新型冠状病毒结合穿山甲ACE2受体的分子机制

仵丽丽^{1,2}, 苏佳岐³, 牛胜^{3,4}, 陈茜^{3,5}, 张艳芳³, 严景华¹, 施一³, 齐建勋³, 高福^{3*}, 王奇慧^{3,5,6*}

1. 中国科学院微生物研究所微生物生理与代谢工程重点实验室, 北京 100101;

2. 中国科学院大学, 北京 100049;

3. 中国科学院微生物研究所病原微生物与免疫学重点实验室, 北京 100101;

4. 山西农业大学动物科技学院, 太谷 030801;

5. 安徽大学物理科学与信息学院, 合肥 230039;

6. 中国科学院大学存济医学院, 北京 100049

* 联系人, E-mail: gaof@im.ac.cn; wangqihui@im.ac.cn

2020-10-23 收稿, 2020-11-05 修回, 2020-11-05 接受, 2020-11-06 网络版发表

中国科学院战略性先导科技专项(XDB29010202)、国家自然科学基金(81922044)和中国科学院“青年创新促进会”项目(2018119)资助

摘要 新型冠状病毒病(coronavirus disease 2019, COVID-19)的病原体为新型冠状病毒(COVID-19 virus), 以下简称为新冠病毒, 也称为severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). 研究认为, SARS-CoV-2可能起源于蝙蝠, 但是其中间宿主仍不清楚. 目前, 除蝙蝠外, 穿山甲被证明是唯一携带SARS-CoV-2相关冠状病毒的哺乳动物, 因而被认为是可能的中间宿主. 本研究利用表面等离子共振实验(surface plasmon resonance, SPR)展现和分析了SARS-CoV-2和两种穿山甲冠状病毒(pangolin-CoVs, GX/P2V/2017和GD/1/2019)分别与穿山甲ACE2(pangolin ACE2, pACE2)和人ACE2(human ACE2, hACE2)受体的结合能力, 解析了SARS-CoV-2刺突蛋白(spike, S)受体结合域(receptor binding domain, RBD)与pACE2复合物的晶体结构, 分辨率为2.3 Å, 发现SARS-CoV-2 RBD结合pACE2的模式与结合hACE2或猫ACE2(cat ACE2, cACE2)相似, 但与hACE2更相近. 通过同源建模的方式, 进一步模拟了GX/P2V/2017 RBD和GD/1/2019 RBD分别与pACE2和hACE2受体的结合, 发现二者结合pACE2和hACE2的方式与SARS-CoV-2 RBD相似, 但晶体结构或同源建模的数据暗示, 上述3种RBDs结合pACE2的能力均稍弱于hACE2, 这与SPR的结论一致. 这些研究结果不仅有助于我们更好地理解SARS-CoV-2的进化, 而且警示我们穿山甲冠状病毒具有感染人的潜力, 强调了持续监测穿山甲携带病毒的重要性, 以防止病毒外溢, 引起人类疾病.

关键词 新型冠状病毒, 穿山甲冠状病毒, 刺突蛋白受体结合域, 血管紧张素转换酶(ACE2), 晶体结构

新发、再发传染病是危害全球公共健康的重要原因之一^[1]. 本次新型冠状病毒病(coronavirus disease 2019, COVID-19)(以下简称新冠病毒病)引起的肺炎疫情是继严重急性呼吸综合征(severe acute respiratory syndrome, SARS)和中东呼吸综合征(Middle East respiratory syndrome, MERS)之后的第3个由冠状病毒引发的人类疫情^[2~4]. 截至2020年11月5日, 全球确诊病例已超过4700万, 死亡人数超过120万([https://www.who.](https://www.who.int/)

[int/](https://www.who.int/)). 目前没有获批的疫苗或药物用于预防和治疗新冠病毒病.

新型冠状病毒(COVID-19 virus, also called severe acute respiratory syndrome coronavirus 2, SARS-CoV-2)是引起COVID-19的病原体, 属于β冠状病毒属, 是已知可以感染人的第7种冠状病毒^[4~6]. 我们及其他研究团队的前期研究结果表明, 同SARS-CoV一样, SARS-CoV-2也是利用血管紧张素转换酶2(angiotensin-con-

引用格式: 仵丽丽, 苏佳岐, 牛胜, 等. 新型冠状病毒结合穿山甲ACE2受体的分子机制. 科学通报, 2021, 66: 73–84

Wu L L, Su J Q, Niu S, et al. Molecular basis of pangolin ACE2 engaged by SARS-CoV-2 (in Chinese). Chin Sci Bull, 2021, 66: 73–84, doi: [10.1360/TB-2020-1372](https://doi.org/10.1360/TB-2020-1372)

verting enzyme 2, ACE2)作为入侵受体,并揭示了SARS-CoV-2刺突蛋白(spike, S)受体结合域(receptor binding domain, RBD)与人ACE2(human ACE2, hACE2)相互作用的分子机制^[5,7]。目前,研究结果提示,SARS-CoV起源于蝙蝠,可能通过一些动物(果子狸或貉)间接传播给人^[8],也可能从蝙蝠直接传播给人^[9]。但是关于SARS-CoV-2的起源和跨种传播仍不清楚。系统进化分析结果表明,SARS-CoV-2与两种蝙蝠冠状病毒RaTG13和RmYN02的序列具有高度的同源性,分别为96.2%和93.3%,因而推测SARS-CoV-2也可能起源于蝙蝠冠状病毒,或者由蝙蝠冠状病毒与其他冠状病毒重组产生^[5,10]。

近期的研究报道,SARS-CoV-2除了可以感染人之外,还可以感染猫、狗、老虎和水貂等动物^[11-14],且我们及多个研究团队通过功能实验发现,SARS-CoV-2 RBD可以广泛结合多个动物的ACE2,包括家畜、宠物、野生动物等^[15,16]。迄今为止,除蝙蝠外,穿山甲被证明是唯一携带SARS-CoV-2相关冠状病毒的哺乳动物^[17]。研究者在2017和2019年分别从中国广西和广东海关缉私行动获取的马来穿山甲样本中分离出冠状病毒,它们属于两个独特的分支,分别命名为广西穿山甲冠状病毒(Guangxi pangolin coronaviruses, GX pangolin-CoVs)和广东穿山甲冠状病毒(Guangdong pangolin coronaviruses, GD pangolin-CoVs)。GD pangolin-CoVs与SARS-CoV-2的基因组序列同源性为92.4%,高于GX pangolin-CoVs(85.5%),尤其在S蛋白的RBD区域,氨基酸序列同源性高达97.4%,高于RaTG13 (89.2%)^[17-20]。研究还发现,SARS-CoV-2与穿山甲冠状病毒、RaTG13等其他蝙蝠冠状病毒在基因组的不同区域内亲缘关系不同,因而提出SARS-CoV-2可能来源于相关冠状病毒的重组^[17,19]。尽管如此,我们对SARS-CoV-2和pangolin-CoVs之间的进化关系仍然知之甚少,并且SARS-CoV-2 RBD结合pACE2,以及pangolin-CoVs结合穿山甲受体或人受体的分子机制仍不清楚。

本研究分析了SARS-CoV-2、1种广西穿山甲冠状病毒(GX/P2V/2017)和1种广东穿山甲冠状病毒(GD/1/2019)分别与穿山甲ACE2(pangolin ACE2, pACE2)和hACE2受体结合的特征,解析了SARS-CoV-2 RBD与pACE2复合物的晶体结构,发现SARS-CoV-2 RBD结合pACE2的模式与结合hACE2或猫ACE2(cat ACE2, cACE2)相似,但与hACE2更相近。我们进一步利用同源建模的方式模拟了GX/P2V/2017 RBD和GD/1/2019

RBD分别与pACE2和hACE2受体的结合,发现二者结合pACE2和hACE2的方式与SARS-CoV-2 RBD相似,但上述3种RBDs结合pACE2的能力均稍弱于hACE2。这些研究结果有助于我们更好地理解SARS-CoV-2的进化和可能来源,也提醒我们穿山甲冠状病毒具有传播给人的潜能,因而我们应该加强对穿山甲携带病毒的持续监测,以防止病毒溢出,引发新的疫情。

1 材料与方法

(i) 克隆构建。基本按照我们前期工作的构建系统完成^[15]。将编码全长pACE2(XP_017505746)的氨基酸序列按哺乳动物密码子偏好性优化合成后克隆至pEGFP-N1载体中,用于流式细胞染色实验;将pACE2胞外段(第1~740位氨基酸)的编码序列与编码小鼠IgG Fc(mFc)的序列拼接在一起克隆至pCAGGS载体中,用于表面等离子共振实验;将pACE2部分胞外段(第18~615位氨基酸)的氨基酸按昆虫细胞密码子偏好性优化合成后克隆至pFastBac1载体中,N端加上GP67信号肽,C端带有6×His标签。

将编码GX/P2V/2017 RBD(GISAID: EPI_ISL_410542,第319~541位氨基酸)和GD/1/2019 RBD(GISAID: EPI_ISL_410721,第319~541位氨基酸)的氨基酸按昆虫细胞密码子偏好性优化合成后分别克隆至pFastBac1载体中,同样N端加上GP67信号肽,C端带有6×His标签。我们前期构建了含有SARS-CoV-2 RBD(GISAID: EPI_ISL_402119,第319~541位氨基酸)编码基因的pFastBac1质粒,以及含有人ACE2(BAJ21180)编码基因的pEGFP-N1质粒^[7]。

根据说明书使用Mut Express[®] II Fast Mutagenesis Kit V2(南京诺维赞公司)进行氨基酸定点突变。

(ii) 蛋白的表达及纯化。为制备含有mFc标签的pACE2和hACE2蛋白,将含有这两种ACE2编码序列的pCAGGS质粒分别瞬时转染至HEK293F细胞中。4~6 d后,收集上清液,离心,经0.22 μm滤膜过滤后,上样至HiTrap rProtein A FF(GE Healthcare)亲和层析柱。用0.1 mol/L glycine-HCl(pH 3.0)洗脱结合的蛋白,洗脱的蛋白立即用1 mol/L Tris-HCl(pH 9.0)中和。蛋白经分子筛层析进一步纯化后浓缩,-80℃保存。

与之前一样^[7],通过bac-to-bac杆状病毒表达系统(Invitrogen)表达带有His标签的pACE2、SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD蛋白。简单地说,将含有目的基因的pFastbac1质粒转化至

DH10Bac感受态细胞中,生成重组杆粒(bacmids)。将bacmids转染至Sf9细胞中进行病毒的扩增,利用Hi5细胞中表达蛋白。表达48 h后,收集Hi5细胞上清液,使用HisTrap HP 5 mL (GE Healthcare)通过镍亲和层析纯化可溶性蛋白。然后用HiLoad 16/600 Superdex™ 200 pg (GE Healthcare)在20 mmol/L Tris-HCl (pH 8.0)和150 mmol/L NaCl缓冲液中凝胶过滤纯化目的蛋白。

(iii) 表面等离子共振实验。在25°C条件下,利用BIAcore 8K (GE Healthcare)的单循环模式测定pACE2和hACE2分别与SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD的亲和力,使用的缓冲液为20 mmol/L HEPES (pH 7.4)、150 mmol/L NaCl和0.005% (v/v) Tween 20,在测定亲和力之前,先将RBD蛋白置换至该缓冲液中。首先,使用氨基偶联法将识别小鼠Fc的抗体(anti-mFc antibody)固定至CM5芯片表面;然后,分别捕获含有mFc标签的pACE2或hACE2蛋白(响应值约500RU);接着,将梯度稀释的RBD蛋白流经芯片,稀释比例按SARS-CoV-2 RBD(25、50、100、200和400 nmol/L), GX/P2V/2017 RBD或GD/1/2019 RBD(12.5、25、50、100和200 nmol/L),分别测定与pACE2的亲和力;稀释比例按SARS-CoV-2 RBD(12.5、25、50、100和200 nmol/L), GX/P2V/2017 RBD或GD/1/2019 RBD(12.5、25、50、100和200 nmol/L),分别测定与hACE2的亲和力。用10 mmol/L Glycine-HCl (pH 1.7)再生芯片,使用BIAcore® 8K分析软件(GE Healthcare)按1:1的Langmuir结合模型计算平衡解离常数(K_D)。

(iv) 流式细胞术。为了检测pACE2的突变体与SARS-CoV-2 RBD的结合情况,将pEGFP-N1-pACE2(WT)、Q42A、K353A或D355A质粒瞬时转染至HEK293T细胞中,24 h后,收集 2×10^5 个细胞并用PBS重悬,与终浓度为1 $\mu\text{g/mL}$ 的SARS-CoV-2 RBD蛋白在37°C孵育30 min;随后用PBS洗涤两次,用anti-His/APC抗体(1:500稀释, Miltenyi Biotec)染色,置37°C再次孵育30 min。洗涤后,用BD FACSCanto分析细胞。

(v) 蛋白结晶、数据收集及结构解析。将高纯度的pACE2和SARS-CoV-2 RBD蛋白按摩尔比1:2混合后,置冰上孵育1 h,随后经分子筛HiLoad 16/600 Superdex™ 200 pg (GE Healthcare)分离出pACE2-SARS-CoV-2 RBD复合物蛋白,缓冲液为20 mmol/L Tris (pH 8.0)和50 mmol/L NaCl,浓缩至 $\sim 12 \text{ mg/mL}$,用于晶体筛选。使用坐滴气相扩散法进行晶体筛选,将0.8 μL 蛋

白与0.8 μL 晶体筛选试剂混合,并置于18°C进行晶体生长。在蛋白浓度为12.7 mg/mL时,获得高质量的晶体,晶体的生长条件为0.05 mol/L磷酸钾和20% (w/v)聚乙二醇8000。

为收集晶体衍射数据,将晶体快速冻至含有20% (v/v)甘油的防冻液中,并转移至液氮罐中。在上海同步辐射光源BL17U线站收集pACE2-SARS-CoV-2 RBD复合物晶体X射线衍射数据。使用HKL2000^[21]对收集的数据进行初步处理,以先前报道的hACE2-SARS-CoV-2 RBD复合物结构(PDB: 6LZG)为模型,利用Phaser^[22]分子置换法解析pACE2-SARS-CoV-2 RBD复合物结构,使用COOT^[23]建立原始模型,并用Phenix^[24]对模型进行修正,用MolProbity^[25]评估最终模型的立体化学质量。表1总结了晶体数据收集、处理和修正参数,使用PyMOL软件进行结构分析并作图(<https://www.pymol.org/>)。本研究获得的SARS-CoV-2 RBD与pACE2的复合物晶体结构已在蛋白数据库(Protein Data Bank, PDB)中注册登记,其注册号为7DHY,同时也在国家微生物科学数据中心(National Microbiology Data Center, NMDC)中注册登记,其注册号为NMDCS0000010。

(vi) 同源结构建模。从蛋白数据库中搜索与GX/P2V/2017 RBD或GD/1/2019 RBD序列高度相似的结构(48个),作为同源建模的模板。使用RosettaCM^[26]生成GX/P2V/2017 RBD和GD/1/2019 RBD结构模型;然后使用HDOCK^[27]进行整体对接,将GX/P2V/2017 RBD和GD/1/2019 RBD模型分别对接至pACE2和hACE2,并使用Rosetta(<https://www.rosettacommons.org/>)对模型进行修正。使用PyMOL软件进行结构分析并作图(<https://www.pymol.org/>)。

2 结果

2.1 SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD分别与pACE2和hACE2的结合特征

为研究SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD与pACE2和hACE2的相互作用,分别表达并纯化了SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD蛋白,利用SPR实验检测了这3种蛋白分别与pACE2和hACE2相互作用的亲和力。

SPR结果显示, SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD蛋白均可以与pACE2和hACE2

表 1 SARS-CoV-2 RBD与pACE2复合物晶体结构数据收集和修正参数统计表**Table 1** Data collection and refinement statistics of the crystal structure of SARS-CoV-2 RBD bound to pACE2

SARS-CoV-2 RBD-pACE2	
数据收集	
空间群	C121
晶胞参数	
a, b, c (Å)	139.996, 97.564, 90.982
α, β, γ (°)	90.00, 129.154, 90.00
分辨率(Å)	50.00~2.30 (2.38~2.30) ^{a)}
独立衍射点	41945
R_{merge} 因子	0.10 (0.85) ^{a)}
R_{pim} 因子	0.041 (0.365) ^{a)}
信噪比	51.40 (2.69) ^{a)}
完整度(%)	99.20 (98.30) ^{a)}
冗余度	6.80 (6.20) ^{a)}
修正	
分辨率(Å)	33.93~2.30
衍射点数	39596
$R_{\text{因子}}/R_{\text{自由因子}}$	0.19/0.23
原子个数	
蛋白	6385
配体	29
水	227
B 因子(Å ²)	
蛋白	37.53
配体	86.26
水	36.74
均方根偏差	
键长(Å)	0.003
键角(°)	0.563
拉氏图	
基本吻合(%)	98.08
允许(%)	1.92
不允许(%)	0.00

a) 括号中的数据表示最高分辨率范围

直接结合. SARS-CoV-2 RBD与pACE2结合的平衡解离常数 K_D 为34.97 nmol/L, 其亲和力与结合hACE2相比低约3.5倍(图1(a), (b)). 值得注意的是, GX/P2V/2017 RBD和GD/1/2019 RBD与pACE2的亲和力与SARS-CoV-2 RBD相似, K_D 值分别为59.5和66.5 nmol/L(图1(c), (e)),

且皆比hACE2低约4倍(图1(d), (f)). 以上结果表明, SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD结合pACE2的能力相似, 且亲和力均低于hACE2 (~4倍).

2.2 SARS-CoV-2 RBD与pACE2相互作用的分子基础

为进一步探究SARS-CoV-2 RBD与pACE2相互作用的机理, 我们解析了SARS-CoV-2 RBD与pACE2的复合物晶体结构(表1), 分辨率为2.3 Å, 发现在一个不对称单元中1个SARS-CoV-2 RBD分子结合1个pACE2分子. 正如预期的那样, SARS-CoV-2 RBD-pACE2复合物的整体结构与我们之前解析的SARS-CoV-2 RBD-hACE2复合物和SARS-CoV-2 RBD-cACE2复合物的结构非常相似(图2(a))^[7,15]. 在SARS-CoV-2 RBD-pACE2结构中, 可以清晰地观察到593个pACE2的氨基酸残基(S19~A614), 其中有3个连续的氨基酸(G136、N137和P138)的电子云密度缺失, N53和N90两个位点可观察到N-糖基化修饰.

进一步分析SARS-CoV-2 RBD与pACE2之间的相互作用细节, 我们列出了沿着二者的结合界面参与形成疏水相互作用或氢键的氨基酸(表2), 其中参与形成氢键或盐桥相互作用的氨基酸在图2(b)、(c)和(d)中展示出来. SARS-CoV-2 RBD的F486和Y489与pACE2的F28和Y83形成疏水界面(图2(b)). SARS-CoV-2 RBD与pACE2的极性相互作用分布于 $\alpha 1'/\beta 1'$ 环、 $\beta 1'/\beta 2'$ 环和 $\beta 2'/\eta 1'$ 环, 包括SARS-CoV-2 RBD的G476与pACE2的S19, N487与E24和Y83(图2(b)), G446、Y449和Q498与Q42, T500和N501与Y41(图2(c)), Q493与K31, G496和G502与K353(图2(d)), 总共形成11对潜在的氢键相互作用. 值得注意的是, D405是SARS-CoV-2 RBD的 $\eta 1$ 螺旋上唯一一个参与相互作用的氨基酸, 与pACE2的H354形成1个相互作用, 此外, SARS-CoV-2 RBD的K417与pACE2的E30也形成盐桥相互作用(图2(d), 表2).

由于hACE2的K353和cACE2的D355已被证明是影响结合SARS-CoV-2 RBD的关键氨基酸^[7,15], 因而我们也重点关注了pACE2的K353和D355. 除此之外, 还包括Q42, 其与SARS-CoV-2 RBD形成3对氢键相互作用, 是形成氢键相互作用最多的氨基酸(图S1). 将pACE2的Q42、K353和D355分别突变成丙氨酸(A)后表达于细胞膜上, 通过流式细胞术(fluorescence-activated cell sorting, FACS)检测它们与SARS-CoV-2

图 1 表面等离子共振(surface plasmon resonance, SPR)方法检测SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD分别与pACE2和hACE2的结合能力. (a), (c), (e) SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD分别与pACE2结合的动力学曲线及平衡解离常数. (b), (d), (f) SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD分别与hACE2结合的动力学曲线及平衡解离常数. 含有mFc标签的pACE2和hACE2蛋白分别被预先固定了识别小鼠mFc抗体的CM5芯片捕获, 随后一系列梯度稀释的SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD蛋白分别流经芯片表面, 响应值被记录下来. 原始虚线和拟合曲线分别用黑色虚线和彩色实线显示, 该实验进行3次独立重复, K_D 值表示3次实验的平均值 $\pm$ 标准差

Figure 1 Binding characterization between SARS-CoV-2 RBD, GX/P2V/2017 RBD or GD/1/2019 RBD and pACE2 or hACE2 by SPR. (a), (c), (e) Binding affinities of SARS-CoV-2 RBD, GX/P2V/2017 RBD or GD/1/2019 RBD to pACE2. (b), (d), (f) Binding affinities of SARS-CoV-2 RBD, GX/P2V/2017 RBD or GD/1/2019 RBD to hACE2. The mFc-tagged pACE2 and hACE2 proteins were captured by the CM5 chip, on which anti-mIgG Fc antibodies were immobilized, and sequentially tested the binding with serially diluted SARS-CoV-2 RBD, GX/P2V/2017 RBD or GD/1/2019 RBD proteins. The raw and fitted curves were displayed in black dotted and colored solid lines, respectively. The assays were independently performed three times. The K_D s shown were the mean $\pm$ SD of three independent experiments

RBD的结合情况. FACS结果显示, K353A或D355A单点突变(pACE2-K353A或pACE2-D355A)足以彻底破坏与SARS-CoV-2 RBD的结合, 而Q42A突变体(pACE2-Q42A)几乎不影响与SARS-CoV-2 RBD的结合(图2(e)).

2.3 SARS-CoV-2 RBD结合pACE2、hACE2和cACE2的比较

pACE2与hACE2和cACE2具有高度同源性, 分别为85.5%和87.9%, 这使得三者在结构上也具有显著的

图2 SARS-CoV-2 RBD与pACE2相互作用的分子基础, 以及pACE2突变体与SARS-CoV-2 RBD的结合特征. (a) SARS-CoV-2 RBD结合pACE2的整体结构, 其中SARS-CoV-2 RBD和pACE2分别用紫色和小麦色表示. (b), (c), (d) pACE2和SARS-CoV-2 RBD的相互作用细节. 参与相互作用的氨基酸被标记出来, 虚线表示氢键相互作用(距离3.3 Å以内), 灰色数字表示键长. (e) FACS检测pACE2突变体与SARS-CoV-2 RBD的结合. 将pEGFP-N1-pACE2(WT)质粒, 或含有Q42A、K353A或D355A突变的质粒转染至HEK293T细胞, 随后将含有His标签的SARS-CoV-2 RBD蛋白与细胞孵育, 再用anti-His/APC抗体检测蛋白与细胞的结合. 图中所示结果为3次独立实验中的一次结果

Figure 2 Molecular basis between SARS-CoV-2 RBD and pACE2 and binding characterization of pACE2 mutants to SARS-CoV-2 RBD. (a) The overall complex structure of SARS-CoV-2 RBD bound to pACE2. SARS-CoV-2 RBD and pACE2 were colored in purple and wheat, respectively. (b), (c), (d) The detailed interaction between pACE2 and SARS-CoV-2 RBD. The residues involved in the interaction were labeled. H-bonds were shown as dotted lines with a cutoff of 3.3 Å, the length of which were shown with gray numbers. (e) Binding of pACE2 mutants to SARS-CoV-2 RBD were tested by flow cytometry. HEK293T cells transfected with pEGFP-N1-pACE2 (WT), or the mutants containing Q42A, K353A or D355A were incubated with His-tagged SARS-CoV-2 RBD protein. Anti-His/APC antibody was used to detect the His-tagged protein binding to the cells. The results shown were one representative data of three independent experiments

相似性(图S2). 整体看来, 将SARS-CoV-2 RBD-pACE2复合物结构整体叠加至SARS-CoV-2 RBD-hACE2上时, 均方根偏差(root mean square deviation, RMSD)为0.673 Å(689个Cα), 而叠加至SARS-CoV-2 RBD-cACE2上时, RMSD为0.923 Å(645个Cα). 仅将SARS-CoV-2 RBD-pACE2复合物中的RBD分别与SARS-CoV-2 RBD-hACE2和SARS-CoV-2 RBD-cACE2中的RBD叠加时, RMSD分别为0.225和0.580 Å, 并且发现pACE2几乎与hACE2重叠, 但与cACE2却呈现出一定的夹角(图3(a)). 由于ACE2中负责结合RBD的主要结合位点位于N端的α1和α2螺旋上, 因此将3种复合物中3个ACE2的α1和α2螺旋叠加, 结果显示, 结合pACE2的RBD(紫色)几乎与结合hACE2的RBD(灰色)重叠, 且两者与结合cACE2的RBD(青色)呈现出约4.7°的夹角(前

期工作中已描述)(图3(b))^[15]. 此外, 我们还发现, pACE2的β4/β5环也几乎与hACE2重叠, 而相比之下, cACE2的β4/β5环却更长、更具有柔性, 并且环的尖端及α12螺旋在远离RBD的方向上移动了约1 Å的距离(图3(c)), 导致cACE2的α12螺旋上的N330失去了与RBD的相互作用. 残基K353和D355是ACE2与SARS-CoV-2 RBD相互作用的关键氨基酸, 位于ACE2的β4/β5环上. 将3种复合物结构中的3个RBDs叠加发现, cACE2往靠近RBD的方向移动, 且β4/β5环的尖端与pACE2和hACE2重叠, 但是相比于pACE2和hACE2, cACE2的α13/η5环却往靠近RBD的方向移动约2.3 Å的距离, 导致两个氨基酸(A386和R393, 紫色)参与了RBD的相互作用, 而pACE2和hACE2的α13/η5环与RBD没有相互作用(图3(d)).

进一步比较3个复合物的相互作用细节发现,

除D405和F490外, SARS-CoV-2 RBD与pACE2相互作用的其余氨基酸均与hACE2相同(表1). D405与pACE2的H354形成1个相互作用, 而与hACE2没有相互作用; F490与hACE2形成1个相互作用, 而与pACE2没有相互作用. 与结合cACE2相比, SARS-CoV-2 RBD与pACE2相互作用的氨基酸有5个不同, 分别为D405、Y473、A475、G476和S477 (表2), D405同样与cACE2也没有相互作用, 而与pACE2形成1个相互作用; Y473与cACE2形成1个相互作用, S477与cACE2形成氢键, 而二者与pACE2没有相互作用; A475和G476与pACE2分别形成偶极相互作用和氢键, 而与cACE2没有相互作用. 虽然SARS-CoV-2 RBD与pACE2、hACE2和cACE2相互作用的氨基酸略有不同, 但参与相互作用

的氨基酸总数是相近的, 分别为19、19和18. SARS-CoV-2 RBD与pACE2、hACE2和cACE2相互作用界面的面积分别为913.8、876.8和960.2 Å², 虽然其与hACE2相互作用界面的面积最小, 但是二者之间的范德华力和氢键相互作用略多于pACE2和cACE2(分别为124、141和143、11、13和11)(表2). 值得注意的是, 由于SARS-CoV-2 RBD的K417与pACE2的E30距离变远, 失去了氢键相互作用, 进而导致其与pACE2的E30相互作用弱于hACE2的D30和cACE2的E30(表2). 这些结果提示SARS-CoV-2 RBD与pACE2的作用力低于hACE2, 这与SPR结果相符. 通过比较SARS-CoV-2 RBD结合pACE2、hACE2和cACE2的足迹, 发现参与相互作用的大部分氨基酸在3种ACE2s中是保守的, 而pACE2和

表 2 SARS-CoV-2 RBD结合pACE2、hACE2和cACE2的结构比较^{a)}

Table 2 Comparison of SARS-CoV-2 RBD binding to pACE2, hACE2 and cACE2^{a)}

SARS-CoV-2 RBD	pACE2	hACE2	cACE2
D405(1/0/0)	H354(1)	—	—
K417(4/3,1/4,1)	E30(4)	D30(3,1)	E30(4,1)
G446(4,1/2,1/1)	Q42(4,1)	Q42(2,1)	Q42(1)
Y449(2,1/7,2/12,1)	E38(1), Q42(1,1)	D38(5,1), Q42(2,1)	E38(9,1), Q42(3)
Y453(1/3,1/5)	S34(1)	H34(3,1)	H34(5)
L455(1/4/5)	S34(1)	H34(4)	E30(1), H34(4)
F456(8/7/13)	T27(5), E30(2), K31(1)	T27(5), D30(1), K31(1)	T27(4), E30(4), K31(5)
Y473(0/0/1)	—	—	T27 (1)
A475(5/5,1/0)	S19(1), E23(2), E24(1), T27(1)	S19(3,1), Q24(1), T27(1)	—
G476(3,1/2/0)	S19(3,1)	S19(2)	—
S477(0/0/4,1)	—	—	Q18(4,1)
F486(8/11/5)	N82(2), Y83(6)	M82(4), Y83(7)	T82(5)
N487(7,2/11,2/7,1)	E24(3,1), Y83(4,1)	Q24(7,1), Y83(4,1)	L24(3), Y83(4,1)
Y489(8/7/10)	T27(2), F28(4), K31(1), Y83(1)	T27(2), F28(4), Y83(1)	T27(1), F28(6), K31(2), Y83 (1)
F490(0/1/0)	—	K31(1)	—
Q493(9,1/6/7)	K31(1,1), S34(7), E35(1)	H34(3), E35(3)	H34(7)
G496(9,1/6,1/9,1)	E38(7), K353(2,1)	D38(1), K353(5,1)	E38(3), K353(6,1)
Q498(10,1/11,2/16,2)	E38(2), Y41(4), Q42(2,1), K353(2)	Y41(5), Q42(5,2), L45(1)	E38(6,1), Y41(7), Q42(3, 1)
T500(18,1/18/18,3)	Y41(6,1), N330(1), K353(3), D355(5), R357(3)	Y41(6), N330(3), D355(6), R357(3)	Y41(3,1), D355(12,2), R357(3)
N501(5,1/10,1/6,1)	Y41(5,1)	Y41(5,1), K353(5)	Y41(1), K353(1), D355(4,1)
G502(5,1/8,1/5)	K353(2,1), H354(3)	K353(3,1), G354(5)	K353(1), D355(4)
Y505(16/19/15)	E37(1), K353(13), H354(2)	E37(3), K353(14), G354(2)	K353(9), G354 (2), A386(1), R393(3)
Total	124, 11	141, 13	143, 11

a) SARS-CoV-2 RBD氨基酸括号内的数字表示与pACE2(前者)、hACE2(中间)和cACE2(后者)形成的范德华力和氢键相互作用数量. ACE2氨基酸括号中的数字表示与SARS-CoV-2 RBD形成的范德华力和氢键相互作用数量. 逗号后的数字表示形成的氢键相互作用数量. 上述分析的范德华力距离为4 Å以内, 氢键距离为3.3 Å以内

图3 SARS-CoV-2 RBD-pACE2、SARS-CoV-2 RBD-hACE2和SARS-CoV-2 RBD-cACE2复合物结构的比较。(a) SARS-CoV-2 RBD-pACE2与SARS-CoV-2 RBD-hACE2和SARS-CoV-2 RBD-cACE2复合物结构的整体比较。SARS-CoV-2 RBD-pACE2复合物结构中，SARS-CoV-2 RBD和pACE2分别用紫色和小麦色表示。SARS-CoV-2 RBD-hACE2和SARS-CoV-2 RBD-cACE2复合物分别用灰色和青色表示。(b) SARS-CoV-2 RBD结合pACE2、hACE2和cACE2的取向比较。(c) 叠加pACE2、hACE2和cACE2 N端的前85个氨基酸($\alpha 1$ 和 $\alpha 2$ 螺旋)，比较3种ACE2s的第320~400位氨基酸($\alpha 12$ 、 $\alpha 13$ 和 $\eta 5$ 螺旋及 $\beta 4$ 和 $\beta 5$ 折叠片)。(d) 叠加3种复合物结构中的3个RBDs，比较pACE2、hACE2和cACE2的第320~400位氨基酸。(e)，(f)，(g) SARS-CoV-2 RBD结合pACE2 (e)、hACE2 (f)和cACE2 (g)的足迹，图中显示为相互作用距离4 Å以内的氨基酸

Figure 3 Structural comparison of SARS-CoV-2 RBD-pACE2 with SARS-CoV-2 RBD-hACE2 and SARS-CoV-2 RBD-cACE2. (a) The overall comparison of SARS-CoV-2 RBD-pACE2 complex to SARS-CoV-2 RBD-hACE2 complex or SARS-CoV-2 RBD-cACE2 complex. SARS-CoV-2 RBD binding to pACE2 was shown in purple and pACE2 was shown in wheat. SARS-CoV-2 RBD-hACE2 complex and SARS-CoV-2 RBD-cACE2 complex were colored in gray and cyan, respectively. (b) The orientation of SARS-CoV-2 RBD bound to pACE2, hACE2 or cACE2. (c) The comparison of residues 320–400 ($\alpha 12$, $\alpha 13$ and $\eta 5$ helices, and $\beta 4$ and $\beta 5$ sheets) of pACE2, hACE2 and cACE2 when the N-terminal 85 residues ($\alpha 1$ and $\alpha 2$ helices) of three ACE2s were aligned. (d) The comparison of residues 320–400 of pACE2, hACE2 and cACE2 when the SARS-CoV-2 RBDs were aligned. (e), (f), (g) The footprints of SARS-CoV-2 RBD on pACE2 (e), hACE2 (f) and cACE2 (g). Residues involved in the interaction are shown with a cutoff of 4 Å

hACE2仅有1个氨基酸不同(图3(e)，(f))，pACE2和cACE2有7个氨基酸的差异(图3(e)，(g))。这些结果表明，相比于SARS-CoV-2 RBD-cACE2复合物结构，SARS-CoV-2 RBD-pACE2更相似于SARS-CoV-2 RBD-hACE2。

2.4 GX/P2V/2017 RBD和GD/1/2019 RBD分别结合pACE2和hACE2的结构模型

GX/P2V/2017和GD/1/2019这两株穿山甲冠状病毒属于两个独特的分支^[17]，其RBD与SARS-CoV-2 RBD的同源性分别为87.1%和97.4%。SARS-CoV-2 RBD上与hACE2具有相互作用的氨基酸(19个，相互作用距离在4 Å以内)用彩色三角形指示(图4(g))。在上述19个氨基酸中，有13个氨基酸与GX/P2V/2017 RBD和GD/1/2019 RBD是相同的，且只有第417位氨基酸在3种

RBDs中是不同的；此外，GD/1/2019 RBD与SARS-CoV-2 RBD相比，只有两个位点(第417和498位)不同；而GX/P2V/2017 RBD与SARS-CoV-2 RBD相比，有5个位点(第417、486、490、493和501位)不同，并且这5个位点也不同于GD/1/2019 RBD(图4(g))。为探究GX/P2V/2017和GD/1/2019与pACE2和hACE2相互作用的分子基础，我们基于已知冠状病毒的RBD结构模拟了这两株穿山甲冠状病毒RBDs的结构，并且将这两个RBDs分别与pACE2和hACE2进行对接，最终获得4个模拟的复合物结构，分别为GX/P2V/2017 RBD-pACE2、GX/P2V/2017 RBD-hACE2、GD/1/2019 RBD-pACE2和GD/1/2019 RBD-hACE2(图4(a)，(b))。将GX/P2V/2017 RBD和GD/1/2019 RBD分别与SARS-CoV-2 RBD叠加对比，RMSD分别为0.462和0.366 Å，而GX/P2V/2017 RBD与GD/1/2019 RBD的RMSD为0.351

图 4 GX/P2V/2017 RBD和GD/1/2019 RBD分别结合pACE2和hACE2的结构模型, 以及与SARS-CoV-2 RBD-pACE2和SARS-CoV-2 RBD-hACE2复合物结构的比较。(a) GX/P2V/2017 RBD和GD/1/2019 RBD分别结合pACE2的结构模型, 并与SARS-CoV-2 RBD结合pACE2进行比较。GX/P2V/2017 RBD、GD/1/2019 RBD和SARS-CoV-2 RBD分别用浅绿色、浅粉色和紫色表示。(b) GX/P2V/2017 RBD和GD/1/2019 RBD分别结合hACE2的结构模型, 并与SARS-CoV-2 RBD结合hACE2进行比较。(c)、(d) GX/P2V/2017 RBD和GD/1/2019 RBD结合pACE2分别与SARS-CoV-2 RBD结合pACE2的细节比较。(e)、(f) GX/P2V/2017 RBD和GD/1/2019 RBD结合hACE2分别与SARS-CoV-2 RBD结合hACE2的细节比较。虚线表示距离3.3 Å以内的氢键相互作用, 键长用相应颜色的数字标出。(g) SARS-CoV-2 RBD、GX/P2V/2017 RBD和GD/1/2019 RBD序列比对。SARS-CoV-2 RBD与hACE2距离4 Å以内的、具有相互作用的氨基酸用彩色三角形表示。序列比对图由ESPrpt 3^[28]产生。

Figure 4 Structural model of GX/P2V/2017 RBD or GD/1/2019 RBD in complex with pACE2 or hACE2, and compared with SARS-CoV-2 RBD-pACE2 and SARS-CoV-2 RBD-hACE2. (a) Model of GX/P2V/2017 RBD or GD/1/2019 RBD bound to pACE2. GX/P2V/2017 RBD, GD/1/2019 RBD and SARS-CoV-2 RBD were colored in palegreen, lightpink and purple, respectively. (b) Model of GX/P2V/2017 RBD or GD/1/2019 RBD bound to hACE2, and compared with SARS-CoV-2 RBD bound to hACE2. (c), (d) The comparison of detailed interaction of GX/P2V/2017 RBD or GD/1/2019 RBD bound to pACE2 with that of SARS-CoV-2 RBD. Hydrogen bonds were shown as dotted lines with a cutoff of 3.3 Å, and the length of the bonds were labeled with the numbers in the corresponding color. (e), (f) The comparison of detailed interaction of GX/P2V/2017 RBD or GD/1/2019 RBD bound to hACE2 with that of SARS-CoV-2 RBD. Hydrogen bonds were shown as dotted lines with a cutoff of 3.3 Å, and the length of the bonds were labeled with the numbers in the corresponding color. (g) Sequence alignment of SARS-CoV-2 RBD, GX/P2V/2017 RBD and GD/1/2019 RBD. Residues of SARS-CoV-2 RBD binding to hACE2 with a cutoff of 4 Å are indicated in colorful triangles. The sequence alignment was generated by ESPrpt 3^[28].

Å。将结合了3种RBDs的pACE2或hACE2叠加比较时, 我们发现3种RBDs的取向是相似的(图4(a), (b)的右图)。

我们进一步分析GX/P2V/2017 RBD和GD/1/2019 RBD与pACE2和hACE2的相互作用细节, 并与SARS-CoV-2 RBD结合pACE2和hACE2进行比较。首先, 对于

GX/P2V/2017 RBD, V417(SARS-CoV-2 RBD为 K417)不仅失去了与pACE2的E30和hACE2的D30的盐桥相互作用, 而且由于其侧链较短, 进而降低了与pACE2和hACE2的相互作用(图4(c), (e)); L486(SARS-CoV-2 RBD为F486)不仅降低了与pACE2的Y83或hACE2的

Y83和M82的疏水相互作用,而且同样由于其侧链较短,也降低了与pACE2和hACE2的相互作用(图4(c), (e));至于Y490,由于SARS-CoV-2 RBD的F490靠主链与hACE2的K31仅形成1个相互作用,且与pACE2没有相互作用(表2),因而Y490可能不改变与hACE2的相互作用(图4(c), (e));E493(SARS-CoV-2 RBD为Q493)与pACE2的K31或hACE2的K31的距离变近,使氢键作用加强,并且增加了盐桥相互作用(图4(c), (e));H498(SARS-CoV-2 RBD为Q498)维持了与hACE2的Q42的氢键相互作用,而相比于结合hACE2,其结合pACE2时侧链发生偏转,降低了与pACE2的Q42的氢键相互作用,但其增加了与pACE2的Y41和hACE2的Y41和L45的疏水作用(图4(d), (f));T501(SARS-CoV-2 RBD为N501)与pACE2和hACE2的距离变远,使其失去了与pACE2的Y41和hACE2的Y41的氢键相互作用,且由于其侧链较短,也降低了与pACE2和hACE2的相互作用(图4(d), (f)).对于GD/1/2019 RBD,其与SARS-CoV-2 RBD相比,仅有两个氨基酸(第417和498位)不同,R417与SARS-CoV-2 RBD的K417相比,虽然静电性没变,但是其侧链的方向往靠近pACE2或hACE2的反方向偏转,进而降低了与pACE2的E30和hACE2的D30的盐桥相互作用(图4(c), (e));H498的作用类似于GX/P2V/2017 RBD的H498,增加了与pACE2的Y41和hACE2的Y41和L45的疏水作用,维持了与hACE2的Q42的氢键相互作用,但降低了与pACE2的Q42的相互作用(图4(d), (f)).这些结果表明, GD/1/2019 RBD与pACE2的亲合力可能略弱于hACE2,这与SPR结果一致。

3 结论与讨论

病毒囊膜蛋白与宿主细胞受体之间的相互作用是介导病毒感染的首要步骤,且在很大程度上决定着病毒的组织嗜性和宿主范围,因而对该相互作用机制的

研究不仅有助于理解病毒感染,而且对探究病毒的自然宿主或中间宿主提供帮助.从SARS-CoV、MERS-CoV到SARS-CoV-2,虽然这三者皆被认为起源于蝙蝠,但是其传播路径可能不同,例如单峰骆驼被认为是MERS-CoV的中间宿主^[29],而果子狸可能将其感染的SARS-CoV传染给人(果子狸作为SARS-CoV的中间宿主存在争议)^[9],目前尚未确定SARS-CoV-2的中间宿主.考虑到中间宿主与蝙蝠在时空环境上的差异,因而在短时间内找到SARS-CoV-2的中间宿主是相当困难的。

穿山甲是目前唯一携带SARS-CoV-2相关冠状病毒的哺乳动物,因而一些学者认为穿山甲是SARS-CoV-2的天然宿主或中间宿主,但是仍存争议.近期的研究报道显示,来自中国的17只穿山甲对SARS-CoV-2的血清学检测均为阴性^[30],这表明中国的穿山甲可能不携带新冠相关的冠状病毒.此外,有研究人员发现,穿山甲冠状病毒在S蛋白的S1/S2剪切位点不存在PRRA基序^[17,20,31],说明SARS-CoV-2可能不直接来自穿山甲.考虑到穿山甲和蝙蝠具有相似的生存环境,二者皆生活在洞穴中,且都携带SARS-CoV-2相关的冠状病毒,表现出趋同进化现象^[17],因此,有学者认为, SARS-CoV-2起源于蝙蝠冠状病毒、穿山甲冠状病毒或二者的重组,然而任何一种假说都需要充分的证据来证明。

在本研究中,我们阐明了SARS-CoV-2 RBD与pACE2相互作用的分子基础,并进一步分析了两株具有代表性的穿山甲冠状病毒分别与pACE2和hACE2受体的结合特征.这些研究结果有助于理解SARS-CoV-2的进化,以及探寻其中间宿主.此外,我们的研究结果表明,穿山甲冠状病毒可以结合人受体ACE2,存在感染人的潜力,强调了监测穿山甲的重要性,以防止病毒外溢.但是穿山甲是否为SARS-CoV-2的天然宿主或中间宿主,仍有待进一步研究。

参考文献

- 1 Gao G F. From "A"IV to "Z"IKV: Attacks from emerging and re-emerging pathogens. *Cell*, 2018, 172: 1157–1159
- 2 Drosten C, Günther S, Preiser W, et al. Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med*, 2003, 348: 1967–1976
- 3 Zaki A M, van Boheemen S, Bestebroer T M, et al. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. *N Engl J Med*, 2012, 367: 1814–1820
- 4 Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*, 2020, 382: 727–733
- 5 Zhou P, Yang X L, Wang X G, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, 2020, 579: 270–273
- 6 Tan W, Zhao X, Ma X, et al. A novel coronavirus genome identified in a cluster of pneumonia cases — Wuhan, China 2019–2020. *China CDC*

- Weekly, 2020, 2: 61–62
- 7 Wang Q, Zhang Y, Wu L, et al. Structural and functional basis of SARS-CoV-2 entry by using human ACE2. *Cell*, 2020, 181: 1–11
- 8 Guan Y. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science*, 2003, 302: 276–278
- 9 Ge X Y, Li J L, Yang X L, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*, 2013, 503: 535–538
- 10 Zhou H, Chen X, Hu T, et al. A novel bat coronavirus closely related to SARS-CoV-2 contains natural insertions at the S1/S2 cleavage site of the spike protein. *Curr Biol*, 2020, 30: 2196–2203.e3
- 11 Zhang Q, Zhang H, Gao J, et al. A serological survey of SARS-CoV-2 in cat in Wuhan. *Emerg Microbes Infect*, 2020, 9: 2013–2019
- 12 Sit T H C, Brackman C J, Ip S M, et al. Infection of dogs with SARS-CoV-2. *Nature*, 2020, 586: 776–778
- 13 Enserink M. Coronavirus rips through Dutch mink farms, triggering culls. *Science*, 2020, 368: 1169
- 14 Wang L, Mitchell P K, Calle P P, et al. Complete genome sequence of SARS-CoV-2 in a tiger from a U.S. zoological collection. *Microbiol Resour Announc*, 2020, 9: e00468–00420
- 15 Wu L, Chen Q, Liu K, et al. Broad host range of SARS-CoV-2 and the molecular basis for SARS-CoV-2 binding to cat ACE2. *Cell Discov*, 2020, 6: 68
- 16 Zhao X, Chen D, Szabla R, et al. Broad and differential animal angiotensin-converting enzyme 2 receptor usage by SARS-CoV-2. *J Virol*, 2020, 94: e00940–00920
- 17 Lam T T Y, Jia N, Zhang Y W, et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins. *Nature*, 2020, 583: 282–285
- 18 Liu P, Chen W, Chen J P. Viral metagenomics revealed Sendai virus and coronavirus infection of Malayan pangolins (*Manis javanica*). *Viruses*, 2019, 11: 979
- 19 Xiao K, Zhai J, Feng Y, et al. Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature*, 2020, 583: 286–289
- 20 Zhang T, Wu Q, Zhang Z. Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Curr Biol*, 2020, 30: 1346–1351.e2
- 21 Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*, 1997, 276: 307–326
- 22 Read R J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystlogr D Biol Crystlogr*, 2001, 57: 1373–1382
- 23 Emsley P, Cowtan K. *Coot*: Model-building tools for molecular graphics. *Acta Crystlogr D Biol Crystlogr*, 2004, 60: 2126–2132
- 24 Adams P D, Afonine P V, Bunkóczi G, et al. *PHENIX*: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystlogr D Biol Crystlogr*, 2010, 66: 213–221
- 25 Chen V B, Arendall Iii W B, Headd J J, et al. *MolProbity*: All-atom structure validation for macromolecular crystallography. *Acta Crystlogr D Biol Crystlogr*, 2010, 66: 12–21
- 26 Song Y, DiMaio F, Wang R Y R, et al. High-resolution comparative modeling with RosettaCM. *Structure*, 2013, 21: 1735–1742
- 27 Yan Y, Tao H, He J, et al. The HDock server for integrated protein–protein docking. *Nat Protoc*, 2020, 15: 1829–1852
- 28 Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*, 2014, 42: W320–W324
- 29 Reusken C B, Haagmans B L, Müller M A, et al. Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: A comparative serological study. *Lancet Infect Dis*, 2013, 13: 859–866
- 30 Deng J, Jin Y, Liu Y, et al. Serological survey of SARS-CoV-2 for experimental, domestic, companion and wild animals excludes intermediate hosts of 35 different species of animals. *Transbound Emerg Dis*, 2020, 67: 1745–1749
- 31 Liu P, Jiang J Z, Wan X F, et al. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog*, 2020, 16: e1008421

补充材料

图S1 pACE2的Q42、K353和D355与SARS-CoV-2 RBD之间的相互作用

图S2 pACE2、hACE2 和cACE2序列比对

本文以上补充材料见网络版csb.scichina.com. 补充材料为作者提供的原始数据, 作者对其学术质量和内容负责.

Summary for “新型冠状病毒结合穿山甲ACE2受体的分子机制”

Molecular basis of pangolin ACE2 engaged by COVID-19 virus

Lili Wu^{1,2}, Jiaqi Su³, Sheng Niu^{3,4}, Qian Chen^{3,5}, Yanfang Zhang³, Jinghua Yan¹, Yi Shi³, Jianxun Qi³, George Fu Gao^{3*} & Qihui Wang^{3,5,6*}¹ CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;² University of Chinese Academy of Sciences, Beijing 100049, China;³ CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China;⁴ College of Animal Science and Veterinary Medicine, Shanxi Agricultural University, Taigu 030801, China;⁵ Institute of Physical Science and Information, Anhui University, Hefei 230039, China;⁶ Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, China* Corresponding authors, E-mail: gaof@im.ac.cn; wangqihui@im.ac.cn

The virus caused coronavirus disease 2019 (COVID-19 virus), also called severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is the seventh coronavirus that can infect humans, and belongs to the genus β -coronavirus in the family *Coronaviridae*. As of 5 November 2020, SARS-CoV-2 has caused >4700000 confirmed cases and >1200000 related deaths in 219 countries and regions, bringing great challenges to global public health. Currently, there is no approved therapeutics or vaccines for the treatment of this disease. Several studies suggested SARS-CoV-2 might have originated from bats based on phylogenetic analysis, but the intermediate host of the virus is still unknown. Besides humans, cats, dogs, tigers, lions, minks, and other species have been reported to be infected by SARS-CoV-2. Several studies have reported pangolins as the only other mammalian species carrying coronaviruses related to SARS-CoV-2 besides bats, and suggested pangolins might be the intermediate host of SARS-CoV-2. Previously, we and other groups identified the angiotensin converting enzyme 2 (ACE2), the receptor of SARS-CoV, also functions as the entry receptor of SARS-CoV-2 and is recognized by the receptor binding domain (RBD) of the spike protein (S). Then, we elucidated the interaction between SARS-CoV-2 RBD and human ACE2 (hACE2) or cat ACE2 (cACE2), and also found that SARS-CoV-2 had broad potential host range, including domestic animals, companion pets and wild animals. In this study, we investigated the binding features of SARS-CoV-2 and two pangolin coronaviruses (pangolin-CoVs, GX/P2V/2017 and GD/1/2019) that recognize the receptors of both pangolin ACE2 (pACE2) and hACE2 using surface plasmon resonance (SPR) and structural methods. We further determined the crystal structure of SARS-CoV-2 S (RBD) in complex with pACE2 at a 2.3 Å resolution, revealing the similarity in the binding mode between SARS-CoV-2 RBD to hACE2 and to cACE2. Interestingly, we found that the SARS-CoV-2 RBD-pACE2 complex is more similar to the SARS-CoV-2 RBD-hACE2 complex than to SARS-CoV-2 RBD-cACE2 complex. Furthermore, we modeled the interactions of GX/P2V/2017 RBD and GD/1/2019 RBD bound to pACE2 and hACE2, respectively, and found that both viruses adopted similar binding mode as SARS-CoV-2 RBD to pACE2 and hACE2. However, crystal structure and homology modeling implied that the interaction between the three RBDs and pACE2 were slightly weaker than their respective binding to hACE2, which was consistent with the receptor-RBD protein interaction measured by SPR. These results could facilitate better understanding of SARS-CoV-2 evolution, indicate the potential of pangolin-CoVs to enter human population that may lead to another outbreak, and highlight the importance of monitoring pangolin CoVs to prevent possible spillovers.

SARS-CoV-2, pangolin coronaviruses, receptor binding domain (RBD), angiotensin converting enzyme 2(ACE2), crystal structure

doi: [10.1360/TB-2020-1372](https://doi.org/10.1360/TB-2020-1372)