like protease HAT [J]. *Bioorg Med Chem Lett*, 2011, 21(16): 4860-4864.
- [6] Szklarczyk D, Franceschini A, Wyder S, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life [J]. *Nucleic Acids Res*, 2015, 43(Database issue): D447-D452.
- [7] Smoot M E, Ono K, Ruscheinski J, et al. Cytoscape 2.8: new features for data integration and network visualization [J]. *Bioinformatics*, 2011, 27(3): 431-432.
- [8] Waterhouse A, Bertoni M, Bienert S, et al. SWISS-MODEL: homology modelling of protein structures and complexes [J]. *Nucleic Acids Res*, 2018, 46(W1): W296-W303.
- [9] Guex N, Peitsch M C, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective [J]. *Electrophoresis*, 2009, 30(Suppl 1): S162-S173.
- [10] Assenov Y, Ramírez F, Schelhorn S E, et al. Computing topological parameters of biological networks [J]. *Bioinformatics*, 2008, 24(2): 282-284.
- [11] Chin C H, Chen S H, Wu H H, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome [J]. *BMC Syst Biol*, 2014, 8(Suppl 4): S11.
- [12] Biasini M, Bienert S, Waterhouse A, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information [J]. *Nucleic Acids Res*, 2014, 42(Web Server issue): W252-W258.
- [13] Guex N, Peitsch M C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling [J]. *Electrophoresis*, 1997, 18(15): 2714-2723.
- [14] Garg V K, Avashthi H, Tiwari A. MFPP-FASTA ProtParam Interface [J]. *bioinformatics*, 2016, 12(2): 74-77. Doi:10.6026/97320630012074.
- [15] Chen T G, Wu S F, Wan P, et al. Gene prediction and function research of SARS-CoV(BJ01) [J]. *Acta Genet Sin*, 2003, 30(8): 773-780.
- [16] Simmons G, Gosalia D N, Rennekamp A J, et al. Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry [J]. *Proc Natl Acad Sci USA*, 2005, 102(33): 11876-11881.
- [17] Kawase M, Shirato K, van der Hoek L, et al. Simultaneous treatment of human bronchial epithelial cells with serine and cysteine protease inhibitors prevents severe acute respiratory syndrome coronavirus entry [J]. *J Virol*, 2012, 86(12): 6537-6545.
- [18] Iwata-Yoshikawa N, Okamura T, Shimizu Y, et al.

- TMPRSS2 contributes to virus spread and immunopathology in the airways of murine models after coronavirus infection [J]. J Virol, 2019, 93(6): e01815-e01818.
- [19] Lin F, Weng M S, Wu J R, et al. Gene variants in four pedigrees with hereditary coagulation factor XI deficiency and one novel mutation identification [J]. Blood Coagul Fibrinolysis, 2020, 31(2): 160-164.
- [20] L Venkatraman K, A Syed A, Indumathi P, et al. VITPOR AI, A coagulation factor XIIa inhibitor from *Porphyra yezoensis*: *in vivo* mode of action and assessment of platelet function analysis [J]. Protein Pept Lett, 2020, 27(3): 243-250.
- [21] Biswas N, Maihofer A X, Mir S A, et al. Polymorphisms at the F12 and KLKB1 loci have significant trait association with activation of the renin-angiotensin system [J]. BMC Med Genet, 2016, 17: 21.
- [22] Khosravi M, Faezeh N G, Abdollah A, et al. miR-193b-3p inhibits PLA2 gene expression in patients with in-stent restenosis [J]. Meta Gene, 2019, 22: 100602.
- [23] Yang H J, Wang L, Wang M, et al. Serine/threonine-protein kinase PFTK1 modulates oligodendrocyte differentiation via PI₃K/AKT pathway [J]. J Mol Neurosci, 2015, 55(4): 977-984.
- [24] Vivanco I, Sawyers C L. The phosphatidylinositol 3-Kinase AKT pathway in human cancer [J]. Nat Rev Cancer, 2002, 2(7): 489-501.
- [25] Chang F, Lee J T, Navolanic P M, et al. Involvement of PI₃K/Akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy [J]. Leukemia, 2003, 17(3): 590-603.
- [26] Huang C L, Wang Y M, Li X W, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China [J]. Lancet, 2020, 395(10223): 497-506.
- [27] Ma Q H, Huang W B, Zhao J, et al. Liu Shen Wan inhibits influenza A virus and excessive virus-induced inflammatory response via suppression of TLR4/NF- κ B signaling pathway *in vitro* and *in vivo* [J]. J Ethnopharmacol, 2020, 252: 112584.
- [28] Chousterman B G, Swirski F K, Weber G F. Cytokine storm and *Sepsis* disease pathogenesis [J]. Semin Immunopathol, 2017, 39(5): 517-528.
- [29] Liu Y Y, Fang Y L, Chen X F, et al. Gasdermin E-mediated target cell pyroptosis by *CAR* T cells triggers cytokine release syndrome [J]. Sci Immunol, 2020, 5(43): 7969.
- [30] Odeh A N, Simecka J W. Regulatory CD4⁺CD25⁺ T cells dampen inflammatory disease in murine *Mycoplasma pneumoniae* and promote IL-17 and IFN- γ responses [J]. PLoS One, 2016, 11(5): e0155648. Doi: 10.1371/journal.pone.0155648.
- [31] Almlöf T, Gustafsson J A, Wright A P. Role of hydrophobic amino acid clusters in the transactivation activity of the human glucocorticoid receptor [J]. Mol Cell Biol, 1997, 17(2): 934-945.
- [32] Pitman I H. Pro-drugs of amides, imides, and amines [J]. Med Res Rev, 1981, 1(2): 189-214.
- [33] Schweinitz A, Steinmetzer T, Banke I J, et al. Design of novel and selective inhibitors of urokinase-type plasminogen activator with improved pharmacokinetic properties for use as antimetastatic agents [J]. J Biol Chem, 2004, 279(32): 33613-33622.
- [34] Bajusz S, Szell E, Bagdy D, et al. Highly active and selective anticoagulants: D-Phe-Pro-Arg-H, a free tripeptide aldehyde prone to spontaneous inactivation, and its stable N-methyl derivative, D-MePhe-Pro-Arg-H [J]. J Med Chem, 1990, 33(6): 1729-1735.
- [35] Tucker T J, Lumma W C, Mulichak A M, et al. Design of highly potent noncovalent thrombin inhibitors that utilize a novel lipophilic binding pocket in the thrombin active site [J]. J Med Chem, 1997, 40(6): 830-832.