

海洋链霉菌差向异构结构域基因的克隆与分析

吴欣园, 朱晨阳, 薛永常*

(大连工业大学生物工程学院, 辽宁大连 116034)

摘要: 为研究差向异构结构域基因的序列特征以及其在非核糖多肽生物合成中的立体异构化特性, 以海洋链霉菌 *Streptomyces* sp. X66 基因组 DNA 为模板, 采用 PCR 技术克隆获得目的基因序列, 经测序及生物信息学分析表明, 该基因序列长 1 099 bp, Blast 序列比对该基因与 *Streptomyces albidoflavus* strain W68 的 E 结构域基因序列相似度达 99.45%, 具有典型的差向异构结构域的功能区域, 且包含差向异构结构域独有的保守基序 HHxxxD, 证明扩增的基因片段为差向异构结构域基因序列。理化分析显示: 其拟编码 366 个氨基酸, 理论等电点为 5.69, 原子组成为 $C_{1752}H_{2738}N_{510}O_{523}S_3$, 不稳定指数为 32.68, 平均亲水系数为 -0.15, 编码产物为酸性亲水稳定蛋白, 且不含信号肽和跨膜结构, 二级结构以 α -螺旋和无规卷曲为主, SDS-PAGE 显示其分子量约为 55 ku, 与预期相符。通过对差向异构结构域基因的研究, 以期为进一步解析差向异构结构域功能及新型非核糖体多肽的研发提供科学依据。

关键词: 链霉菌; 非核糖体肽合成酶; 差向异构结构域; 基因克隆; 生物信息学

文章编号: 1673-9078(2022)11-90-97

DOI: 10.13982/j.mfst.1673-9078.2022.11.0093

Cloning and Bioinformatics Analysis of Epimerization Domain Gene of Marine *Streptomyces* Species

WU Xinyuan, ZHU Chenyang, XUE Yongchang*

(School of Biological Engineering, Dalian Polytechnic University, Dalian 116034, China)

Abstract: To study the sequence characteristics of epimerization domain genes and their stereo-isomerization in the biosynthesis of non-ribosomal peptides, an epimerization domain gene was cloned from the genomic DNA of *Streptomyces* sp. X66. Bioinformatics analysis showed that the gene sequence length was 1,099 bp. Blast sequence alignment showed that the similarity between the cloned gene and the epimerization domain gene of *Streptomyces albidoflavus* strain W68 is 99.45%, and that the former has a typical functional region of the epimerization domain. It contains a unique conserved motif (HHxxxD) of the epimerization domain, which proves that the amplified gene fragment is an epimerization domain gene sequence. Physical and chemical analysis showed that it can encode 366 amino acids, and that its theoretical isoelectric point is 5.69. Its atomic composition, instability coefficient, and total average hydrophilicity were found to be $C_{1752}H_{2738}N_{510}O_{523}S_3$, 32.68, and -0.15, respectively. Its coding product is an acidic hydrophilic stable protein without a signal peptide and transmembrane structure. The secondary structure of the protein comprises an α helix and random coil. SDS-PAGE showed that its molecular weight is about 55 ku, consistent with the expected value. By studying such epimerization domain genes, we hope to provide a scientific basis for the follow-up study of epimerization domain functions and the development of new non-ribosomal peptides.

Key words: *Streptomyces*; non-ribosomal peptide synthetase; epimerization domain; gene cloning; bioinformatics

引文格式:

吴欣园,朱晨阳,薛永常.海洋链霉菌差向异构结构域基因的克隆与分析[J].现代食品科技,2022,38(11):90-97

WU Xinyuan, ZHU Chenyang, XUE Yongchang. Cloning and bioinformatics analysis of epimerization domain gene of marine *Streptomyces* species [J]. Modern Food Science and Technology, 2022, 38(11): 90-97

放线菌是诸多天然生物活性物质的重要来源, 目

收稿日期: 2022-01-26

基金项目: 辽宁省自然科学基金项目 (20180550858)

作者简介: 吴欣园 (1997-), 女, 硕士研究生, 研究方向: 微生物分子生物学,

E-mail: wxyxinyuan@163.com

通讯作者: 薛永常 (1966-), 男, 教授, 研究方向: 植物及微生物分子生物学,

E-mail: xueych@dlpu.edu.cn

前已从放线菌中发现了 10 000 多种生物活性化合物^[1], 具有广泛实际用途和巨大经济价值。海洋放线菌尤其是链霉菌具有陆地微生物所不具备的适应性^[2,3], 能在高盐、高压、低温和寡营养等条件下产生具有特殊化学结构的次级代谢产物^[4,5], 其大都是通过非核糖体肽合成酶 (Non-Ribosomal Peptide Synthetase, NRPS) 和聚酮合酶途径产生^[6]。NRPS 是一种多功能大型合成

酶,由执行不同功能的结构域组成的模块按特定顺序排列而成^[7-10](图1),NRPS体系除了腺苷化(Adenylation, A)结构域、肽基载体蛋白(Peptidyl Carrier Protein Domain, PCP)、缩合(Condensation, C)结构域等三个必需的核心结构域外,还有差向异构(Epimerization, E)结构域、硫酯酶(Thioesterase, TE)等修饰性结构域^[11],共同完成非核糖多肽(Non-Ribosomal Peptide, NRP)的生物合成及组装。在细菌、蓝藻和真菌中NRPs具有抗真菌、抗病毒、抗肿瘤和免疫抑制剂等非常重要的药理活性^[12-14],广泛应用于医药领域。

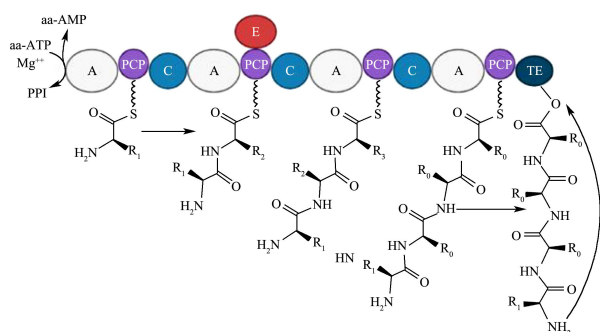


图1 NRPS 模块组成及合成机制^[10]

Fig.1 Module composition and synthetic mechanism of NRPS

注: A 为腺苷化结构域,负责激活和识别氨基酸; PCP 为肽基载体蛋白,转移产物中间体; C 为缩合结构域,催化氨基酸之间肽键的形成; TE 为硫酯酶结构域,负责多肽链的水解释放与环化; E 为差向异构结构域,催化 L 型氨基酸异构化为 D 型氨基酸。

E 结构域作为 NRPS 组成模块中常见的修饰性结构域,能将连接在 PCP 结构域上的 L-氨基酸(L-Amino Acid, L-AA)立体异构化为 D-氨基酸(D-Amino Acid, D-AA),掺入到正在延伸的 NRP 上,从而增加了 NRP 组成的多样性^[15]。由 E 结构域催化引入的 D-AA 不仅影响产物肽的最终构象以及下游加工单元模块(如硫酯酶结构域)的接受性^[16],还在确保延伸模块的正常工作方面发挥重要作用^[17]。在增加 NRP 功能和结构多样性的同时还降低其蛋白水解敏感性,提高了 NRP 的抗癌抑菌活性^[18]。目前,在 E 结构域异构化机制及其晶体结构已有所研究,Stachelhaus 等^[19]研究短杆菌素 S 合成酶 GrsA E 结构域的突变时发现⁷⁵³His 和⁸⁹²Glu 影响异构化活性。而 Samel 等^[16]解析了酪氨酸合成酶的 E 结构域晶体结构,发现⁸²²Glu 可能作为酸碱催化剂起作用,⁷⁴³His 则是用来稳定异构化过程中产生的瞬时烯醇中间体。在 GrsA 的 PCP-E 双结构域中,⁷⁵³His 的咪唑侧链与磷酸泛酰巯基乙胺基(4'-Phosphopantetheinyl, Ppant)的硫原子相连^[20]。Kim 等^[21]则在研究 GrsA 的 E 结构域催化机制发现,只有⁷⁵³His 能与氯乙烯基甘氨酸

形成稳定的酶探针加合物,进一步证实了 His 在异构化过程中的重要地位。随着对 NRPS 高级结构研究的深入,各结构域间的相互作用机制受到广泛关注。Chen 等^[20]研究 GrsA 的 PCP-E 双结构域时首次提出了 PCP 与 E 结构域之间是通过⁶¹³Arg/⁷⁸⁸Asp、⁶¹⁴Arg/⁷⁸⁵Glu 两个盐桥提供的静电相互作用发生结合的。由于 E 结构域与上下游结构域的供体位置相互作用以及其底物识别和催化机理尚不透彻,E 结构域与 PCP 结构域的具体作用还需进一步剖析。而对 E 结构域的分子机理和生物学作用的研究,对揭示 NRPS 模块间的通信机制、NRPS 的结构域操纵和重排以及新药的研发都有重要意义。本研究通过从实验室保存的具有完整 NRPS 基因簇的海洋链霉菌基因组 DNA 中扩增 E 结构域基因,并进行生物信息学分析,以期对随后深入研究其结构与功能提供参考。

1 材料和方法

1.1 试验材料

1.1.1 菌株与质粒

本文链霉菌 *Streptomyces* sp. X66 是由实验室保藏菌株^[22],大肠埃希菌 (*Escherichia coli*) DH5 α 、BL21 (DE3),表达载体 pET-32a(+)均为实验室保存。pMDTM19-T 载体购自宝生物工程(大连)有限公司。

1.1.2 培养基

LB 液体培养基、LB 固体培养基、高氏一号固体培养基配置参考文献^[23]。

1.1.3 试剂与仪器

细菌基因组 DNA 提取试剂盒、柱式 DNA 胶回收试剂盒,生工生物工程(上海)有限公司; FastPure 质粒提取试剂盒、DNA Marker DL2000,南京诺唯赞生物科技有限公司; QuickCutTMEcoR I (15 U/ μ L)、QuickCutTMHind III (15 U/ μ L)、PrimeSTAR[®] Max DNA Polymerase,宝生物工程(大连)有限公司; DNA Marker IV,天根生化科技(北京)有限公司; PCR Thermal Cycler Dice TP610,日本 TaKaRa。

1.2 试验方法

1.2.1 引物设计

参考 GenBank 数据库中 *Streptomyces albidoflavus* strain W68 (登录号: CP064783.1) NRPS 中 E 结构域基因序列,使用 Primer Premier 5.0 软件设计特异性引物。同时在引物的 5'端添加 EcoR I、Hind III 酶切位点,所设计引物为:

E-F: 5'-CGGAATTCATGCTGCGCGCCGTCTAC-3'

E-R: 5'-CCAAGCTTACTGAGCGCGGGAGTAGGAC-3'

由北京六合华大基因科技有限公司合成。

1.2.2 E 结构域基因克隆及鉴定

以 *Streptomyces* sp. X66 基因组 DNA 为模板, E-F 和 E-R 为特异引物进行 PCR 扩增。PCR 扩增体系为 20 μ L: ddH₂O 8.5 μ L、E-F (10 μ mol/L) 及 E-R (10 μ mol/L) 各 0.5 μ L, PrimeSTAR[®]Max Premix 10 μ L, 基因组 DNA 0.5 μ L。1wt%琼脂糖凝胶电泳检测扩增产物, 回收目的片段与 pMD[™]19-T 载体连接, 并转化大肠杆菌 DH5 α 感受态细胞。阳性重组质粒经特异引物 PCR、双酶切及测序鉴定。

1.2.3 E 结构域基因生物信息学分析

利用 NCBI (<http://www.ncbi.nlm.nih.gov/>) 数据库 Nucleotide Blast 对测序结果进行比对; ORF Finder 对该扩增序列进行翻译, 获得其拟编码的氨基酸序列; 使用 ExPASy ProtParam^[24]和 ProtScale^[25]对拟编码的氨基酸序列的理化性质及亲/疏水性分析; 采用 TMHMM 2.0 Server^[26]、SignalP 6.0 Server^[27]在线工具对蛋白的跨膜结构域、信号肽进行预测; 采用 SOPMA^[28]和 SWISS-MODEL^[29]预测其二级结构以及三级结构模型; NCBI 数据库中的 CDD 工具预测该蛋白的保守区域; Protein Blast 搜索 E 结构域的同源蛋白, 选取相似性 > 85%的蛋白利用 DNAMAN 软件进行氨基酸多重序列比对分析; 利用 MEGA X^[30]软件基于邻接法分析 E 结构域蛋白和其他物种 E 结构域蛋白的系统发生关系。

1.2.4 E 结构域蛋白的异源表达

利用 QuickCut[™]EcoR I 和 QuickCut[™]Hind III 对 pET-32a(+)及构建的重组质粒 pMD[™]19-T-E 进行双酶切, 回收目的片段, 连接并转入大肠杆菌 BL21 (DE3) 感受态细胞。分别培养只含 pET-32a(+)单克隆菌及含有重组子 pET-32a(+)-E 单克隆菌落, 待 OD₆₀₀ 为 0.6 时, 加入 IPTG 使反应终浓度达到 0.5 mmol/L, 16 $^{\circ}$ C 诱导 12 h 后, 进行 SDS-PAGE 分析目的基因的表达。

2 结果与分析

2.1 E 结构域基因的克隆

以 *Streptomyces* sp. X66 基因组 DNA 为模板, E-F 和 E-R 为特异引物扩增目的基因片段。经 1%琼脂糖凝胶电泳检测 (图 2), 在 1 100 bp 左右有单一清晰扩增条带, 与预期的 E 结构域基因片段大小基本一致。

回收目的片段与 pMD[™]19-T 载体连接, 转化大肠杆菌 DH5 α 感受态细胞, 选取阳性菌落提取质粒。电泳发现能够在 3 500 bp 左右出现了一条清晰明亮条带 (图 3), 这与预期的质粒大小一致。

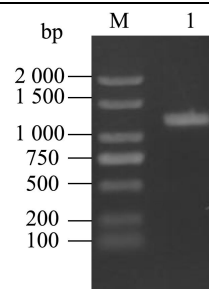


图 2 E 结构域基因扩增产物凝胶电泳图

Fig.2 Agarose gel electrophoresis of E domain gene fragment

注: M 为 DNA Marker DL2000; 1 为扩增的 E 结构域产物。

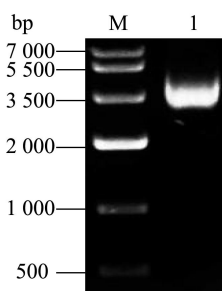


图 3 质粒 DNA 电泳图

Fig.3 Agarose gel electrophoresis of recombinant plasmid

注: M 为 DNA Marker IV; 1 为重组质粒。

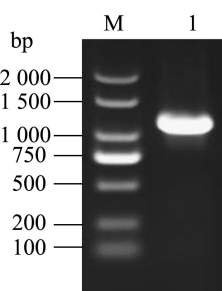


图 4 质粒 DNA 中扩增 E 结构域基因电泳图

Fig.4 Agarose gel electrophoresis of E domain gene amplified from recombinant plasmid DNA

注: M 为 DNA Marker DL2000; 1 为质粒 DNA 扩增 E 结构域基因。

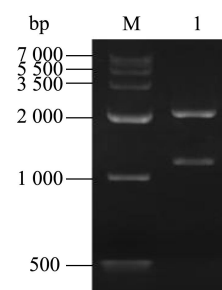


图 5 重组质粒双酶切电泳图

Fig.5 Recombinant plasmid cut by Hind III and EcoR I

注: M 为 DNA Marker IV; 1 为重组质粒 Hind III/EcoR I 双酶切图。

以重组质粒 DNA 为模板进行特异性扩增 (图 4), 扩增的条带大小在 1 100 bp 左右, 与目的条带的大小

相符,说明该重组质粒含有目的片段。将该重组质粒命名为 pMDTM19-T-E。

用限制性核酸内切酶 *EcoR* I、*Hind* III 双酶切重组质粒(图 5),在 3 000 bp 左右与 1 100 bp 左右各得到一条清晰条带,符合预期的结果,证明目的基因成功插入 T-载体中。将重组质粒送至生工生物

工程(上海)股份有限公司测序。

测序结果得知,该插入片段为 1 099 bp。Blast 序列比对表明其与 Yao 等^[31]公布的 *Streptomyces albidoflavus* strain W68 的 E 结构域基因序列相似度达 99% 以上(图 6),证明扩增的基因片段为 E 结构域基因序列。

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Streptomyces sp WA1-19 chromosome, complete genome	Streptomyces sp...	2019	2019	100%	0.0	99.82%	6850369	CP085039.1
<input checked="" type="checkbox"/> Streptomyces albidoflavus strain UYFA156 chromosome, complete genome	Streptomyces al...	2017	2017	100%	0.0	99.82%	6823997	CP040466.1
<input checked="" type="checkbox"/> Streptomyces albidoflavus strain W68 chromosome, complete genome	Streptomyces al...	1997	1997	100%	0.0	99.45%	6796629	CP064783.1

图 6 E 结构域基因序列比对

Fig.6 Alignment of the predicted sequences of E domain gene

2.2 E 结构域基因生物信息学分析

通过 ProtParam 工具对该扩增序列编码的氨基酸进行分析,该 E 结构域基因共拟编码 366 个氨基酸,分子式为 C₁₇₅₂H₂₇₃₈N₅₁₀O₅₂₃S₃,理论等电点为 5.69,不稳定指数为 32.68,平均亲水系数为-0.15。该 E 结构域蛋白的等电点小于 7,不稳定指数小于 40 且平均亲水系数为负值,推测该蛋白可能是一个酸性亲水稳定蛋白。

蛋白质的亲疏水性决定了它在水中的可溶性,高度可溶的蛋白质一般都是亲水性的部分裸露在最外面,而疏水性氨基酸在内部。利用 ProtScale 软件,选用默认 Kyte & Doolittle 计算方法分析该蛋白的亲疏水性(图 7),发现在 E 结构域蛋白序列中,在第 143 位氨基酸处存在最小值-2.25,此处亲水性最强,疏水区在第 204 位氨基酸处有最大值 2.33,亲水区域比疏水区域稍多一些,整体呈现亲水性,预测该蛋白为可溶性蛋白,这也与 ProtParam 中的分析结果相吻合。

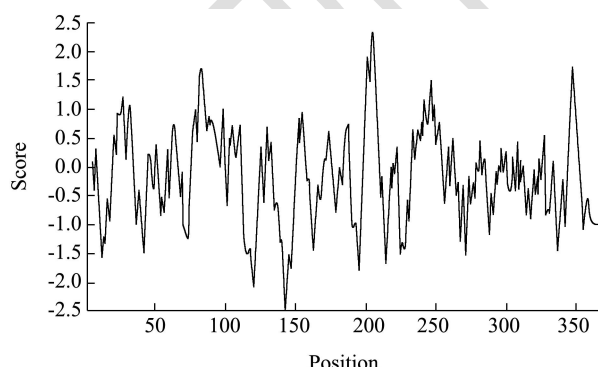


图 7 E 结构域蛋白亲水性/疏水性预测

Fig.7 Prediction of hydrophobicity/hydrophilicity of E domain protein

跨膜结构域多为强疏水性氨基酸组成,根据氨基酸亲疏水性结果推测,拟翻译的蛋白不具备跨膜结

构。为进一步验证该猜想,利用 TMHMM 软件预测 E 结构域蛋白跨膜结构域(图 8),跨膜信号没有波动,编码产物中不存在跨膜螺旋结构,故该蛋白为非跨膜类蛋白质。

信号肽通常用于指导蛋白质的跨膜转移或定位,运用 SignalP 预测该蛋白信号肽(图 9),结果显示该蛋白不具备信号肽。据此推测可能是因为它不是分泌蛋白或者异构化反应中不涉及蛋白质的转移,所以也不会存在跨膜结构。这一结果与 E 结构域基因拟表达蛋白的跨膜结构预测结果相符合。

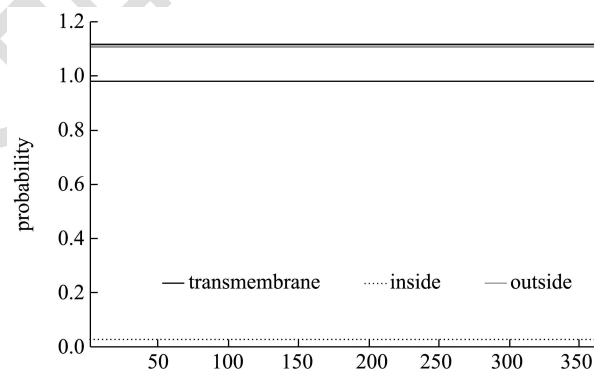


图 8 E 结构域蛋白跨膜结构域预测

Fig.8 Transmembrane domain prediction of E domain protein

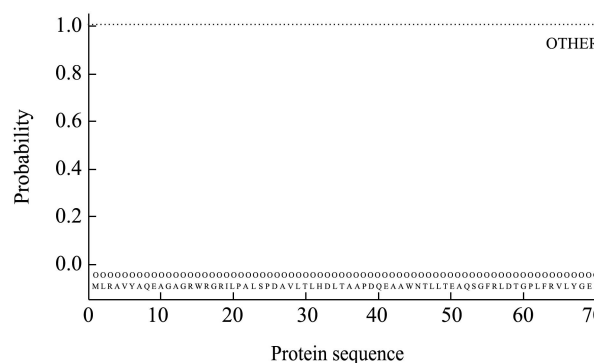


图 9 E 结构域蛋白的信号肽预测

Fig.9 Signal peptide prediction of E domain protein

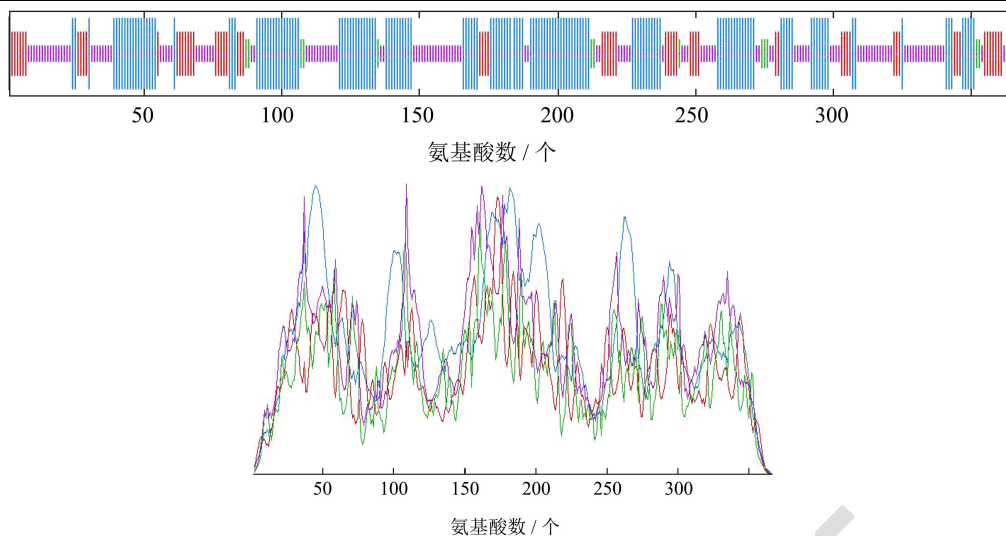


图 10 E 结构域蛋白二级结构预测

Fig.10 Prediction of the secondary structure of E domain protein

应用 SOPM 软件对该蛋白的二级结构进行预测 (图 10)。该蛋白的二级结构由 41.69%的 α -螺旋、17.44%的延伸链、3.81%的 β 转角和 37.06%无规则卷曲组成。其中以稳定的 α -螺旋结构为主,说明 E 结构域基因拟表达蛋白的二级结构较为稳定,这也与 ProtParam 中的理化性质结果相吻合。以酪氨酸合成酶 B 第三个模块的差向异构化结构域 TycB₃(E)的晶体结构为模板构建 E 结构域蛋白的三维结构,运用 SWISS-MODEL 进一步分析其三级结构 (图 11),发现序列相似性为 0.37,这与 Samel 等^[16]和 Chen 等^[20]的结果基本一致,左右两个与氯霉素乙酰转移酶结构类似的亚域构成 E 结构域的主体部分,活性位点位于两个亚域的交界处,整体均呈“V”字型结构。该蛋白结构整体以 α -螺旋为主,还包含少量的无规卷曲和延伸链,符合二级结构的预测结果。

通过 CDD 工具预测了 E 结构域的保守区域 (图 12),可以看出该序列中含有 E 结构域催化的特有结构域,属于 C 结构域超家族成员,该超家族还包括环化结构域,双重 E/C 结构域,因此可以判定克隆的目的基因片段属于 E 结构域基因。

用 NCBI Blast 检索工具比对 E 结构域基因拟编码

的氨基酸序列,选取数据库中相似性较高的 9 个蛋白质序列,使用 DNAMAN 进行多重序列比对分析 (图 13),发现该 E 结构域的氨基酸序列与同属其他物种的相似度较高,氨基酸序列较为保守,且含有一段 E 结构域保守基序 HHxxxD。

MEGA X 构建的 E 结构域蛋白的系统发育进化树 (图 14)显示该 E 结构域蛋白质序列与 *Streptomyces albidoflavus* 的 NRPS 处于同一分支,同源性达到 95%,亲缘关系较近;与 *Streptomyces scopuliridis* RB72 的 E 结构域蛋白序列同源性较低,亲缘关系相对较远。

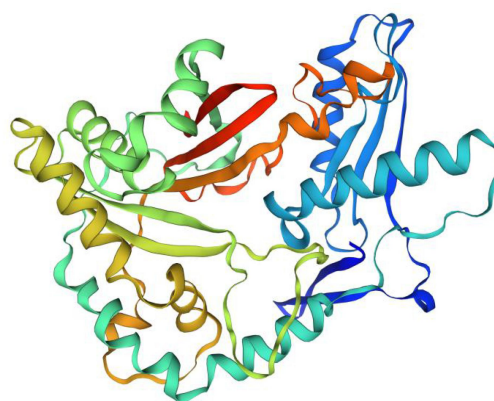


图 11 E 结构域蛋白三级结构预测

Fig.11 The tertiary structure prediction of E domain protein

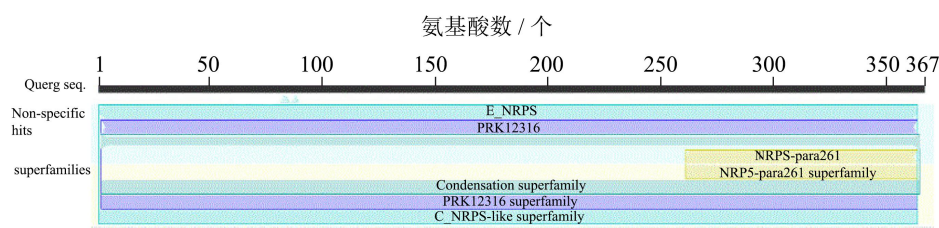


图 12 E 结构域蛋白保守结构域分析

Fig.12 Domain analysis of E domain protein

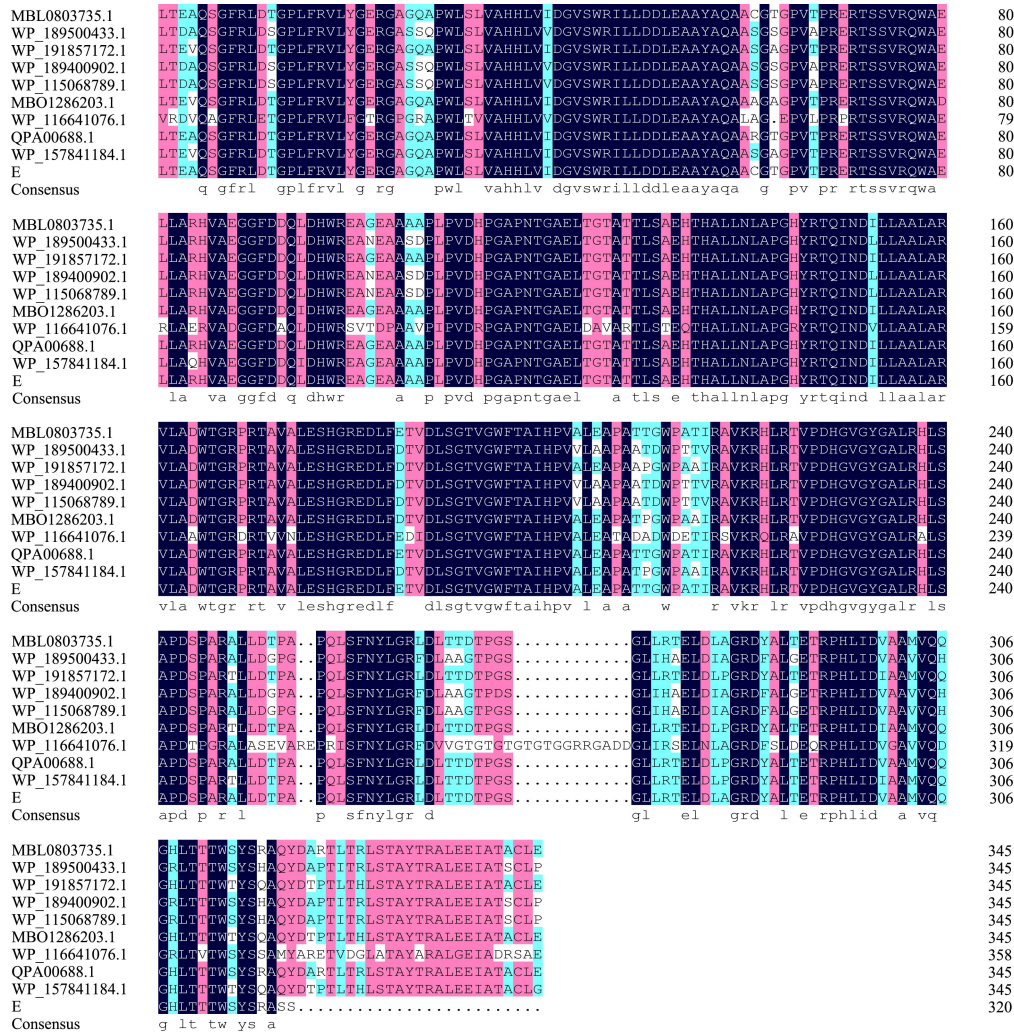


图 13 氨基酸多序列比对

Fig.13 Alignment of the predicted amino acid sequences

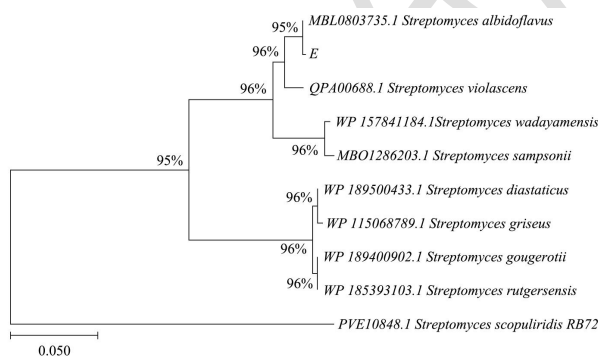


图 14 E 结构域蛋白进化树分析

Fig.14 Phylogenetic tree of E domain protein

2.3 E 结构域基因的异源表达

选取含 pET-32a(+) 的单菌落及构建成功的 pET-32a(+)-E 阳性表达质粒的单菌落进行低温诱导培养, 样品处理后用 SDS-PAGE 进行检测, 结果如图 15。

作为对照的 pET-32a(+) 质粒对诱导物 IPTG 添加与否对其表达没有影响。而含有 pET-32a(+)-E 的菌株在

未加入 IPTG 时, 没有目的蛋白的表达; 在加入 IPTG 低温诱导 12 h 后, 在 55 ku 左右出现明显的条带, 这与预测的融合蛋白大小相近, 说明 E 结构域基因在大肠杆菌 BL21 (DE3) 中能够表达。

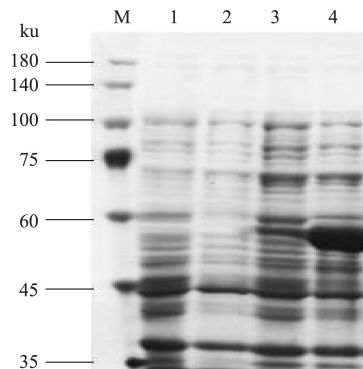


图 15 E 结构域基因拟表达蛋白 SDS-PAGE 分析

Fig.15 SDS-PAGE analysis of E domain gene expression

注: M 为蛋白 Marker; 1 为无 IPTG 诱导下的 pET-32a(+) 表达; 2 为 IPTG 诱导下的 pET-32a(+) 表达; 3 为无 IPTG 诱导下的 pET-32a(+) 表达; 4 为 IPTG 诱导下的 pET-32a(+) 表达。

3 结论

本研究从海洋链霉菌 *Streptomyces* sp. X66 的基因组中扩增获得 1 099 bp 的 E 结构域基因序列, 其编码 366 个氨基酸, 生物信息学分析发现它具有差向异构结构域特有的保守基序 HHxxxD, 理论等电点为 5.69, 不稳定指数为 32.68, 故该基因拟翻译的蛋白为酸性稳定的非分泌型蛋白质。平均亲水系数为-0.15, 结合其亲疏水性预测结果来看, 该蛋白为亲水性蛋白, 不存在信号肽和跨膜结构域; 其二级结构和三级结构显示主要以 α -螺旋和无规则卷曲结构为主, 与已报道的酪氨酸合成酶 B (PDB ID: 6TA8) 的 E 结构域晶体结构具有高度相似性。同源比对和系统发育进化树发现, 在链霉菌属中 E 结构域基因编码的氨基酸序列相似性较高, 能看出 E 结构域基因在不同种属内具有极高的保守性, 这也对应了 E 结构域蛋白保守的立体异构功能。在 0.5 mmol/L 的 IPTG 诱导下, 目的基因能表达得到大小为 55 ku 左右的融合蛋白, 这为深入开展 E 结构域蛋白的功能研究提供参考。

参考文献

- [1] Girão M, Ribeiro I, Ribeiro T, et al. Actinobacteria isolated from *Laminaria ochroleuca*: a source of new bioactive compounds [J]. *Frontiers in Microbiology*, 2019, 10: 683
- [2] Bredholt H, Fjærvik E, Johnsen G, et al. *Actinomycetes* from sediments in the Trondheim Fjord, Norway: diversity and biological activity [J]. *Marine Drugs*, 2008, 6(1): 12-24
- [3] Lam K S. Discovery of novel metabolites from marine *Actinomycetes* [J]. *Current Opinion in Microbiology*, 2006, 9(3): 245-251
- [4] 薛永常,蔡宏明,李根.海洋链霉菌 *Streptomyces* sp. B-17 卤化酶基因的克隆及生物信息学分析[J].微生物学杂志,2014,34(2):12-16
- [5] Yang Z, He J, Wei X, et al. Exploration and genome mining of natural products from marine *Streptomyces* [J]. *Applied Microbiology Biotechnology*, 2020, 104(1): 67-76
- [6] Niu X Y, Thaochan N, Hu Q B. Diversity of linear non-ribosomal peptide in biocontrol fungi [J]. *Journal of Fungi (Basel)*, 2020, 6(2): 61
- [7] Bonhomme S, Dessen A, Macheboeuf P. The inherent flexibility of type I non-ribosomal peptide synthetase multienzymes drives their catalytic activities [J]. *Open Biology*, 2021, 11(5): 200386
- [8] Bloudoff K, Fage C D, Marahiel M A, et al. Structural and mutational analysis of the non-ribosomal peptide synthetase heterocyclization domain provides insight into catalysis [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(1): 95-100
- [9] Sung C T, Chang S L, Entwistle R, et al. Overexpression of a three-gene conidial pigment biosynthetic pathway in *Aspergillus nidulans* reveals the first NRPS known to acetylate tryptophan [J]. *Fungal Genetics and Biology*, 2017, 101: 1-6
- [10] 左佃光,马俊英,王博,等.非核糖体肽类化合物的组合生物合成策略研究进展[J].中国抗生素杂志,2012,37(3):168-175
- [11] Bruner S D, Weber T, Kohli R M, et al. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE [J]. *Structure*, 2002, 10(3): 301-310
- [12] Matsunaga S, Fusetani N. Non-ribosomal peptides from marine sponges [J]. *Current Organic Chemistry*, 2003, 7(10): 945-966
- [13] Nikolouli K, Mossialos D. Bioactive compounds synthesized by non-ribosomal peptide synthetases and type-I polyketide synthases discovered through genome-mining and metagenomics [J]. *Biotechnology Letters*, 2012, 34(8): 1393-1403
- [14] Felnagle E A, Jackson E E, Chan Y A, et al. Non-ribosomal peptide synthetases involved in the production of medically relevant natural products [J]. *Molecular Pharmaceutics*, 2008, 5(2): 191-211
- [15] Wheadon M J, Townsend C A. Evolutionary and functional analysis of an NRPS condensation domain integrates β -lactam, D-amino acid, and dehydroamino acid synthesis [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2021, 118(17): e2026017118
- [16] Samel S A, Czodrowski P, Essen L. Structure of the epimerization domain of tyrocidine synthetase A [J]. *Acta Crystallographica Section D: Biological Crystallography*, 2014, 70(5): 1442-1452
- [17] Linne U, Marahiel M A. Control of directionality in non-ribosomal peptide synthesis: role of the condensation domain in preventing misinitiation and timing of epimerization [J]. *Biochemistry*, 2000, 39(34): 10439-10447
- [18] Süssmuth R D, Mainz A. Non-ribosomal peptide synthesis-principles and prospects [J]. *Angewandte Chemie*, 2017, 56(14): 3770-3821
- [19] Stachelhaus T, Walsh C T. Mutational analysis of the epimerization domain in the initiation module PheATE of gramicidin S synthetase [J]. *Biochemistry*, 2000, 39(19): 5775-5787
- [20] Chen W H, Li K, Guntaka N S, et al. Interdomain and intermodule organization in epimerization domain containing

- non-ribosomal peptide synthetases [J]. ACS Chemical Biology, 2016, 11(8): 2293-2303
- [21] Kim W E, Patel A, Hur G H, et al. Mechanistic probes for the epimerization domain of non-ribosomal peptide synthetases [J]. Chembiochem, 2019, 20(2): 147-152
- [22] 薛永常,蔡宏明,唐婧媛,等.大连海域海洋放线菌多样性及分类鉴定初探[J].微生物学杂志,2013,33(4):30-33
- [23] 杜连祥,路福平.微生物学实验技术[M].北京:中国轻工业出版社,2015
- [24] Gasteiger E, Gattiker A, Hoogland C, et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis [J]. Nucleic Acids Research, 2003, 31(13): 3784-3788
- [25] Gasteiger E, Hoogland C, Gattiker A, et al. Protein identification and analysis tools on the ExPASy server [A]. in: Walker J M. (ed) The Proteomics Protocols Handbook [M]. ToTowa, NJ: Humana Press Inc., 2005
- [26] Krogh A, Larsson B, Von Heijne G, et al. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes [J]. Journal of Molecular Biology, 2001, 305(3): 567-580
- [27] Teufel F, Almagro Armenteros J J, Johansen A R, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models [J]. Nature Biotechnology, 2022, 40: 1023-1025
- [28] Geourjon C, Deléage G. Sopma: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments [J]. Computer Applications in the Biosciences, 1995, 11(6): 681-684
- [29] Guex N, Peitsch M C. Swiss-model and the Swiss-PdbViewer: An environment for comparative protein modeling [J]. Electrophoresis, 1997, 18(15): 2714-2723
- [30] Kumar S, Stecher G, Li M, et al. MEGA X: molecular evolutionary genetics analysis across computing platforms [J]. Molecular Biology and Evolution, 2018, 35(6): 1547-1549
- [31] Yao X, Zhang Z, Huang J, et al. Candicidin isomer production is essential for biocontrol of cucumber *Rhizoctonia* rot by *Streptomyces albidoflavus* W68 [J]. Applied and Environmental Microbiology, 2021, 87(9): e03078