

蚯蚓纤溶酶新基因 PV₂₄₂ 的克隆与表达*

徐义辉 梁国栋** 孙兆军 陈 飞 付士红 柴玉波 侯云德

(中国预防医学科学院病毒学研究所, 病毒基因工程国家重点实验室, 北京 100052)

摘要 从我国特有的双胸蚓属 (*Lumbricus bimastus*) 蚯蚓体内提取到一种有纤维蛋白平板溶解活性的蛋白质, 采用聚偏二氟乙烯膜 (PVDF) 微量测序法测得蛋白质的 N 端氨基酸序列, 依此设计简并引物, 通过 RT-PCR 获得其对应 cDNA. 该 cDNA 全长 888 bp, 1~726 位核苷酸对应读码框架 (ORF), 编码含 242 个氨基酸的成熟肽. 终止子 (TAG) 位于 cDNA 的 727~729 位, 其余核苷酸序列为 3' 端非编码区. 成熟肽命名为蚯蚓纤溶酶 PV₂₄₂. 蛋白质预测得知蛋白质等电点为 4.33, 含组氨酸 (His44) 及丝氨酸 (Ser191) 两个活性位点; 蛋白质由两个结构域组成, 表面有活性裂隙; 该蛋白质属丝氨酸蛋白酶超家族胰蛋白酶类. 经国际基因库等多种查询比较未见 PV₂₄₂ 基因的报道, 为首次发现的新基因, 在国际基因库的输入号为 GenBank AF109648. 随后构建了 pTrxFUS-PV₂₄₂ 重组质粒, 并在大肠杆菌 GI724 中获得融合蛋白 TrxA/PV₂₄₂ 的可溶性表达, 采用离子交换层析法纯化表达蛋白, 融合蛋白有纤维蛋白溶解活性.

关键词 双胸蚓属, 蚯蚓纤溶酶, 蛋白质测序, 克隆, 重组表达

学科分类号 Q812, R972

血栓形成可导致心肌梗塞、中风、肺栓塞等严重威胁人类健康和生命的疾病. 目前应用的溶血栓药最常用的是组织纤溶酶原激活物 (tPA) 与尿激酶型纤溶酶激活物 (uPA), 此外还有链激酶、葡激酶等. 但这些药物在体内半衰期很短, 临床使用时可能发生出血性不良反应或免疫反应. 蚯蚓纤溶酶是从蚯蚓中提取的具有溶血栓活性的一类蛋白水解酶. 该类蛋白酶富含酸性氨基酸、分子质量在 20~40 ku, 具有半衰期长、无出血等不良反应. 蚯蚓纤溶酶前期实验表明溶栓效果比较理想, 具有广泛的临床应用前景, 但是还存在一系列问题有待解决. 所以, 对蚯蚓纤溶酶的研究将具有重要的意义.

从我国特有的正蚓科双胸属蚯蚓 (*Lumbricus bimastus*) 中分离出一种纤溶酶, 经实验证明有纤维蛋白平板溶解活性^[1]. 在此基础上克隆到该纤溶酶的基因, 基因命名为蚯蚓纤溶酶 PV₂₄₂, 是首次发现的新基因. 实验同时对 PV₂₄₂ 基因进行了重组表达及活性测定.

1 材料和方法

1.1 材料

蚯蚓纤溶酶提取物由辽宁省微生物研究所提供. 尿激酶检测试剂盒购自卫生部生物制品检定研究所; 聚偏二氟乙烯膜 (PVDF) 为 Bio-Rad 公司产品; mRNA 分离纯化试剂盒、pGET-Vector 为

Promega 公司产品; 色氨酸购自华美生物技术公司. 强阴离子纯化柱 SepharoseFF 为 Pharmacia 公司产品; 蚯蚓纤溶酶 PV₂₄₂ 多克隆抗体由本室制备; Taq 酶、Pst I、Kpn I 为 Boehringer 公司产品; 感受态 *E. coli* (DH5α) 由本室制备; 大肠杆菌 TOP10、pTrxFUS、GI724 为 Invitrogen 公司产品; RM 培养基、M9 诱导培养基、细菌裂解液按说明书配制.

1.2 方法

1.2.1 蛋白质 N 端氨基酸序列测定: 用 1 倍的上样缓冲液将从正蚓科双胸蚓属 (*Lumbricus bimastus*) 蚯蚓中提取的纤溶酶干粉溶解成 1 g/L. 按 10 μl/孔上样进行聚丙烯酰胺凝胶电泳 (PAGE) 3 h, 按文献 [2] 所述方法进行亲水性聚偏二氟乙烯膜 (PVDF) 转印, 和 PVDF 膜的染色、脱色. 将上述 PVDF 膜上的蛋白质条带按带形剪下, 用 Applied Biosystems 的 491 型 Protein Sequencer 自动测序仪进行蛋白质 N 端测序^[2].

1.2.2 mRNA 的反转录及反转录产物的 PCR 扩增: 根据纯化所得的蚯蚓纤溶酶 N 端氨基酸序列设计合成简并寡核苷酸扩增引物, 引物 1 序列

* 国家高技术研究发展计划 (863) 生物技术资助项目 (102-08-02-03).

** 通讯联系人.

Tel: 010-63581367, E-mail: gdliang@95777.com

收稿日期: 2002-01-29, 接受日期: 2002-03-28

为 5' GTNATHGGNGGNCANAAYGC3' (其中 N = A、C、G、T; H = A、C、T; Y = C、T). 另一引物为 Oligo dT (12~18). 取蚯蚓 mRNA 3.0 μ l, Oligo dT 1.0 μ l (10 mmol/L), 加入反转录缓冲液, 按常规方法进行反转录. 随后, 取反转录产物进行 PCR 扩增, T_m 为 60 $^{\circ}$ C.

1.2.3 PCR 产物的克隆与克隆片段序列测定: PCR 产物的酶切、电泳回收、酶切产物与 pGEM-T 载体连接与转化等基因操作参照文献 [2] 的方法. 提取重组 pGEM-T 质粒, 测定重组片段的核苷酸序列.

1.2.4 重组片段序列分析与蛋白质结构及特性预测: 根据重组片段核酸序列及其对应的氨基酸序列, 采用多种蛋白质分析软件和算法 (DNAsis、pcgene 软件、SWISS-MODEL 法、Insight II 97 软件)、NCBI (美国国立生物技术信息中心) 的多种 Blast 功能和 PHD (Profile fed neural network systems from HeiDelberg) prediction 分析功能对克隆片段的组成及其对应蛋白质的理化性质、结构与活性进行预测和分析.

1.2.5 pTrxfus-PV₂₄₂ 重组质粒的构建: 根据克隆片段和载体酶切位点的特性及纯化需要, 设计扩增重组片段的 PCR 扩增引物. 引物 2: 5' c ggg gta cct cat cat cat cat cat cat gtg ata ggg ggc act aac g 3' (内设 *Kpn* I 酶切位点); 引物 3: 5' aactg cag tcgacaagcttactaacg aga gtt gtc tcc gat c3' (内设 *Pst* I 酶切位点). 带下划线部分为终止子; 纯化用多 His 位点设计在表达蛋白 N 端.

以重组 pGEM-T 质粒为模板, 引物 2、3 进行 PCR 扩增, T_m 为 55 $^{\circ}$ C. 扩增产物和 pTrxfus 质粒分别用 *Kpn* I 和 *Pst* I 双酶切, 回收酶切片段. 回收产物采用 T4 连接酶连接. 将连接产物转化感受态大肠杆菌 TOP10, 取 100 μ l 转化菌液涂 Amp⁺ LB 平板, 37 $^{\circ}$ C 培养 12 h.

1.2.6 pTrxfus-PV₂₄₂ 在 GI724 中的可溶性表达与检测: 挑取上述单菌落 (GI724/pTrxFUS-PV₂₄₂) 接种于 M9 丰富培养基中, 30 $^{\circ}$ C 过夜培养后转接于诱导培养基中, 30 $^{\circ}$ C 培养至 A_{600} 为 0.5, 加色氨酸至终浓度为 100 mg/L, 在 37 $^{\circ}$ C 继续培养 3 h. 取细菌培养物, 离心、收集菌体, 用细胞裂解液重悬, 超声破碎. 离心后取上清进行 SDS-PAGE 和蛋白质印迹检测, 并对 SDS-PAGE 胶进行扫描测定表达量. 用微量蛋白质测序法检测重组表达蛋白 N 端氨基酸序列.

1.2.7 Trx A/PV₂₄₂ 融合蛋白的纯化与活性检测: 取 300 ml 上述诱导表达培养液离心、取菌体进行超声破碎. 上清液用去离子水透析 8 h、低温冻干, 冻干粉用上样缓冲液 (50 mmol/L Tris pH 7.23) 溶解, 上样于经过平衡处理的强阴离子柱 QsepharoseFF, 以平衡液 (50 mmol/L Tris, pH 7.23) 将不保留组分洗尽. 然后以 2 mol/L NaCl 洗脱液进行梯度洗脱, 收集洗脱峰进行蛋白质印迹分析.

洗脱峰所得样品经透析后, 用 PBS (pH 7.8) 溶解成 100 μ l. 挑取 GI724/pTrxFUS、GI724/pTrxFUS-PV₂₄₂ 单菌落分别接种于 5 ml M9 丰富培养基中, 30 $^{\circ}$ C 过夜培养. 次日转接于 50 ml 诱导培养基中, 30 $^{\circ}$ C 培养、色氨酸诱导 4 h. 表达菌体的超声破碎上清经透析、冻干后用去离子水溶解成 30 μ l, 进行纤维蛋白平板活性检测. 同时各取 30 μ l 上述纯化样品进行纤维蛋白平板活性检测.

2 结 果

2.1 蚯蚓纤溶酶提取物的 SDS PAGE 分析

从蚯蚓纤溶酶提取物中分离出单一蛋白质, 蛋白质的 SDS-PAGE 电泳分子质量大约为 30 ku. 该蛋白质经检测有纤维蛋白平板溶解活性^[1].

2.2 提取物蛋白质成分的 N 端氨基酸序列分析

分离蛋白经 SDS-PAGE 电泳、PVDF 膜转印、染色后, 切下相应蛋白质条带进行蛋白质 N 端测序. 测得蛋白质的 N 端前 10 个氨基酸序列为: VIGGTNASPG.

2.3 蚯蚓纤溶酶 PV₂₄₂ 的基因克隆

根据蛋白质带 N 端氨基酸序列设计简并引物, 以蚯蚓的总 mRNA 为模板经 RT-PCR 获得约 1 kb 的特异 cDNA 扩增片段, 将扩增片段克隆到 pGEMT 载体后, 经酶切鉴定为目的片段.

2.4 PV₂₄₂ 基因分析及其编码蛋白质的预测分析

对上述 pGEM-T 重组质粒上的 cDNA 片段进行核酸序列测定和分析, 结果是: cDNA 全长 888 bp, 其中 1~726 位核苷酸对应基因的读码框架 (ORF), 编码成熟肽序列, 其余核苷酸序列为 3' 端非编码区. 成熟肽终止子位于该 cDNA 片段的 727~729 位之间, 终止密码子为 TAG. 该 cDNA 序列已输入 GenBank, 序列号为 AF109648.

上述 cDNA 序列按照三联密码子原则, 采用 DNAsis 推测出编码的成熟肽氨基酸序列. 该序列编码 242 个氨基酸, N 端第一个氨基酸为 V, 因此将该蛋白质命名为蚯蚓纤溶酶 PV₂₄₂, 该蛋白质的

基因称作蚯蚓纤溶酶 PV₂₄₂ 基因. PV₂₄₂ 经国际基因库检索未见相同序列报道, 为新基因. Blast 分析 PV₂₄₂ 蛋白与海沙蝎的糜蛋白酶原同源性最高为 44%, 该蛋白质与日本报道的蚯蚓纤溶酶 F-III 1、F-III 2 及中国发现的另一蚯蚓纤溶酶的氨基酸序列同源性在 30%~35%.

使用 pcgene 软件分析 PV₂₄₂ 氨基酸序列组成, 预测其 pI 为 4.33, 含组氨酸 (40 LSASHC 45) 及丝氨酸 (189 GDSGG 193) 两个活性位点, PHD 分析 PV₂₄₂ 是丝氨酸蛋白酶超家族中的胰蛋白酶家

族成员. 三维结构预测显示 (图 1): PV₂₄₂ 含两个相对独立的结构域 (domain). 两个结构域均含多个 β 折叠, 属于 β 折叠型. 两结构域交接处为活性区, 活性位点 His44、Ser191 位于两个结构域交界处的两个不同 loop 上. 预测认为 PV₂₄₂ 形成三个二硫键, 即 Cys158~Cys176、Cys127~Cys197、Cys29~Cys45. PV₂₄₂ 三维结构显示蛋白质外形呈球形, 表面有一南北纵向疏水裂隙 (图 1b 箭头所示), 活性位点 His44、Ser191 位于裂隙的中心 (图 1), 为典型的蛋白酶结构形态.

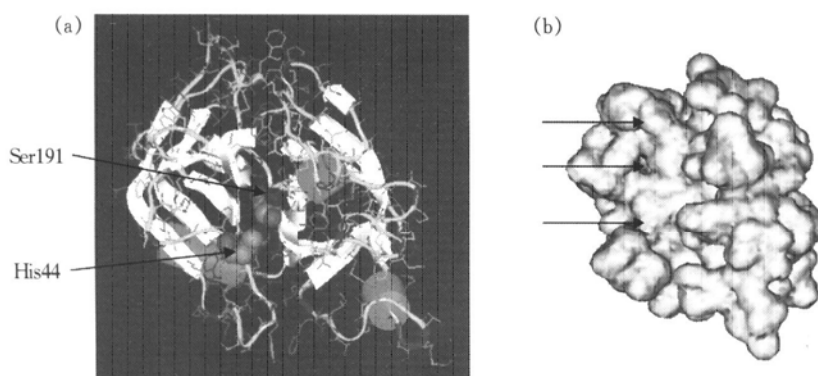


Fig. 1 The predicted structure of PV₂₄₂

(a) The inside structure of PV₂₄₂: the molecular had two domains and one was in left, another in right. The balls represented the active amino acids His44 and Ser191; (b) The surface form of PV₂₄₂: there was a cleft on the surface of PV₂₄₂ (where the arrows pointed).

2.5 TrxA-PV₂₄₂ 融合蛋白在大肠杆菌中的表达

将 PV₂₄₂ 基因克隆入 pTrxfus 质粒构成 pTrxfus-PV₂₄₂ 重组质粒. 其中 PV₂₄₂ 上游为分子伴侣硫氧还蛋白 TrxA 的基因, PV₂₄₂ 与 TrxA 以融合形式表达. TrxA 分子质量为 11.7 ku, 如果融合 30 ku 的 PV₂₄₂, 表达蛋白应为 40 ku 左右. 将 pTrxfus-PV₂₄₂ 重组质粒转化大肠杆菌 G1724, 随后进行色氨酸诱导. 表达产物经 SDS-PAGE 检测,

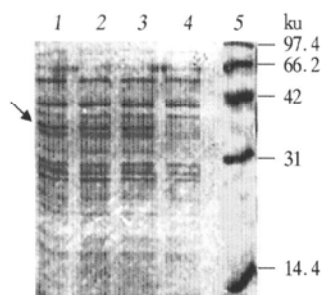


Fig. 2 SDS PAGE analysis of TrxA-PV₂₄₂ fusion protein expression in G1724

1, 2, 3, 4: G1724/pTrxfus PV₂₄₂ induced for 3 h, 2 h, 1 h, 0 h (The arrowhead pointed the target fusion protein); 5: protein molecular mass marker.

显示有 40 ku 左右大小的诱导表达蛋白带 (图 2). 用本室制备的蚯蚓纤溶酶 PV₂₄₂ 多克隆抗体对诱导表达上清进行蛋白质印迹检测, 诱导样品在 40 ku 左右同样位置有阳性显色带 (图 3 中 3、4、5), 阴性对照 (图 3 中 1、2) 没有. 对上述诱导表达带进行 PVDF 微量蛋白测序, 该蛋白质 N 端前 5 个氨基酸序列为: SDKII, 同融合蛋白的 N 端序列一

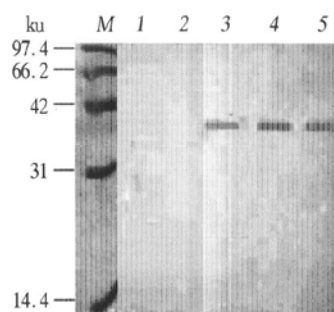


Fig. 3 Western blot analysis of TrxA-PV₂₄₂ fusion protein expression in G1724

1: G1724/pTrxfus induced for 4 h; 2~5: G1724/pTrxfus PV₂₄₂ induced for 0 h, 1 h, 2 h, 3 h; M: protein molecular mass marker.

致. 认为上述蛋白带是融合蛋白 TrxA/PV₂₄₂. 融合蛋白表达量占菌体可溶性蛋白的 8%.

2.6 TrxA/PV₂₄₂融合蛋白的纯化与活性鉴定结果

挑取 pTrxfus PV₂₄₂/GI724 单克隆, 进行诱导表达. 取 300 ml 诱导液离心, 取沉淀. 用细菌裂解液重悬进行超声破碎. 破碎上清低温冻干后, 用离子交换法纯化. 纯化产生 3 个洗脱峰. 蛋白质印迹检测纯化峰样, 结果显示 (图 4): 峰 1、峰 2 含融合表达蛋白带, 峰 3 是杂质峰. 样品活性检测结果 (图 5): GI724/pTrxfus PV₂₄₂诱导表达 4 h

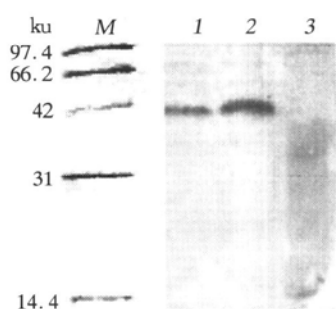


Fig. 4 Western blot analysis of purified TrxA-PV₂₄₂ fusion protein expression in GI724

1: eluted peak 1; 2: eluted peak 2; 3: eluted peak 3; M: protein molecular marker.

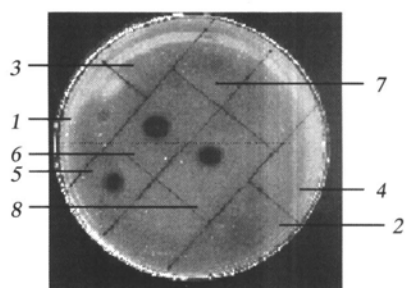


Fig. 5 The analysis of TrxA-PV₂₄₂ fusion protein activity

1: 30 μ l aliquote of sonicated bacteria (pTrxfus/PV₂₄₂ in GI724 induced for 4 h); 2: 30 μ l aliquote of sonicated bacteria (pTrxfus in GI724 induced for 4 h); 3: 30 μ l PBS; 4: 30 μ l Try solution; 5: 30 μ l eluted peak 1 production; 6: 30 μ l eluted peak 2 production; 7: eluted peak 3 production 30 μ l; 8: 1 U urokinase.

的菌体破碎上清样, 第 1 号峰样, 2 号峰样, 尿激酶样 (依次对应图 5 中 1、5、6、8 号样) 有纤维平板溶解活性, 其中 2 号洗脱峰样品 (图 5 的 6 号样) 产生的溶圈最大, GI724/pTrxfus-PV₂₄₂诱导表达 4 h 的菌体破碎上清样 (图 5 的 1 号样) 的溶圈最弱, 只是隐约可见. 而 GI724/pTrxfus 诱导表达 4 h 的菌体破碎上清样没有产生溶圈. 由此认为表达目的蛋白有纤溶活性. 设 PBS、色氨酸、

GI724/pTrxfus 色氨酸诱导 4 h 后菌体的超声破碎上清 (50 ml 培养物) 为阴性对照 (图 5 中 2、3、4、7 号样). 纯化后的活性峰样经活性检测和计算得知: 每升培养液纯化到的融合蛋白活性为 10 U.

3 讨 论

蚯蚓纤溶酶是一种新的溶栓酶药物, 对其理化特性、活性特点、分离纯化已有大量的报道^[3~5]. 对其基因结构、分子结构、活性机理研究也日渐深入^[6~9]. 目前, 日本和韩国已经克隆了几种蚯蚓纤溶酶的基因. 本课题所得到的蚯蚓纤溶酶 PV₂₄₂来自我国特有的正蚓科双胸蚓属蚯蚓 (*Lumbricus bimastus*), 国际基因库检索证明 PV₂₄₂基因属于新基因, 与同类基因比较同源性不高, 蛋白质的氨基酸序列与同类蛋白同源性仅在 30%~40% 左右.

我们采用融合蛋白表达系统对 PV₂₄₂ 进行表达, 其优点是通过在目的基因的上游连接一个融合伴侣分子 TrxA, 避免了目的基因的结构对载体启动子的影响. 同时, TrxA 是一种能辅助蛋白质正确折叠的分子伴侣, 往往可使不可溶蛋白变为可溶性表达^[10,11]. PV₂₄₂的可溶性表达避免了繁琐的蛋白质复性过程.

总之, 本研究以 SDS-PAGE 分析了我国特有的双胸蚓属中提取的纤溶酶蛋白, 获得其 N 端序列, 并首次克隆到该蚯蚓纤溶酶的基因——PV₂₄₂. 该基因经重组操作后在大肠杆菌中与 TrxA 融合表达. 实验还同时预测了 PV₂₄₂蛋白的结构与特性, 为蚯蚓纤溶酶 PV₂₄₂的进一步分子生物学研究和基因工程开发奠定了基础.

参 考 文 献

- 1 陈 飞, 刘艳玲, 吴红艳, 等. 蚯蚓有效成分研究进展. 微生物学杂志, 2000, 20 (1): 45~46
Chen F, Liu Y L, Wu H Y, et al. Chin Microbiol J, 2000, 20 (1): 45~46
- 2 F. 奥斯伯, R. F. 金斯顿, I. G. 塞德曼, 等著. 颜子颖, 王涪林, 译. 精编分子生物学实验指南. 北京: 科学出版社, 1998. 406~407
Ausubel F M, Kingston R E, Seidman J G, et al. Shot Protocols in Molecular Biology. 3rd. John Wiley & Sons Inc, 1998. 406~407
- 3 程牛亮, 牛 勃, 单鸿仁, 等. 地龙提取物对试验性血栓的溶解作用及有效成分的研究. 山西医学院学报, 1989, 20 (4): 203~206
Chen N L, Niu B, Shan H R, et al. Acta Shanxi Medical University, 1989, 20 (4): 203~206
- 4 路英华, 金汝成, 吴应文, 等. 钜齿远蚓 (*Amyntas dancataa*) 纤溶酶的分离纯化极其若干性质. 生物化学杂志,

- 1988, **4** (2): 166~ 172
- Lu Y H, Jin Y C, Wu Y W, *et al.* Chin Biochem J, 1988, **4** (2): 166~ 172
- 6 Nakajima N, Mihara H, Sumi H. Characterization of potent fibrinolytic enzymes in earthworm, *lumbricus rubellus*. Bioscience Biotech Biochem, 1993, **57** (10): 1726~ 1723
- 7 Nakajima N, Ishihara K, Sugimoto M, *et al.* Chemical modification of eathworm fibrinolytic enzyme with human serum albuin fragment and characterization of the protease as a therapeutic enzyme. Biosci Biotech Biochem, 1996, **60** (2): 293~ 300
- 8 Sugimoto M, Nakajima N. Molecular cloning, sequencing, and expression of cDNA encoding serine protease with fibrinolytic activity from earthworm. Biotech Biochem, 2001, **65** (7): 1575 ~ 1580
- 9 Tang Y, Zhang J, Gui L, *et al.* Crystallization and preliminary X-ray analysis of earthworm fibrinolytic enzyme component a from *eisenia fetida*. Acta Crystallogr D Biol Crystallogr, 2000, **56** (12): 1659~ 1661
- 10 LaVallie E R, DiBlasio E A, Kovacic S, *et al.* A thioredoxin gene fusion expression system that circumvents inclusion body formation in the *E. coli* cytoplasm. Biotechnology (NY), 1993, **11** (2): 187~ 193
- 11 Holmgren A. Thioredoxin. Ann Rev Biochem, 1985, **54**: 237~ 271

Cloning and Expression of The Novel Gene ——PV₂₄₂ of Earthworm Fibrinolytic Enzyme^{*}

XU Yǎ-Hui, LIANG Guo-Dong^{**}, SUN Zhao-Jun, CHEN Fei,
FU Shi-Hong, CHAI Yu-Bo, HOU Yun-De

(The Key Laboratory for Molecular Virology and Genetic Engineering, The Chinese Academy of Prevent Medicine, Beijing 100052, China)

Abstract One kind of protein from earthworm (*Lumbricus bimastus*) with the molecular mass about 30 ku was extracted by SDS-PAGE. Then the N-terminal of the protein was sequenced. The PCR primer was designed according to the N-terminal sequence, and its cDNA fragment was obtained by RT-PCR. The cDNA was cloned into pGEMT-vector and sequenced. The sequence showed that the fragment was 888 bp, with the ORF of 726 bp and the 3' untranslation terminal area. The ORF encoded a protein of 242 amino acids, so the protein was named as PV₂₄₂. The prediction of protein structure shows that the PV₂₄₂ has two domains, between which is the active sites-His44 and Ser191. At the same time, the pI 4.33 of PV₂₄₂ was detected, and it is a protease in trypsin family of serine protease superfamily. PV₂₄₂ was expressed in pTrxFUS expression system as the fusion protein TrxA-PV₂₄₂ in soluble form, the TrxA was a molecular chaperon which made the expressed protein soluble. The fusion protein was purified by ion exchange chromatography. The purified fusion protein TrxA-PV₂₄₂ has fibrinolytic activity.

Key words *Lumbricus bimastus*, fibrinolytic enzyme of earthworm, cloning, recombination and expression

^{*} This work was supported by a grant from State 863 High Technology R & D Project of China (102-08-02-03).

^{**} Corresponding author. Tel: 86-10-63581367, E-mail: gdliang@95777.com

Received: January 29, 2002 Accepted: March 28, 2002