

## • 医药生物技术 •

# 基于密码子优化策略的新型冠状病毒主蛋白酶在大肠杆菌中的表达条件优化与活性鉴定

陈云雨<sup>1</sup>, 付正豪<sup>1</sup>, 闫干干<sup>1</sup>, 林媛<sup>2</sup>, 刘晓平<sup>1</sup>

1 皖南医学院 药物筛选与评价研究所, 安徽 芜湖 241002

2 中国医学科学院-北京协和医学院 药物研究所 天然药物活性物质与功能国家重点实验室, 北京 100050

陈云雨, 付正豪, 闫干干, 等. 基于密码子优化策略的新型冠状病毒主蛋白酶在大肠杆菌中的表达条件优化与活性鉴定. 生物工程学报, 2021, 37(4): 1334-1345.

Chen YY, Fu ZH, Yan GG, et al. Optimization of expression conditions and determination the proteolytic activity of codon-optimized SARS-CoV-2 main protease in *Escherichia coli*. Chin J Biotech, 2021, 37(4): 1334-1345.

**摘要:** 新型冠状病毒主蛋白酶 (Main protease, Mpro) 在调控新冠病毒 RNA 复制中具有重要的生物学功能, 且 Mpro 在冠状病毒中的进化高度保守并不易突变, 已成为新型广谱抗冠状病毒药物开发的理想靶标之一。为了制备高纯度、高活性的 Mpro, 根据密码子偏爱性原则, 将优化的 Mpro 基因分别连接到 pET-21a 与 pET-28a 表达载体中构建重组质粒。将重组质粒转化到大肠杆菌 *Escherichia coli* Rosetta(DE3) 感受态细胞中, 分别进行原核表达条件优化, 所表达的重组蛋白质命名为 Mpro 与 Mpro-28。Mpro 与 Mpro-28 经 HisTrap<sup>TM</sup> 亲和层析法分离纯化后, 以荧光共振能量转移 (Fluorescence resonance energy transfer, FRET) 实验进行生物学活性鉴定。FRET 实验结果表明, 纯化的 Mpro 具有良好的水解活性,  $K_m$  值为 11.68  $\mu\text{mol/L}$ ,  $k_{\text{cat}}$  值为 0.037/s, 比活力不低于 25 000 U/mg, 约为 Mpro-28 的 25 倍, 说明天然的氨基端对 Mpro 的生物学功能是必需的, 羧基端残留的多聚组氨酸标签对其水解活性影响较小。文中基于密码子优化策略, 成功地进行了新冠病毒 Mpro 在大肠杆菌中的表达条件优化与活性鉴定, 为靶向 Mpro 广谱抗冠状病毒药物高通量筛选模型的建立奠定了实验基础。

**关键词:** 新型冠状病毒, 主蛋白酶, 密码子优化, 原核表达, 荧光共振能量转移

**Received:** July 8, 2020; **Accepted:** August 25, 2020

**Supported by:** National Natural Science Foundation of China (Nos. 81703546, 81773784), Natural Science Foundation of Anhui Province, China (No. 1808085QH265), Jilin Scientific and Technological Development Program (No. 20160520045JH), University Natural Science Research Project of Anhui Province, China (No. KJ2019ZD30), Key Technologies Research and Development Program of Anhui Province, China (No. 202004a07020041), Beijing Nova Program (No. Z181100006218075).

**Corresponding authors:** Yuan Lin. Tel: +86-10-63017906; E-mail: linyuan@imm.ac.cn

Xiaoping Liu. Tel: +86-553-3932601; E-mail: liuxiaoping@wnmc.edu.cn

国家自然科学基金 (Nos. 81703546, 81773784), 安徽省自然科学基金 (No. 1808085QH265), 吉林省科技发展计划项目 (No. 20160520045JH), 安徽省高校自然科学研究重大项目 (No. KJ2019ZD30), 安徽省重点研究与开发计划项目 (No. 202004a07020041), 北京市科技新星计划项目 (No. Z181100006218075) 资助。

网络出版时间: 2020-09-04

网络出版地址: <https://kns.cnki.net/kcms/detail/11.1998.Q.20200903.1710.001.html>

# Optimization of expression conditions and determination the proteolytic activity of codon-optimized SARS-CoV-2 main protease in *Escherichia coli*

Yunyu Chen<sup>1</sup>, Zhenghao Fu<sup>1</sup>, Gangan Yan<sup>1</sup>, Yuan Lin<sup>2</sup>, and Xiaoping Liu<sup>1</sup>

<sup>1</sup> Institute for Drug Screening and Evaluation, Wannan Medical College, Wuhu 241002, Anhui, China

<sup>2</sup> State Key Laboratory of Bioactive Substances and Function of Natural Medicine, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China

**Abstract:** The main protease (Mpro) of SARS-CoV-2 is a highly conserved and mutation-resistant coronaviral enzyme, which plays a pivotal role in viral replication, making it an ideal target for the development of novel broad-spectrum anti-coronaviral drugs. In this study, a codon-optimized *Mpro* gene was cloned into pET-21a and pET-28a expression vectors. The recombinant plasmids were transformed into *E. coli* Rosetta(DE3) competent cells and the expression conditions were optimized. The highly expressed recombinant proteins, Mpro and Mpro-28, were purified by HisTrap<sup>TM</sup> chelating column and its proteolytic activity was determined by a fluorescence resonance energy transfer (FRET) assay. The FRET assay showed that Mpro exhibits a desirable proteolytic activity (25 000 U/mg), with  $K_m$  and  $k_{cat}$  values of 11.68  $\mu\text{mol/L}$  and 0.037/s, respectively. The specific activity of Mpro is 25 times that of Mpro-28, a fusion protein carrying a polyhistidine tag at the N and C termini, indicating additional residues at the N terminus of Mpro, but not at the C terminus, are detrimental to its proteolytic activity. The preparation of active SARS-CoV-2 Mpro through codon-optimization strategy might facilitate the development of the rapid screening assays for the discovery of broad-spectrum anti-coronaviral drugs targeting Mpro.

**Keywords:** severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), main protease, codons optimization, prokaryotic expression, fluorescence resonance energy transfer (FRET)

自 2019 年 12 月以来,新型冠状病毒肺炎 (Coronavirus disease 2019, COVID-19) 席卷全球。截至 2020 年 7 月 31 日,全球 COVID-19 确诊病例超 1 779 万例,罹难病例逾 68.3 万例。COVID-19 是近百年来人类遭遇的影响范围最广的全球性大流行病,因其具有较高的传染性和死亡率,已被世界卫生组织宣布为国际突发公共卫生事件,给人类生命安全和健康带来了重大威胁<sup>[1]</sup>。然而,目前尚缺乏安全有效的 COVID-19 疫苗与治疗药物。

新型冠状病毒 (Severe acute respiratory syndrome coronavirus 2, SARS-CoV-2) 为单股正链 RNA 病毒,由 29 891 个碱基对组成,共编码 9 860 个氨基酸。SARS-CoV-2 与 SARS-CoV 同属  $\beta$ -冠状病毒,核苷酸序列同源性达 82%<sup>[2-4]</sup>。从 SARS-CoV-2 多顺反子基因组的 5'端开始,依次编码可被加工为多种非结构蛋白质 (Non-structural proteins, NSP) 的多聚蛋白质体 (Polyproteins)

ppla 和 pplab 以及刺突蛋白 (Spike protein, S)、包膜蛋白 (Envelope protein, E)、膜蛋白 (Membrane protein, M) 和核衣壳蛋白 (Nucleocapsid protein, N) 等结构蛋白质 (Structural proteins)。研究表明, SARS-CoV-2 与 SARS-CoV 相同,均通过冠状病毒的 S 蛋白与血管紧张素转化酶 2 (Angiotensin converting enzyme 2, ACE2) 受体结合,感染支气管上皮细胞和 II 型肺泡上皮细胞,引发人呼吸道感染<sup>[5]</sup>。在 SARS-CoV-2 侵入宿主细胞后,释放遗传物质 RNA,借助宿主细胞核糖体翻译为多聚蛋白质体,在新冠病毒主蛋白酶 (Main protease, Mpro) 的水解作用下,多聚蛋白质体被切割成新冠病毒 RNA 复制所必需的功能分子 (如 RNA 聚合酶、RNA 解旋酶等),进而完成子代病毒的装配与释放<sup>[6]</sup>。

SARS-CoV-2 Mpro 是相对分子量为 34 kDa 的半胱氨酸蛋白酶,与 SARS-CoV Mpro 相似性

高达 96%。Mpro 的活性形式为二聚体结构,其活性中心是由第 41 位组氨酸 (His41) 和第 145 位半胱氨酸 (Cys145) 形成的二聚体催化结构<sup>[7-8]</sup>。鉴于进化高度保守的 Mpro 在调控冠状病毒 RNA 复制中具有重要的生物学功能,且人体缺乏与其同源的蛋白酶, Mpro 已成为新型广谱抗冠状病毒药物开发的理想靶标之一<sup>[6,9]</sup>。

本研究基于密码子优化策略,通过优化新冠病毒 Mpro 在大肠杆菌中的表达条件与活性鉴定,制备了高纯度、高活性的 Mpro,为靶向 Mpro 广谱抗冠状病毒药物高通量筛选模型的建立奠定了实验基础。

## 1 材料与方法

### 1.1 材料

#### 1.1.1 试剂

DNA 标准分子量、质粒提取试剂盒、蛋白质标准分子量、*Nde* I、*Xho* I、大肠杆菌 *Escherichia coli* Rosetta(DE3) 感受态细胞购自 TransGen 公司;蛋白胨、酵母粉、琼脂粉购自 Oxide 公司;7-甲氧基香豆素-4-乙酸 (7-methoxycoumarin-4-acetic acid, MCA)、氨苄西林、卡那霉素、异丙基硫代半乳糖苷 (Isopropyl  $\beta$ -D-thiogalactoside, IPTG) 购自 Aladdin 公司;免疫印迹膜、HisTrap<sup>TM</sup> 层析柱购自 GE 公司;BCA (Bicinchoninic acid) 蛋白质定量试剂盒购自 Thermo 公司;小鼠抗组氨酸 (Histidine, His) 标签单克隆抗体、辣根过氧化物酶 (Horseradish peroxidase, HRP) 标记的羊抗小鼠 IgG 购自中杉金桥公司;超敏 ECL 化学发光液购自 NCM Biotech 公司;96 孔板购自 Corning 公司;MCA 标记的多肽底物 [MCA-substrate: MCA-AVLQSGFR-Lys (Dnp)-Lys-NH<sub>2</sub>,  $\lambda_{ex}/\lambda_{em}$ : 320 nm/405 nm] 购自吉尔生化 (上海) 有限公司;pET-21a 与 pET-28a 表达载体由本室保存;其他生化试剂为国产分析纯试剂。

#### 1.1.2 仪器

超净工作台 (AIRTECH 公司);加热制冷型

金属浴 (Thermo 公司);高速冷冻离心机 (艾本德公司);琼脂糖水平电泳仪 (北京六一生物科技有限公司);蛋白质电泳仪 (WIX 公司);快速转印仪 (Bio-Rad 公司);细菌振荡培养箱 (上海知楚仪器有限公司);超声波细胞粉碎机 (SCIENTZ 公司);AKTA Pure (GE 公司);凝胶成像系统 (CLINX 公司);多功能酶标仪 (BioTek 公司)。

## 1.2 方法

### 1.2.1 Mpro 基因序列优化

在 NCBI 数据库中检索 SARS-CoV-2 Mpro 基因 (GenBank 登录号: YP\_009725301.1), 将其录入到 JAVA Condon Adaption Tool 软件 (JCat, www.jcat.de) 中, 选择表达宿主为 *Escherichia coli*, 优化的序列中避免 *Nde* I 和 *Xho* I 酶切位点, 然后对 Mpro 基因序列进行密码子适应指数 (Codon adaption index, CAI) 分析、GC 含量计算与基因序列优化。

### 1.2.2 Mpro 原核表达质粒的构建与鉴定

在优化的 Mpro 基因序列 5'端和 3'端分别加入 *Nde* I 与 *Xho* I 酶切位点, 由 Gene Universal 公司进行基因合成。将合成的 Mpro 基因分别连接到 pET-21a 和 pET-28a 表达载体中, 将其表达的重组蛋白质命名为 Mpro 和 Mpro-28。构建的上述重组质粒经双酶切法鉴定后, 再由 Gene Universal 公司进行测序, 以 SnapGene 软件进行序列比对分析。

### 1.2.3 Mpro 原核表达

将重组质粒 pET-21a-Mpro 以冷 CaCl<sub>2</sub> 法转化至 *E. coli* Rosetta(DE3) 感受态细胞中, 再以氨苄西林抗性筛选法进行阳性重组子筛选。随机挑取 6 个重组子接种至含 100  $\mu$ g/mL 氨苄西林抗性的 LB 液体培养基中, 37  $^{\circ}$ C 培养 7 h, 再加入 1 mmol/L IPTG, 30  $^{\circ}$ C 诱导 10 h, 进行 Mpro 原核表达, 再以 12% SDS-PAGE 进行分析。以转化 pET-21a 质粒的 *E. coli* Rosetta(DE3) 为对照组。

以卡那霉素抗性筛选法进行 pET-28a-Mpro 重组子的筛选, 同法进行 Mpro-28 原核表达。

### 1.2.4 Mpro 原核表达条件优化

Mpro 最佳诱导时间的确定: 以 2% 接种率, 将工程菌接种到含 100  $\mu\text{g}/\text{mL}$  氨苄西林抗性的 LB 液体培养基中, 37  $^{\circ}\text{C}$  培养 7 h, 加入 1 mmol/L IPTG, 30  $^{\circ}\text{C}$  诱导 Mpro 原核表达。在第 0、2、4、6、8、10、12 h 时间点分别收集 800  $\mu\text{L}$  菌液, 以 SDS-PAGE 鉴定 Mpro 在各诱导时间点的表达量。设置诱导温度为 25  $^{\circ}\text{C}$ 、20  $^{\circ}\text{C}$ , 重复上述操作。同法进行 Mpro-28 最佳诱导时间的确定。

Mpro 最佳诱导温度的确定: 将工程菌接种到 50 mL LB 液体培养基中, 37  $^{\circ}\text{C}$  培养 7 h, 加入 1 mmol/L IPTG, 设定诱导条件为 30  $^{\circ}\text{C}$  诱导 8 h, 收集菌体。同法设定诱导条件为 25  $^{\circ}\text{C}$  诱导 10 h 和 20  $^{\circ}\text{C}$  诱导 12 h。收集的菌体以 8 mL TBS (50 mmol/L Tris, 150 mmol/L NaCl, pH 8.0) 溶液重悬, 冰浴条件下, 以超声波法破碎。离心收集上清液和沉淀, 再以 SDS-PAGE 鉴定 Mpro 可溶表达量。同法进行 Mpro-28 最佳诱导温度的确定。

最佳 IPTG 诱导浓度的确定: 将工程菌接种到 5 mL LB 液体培养基中, 37  $^{\circ}\text{C}$  培养 7 h, 分别加入 0、0.2、0.4、0.6、0.8、1 mmol/L IPTG, 30  $^{\circ}\text{C}$  诱导 8 h。等量收集菌体, 以 SDS-PAGE 鉴定 IPTG 诱导浓度对 Mpro 表达量的影响。同法进行 Mpro-28 最佳 IPTG 诱导浓度的确定。

### 1.2.5 Mpro 分离纯化与鉴定

将工程菌接种到 1 L LB 液体培养基 (含 100  $\mu\text{g}/\text{mL}$  氨苄西林) 中, 37  $^{\circ}\text{C}$  培养 7 h, 加入 0.2 mmol/L IPTG, 30  $^{\circ}\text{C}$  诱导 8 h。菌体以超声波法破碎后, 上清液再以 25% 饱和硫酸铵溶液沉淀制备 Mpro 粗提液。按照参考文献[10]所述的 HisTrap<sup>TM</sup> 亲和层析法进行 Mpro 分离纯化。纯化的 Mpro 以 TBS 溶液透析和蛋白质浓缩超滤管浓缩, 再以 BCA 法定量。同法进行 Mpro-28 的分离纯化。

将 1、2、3、4  $\mu\text{g}$  Mpro 以半干法转膜, 免疫印迹膜分别以小鼠抗 His 标签单抗 (1:2 000) 和

HRP-羊抗小鼠 IgG (1:4 000) 室温孵育 1 h, 以 NCM ECL Ultra 溶液显影成像。

### 1.2.6 Mpro 生物学活性鉴定

MCA 标准曲线的绘制: 将 10  $\mu\text{mol}/\text{L}$  MCA 在 TBS 溶液中以 2 倍倍比法稀释 5 个浓度梯度, 以只含 TBS 溶液孔为阴性对照孔。上述 MCA 稀释液以 50  $\mu\text{L}/\text{孔}$  加入到 96 孔板中, 各孔中 MCA 总量分别为 0、31.25、62.5、125、250、500 pmol, 设置自动增益模式, 以多功能酶标仪 (BioTek) 检测相对荧光强度值 (Relative fluorescence units, RFU)。根据各孔 MCA 总量和  $\Delta\text{RFU}$  值 ( $\Delta\text{RFU}=\text{RFU}_{\text{MCA}}-\text{RFU}_0$ ) 拟合回归方程, 绘制 MCA 标准曲线。

Mpro 水解活性的鉴定: 将 2 mmol/L MCA-substrate 以 TBS 溶液稀释至 10  $\mu\text{mol}/\text{L}$ , 加入 Mpro (10  $\mu\text{mol}/\text{L}$ ) 使其终浓度分别为 0、0.125、0.25、0.5  $\mu\text{mol}/\text{L}$ 。上述反应液以 50  $\mu\text{L}/\text{孔}$  加入到 96 孔板中, 设置增益值为 58, 检测温度为 25  $^{\circ}\text{C}$ , 激发光为 320 nm, 发射光为 405 nm, 检测间隔为 1 s, 检测总时间为 3 min, 以多功能酶标仪检测 RFU 值。在 30 s 水解反应时间内, 第  $T_1$  和  $T_2$  时间点的 RFU 值分别记录为  $\text{RFU}_{T_1}$  和  $\text{RFU}_{T_2}$ , 计算  $\Delta\text{RFU}=\text{RFU}_{T_2}-\text{RFU}_{T_1}$ , 通过标准曲线计算在  $\Delta T=T_2-T_1$  时间内, Mpro 水解 MCA-substrate 产生的 MCA-AVLQ 总量 (pmol)。同法检测 Mpro-28 的水解活性。活力单位 (U) 定义: 在 25  $^{\circ}\text{C}$  和 pH 8.0 条件下的水解反应中, Mpro 每分钟水解 MCA-substrate 生成 1 pmol MCA-AVLQ (产物, P) 所需的酶量。Mpro 比活力计算公式如下:

$$\text{Mpro 比活力}(\text{U}/\text{mg})=\frac{\text{P}(\text{pmol})}{(\text{T}_2-\text{T}_1)/60\times\text{Mpro}(\text{mg})}$$

Mpro 二聚体结构的预测分析: 输入 Mpro 氨基酸序列, 利用 SWISS-MODEL 软件 (www.swissmodel.expasy.org) 对 Mpro 二聚体结构进行预测分析。

### 1.2.7 Mpro 米氏常数与催化常数值的测定

将 Mpro (10  $\mu\text{mol}/\text{L}$ ) 以 TBS 溶液稀释至

0.25  $\mu\text{mol/L}$ , 加入 MCA-substrate (2 mmol/L) 使其终浓度分别为 1.25、2.5、5、10、20、30、40、45  $\mu\text{mol/L}$ , 每组设置 3 组复孔。按照上述方法进行 RFU 值检测, 计算在不同 MCA-substrate 浓度条件下, 各组水解反应的初速度 ( $\Delta\text{RFU/s}$ )。以 GraphPad Prism 5.0 软件拟合米氏方程, 计算 Mpro 的米氏常数值 (Michaelis constant,  $K_m$ ) 与催化常数 (Catalytic number,  $k_{\text{cat}}$ )。

## 2 结果与分析

### 2.1 Mpro 基因的优化

应用 JCat 软件分析, Mpro 原始序列的 CAI 值为 0.32 (<0.8), GC 含量为 38%, 且含有较多的低翻译率密码子 (图 1A)。理论上 CAI 值低于 0.8 时, 不利于外源基因在大肠杆菌中的异源表达, 因此推测 Mpro 原始序列的理论表达量较低。利用大肠杆菌密码子偏爱性原则, 上述 Mpro 原始序列经 JCat 软件优化后, CAI 值提高到 0.95 (>0.8), GC 含量为 47% 且均匀分布。在优化的 Mpro 序列中, 大肠杆菌偏爱密码子比率有了极大的提高 (图 1B)。Mpro 基因序列的优化有利于提高其在大肠杆菌中的表达量。

### 2.2 Mpro 原核表达质粒的构建与鉴定

将优化的 Mpro 基因合成后, 以 Nde I 和

Xho I 为酶切位点, 连接到 pET-21a 表达载体中构建重组质粒 pET-21a-Mpro。在此种 Mpro 构建策略中, Mpro 的氨基端仅残留一个甲酰甲硫氨酸 (Formylmethionine, fMet), 羧基端仅融合多聚组氨酸标签 (图 2A)。重组质粒以双酶切法进行鉴定, 得到了与预期大小 (918 bp) 一致的 Mpro 基因片段 (图 2B)。上述序列经测序和 BLAST 比对分析显示, 其基因序列与预期优化序列完全一致, 氨基酸序列与天然 Mpro 一级结构完全相同, 表明成功地构建了重组质粒 pET-21a-Mpro。

### 2.3 Mpro 原核表达

6 株经氨苄西林抗性筛选的重组子经 IPTG 诱导后, 收集的菌体经 SDS-PAGE 分析表明, 与对照组相比, 诱导的重组子在 34 kDa 位置均有明显的重组蛋白质表达条带, 其与 Mpro 理论分子量基本一致, 表达量约为 15% (图 3), 这表明成功进行了 Mpro 原核表达。

### 2.4 Mpro 原核表达条件的优化

工程菌分别在 30  $^{\circ}\text{C}$ 、25  $^{\circ}\text{C}$ 、20  $^{\circ}\text{C}$  以 IPTG 诱导培养, 设定 2 h 为诱导时间间隔, 在各诱导时间点等量收集菌体。SDS-PAGE 结果表明, 在 30  $^{\circ}\text{C}$ 、25  $^{\circ}\text{C}$ 、20  $^{\circ}\text{C}$  诱导温度下, Mpro 最佳诱导时间分别为 8、10、12 h, 此时 Mpro 表达量约为 15%, 达到表达量峰值 (图 4A-C)。等量收集以

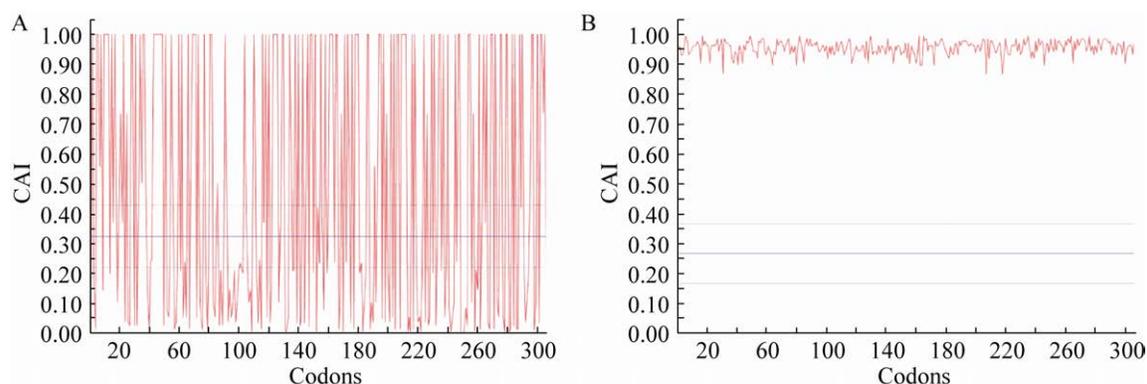


图 1 Mpro 基因序列的优化

Fig. 1 Optimization of Mpro gene sequence using JCat analysis. (A) Plot of the codons usage of the wild-type Mpro gene. (B) Plot of the codons usage of the optimized Mpro gene.

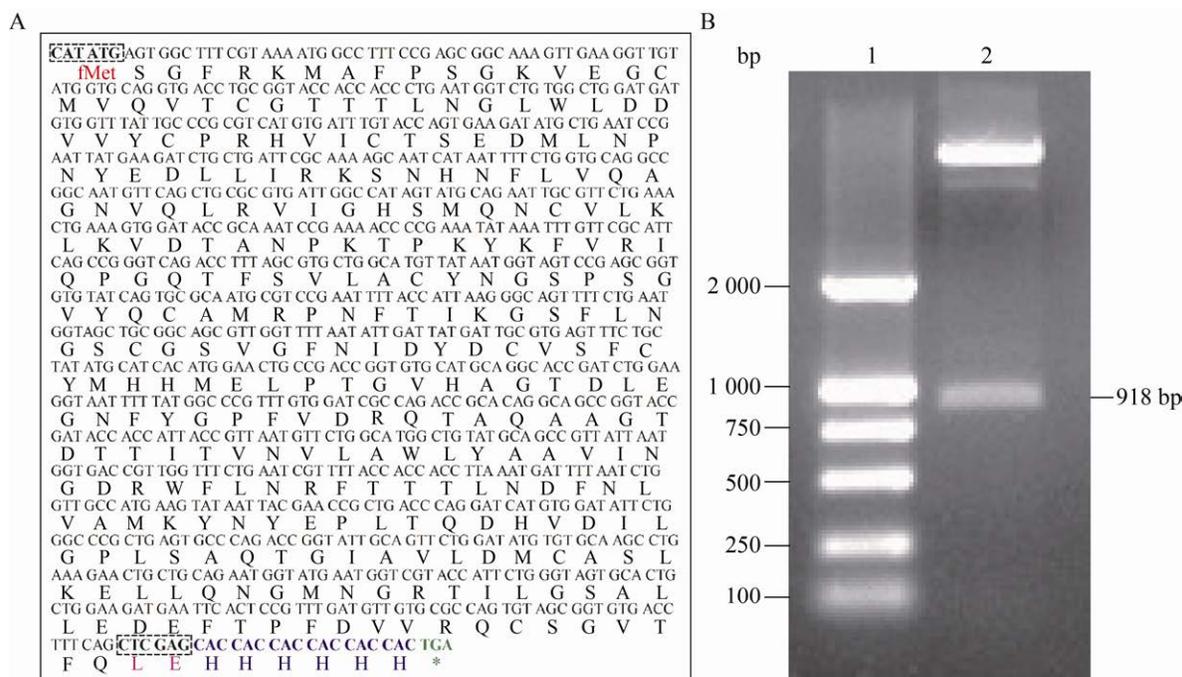


图 2 Mpro 原核表达质粒的构建与鉴定

Fig. 2 Construction and digestion map for pET-21a-Mpro plasmid. (A) The DNA sequence of the codon-optimized *Mpro* gene. The codon-optimized *Mpro* gene was shown and two restriction sites of *Nde* I and *Xho* I were labeled in dotted boxes. The primary structure of *Mpro* encoded by pET-21a-Mpro plasmid was also shown, and the first amino acid fMet and the C-terminal polyhistidine tag were highlighted in red and purple, respectively. (B) The double-restriction enzyme digestion map of pET-21a-Mpro plasmid. 1: DNA marker; 2: *Mpro* fragment (918 bp).

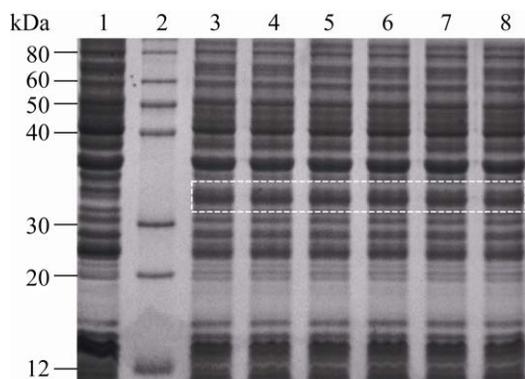


图 3 Mpro 原核表达

Fig. 3 The prokaryotic expression of *Mpro* using the constructed pET-21a-Mpro plasmid. 1: negative control; 2: protein marker; 3-8: total cell extracts after IPTG induction overnight.

上述最佳诱导时间培养的工程菌，再以超声波法破碎。SDS-PAGE 结果表明，在上述设定的诱导温度和时间下，*Mpro* 均大量存在于上清液中，呈可溶表达 (图 4D)。另外，IPTG 浓度对 *Mpro* 表达量

影响很小，0.2 mmol/L IPTG 即可充分诱导 *Mpro* 在大肠杆菌中的高效表达 (图 4E)。

综上所述，确定诱导温度 30 °C、诱导时间 8 h 和 0.2 mmol/L IPTG 作为 *Mpro* 在大肠杆菌中的最佳表达条件。

## 2.5 Mpro 分离纯化与鉴定

收集的菌体以超声波法破碎，裂解上清液首先以 25% 饱和硫酸铵溶液沉淀制备 *Mpro* 粗提液。由于 *Mpro* 的羧基端融合有多聚组氨酸标签，因此采用 HisTrap<sup>TM</sup> 亲和层析法纯化 (图 5A)。SDS-PAGE 分析结果表明，纯化的 *Mpro* 经考马斯亮蓝染色后，在预期分子量 34 kDa 位置呈单一条带，说明纯化的 *Mpro* 具有较高的纯度 (图 5B)。CLINX Image Analysis 软件分析显示，纯化的 *Mpro* 纯度为 96%。Western blotting 实验结果显示，在相同分子量位置仍有单一条带，且具有明显的

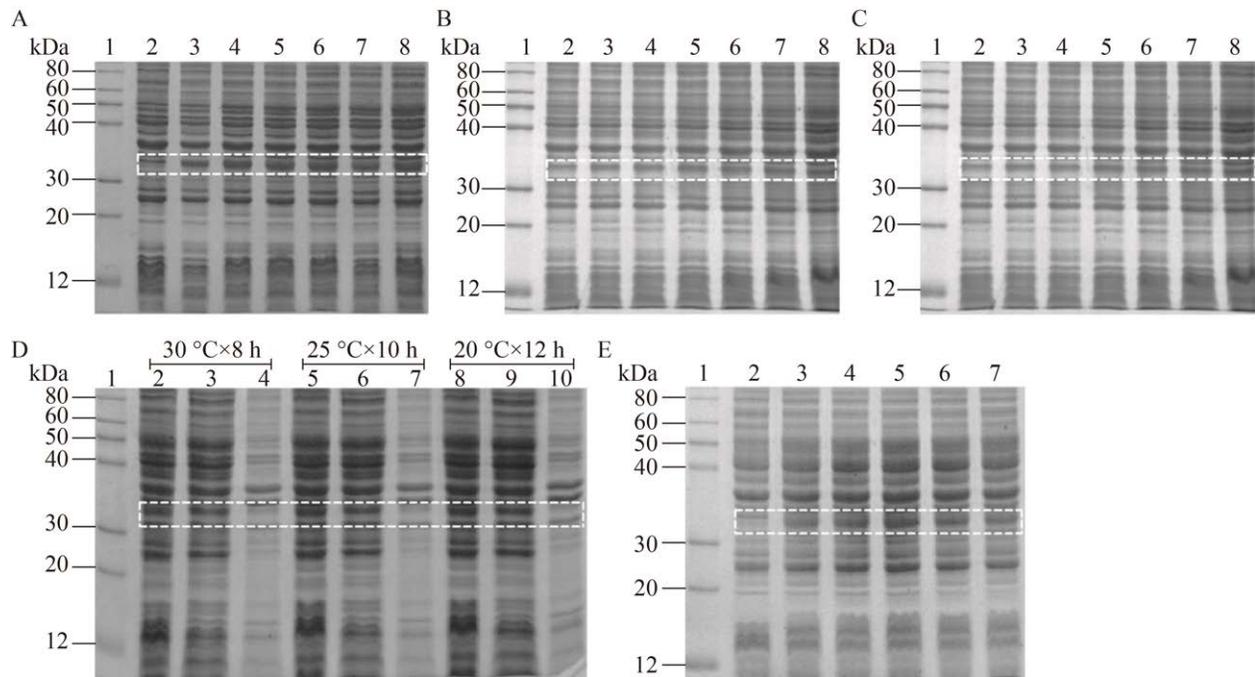


图4 Mpro 原核表达条件的优化

Fig. 4 The prokaryotic expression condition optimizations for Mpro in *E. coli* cells. (A–C) Expression of Mpro at 30 °C (A), 25 °C (B) and 20 °C (C) from 0 h to 12 h. 1: protein marker; 2: 0 h; 3: 2 h; 4: 4 h; 5: 6 h; 6: 8 h; 7: 10 h; 8: 12 h. (D) Analysis of the soluble expression of Mpro at the indicated conditions. 1: protein marker; 2, 5, 8: total cell extracts after IPTG induction; 3, 6, 9: supernatant of the cell lysate; 4, 7, 10: pellet of the cell lysate. (E) Analysis of the Mpro expression at the indicated IPTG concentrations. 1: protein marker; 2: 0 mmol/L IPTG; 3: 0.2 mmol/L IPTG; 4: 0.4 mmol/L IPTG; 5: 0.6 mmol/L IPTG; 6: 0.8 mmol/L IPTG; 7: 1 mmol/L IPTG.

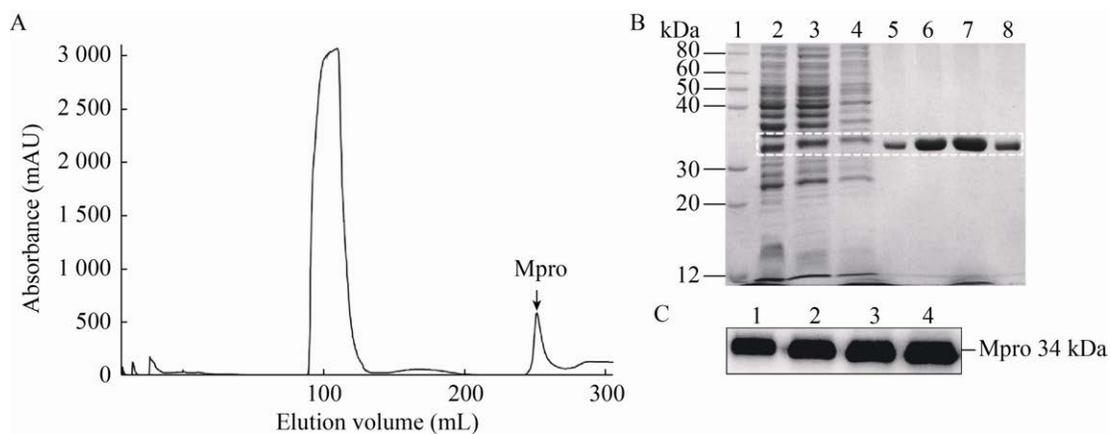


图5 Mpro 分离纯化与鉴定

Fig. 5 One step purification of Mpro by a HisTrap™ chelating column. (A) A typical purification profile of Mpro. A HisTrap™ chelating column was used to purify polyhistidine tagged Mpro from *E. coli* cell extracts. The targeted  $A_{280}$  peak of Mpro was labeled by a black arrow. (B) Analysis of the purified Mpro using SDS-PAGE assay. 1: protein marker; 2: total cell extracts after IPTG induction; 3: supernatant of the cell lysate; 4: precipitated Mpro in 25% saturated ammonium sulfate solution; 5–8: purified Mpro band (34 kDa). (C) Analysis of the purified Mpro using Western blotting assay. The imprinted Mpro in blotting membrane was probed by anti-His tag monoclonal antibody and HRP-secondary antibody, and then Mpro bands were visualized by the enhanced chemiluminescence (ECL) solution. 1: 1  $\mu$ g Mpro band; 2: 2  $\mu$ g Mpro band; 3: 3  $\mu$ g Mpro band; 4: 4  $\mu$ g Mpro band.

量效关系,证实了 Mpro 的正确表达 (图 5C)。纯化的 Mpro 经透析和浓缩后,其浓度为 0.35 mg/mL (10  $\mu\text{mol/L}$ )。

## 2.6 Mpro 生物学活性鉴定

Mpro 为半胱氨酸蛋白酶,具有典型的蛋白质水解活性。根据 Mpro 在新冠病毒生命周期中切割的天然底物序列,以 MCA-substrate 作为其模拟底物,基于荧光共振能量转移 (Fluorescence resonance energy transfer, FRET) 的原理鉴定 Mpro 的水解活性 (图 6A)。FRET 实验结果表明,

Mpro 对 MCA-substrate 具有良好的水解活性,比活力不低于 25 000 U/mg (图 6B-C)。通过 SWISS MODEL 软件对 Mpro 二聚体结构进行模拟分析,显示 Mpro 可形成良好的二聚体结构 (图 6D)。上述结果说明,纯化的 Mpro 具有良好的生物学活性,羧基端残留的多聚组氨酸标签对 Mpro 水解活性影响较小。

## 2.7 Mpro-28 原核表达、分离纯化与活性鉴定

将优化的 *Mpro* 基因以 *Nde* I 和 *Xho* I 为酶切位点连接到 pET-28a 表达载体中,成功地构建了原

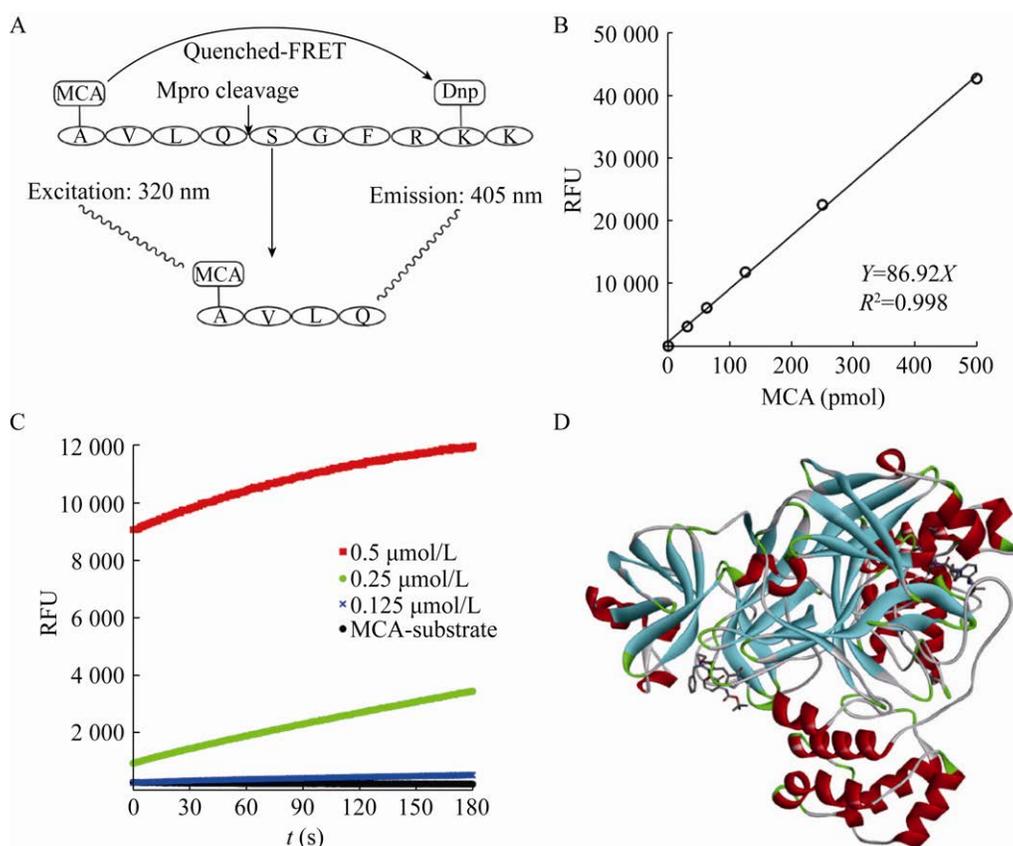


图 6 Mpro 生物学活性鉴定

Fig. 6 The proteolytic activity analysis of Mpro using a quenched-FRET assay. (A) Schematic illustration for a quenched-FRET assay principle. Mpro cleavage site in MCA-substrate was indicated by a black arrow. MCA: 7-methoxycoumarin-4-acetic acid; Dnp: 2,4-dinitrophenyl. (B) Plot MCA standard curve and the equation is fitting by a linear curve. (C) The proteolytic activity of Mpro in a quenched-FRET assay. The reaction was initiated by adding 10  $\mu\text{mol/L}$  MCA-substrate to a solution containing different concentrations of Mpro (0, 0.125, 0.25, 0.5  $\mu\text{mol/L}$ ). After that, the fluorescence signal (RFU) at 320 nm/405 nm was immediately measured every 1 s for 3 min by a microplate reader (BioTek). (D) Bioinformatics analysis of Mpro homodimer structure. The Mpro homodimer structure was predicted using SWISS MODEL analysis. One protomer of the homodimer was shown in light blue, the other one in red.

核表达质粒 pET-28a-Mpro (图 7A)。按照 Mpro 所述的原核表达条件优化方法, 确定了 Mpro-28 (37 kDa) 可溶表达的最佳条件为: 0.2 mmol/L IPTG、30 °C 诱导 8 h, 此时表达量约为 35%。由于 Mpro-28 的氨基端和羧基端均融合有多聚组氨酸标签, 因此以 HisTrap™ 亲和层析法成功地制备了高纯度的 Mpro-28 (图 7B)。FRET 实验结果表明, 相较于 Mpro, Mpro-28 对 MCA-substrate 的水解活性较弱, 比活力为 1 000 U/mg, 仅为 Mpro 比活力的 1/25

(图 7C), 说明氨基端残留的多聚组氨酸标签极大地减弱了 Mpro-28 的水解活性。

## 2.8 Mpro 米氏常数与催化常数的测定

0.25 μmol/L Mpro 分别与不同浓度的 MCA-substrate 进行水解反应时, 在酶促反应前 30 s 时间内, 酶促反应呈典型的一级反应, 其反应曲线斜率即为酶促反应的初速度 (图 8A)。分别测定各组底物浓度条件下酶促反应的初速度, 以 GraphPad Prism 5.0 软件拟合米氏方程 (Michaelis-Menten

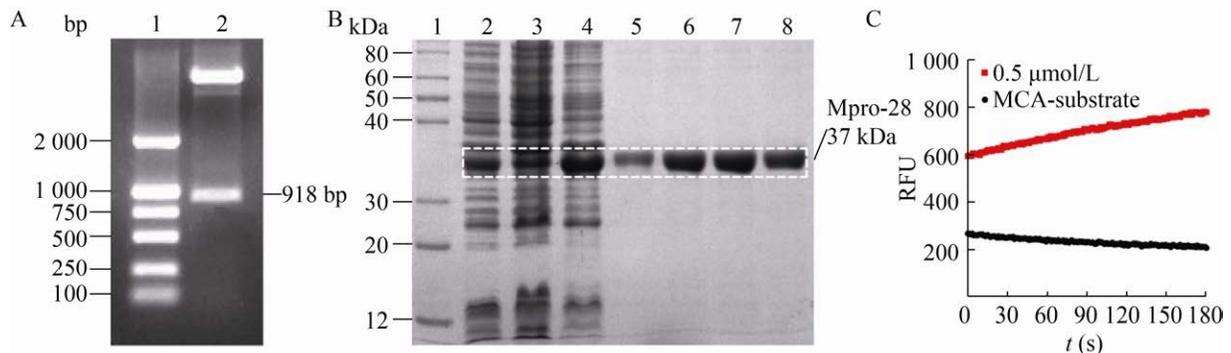


图 7 Mpro-28 原核表达、分离纯化与活性鉴定

Fig. 7 Expression, purification and proteolytic activity analysis of Mpro-28. (A) The double-restriction enzyme digestion map of pET-28a-Mpro plasmid. 1: DNA marker; 2: *Mpro* fragment (918 bp). (B) Analysis of the purified Mpro-28 using SDS-PAGE assay. 1: protein marker; 2: total cell extracts after IPTG induction; 3: supernatant of the cell lysate; 4: precipitated Mpro-28 in 25% saturated ammonium sulfate solution; 5–8: purified Mpro-28 band (37 kDa). (C) The proteolytic activity of Mpro-28 in a quenched-FRET assay. This FRET assay was performed as described above.

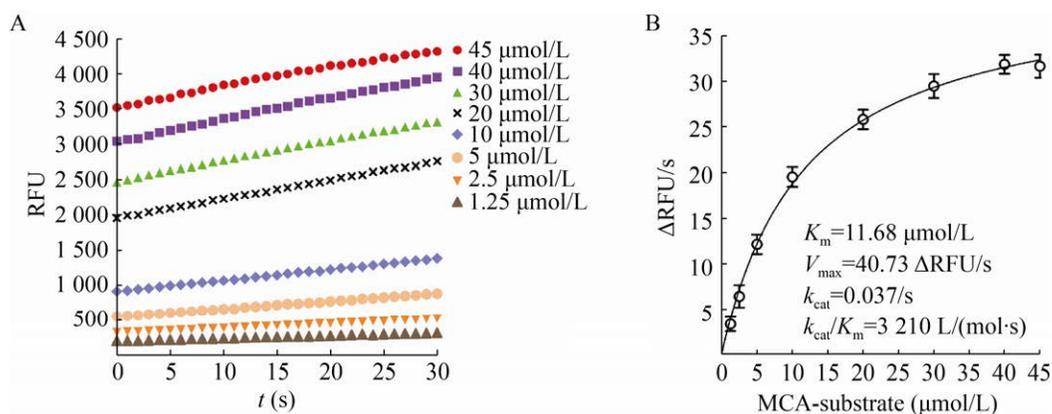


图 8 Mpro 米氏常数与催化常数的测定

Fig. 8 Determination of Mpro kinetic constant  $K_m$  and  $k_{cat}$  values using a quenched-FRET assay. (A) Analysis of the initial velocity of Mpro proteolytic kinetics. All initial rates were obtained by fitting the linear portion of the curves to a straight line. (B) Calculation of Mpro kinetic constant  $K_m$  and  $k_{cat}$  values. The Michaelis-Menten equation of Mpro proteolytic kinetics was plotted using GraphPad Prism 5.0 according to the initial rates, and then  $K_m$ ,  $V_{max}$  and  $k_{cat}$  values were determined.

equation), 计算 Mpro 的  $K_m$  值为 11.68  $\mu\text{mol/L}$ , 最大反应速度  $V_{\max}$  为 40.73  $\Delta\text{RFU/s}$ ,  $k_{\text{cat}}$  值为 0.037/s, 专一性常数值 ( $k_{\text{cat}}/K_m$ ) 为 3 210  $\text{L}/(\text{mol}\cdot\text{s})$  (图 8B)。

### 3 讨论

目前, COVID-19 疫情呈全球扩大趋势, 感染与罹难病例与日俱增, 已成为全球公共卫生关注的焦点与核心问题。COVID-19 疫情来势汹汹, 并且有与人类长期共存和新冠病毒 D614G 变异导致的高传染性风险<sup>[11]</sup>。鉴于现在已知的多种冠状病毒, 如 SARS、中东呼吸综合征 (Middle east respiratory syndrome, MERS) 都具有极强的传染性和致病性, 积极开发广谱抗冠状病毒药物和新型疫苗刻不容缓。Mpro 作为调控新冠病毒 RNA 复制的核心功能分子, 因其在冠状病毒中的进化高度保守且不易突变, 已成为新型广谱抗冠状病毒药物开发的理想靶标之一<sup>[6,9]</sup>。

大肠杆菌原核表达系统是高活性重组蛋白质制备的重要方法之一, 具有操作简便、成本低廉、表达量高等诸多优点<sup>[12]</sup>。但不同物种的基因在密码子使用上存在着明显的偏爱性, 密码子偏爱性对重组蛋白质的异源表达具有重要影响<sup>[13-14]</sup>。另外, 诱导时间、诱导温度、IPTG 浓度也是影响重组蛋白质异源表达的重要因素<sup>[15-16]</sup>。为了提高 Mpro 在大肠杆菌中的表达量, 在实验室条件下优化 Mpro 基因后再优化原核表达条件是极其必要和重要的。

本研究首先利用 JCat 软件对 Mpro 基因进行密码子优化, 使其基因的 CAI 值由 0.32 提高到 0.95, 大肠杆菌偏爱密码子比例和 GC 含量显著提高。优化的 Mpro 基因经 *Nde* I 和 *Xho* I 双酶切后, 连接到 pET-21a 表达载体中。这种构建策略可使 Mpro 的氨基端仅残留一个 fMet, 羧基端融合有多聚组氨酸标签。一方面, 可以最大程度地保持 Mpro 的天然构象和生物学活性; 另一方面,

融合的多聚组氨酸标签也方便了 Mpro 的分离纯化与生化检测。Mpro 原核表达条件优化实验结果表明, 应用上述构建策略, 以 0.2 mmol/L IPTG 在 30  $^{\circ}\text{C}$  诱导 8 h 时, Mpro 大量存在于上清液中, 基本以可溶形式表达。利用 HisTrap<sup>TM</sup> 亲和层析法即可快速制备高纯度的 Mpro, 其表达量约为 5 mg/L。利用相同的构建策略和原核表达条件优化方案, 我们也成功构建了重组质粒 pET-28a-Mpro, 制备了可溶表达的高纯度 Mpro-28。与 Mpro 构建策略稍有不同, 在 Mpro-28 的氨基端和羧基端均融合有多聚组氨酸标签。相较于 Mpro, Mpro-28 的表达量却有较大的提高, 约为 15 mg/L。

在新冠病毒的生命周期中, 多聚蛋白质体需要在 Mpro 水解作用下切割成新冠病毒 RNA 复制所必需的功能分子, 如 RNA 聚合酶等<sup>[6]</sup>。因此, Mpro 对特异底物的水解作用是其生物学活性评价的核心指标之一。荧光共振能量转移 (FRET) 是一种非辐射能量跃迁, 通过分子间的电偶极相互作用, 将供体激发态能量转移到受体激发态的过程<sup>[17]</sup>。该分析方法具有操作简单、检测灵敏、分析快速等优点。根据 Mpro 水解新冠病毒多聚蛋白质体的天然切割位点, 将 MCA 标记的多肽作为 Mpro 模拟底物, 淬灭基团为 Dnp, 荧光基团 MCA 和淬灭基团 Dnp 在多肽底物中由 Mpro 剪切序列隔开。当 MCA-substrate 被 Mpro 水解后, MCA-AVLQ 多肽片段因失去 FRET 效应而产生增强型荧光。FRET 实验结果表明, Mpro 具有良好的水解活性, 比活力可达 25 000 U/mg, 其  $K_m$  值为 11.68  $\mu\text{mol/L}$ ,  $k_{\text{cat}}$  值为 0.037/s, 专一性常数值 ( $k_{\text{cat}}/K_m$ ) 为 3 210  $\text{L}/(\text{mol}\cdot\text{s})$ 。但 Mpro-28 的水解活性较弱, 比活力仅为 Mpro 比活力的 1/25, 再次证实了天然的氨基端对 Mpro 生物学功能是必需的, 羧基端残留的多聚组氨酸标签对其水解活性影响较小<sup>[18-20]</sup>。上述建立的 FRET 方法不仅可以用于 Mpro 生物学活性的定量评价, 还可以用于 Mpro 小分子抑制剂的筛选与体外抑制活性评价,

因此所建立的 FRET 方法具有良好的实用性和推广性。

虽然基于密码子偏爱性策略, *Mpro* 基因经优化后在一定程度上提高了其原核表达量, 但翻译动力学、翻译后折叠及代谢水平等因素也是影响 *Mpro* 表达量的要素<sup>[14]</sup>。在后续的研究中, 我们将综合考虑密码子偏爱性、协调性、敏感性与基因序列对 *Mpro* 原核表达的协同影响, 进一步优化 *Mpro* 基因序列, 更大幅度地提高其表达量, 以便快速制备高产量、高纯度、高活性的 *Mpro*, 加速广谱抗新冠病毒药物的转化医学研究。

综上所述, 本研究基于密码子优化策略成功地进行了新冠病毒 *Mpro* 在大肠杆菌中的表达条件优化与生物学活性鉴定, 为靶向 *Mpro* 广谱抗冠状病毒药物高通量筛选模型的建立奠定了实验基础。

致谢: 感谢通用生物系统 (安徽) 有限公司武新春在新冠病毒 *Mpro* 密码子优化中给予的宝贵意见! 感谢上海科技大学免疫化学研究所杨海涛研究员、雷颖博士和 BioTek 上海代表处杨菁喆在 *Mpro* 原核表达与活性鉴定中给予的悉心指导和无私帮助!

## REFERENCES

- [1] Zheng J. SARS-CoV-2: an emerging coronavirus that causes a global threat. *Int J Biol Sci*, 2020, 16(10): 1678-1685.
- [2] 杨利敏, 田德雨, 刘文军. 新型冠状病毒疫苗研究策略分析. *生物工程学报*, 2020, 36(4): 593-604.  
Yang LM, Tian DY, Liu WJ. Strategies for vaccine development of COVID-19. *Chin J Biotech*, 2020, 36(4): 593-604 (in Chinese).
- [3] 刘艾林, 杜冠华. 基于靶点的抗新型冠状病毒病 COVID-19 药物发现. *药学学报*, 2020, 55(6): 1073-1080.  
Liu AL, Du GH. Drug discovery for COVID-19 treatment based on drug targets. *Acta Pharm Sin*, 2020, 55(6): 1073-1080 (in Chinese).
- [4] Chan JFW, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect*, 2020, 9(1): 221-236.
- [5] Xu XT, Chen P, Wang JF, et al. Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission. *Sci Chin Life Sci*, 2020, 63(3): 457-460.
- [6] Jin ZM, Du XY, Xu YC, et al. Structure of  $M^{pro}$  from SARS-CoV-2 and discovery of its inhibitors. *Nature*, 2020, 582(7811): 289-293.
- [7] Morse JS, Lalonde T, Xu SQ, et al. Learning from the past: possible urgent prevention and treatment options for severe acute respiratory infections caused by 2019-nCoV. *Chem Bio Chem*, 2020, 21(5): 730-738.
- [8] Zhang LL, Lin DZ, Sun XY, et al. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved  $\alpha$ -ketoamide inhibitors. *Science*, 2020, 368(6489): 409-412.
- [9] Goyal B, Goyal D. Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. *ACS Comb Sci*, 2020, 22(6): 297-305.
- [10] 陈云雨, 牛夏忆, 李妍, 等. 基于  $\beta$ -catenin/Lef1 相互作用为靶标的新型抗肿瘤药物高通量筛选模型的建立. *生物工程学报*, 2019, 35(4): 707-717.  
Chen YY, Niu XY, Li Y, et al. Development of an ELISA-based high throughput screening method for novel anticancer agents targeting  $\beta$ -catenin/Lef1 interaction. *Chin J Biotech*, 2019, 35(4): 707-717 (in Chinese).
- [11] Korber B, Fischer WM, Gnanakaran S, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell*, 2020, 182(4): 812-827.
- [12] Kaur J, Kumar A, Kaur J. Strategies for optimization of heterologous protein expression in *E. coli*: roadblocks and reinforcements. *Int J Biol Macromol*, 2018, 106: 803-822.
- [13] 蔡海莺, 李杨, 张辉, 等. 基因设计对重组蛋白表达的影响研究进展. *生物工程学报*, 2013, 29(9): 1201-1213.  
Cai HY, Li Y, Zhang H, et al. Effects of gene design on recombinant protein expression: a review. *Chin J Biotech*, 2013, 29(9): 1201-1213 (in Chinese).
- [14] 杨云彭, 马晓焉, 霍毅欣. 密码子优化策略在异源蛋白表达中的应用. *生物工程学报*, 2019, 35(12):

2227-2237.

Yang YP, Ma XY, Huo YX. Application of codon optimization strategy in heterologous protein expression. *Chin J Biotech*, 2019, 35(12): 2227-2237 (in Chinese).

- [15] 陈云雨, 牛夏忆, 李森, 等. 大肠杆菌 FtsZ 蛋白原核表达及多克隆抗体的制备与鉴定. *生物工程学报*, 2019, 35(6): 1117-1125.

Chen YY, Niu XY, Li M, et al. Bacterial expression, preparation and identification of polyclonal antibody against *Escherichia coli* FtsZ. *Chin J Biotech*, 2019, 35(6): 1117-1125 (in Chinese).

- [16] 付正豪, 王慷, 胡克, 等. 重组人 GST-TCF4  $\beta$ BD 在大肠杆菌中的表达条件优化与活性鉴定. *生命的化学*, 2020, 40(4): 614-621.

Fu ZH, Wang K, Hu K, et al. Expression condition optimization and bioactivity analysis of recombinant human GST-TCF4  $\beta$ BD in *Escherichia coli* cells.

*Chem Life*, 2020, 40(4): 614-621 (in Chinese).

- [17] Blanchard JE, Elowe NH, Huitema C, et al. High-throughput screening identifies inhibitors of the SARS coronavirus main proteinase. *Chem Biol*, 2004, 11(10): 1445-1453.

- [18] Yang HT, Yang MJ, Ding Y, et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc Natl Acad Sci USA*, 2003, 100(23): 13190-13195.

- [19] Xue XY, Yang HT, Shen W, et al. Production of authentic SARS-CoV M<sup>pro</sup> with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *J Mol Biol*, 2007, 366(3): 965-975.

- [20] Xia B, Kang X. Activation and maturation of SARS-CoV main protease. *Protein Cell*, 2011, 2(4): 282-290.

(本文责编 郝丽芳)