



天津科技大学学报

Journal of Tianjin University of Science & Technology

ISSN 1672-6510, CN 12-1355/N

《天津科技大学学报》网络首发论文

题目： 枯草芽孢杆菌碱性蛋白酶的系统生物信息分析
作者： 袁宇峰，白琳，毛淑红
DOI： 10.13364/j.issn.1672-6510.20230223
收稿日期： 2023-11-28
网络首发日期： 2024-09-30
引用格式： 袁宇峰，白琳，毛淑红. 枯草芽孢杆菌碱性蛋白酶的系统生物信息分析
[J/OL]. 天津科技大学学报. <https://doi.org/10.13364/j.issn.1672-6510.20230223>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。



DOI: 10.13364/j.issn.1672-6510.20230223

枯草芽孢杆菌碱性蛋白酶的系統生物信息分析

袁宇峰, 白琳, 毛淑红

(天津科技大学生物工程学院, 天津 300457)

摘要: 利用 NCBI 等数据库, 对枯草芽孢杆菌碱性蛋白酶的生物信息特征进行分析。结果表明, 甘氨酸 (G)、缬氨酸 (V)、丙氨酸 (A) 常常出现在碱性蛋白酶的保守位点。大部分碱性蛋白酶属于肽酶 S8 家族, 包括两种常见的类型: 枯草菌素亚家族型, 长度 379~382 aa, 等电点在 9 左右, 不稳定指数在 11~25 之间, 脂肪指数在 80~83 之间, 亲水蛋白质; S8-1 亚家族型, 长度 400~442 aa, 等电点在 5 左右, 不稳定指数在 30~36 之间, 脂肪指数在 78~80 之间, 亲水蛋白质。部分碱性蛋白酶含有抑制因子 i9 超家族结构域。本研究分析碱性蛋白酶的结构域与功能, 为提升其工业应用价值提供理论依据。

关键词: 碱性蛋白酶; 枯草芽孢杆菌; 序列比对; 生物信息学

中图分类号: Q814

文献标志码: A

文章编号: 1672-6510 (0000)00-0000-00

Systematic Bioinformatics Analysis of Alkaline Proteases from *Bacillus subtilis*

YUAN Yufeng, BAI Lin, MAO Shuhong

(College of Biotechnology, Tianjin University of Science & Technology, Tianjin 300457, China)

Abstract: The bioinformatic characteristics of *Bacillus subtilis* were analyzed using databases such as NCBI. The results indicate that G, V, and Amino acids often appear as conserved sites in alkaline proteases. Most alkaline proteases belong to the peptidase S8 family, including two common types: one is the subtilisin subfamily with a length of 379-382 aa, an isoelectric point of around 9, an instability index of 11-25, and a fat index of 80-83; the other is S8-1 subfamily with a length of 400-442 aa, an isoelectric point of around 5, an instability index between 30-36, and a fat index between 78-80. Both of them are hydrophilic proteins. In addition, some alkaline proteases contain inhibitory factor i9 superfamily domains. Further analysis of the domain and function of alkaline protease in this study is of great significance for enhancing its industrial application.

Key words: alkaline protease; *Bacillus subtilis*; sequence alignment; bioinformatics

碱性蛋白酶是一种能够在碱性条件下水解蛋白肽键的酶, 其最适合的 pH 范围为 9~11。广泛应用于洗涤剂、食品、医疗保健、酿酒、丝绸、皮革制造等行业。许多微生物都能产碱性蛋白酶。工业生产中使用的主要菌株是枯草芽孢杆菌和链霉菌。枯草芽孢杆菌是一种重要的原核表达宿主, 具有强大的蛋白质表达和分泌系统, 且具有蛋白酶耐受性和相对完整的载体/宿主系统, 因此在碱性蛋白酶的异

源表达方面具有独特的优势。大多数枯草芽孢杆菌产生的碱性蛋白酶具有含有丝氨酸的活性中心, 属于丝氨酸蛋白酶。目前枯草芽孢杆菌是工业上最常用的芽孢杆菌菌株, 也是市场上大多数商业碱性蛋白酶的生产菌株^[1-2]。全球商品酶巨头诺维信 (Novozymes) 的最具代表性碱性蛋白酶制剂 Alcalase 即用枯草芽孢杆菌生产^[3]。Emon 等^[4]从家禽粪便混合土壤中分离出一株枯草芽孢杆菌 AKAL7 菌株, 具

收稿日期: 2023-11-28; 修回日期: 2024-04-29

基金项目: 国家自然科学基金(32001657), 国家重点研发计划(2021YFC2100400)

作者简介: 袁宇峰(1998—), 男, 河北邢台人, 硕士研究生; 通信作者: 毛淑红, 教授, shuhongmao@tust.edu.cn

有角蛋白溶解活性，能够溶解鸡毛等天然的角蛋白底物，可用于洗衣洗涤剂，还在皮革加工业中具有脱毛酶的潜力。Yang 等^[5]利用枯草芽孢杆菌 WB600，成功表达了高地芽孢杆菌 W3 的 *AprE* 基因，碱性蛋白酶产物具有一定热稳定性，可在乳制品蛋白脱敏过程中发挥作用。

针对碱性蛋白酶的生物信息学分析已有相关研究^[6-7]，但这些研究只是针对一种酶，并没有对一类酶的系统性比较研究。目前已有其他酶的比较研究^[8]，但还没有碱性蛋白酶系统的生物信息学分析，缺乏对碱性蛋白酶进化保守规律、构效关系等的系统性了解，制约了对碱性蛋白酶的挖掘与设计。本研究对具有代表性的若干枯草芽孢杆菌碱性蛋白酶的序列信息、理化性质、结构域、保守序列、进化系谱图等进行生物信息学统计分析，并对比枯草芽孢杆菌的中性蛋白酶以及链霉菌的碱性蛋白酶，为研究碱性蛋白酶的性质和碱性蛋白酶改良提供参考。

1 材料与方法

1.1 蛋白酶生物信息的收集

从 NCBI 数据库(<https://www.ncbi.nlm.nih.gov>)中检索“alkaline protease”和“neutral protease”，筛选枯草芽孢杆菌 (*Bacillus subtilis*) 具有代表性的蛋白酶和链霉菌 (*Streptomyces*) 的碱性蛋白酶作为样本。

1.2 序列信息比对和理化性质研究

运用 Clustal X 软件分析比对所有样本的氨基酸序列，寻找保守程度较高的位点^[9]，使用 GeneDoc^[10]优化结果。使用 Expasy 的 ProtParam 工具 (<https://web.expasy.org/protparam/>) 预测蛋白酶的物理化学性质^[11]。

1.3 进化分析

使用 Clustal Omega 在线网站 (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) 对样本进行多重序列比对^[12]。使用 MEGA 11 构建进化树^[13]，采用 neighborjoining (NJ) 算法。

1.4 motif (基序) 和保守结构域预测

采用 MEME (<http://memesuite.org/tools/meme>) 预测蛋白酶的 motif 结构，参数设为 12^[14-15]。使用 Batch CD-Search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) 预测保守结构域。最后使用 Ttools 将预测结果进行可视化输出^[16]。

2 结果与分析

2.1 蛋白酶样本的选取

从 NCBI 数据库选取了具有代表性的枯草芽孢杆菌碱性蛋白酶 26 项、中性蛋白酶 6 项以及链霉菌碱性蛋白酶 6 项 (表 1—表 3)，酶 ID 系 NCBI 数据库的 GenBank 编号 (Accession) 即检索号。

表 1 26 种枯草芽孢杆菌碱性蛋白酶及其来源

Tab. 1 26 alkaline proteases from *B.subtilis*

酶 ID	来源	酶 ID	来源
AAA22742.1	<i>B. subtilis</i> (strain 168)	AOA53643.1	<i>B. subtilis</i> HJ0-6
AAC63365.1	<i>B. subtilis</i> K-54	BAA00186.1	<i>B. subtilis</i> subsp.amylosacchariticus
AAX35771.1	<i>B. subtilis</i> B4069	BAJ41479.1	<i>B. subtilis</i> FP-133
ABY83469.1	<i>B. subtilis</i> ZJ06	BAM57214.1	<i>B. subtilis</i> BEST7003
ADK11043.1	<i>B. subtilis</i> HT24	CCU57488.1	<i>B. subtilis</i> E1
ADM37095.1	<i>B. subtilis</i> subsp. spizizeniistr.W23	EFG91258.1	<i>B. subtilis</i> subsp.spizizenii ATCC 6633
AFH88394.1	<i>B. subtilis</i> D-2	GAK79858.1	<i>B. subtilis</i> Miyagi 4
AFQ56946.1	<i>B. subtilis</i> QB928	KIN27816.1	<i>B. subtilis</i> B4069
AGE62884.1	<i>B. subtilis</i> XF-1	OAZ67669.1	<i>B. subtilis</i> SRCM101280
AGE63567.1	<i>B. subtilis</i> XF-1	QDJ95743.1	<i>B. subtilis</i> R5
AGL34969.1	<i>B. subtilis</i> 50a	QIN90910.1	<i>B. subtilis</i> K3
AGV98709.1	<i>B. subtilis</i> DMA-09	RAP06595.1	<i>B. subtilis</i> MG27
AKD34420.1	<i>B. subtilis</i> HJ5	SCV44579.1	<i>B. subtilis</i> Ia1a

表 2 6 种枯草芽孢杆菌中性蛋白酶及其来源

Tab. 2 6 neutral proteases from *B.subtilis*

酶 ID	来源	酶 ID	来源
AAA82609.1	<i>Bacillus subtilis</i> NS15-4	BAJ41480.1	<i>Bacillus subtilis</i> FP-133
AGA21175.1	<i>Bacillus subtilis</i> subsp. subtilis str. BSP1	OAZ67749.1	<i>Bacillus subtilis</i> SRCM101280
AOR97542.1	<i>Bacillus subtilis</i> BS38	POD83287.1	<i>Bacillus subtilis</i> subsp. Subtilis SRCM101384

表 3 6 种链霉菌碱性蛋白酶及其来源

Tab. 3 6 alkaline proteases from *Streptomyces*

酶 ID	来源	酶 ID	来源
AQT71043.1	<i>Streptomyces</i> sp. fd1-xmd	RST05026.1	<i>Streptomyces</i> sp. WAC05374
QEV16121.1	<i>Streptomyces</i> fradiae ATCC 10745	THA86047.1	<i>Streptomyces</i> sp. A0592
QNE28904.1	<i>Streptomyces</i> sp. INR7	WKD36329.1	<i>Streptomyces xanthophaeus</i> GA114 Liu

2.2 蛋白酶的保守序列信息

对 38 种蛋白酶氨基酸序列进行分析, 结果如图 1 所示, 图中亮黄色位点为高度保守区域。

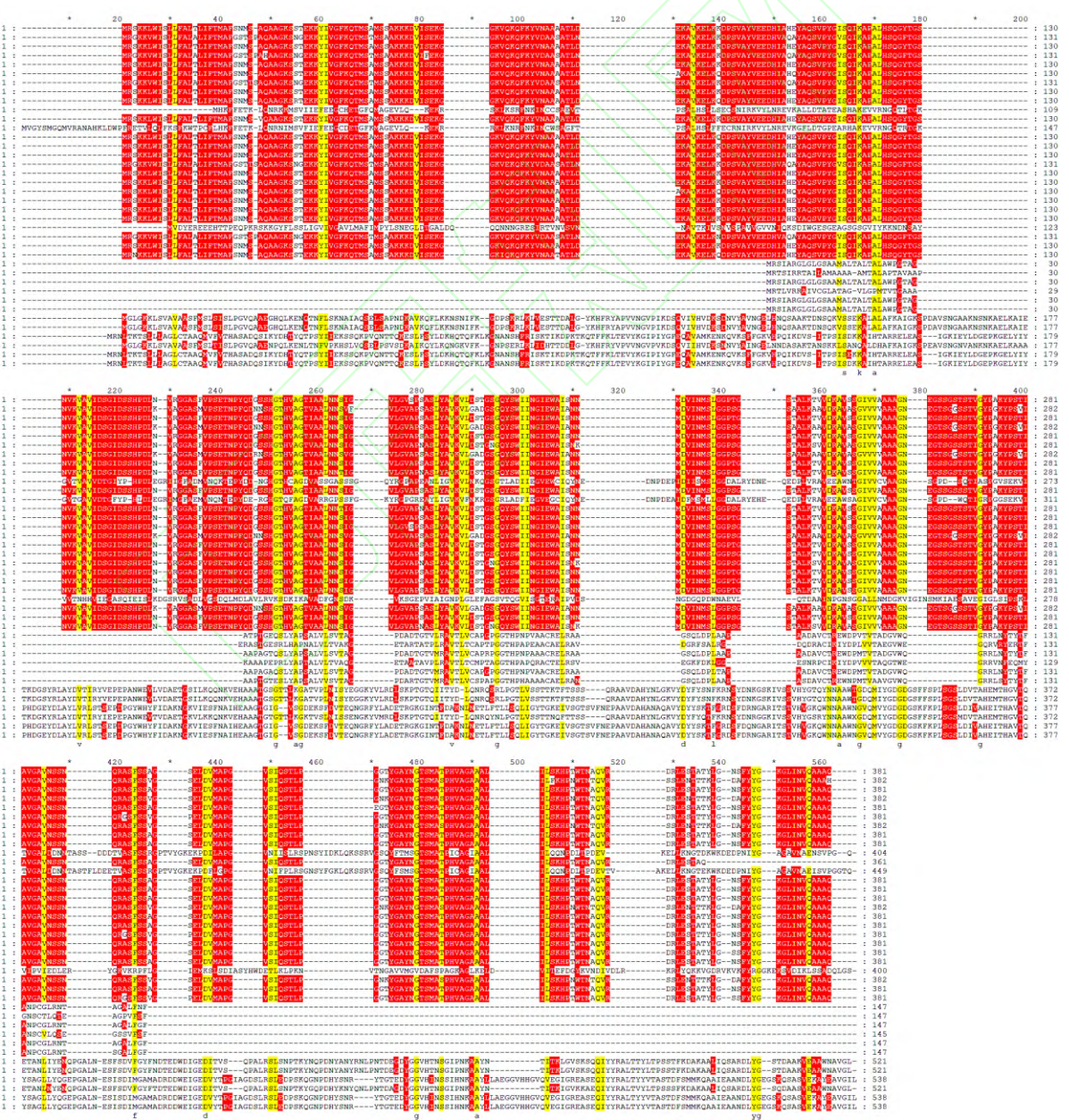


图 1 38 种蛋白酶氨基酸序列比对图

Fig. 1 Comparison of 38 proteases sequences

枯草芽孢杆菌所有碱性蛋白酶在第 63、218、303、367、427、489 位都有 G（甘氨酸），71 位都有 M（甲硫氨酸），133、234、259、488 位都有 A（丙氨酸），210、237 位都有 V（缬氨酸），226、493、531 位都有 L（亮氨酸），301、308 位都有 I（异亮氨酸），305 位都有 E（谷氨酸），即以上是它们的高度保守位点，此外存在大量的红色连续保守区域。G、V、A 三个氨基酸出现最多，推测其与酶催化活性有关。甘氨酸（G）空间位阻最小，在三维结构中较为灵活，且对 β 转角形成具有重要作用。缬氨酸（V）是疏水氨基酸，常常位于蛋白质核心，丙氨酸（A）是无特殊侧链的中性氨基酸，都对稳定蛋白质结构和发挥活性有着关键作用。除了 AGE63567.1、AGV98709.1 和 QDJ95743.1，其他蛋白酶的序列都高度相似。

链霉菌碱性蛋白酶的序列明显与枯草芽孢杆菌碱性蛋白酶序列相似度低，但保守位点 A-L（170-171 位）、A-V（285-286 位）和 G（266、295 和 391 位）均与后者高度相同，表明 G、V、A 三个氨基酸对两种生物来源的碱性蛋白酶都有关键影响。

枯草芽孢杆菌中性蛋白酶的序列也明显不同于枯草芽孢杆菌碱性蛋白酶，前者与链霉菌碱性蛋白酶一样，只存在分散的一些位点，几乎没有连续的保守区域。枯草芽孢杆菌碱性蛋白酶有自己专属的保守结构。

2.3 蛋白酶的进化树与分类

38 种蛋白酶的系统发育进化树如图 2 所示。枯草芽孢杆菌的碱性蛋白酶、中性蛋白酶以及链霉菌的碱性蛋白酶各自聚类，且枯草芽孢杆菌的碱性蛋白酶和中性蛋白酶亲缘关系更近，QDJ95743.1 与其他蛋白酶亲缘较远。全部样本大致可以分为 3 组，每组之间亲缘性与来源菌株亚种无明显关系。

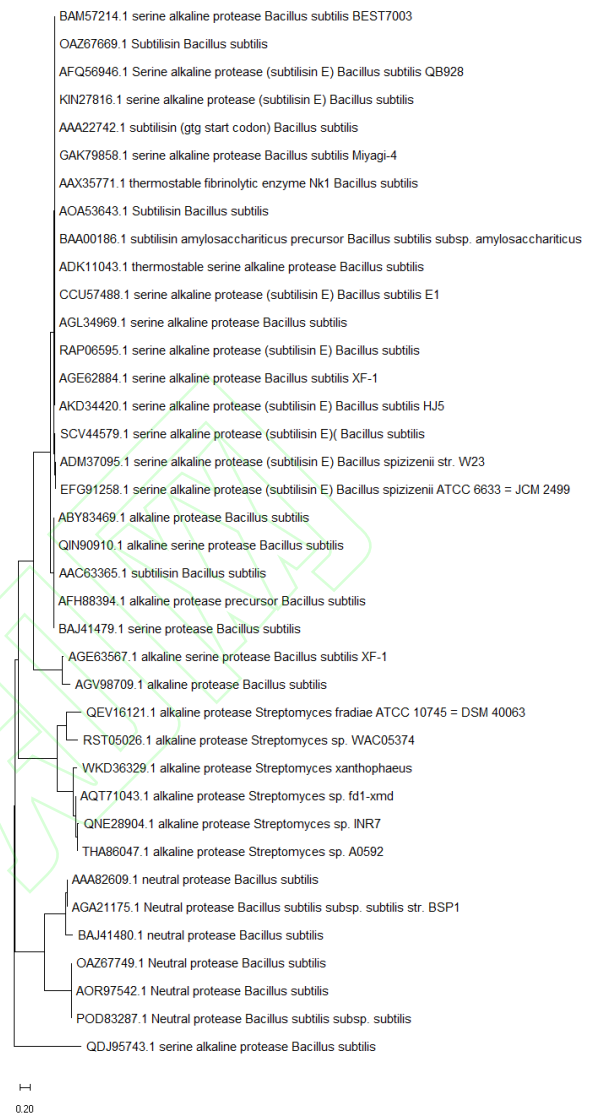


图 2 38 种蛋白酶的系统发育进化树

Fig. 2 Phylogenetic evolution tree of 38 proteases

2.4 蛋白酶的理化性质

26 种枯草芽孢杆菌碱性蛋白酶的理化性质见表 4。除了 AGE63567.1、AGL34969.1、AGV98709.1 和 QDJ95743.1，大多数酶的氨基酸残基数是 381 aa 或 382 aa，等电点在 9.04~9.31 之间，脂肪指数在 79.41~82.28 之间，而且都是稳定的亲水性蛋白。

6 种枯草芽孢杆菌中性蛋白酶的理化性质见表 5。6 种蛋白酶的氨基酸残基数是 521 aa 或 538 aa，等电点虽然相差较大，但是酶发挥最高活性仍在中性 pH，且都为稳定的亲水性蛋白。

6 种链霉菌碱性蛋白酶的理化性质见表 6。氨基酸残基数为 147 aa 或 145 aa，多数酶的等电点大于 7，

稳定性并不同一, 如 QEV16121.1 和 RST05026.1 并 性的蛋白质。
不稳定, 而其他酶稳定。链霉菌碱性蛋白酶是疏水

表 4 26 种枯草芽孢杆菌碱性蛋白酶的理化性质

Tab. 4 Physicochemical properties of 26 *B. subtilis* alkaline proteases

酶 ID	残基数量/aa	相对分子质量	等电点	不稳定指数	脂肪指数	总平均亲水性
AAA22742.1	381	3.950×10^4	9.04	29.71	80.97	-0.068
AAC63365.1	382	3.925×10^4	9.23	25.07	82.02	0.000
AAX35771.1	381	3.947×10^4	9.04	30.07	81.23	-0.061
ABY83469.1	382	3.918×10^4	9.23	24.29	82.28	-0.008
ADK11043.1	381	3.959×10^4	9.04	30.93	81.23	-0.080
ADM37095.1	381	39.55×10^4	9.25	30.60	82.23	-0.059
AFH88394.1	382	3.914×10^4	9.31	23.94	82.02	-0.019
AFQ56946.1	381	3.948×10^4	9.04	29.06	81.23	-0.061
AGE62884.1	381	3.958×10^4	9.05	30.19	81.23	-0.077
AGE63567.1	404	4.332×10^4	4.87	38.13	79.41	-0.402
AGL34969.1	361	37.35×10^4	8.95	31.65	81.39	-0.068
AGV98709.1	449	4.959×10^4	6.08	32.87	68.40	-0.432
AKD34420.1	381	3.951×10^4	8.91	30.19	81.23	-0.067
AOA53643.1	381	3.947×10^4	9.04	30.07	81.23	-0.061
BAA00186.1	382	3.947×10^4	9.23	31.45	80.97	-0.069
BAJ41479.1	382	3.910×10^4	9.14	23.81	82.02	-0.016
BAM57214.1	381	3.947×10^4	9.04	30.07	81.23	-0.061
CCU57488.1	381	3.945×10^4	9.04	30.80	81.23	-0.062
EFG91258.1	381	3.955×10^4	9.25	30.60	82.23	-0.059
GAK79858.1	381	3.951×10^4	9.04	28.84	81.73	-0.055
KIN27816.1	381	3.945×10^4	9.04	29.79	81.23	-0.054
OAZ67669.1	381	3.948×10^4	9.04	29.06	81.23	-0.061
QDJ95743.1	400	4.273×10^4	5.38	22.46	95.92	-0.194
QIN90910.1	382	3.910×10^4	9.16	25.02	82.28	-0.017
RAP06595.1	381	3.945×10^4	9.04	30.80	81.23	-0.062
SCV44579.1	381	3.963×10^4	9.16	31.32	82.23	-0.072

表 5 6 种枯草芽孢杆菌中性蛋白酶的理化性质

Tab. 5 Physicochemical properties of 6 *B. subtilis* neutral proteases

酶 ID	残基数量/aa	相对分子质量	等电点	不稳定指数	脂肪指数	总平均亲水性
AAA82609.1	521	5.652×10^4	7.16	30.09	74.38	-0.510
AGA21175.1	521	5.652×10^4	7.16	30.09	74.38	-0.510
AOR97542.1	538	5.934×10^4	5.66	25.86	73.48	-0.488
BAJ41480.1	521	5.686×10^4	8.48	25.52	73.22	-0.551
OAZ67749.1	538	5.932×10^4	5.60	26.29	73.48	-0.487
POD83287.1	538	5.934×10^4	5.66	25.86	73.48	-0.488

表 6 6 种链霉菌碱性蛋白酶的理化性质

Tab. 6 Physicochemical properties of 6 *Streptomyces* alkaline proteases

酶 ID	残基数量/aa	相对分子质量	等电点	不稳定指数	脂肪指数	总平均亲水性
AQT71043.1	147	1.488×10^4	6.53	29.60	83.88	0.194

QEV16121.1	147	1.541×10^4	8.88	42.92	79.25	0.003
QNE28904.1	147	1.486×10^4	8.49	29.47	79.32	0.195
RST05026.1	145	1.518×10^4	8.97	56.76	82.14	0.128
THA86047.1	147	1.479×10^4	7.68	29.18	80.61	0.239
WKD36329.1	147	1.473×10^4	8.88	32.49	79.93	0.288

2.5 蛋白酶的 motif 结构和保守结构域分析

蛋白酶 motif 结构和保守结构域分析图如图 3 所示。枯草芽孢杆菌碱性蛋白酶 AGE63567.1 和 AGV98709.1 只包含 4 个 motif 结构, QDJ95743.1 不包含任何 motif 结构, 其他碱性蛋白酶都包含 motif 结构 1—11。有趣的是, 枯草芽孢杆菌中性蛋白酶和链霉菌碱性蛋白酶都只包含 motif 12。枯草芽孢杆菌碱性蛋白酶与其他两种酶之间不存在相同的 motif 结构。枯草芽孢杆菌碱性蛋白酶 AGE63567.1 和 AGV98709.1 只有 Peptidases_S8_1 结构域, QDJ95743.1 只有 DegQ 结构域, 其余碱性蛋白酶都含有 Peptidases_S8_Subtilisin_subset 结构域, 大多数碱性蛋白酶还有 Inhibitor_i9 superfamily 结构域。枯草芽孢杆菌中性蛋白酶均含有 LasB 结构域, 链霉菌碱性蛋白酶大都涉及 SSI 结构域。3 种类型的酶之间不存在相同的结构域。

绝大多数枯草芽孢杆菌碱性蛋白酶属于肽酶 S8 家族, 其中典型的都属于肽酶 S8 家族的枯草菌素亚家族, 有的还属于抑制因子 i9 超家族。肽酶 S8 枯草菌素亚家族 (Peptidases_S8_Subtilisin_subset) 结构域是肽酶 S8 家族酶的一个显著特征, 其组成为 Asp (天冬氨酸) /His (组氨酸) /Ser (丝氨酸) 催化三联体, 此三联体作为酶的电荷中继系统为酶提供活性, 且需要与 Ca^{2+} 结合才能诱导酶的折叠进而稳定结构、发挥活性。该三联体也在胰蛋白酶类蛋白酶中发现, 但三维结构不同, 并且二者不同源。肽酶 S8-1 家族 (Peptidases_S8_1) 结构域也是属于肽酶 S8 家族酶的一个特征。它也有与肽酶 S8 枯草菌素亚家族类似的 Asp/His/Ser 催化三联体, 但相比前者 Asp 位点靠后。 Ca^{2+} 可以增强酶的稳定性, 酶可以通过不

同亲和力的结合位点结合多达 4 个 Ca^{2+} 。这个家族的一些酶含有二硫键, 使得其可在极端温度和 pH 下发挥活性^[17]。结合前面对理化性质的分析, 发现, 样本的理化性质分为两种类型: 枯草菌素亚家族型, 长度 379~382 aa, 等电点在 9 左右, 不稳定指数在 11~25 之间, 脂肪指数在 80~83 之间, 都是亲水蛋白质; S8-1 亚家族型, 长度 400~442 aa, 等电点在 5 左右, 不稳定指数在 30~36 之间, 脂肪指数在 78~80 之间, 都是亲水蛋白质。抑制因子 i9 超家族 (Inhibitor I9) 结构域在酿酒酵母的蛋白酶 B 抑制因子和来自枯草杆菌蛋白酶家族肽酶的激活肽中较为常见, 其参与了枯草菌素蛋白酶成熟肽酶的折叠过程^[18]。DegQ (sacQ) 为调节肽, 属于 cl06985 超家族, DegQ 肽家族控制枯草芽孢杆菌细胞内降解酶的合成速率。DegQ 包含 46 个氨基酸, 可激活降解酶的合成, 该肽的表达受到降解酶的产物抑制和上游 DegS-DegU 基因产生的双组分调节蛋白的介导, 主要是通过 DegQ 上游-393~-186 位区间的靶点正调控, 该区域可能是位于上游的基因转录终止子或参与调节 DegQ 基因表达的结构^[19-20]。此外, 在米曲霉碱性蛋白酶中也发现其具有典型 Peptidase_S8 保守结构域, 且催化三联体线性排列也与枯草芽孢杆菌碱性蛋白酶一样, 为 DHS。因此可以认为二者较为相似, 且催化机理同为酸碱催化和同源催化协同作用^[21]。

此外, 对具有代表性的枯草芽孢杆菌碱性蛋白酶 AAX35771.1 进行了建模, 三维结构如图 4 所示, 主要由 α -螺旋 (红色部分), β -折叠 (黄色部分) 构成。

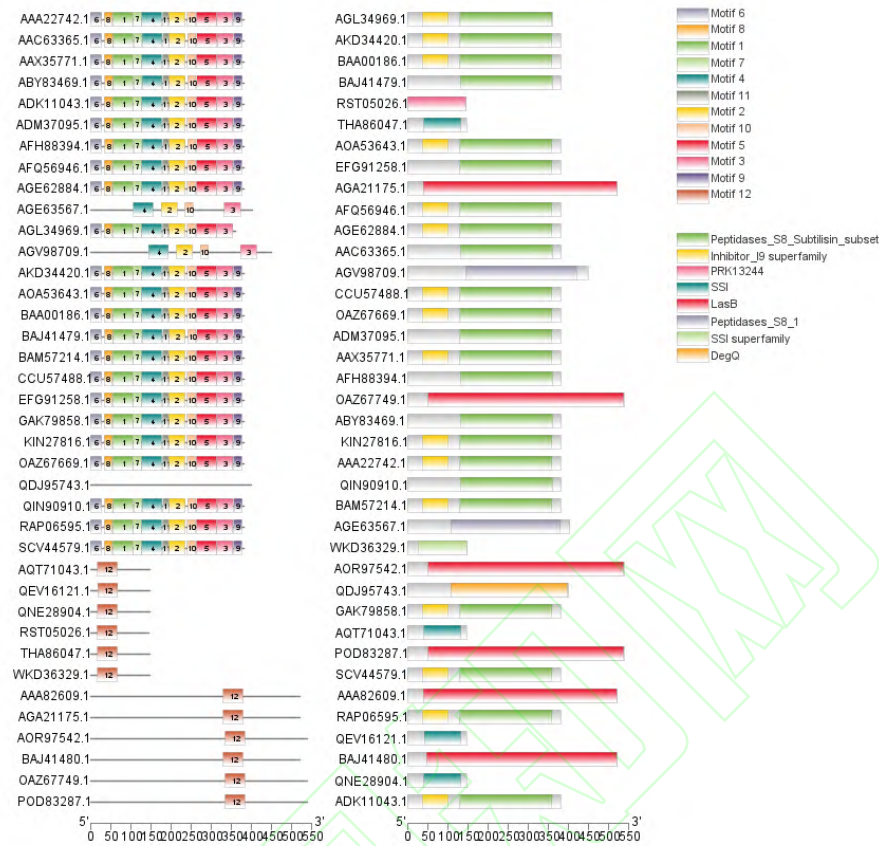


图 3 蛋白酶的 motif 结构和保守结构域分析图

Fig. 3 Schematic representation of motif and conserved domains of proteases

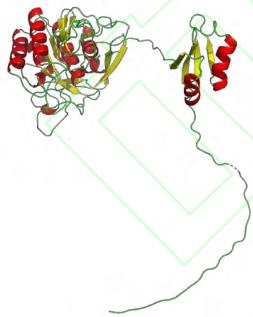


图 4 枯草芽孢杆菌碱性蛋白酶 AAX35771.1 三维结构图示

Fig. 4 Graphical representation of the structure of *B. subtilis* alkaline protease AAX35771.1

3 讨论

碱性蛋白酶广泛应用于各种工业生产领域，为适应不同的应用场景，利用蛋白质工程提升酶的催化性能成为酶研究领域的热点^[2]。在碱性蛋白酶的改进方案中，定点突变是一个很好的改良手段。有研

究将短小芽孢杆菌碱性蛋白酶位于中心的缬氨酸 (V) 替换为异亮氨酸 (I)，可改变其与底物的结合能力，增加了底物亲和性^[22]。甘氨酸 (G) 常常参与 β -转角的形成，有研究将较为灵活的脯氨酸 (P) 取代碱性蛋白酶中 β -转角的 G，提高了酶分子的稳定性^[23]。肽酶 S8 家族也称为枯草菌素家族，是丝氨酸肽酶的第二大家族，分布广泛。

本研究选用的碱性蛋白酶大多也属于 S8 家族。枯草菌素家族蛋白中常常存在多个保守结构域，如抑制因子 i9、肽酶 S8-3、蛋白酶相关 (PA) 枯草菌素蛋白样 (PA-subtilisin-like)、肽酶 S8-S53 等^[24]。这些独特的结构域组合赋予蛋白质各种各样的功能。抑制因子 i9 结构域可以阻断蛋白质的催化中心，从而调节肽酶前体的折叠和活性。PA 结构域能够形成同型二聚体，与自抑制 β -发夹结构域结合，从而解除自抑制，产生活性酶，还能决定最佳底物长度。肽酶 S8-3 和肽酶 S8-S53 结构域可能在降解特定底物方面具有潜在的作用^[25]。

本研究通过对枯草芽孢杆菌来源的碱性蛋白酶生物信息挖掘,系统分析了碱性蛋白酶序列、理化性质、结构域、保守序列、进化系谱图等信息。结果表明,G、V、A三个氨基酸通常出现在枯草芽孢杆菌碱性蛋白酶的保守位点。系统发育树上越接近的酶,亲缘关系越近,同源性更高。枯草芽孢杆菌碱性蛋白酶具有保守结构域的特征,进一步研究不同结构域与其功能的关系,对于碱性蛋白酶的改造以满足其工业应用具有重要的价值。

参考文献:

- [1] 刘逸寒,王洪彬,刘夫锋,等.微生物碱性蛋白酶的研究与开发[J].天津科技大学学报,2017,32(2): 1-6.
- [2] 邓菊云.微生物碱性蛋白酶研究进展[J].现代食品科技,2008,105(3): 293-296.
- [3] 李捷,吴颖.诺维信在华主要蛋白酶制剂专利技术综述[J].山东化工,2019,48(16): 67-68.
- [4] EMON TH, HAKIM A, CHAKRABORTHY D, et al. Kinetics, detergent compatibility and feather-degrading capability of alkaline protease from *B.subtilis* AKAL7 and *Exiguobacterium indicum* AKAL11 produced with fermentation of organic municipal solid wastes[J]. Journal of Environmental Science and Health, 2020, 55: 1339-1348.
- [5] YANG S, ZHAI L, HUANG L, et al. Mining of alkaline proteases from *Bacillus altitudinis* W3 for desensitization of milk proteins: their heterologous expression, purification, and characterization[J]. International Journal of Biological Macromolecules, 2020, 153: 1220-1230.
- [6] 郝建华,袁翠,王跃军,等.低温碱性蛋白酶 QDAPr 的生物信息学分析及同源建模[J].生物信息学,2006(2):62-64.
- [7] 黄磊,董自星,金鹏,等.地衣芽孢杆菌碱性蛋白酶嗜碱突变体的特征分析[J].食品与发酵工业,2017,43(6):34-40.
- [8] 艾嘉,刘坤,张蕾,等.细菌漆酶的生物信息学分析[J].基因组学与应用生物学,2019,38(3):1070-1078.
- [9] LARKIN MA, BLACKSHIELDS G, BROWN NP, et al. Clustal W and Clustal X version 2.0[J]. Bioinformatics, 2007, 23(21): 2947-2948.
- [10] 章莹,蔡国斌,何立,等.寄生性吸虫谷胱甘肽过氧化物酶生物信息学分析[J].中国人兽共患病学报,2010,26(4): 293-298.
- [11] DUVAUD S, GABELLA C, LISACEK F, et al. ExPasy, the Swiss bioinformatics resource portal, as designed by its users[J]. Nucleic acids research, 2021, 49(1): 216-227.
- [12] MADEIRA F, PEARCE M, TIVEY ARN, et al. Search and sequence analysis tools services from EMBL-EBI in 2022[J]. Nucleic acids research, 2022, 50(1): 276-279.
- [13] TAMURA K, STECHER G, KUMAR S. MEGA11: Molecular evolutionary genetics analysis version 11[J]. Molecular biology and evolution,2021, 38(7): 3022-3027.
- [14] 赵冰雪,聂功平,张小惠,等.芝麻 AP2/ERF 基因家族生物信息学鉴定与分析[J].应用技术学报,2022,22(4):399-412.
- [15] BAILEY TL, BODEN M, BUSKE FA, et al. MEME SUITE: tools for motif discovery and searching[J]. Nucleic acids research, 2009, 37: 202-208.
- [16] CHEN C, CHEN H, ZHANG Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data[J]. Molecular plant, 2020, 13(8): 1194-1202.
- [17] SIEZEN RJ, DE VOS WM, LEUNISSEN JA, et al. Homology modelling and protein engineering strategy of subtilases, the family of subtilisin-like serine proteinases[J]. Protein engineering, 1991, 4(7):719-737.
- [18] LI Y, HU Z, JORDAN F, et al. Functional analysis of the propeptide of subtilisin E as an intramolecular chaperone for protein folding. Refolding and inhibitory abilities of propeptide mutants[J]. Journal of biological chemistry, 1995, 270(42):25127-25132.
- [19] MSADEK T, KUNST F, HENNER D, et al. Signal transduction pathway controlling synthesis of a class of degradative enzymes in *B.subtilis*: expression of the regulatory genes and analysis of mutations in degS and degU[J]. Journal of bacteriology, 1990, 172(2):824-834.
- [20] MSADEK T, KUNST F, HENNER D, et al. DegS-DegU and ComP-ComA modulator-effector pairs control expression of the *B.subtilis* pleiotropic regulatory gene degQ[J]. Journal of bacteriology, 1991, 173(7):2366-2377.
- [21] 郭继平,马光,朱会霞,等.米曲霉碱性蛋白酶三维结构的同源建模[C]. International Conference on Cellular, Molecular Biology, Biophysics & Bioengineering. 2010.
- [22] ZHAO HY, FENG H. Engineering *Bacillus pumilus* alkaline serine protease to increase its low-temperature proteolytic activity by directed evolution[J]. BMC Biotechnology, 2018, 18(1):34-45.
- [23] ZHANG H, SANG J, ZHANG Y, et al. Rational design of a *Yarrowialipolytica* derived lipase for improved thermostability[J]. International journal of biological macromolecules, 2019, 137: 1190-1198.
- [24] CAOJ, HANX, ZHANGT, et al. Genome-wide and molecular evolution analysis of the subtilase gene family in

Vitis vinifera[J]. BMC Genomics, 2014, 15, 1116.

- [25] CUI H, ZHOU G, RUAN H, et al. Genome-wide identification and analysis of the maize serine peptidase S8 family genes in response to drought at seedling stage[J]. Plants-basel, 2023, 12(2):369-376.

