

链霉菌 MY0504 纤溶酶 YG4 基因的克隆 及生物信息学分析

樊丹^{1,2}, 马萱², 麻云莲², 李亚璞², 董超², 史延茂², 张小兵²

(1. 河北工业大学化工学院生物工程系, 天津 300130) (2. 河北省科学院生物研究所, 河北石家庄 050081)

摘要: 本研究获得了海洋链霉菌 MY0504 纤溶酶 YG4 基因, 并对其进行了生物信息学分析。利用依据同源序列设计的兼并引物, 对提取的链霉菌 MY0504 的基因组 DNA 进行 PCR 扩增, 将得到的 DNA 片段连接到 pMD18-T 载体后转化感受态 Trans10 细胞, 然后对阳性克隆进行测序, 并对序列进行生物信息学分析。最终克隆了 1083 bp 的海洋链霉菌 MY0504 纤溶酶 YG4 基因。将获得的 YG4 基因输入 Gen Bank 网站, 进行检索比对, 结果显示与丝氨酸蛋白酶基因碱基序列同源性 100%。对 16s rDNA 序列分析结果表明, 该菌株与达格斯地链霉菌、氯化链霉菌、嫩白黄链霉菌、浅紫链霉菌的亲缘关系较近; 通过对 YG4 基因编码蛋白的一级结构、二级结构、亚细胞定位和三级结构的预测和分析, 结果表明: 该基因编码 360 个氨基酸, 其编码产物为稳定的亲水蛋白; 二级结构以 α 螺旋和 β 折叠为主, 无信号肽和跨膜结构域, 有 40 个磷酸化位点; 高级结构以无规则卷曲为主。本研究结果为研究海洋链霉菌 MY0504 纤溶酶 YG4 基因的表达机制及目的蛋白表达水平的提高提供了重要信息。

关键词: 海洋链霉菌; 纤溶酶; 基因克隆; 生物信息学

文章编号: 1673-9078(2019)03-65-72

DOI: 10.13982/j.mfst.1673-9078.2019.3.011

Cloning and Bioinformatics Analysis of Fibrinolytic Enzyme YG4 Gene from *Streptomyces* sp. MY0504

FAN Dan^{1,2}, MA Xuan², MA Yun-lian², LI Ya-pu², DONG Chao², SHI Yan-mao², ZHANG Xiao-bing²

(1.College of Chemistry and Engineering Hebei University of Technology, Tianjin 300130, China)

(2.Institute of Biology Hebei Academy of Sciences, Shijiazhuang 050081, China)

Abstract: In this study, the fibrinolytic enzyme YG4 gene from marine *Streptomyces* sp. MY0504 was obtained and analyzed by bioinformatics methods. The genomic DNA of the extracted *Streptomyces* MY0504 was amplified by PCR using the degenerate primers designed based on the homologous sequences. The obtained DNA fragments were ligated to the pMD18-T vectors and transformed into competent Trans10 cells. The positive clones were then sequenced, and the obtained sequences were analyzed bioinformatically. Finally, the fibrinolytic enzyme YG4 gene from marine *Streptomyces* strain MY0504 sized 1083 bp was cloned. The obtained YG4 gene was introduced into the GenBank website, and then an alignment search and comparison was performed. The result showed that the YG4 gene exhibited 100% homology with the base sequence of the serine protease gene. The results of the 16s rDNA sequence analysis showed that the strain was closely related to *Streptomyces daghestanicus*, *Streptomyces hydrogenans*, *Streptomyces albidoflavus* and *Streptomyces violascens*. The primary structure, secondary structure, subcellular localization and tertiary structure of the protein encoded by YG4 gene were predicted and analyzed. The analysis results showed that the gene encoded 360 amino acids, and the encoded product was a stable hydrophilic protein. The secondary structure was mainly composed of alpha helix and beta sheet without signal peptides and transmembrane domains but with 40 phosphorylation sites. The high-order structure was dominated by random coils. The results of this study provided important information for investigating the expression mechanism of the fibrinolytic enzyme YG4 gene from marine *Streptomyces* MY0504 and for improving the expression level of a target protein.

Key words: marine *Streptomyces*; fibrinolytic enzyme; gene cloning; bioinformatics

收稿日期: 2018-10-09

基金项目: 河北省科学院科技计划项目 (15303; 16303; 17303)

作者简介: 樊丹 (1992-), 女, 硕士研究生, 研究方向: 分子生物学

通讯作者: 史延茂 (1962-), 研究员, 研究方向: 发酵工程; 共同通讯作者: 张小兵 (1973-), 男, 博士, 高级工程师, 研究方向: 分子生物学

心脑血管疾病严重危害人类健康,以其高发病率、高致残率、高死亡率,给家庭、社会和国家都带来极其严重的精神压力和经济负担,血栓栓塞症是其主要一类^[1],此类疾病主要包括三种类型:急性心肌梗死、脑血栓中风、末梢动静脉形成血栓,即脉管栓塞。随着人类生活水平和医疗条件的改善和提高,人口平均寿命日益延长,此外,饮食习惯和生活环境改变,致使心脑血管疾病的发病率也随之逐渐提高。目前,被广泛用于治疗血栓病的疗法是溶栓疗法^[2,3],溶栓剂主要分为两种类型,一种是纤溶酶原激活物,如组织型纤溶酶原激活物(t-PA)和尿激酶(u-PA);另一种是纤溶酶类物质,它们可以直接作用于血纤维蛋白,从而迅速溶解血栓,如纳豆激酶、蚓激酶等。u-PA 和 t-PA 仍广泛应用于溶栓治疗,但这些药物副作用较大,容易导致出血倾向,价格较贵,半衰期短,口服无效,而开发新型高效无副作用的溶栓制剂迫在眉睫^[4]。

随着海洋空间的开发,海洋技术手段的发展,海洋微生物日益成为研究的焦点^[4,5],而放线菌早已被证实能产生种类繁多的活性物质,因此,海洋来源放线菌逐渐成为生产新的天然产物的最具潜力的类群,其产生的代谢物质很可能更加符合人类对有效药物的需求^[6,7]。链霉菌分泌多种蛋白水解酶,其中相当一部分属于丝氨酸蛋白酶,而人血液中的纤溶酶及纤溶酶原激活剂也均属于丝氨酸蛋白酶^[8,9]。因此,对该种菌产生的活性物质研究具有重要的现实意义,利用链霉菌产生的新型纤溶酶制备溶栓药物,亦是一种新的尝试。本研究基于以前的研究结果^[10~12],克隆了链霉菌 MY0504 的纤溶酶 YG4 基因(以下简称 MY-YG4),

再借助相关生物信息学知识及相关软件,分析预测该基因编码蛋白的理化性质、结构和功能等,以期为后续该基因的功能研究奠定基础。

1 材料与方法

1.1 材料

1.1.1 菌株与质粒

菌株 MY0504,由本实验室自渤海海域海水中分离,菌株已保藏在中国微生物菌种保藏管理委员会普通微生物中心(CGMCC),保藏编号 No.10664; pMD18-T 载体购 TaKaRa 公司。

1.1.2 培养基

LB 培养基(g/L):胰蛋白胨 10 g,酵母抽提物 5 g,氯化钠 10 g。

1.1.3 实验仪器

恒温振荡器,上海苏坤设备有限公司;TC-96/G/H (b) C 型基因扩增仪器,杭州博日科技有限公司;VS-1300L-U 洁净工作台,苏净安泰;H2O3-PRO III 加热制冷型金属浴,卡尤迪生物科技;DYY-7C 型电泳仪,北京市六一仪器厂;FE28 PH 计,梅特勒-托利多仪器(上海)有限公司。

1.2 方法

1.2.1 菌株的培养与基因组 DNA 提取

挑取斜面菌体接种于液体 LB 培养基中,于 37 ℃ 200 r/min 摇床培养 48 h。DNA 的提取采用 TIANGEN 科技有限公司的“细菌基因组 DNA 提取试剂盒”^[13]。

表 1 蛋白质测序结果

Table 1 The result of protein sequencing

登录号	描述	得分
A0A0D6WUN7	Streptogrisin OS=Streptomyces sp. MBRL 601 GN=SF12_07495 PE=4 SV=1-[A0A0D6WUN7_9ACTN]	1143.41
A0A0S1UK64	Serine protease OS=Streptomyces sp. FR-008 GN=SFR_2227 PE=3 SV=1-[A0A0S1UK64_9ACTN]	689.23
A0A0S1UHF4	Peptidase M28 OS=Streptomyces sp. FR-008 GN=SFR_1292 PE=4 SV=1-[A0A0S1UHF4_9ACTN]	200.33

表 2 PCR 引物序列

Table 2 Nucleotide sequences of PCR primers

Primer	Primer sequences(5'-3')
Pla-BamH I -F	GGATCCGTGAACCAACCGACGCATACCCAAGC
Pla-Xho I -R	CTCGAGTCAGCCGATCTCGACGCCGTACGCG

注:下划线为酶切位点 BamH I (GGATCC)和 Xho I (CTCGAG)。

1.2.2 PCR 扩增

据北京蛋白质中心所测得的蛋白质氨基酸序列信息(表 1),在 UniProt 网站搜索获得相关纤溶酶的氨基酸序列,根据氨基酸序列检索到相应的基因 CDS 区序列;通过比对所获得 CDS 区序列,进行全序列相

似性分析,在相似性高的基因位置设计引物,并以 Prime Primer 5.0 软件和 DNAMAN 软件辅助设计简并引物^[14](表 2)。以基因组 DNA 为模板,PCR 扩增条件为:94 ℃ 预变性 5 min,94 ℃ 变性 45 s,退火温度 53 ℃ 退火 35 s,72 ℃ 延伸 1 min,32 个循环后,72 ℃

延伸 8 min, 4 °C 保存。

1.2.3 序列测定

扩增产物经 1%琼脂糖凝胶电泳检测, 将与预期结果大小相符的 PCR 扩增片段回收, 回收片段与克隆载体 pMD18-T 连接, 16 °C 过夜连接, 转化 Trans10 感受态细胞, 进行菌液 PCR 验证, 阳性样品送北京华大基因测序。

1.2.4 生物信息学分析

通过 NCBI 的 ORF finder 及 DNAMAN 对基因的 ORF 及氨基酸序列进行预测; 通过 Ex PASy 数据库 (<http://www.Expasy.org/>) 提供的 Prot Param 在线程序, 预测该基因编码蛋白的氨基酸组成、等电点、正/负电荷残基数、原子总数、摩尔消光系数、半衰期、不稳定指数、脂肪族氨基酸指数、总平均亲水性等; 通过 Prot Scale 软件预测疏水性。通过 CBS Prediction Servers 数据库 (<http://www.cbs.dtu.dk/services/>) 提供的 Signal P 4.1 Server^[15,16]在线程序, 判断是否存在信号肽; 通过 TMHMM Server v. 2.0^[17]软件, 预测跨膜结构区域; 通过 Protfun 2.2 Server 软件, 预测编码蛋白的功能分类。利用 NetPhos 3.1 Server 软件对该基因编码的蛋白磷酸化位点进行预测和分析。通过 Predict Protein 软件 (<https://www.predictprotein.org/>), 预测蛋白质的二级结构。应用 NCBI conserved domains (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 预测蛋白结构域。通过 SWISS-MODEL, 对该编码蛋白的三维结构进行同源建模。

2 结果与分析

2.1 YG4 基因克隆

根据 NCBI 网站检索到的纤溶酶序列, 预计 *Streptomyces* sp. MY0504 纤溶酶 YG4 基因的大小为 900~1200 bp。通过“细菌基因组 DNA 提取试剂盒”提取链霉菌基因组 DNA, 经 Nanodrop 和电泳检测, 表明所提取的 DNA 质量较高, 可用于目的基因扩增的模板制备。应用 Pla-BamH I -F 上游引物及 Pla-Xho I -R 下游引物, 进行 YG4 基因扩增, 获得 1083 bp 的特异性条带 (图 1), 经 PCR 鉴定的阳性菌送至北京华大进行测序。测序结果表明: 扩增出的 YG4 基因与 Gen Bank 上发表的丝氨酸蛋白酶基因比较, 碱基序列同源性 100%。经 1%琼脂糖凝胶电泳检测, 与预期大小一致。拼接阳性克隆的测序结果, 最终所获得目标序列 (图 5)。

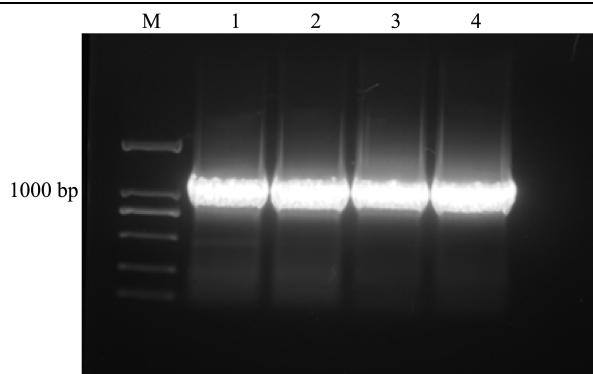


图 1 YG4 基因的扩增

Fig.1 YG4 amplification

注: M: Trans 2K DNA Marker; 1~4: PCR 产物。

2.2 16S rDNA 鉴定亲缘关系

用通用引物 27 F 和 1492 R、1525 R, 扩增链霉菌 MY0504 的 16S rDNA (图 2)。扩增片段大小为 1500 bp 左右, 无明显杂带、拖带现象。扩增获得的 16S rDNA 序列如图 3。

将获得的 16S rDNA 序列, 提交 NCBI 数据库 (<https://submit.ncbi.nlm.nih.gov/subs/genbank/>), 获得登录号: MG203935。用所获得的 16S rDNA 序列, 在 NCBI 和 EzBioCloud 网站 Blast, 应用 MEGA (MEGA 5.1) 软件进行分子进化遗传分析, Neighbor-Joining 法完成分子系统学分析, 同时进行 1000 次 bootstrap 统计学检验, 构建系统发育树 (图 4), 展示细菌的遗传多样性和菌株间的亲缘关系^[18]。结果表明, 该菌株与达格斯地链霉菌、氢化链霉菌、嫩白黄链霉菌、浅紫链霉菌的亲缘关系较近。

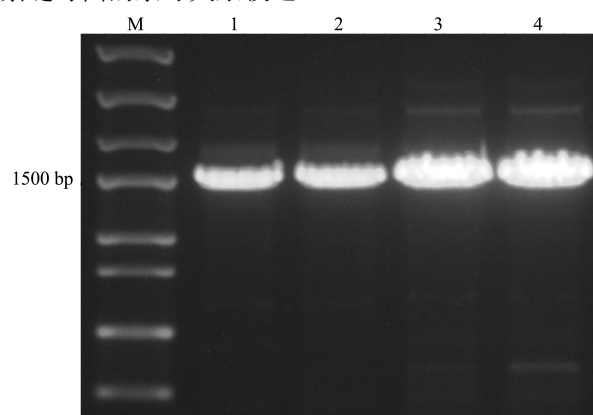


图 2 链霉菌 MY0504 16S rDNA 电泳图谱

Fig.2 Electrophoretogram of *Streptomyces* sp. MY0504 16S rDNA gene

注: M: DNA Marker: trans 5K; 1~2: 引物 27F-1492R; 3~4: 引物 27F-1525R。

1	GACGAAACGCT	GGCGGCGTGC	TTAACACATG	CAAGTCGAAC	GATGAAOCGC	TTTGGGGCGG
61	GGATTAGTGG	CGAACGGGTG	AGTAACACGT	GGGCAATCTG	CCCTGCACCTC	TGGGCAAGC
121	CCTGGAACG	GGGTCTAATA	CCGGATATGA	CTGTCCATCG	CATGTTGGAT	GGTGTAAAGC
181	TCCGGCGGTG	CAGGATGAGC	CCGGGGCTTA	TCAGCTTGTT	GGTGAGGTAG	TGGCTCAGCA
241	AGGCGACGAC	GGGTAGCCGG	CCTGAGAGGG	CGAOCGGCCA	CACTGGGACT	GAGACACGGC
301	CCAGACTCCT	ACGGGAGGCA	GCAGTGGGA	ATATTGCACA	ATGGCGGAAA	GCCTGATGCA
361	GCGACGCGC	GTGAGGGATG	ACGGCCTTCG	GGTGTAAAC	CTCTTTCAGC	AGGGAAGAAG
421	CGAAGTGAC	GGTACCTGCA	GAAGAAGCGC	CGGCTAACTA	CGTGCCAGCA	GCCGCGGTAA
481	TACGTAGGGC	GCAAGCGTTG	TCCGGAATTA	TGGGCGTAA	AGAGCTGTGA	GGCGGCTTGT
541	CACGTCCGTT	GTGAAAGCCC	GGGGCTTAAC	CCCGGGTCTG	CAGTCGATAC	GGGCAGGCTA
601	GAGTTCGGTA	GGGGAGATCG	GAATTCCTGG	TGTAGCGGTG	AAATGCGCAG	ATATCAGGAG
661	GAACACCGGT	GGCGAAGGCG	GATCTCTGGG	CCGATACTGA	CGCTGAGGAG	CGAAAGCGTG
721	GGGAGCGAAC	AGGATTAGAT	ACCTCTGGTAG	TCCACGCGGT	AAAAGTGGG	CACTAGGTGT
781	GGGCAACATT	CCAOGTTGTC	CGTGCCGAG	CTAACGCATT	AAGTGCCCCG	CCTGGGGAGT
841	ACGCGCGCAA	GGCTAAACT	CAAAGGAATT	GACGGGGGCC	CGCACAGCG	GCGGAGCATG
901	TGGCTTAATT	CGAOCGAACG	CGAAGAACCT	TACCAAGGCT	TGACATACAC	CGGAAACGTC
961	TGGAGACAGG	CGCCCCCTTG	TGGTCGGTGT	ACAGGTGGTG	CATGGCTGTC	GTCAGCTCGT
1021	GTCGTGAGAT	GTGGGGTAA	GTCCCGCAAC	GAGCGCAACC	CTTGTCOCGT	GTTGCCAGCA
1081	GGCCCTTGTG	GTGCTGGGGA	CTCACGGGAG	ACCGCCGGGG	TCAACTGCGA	GGAAGGTGGG
1141	GACGACGTCA	AGTCATCATG	CCCTTATGT	CTTGGGCTGC	ACAOGTGCTA	CAATGGCCGG
1201	TACAATGAGC	TGCGATACCG	CGAGGTGGAG	CGAATCTCAA	AAAGCCGCTC	TCAGTTGCGA
1261	TTGGGGTCTG	CAACTCGACC	CCATGAAGTC	GGAGTCGCTA	GTAATCGCAG	ATCAGCATTG
1321	CTGCGGTGAA	TACGTTCCCG	GGCCTTGAC	ACAOCGCCCG	TCAOGTCACG	AAAGTCGGTA
1381	ACAACCGAAG	CCGGTGCCCC	AACCCCTTGT	GGGAGGGAGC	TGTGAAAGGT	GGGACTGGCG
1441	ATTGGGACGA	AGTOGTAACA	AGGTAGCCGT	ACCGGAAGG		

图3 16S rDNA 序列

Fig.3 The sequences of 16S rDNA

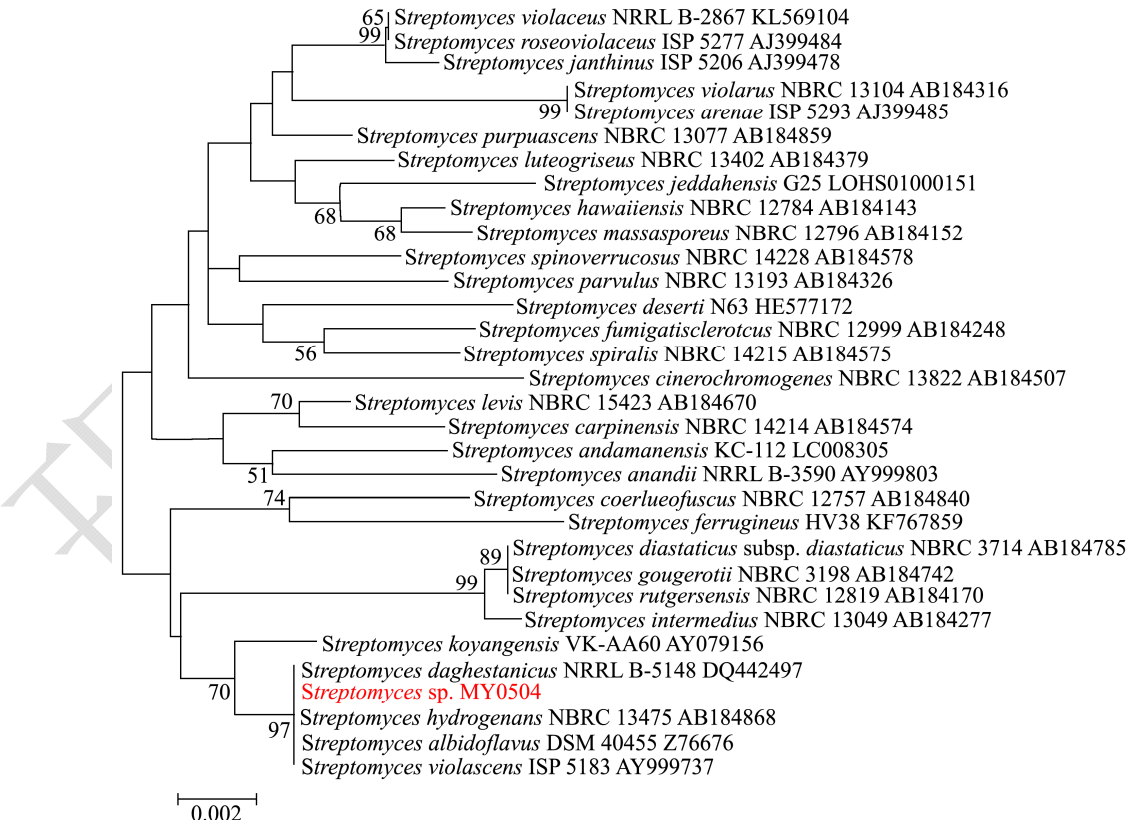


图4 基于 *Streptomyces* sp. MY0504 和相关菌株的 16S rDNA 序列构建的系统发育树

Fig.4 The phylogenetic tree based on 16S rDNA sequences of *Streptomyces* sp. MY0504 and related strains

注：建树采用 Neighbor-Joining 法，并进行 1000 次 bootstrap 统计学检验。

```

1  GTGAACCAACGACGATACCAAGCGCGGGTCGCCGTGACCGGTGCCGGCATCGCCGCACTGGTCGCCGCGGG
1  V N H R R I P K R R V A V T G A G I A A L V A A G
85 95 105 115 125 135 145
76 CTGACGTTCCAGACGGCCAACGCCAGCGAGCGCGCAACCGCGCACCTCTCGGTGCGCTCCGCGGGC
26 L T F Q T A N A S E P Q S A E P R T L S V A S A G
160 170 180 190 200 210 220
151 AAGCTCGCCGACACCTCGCCGGTGACCTCGGCAAGCGCGCGCGGACGTACTACGACGCGCAAGCGCAAGAAC
51 K L A D T L A G D L G K A A A G T Y Y D A K A K N
235 245 255 265 275 285 295
226 CTGGTGGTGAACGTCCTCGACGAGGCGCGCCACCGAGGTGAGGCGTCCGGTGCCAGGCCAGAATCGTCGAG
76 L V V N V L D E A A A T E V E A S G A Q A K I V E
310 320 330 340 350 360 370
301 CACTCCCTCACC CGCTGAACGACGCGCGCGGCCACCTCAAGGACCGGCCAGCATCGCGGGCACCTCGTGGGCG
101 H S L T A L N D A R A T L K D R A S I A G T S W A
385 395 405 415 425 435 445
376 ATCGACCGGACGACCAAGGTCGTCTGTCACCGCCGACCGCACGGTCGAGGCGAAGGAGCTGGCGAAGCTCACC
126 I D P T S N K V V V T A D R T V E G K E L A K L T
460 470 480 490 500 510 520
451 AAGGTGGTGGACGCGCTCGGTTCGAAGCGCAACTCAAGCAGACCAAGGTTGAGTTCAAGCGGTTTCATCGCGGT
151 K V V D A L G S K A E L K Q T K G E F K R F I A G
535 545 555 565 575 585 595
526 GCGGACGCCATCCACGCGGACGCGCGCGCTGCTCCCTCGGCTTCAACGTGGTGAAGGCGCGGACGCGTACTTC
176 G D A I H G D G G R C S L G F N V V K G G E P Y F
610 620 630 640 650 660 670
601 CTCACCGCGGTCACTGCACCGAGGGGATCTCCAGCTGGTCCGAGACCTCGGGCGGTCGGAGATCGGCACACG
201 L T A G H C T E G I S S W S E T S G G S E I G T T
685 695 705 715 725 735 745
676 GTGAGTCAAGCTTCCCGCGGACGACTACGGGTGGTCCAGTACACCGGCTCCACCGAGCACCCACGAGGTG
226 V E S N F P G D D Y G L V Q Y T G S T E H P S E V
760 770 780 790 800 810 820
751 AACCTGTACGACGCGACCCAGGCCATCACCGGTGCCGGTGACGCCACCGTCGGCCAGTCCGGTGACCGCTCC
251 N L Y D G S T Q A I T G A G D A T V G Q S V T R S
835 845 855 865 875 885 895
826 GGCTCCACCGAGGTCTCCACCGGAGGTACGCGGCTCGACGCCACCGTGAACACGCGCGCGGACGTC
276 G S T T Q V S T G E V T A L D A T V N Y G G G D V
910 920 930 940 950 960 970
901 CTCGAGGGCTGATCCAGACCACCGTCTCGCGGAGCGCGGACAGCGCGGTTCCGCTCTTCGACGCGCTCCAG
301 V E G L I Q T T V C A E P G D S G G S L F D G S T
985 995 1005 1015 1025 1035 1045
976 GCGGTGGGCTACCTCCGCGGCGAGCGGTGACTGCTCTCGGGCGGAGACCTTCTCCAGCGGTGCCGGAG
326 A L G L T S G G S G D C S S G G E T F F Q P V P E
1060 1070 1080
1051 GCGCTCTCCGCGTACGCGCTCGAGATCGGCTGA
351 A L S A Y G V E I G *

```

图5 MY-YG4基因全长及编码氨基酸序列预测

Fig.5 Complete DNA and deduced encoding amino acid sequences of MY-YG4 gene

2.3 生物信息分析

2.3.1 MY-YG4基因的ORF及其氨基酸序列的预测

按照 NCBI 数据库及 DNAMAN 软件处理所得结果, MY-YG4 基因的最大 ORF 从第一个碱基开始, 到第 1083 个碱基终止, 即所有序列都为编码区, 共编码 360 个氨基酸 (图 5)。

2.3.2 一级结构及其理化性质分析

利用 Prot Param 软件, 分析链霉菌 MY0504 纤溶酶 YG4 (MY-YG4) 基因编码蛋白的氨基酸组成 (图 6) 及其理化性质。链霉菌 MY0504 纤溶酶 YG4 (MY-YG4) 基因共编码 360 个氨基酸, 其中甘氨酸 (Gly) 含量最高, 占全部氨基酸的 13.90%。带负电氨基酸 (Asp+Glu)、带正电氨基酸 (Arg+Lys) 残基的数量分别为 43 和 27 个。分子式为 $C_{1569}H_{2503}N_{443}O_{543}S_4$, 原子总数为 5062, 相对分子质量为 36389.02 u, 理论等电点为 4.74, 是酸性蛋白。在人体内正常生理环境下, 该酶带正电, 而纤维蛋白通常带负电荷, 所以链霉菌纤溶酶很容易靠近纤维蛋白, 并与之结合。因此, 链霉菌纤溶酶对血栓有一定的靶向作用。在 280 nm 波长下, 所有半胱氨酸形成胱氨酸 (cystines) 时的消

光系数为 23170 L/(mol·cm), 对应的吸光度为 0.637; 所有半胱氨酸均不形成胱氨酸时的消光系数为 22920 L/(mol·cm), 对应的吸光度为 0.630。哺乳动物网织红细胞的半衰期为 100 h, 不稳定指数为 26.16, 表明该基因翻译出来的纤溶酶是一个稳定蛋白 (不稳定系数小于 40 时为稳定蛋白), 脂肪族氨基酸指数为 77.83。

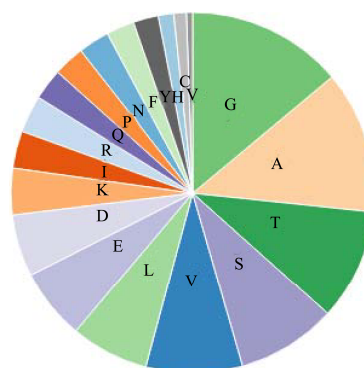


图6 MY-YG4基因编码蛋白的氨基酸组成

Fig.6 Amino acid composition of MY-YG4 gene coding protein

2.3.3 亲水性/疏水性、跨膜结构

蛋白质的折叠主要由氨基酸的亲、疏水性驱动, 是每种氨基酸固有的特性^[19]。蛋白质在折叠时形成疏水的内核和亲水的表面, 同时潜在跨膜区会出现高疏水性结构域, 通过对亲疏水性分析可以反映蛋白质表

面氨基酸的分布和跨膜结构域^[20]。

通过 ProtScale 程序分析,预测 MY-YG4 基因编码氨基酸序列的亲水性/疏水性,结果如图 7 所示。多肽链的第 22 位具有最大值 2.567,疏水性最强;第 6 位存在最小值-2.856,为亲水性氨基酸。平均疏水性通过理化性质分析显示为-0.141,在整条肽链中,亲水氨基酸数量较多,表明整条多肽链表现为亲水性,推测该蛋白为可溶性蛋白。

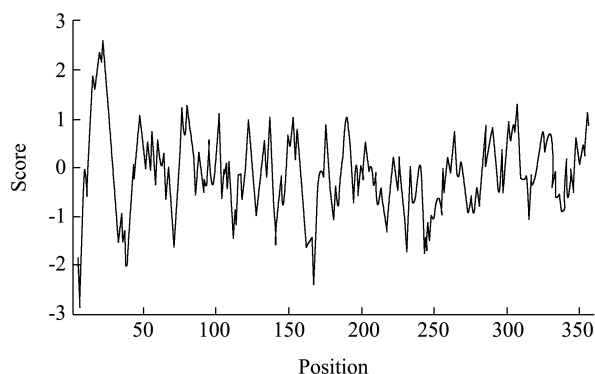


图 7 MY-YG4 基因编码蛋白的疏水性和亲水性

Fig.7 The hydrophobicity/hydrophilicity of MY-YG4 gene coding protein

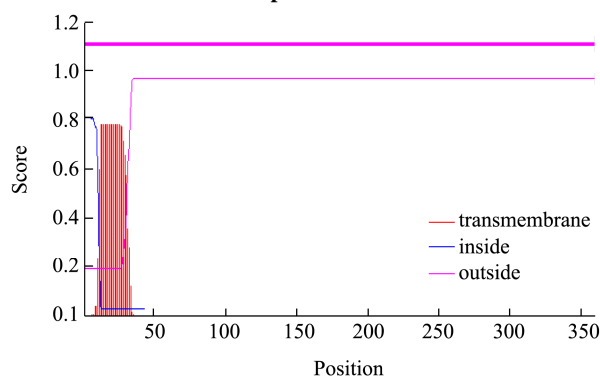


图 8 MY-YG4 基因编码蛋白的跨膜区预测结果

Fig.8 The transmembrane domain of MY-YG4 gene coding protein

跨膜区必须由强疏水的氨基酸组成,才能使膜蛋白穿过膜的磷脂双分子层。通过蛋白亲、疏水性分析发现,该蛋白为亲水性蛋白,推测不存在跨膜区。进一步利用 TMHMM 程序对该蛋白跨膜区进行了分析,结果如图 8 所示。表明确实不存在跨膜区,这与亲水性的分析的结果是一致的。

2.3.4 信号肽

信号肽是引导前体蛋白质通过细胞膜分泌到胞外的一段序列,对其预测和分析有助于了解蛋白质的细胞定位并区分蛋白质的功能域^[21]。分析发现,纤溶酶蛋白典型的信号肽的 C 值和 Y 值趋向于 1,且 S 值在切割位点之前高,而在切割位点之后降低。信号肽预测结果如图 9,该蛋白最高原始剪切位点分值 (C)、

最高信号肽分值 (S) 以及最高综合剪切位点分值 (Y) 分别为 0.467、0.420、0.275,该结果不具备信号肽的要求。该蛋白不存在信号肽,为非分泌蛋白。

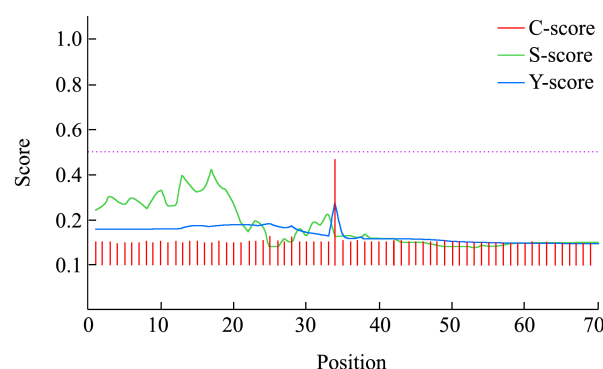


图 9 MY-YG4 基因编码蛋白的信号肽预测结果

Fig.9 The signal peptide of MY-YG4 gene coding protein

2.3.5 磷酸化位点预测与分析

利用 NetPhos 3.1 Server 软件对该基因编码的蛋白磷酸化位点进行预测和分析(图 10)。结果表明,多肽链分值大于 0.5 的有 40 个氨基酸位点,说明该基因有 40 个磷酸化位点,且在多肽链中分布不均匀。其中,有 23 个丝氨酸残基(Ser)可能发生磷酸化;有 14 个苏氨酸残基(Thr)可能发生磷酸化;有 3 个酪氨酸残基(Tyr)可能发生磷酸化。

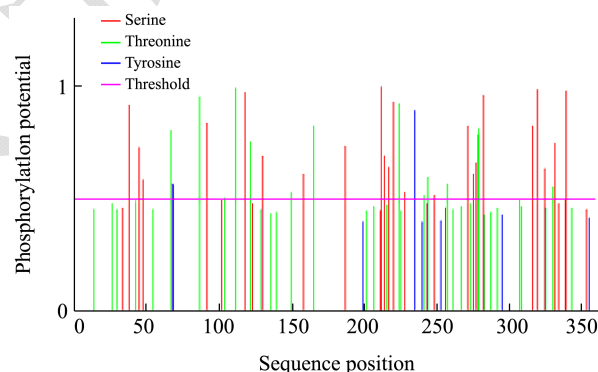


图 10 MY-YG4 磷酸化位点预测与分析

Fig.10 Phosphorylation site prediction and analysis of MY-YG4 amino acid

2.3.6 二级结构

蛋白质分子的多肽链通常折叠和盘绕,形成比较稳定的空间结构,具有特有的生物活性和理化性质。因此,蛋白质二级结构的预测和分析对其空间结构的了解有着重要意义预测。常见的二级结构元件主要有 α -螺旋 (Alpha-helix)、 β -转角 (Beta-turn)、 β -折叠 (Beta-sheet)、无规则卷曲 (Random coil) 以及伸展链 (Extended strand) 等。通过 SOPMA 对纤溶酶蛋白二级结构进行预测,结果如图 11 所示,该蛋白中存在 114 个无规则卷曲占 31.67%、112 个 α -螺旋占 31.11%、38 个 β 转角占 10.56%、96 个伸展链占 26.67%。

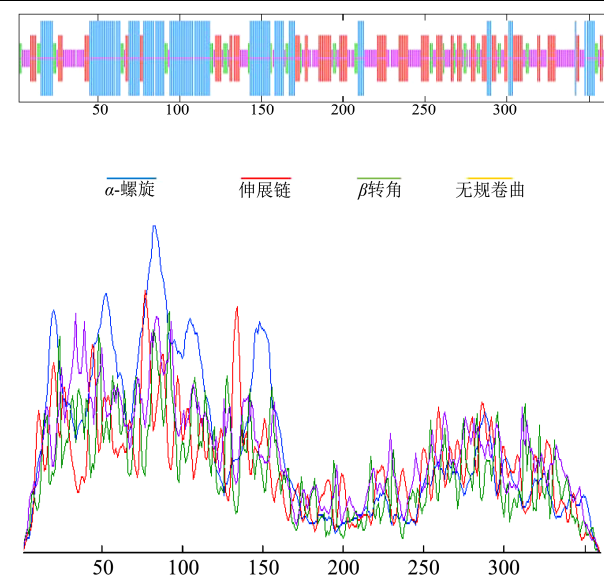


图 11 MY- YG4 基因编码蛋白的折叠盘绕方式

Fig.11 The folded and coiled ways of MY-YG4 gene coding protein

2.3.7 三级结构

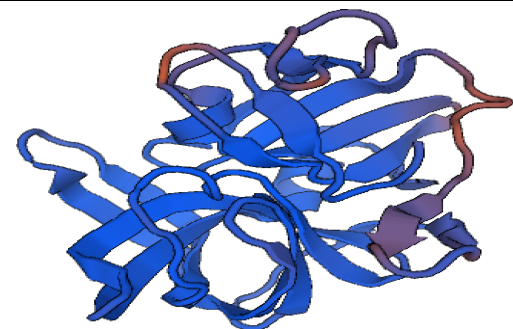


图 12 MY- YG4 基因编码蛋白的三级结构

Fig.12 The tertiary structure of MY-YG4 gene coding protein

将氨基酸序列提交 SWISS-MODEL，同源建模得到蛋白质的三维结构，如图 12 所示。该蛋白存在较多的卷曲和螺旋，结构较为丰富，而这些结构对其生物学功能的发挥有重要作用。

2.4 其他纤溶酶的特征与功能

通过比较其他纤溶酶的特征与功能，我们总结出链霉菌纤溶酶有很大的开发和利用空间（表 3）。

表 3 不同纤溶酶的特征与功能

Table 3 Characteristic and function of different plasmin

细菌名称	细胞壁结构	菌株分类	纤溶酶名称和溶栓功能方式	优缺点	与人类关系	应用开发状态
β-溶血链球菌	革兰氏阳性	厚壁菌门 芽孢杆菌纲 乳杆菌目 链球菌科 链球菌属	链激酶 streptokinase, 与纤溶酶原结合使成为纤溶酶	价格低廉，易引起免疫反应	致病菌	以重组链激酶为主
金黄色葡萄球菌	革兰氏阳性	厚壁菌门 芽孢杆菌纲 芽孢杆菌目 葡萄球菌科 葡萄球菌属	葡激酶 staphylokinase, 与纤溶酶原结合使成为纤溶酶	纤维蛋白特异性强，有一定的免疫反应	致病菌	基因工程技术，蛋白融合等技术
纳豆芽孢杆菌	革兰氏阳性	厚壁菌门 芽孢杆菌纲 芽孢杆菌目 芽孢杆菌科 芽孢杆菌属	纳豆激酶，可以直接水解纤维蛋白，也有激活纤溶酶原的作用	溶栓功能较多，使用安全，纯化后酶制剂功能研究较少	非致病菌	口服保健食品较多，尚未药物开发
产芽孢梭菌	革兰氏阳性	厚壁菌门 梭菌纲 梭菌科 梭菌目 梭菌属	可以直接水解纤维蛋白，也有激活纤溶酶原的作用	/	致病菌	未开发
链霉菌	革兰氏阳性	放线菌门 放线菌纲 放线菌亚纲 放线菌目 链霉菌亚目 链霉菌科 链霉菌属	可以直接水解纤维蛋白，多不能激活纤溶酶原。	代谢产物特殊，放线菌代谢产物复杂	/	未开发

3 结论

3.1 血栓性疾病严重威胁着人类的健康，溶栓疗法是目前最安全有效的方法之一。虽然随着分子生物学技术的发展和药物筛选手段的不断提高，各种不同来源的溶栓药物在治疗血栓病上均已取得了很大的进展，但都还存在着一些缺陷。由于利用微生物进行生产，具有周期短、产量高、生产工艺简单、成本低等特点，从微生物寻找新型天然来源的溶栓药已经成为生物医药发展的一个重要方向。海洋生物的高效代谢、繁殖

及具有特殊防御方法的生存方式，这些活动都与其体内的各种生理活性物质相关，而且这些活性物质极有可能有重要的药理作用，所以我们推测海洋链霉菌的代谢物质很可能更加符合人类对有效药物的需求。当然溶栓药物的研究不仅要从自然界中筛选目的菌，更重要的是用分子生物学和基因工程技术对菌种进行改造，以提高溶栓药物的应用专一性。

3.2 本研究主要从海洋链霉菌 MY0504 克隆了纤溶酶 YG4 (MY-YG4) 基因，经 NCBI 网站 BLAST 搜索后，发现与马怡茗^[22]等报道的角蛋白酶基因 gm2886

(GenBank Accession Number: KY368946) 序列完全一致。多种底物检测表明, 重组蛋白 GM2886-His6 具有蛋白酶活性, 可以降解水不溶性的天青角蛋白和羽毛粉。

3.3 研究发现已知蛋白的新功能, 对蛋白的开发利用意义巨大。于是, 我们对 MY-YG4 基因进行了生物信息学分析。结果表明, 该基因编码的蛋白由 360 个氨基酸组成, 属于丝氨酸蛋白酶, 为稳定的亲水蛋白, 无信号肽和跨膜结构域, 高级结构以无规则卷曲为主。该蛋白三级结构较为丰富, 而这些结构对其生物学功能的发挥有重要作用。这些性质对该基因及其家族的结构和生物学功能的研究, 奠定了理论基础。但是为了获取更准确的研究结果, 仍须克隆验证, 因此关于该基因的分子克隆和功能鉴定, 我们还在进行更深层次的试验和研究。

参考文献

- [1] 鲁艳丽, 宁喜斌. 血栓形成机理及溶血栓药物的研究进展[J]. 食品研究与开发, 2006, 1: 169-172
LU Yan-li, NING Xi-bin. The thrombus forms the mechanism and dissolves the thrombus medicine the research to progress [J]. Food Research and Development, 2006, 1: 169-172
- [2] Narasimhan M K, Chandrasekaran M, Rajesh M. Fibrinolytic enzyme production by newly isolated *Bacillus cereus* SRM-001 with enhanced *in vitro* blood clot lysis potential [J]. J Gen Appl Microbiol, 2015, 61(5): 157-164
- [3] Banerjee A, Chisti Y, Banerjee U C. Streptokinase-a clinically useful thrombolytic agent [J]. Biotechnol Adv, 2004, 22(4): 287-307
- [4] 郝树站, 王素英. 海洋微生物生物活性物质的研究进展[J]. 生命科学研究, 2005, S1: 35-38
HAO Shu-zhan, WANG Su-ying. The sea microorganism biological activity material research progresses [J]. Life Sciences Research, 2005, S1: 35-38
- [5] 田新朋, 张恩, 李文均. 海洋放线菌研究进展[J]. 微生物学报, 2011, 51(2): 161-169
TIAN Xin-peng, ZHANG Si, LI Wen-jun. The sea ray fungi research progresses [J]. Microorganism Journal, 2011, 51(2): 161-169
- [6] SV, SSS, OSA, et al. Antimicrobial potential of *Actinomycetes* species isolated from marine environment [J]. Asian Pacific Journal of Tropical Biomedicine, 2012, 6: 469-473
- [7] Elich E, Schreinemakers P, Vullings M, Partha N. Ultrasound induced production of thrombinase by marine actinomycetes: kinetic and optimization studies [J]. Biochemical Engineering Journal, 2012, 61(8): 34-42
- [8] Zotchev S B. Marine actinomycetes as an emerging resource for the drug development pipelines [J]. J Biotechnol, 2012, 158(4): 168-175
- [9] Ju X, Cao X, Yong S, et al. Purification and characterization of a fibrinolytic enzyme from *Streptomyces* sp. XZNUM 00004 [J]. World J Microbiol Biotechnol, 2012, 28(7): 2479-2486
- [10] 侯正欣, 董超, 马萱, 等. 海洋来源链霉菌 MY0504 产纤溶酶的发酵条件优化[J]. 微生物学通报, 2017, 5: 1009-1016
HOU Zheng-xin, DONG Chao, MA Xuan, et al. Sea origin streptomycete MY0504 produces the plasmin the fermentation condition to optimize [J]. Microbiology Notification, 2017, 5: 1009-1016
- [11] 米阳, 董超, 侯正欣, 等. 海洋链霉菌发酵纤溶酶的分离纯化和酶学性质研究[J]. 中国海洋药物, 2016, 35(3): 43-48
MI Yang, DONG Chao, HOU Zheng-xin, et al. The sea streptomycete fermentation plasmin separation purification and the zymology nature study [J]. Chinese Sea Medicine, 2016, 35(3): 43-48
- [12] 董超, 米阳, 原晋波, 等. 产纤溶酶海洋放线菌的筛选及初步鉴定[J]. 中国酿造, 2015, 7: 59-64
DONG Chao, MI Yang, YUAN Jin-bo, et al. Produces the plasmin sea ray fungi screening and appraises [J]. China Brews, 2015, 7: 59-64
- [13] Xiang L, Moore BS. Characterization of benzoyl coenzyme a biosynthesis genes in the enterocin-producing bacterium "*Streptomyces maritimus*" [J]. Journal of Bacteriology, 2003, 185(2): 399-404
- [14] 敬俊锋, 陈斌, 李莹, 等. 纳豆激酶基因的克隆及其在毕赤酵母中的表达[J]. 生物学杂志, 2011, 5: 55-57
JIN Jun-feng, CHEN Bin, LI Ying, et al. Accepts the bean activating enzyme gene the clone and in finishes in the red yeast expression [J]. Biology Magazine, 2011, 5: 55-57
- [15] 王志坤, 常健敏, 李丹丹, 等. 大豆 *GmWRI1a* 基因克隆及生物信息学分析[J]. 东北农业大学学报, 2013, 7: 11-16
WANG Zhi-kun, CHANG Jian-min, Li Dan-dan, et al. Soybean *GmWRI1a* gene clone and biological information study analysis [J]. Northeast Agricultural College Journal, 2013, 7: 11-16

(下转第 147 页)