

## 论著·基础研究

## 蓝藻丝氨酸蛋白酶抑制剂 arthropin 的制备及其靶蛋白酶的筛选

许佳伟<sup>1</sup>, 周爱武<sup>2</sup>, 杨愈丰<sup>1</sup>

1. 遵义医科大学珠海校区生物工程系, 珠海 519040; 2. 上海交通大学基础医学院病理生理学系, 上海 200025

**[摘要]** **目的**·制备高纯度的蓝藻丝氨酸蛋白酶抑制剂 (serine protease inhibitor, serpin), 筛选其靶蛋白酶, 并探究其抑制活性。**方法**·通过氨基酸序列分析, 在人类口腔微生物组数据库 (eHOMD) 发现了一种钝顶节旋藻 (*Arthrospira platensis*) 的 serpin, 将其命名为 arthropin; 构建融合表达载体 pSUMO3-arthropin, 将其转入大肠埃希菌 BL21 (DE3) 诱导表达; 采用镍离子亲和层析、酶切、反向镍离子亲和层析、阴离子交换层析的 4 步层析纯化方案将 arthropin 重组蛋白分离纯化; 将 arthropin 重组蛋白分别与活化的凝血因子 IX (activated factor IX, FIXa)、FIXa、FXIa、活化的蛋白 C (activated protein C, APC)、激肽释放酶 1 (kallikrein 1, KLK1) 等 14 种丝氨酸蛋白酶共孵育, 再利用 SDS-PAGE 分析共价复合物形成情况, 利用反应动力学方法检测 arthropin 对 KLK1 的抑制速率。对 arthropin 重组蛋白的结晶条件进行筛选, 选取合适的晶体进行 X 射线衍射并采集数据。利用 AlphaFold Colab 预测 arthropin 亚稳态结构模型。**结果**·SDS-PAGE 分析发现, arthropin 融合蛋白在大肠埃希菌 BL21 (DE3) 中表达; 纯化后, 获得了高纯度的与理论相对分子质量 (45 800) 相近的 arthropin 重组蛋白。将 arthropin 重组蛋白与 14 种丝氨酸蛋白酶共孵育后发现, arthropin 能与 FIXa、APC、FIXa、FXIa、胰蛋白酶 (trypsin)、组织蛋白酶 G (cathepsin G)、KLK1、KLK7 和凝血酶 (thrombin) 9 种蛋白酶形成稳定的共价复合物。Arthropin 抑制 KLK1 的二级反应抑制速率常数为  $1.7 \times 10^3 \text{ L}/(\text{mol} \cdot \text{s})$ 。Arthropin 重组蛋白可以在 25% 聚乙二醇单甲醚 550 (PEG MME 550)、0.1 mol/L 吗啉乙磺酸 (MES, pH 6.5)、0.01 mol/L  $\text{ZnCl}_2$  条件下形成片状晶体, X 射线衍射分辨率为  $10 \text{ \AA}$  ( $1 \text{ \AA} = 1 \times 10^{-10} \text{ m}$ )。AlphaFold Colab 预测结果发现, arthropin 具有经典的抑制型 serpin 的结构特点。**结论**·成功获得高纯度的钝顶节旋藻丝氨酸蛋白酶抑制剂 arthropin, 其具有较广谱的丝氨酸蛋白酶抑制能力, 但抑制速率较低。

**[关键词]** 丝氨酸蛋白酶抑制剂; 钝顶节旋藻; 原核表达; 蛋白结晶**[DOI]** 10.3969/j.issn.1674-8115.2023.04.004 **[中图分类号]** Q71 **[文献标志码]** A

## Preparation and target protease identification of a cyanobacterial serine protease inhibitor, arthropin

XU Jiawei<sup>1</sup>, ZHOU Aiwu<sup>2</sup>, YANG Yufeng<sup>1</sup>

1. Department of Biological Engineering, Zhuhai Campus of Zunyi Medical University, Zhuhai 519040, China; 2. Department of Pathophysiology, College of Basic Medical Sciences, Shanghai Jiao Tong University, Shanghai 200025, China

**[Abstract]** **Objective**·To prepare a high-purity cyanobacterial serine protease inhibitor, screen its target proteases, and detect its inhibitory activity. **Methods**·A novel serine protease inhibitor from *Arthrospira platensis* was identified in the Expanded Human Oral Microbiome Database (eHOMD) by amino acid sequence alignment and named as arthropin. The fusion expression vector pSUMO3-arthropin was constructed and transferred into *Escherichia coli* (*E. coli*) BL21(DE3) system for fusion protein expression. The recombinant arthropin was purified by a four-step chromatographic purification approach of nickel affinity chromatography, enzymatic digestion, reverse nickel affinity chromatography, and anion exchange chromatography. In addition, the recombinant arthropin was co-incubated with 14 serine proteases such as activated factor IX (FIXa), FIXa, FXIa, activated protein C (APC) and kallikrein 1 (KLK1), respectively, and then analyzed by SDS-PAGE. The inhibitory rate of arthropin on KLK1 was assayed with kinetic methods. The crystallization conditions of the recombinant arthropin were screened preliminarily, and the suitable crystals were picked for X-ray diffraction to collect the data. Finally, a sub-stable structure model of arthropin was predicted with AlphaFold Colab. **Results**·SDS-PAGE showed that the fused arthropin was successfully expressed in the *E. coli* BL21(DE3) system, and following purification, the high-purity recombinant arthropin, the relative molecular mass of which was similar to the theoretical

[作者简介] 许佳伟 (1994—), 男, 硕士生; 电子信箱: 15820572242@163.com。

[通信作者] 杨愈丰, 电子信箱: yfyang@zmc.edu.cn。

[Corresponding Author] YANG Yufeng, E-mail: yfyang@zmc.edu.cn。



value (45 800), was obtained. The co-incubation analysis of recombinant arthropin with 14 serine proteases revealed that arthropin was able to form stable covalent complexes with 9 proteases, including FXa, APC, FIXa, FXIa, trypsin, cathepsin G, KLK1, KLK7 and thrombin. Arthropin inhibited KLK1 with a second-order association rate constant of  $1.7 \times 10^3 \text{ L}/(\text{mol} \cdot \text{s})$ . Moreover, the recombinant arthropin crystallised under the condition of 25% PEG MME 550, 0.1 mol/L MES (pH 6.5) and 0.01 mol/L  $\text{ZnCl}_2$ , and the crystals preliminarily diffracted to a resolution of 10 Å ( $1 \text{ Å} = 1 \times 10^{-10} \text{ m}$ ). The analysis of the structure predicted by AlphaFold Colab revealed that arthropin had the classical structural features of the inhibitory serpin. **Conclusion** Arthropin, a serpin from *Arthrospira platensis*, was successfully obtained with high purity and a broad-spectrum of serine protease inhibition, but at a low inhibitory rate.

**[Key words]** serine protease inhibitor (serpin); *Arthrospira platensis*; prokaryotic expression; protein crystallization

丝氨酸蛋白酶抑制剂 (serine protease inhibitor, serpin) 是规模最大、分布最广的一类蛋白酶抑制剂超家族, 广泛存在于动物、植物、细菌、病毒和古菌中<sup>[1-6]</sup>。Serpins 家族成员多数由 330~500 个氨基酸残基组成, 它们的一级序列同源性不高 (约 25%), 但二级和三级结构极其相似。典型的 serpin 主要由 1 个包含 8~9 个  $\alpha$  螺旋 (helix A~I)、3 个  $\beta$  片层 ( $\beta$ -sheet A~C) 的保守主体结构 and 1 个暴露在分子表面的反应中心环 (reactive center loop, RCL) 组成<sup>[6]</sup>。RCL 的长度和序列的多样性决定着 serpin 靶蛋白酶的特异性。

不同于其他蛋白酶抑制剂家族对蛋白酶非共价抑制作用的锁钥 (lock-and-key) 机制, serpin 对蛋白酶抑制是不可逆的。在天然状态下, 处于亚稳态构象的 serpin 利用 RCL 捕获靶蛋白酶并与其形成以共价酰胺键连接的复合物。随后, RCL 被水解 (被水解肽键往 N 端方向氨基酸依次命名为 P1、P2、P3……, 往 C 端方向依次命名为 P1'、P2'、P3'……) 并部分插入至 s3A ( $\beta$ -sheet A 第 3 条折叠股链) 与 s5A 之间作为新的 s4A, 这导致  $\beta$ -sheet A 扩展而主体结构松弛, 此时能级达到最低。与 s4A 相连的蛋白酶被拉扯至 serpin 底部的同时, 也因活性中心形变而失去活性<sup>[7]</sup>。Serpins 这种由紧凑 (S 型) 向松弛 (R 型) 状态转变的不可逆抑制作用被称为“自杀性抑制”。

在过去的数十年里, 人们对哺乳动物 serpin 的结构、功能和调控机制已经做了相当细致的研究<sup>[8]</sup>。这些 serpin 分为抑制型和非抑制型 2 类, 参与调节人体的很多生理过程, 例如血液凝集<sup>[9]</sup>、炎症反应<sup>[10]</sup>、组织重塑<sup>[11]</sup>、细胞凋亡<sup>[12]</sup>、激素转运<sup>[13-14]</sup>、肿瘤抑制<sup>[15]</sup>、分子伴侣<sup>[13]</sup>等, 而且这些 serpin 结构与生理功能的失衡, 以及多种疾病的发生和发展密切相关<sup>[16-17]</sup>。

相较于对哺乳动物 serpin 的广泛研究, 其他种属

serpin 的研究相对较少。前期已经有少量针对原核生物 serpin 的报道<sup>[4, 18-21]</sup>。例如, 2002 年 IRVING 等<sup>[4]</sup>首次在细菌和古菌的基因组中鉴定了 12 个 serpin 的基因序列。随后人们又在极端嗜热的 *Thermococcus kodakaraensis* 中发现了具有耐酸和热稳定性的 Tk-serpin, 它能够抑制环境中的蛋白酶从而保护细菌免受侵害<sup>[18]</sup>。最近人们也从人体牙周寄生细菌 *Tannerella forsythia* 和肠道寄生菌中发现相关 serpin, 它们可能通过抑制炎症相关丝氨酸蛋白酶参与人体的炎症反应<sup>[19-20]</sup>。而针对原核生物蓝藻来源 serpin 的结构、功能研究目前尚未有报道。

钝顶节旋藻 (*Arthrospira platensis*) 是一种富含天然活性物质的光合自养蓝藻; 近些年来, 钝顶节旋藻作为高营养食品和膳食添加剂被广泛研究, 极具经济价值<sup>[22-23]</sup>。在医药领域, 钝顶节旋藻富含高活性蛋白、脂多糖和亚麻酸, 其水提取物被发现具有抗病毒<sup>[24-25]</sup>、抗氧化<sup>[26]</sup>、抗疟疾<sup>[27]</sup>、促进伤口愈合<sup>[28]</sup>和抗癌<sup>[29-31]</sup>等活性; 其高表达的藻蓝蛋白能显著抑制癌细胞的生长且不良反应较小<sup>[32-34]</sup>。本研究利用基因组与氨基酸序列比对分析, 在钝顶节旋藻中发现了 1 个新的 serpin, 并命名为 arthropin; 通过构建 pSUMO3-arthropin 融合表达载体, 表达、纯化 arthropin 蛋白, 尝试了对其结晶条件的初步筛选, 同时检测了它对 14 种不同丝氨酸蛋白酶的抑制活性, 为后续针对其结构和功能的研究奠定了基础。

## 1 材料与方法

### 1.1 材料

**1.1.1 菌种与质粒** 大肠埃希菌 (*E. coli*) BL21 (DE3) 购自天根生化科技 (北京) 有限公司。pSUMO3 表达载体由本实验室改造并保存。原 pSUMO3 载体表达的融合蛋白携带小分子泛素样修饰

蛋白 (small ubiquitin-like modifier, SUMO), 其在提高融合蛋白的表达水平同时, 可被小泛素修饰特异性蛋白酶 2 (sentrin-specific protease 2, SENP2) 完整切割<sup>[35]</sup>。本实验室对原 pSUMO3 表达载体 [购自安升达 (天津) 生物科技有限公司] 进行优化, 删除了 SUMO 标签末端甘氨酸至限制性内切酶 *Bam*H I 酶切位点前的碱基序列, 这使得融合蛋白被酶切后仅残留 1 个丝氨酸, 减少对后续蛋白质结晶的影响。

**1.1.2 主要试剂与仪器** 氨苄青霉素及缓冲液配制试剂购自生工生物工程 (上海) 股份有限公司, 2×YT 培养基购自英国 Oxoid 公司, 异丙基-β-D-1-硫代半乳糖苷 (IPTG) 购自上海麦约尔生物技术有限公司, 十二烷基硫酸钠-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 预制胶购自美国 Invitrogen 公司, 14 种丝氨酸蛋白酶购自美国 Haematologic Technologies 公司, 激肽释放酶 1 (kallikrein 1, KLK1) 荧光底物 (H-D-Val-Leu-Arg-AMC) 购自国肽生物科技 (北京) 有限公司, HisTrap™ FF 预装柱 (His 标签蛋白纯化柱)、HiTrap™ Q HP (离子交换预装柱) 购自瑞典 GE Healthcare 公司, SENP2 由实验室纯化自用。Crystal Screen™ 蛋白结晶试剂盒购自美国 Hampton Research 公司。

UH-12 型高压细胞破碎仪购自永联生物科技 (上海) 有限公司, AKTA purifier 蛋白纯化仪购自瑞典 GE Healthcare 公司, Gel Doc XR 凝胶成像仪购自美国 Bio-Rad 公司, 96 孔坐滴式结晶板购自美国 Hampton Research 公司, Screenmaker 96+8 蛋白结晶配液工作站购自瑞士 Tecan 公司, LG-P52 倒置显微镜购自日本 Olympus 公司。

## 1.2 方法

**1.2.1 新型 serpin 的筛选与融合表达载体的构建** 利用经典的 serpin α1-抗胰蛋白酶 (α1-antitrypsin) 的氨基酸序列在人类口腔微生物组数据库 (eHOMD, <http://www.homd.org/>) 进行氨基酸序列比对; 选取序列同源性较高的疑似 serpin (score>100), 并与 α1-抗胰蛋白酶、人源性激肽释放酶结合蛋白 (kallistatin, KAL)、miropin 等 serpin 的氨基酸序列再次比对, 着重分析活性中心与保守区域的氨基酸序列; 使用 SWISS 软件对初步筛选获得的疑似 serpin 同源建模并分析结构。将新型 serpin 的基因序列交由安升达 (天津) 生物科技有限公司进行全基因技术合

成, 同时委托其完成密码子优化与融合表达载体的构建。

**1.2.2 融合蛋白的表达** 将融合表达载体 pSUMO3-arthropin 转化至大肠埃希菌 BL21 (DE3), 并涂布于含氨苄青霉素的 LB 固体平板。37 °C 倒置培养过夜后, 挑取单个菌落至含氨苄青霉素的 2×YT 液体培养基, 于 37 °C、220 r/min 摇床上扩菌, 直至菌液在 600 nm 处的光密度值达到 0.6。加入终浓度为 0.1 mmol/L 的 IPTG 诱导表达, 并将温度降至 20 °C 继续培养 12 h, 结束后 4 °C 离心收集菌体。另留取少量诱导前、后的菌液, 利用 SDS-PAGE 验证融合蛋白的表达情况。

**1.2.3 融合蛋白的酶切与纯化** 将收集的细菌用缓冲液 A [20 mmol/L Tris-HCl (pH=7.4)、0.5 mol/L NaCl、20 mmol/L 咪唑] 重悬, 并在 6×10<sup>4</sup> kPa 压强下破碎细胞。用 HisTrap™ FF 预装柱挂载离心后的细菌裂解上清液, 而后进行 20~300 mmol/L 咪唑的连续梯度洗脱 (缓冲液 B: 其他成分与缓冲液 A 相同, 仅咪唑浓度变化)。将收集的 280 nm 吸收峰的峰尖洗脱液透析去除咪唑, 再使用蛋白酶 SENP2 消化融合蛋白 12 h (融合蛋白: SENP2=100: 1), 使 His-SUMO3 标签游离。将消化液再次挂载于 HisTrap™ FF 预装柱, 收集流穿液中的 arthropin。最后, 将流穿液透析至缓冲液 C [20 mmol/L Tris-HCl (pH=8.0)、0.1 mol/L NaCl], 再挂载于 HiTrap™ Q HP 预装柱, 进行 100~1 000 mmol/L NaCl 的连续梯度洗脱, 收集 280 nm 吸收峰的峰尖洗脱液, 即可得到高纯度的 arthropin 蛋白。每步纯化均使用 SDS-PAGE 分析验证。

**1.2.4 Arthropin 的靶蛋白酶筛选与反应动力学检测** 选取本实验室常用的 14 种丝氨酸蛋白酶, 包括活化的凝血因子 IX (activated factor IX, FIXa)、活化的凝血因子 X (FXa)、活化的凝血因子 XI (FXIa)、活化的蛋白 C (activated protein C, APC)、组织型纤溶酶原激活剂 (tissue-type plasminogen activator, tPA)、人嗜中性粒细胞弹性蛋白酶 (human neutrophil elastase, HNE)、猪胰腺弹性蛋白酶 (pancreatic elastase, PE)、胰蛋白酶 (trypsin)、KLK1、KLK7、蛋白酶 3 (proteinase 3, PR3)、凝血酶 (thrombin, 又称 FIIa)、组织蛋白酶 B (cathepsin B, CaB)、组织蛋白酶 G (cathepsin G, CaG)。利用抑制型 serpin 能与蛋白酶形成共价复合物的特点, 将



纯化好的 arthropin (3  $\mu\text{g}$ ) 分别与 14 种蛋白酶 (0.4  $\mu\text{g}$ ) 混合并于 37  $^{\circ}\text{C}$  孵育 30 min, 利用 SDS-PAGE 分析可直观显示 arthropin 与 14 种蛋白酶的结合情况。

在 20  $\mu\text{L}$  反应体系中, 将 KLK1 (终浓度 0.1  $\mu\text{mol/L}$ ) 分别与不同浓度 arthropin (终浓度 0~3.2  $\mu\text{mol/L}$ ) 混合并于室温孵育 20 h, 再添加 180  $\mu\text{L}$  含有 KLK1 荧光底物 (终浓度 0.1 mmol/L) 的缓冲液 [PBS、0.1 mg/mL 牛血清白蛋白、0.1% 聚乙二醇 8000 (PEG-8000)], 检测 460 nm 处荧光信号强度。当荧光底物完整时, 其 C 端 7-氨基-4-甲基香豆素 (7-amino-4-methylcoumarin, AMC) 的荧光信号被 N 端  $\text{H}_2\text{N}$  基团吸收; 而当底物被 KLK1 水解, 释放的 AMC 在 380 nm 处被激发, 继而在 460 nm 处显示荧光信号, 根据反应物在 460 nm 处荧光信号强度的变化即可计算出 KLK1 的剩余活性。将 arthropin/KLK1 浓度比与体系中 KLK1 剩余活性进行线性拟合即可得到该反应的化学计量比 (stoichiometry inhibition, SI) [36]。

二级反应表观速率常数 (apparent second-order rate constant,  $k_{\text{app}}$ ) 的检测是在伪一级反应条件 (即 arthropin 浓度远大于 KLK1 浓度) 下进行的。将 0.1  $\mu\text{mol/L}$  KLK1 与一定浓度的 arthropin 混合并孵育不同时间, 再利用 0.1 mmol/L KLK1 荧光底物缓冲液检测反应体系中 KLK1 的剩余活性, 将剩余活性的负对数与孵育时间线性拟合, 其斜率即是一级反应速率常数 (first order rate constant,  $k_{\text{obs}}$ )。将不同 arthropin 浓度 (3~12  $\mu\text{mol/L}$ ) 反应的  $k_{\text{obs}}$  与 arthropin 浓度进行线性拟合, 其斜率即是  $k_{\text{app}}$ , 而二级反应速率常数 (second-order association rate constant,  $k_2$ ) =  $k_{\text{app}} \times \text{SI}$  [36]。

**1.2.5 Arthropin 的结晶条件筛选** 将 Crystal Screen<sup>TM</sup> 蛋白结晶试剂盒包含的 1 600 种结晶试剂分

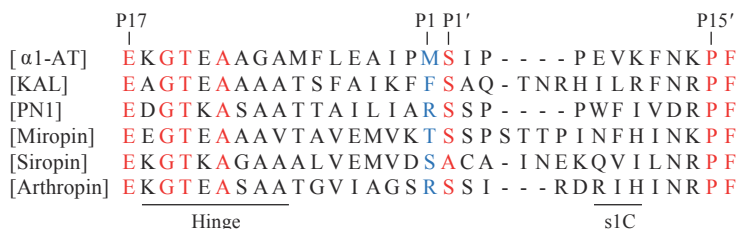
别以 65 nL/孔分装至 96 孔坐滴式结晶板上, 在 20  $^{\circ}\text{C}$  下, 将 200 nL 高纯度、性质均一的 arthropin 蛋白溶液 (13 mg/mL) 与 200 nL 结晶试剂混合并接种于 96 孔坐滴式结晶板, 密封。所有操作均由自动化分液机器人完成。针对初筛结果, 选取晶型较好的结晶条件进行优化, 即调整蛋白质浓度、沉淀剂浓度、体系 pH 值和结晶温度等, 以期获得晶型更好的蛋白晶体。另一方面, 用冷冻环 (cryoloop) 将初筛所得晶体捞取并放入冷冻保护剂 (在原结晶条件基础上加入 25% 甘油)。液氮中瞬间冷冻后, 于上海同步辐射光源 (Shanghai Synchrotron Radiation Facility, SSRF) 18U 进行 X 射线衍射并采集数据。

**1.2.6 Arthropin 三维结构的预测** 将 arthropin 的氨基酸序列导入至蛋白预测结构数据库 (AlphaFold Protein Structure Database, <https://alphafold.ebi.ac.uk>) 进行自动化的三维结构预测, 再对生成的 arthropin 结构模型进行分析。

## 2 结果

### 2.1 新型 serpin 的筛选与融合表达载体的构建

通过对氨基酸序列的比对与分析, 在钝顶节旋藻中发现了一个新的 serpin, 并把它命名为 arthropin。与经典 serpin 的 RCL 序列比对结果 (图 1) 显示, arthropin 具有典型的抑制型 serpin 反应中心环, 其 P9~P17 基本符合 ExGTEAAAAA ( $x=\text{E/K/R}$ ) 的保守序列规律, 而活性中心 P1-P1' 是抑制型 serpin 最常见的 R-S, 提示 arthropin 可能具有抑制丝氨酸蛋白酶的活性。Arthropin 基因序列插入优化后的 pSUMO3 质粒 (*Bam*H I 和 *Xho* I 酶切位点之间) 构建融合表达载体 pSUMO3-arthropin。



**Note:**  $\alpha$ 1-AT— $\alpha$ 1-antitrypsin (Uniprot ID: A0A024R6I7); KAL—kallistatin (Uniprot ID: P29622); PN1—protein nexin 1 (Uniprot ID: P07093); miropin (Uniprot ID: G8UQY8); siropin (UniProt ID: R6RSK2); arthropin (GenBank ID: AMW27298.1). s1C—strand 1 of  $\beta$ -sheet C.

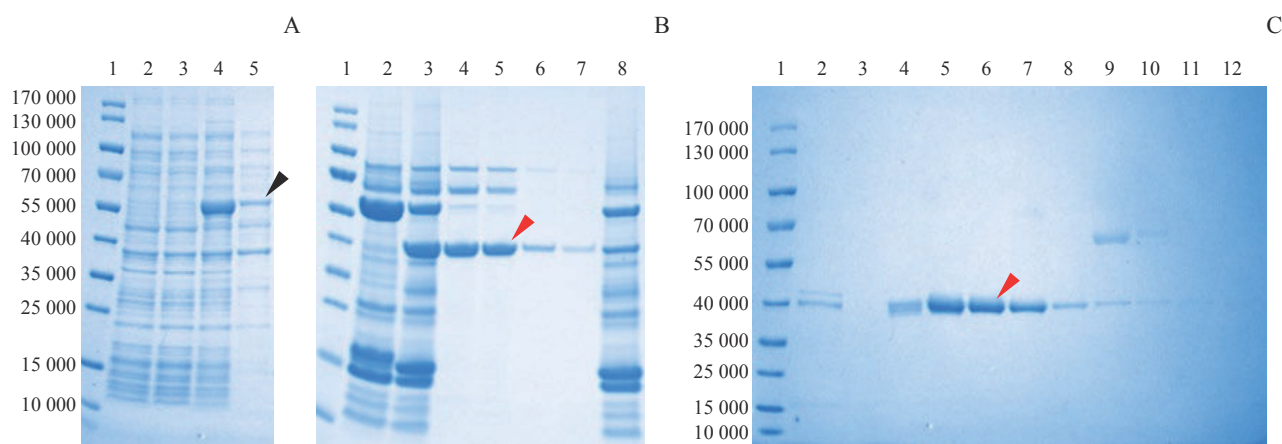
图1 Arthropin与多种serpin的RCL序列比对

Fig 1 RCL sequence comparison of arthropin with various serpins

## 2.2 融合蛋白的表达与纯化

将融合表达质粒 pSUMO3-arthropin 转入大肠埃希菌 BL21 (DE3) 后, 利用 IPTG 进行诱导表达。结果显示, 诱导后的大肠埃希菌裂解总蛋白中检测到大量相对分子质量约为 60 000 的蛋白, 与 His-SUMO3-arthropin 理论相对分子质量 (60 800) 相近, 推测为 arthropin 融合蛋白 (图 2A); 而细菌裂解上清液中只有少量该蛋白, 提示大部分融合蛋白以包涵体的形式

存在, 仅有少部分以可溶性蛋白的形式存在。后续我们采用 HisTrap™ FF 预装柱 (镍离子亲和层析) 从细菌裂解上清液中纯化 His-SUMO3-arthropin 融合蛋白, 然后通过 SENP2 酶切及第 2 次镍离子亲和层析去除 His-SUMO3 标签, 并通过 HiTrap™ Q HP 预装柱 (阴离子交换层析) 进一步纯化 arthropin, 最终获得了高纯度的、与理论相对分子质量 (45 800) 相近的重组 arthropin (图 2B、C)。



**Note:** A. Protein expression in the whole bacteria and lysis supernatant of *E. coli* before and after induction. Lane 1—marker; lane 2—total bacterial protein sample before induction; lane 3—supernatant protein sample before induction; lane 4—total bacterial protein sample after induction for 12 h; lane 5—supernatant protein sample after induction for 12 h. B. Enzymatic digestion and purification by nickel ion affinity chromatography. Lane 1—marker; lane 2—protein sample before SENP2 digestion; lane 3—protein sample after SENP2 digestion; lane 4—5—flow-through solution of the nickel column after SENP2 digestion; lane 6—7—washing solution of the nickel column after SENP2 digestion; lane 8—elution solution of the nickel column after SENP2 digestion. C. Purification by anion-exchange chromatography. Lane 1—marker; lane 2—12—protein samples eluted by a continuous gradient of 145–685 mmol/L NaCl. The black arrow indicates the fusion protein and the red arrows indicate the monomeric protein.

图2 SDS-PAGE分析 arthropin 融合蛋白的表达、酶切与纯化

Fig 2 Expression, enzymatic digestion and purification of arthropin fusion protein analyzed by SDS-PAGE

## 2.3 Arthropin 的靶蛋白酶筛选与动力学检测

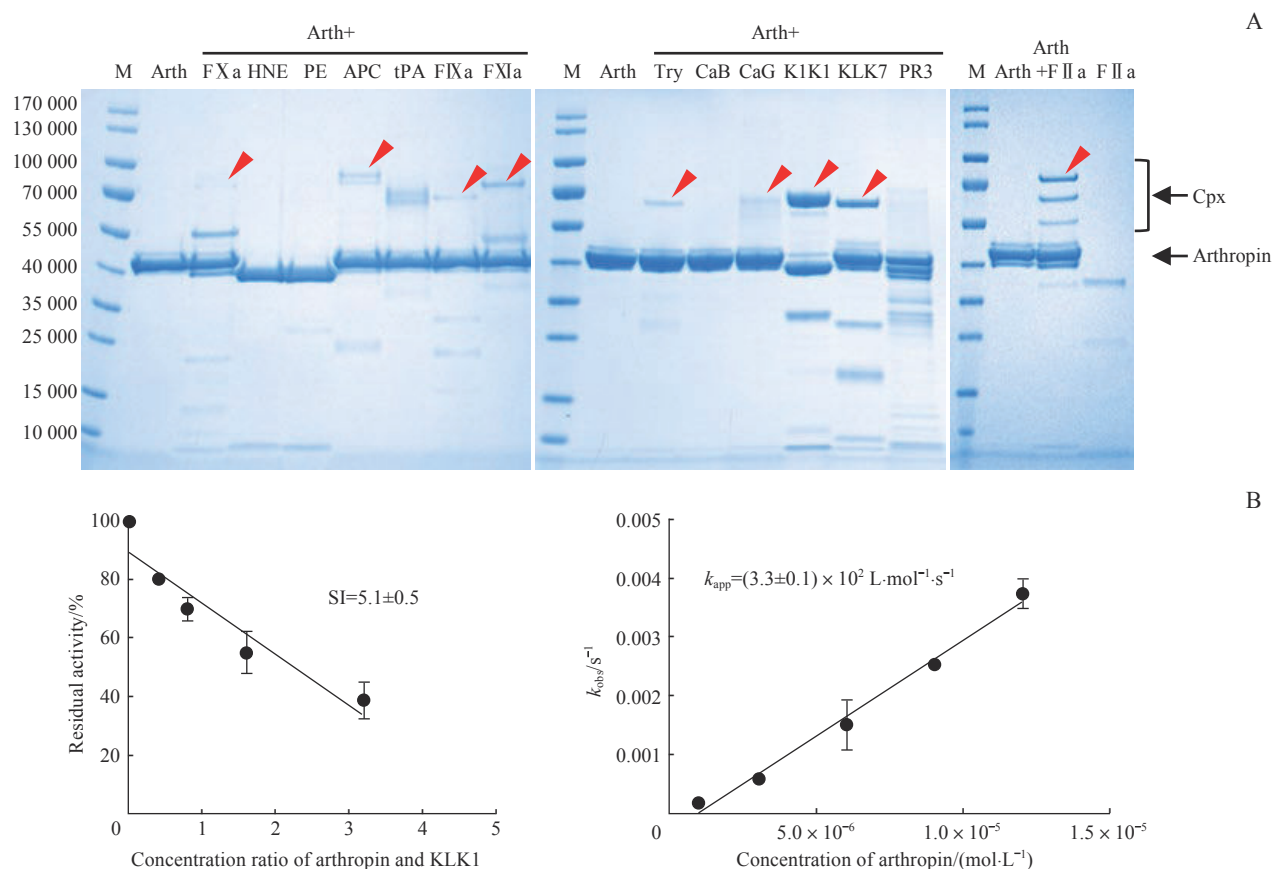
为了鉴定 arthropin 抑制丝氨酸蛋白酶的特异性, 我们将重组 arthropin 与 14 种丝氨酸蛋白酶分别反应, 然后通过 SDS-PAGE 检测其与丝氨酸蛋白酶形成共价复合物的情况。图 3A 结果显示, arthropin 不能与 HNE 和 PE 形成复合物, 反而是作为底物被这 2 种蛋白酶降解, 变成 1 条相对分子质量稍小的蛋白条带及 1 条低于 10 000 的条带。Arthropin 能与 FXa、APC、F IX a、F XI a、胰蛋白酶、CaG、KLK1、KLK7 和 F II a 形成稳定的共价复合物条带, 其中与 KLK1 形成的复合物条带最深, 这说明 arthropin 对这些蛋白酶有抑制活性, 可能对 KLK1 的抑制作用最为显著。

为了更直观地了解 arthropin 对蛋白酶的抑制能力, 我们利用动力学方法检测了 arthropin 抑制 KLK1

的反应速率。图 3B 结果显示, arthropin 抑制 KLK1 的 SI 值为  $5.1 \pm 0.5$ ,  $k_{app}$  为  $(3.3 \pm 0.1) \times 10^2 \text{ L}/(\text{mol} \cdot \text{s})$ 。所以, arthropin 对 KLK1 的  $k_2$  为  $1.7 \times 10^3 \text{ L}/(\text{mol} \cdot \text{s})$ , 低于其他细菌来源的 serpin (如 miropin、siropin 等) [19-20, 37-40]。以上结果表明, arthropin 是一种抑制型 serpin, 对多种蛋白酶有抑制活性, 但反应速率较低。

## 2.4 Arthropin 的结晶条件筛选与结构预测分析

为了明确 arthropin 的结构特点并更直观地阐明其抑制机制, 我们尝试筛选 arthropin 的结晶条件, 为后续晶体结构的解析奠定基础。在以聚乙二醇或磷酸盐作为沉淀剂, 溶液 pH 为 5.6~7.6 的结晶试剂中, arthropin 可以形成碎片状、针状和小颗粒状的蛋白质晶体 (图 4)。其中, 片状晶体棱角分明且光泽较好

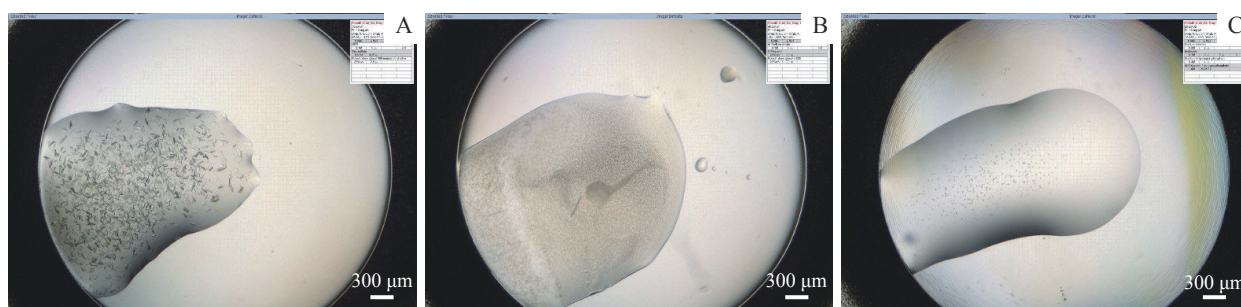


**Note:** A. Arthropin (3  $\mu\text{g}$ ) was incubated with the proteases (0.4  $\mu\text{g}$ ), respectively, and the complex formation was analyzed by SDS-PAGE. The red arrows indicate the complexes of arthropin with the target proteases. M—marker; Try—trypsin; Arth—arthropin; Cpx—arthropin-proteinase covalent complex. B. Determination of the SI (left panel) and the  $k_{app}$  (right panel) of arthropin's inhibition of KLK1.

**图3** Arthropin 靶蛋白酶的筛选与反应动力学检测

**Fig 3** Screening and reaction kinetics assays for target proteases of arthropin

[结晶条件: 25% 聚乙二醇单甲醚 550 (PEG MME 550)、0.1 mol/L 吗啉乙磺酸 (MES, pH 6.5)、0.01 mol/L ZnCl<sub>2</sub>], 但由于该晶体较小, X 射线衍射后分辨率仅为 10 Å (1 Å = 1 × 10<sup>-10</sup> m)。



**Note:** A. 25% PEG MME 550, 0.1 mol/L MES (pH 6.5), and 0.01 mol/L ZnCl<sub>2</sub>. B. 20% PEG 4000, 20% iso-propanol, and 0.1 mol/L sodium citrate (pH 5.6). C. 0.4 mol/L NaH<sub>2</sub>PO<sub>4</sub>, 1.6 mol/L K<sub>2</sub>HPO<sub>4</sub>, 0.1 mol/L imidazole (pH 8.0), and 0.2 mol/L NaCl (final pH 7.6).

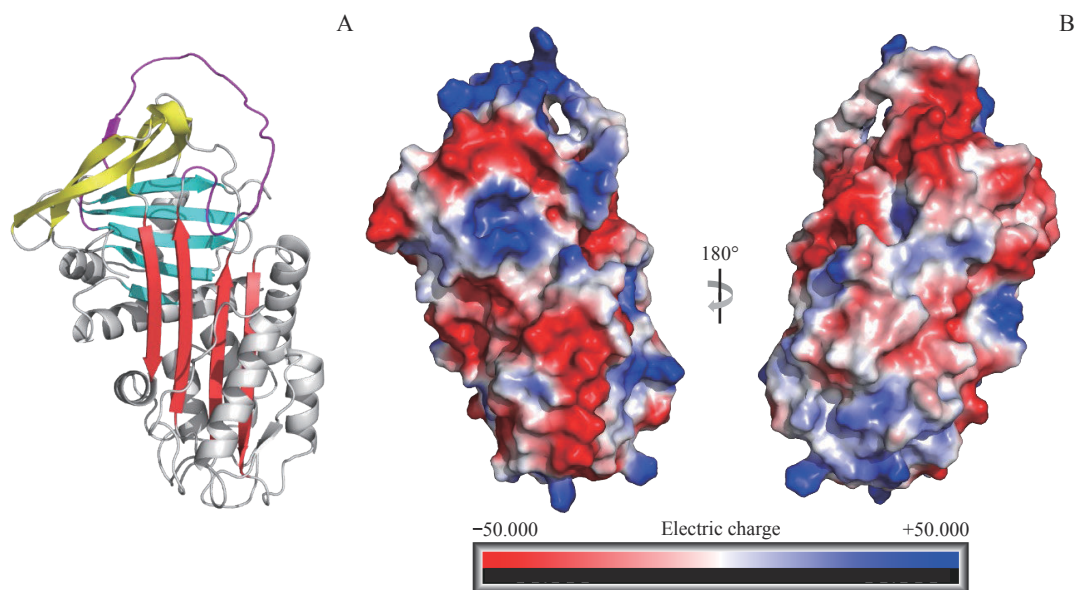
**图4** 不同结晶条件筛选得到的 arthropin 蛋白质晶体(×6.3)

**Fig 4** Arthropin crystals from different crystallization conditions (×6.3)

为了获得一些关于 arthropin 的三维空间结构信息, 我们利用 AlphaFold Colab 预测了 arthropin 亚稳态结构模型; 结果 (图 5) 显示, arthropin 具有经典的抑制型 serpin 的结构特点, 具有中间完整的  $\beta$ -sheet A 和暴露在

分子表面的 RCL, 这是其抑制活性的结构基础。Arthropin 的表面电荷分析结果显示, 其分子表面多数携带负电荷, 表面正电荷较少, 结合其理论等电点为 5.35, 我们认为 arthropin 是一种偏酸性蛋白。





**Note:** A. The arthropin model created with AlphaFold Colab.  $\beta$ -sheets A, B, and C are indicated in red, blue, and yellow, respectively. RCL is in purple. B. Surface charge diagram of arthropin. The red colour indicates the negative charge and the blue colour indicates the positive charge.

图5 Arthropin的结构分析

Fig 5 Structural analysis of arthropin

### 3 讨论

丝氨酸蛋白酶家族成员数量庞大,其序列同源性不高,但结构高度一致。无论是真核或原核生物 serpin 总能以多样的 RCL 抑制特定的丝氨酸蛋白酶。事实上,抑制型 serpin 中多样化的 RCL 同样存在一定的保守规律<sup>[41]</sup>。RCL 的 P1~P7 基本决定了 serpin 的特异性,它们的保守程度较低,多数是疏水氨基酸,其中作为活性中心的 P1 大多数是精氨酸或亮氨酸。P8 多数是小侧链氨基酸。而 P9~P17 是 RCL 的铰链区,通常富含丙氨酸,并且序列较为保守<sup>[2,4]</sup>。P14 作为第一个侧链被掩藏的氨基酸尤为重要,必须是不带电氨基酸;若将 P14 突变为带电氨基酸,虽不影响纤溶酶原激活物抑制物 1 (plasminogen activator inhibitor-1, PAI-1) 对蛋白酶的识别,但将使其丧失对蛋白酶的抑制活性<sup>[41]</sup>。本研究中,arthropin 的 P1 为精氨酸,其 RCL 的序列亦符合典型的抑制型 serpin 的保守规律,因此属于抑制型 serpin 的可能性较大。

为了进一步验证和探究 arthropin 潜在的生理功能,我们利用原核表达系统对 arthropin 进行了重组表达,arthropin 融合蛋白多数以无活性的包涵体形式存在,仅有少部分以可溶性蛋白的形式存在。目前报道的其他原核生物 serpin,如 miropin、siropin 和 TK-serpin 的纯化后产量为每升菌液 25  $\mu$ g、12 mg 和

10 mg<sup>[18-20]</sup>,而 arthropin 纯化后产量为每升菌液 1.8 mg,该产量高于 miropin,但与其他原核生物 serpin 相比处于较低水平。考虑到在常规表达条件下,上清液中的融合蛋白多数以有活性的亚稳态构象存在,且常规产量已满足后续实验需求,本研究未对表达条件进一步优化。由于表达的融合蛋白溶液中含有较多杂蛋白,因此本研究采用了镍离子亲和层析、酶切、反向镍离子亲和层析、阴离子交换层析的 4 步层析纯化,最终得到了高纯度的重组 arthropin 蛋白。

将纯化后的 arthropin 重组蛋白用于晶体结构解析,仅能得到分辨率为 10 Å 的蛋白质晶体,尚不足以解析其结构。分析其主要原因有二:一是重组蛋白均一性不足,后续研究可通过增加凝胶过滤层析等纯化步骤进一步提高蛋白均一性;二是重组蛋白自身结构过于灵活,不利于蛋白质晶体的形成。以往对真核生物 serpin 的结构与功能研究已知,保证 serpin 的主体结构及 RCL 完整是其发挥抑制功能的前提,而 serpin N 端可能存在的非二级结构片段一般不具有生理功能<sup>[1]</sup>。基于 arthropin 的结构模型,并将其氨基酸序列与  $\alpha$ 1-抗胰蛋白酶进行比对,我们发现 arthropin 的 N 端可能存在长度为 52 个氨基酸的非二级结构片段,该片段长度与经典 serpin—— $\alpha$ 1-抗胰蛋白酶相近,将此片段截短可能有利于重组蛋白结晶。

Arthropin 与丝氨酸蛋白酶共孵育的 SDS-PAGE 结

果显示, arthropin 对 9 种蛋白酶发挥抑制活性, 其中对 KLK1 的抑制作用可能最为显著。Arthropin 对 KLK1 的抑制速率 ( $k_2$ ) 为  $1.7 \times 10^3 \text{ L}/(\text{mol} \cdot \text{s})$ , 略低于其他原核生物 serpin, 远低于真核生物 serpin。天然状态下, 原核生物 serpin 对丝氨酸蛋白酶的抑制速率为  $10^4 \sim 10^5 \text{ L}/(\text{mol} \cdot \text{s})$ 。例如, thermopain 抑制胰凝乳蛋白酶 (chymotrypsin) 的速率为  $8.4 \times 10^4 \text{ L}/(\text{mol} \cdot \text{s})$  [37], 双歧杆菌中 serpin 抑制中性粒细胞弹性蛋白酶 (neutrophil elastase) 的速率为  $4.7 \times 10^4 \text{ L}/(\text{mol} \cdot \text{s})$  [38], aeropain 抑制胰凝乳蛋白酶的速率为  $1.8 \times 10^5 \text{ L}/(\text{mol} \cdot \text{s})$  [39] 等。而真核生物 serpin 抑制作用的速率为  $10^5 \sim 10^6 \text{ L}/(\text{mol} \cdot \text{s})$  [40]。Arthropin 对蛋白酶较低的抑制速率可能与其较广谱的抑制能力有关。

Miropain 是首个被发现具有广谱抑制能力的原核生物 serpin, 其来源于口腔牙周疾病相关细菌; 在应激时, miropain 通过抑制多种蛋白酶保护细菌自身的存活 [19,42]。而蓝藻作为最古老的原核生物之一, 同样具有抵御各种生物和环境侵袭的生命机制 [43]。蓝藻天然生产的次生代谢化合物便多数具有显著的蛋白酶抑制活性 [44], 例如铜绿菌素 (aeruginosin) [45]、微囊藻毒素 (microcystins) [46-47]、草抑制素 A~C (grassyatin A~C) [29] 等。Arthropin 来源于钝顶节旋藻, 基于 arthropin 拥有比 miropain 更为广谱的丝氨酸蛋白酶抑制能力, 我们推测 arthropin 极有可能是作为蓝藻机体防御中重要的一员。在本研究筛选范围内, arthropin 抑制的蛋白酶多数与凝血相关。虽然目前

arthropin 的天然靶蛋白酶尚不清楚, 但是它可能通过抑制多种外源性的丝氨酸蛋白酶, 发挥保护自身免受侵害的作用。

综上所述, 本研究发现并制备了来源于钝顶节旋藻的新型丝氨酸蛋白酶抑制剂 arthropin; 并对其初步的靶蛋白酶筛选揭示 arthropin 拥有广谱的丝氨酸蛋白酶抑制活性, 这可能对其发挥保护自身免受环境蛋白酶侵害意义重大。同时, 我们对 arthropin 结晶条件的筛选也为更深入地探索 arthropin 结构与作用机制奠定了基础。

#### 利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors disclose no relevant conflict of interests.

#### 作者贡献/Authors' Contributions

周爱武主持并设计本研究项目; 许佳伟在周爱武、杨愈丰的指导下完成所有实验工作和数据整理; 许佳伟负责撰写及修改论文; 杨愈丰、周爱武负责论文修改的指导。所有作者均阅读并同意了最终稿件的提交。

ZHOU Aiwu conducted and designed this study. XU Jiawei completed all the experiments and the data compilation with the guidance of ZHOU Aiwu and YANG Yufeng. XU Jiawei wrote and revised the thesis, and YANG Yufeng and ZHOU Aiwu provided guidance on the revision of the thesis. All the authors have read the last version of paper and consented for submission.

- Received: 2022-12-08
- Accepted: 2023-04-03
- Published online: 2023-04-28

#### 参 考 文 献

- [1] LAW R H, ZHANG Q W, MCGOWAN S, et al. An overview of the serpin superfamily[J]. Genome Biol, 2006, 7(5): 216.
- [2] HUBER R, CARRELL R W. Implications of the three-dimensional structure of  $\alpha_1$ -antitrypsin for structure and function of serpins[J]. Biochemistry, 1989, 28(23): 8951-8966.
- [3] GETTINS P G W. Serpin structure, mechanism, and function[J]. Chem Rev, 2002, 102(12): 4751-4804.
- [4] IRVING J A, STEENBAKKERS P J, LESK A M, et al. Serpins in prokaryotes[J]. Mol Biol Evol, 2002, 19(11): 1881-1890.
- [5] SPENCE M A, MORTIMER M D, BUCKLE A M, et al. A comprehensive phylogenetic analysis of the serpin superfamily[J]. Mol Biol Evol, 2021, 38(7): 2915-2929.
- [6] HEIT C, JACKSON B C, MCANDREWS M, et al. Update of the human and mouse serpin gene superfamily[J]. Hum Genomics, 2013, 7(1): 22.
- [7] HUNTINGTON J A, READ R J, CARRELL R W. Structure of a serpin-protease complex shows inhibition by deformation[J]. Nature, 2000, 407(6806): 923-926.
- [8] IRVING J A, PIKE R N, DAI W W, et al. Evidence that serpin architecture intrinsically supports papain-like cysteine protease inhibition: engineering  $\alpha_1$ -antitrypsin to inhibit cathepsin proteases[J]. Biochemistry, 2002, 41(15): 4998-5004.
- [9] IKEZOE T. Advances in the diagnosis and treatment of disseminated intravascular coagulation in haematological malignancies[J]. Int J Hematol, 2021, 113(1): 34-44.
- [10] GREENE C M, MARCINIAK S J, TECKMAN J, et al.  $\alpha_1$ -Antitrypsin deficiency[J]. Nat Rev Dis Primers, 2016, 2: 16051.
- [11] FRANÇOIS D, VENISSE L, MARCHAL-SOMME J, et al. Increased expression of protease nexin-1 in fibroblasts during idiopathic pulmonary fibrosis regulates thrombin activity and fibronectin expression[J]. Lab Invest, 2014, 94(11): 1237-1246.
- [12] MAAS C, DE MAAT S. Therapeutic SERPINS: improving on nature[J]. Front Cardiovasc Med, 2021, 8: 648349.
- [13] KLIEBER M A, UNDERHILL C, HAMMOND G L, et al. Corticosteroid-binding globulin, a structural basis for steroid transport and proteinase-triggered release[J]. J Biol Chem, 2007, 282(40): 29594-29603.
- [14] ZHOU A W, WEI Z Q, READ R J, et al. Structural mechanism for



- the carriage and release of thyroxine in the blood[J]. *Proc Natl Acad Sci U S A*, 2006, 103(36): 13321-13326.
- [15] ZHANG M, MAGIT D, SAGER R. Expression of maspin in prostate cells is regulated by a positive Ets element and a negative hormonal responsive element site recognized by androgen receptor[J]. *Proc Natl Acad Sci U S A*, 1997, 94(11): 5673-5678.
- [16] HUNT L T, DAYHOFF M O. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and  $\alpha_1$ -proteinase inhibitor[J]. *Biochem Biophys Res Commun*, 1980, 95(2): 864-871.
- [17] LOMAS D A, EVANS D L, FINCH J T, et al. The mechanism of Z  $\alpha_1$ -antitrypsin accumulation in the liver[J]. *Nature*, 1992, 357(6379): 605-607.
- [18] TANAKA S, KOGA Y, TAKANO K, et al. Inhibition of chymotrypsin- and subtilisin-like serine proteases with Tk-serpin from hyperthermophilic archaeon *Thermococcus kodakaraensis*[J]. *Biochim Biophys Acta*, 2011, 1814(2): 299-307.
- [19] KSIAZEK M, MIZGALSKA D, ENGHILD J J, et al. Miropin, a novel bacterial serpin from the periodontopathogen *Tannerella forsythia*, inhibits a broad range of proteases by using different peptide bonds within the reactive center loop[J]. *J Biol Chem*, 2015, 290(1): 658-670.
- [20] MKAOUAR H, AKERMI N, MARIAULE V, et al. Siropins, novel serine protease inhibitors from gut microbiota acting on human proteases involved in inflammatory bowel diseases[J]. *Microb Cell Fact*, 2016, 15(1): 201.
- [21] NAGY Z A, SZAKÁCS D, BOROS E, et al. Ecotin, a microbial inhibitor of serine proteases, blocks multiple complement dependent and independent microbicidal activities of human serum[J]. *PLoS Pathog*, 2019, 15(12): e1008232.
- [22] WOLLINA U, VOICU C, GIANFALDONI S, et al. *Arthrospira platensis*-potential in dermatology and beyond[J]. *Open Access Maced J Med Sci*, 2018, 6(1): 176-180.
- [23] CLEMENT G. Production and characteristic constituents of the algae *Spirulina platensis* and maxima[J]. *Ann Nutr Aliment*, 1975, 29(6): 477-488.
- [24] AYEHEUNIE S, BELAY A, BABA T W, et al. Inhibition of HIV-1 replication by an aqueous extract of *Spirulina platensis* (*Arthrospira platensis*)[J]. *J Acquir Immune Defic Syndr Hum Retrovirol*, 1998, 18(1): 7-12.
- [25] TEAS J, IRHIMEH M R. Dietary algae and HIV/AIDS: proof of concept clinical data[J]. *J Appl Phycol*, 2012, 24(3): 575-582.
- [26] ROMAY C H, GONZÁLEZ R, LEDÓN N, et al. C-phycocyanin: a biliprotein with antioxidant, anti-inflammatory and neuroprotective effects[J]. *Curr Protein Pept Sci*, 2003, 4(3): 207-216.
- [27] PORTMANN C, BLOM J F, KAISER M, et al. Isolation of aerucyclamides C and D and structure revision of microcyclamide 7806A: heterocyclic ribosomal peptides from *Microcystis aeruginosa* PCC 7806 and their antiparasite evaluation[J]. *J Nat Prod*, 2008, 71(11): 1891-1896.
- [28] SYARINA P N, KARTHIVASHAN G, ABAS F, et al. Wound healing potential of *Spirulina platensis* extracts on human dermal fibroblast cells[J]. *EXCLI J*, 2015, 14: 385-393.
- [29] KWAN J C, EKSIOGLU E A, LIU C, et al. Grassystatins A-C from marine cyanobacteria, potent cathepsin E inhibitors that reduce antigen presentation[J]. *J Med Chem*, 2009, 52(18): 5732-5747.
- [30] GRAWISH M E. Effects of *Spirulina platensis* extract on Syrian hamster cheek pouch mucosa painted with 7, 12-dimethylbenz[ $\alpha$ ]anthracene[J]. *Oral Oncol*, 2008, 44(10): 956-962.
- [31] MISHIMA T, MURATA J, TOYOSHIMA M, et al. Inhibition of tumor invasion and metastasis by calcium spirulan (Ca-SP), a novel sulfated polysaccharide derived from a blue-green alga, *Spirulina platensis*[J]. *Clin Exp Metastasis*, 1998, 16(6): 541-550.
- [32] BRAUNE S, KRÜGER-GENGE A, KAMMERER S, et al. Phycocyanin from *Arthrospira platensis* as potential anti-cancer drug: review of *in vitro* and *in vivo* studies[J]. *Life (Basel)*, 2021, 11(2): 91.
- [33] LIU Q, HUANG Y H, ZHANG R H, et al. Medical application of *Spirulina platensis* derived C-phycocyanin[J]. *Evid Based Complement Alternat Med*, 2016, 2016: 7803846.
- [34] CATASSI A, CESARIO A, ARZANI D, et al. Characterization of apoptosis induced by marine natural products in non small cell lung cancer A549 cells[J]. *Cell Mol Life Sci*, 2006, 63(19-20): 2377-2386.
- [35] MALAKHOV M P, MATTERN M R, MALAKHOVA O A, et al. SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins[J]. *J Struct Funct Genomics*, 2004, 5(1-2): 75-86.
- [36] WEI H D, CAI H Y, WU J W, et al. Heparin binds lamprey angiotensinogen and promotes thrombin inhibition through a template mechanism[J]. *J Biol Chem*, 2016, 291(48): 24900-24911.
- [37] IRVING J A, CABRITA L D, ROSSJOHN J, et al. The 1.5 Å crystal structure of a prokaryote serpin: controlling conformational change in a heated environment[J]. *Structure*, 2003, 11(4): 387-397.
- [38] IVANOV D, EMONET C, FOATA F, et al. A serpin from the gut bacterium *Bifidobacterium longum* inhibits eukaryotic elastase-like serine proteases[J]. *J Biol Chem*, 2006, 281(25): 17246-17252.
- [39] CABRITA L D, IRVING J A, PEARCE M C, et al. Aeropin from the extremophile *Pyrobaculum aerophilum* bypasses the serpin misfolding trap[J]. *J Biol Chem*, 2007, 282(37): 26802-26809.
- [40] DAHLEN J R, JEAN F, THOMAS G, et al. Inhibition of soluble recombinant furin by human proteinase inhibitor 8[J]. *J Biol Chem*, 1998, 273(4): 1851-1854.
- [41] SANRATTANA W, MAAS C, DE MAAT S. SERPINs: from trap to treatment[J]. *Front Med (Lausanne)*, 2019, 6: 25.
- [42] GOULAS T, KSIAZEK M, GARCIA-FERRER I, et al. A structure-derived snap-trap mechanism of a multispecific serpin from the dysbiotic human oral microbiome[J]. *J Biol Chem*, 2017, 292(26): 10883-10898.
- [43] ANAS A R, NAKAJIMAA, NARUSE C, et al. Determination of FVIIa-sTF inhibitors in toxic *Microcystis* cyanobacteria by LC-MS technique[J]. *Mar Drugs*, 2015, 14(1): 7.
- [44] WELKER M, BRUNKE M, PREUSSEL K, et al. Diversity and distribution of *Microcystis* (Cyanobacteria) oligopeptide chemotypes from natural communities studied by single-colony mass spectrometry[J]. *Microbiology (Reading)*, 2004, 150(Pt 6): 1785-1796.
- [45] AHMED M N, WAHLSTEN M, JOKELA J, et al. Potent inhibitor of human trypsin from the aeruginosin family of natural products[J]. *ACS Chem Biol*, 2021, 16(11): 2537-2546.
- [46] LIANG J, LI T, ZHANG Y L, et al. Effect of microcystin-LR on protein phosphatase 2A and its function in human amniotic epithelial cells[J]. *J Zhejiang Univ Sci B*, 2011, 12(12): 951-960.
- [47] VON ELERT E, OBERER L, MERKEL P, et al. Cyanopeptolin 954, a chlorine-containing chymotrypsin inhibitor of *Microcystis aeruginosa* NIVA Cya 43[J]. *J Nat Prod*, 2005, 68(9): 1324-1327.

[本文编辑] 瞿麟平