

新冠病毒 Nsp1 蛋白结构与功能的生物信息学分析及原核表达

刘玲¹, 李璟², 范蕾², 宣焱¹, 杜淼¹, 张德玖³, 李卓禧², 陈晓聪², 杨娅男¹, 徐本锦¹

摘要:目的 Nsp1 是新冠病毒的主要毒力因子,介导了病毒在宿主细胞中的免疫逃逸,扩大了病毒感染范围。本研究对 Nsp1 进行系统生信分析及原核表达,有助于全面理解该蛋白在新冠病毒感染宿主细胞中的分子机理。方法 采用 Pfam、TMHMM、ProtScale、ExpASy 和 SignalP 4.0 等工具对 Nsp1 的翻译后修饰、理化性质、跨膜螺旋、相互作用网络、同源性及进化特点等进行系统分析;利用分子克隆技术构建重组表达载体 pET-22b-Nsp1 并进行原核表达。结果 Nsp1 由 180 个氨基酸组成,分子量 19.78 kDa,等电点为 5.36,不稳定指数 28.83,在哺乳动物中的半衰期约 30 h,在大肠杆菌内超过 10 h,有 12 个可能的磷酸化位点和 3 个 O-糖基化位点,无信号肽和跨膜螺旋,亲水性较强;二级结构分析显示,Nsp1 中无规则卷曲占比最高(45.00%),其次是 α 螺旋(25.56%)和延伸链(20.56%), β -转角占比最低(8.89%);序列对比和进化分析显示,与 SARS-CoV-2 Nsp1 序列一致性最高的是蝙蝠非典型冠状病毒 WIV1(85.0%);原核表达发现 Nsp1 主要在沉淀中表达,经质谱鉴定目的蛋白就是 Nsp1。结论 本研究为新冠病毒 Nsp1 的表达、纯化及功能分析提供了重要借鉴,有助于进一步揭示 Nsp1 的生物学功能及开展相关抑制剂和抗病毒药物的研发。

关键词:新冠病毒;Nsp1 蛋白;生信分析;结构与功能;进化分析

中图分类号:R373.1 文献标识码:A 文章编号:1002-2694(2022)07-0566-11

Bioinformatics analysis of the structure, function and prokaryotic expression of SARS-CoV-2 nonstructural protein

LIU Ling¹, LI Jing², FAN Lei², XUAN Yan¹, DU Miao¹, ZHANG De-jiu³,
LI Zhuo-xi², CHEN Xiao-cong², YANG Ya-nan¹, XU Ben-jin¹

(1. Department of Medical Laboratory, Fenyang College of Shanxi Medical University, Fenyang 032200, China;

2. Department of Basic Medicine, Fenyang College of Shanxi Medical University, Fenyang 032200, China;

3. Institute for Translational Medicine, Qingdao University, Qingdao 266021, China)

Abstract: The Nsp1 protein, the main virulence factor of the virus of SARS-CoV-2, mediates viral immune escape in host cells and expands the viral infection range. Systematic bioinformatics analysis and prokaryotic expression of Nsp1 protein were performed to aid the understanding of the molecular mechanism of SARS-CoV-2 infection in host cells. Pfam, TMHMM, ProtScale, ExpASy, SignalP 4.0 and other tools were used to systematically analyze the post-translational modifications, physicochemical properties, transmembrane helix, interaction network, homology and evolutionary characteristics of Nsp1 protein. The recombinant expression vector pET-22b-Nsp1 was constructed through molecular cloning technology and expressed in a

prokaryotic system. Nsp1 is composed of 180 amino acids, with a molecular weight of 19.78 kDa, an isoelectric point of 5.36 and an instability index of 28.83. It has a half-life of 30 h in mammalian reticulocytes and more than 10 h in *E. coli*. It has 12 potential phosphorylation sites, three potential O-glycosylation sites and no signal peptide. It is a hydrophilic protein without a transmembrane helix. Secondary structure analysis indicated that the highest proportion of structural components comprised random coils (45.00%), followed by α -helices (25.56%) and extended chains (20.56%); the lowest proportion comprised β -turns (8.89%). Multiple sequence alignment and evolutionary analysis revealed that Bat SARS-like corona-

山西省基础研究计划(自由探索类)项目(No.20210302123397)、国家级大学生创新创业训练计划项目(No.202117114001)、山西省高等学校大学生创新创业训练计划重点项目(No.S202117114008)、山西省高等学校科技创新项目(No.2020L0749)、山西医科大学汾阳学院引进人才科研启动金项目(No.2020A01)、吕梁市科技计划项目(No.2020SHFZ29)和国家自然科学基金项目(No.31870816)联合资助。刘玲、李璟、范蕾对本文有同等贡献。

通讯作者:徐本锦,Email:bj0726@sxmu.edu.cn;

ORCID: 0000-0002-2480-0040

作者单位:1. 山西医科大学汾阳学院医学检验系,汾阳 032200;

2. 山西医科大学汾阳学院基础医学部,汾阳 032200;

3. 青岛大学转化医学研究院,青岛 266021

virus WIV1 had the highest sequence identity (85.00%) with the Nsp1 protein of SARS-CoV-2. After prokaryotic expression, Nsp1 protein was found to be mainly expressed in the precipitate after centrifugation of bacterial lysate. The target protein was further identified as Nsp1 by mass spectrometry. This study provides an important reference for the expression, purification and functional analysis of SARS-CoV-2 Nsp1 protein, and further reveals the biological functions of Nsp1, thus providing a reference for research and development of related inhibitors and antiviral drugs.

Keywords: SARS-CoV-2; Nsp1 protein; bioinformatics analysis; structure and function; evolutionary analysis

Supported by Fundamental Research Program of Shanxi Province (No.20210302123397), the National College Students' Innovation and Entrepreneurship Training Program (No.202117114001), Key Projects of Innovation and Entrepreneurship Training Program for College Students in Shanxi Province (No.S202117114008), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (No.2020L0749), Starting Fund for Talent Introduction of Fenyang College of Shanxi Medical University (No.2020A01), Project of Lvliang City Science and Technology Program (No.2020SHFZ29), and National Natural Science Foundation of China (No.31870816). Liu Ling, Li Jing and Fan Lei contributed equally to this article. Corresponding author: Xu Ben-jin, Email: bj0726@sxmu.edu.cn

2019年12月,由新冠病毒(SARS-CoV-2)引起的肺炎疫情在中国武汉暴发,患者表现为发热、咳嗽、肺部磨玻璃样病变等呼吸系统病变,或伴随腹泻等消化系统症状^[1-2]。截止2021年9月28日,新冠肺炎累计确诊病例已超过2.31亿例,累计死亡病例超过475万例^[3]。新冠疫情的迅速蔓延给全球医疗体系带来了严峻考验。因此,揭示SARS-CoV-2的传播机制及其关键蛋白的结构与功能成为遏制疫情的关键。

SARS-CoV-2属 β 冠状病毒,是一种包膜单股正链RNA病毒^[4-5]。其基因组约30 kb,由一个5'帽子和5'非翻译区(5' UTR),10个开放读码框(open reading frames, ORFs)以及一个多聚腺苷酸化的3'非翻译区(3' UTR)组成^[6]。ORF1a和ORF1b是SARS-CoV-2的2个主要ORF,约占病毒基因组的70%(图1)。ORF1a通过核糖体移码被翻译成多聚蛋白pp1a或pp1ab^[7-8],最后被加工成16种非结构蛋白(non-structural protein, NSPs),NSPs具有多种酶活性,参与病毒RNA复制和转录调控。

Nsp1由ORF1a的5'末端编码(图1),是感染细胞中产生的第1种冠状病毒蛋白^[9],其主要功能是通过结合核糖体抑制宿主基因表达^[10-11]。目前,国内外有关SARS-CoV-2 Nsp1蛋白结构与功能方面的研究报道仍然较少。本研究利用生信分析手段对SARS-CoV-2 Nsp1的性质、翻译后修饰、结构域等进行了系统分析,还对其进行了同源性分析、进化分析及原核表达。本研究有助于揭示Nsp1在SARS-

CoV-2 侵染宿主细胞中的作用机制,同时有助于加快针对Nsp1的靶向药物研发。

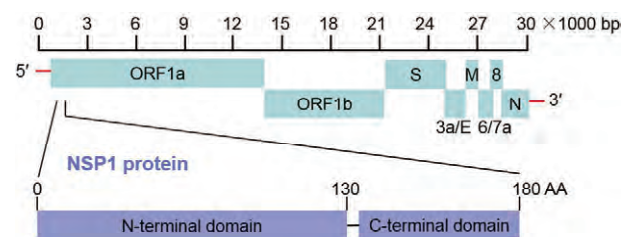


图1 SARS-CoV-2基因组结构示意图

Fig.1 Schematic of SARS-CoV-2 genome organisation

1 材料与方法

1.1 材料 限制性核酸内切酶 *Nde*I、*Xho*I 购自 NEB, pET-22b 空质粒为本实验室保存, DNA 分子量标准购自宝生物(大连), 蛋白分子量标准、质粒提取、凝胶回收试剂盒以及感受态细胞购自全式金公司, 其它生化试剂购自国药。

1.2 载体构建及原核表达 用 *Xho*I 和 *Nde*I 对 pET-22b 空质粒和 *Nsp1* 基因同时双酶切, 回收酶切产物后将 *Nsp1* 基因与载体连接, 转化 Top10 感受态细胞并进行菌落 PCR 验证。Nsp1 蛋白的诱导表达步骤参照文献执行^[12]。

1.3 生信分析 对 Nsp1 蛋白的理化特性、跨膜螺旋、亲/疏水性等进行全面分析, 相关分析网站如表 1 所示。Nsp1 蛋白的多序列比对与系统进化分析参照文献执行^[12]。

表1 生物信息分析网址

Tab.1 Bioinformatics analysis websites

分析项目	分析网站
Analysis items	Analysis websites
理化性质	https://web.expasy.org/protparam/
跨膜螺旋	https://services.healthtech.dtu.dk/service.php?TMHMM-2.0
亲/疏水性	https://web.expasy.org/protscale/
磷酸化位点	https://services.healthtech.dtu.dk/service.php?NetPhos-3.1
糖基化位点	https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0 https://services.healthtech.dtu.dk/service.php?YinOYang-1.2
亚细胞定位	https://psort.hgc.jp/form2.html
信号肽	http://www.cbs.dtu.dk/services/SignalP-4.0/
二级结构分析	https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopm.html
结构域分析	https://www.uniprot.org/uniprot/P0DTC1#PRO_Q000449635
三维结构	https://swissmodel.expasy.org/interactive

2 结果

2.1 Nsp1 的理化特性 Nsp1 由 180 个氨基酸残基构成(共 20 种),相对分子量 19.78 kDa,等电点 5.36,负电性氨基酸残基数(Asp + Glu)为 27,正电性氨基酸残基数(Arg+Lys)为 19,占比最高的残基是甘氨酸(12.20%),其次是缬氨酸和亮氨酸(均为 11.70%)。见表 2。Nsp1 的分子式为 $C_{872}H_{1383}N_{247}O_{270}S_4$,原子总数为 2 776,消光系数 $12\ 950\ (\text{mol/L})^{-1}\text{cm}^{-1}$ 。在哺乳动物细胞中的半衰期约 30 h,在酵母中超过 20 h,在 *E.coli* 内超过 10 h。Nsp1 的不稳定指数约 28.83,脂肪系数达 89.72,平均亲水性指数 -0.378 。

2.2 Nsp1 蛋白跨膜结构分析 在线预测 Nsp1 的跨膜结构,结果显示该蛋白无跨膜区(图 2)。

2.3 Nsp1 的亲水与疏水特性 结果表明,Nsp1 的 L88 疏水特性最强,分值 1.878,N162 亲水特性最强,分值 -2.611 ;含有亲水氨基酸 112 个,疏水氨基酸 60 个,平均亲水指数 -0.378 ,提示 Nsp1 是一个亲水蛋白(图 3)。

表2 新冠病毒 Nsp1 蛋白的氨基酸组成

Tab.2 Amino acid composition of SARS-CoV-2 Nsp1 protein

氨基酸	数量/个	频率/%	氨基酸	数量/个	频率/%
Amino acids	Number	Frequency	Amino acids	Number	Frequency
Ala (A)	7	3.90	Leu (L)	21	11.70
Arg (R)	10	5.60	Lys (K)	9	5.00
Asn (N)	6	3.30	Met (M)	3	1.70
Asp (D)	9	5.00	Phe (F)	5	2.80
Cys (C)	1	0.60	Pro (P)	8	4.40
Gln (Q)	7	3.90	Ser (S)	10	5.60
Glu (E)	18	10.00	Thr (T)	7	3.90
Gly (G)	22	12.20	Trp (W)	1	0.60
His (H)	7	3.90	Tyr (Y)	5	2.80
Ile (I)	3	1.70	Val (V)	21	11.70

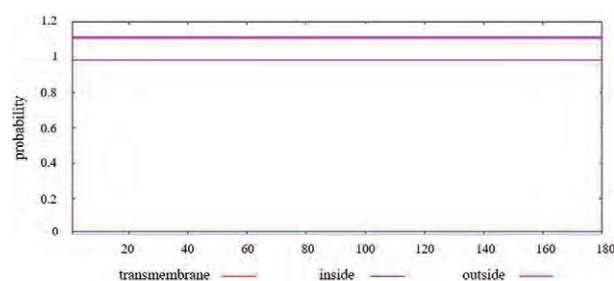


图2 新冠病毒 Nsp1 蛋白的跨膜结构预测

Fig.2 Transmembrane structure prediction of SARS-CoV-2 Nsp1 protein

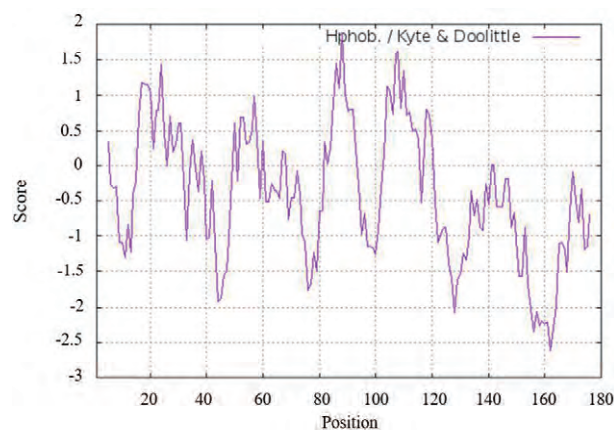


图3 新冠病毒 Nsp1 蛋白亲水/疏水性分析

Fig.3 Hydrophobic/hydrophobic analysis of SARS-CoV-2 Nsp1 protein

2.4 Nsp1 的磷酸化修饰预测 Nsp1 存在 12 个磷酸化修饰位点,其中 S17、S34、S40、S74、S100、S135、S142 和 S166 为丝氨酸修饰位点,T151 为苏氨酸修饰位点,Y68、Y97 和 Y154 为酪氨酸修饰位点(图 4、表 3)。

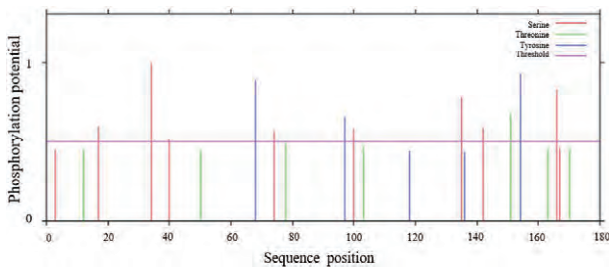


图 4 新冠病毒 Nsp1 蛋白的磷酸化位点

Fig.4 Phosphorylation sites of SARS-CoV-2 Nsp1 protein

表 3 Nsp1 蛋白的磷酸化位点及对应的激酶

Tab.3 Phosphorylation sites of Nsp1 protein and corresponding kinases

位点	分值	激酶	位点	分值	激酶	位点	分值	激酶
Site	Score	Kinase	Site	Score	Kinase	Site	Score	Kinase
S17	0.592	PKA	S74	0.567	PKG	S142	0.588	CKII
S34	0.990	unsp	Y97	0.654	unsp	T151	0.669	CKII
S40	0.511	CKII	S100	0.577	unsp	Y154	0.961	unsp
Y68	0.885	unsp	S135	0.775	unsp	S166	0.828	unsp

注: unsp 表示未确定的激酶。

2.5 Nsp1 蛋白亚细胞定位与信号肽预测 预测得知 56.5%的 Nsp1 在细胞质,30.4%存在于细胞核,线粒体、分泌小泡和细胞骨架中各占 4.3%。对 N 端前 70 个残基进行分析后发现,Nsp1 蛋白不含信号肽(图 5)。

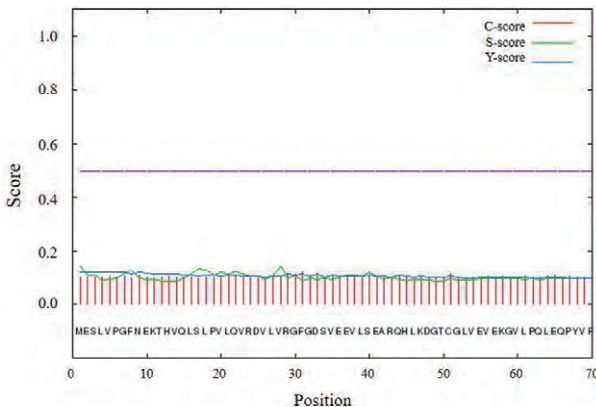


图 5 新冠病毒 Nsp1 蛋白信号肽预测

Fig.5 Prediction of SARS-CoV-2 Nsp1 protein signal peptide

2.6 Nsp1 蛋白糖基化位点预测 没有信号肽的蛋白质不太可能暴露于 N-糖基化机制,因此可能不会被糖基化(在体内)。利用 NetNGlyc 1.0 Server 预测后发现,Nsp1 蛋白可能不会被 N-糖基化修饰(图 6)。

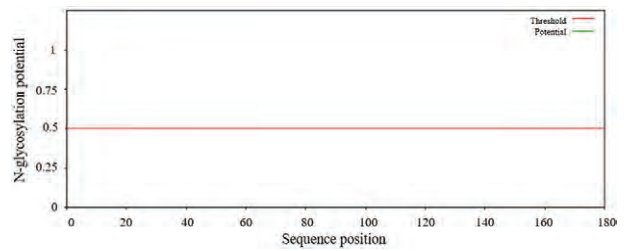


图 6 新冠病毒 Nsp1 蛋白 N-糖基化位点

Fig.6 N-glycosylation site of SARS-CoV-2 Nsp1 protein

O-糖基化修饰预测发现,Nsp1 有 3 个可能的 O-糖基化修饰位点:S40、S74 和 T163(图 7)。

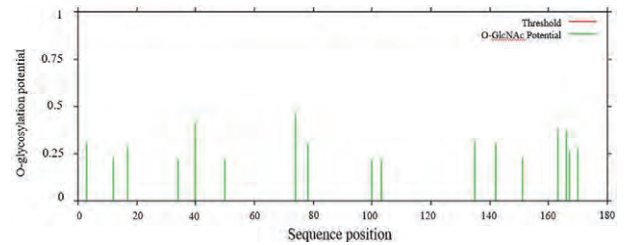


图 7 新冠病毒 Nsp1 蛋白 O-糖基化位点

Fig.7 O-glycosylation sites of SARS-CoV-2 Nsp1 protein

2.7 Nsp1 蛋白二级结构分析 使用 Prabi 网站对 Nsp1 蛋白的二级结构进行分析。结果显示,Nsp1 中 α-螺旋占 25.56%,延伸链占 20.56%,β-转角占 8.89%,无规则卷曲占 45.00%(图 8)。

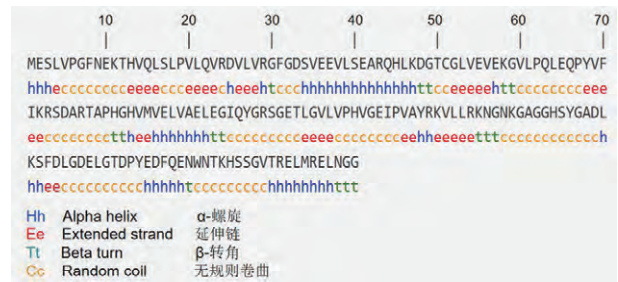


图 8 新冠病毒 Nsp1 蛋白二级结构

Fig.8 Secondary structure of SARS-CoV-2 Nsp1 protein

2.8 Nsp1 蛋白的相互作用网络 最新的蛋白质组学分析显示^[13],Nsp1 与 6 个蛋白存在相互作用,分别是:DNA 引物酶小亚基(PRIM1),DNA 引物酶大亚基(PRIM2),DNA 聚合酶 α 复合物催化亚基(POLA1),DNA 聚合酶 α 复合物辅助亚基(POLA2),Plakophilin-2(PKP2)及前胶原半乳糖基转移酶 1(COLGALT1)(图 9)。

2.9 Nsp1 蛋白结构域分析 Nsp1 含有 2 个结构域,N 端为冠状病毒 Nsp1 球状结构域(T12-G133),

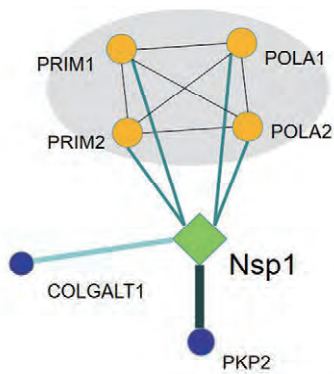


图9 新冠病毒 Nsp1 蛋白相互作用网络
Fig.9 Interaction network of SARS-CoV-2 Nsp1 protein

C 端为 β 冠状病毒 Nsp1 C 末端结构域 (E148-G179)(图 10)。Nsp1 是 α 冠状病毒和 β 冠状病毒的一个特征性蛋白,在抑制宿主基因表达和抗病毒反应方面表现出功能保守性和机制多样性。尽管冠状病毒 Nsp1 蛋白之间的序列同源性较低,但其核心结构共享一个相对保守的球状结构域。

Sarbe 冠状病毒 Nsp1 的 C 末端结构域通过与核糖体 40S 亚基的 mRNA 通道结合,干扰 mRNA 结合进而抑制宿主蛋白质翻译。这种抑制机制可能是 Sarbe 冠状病毒特有的,因为 Nsp1 的 C 端在 α 冠状病毒中较短,并且在包括 MERS 冠状病毒在内的其他 β 冠状病毒中并不高度保守。

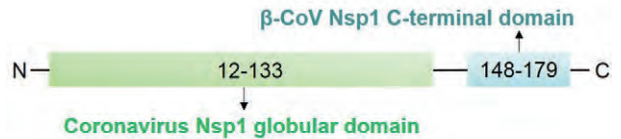


图 10 新冠病毒 Nsp1 蛋白的结构域分析
Fig.10 Domain analysis of SARS-CoV-2 Nsp1 protein

2.10 Nsp1 蛋白的三维结构 根据蛋白质三级结构建立网站构建 Nsp1 蛋白的三级结构。结果只显示 10~126 位点,148~180 位点的三级结构模型(图 11)。

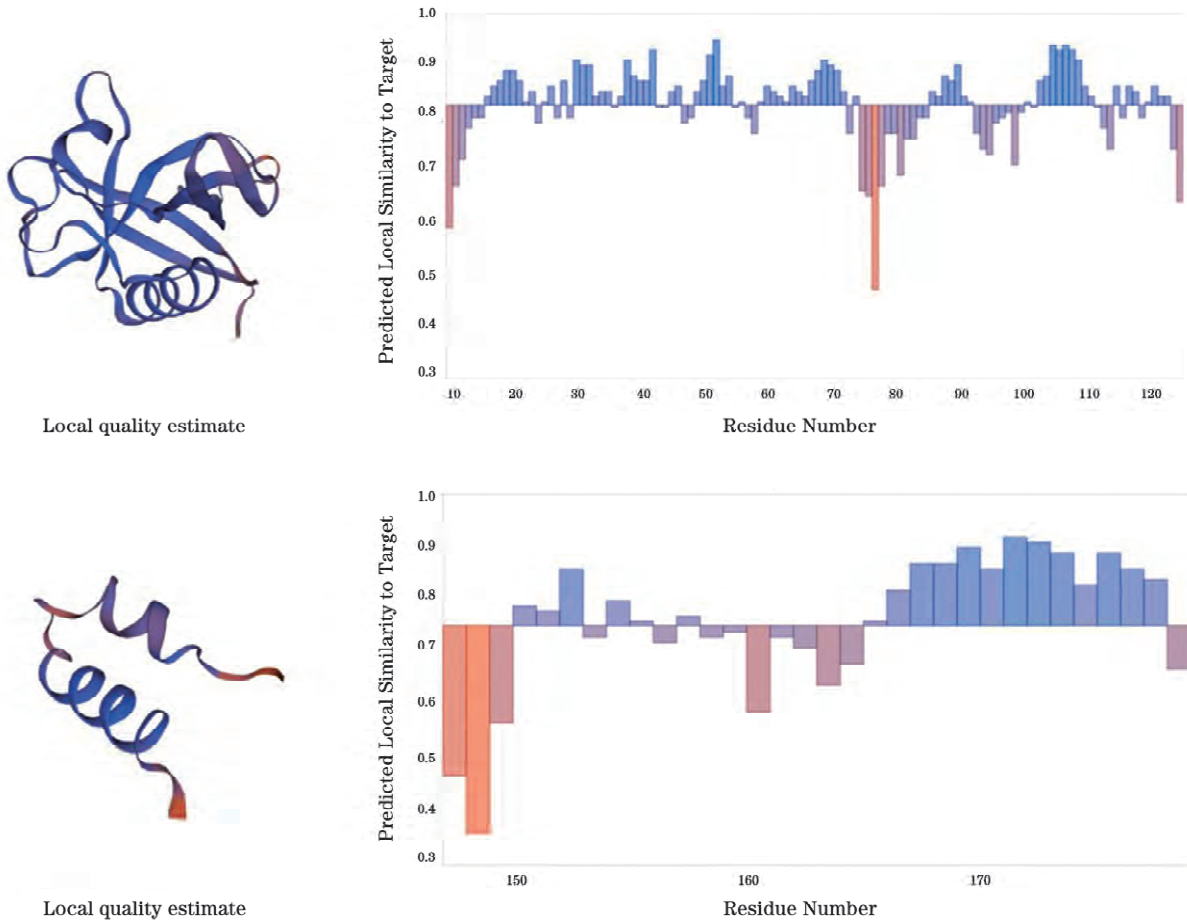


图 11 新冠病毒 Nsp1 蛋白的三维结构分析
Fig.11 Three-dimensional structure analysis of SARS-CoV-2 Nsp1 protein

2.11 Nsp1 蛋白序列同源性分析 通过 Uniprot 网站检索发现,与新冠 Nsp1 序列最接近的是 Bat SARS-like coronavirus WIV1(85.0%),接着是 Bat coronavirus Rp/Shaanxi 2011、严重急性呼吸综合

冠 状 病 毒、SARS coronavirus WH20 和 SARS coronavirus PUMC02,序列一致性均为 84.4%,与 Bat Hp-beta-coronavirus/Zhejiang2013 的序列相似 度为 28.7%。所有序列中相同的残基占比 21.1% (* 所示),性质相似的占比 19.4%(.:所示),相似性 较低的占比 16.1%(.:所示)(图 12)。

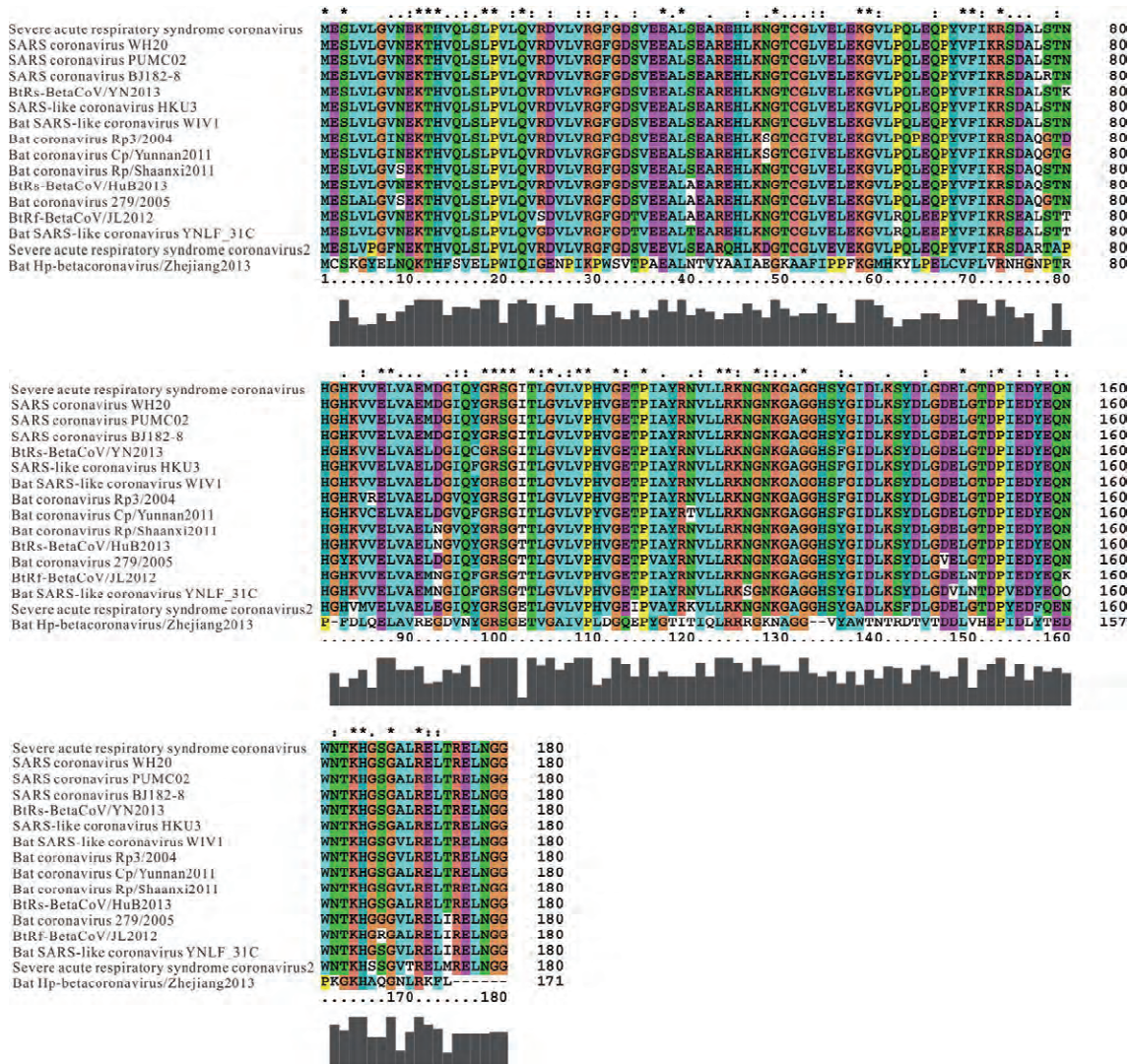


图 12 新冠病毒 Nsp1 蛋白序列同源性

Fig.12 Homology of SARS-CoV-2 Nsp1 protein sequence

2.12 Nsp1 的进化树构建 结果显示,这 15 种病毒的蛋白与 SARS-CoV-2 的 Nsp1 蛋白亲缘关系较远,BtRs-BetaCoV/YN2013 与非典型冠状病毒 HKU3 聚为一支,置信度为 72,聚为一支后与蝙蝠非典型冠状病毒 WIV1 聚为一支,置信度为 39。此外,BtRf-BetaCoV/JL2012 与 Bat SARS-like coronavirus YNLF_31C 聚类为同一支,置信度 78;Bat coronavirus Rp/Shaanxi2011 与 BtRs-BetaCoV/HuB2013 聚类为同一支,置信度 48(图 13)。

2.13 表达载体 pET-22b-Nsp1 的构建

2.13.1 Nsp1 的酶切与 PCR 验证 用 Nde I 与 Xho I 对 Nsp1 进行酶切(图 14A);PCR 扩增 Nsp1 序列,电泳检测显示片段大小合适(图 14B)。

2.13.2 pET-22b-Nsp1 的酶切及转化 用 Nde I 与 Xho I 同时酶切重组载体,电泳检测显示片段大小正确(图 15A);将 pET-22b-Nsp1 转化 Top10,于 37 °C 恒温培养 13 h(图 15B 左),提取质粒后测序验证,选取正确的转化 BL21(图 15B 右)。

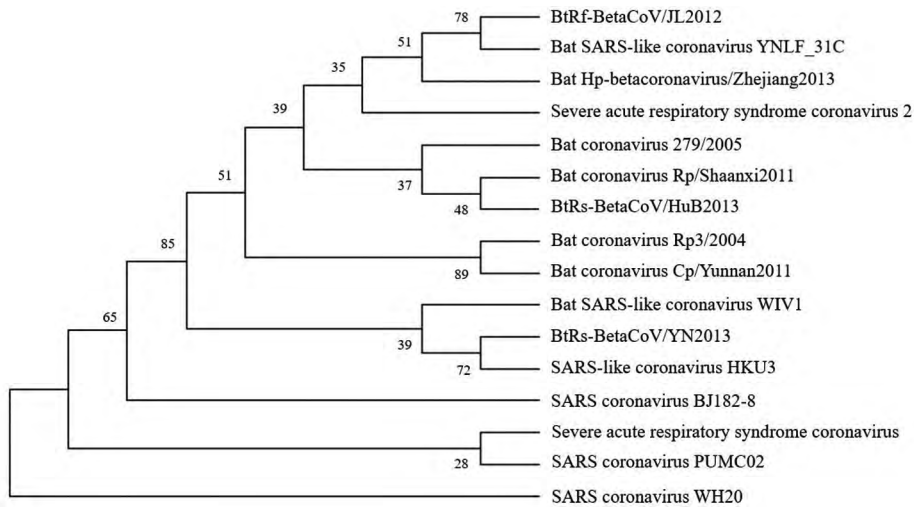
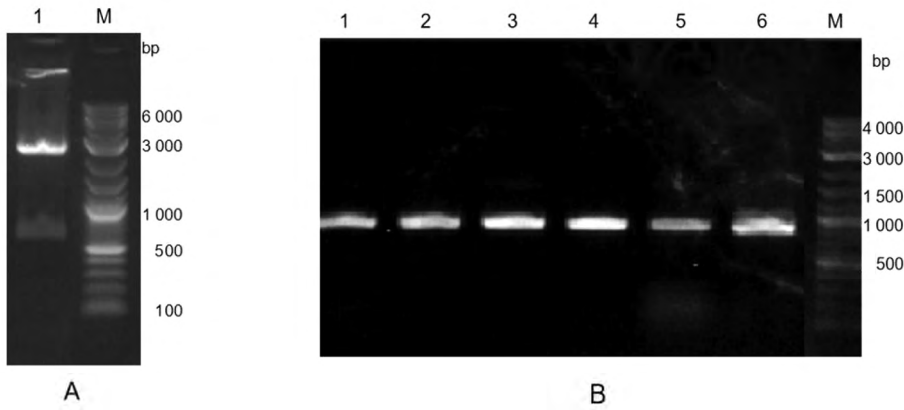


图 13 新冠病毒 Nsp1 蛋白系统进化树

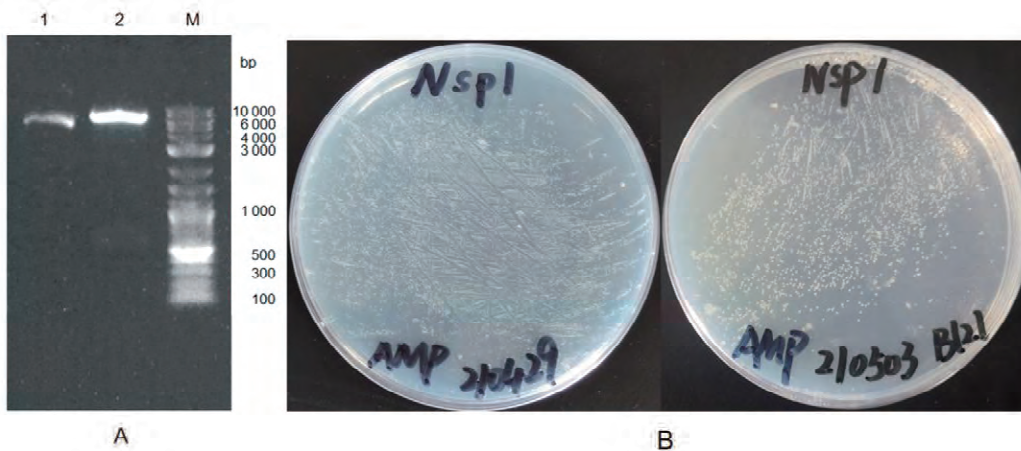
Fig.13 Phylogenetic tree of SARS-CoV-2 Nsp1 protein



A: Nsp1 双酶切实验 1: Nsp1 酶切产物 M: 2-log DNA marker B: 菌落 PCR 1~6 阳性单克隆

图 14 Nsp1 的酶切与菌落 PCR

Fig.14 Restriction endonuclease digestion and colony PCR validation of Nsp1



A: pET-22b-Nsp1 酶切实验 1: 不酶切 2: 双酶切 B: pET-22b-Nsp1 转化实验 (左: Top10, 右: BL21)

图 15 pET-22b-Nsp1 的酶切和转化

Fig.15 Enzyme digestion and transformation of pET-22b-Nsp1

2.14 Nsp1 的蛋白表达 菌体培养至 OD₆₀₀ 为 0.4~0.6 时用 0.5 mmol/L IPTG 去诱导,12% SDS-PAGE 检测显示,Nsp1 蛋白表达量经诱导之后显著上调(泳道 4 和泳道 10);对诱导前和诱导后的全细胞裂解液、上清液、沉淀分别检测后发现,Nsp1 主要表达在沉淀(泳道 3、6;泳道 9、12)(图 16)。对泳道 6 和 12 的相应位置进行切胶并质谱鉴定(图 17),发现沉淀中表达的确是目标蛋白 Nsp1。

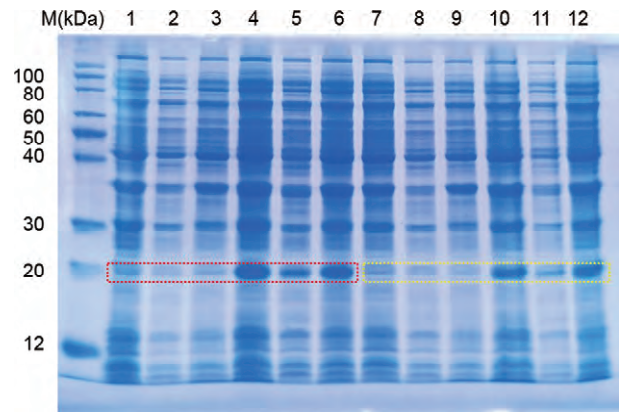


图 16 Nsp1 蛋白的诱导表达检测

Fig.16 Detection of induced expression of Nsp1 protein

Protein View: pdb|7K5I|1

Chain 1, Host translation inhibitor nsp1

Score: 137
Expect: 4.9e-12
Nominal mass (Mr): 19820
Calculated pI: 5.36

Protein sequence coverage: 85%

Matched peptides shown in **bold red**.

1 MESLVPGFNE KTHVQLSLPV LQVRDVLVVRG FGDSEVEVLS EARQHLKDGTT
51 CGLVEVEKGV LPQLQPYVVF IKRSDARTAP HGHVMVELVA ELEGIQYGRS
101 GETLGVLPVPH VGEIPVAYRK VLLRKNKNGK AGGHSYGADL KSFDLGDELG
151 TDPYEDFQEN WNTKHSSTGVT RELMRELNGG

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	M	Peptide
1 - 11	1250.5994	1249.5921	1249.6013	-7.31	0	-.MESLVPGFNEK.T
1 - 11	1266.5950	1265.5877	1265.5962	-6.70	0	-.MESLVPGFNEK.T + Oxidation (M)
12 - 24	1489.8739	1488.8666	1488.8777	-7.39	0	K.THVQLSLPVLQVR.D
30 - 43	1494.6944	1493.6871	1493.6998	-8.48	0	R.FGDSEVEVLSSEAR.Q
48 - 58	1206.5585	1205.5513	1205.5598	-7.09	0	K.DGTCGLVEVEK.G
59 - 72	1630.9094	1629.9022	1629.9130	-6.67	0	K.GVLPQLQPYVFIK.R
59 - 73	1787.0157	1786.0084	1786.0141	-3.20	1	K.GVLPQLQPYVFIKR.S
78 - 99	2406.2100	2405.2027	2405.2162	-5.58	0	R.TAPHGHVMVELVAELEGIQYGR.S
78 - 99	2422.2316	2421.2243	2421.2111	5.48	0	R.TAPHGHVMVELVAELEGIQYGR.S + Oxidation (M)
100 - 119	2093.1190	2092.1117	2092.1317	-9.56	0	R.SGETLGVLPVPHVGEIPVAYR.K
100 - 120	2221.2261	2220.2188	2220.2267	-3.53	1	R.SGETLGVLPVPHVGEIPVAYRK.V
130 - 141	1132.5284	1131.5211	1131.5309	-8.62	0	K.GAGGHSYGADLK.S
142 - 164	2720.1802	2719.1729	2719.1562	6.15	0	K.SFDLGLDELGTDPYEDFQENWNTK.H
165 - 175	1272.5850	1271.5777	1271.6404	-49.3	1	K.HSSGVTRLMR.E
165 - 175	1288.5582	1287.5509	1287.6354	-65.6	1	K.HSSGVTRLMR.E + Oxidation (M)

图 17 Nsp1 蛋白的质谱鉴定

Fig.17 Identification of nsp1 protein by mass spectrometry

3 讨论

新冠疫情的全球蔓延给世界公共卫生安全带来了巨大挑战,尽快阐明 SARS-CoV-2 的传播机制并开发有效的抗病毒药物是遏制疫情蔓延的关键所在^[14]。结构分析显示,Nsp1 的 C 末端结构域阻碍 mRNA 在核糖体上的进入通道(mRNA entry tunnel),加速宿主细胞 mRNA 降解,导致宿主蛋白合成减少并促进病毒存活^[11,15]。Nsp1 还能有效阻断 RIG-I 依赖的先天免疫反应,抑制宿主细胞对病毒

的清除^[11,16]。由于 Nsp1 参与病毒生命周期的多个阶段,因此它是一个潜在的抗病毒靶点和重要的毒力因子。

生物信息学分析对于目的蛋白的表达具有很强的指导意义,*E.coli* 表达系统以其安全性好、操作简便、表达量高、容易放大培养等优点,被广泛用于基因工程药物生产^[17]。此外,原核表达对于快速制备抗体和体外功能研究具有较大优势。目的蛋白的理化性质、分子量、氨基酸组成、等电点、亲疏水性、序

列保守程度、GC含量、密码子组成等对于原核表达都具有重要影响。通常情况下,蛋白分子量超过100 kD或低于5 kD时均难以表达。分子量太小的易被降解,通过加入融合标签 GST、Trx、MBP或者较大的促融合蛋白标签有助于使目的蛋白正确折叠,并以可溶形式表达;对于密码子进行优化,也可提高目的蛋白的翻译效率和表达量。本文生物信息学分析显示,Nsp1蛋白分子量19.78 kDa,等电点5.36,性质较稳定,在哺乳动物中半衰期约30 h,在*E. coli*内大于10 h,无信号肽和跨膜螺旋,亲水性较强,不属于分泌蛋白。无规则卷曲在Nsp1蛋白中占比最高(45.00%),其次是 α 螺旋(25.56%)和延伸链(20.56%),以上结果为进一步揭示该蛋白的结构及功能提供了数据基础。

糖基化和磷酸化修饰是生物体内最重要的翻译后修饰方式,糖基化对蛋白的折叠、免疫原性、稳定性及定位具有重要影响。磷酸化修饰可改变蛋白的生物活性,从而调控细胞内信号转导。蛋白的磷酸化修饰与病毒粒子增殖、组装和复制密切相关,在调控病毒与宿主的代谢中起重要作用。本研究分析显示,Nsp1有3个O-糖基化位点及12个磷酸化位点,然而仍需进一步实验验证。系统鉴定病毒蛋白的磷酸化位点并揭示其调控机制将有助于抗病毒药物的研发和病毒感染机制的阐明。

SARS-CoV的Nsp1蛋白通过“双管齐下”的策略抑制宿主基因表达,它首先结合核糖体40S亚基,在翻译起始的不同阶段使经典的mRNA翻译(canonical mRNA translation)停滞^[18-19]。其次,Nsp1与核糖体结合导致宿主mRNA的酶切和降解。然而,Nsp1与病毒mRNA的5' UTR保守区相互作用进而避免自身蛋白翻译停滞的机制还有待进一步研究^[20]。本文多序列比对显示,SARS-CoV-2与SARS-CoV的Nsp1序列一致性为84.4%(图12),提示二者具有相似的理化性质和生物学功能。研究发现SARS-CoV Nsp1的C末端残基K164和H165在 β 冠状病毒中高度保守,二者对Nsp1与40S亚基的相互作用至关重要,K164A和H165A的突变导致Nsp1无法结合40S并抑制翻译起始^[19]。近期,Schubert等^[21]和Thoms等^[11]的电镜结构分析显示,SARS-CoV-2 Nsp1的C末端区域折叠成两个螺旋,插入mRNA入口通道,阻碍mRNA进入进而抑制宿主蛋白翻译。第1个螺旋(P153-N160)与40S核糖体蛋白uS3和uS5形成疏水作用,第2个螺旋(S166-N178)与eS30和18S rRNA的h18相互作用。保守的KH二肽(K164和

H165)与h18形成关键相互作用,这种相互作用是基于H165与U607和U630之间的碱基堆积力,以及K164与G625和U630的磷酸骨架之间的静电相互作用。此外,SARS-CoV-2 Nsp1的K164和H165残基对于核糖体结合和抑制宿主基因表达至关重要^[11, 22]。以上结论与本文多序列比对结果相吻合。

结构域分析显示,Nsp1的N端(T12-G133)为冠状病毒Nsp1球状结构域,C端(E148-G179)为 β 冠状病毒Nsp1 C末端结构域(图10)。Zhao K等^[23]对SARS-CoV-2 Nsp1 N末端(K11-K125)的结构与功能研究发现,单独的N末端不能与核糖体共定位并抑制宿主蛋白翻译,单独的C端可与核糖体共定位但其抑制蛋白质翻译能力明显减弱。有趣的是,将Nsp1的C端与增强型绿色荧光蛋白(EGFP)或其他蛋白融合以取代其N端,可将蛋白翻译抑制能力恢复到与全长Nsp1相当水平。这就说明Nsp1的N末端能够稳定C末端同核糖体的结合,并作为非特异性屏障阻断mRNA通道从而抑制宿主细胞翻译,因此SARS-CoV-2 Nsp1主要通过其C端抑制宿主翻译,但N末端是Nsp1发挥功能所必需的。K141、S142和F143残基靠近SARS-CoV-2 Nsp1的C末端,参与40S核糖体亚基的结合^[11, 21, 24]。本研究发现,K141-F143序列保守性不高,除SARS-CoV-2之外,其它序列中有14条对应残基均为K141-Y143。此外,Benedetti等^[25]鉴定到一个K141-F143缺失的SARS-CoV-2基因组,这种缺失变异可能改变Nsp1蛋白的结构,抑制Nsp1同40S核糖体结合,影响病毒和宿主基因表达调控的活性^[26]。以上结果表明,SARS-CoV-2的Nsp1基因正在经历一个进化过程,这可能有助于病毒更好地适应人类宿主。

最后,本文构建了SARS-CoV-2 Nsp1蛋白的表达载体pET-22b-Nsp1并完成原核表达,初步建立了Nsp1蛋白表达条件。当培养至OD₆₀₀为0.4~0.6时加入IPTG(0.5 mmol/L)诱导,先37℃、220 r/min培养2.5 h,接着25℃、160 r/min继续培养9 h。结果显示,Nsp1蛋白出现过表达且主要在细菌裂解液离心之后的沉淀中表达(图16),经质谱鉴定(图17),沉淀中主要表达的确实是目标蛋白Nsp1。因此大规模纯化Nsp1用于疫苗研发或结构分析应考虑优先从沉淀中进行。另一方面,从功能研究的角度考虑,可采取改变培养温度16℃~30℃、降低IPTG浓度(0.01~0.1) mmol/L并延长诱导时间使蛋白更多的在上清表达,因为上清表达的蛋白要比

经包涵体纯化的蛋白活性好。需要指出的是,本文还预测到 Nsp1 存在糖基化和磷酸化修饰,鉴于原核表达系统没有翻译后修饰功能,因此功能研究实验应考虑从真核系统去表达该蛋白。

综上,本文对 SARS-CoV-2 的 Nsp1 进行了较为全面的生信分析、表达载体构建及原核表达,研究结果将有助于深入了解 Nsp1 的生物学功能和设计靶向药。基于本研究结果,将来可针对以下方向进行重点攻关:1)Nsp1 蛋白有效抑制宿主细胞翻译但不影响自身蛋白表达的分子机制的阐明;2)针对 Nsp1 蛋白的靶向药物设计;3)针对 Nsp1 蛋白所涉及的关键信号通路开发特异性的抑制剂;4)针对 Nsp1 蛋白或基因的高灵敏度病毒检测方法的开发;5)系统鉴定 Nsp1 蛋白的翻译后修饰位点并阐明其调控机理。

4 结 论

Nsp1 性质稳定,亲水性较强,有多个翻译后修饰位点,原核表达后发现其主要在细菌裂解液离心后的沉淀中表达。因此,其功能研究可能采用真核系统的可溶性表达更适宜。此外,该蛋白序列保守性高,且主要定位于宿主细胞的细胞质与细胞核中,提示其功能重要。本研究为针对该蛋白的纯化、结晶、结构分析、抗体制备和体外功能研究提供了重要的较好的数据基础,有助于加快针对 Nsp1 的靶向药物研发。

(感谢中国科学院生物物理研究所质谱平台为本文 Nsp1 蛋白鉴定所提供的大力支持。)

利益冲突:无

引用本文格式:刘玲,李璟,范蕾,等. 新冠病毒 Nsp1 蛋白结构与功能的生物信息学分析及原核表达[J]. 中国人兽共患病学报, 2022, 38(7): 566-576. DOI:10.3969/j.issn.1002-2694.2022.00.074

参考文献:

- [1] 吴雅玲, 陈骥, 王德民. SARS-CoV-2 和新型冠状病毒肺炎:发病机制与药物开发研究进展[J]. 福建师范大学学报(自然科学版), 2020, 36(5): 102-116. DOI: 10.12046/j.issn.1000-5277.2020.05.014
- [2] Peng Y, Du N, Lei Y, et al. Structures of the SARS-CoV-2 nucleocapsid and their perspectives for drug design[J]. EMBO J, 2020, 39(20): e105938. DOI: 10.15252/embj.2020105938
- [3] World Health Organization. Weekly operational update on COVID-19-28 September 2021[EB/OL]. (2021-09-28)[2021-10-08]. <https://www.who.int/publications/m/item/weekly-operational-update-on-covid-19-28-september-2021>
- [4] Zhou P, Yang XL, Wang XG, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin[J]. Nature, 2020, 579(7798): 270-273. DOI: 10.1038/s41586-020-2012-7
- [5] Coronaviridae study group of the international committee on taxonomy of V. The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2[J]. Nat Microbiol, 2020, 5(4): 536-544. DOI: 10.1038/s41564-020-0695-z
- [6] Simeoni M, Cavinato T, Rodriguez D, et al. I(nsp1)ecting SARS-CoV-2-ribosome interactions[J]. Commun Biol, 2021, 4(1): 715. DOI: 10.1038/s42003-021-02265-0
- [7] Arya R, Kumari S, Pandey B, et al. Structural insights into SARS-CoV-2 proteins[J]. J Mol Biol, 2021, 433(2): 166725. DOI: 10.1016/j.jmb.2020.11.024
- [8] Helmy YA, Fawzy M, Elasad A, et al. The covid-19 pandemic: A comprehensive review of taxonomy, genetics, epidemiology, diagnosis, treatment, and control[J]. J Clin Med, 2020, 9(4): 1225. DOI: 10.3390/jcm9041225
- [9] Gustin KM, Guan BJ, Dziduszko A, et al. Bovine coronavirus nonstructural protein 1 (p28) is an RNA binding protein that binds terminal genomic cis-replication elements[J]. J Virol, 2009, 83(12): 6087-6097. DOI: 10.1128/JVI.00160-09
- [10] Bai C, Zhong Q, Gao GF. Overview of SARS-CoV-2 genome-encoded proteins[J]. Sci China Life Sci, 2021: 1-15. DOI: 10.1007/s11427-021-1964-4
- [11] Thoms M, Buschauer R, Ameismeier M, et al. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2[J]. Science, 2020, 369(6508): 1249-1255. DOI: 10.1126/science.abc8665
- [12] 徐本锦, 刘玲. 大肠杆菌翻译起始因子 IF1 的生物信息学分析、表达载体构建及蛋白纯化[J]. 基因组学与应用生物学, 2020, 39(12): 5606-5616. DOI: 10.13417/j.gab.039.005606
- [13] Gordon DE, Jang GM, Bouhaddou M, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing[J]. Nature, 2020, 583(7816): 459-468. DOI: 10.1038/s41586-020-2286-9
- [14] 马睿忆, 董晓婧, 肖霞, 等. SARS-CoV-2 非结构蛋白 1 抑制 I 型干扰素应答反应[J]. 国际病毒学杂志, 2021, 28(2): 116-119. DOI: 10.3760/cma.j.issn.1673-4092.2021.02.007
- [15] Kamitani W, Narayanan K, Huang C, et al. Severe acute respiratory syndrome coronavirus nsp1 protein suppresses host gene expression by promoting host mRNA degradation[J]. Proc Natl Acad Sci U S A, 2006, 103(34): 12885-12890. DOI: 10.1073/pnas.0603144103
- [16] Hu Y, Li W, Gao T, et al. The severe acute respiratory syndrome coronavirus nucleocapsid inhibits type I interferon production by interfering with TRIM25-mediated RIG-I ubiquitination[J]. J Virol, 2017, 91(8): e02143-16. DOI: 10.1128/JVI.02143-16
- [17] 李生军, 张海霞, 冯若飞. 大肠杆菌表达病毒样颗粒的研究进展[J]. 中国免疫学杂志, 2021, 37(12): 1526-1532. DOI: 10.3969/j.issn.1000-484X.2021.12.024
- [18] Lokugamage KG, Narayanan K, Huang C, et al. Severe acute

- respiratory syndrome coronavirus protein nsp1 is a novel eukaryotic translation inhibitor that represses multiple steps of translation initiation[J]. *J Virol*, 2012, 86(24):13598-13608. DOI: 10.1128/JVI.01958-12
- [19] Kamitani W, Huang C, Narayanan K, et al. A two-pronged strategy to suppress host protein synthesis by SARS coronavirus Nsp1 protein[J]. *Nat Struct Mol Biol*, 2009, 16(11):1134-1140. DOI: 10.1038/nsmb.1680
- [20] Huang C, Lokugamage KG, Rozovics JM, et al. SARS coronavirus nsp1 protein induces template-dependent endonucleolytic cleavage of mRNAs; viral mRNAs are resistant to nsp1-induced RNA cleavage [J]. *PLoS Pathog*, 2011, 7(12): e1002433. DOI: 10.1371/journal.ppat.1002433
- [21] Schubert K, Karousis ED, Jomaa A, et al. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation[J]. *Nat Struct Mol Biol*, 2020, 27(10):959-966. DOI: 10.1038/s41594-020-0511-8
- [22] Shen Z, Zhang G, Yang Y, et al. Lysine 164 is critical for SARS-CoV-2 Nsp1 inhibition of host gene expression[J]. *J Gen Virol*, 2021, 102(1): jgv001513. DOI: 10.1099/jgv.0.001513
- [23] Zhao K, Ke Z, Hu H, et al. Structural basis and function of the N terminus of SARS-CoV-2 nonstructural protein 1[J]. *Microbiol Spectr*, 2021, 9(1):e0016921. DOI: 10.1128/Spectrum.00169-21
- [24] Banerjee AK, Blanco MR, Bruce EA, et al. SARS-CoV-2 disrupts splicing, translation, and protein trafficking to suppress host defenses[J]. *Cell*, 2020, 183(5):1325-1339. DOI: 10.1016/j.cell.2020.10.004
- [25] Benedetti F, Snyder GA, Giovanetti M, et al. Emerging of a SARS-CoV-2 viral strain with a deletion in nsp1[J]. *J Transl Med*, 2020, 18(1):329. DOI: 10.1186/s12967-020-02507-5
- [26] Nakagawa K, Makino S. Mechanisms of coronavirus Nsp1-mediated control of host and viral gene expression[J]. *Cells*, 2021, 10(2):300. DOI: 10.3390/cells10020300

收稿日期:2021-10-11 编辑:梁小洁

(上接第 565 页)

- [17] Kupferschmidt K, Vogel G. How bad is Omicron some clues are emerging[J]. *Science*, 2021, 374(6573): 1304-1305. DOI: 10.1126/science.acx9782
- [18] 孙丹, 杨金燕, 夏婷婷, 等. 新型冠状病毒 Omicron 变异株的流行病学特征及防控研究进展[J]. *中华医院感染学杂志*, 2022, 32(8):1266-1270. DOI:10.11816/cn.ni.2022-21376
- [19] Ozono S, Zhang Y, Ode H, et al. SARS-CoV-2 D614G spike mutation increases entry efficiency with enhanced ACE2-binding affinity[J]. *Nat Commun*, 2021, 12(1): 848. DOI: 10.1038/s41467-021-21118-2
- [20] Petersen E, Ntoumi F, Hui DS, et al. Emergence of new SARS-CoV-2 variant of concern Omicron (B.1.1.529) - highlights Africa's research capabilities, but exposes major knowledge gaps, inequities of vaccine distribution, inadequacies in global COVID-19 response and control efforts[J]. *Int J Infect Dis*, 2021, 114:268-272. DOI:10.1016/j.ijid.2021.11.040
- [21] Wang R, Chen J, Hozumi Y, et al. Emerging vaccine-breakthrough SARS-CoV-2 variants[J]. *ArXiv*, 2021, 62(2): 412-422. DOI:10.1021/acs.jcim.1c01451
- [22] Jenkins GM, Holmes EC. The extent of codon usage bias in human RNA viruses and its evolutionary origin[J]. *Virus Res*, 2003, 92(1):1-7. DOI:10.1016/s0168-1702(02)00309-x
- [23] Li Y, Yang X, Wang N, et al. GC usage of SARS-CoV-2 genes might adapt to the environment of human lung expressed genes [J]. *Mol Genet Genomics*, 2020, 295(6):1537-1546. DOI:10.1007/s00438-020-01719-0
- [24] V'Kovski P, Kratzel A, Steiner S, et al. Coronavirus biology and replication: implications for SARS-CoV-2[J]. *Nat Rev Microbiol*, 2021, 19(3): 155-170. DOI:10.1038/s41579-020-00468-6

收稿日期:2022-01-06 编辑:张智芳