DOI:10. 13880/j. cnki. 65-1174/n. 2026. 22. 015 文章编号:1007-7383(2026)05-0651-10

肠道卵形拟杆菌基因组中的多糖合成酶挖掘与验证

李雯,刘雪如,韩林,郭宇昕,余金明,刘开平,胡圣伟,李晓悦*
(石河子大学生命科学学院,新疆 石河子 832003)

摘要:目的 多糖是由单糖通过糖苷键连接而成的高分子碳水化合物,可通过促进有益菌增殖、调节菌群结构及促进肠道菌群代谢生成短链脂肪酸等机制发挥显著的益生元功能。卵形拟杆菌属是动物肠道微生物基因组中挖掘高效碳水化合物活性酶(CAZyme)的重要研究对象。**方法** 为深入探究该菌属的酶系特征,本研究分析了 1 780 个不同来源的卵形拟杆菌宏基因组组装基因组(MAG),发现其是 CAZyme 多样性的丰富储库,共预测到糖苷水解酶(GH)编码基因 506 389 条和糖基转移酶(GT)编码基因 103 975 条。后续实验结果表明,卵形拟杆菌不仅有降解多糖的功能,还有合成多糖的能力。通过多糖利用位点(PUL)注释,筛选得位点最多为 119 个的卵形拟杆菌基因组序列,利用分子对接与分子动力学模拟技术,预测鉴定出一个潜在的 CAZyme,该酶编码 452 个氨基酸,命名为 V6248。**结果** 体外酶活验证结果表明,V6248 是一种金属酶(Zn^{2+}可提高其合成效率),能够催化纤维二糖、麦芽三糖、α-乳糖、麦芽糖、葡萄糖和木糖合成多糖,为多糖益生元的合成提供了新途径。**结论** 本研究不仅为新型 CAZyme 的高效挖掘提供了思路,也为开发新型生物催化剂提供了宝贵资源。

关键词:多糖;卵形拟杆菌;碳水化合物活性酶;糖基转移酶;分子对接
中图分类号:Q936 **文献标志码:**A

Mining and verification of polysaccharide synthases in *Bacteroides ovatus* Microbial MAGs

LI Wen, LIU Xueru, HAN Lin, GUO Yuxin, YU Jinming, LIU Kaiping,
HU Shengwei, LI Xiaoyue*
(College of Life Sciences, Shihezi University, Shihezi, Xinjiang 832003, China)

Abstract: Objective To polysaccharides are high-molecular-weight carbohydrates formed by monosaccharides connected via glycosidic bonds, which exert significant prebiotic functions through multiple mechanisms, including promoting the proliferation of beneficial bacteria, regulating the structure of intestinal flora, and facilitating the metabolism of intestinal flora to produce short-chain fatty acids. Bacteroides ovatus is a critical research target for mining efficient carbohydrate-active enzymes (CAZymes) from the animal intestinal microbial genome. **Methods** To deeply explore the enzymatic characteristics of this genus, this study analyzed 1 780 metagenome-assembled genomes (MAGs) of Bacteroides ovatus from different sources, and revealed that Bacteroides ovatus is a rich reservoir of CAZyme diversity. In total, 506 389 glycoside hydrolase (GH) coding genes and 103 975 glycosyltransferase (GT) coding genes were predicted. Subsequent experimental results show that Bacteroides ovatus not only possesses the function of degrading polysaccharides, but also has the capacity to synthesize polysaccharides. Through polysaccharide utilization loci (PUL) annotation, the genomic sequence of Bacteroides ovatus with the maximum number of PULs (119) was screened. Using molecular docking and molecular dynamics simulation, a potential CAZyme was predicted and identified. This enzyme encodes 452 amino acids and is named V6248. **Results** In vitro enzyme activity verification demonstrates that V6248 is a metalloenzyme, and Zn^{2+} can improve its synthetic efficiency. It can catalyze the synthesis of polysaccharides from cellobiose, maltotriose, α-lactose, maltose, glucose and xylose, which provides a novel approach for the synthesis of polysaccharide prebiotics. **Conclusion** This study not only provides insights for the efficient mining of novel CAZymes, but also offers valuable resources for the development of new biocatalysts.

Key words: polysaccharides; *Bacteroides ovatus*; carbohydrate-active enzymes; glycosyltransferase; molecular docking

收稿日期:2026-04-20
基金项目:国家自然基金项目(32560017);新疆维吾尔自治区天池英才工程项目
作者简介:李雯(2000—),女,硕士研究生,专业方向为生化与分子学。
*通信作者:李晓悦(1994—),女,副教授,从事肠道宏基因组学研究,e-mail:1975226215@qq.com。

多糖是由多个单糖通过糖苷键连接而成的高分子碳水化合物,广泛存在于自然界中。近年来,随着人们对功能性食品与肠道健康关注度的持续提升,多糖在食品、医药及生物制造等领域展现出日益重要的应用价值。作为一类重要的膳食生物活性成分,多糖可通过促进有益菌增殖、抑制条件致病菌定植、调节菌群结构多样性以及调控短链脂肪酸等微生物代谢产物的合成,发挥显著的益生元功能[1]。在代谢性疾病方面,多糖作为益生元能够改善肠道完整性并调节宿主代谢,从而对肥胖、糖尿病等疾病起到积极的干预作用[2]。此外,发酵多糖在维持代谢稳态和干预相关疾病方面也展现出良好潜力。

抗生素抗性基因(antibiotic resistance genes, ARGs)作为微生物产生耐药表型的关键遗传元件,也与菌群多糖代谢及抗性调控存在密切关联[3]。多糖代谢可为 ARGs 的表达与水平基因转移提供物质和能量支撑,而 ARGs 则能调控微生物代谢网络,影响胞内及胞外多糖的合成;二者通过代谢通路的交叉调控,共同影响肠道菌群的适应性与耐药特征[4]。

相较于传统的植物提取法与化学合成法,酶法合成多糖具有反应条件温和、产物结构可控、环境友好等显著优势。利用碳水化合物活性酶(CAZymes)进行多糖的酶法合成,已成为可持续生物制造领域的研究热点[5]。CAZymes 中的糖基转移酶(GTs)能够将单糖从活化供体转移至受体分子,催化糖苷键形成,从而实现多糖链的定向延伸与组装[6]。在众多微生物来源中,卵形拟杆菌(*Bacteroides ovatus*, *B. ovatus*)作为人类和动物肠道核心共生菌之一,其基因组中蕴含着丰富的 CAZymes 资源,尤其以多糖利用位点(polysaccharide utilization loci, PULs)为典型代表[7]。PUL 是拟杆菌门特有的基因簇结构,通常编码多糖识别、结合、转运及水解所需的一系列酶;值得注意的是,近年来研究发现,部分 PUL 中也存在编码糖基转移酶的基因,可能参与细菌表面多糖或胞外多糖的合成。这些发现为从肠道共生菌中挖掘新型多糖合成酶提供了重要线索[8]。

本研究系统收集了已报道的卵形拟杆菌基因组信息,通过基因组注释与比较基因组学分析,聚焦其 PUL 位点,从中筛选并鉴定一个新型高效的、具有多糖合成活性的碳水化合物酶,以期为功能性多糖的绿色酶法合成提供新酶资源。

1 材料与方法

1.1 卵形拟杆菌基因组组装和注释

基于在 NCBI 平台上对“卵形拟杆菌(*Bacteroides ovatus*)”的基因组检索,下载卵形拟杆菌基因组序列(MAGs)。使用 CheckM2 评估这些基因组序列的质量[9]。丢弃低质量的基因组序列,并使用99.5%的阈值(dRep v. 3.4.3)去除了冗余序列[10]。所有基因组序列均使用 GTDB-TK(版本 2.3.2,修订版 214)进行了分类学注释[11]。

1.2 潜在的新型 CAZymes

为进一步挖掘潜在的新型高效碳水化合物活性酶(CAZymes),我们首先使用 run_dbcan2 对所有宏基因组组装基因组(MAGs)进行了 CAZymes 注释,同时利用 PULpy 对多糖利用位点(PULs)进行了预测[12]。提取在 PULs 上未被注释为 CAZymes,将其与已知碳水化合物酶库进行 blastp 比对,保留相似度为 10%~30%的蛋白作为候选蛋白。此外,我们使用 BLASTP 将这些候选酶的序列与 NR 数据库进行比对,进一步分析其相似性,并确认这些序列在现有数据库中均被注释为未知蛋白。同时满足上述所有条件的酶,即被定义为潜在的新型碳水化合物酶[13]。使用 AlphaFold 3 将这些蛋白序列转化为3D 的 pdb 结构,多种碳水化合物底物分别从 PubChem 数据库(https://pubchem.ncbi.nlm.nih.gov/)和 RCSB 蛋白质数据库(https://www.rcsb.org/)获取,利用 autoduck vina 将其与多种碳水化合物小分子进行分子对接,对接结果借助 PyMOL 实现可视化[14]。

1.3 CAZymes 理化性质预测

挑选具有最强结合能的蛋白作为候选蛋白。运用 ProtParam 在线工具(https://web.expasy.org/protparam)估算蛋白的理论分子量(Mw)与等电点(pI)[15];借助 ProtScale 在线软件(https://web.expasy.org/protscale/)分析氨基酸序列的亲水/疏水性及水溶性[16]。信号肽和跨膜结构域的预测分别采用 SignalP 6.0 与 DeepTMHMM-1.0[17]。此外,利用 NetPhos 3.1(https://services.healthtech.dtu.dk/services/NetPhos-3.1/)预测候选蛋白的磷酸化位点,并通过 NetNGlyc 1.0(https://services.healthtech.dtu.dk/services/NetNGlyc-1.0/)分析糖基化位点[18]。

1.4 分子动力学模拟

采用 GROMACS 2023.2 软件包进行全原子分

子动力学模拟[19]。蛋白质-碳水化合物复合物由 CHARMM36m 力场描述,水分子采用 TIP3P 显式水模型[20]。体系经能量最小化和多步平衡(NVT 与 NPT)后,在恒温恒压系综(NPT)下进行 100 ns 的生产模拟,温度控制为 300 K,压力控制为 1 bar。轨迹分析主要基于平衡阶段(40~100 ns)的数据。通过计算均方根偏差(RMSD)、回转半径(Rg)及溶剂可及表面积(SASA)评估复合物的整体稳定性和结构致密性。残基均方根波动(RMSF)用于识别柔性区域。氢键分析采用标准几何准则:供体-受体距离小于 3.5 Å,且供体-氢-受体角度大于 120°。为解析复合物在构象空间中的运动模式,在去除整体平移和旋转后,对 Cα 原子进行主成分分析(PCA),并基于前两个主成分(PC1/PC2)构建自由能形貌。配体结合收敛性通过轨迹构象分布图(即木聚糖的构象聚类密度图)验证。此外,通过追踪水分子的运动轨迹,分析了蛋白质-木聚糖界面处的水合模式。

1.5　原核表达与蛋白纯化

编码候选蛋白的基因序列由生工生物工程(上海)股份有限公司合成,并克隆至 pET-28a$^+$载体构建重组表达质粒。将重组质粒转化至 DH5α 菌株,接种于含卡那霉素的 LB 液体培养基中,在 37 ℃、180 r·min^{-1}条件下过夜扩增。采用质粒提取试剂盒(生工生物,上海)抽提质粒,并将纯化后的质粒送测序验证。序列正确的重组质粒随后转化至大肠杆菌 DE3 表达菌株,用于蛋白表达[21]。

为确定候选蛋白的最优表达条件,分别测试了不同诱导温度(16、23、30、37、44、51 ℃)。在最佳温度下,设置 IPTG 浓度梯度(0.5‰、1.0‰、1.5‰、2.0‰、2.5‰)以筛选最适诱导剂浓度。在此基础上,进一步考察不同诱导时间(0、4、8、12、16、20、24 h)对表达量的影响。采用 12%分离胶进行 SDS-PAGE 分析,以空载体转化的 DE3 菌株作为阴性对照。当 DE3 培养物的 OD_{600} 达到 0.6~0.8 时,加入最适浓度的 IPTG,继续培养 16 h 以诱导目标蛋白表达[22]。诱导结束后,于 8 000 r·min^{-1}离心 10 min 收集菌体,加入 30 mL 裂解缓冲液(50 mmol·L^{-1} Tris,500 mmol·L^{-1} NaCl,pH 7.5),于 4 ℃裂解 12 h。随后将裂解液在液氮与 37 ℃水浴之间反复冻融 3 次。最后,使用超声破碎仪(功率 90 W)对样品进行超声破碎。破碎后的产物于 8 000 r·min^{-1}离心 10 min,上清液经 0.22 μm 滤膜过滤,利用 HisTrap HP 预装柱(GE Healthcare Life Sciences, Logan, UT)纯化目标蛋白。纯化后的蛋白样品经 12% SDS-PAGE 分离,考马斯亮蓝染色显影。将纯化的候选蛋白置换至 PBS 缓冲液中,用于后续酶活性测定。

1.6　多糖合成酶活力测定

纯化蛋白的浓度采用 BCA 试剂盒(生工生物工程(上海)股份有限公司)参照制造商说明书进行定量。酶活性测定采用 3,5-二硝基水杨酸(DNS)法。首先建立还原糖标准曲线:分别取 1、2、3、4、5 mg·mL^{-1}的还原糖溶液各 300 μL,依次加入 200 μL 去离子水和 500 μL DNS 试剂,混匀后测定吸光度。酶促反应体系如下:取 200 μL 纯化酶液,加入 300 μL 底物溶液(底物溶于 0.05 mol·L^{-1}柠檬酸-柠檬酸钠缓冲液,pH 4.8),于 37 ℃反应 30 min。反应终止时迅速加入 500 μL DNS,沸水浴加热 5~10 min,立即置于冰上冷却至室温,随后在 540 nm 波长处测定吸光度(OD_{540})[23]。根据标准曲线计算反应生成的还原糖量。

酶活性单位(U)定义为:在上述最适条件下,每分钟催化 1 μmol 底物转化所需的酶量。酶比活力以每毫克蛋白所具有的酶活性单位数(U/g)表示。比活力的计算公式如下:

$$比活力\ (U/g)=(c\times K\times V\times 1\ 000)/(M\times t\times m)。$$

式中:c 为酶产物浓度(mg·mL^{-1});K 为底物溶液的稀释倍数;V 为底物体积(mL);M 为产物的摩尔质量(g·mol^{-1});t 为酶反应时间(min);m 为酶的质量(g);1 000 为毫摩尔(mmol)与微摩尔(μmol)之间的换算系数[24]。所有酶活性测定结果均为 3 次重复实验的平均值。

1.7　碳水化合物酶活性测定

使用 pH 为 4.8 的柠檬酸-柠檬酸钠缓冲液制备浓度为 1%的碳水化合物溶液,底物包括葡萄糖、木糖、蔗糖、阿拉伯糖、乳糖、麦芽糖、麦芽三糖、蜜二糖、海藻糖、水杨苷、熊果苷等。在金属离子的终浓度为 5 mmol·L^{-1}时,评估 Zn^{2+}离子对 CAZymes 的活性影响。

1.8　统计分析及可视化

酶活性测定所得数据均为 3 次重复测量的平均值,所有结果以 3 次独立试验($n=3$)的均值±标准差(SD)表示。采用 SPSS 23.0 进行统计分析。相关图表则借助在线工具 CNSknowall (https://cnsknowall.com/index.html#/HomePage) 与 chiplot (https://www.chiplot.online/) 完成可视化及美化处理[25]。

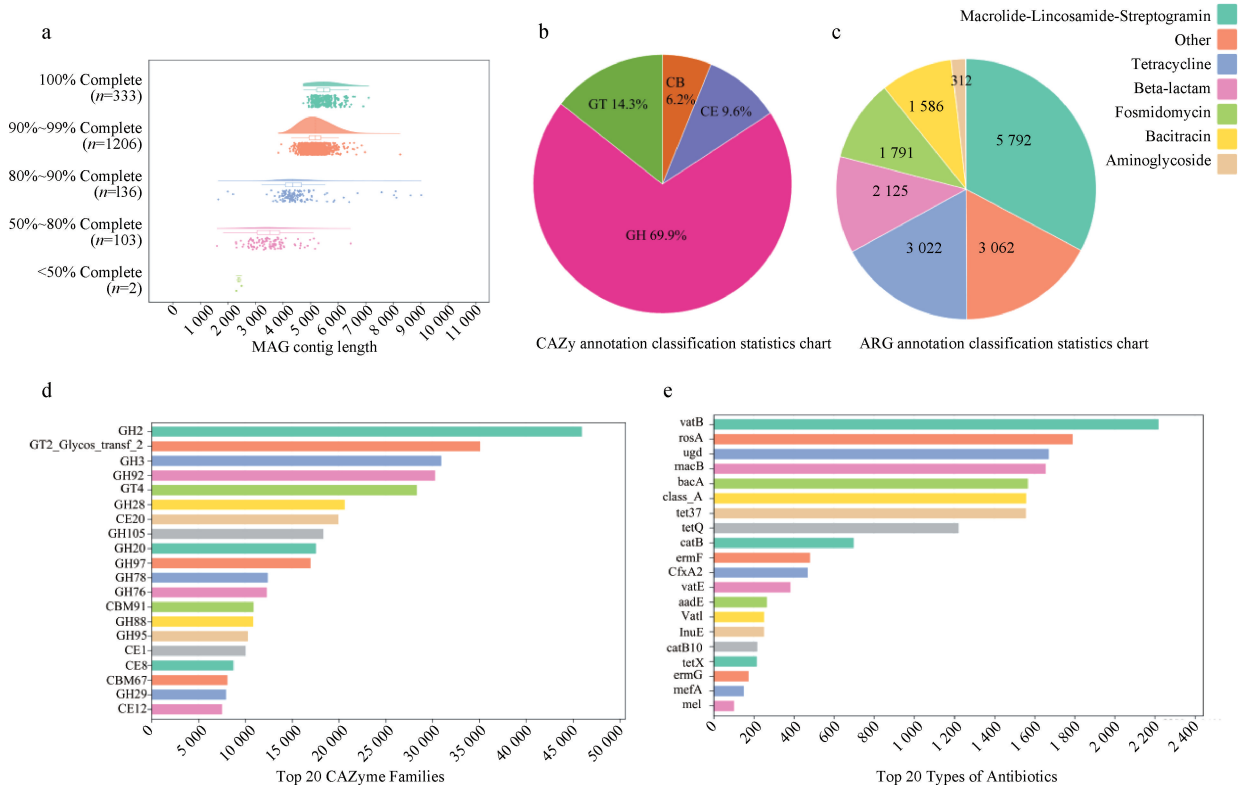
2 结果与分析

2.1 卵形拟杆菌 MAGs 的 CAZymes 与 ARGs 概述

我们使用 CheckM2 对 1 780 个 *B. ovatus* 进行质量评估,其中有 334 个 MAGs 归类为高质量(完整性>90%且污染<5%)、1 445 个为中等质量(完整性 50%~90%且污染<10%)、2 个为低质量(完整性<50%)(图 1a)。中、高质量 MAGs 占总 MAGs 的 99.9%。

我们对 *B. ovatus* 的 MAGs 进行了 CAZymes 注释,结果表明,这些 *B. ovatus* 共编码 724 567 个 CAZymes 基因,其中糖苷水解酶(GH)占比最高(506 389,69.9%),其次为糖基转移酶(GT,103 975,14.3%)(图 1b),GT 作为第二大类别,该菌具备可观的多糖合成能力,可能参与胞外多糖、荚膜多糖或脂多糖等结构多糖的生物合成。家族丰度分析显示,GH9(230 946)、GH2(45 972)、GH3(35 101)、GH4(30 313)和 GH28(28 354)是数量最多 GH 亚家族;此外,CE20(20 632)、GH105(19 974)、GH20(18 397)、GH97(17 576)、GH18(16 990)及多个 CBM 家族(CBM91:12 452;CBM67:10 071)也呈现高丰度(图 1d)。这些酶类主要参与植物多糖(纤维素、半纤维素、果胶)及宿主聚糖降解,与拟杆菌属适应肠道生态位特征一致。同时,高丰度 GT 家族成员(如 GT2、GT4 等)为多糖链延伸与合成提供了酶学基础,表明该菌在降解多糖获取能量同时,也可能通过合成多糖参与生物被膜形成或宿主互作。

B. ovatus 是人类肠道中重要的潜在益生菌,但是,我们发现 *B. ovatus* 中存在着大量的抗生素抗性基因(ARG)。从 *B. ovatus* MAGs 中共鉴定出了 17 667 个 ARG,分属 8 大类抗生素抗性。其中,大环内酯-林可酰胺-链阳菌素(MLS)抗性基因数量最多(5 792 个,32.8%),其次为四环素类(3 022 个,17.1%)、β-内酰胺类(2 125 个,12.0%)等(图 1c)。在 ARG 亚型水平上,vatB(2 218 个)、rosa(1 791 个)和 ugd(1 671 个)等为优势亚型(图 1e)。rosa 为磷霉素抗性基因,vatB 属于 MLS 抗性基因,而 bacA 与杆菌肽抗性相关,这些抗性基因极大的促进了 *B. ovatus* 在肠道中的定植。值得注意的是,ugd(UDP-葡萄糖脱氢酶)和 bacA 不仅参与抗生素耐药性,还在细菌多糖合成中发挥关键作用,这进一步印证了该菌在多糖代谢与抗性调控之间的潜在联系。



a:拟杆菌属 MAGs 的质量评估和长度分布,横轴表示基因组长度,纵轴表示基因组的完整性;b:拟杆菌属物种的碳水化合物活性酶(CAZymes)注释;c:抗生素抗性基因(ARG)的注释;d:物种丰度最高的 20 个 CAZyme 家族;e:物种注释频率前 20 位的抗生素类别。

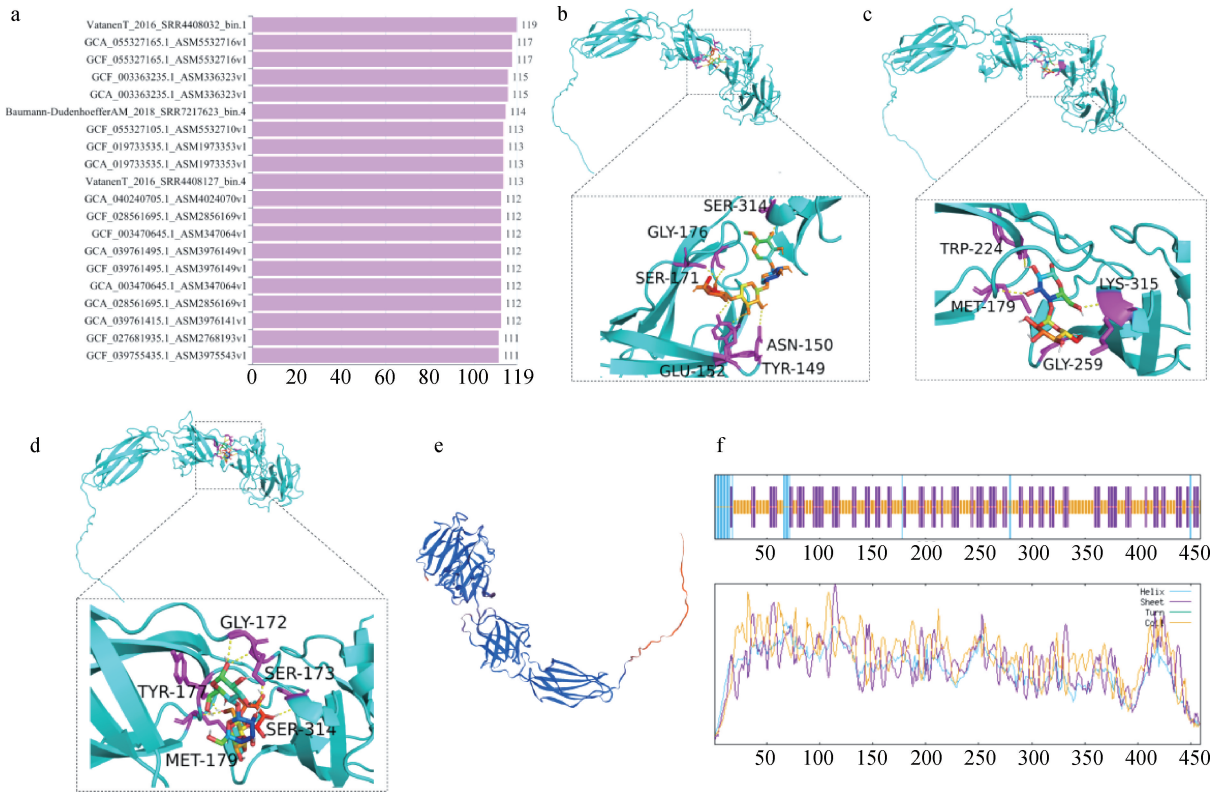
图 1 拟杆菌属基因组质量评估及其碳水化合物活性酶与抗生素分析

Figure 1 Quality assessment of the genomes of the *Bacteroides* genus and analysis of its carbohydrate-active enzymes and antibiotic resistance genes

2.2　候选蛋白的筛选

我们利用 PULpy 对 *B. ovatus* 进行了 PULs 注释,发现共计编码了 167 560 个 PULs,其中 VatanenT_2016_SRR4408032_bin.1 编码了最多的 PULs,共计 119 个(图 2a)。我们提取了 VatanenT_2016_SRR4408032_bin.1 编码的 PULs 上的所有未知蛋白,并将它们与 NR 数据库进行比对得到了 12 条潜在的蛋白序列。使用 AlphaFold 3 将这 12 个蛋白序列转化为 3D 的 pdb 结构,使用 AutoDock Vina 将这 12 个蛋白分别与多种代表性碳水化合物配体进行了分子对接。

我们发现 VatanenT_2016_SRR4408032_bin.1_k119_6428_3 具有最强的碳水化合物结合能力(表 1),命名为 V6428。



a:编码最多潜在新型 CAZymes 的前 20 个基因组;b:V6428 与木聚糖分子对接的可视化图像;c:V6428 与麦芽糖分子对接的可视化图像;d:V6428 与麦芽四糖分子对接的可视化图像;e:V6428 的三维结构;d:V6428 的二级结构预测。

图 2　蛋白 V6428 的筛选及结构预测

Figure 2　Screening and structural prediction of protein V6428

表 1　分子对接结果

Table 1　Molecular docking results

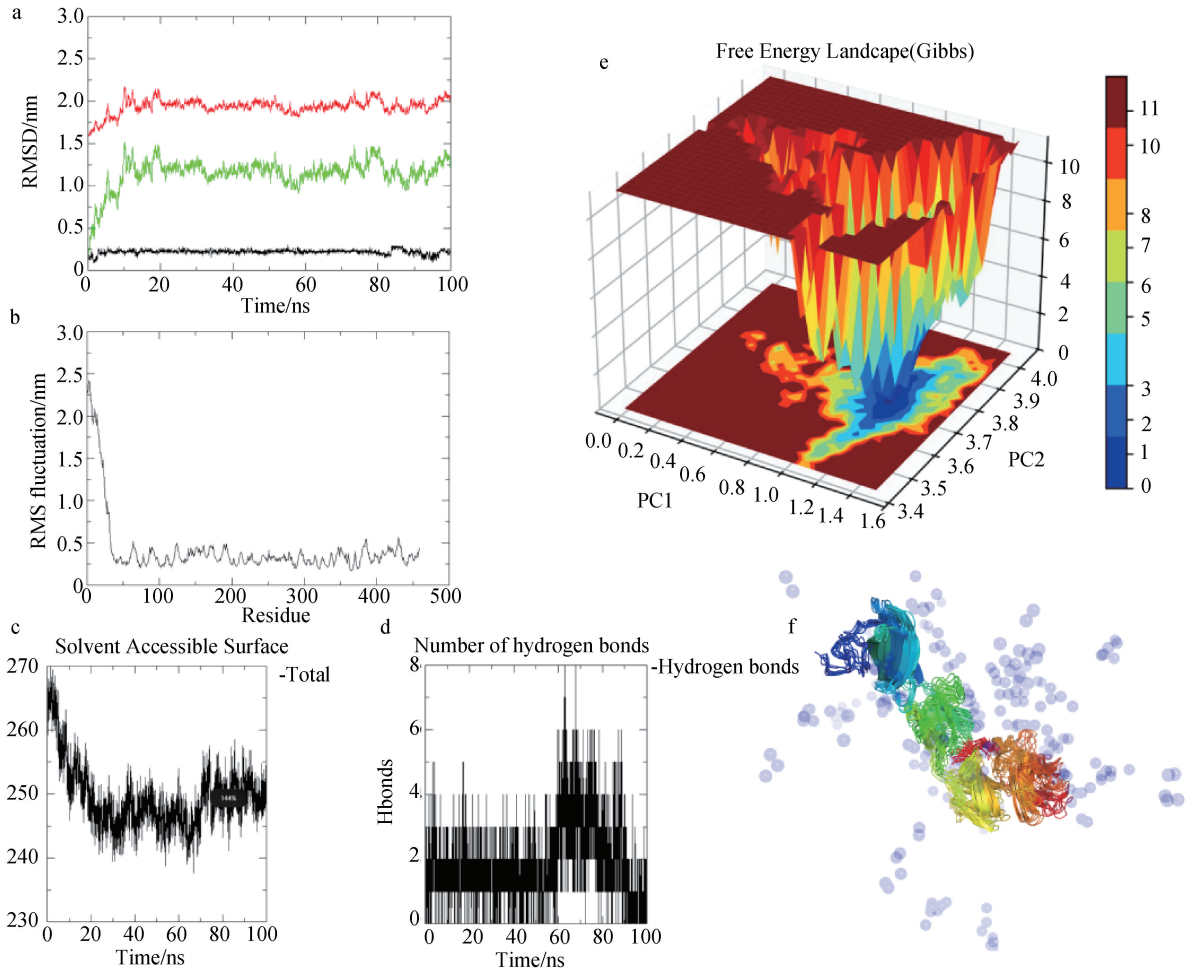
ID	支链淀粉	麦芽三糖	麦芽四糖	木聚糖	棉子糖	蔗糖	海藻糖	cmc-na
VatanenT_2016_SRR4408032_bin.1_k119_7990_8	-6.453	-6.512	-7.296	-7.156	-6.208	-5.702	-6.279	-4.947
VatanenT_2016_SRR4408032_bin.1_k119_6428_3	-6.481	-6.376	-6.772	-7.269	-6.869	-6.053	-6.069	-3.938
VatanenT_2016_SRR4408032_bin.1_k119_48103_9	-7.476	-5.899	-6.901	-5.299	-5.302	-5.549	-5.275	-4.295
VatanenT_2016_SRR4408032_bin.1_k119_38249_2	-5.962	-5.981	-5.839	-5.93	-5.698	-4.651	-5.15	-2.334
VatanenT_2016_SRR4408032_bin.1_k119_37670_6	-6.555	-6.256	-5.364	-6.011	-6.269	—	-2.982	-1.303
VatanenT_2016_SRR4408032_bin.1_k119_36805_8	-6.126	-5.588	-6.659	-6.115	-5.42	-5.459	-5.269	-3.845
VatanenT_2016_SRR4408032_bin.1_k119_27061_3	-5.464	-5.513	-5.72	-5.473	-5.329	-4.455	-4.228	-4.113
VatanenT_2016_SRR4408032_bin.1_k119_25291_16	-4.864	-4.871	-5.431	-5.402	-4.699	-3.165	-5.054	—
VatanenT_2016_SRR4408032_bin.1_k119_24282_11	-5.484	-5.273	-5.483	-5.114	-4.491	-4.546	-4.167	-3.425
VatanenT_2016_SRR4408032_bin.1_k119_13751_27	-6.396	-6.13	-6.658	-7.358	-5.689	-6.292	-5.802	-3.687
VatanenT_2016_SRR4408032_bin.1_k119_13751_25	-5.111	-5.097	-5.471	-4.857	-4.105	-3.934	-3.853	-2.693

其中 V6428 与木聚糖的结合能最强,为-7.269 kcal/mol(图 2b),其次是麦芽三糖(-6.25 kcal·mol^{-1})(图 2c)和麦芽四糖(-6.1 kcal·mol^{-1})(图 2d)。利用 SOPMA 和 AlphaFold 3 分别预测了 V6428 的二

级和三级结构(图 2e)。结果表明,该蛋白由 459 个氨基酸组成,其中 28 个氨基酸形成 α 螺旋结构、264 个氨基酸形成无规则卷曲结构、167 个氨基酸形成延伸链结构,分别占二级结构的 6.10%、57.52%和 36.38%(图 2f),我们观察到 V6428 含有 Zn^{2+} 保守结构域,此外 V6428 还含有一个 DUF5016 结构域,这是来源于植物的一个未知结构域,进一步证实了 V6428 的潜在活性,我们这些结果为后续 V6428 酶功能的实验验证提供了重要的理论基础。

2.3 候选蛋白分子动力学模拟

根据分子对接结果,蛋白 V6428 与木聚糖的结合能最低,表明两者具有最强的亲和力。为了进一步从原子运动角度解释该复合物的动态稳定性及相互作用机制,我们对该对接复合物进行了 100 ns 的分子动力学模拟。在整个模拟过程中,复合物的骨架 RMSD 在 0.5~1.0 nm 之间平稳波动,并在约 40 ns 后趋于平衡,未出现剧烈偏离,说明蛋白-木聚糖复合物结构稳定(图 3a)。此外,大部分残基的波动幅度低于 0.2 nm,少数环区残基(约第 50~70 位及第 120~140 位)表现出较高的柔性(>0.5 nm),这些区域可能参与底物识别或诱导契合(图 3b)。溶剂可及表面积在模拟期间维持在相对恒定的水平(总表面积约 220 nm^2),未发生明显增大或减小,提示蛋白折叠核心完整,且木聚糖结合未引起蛋白去折叠(图 3c)。并且,复合物在模拟中平均形成约 10~15 个氢键,且数目波动较小,表明存在稳定且持续的氢键网络(图 3d)。以主成分 PC1 和 PC2 为坐标,呈现出单一且陡峭的低能谷(最小吉布斯自由能约 0.0~0.2),复合物的构象主要集中在一个全局能量最小值附近,动力学稳定性良好(图 3e)。小分子轨迹构象分布图进一步显示木聚糖在模拟过程中主要聚集于单一低能构象盆地,未出现多态分布,证实其结合模式高度收敛(图 3f)。



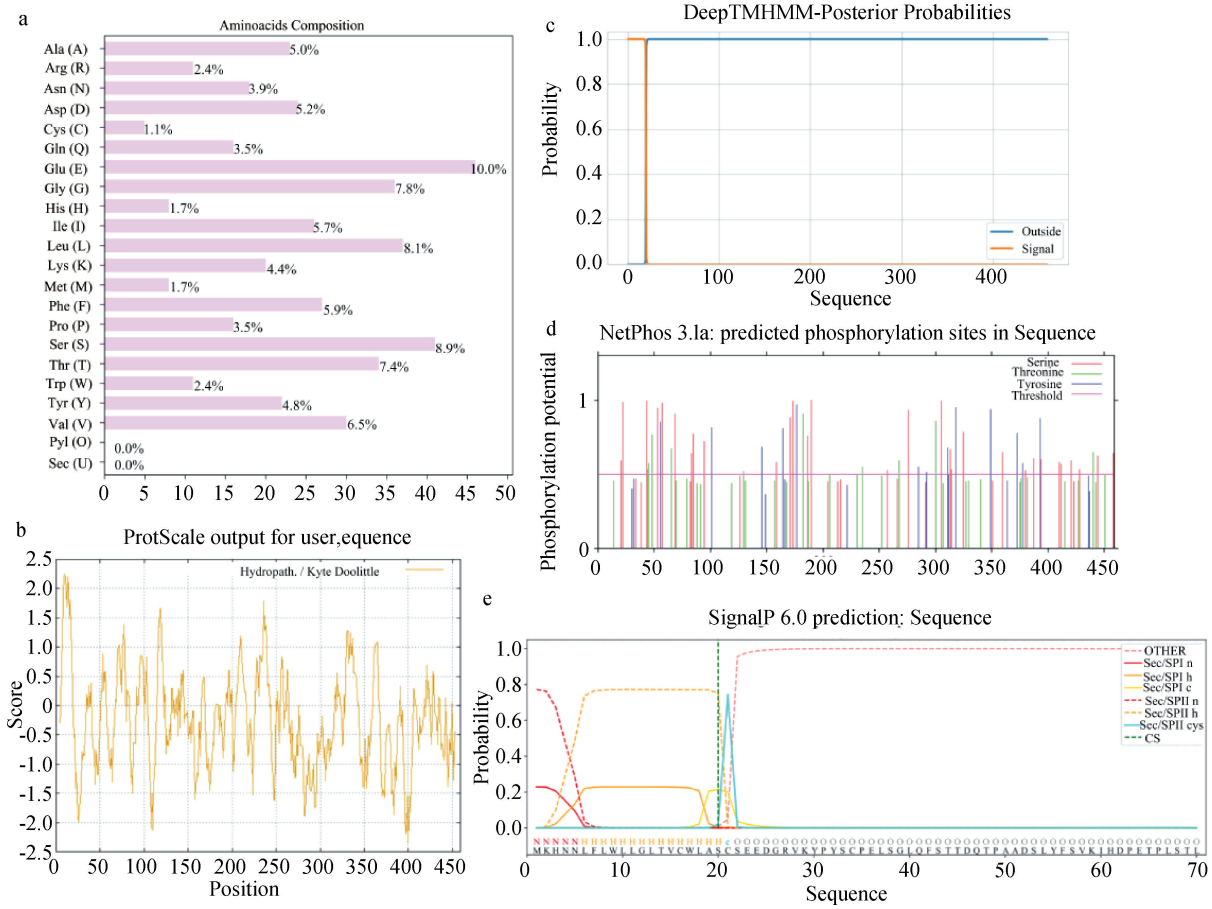
a:复合物骨架均方根偏差(RMSD)随时间的变化;b:复合物中氨基酸残基的均方根波动(RMSF)分布;c:溶剂可及表面积(SASA)随模拟时间的变化;d:复合物间氢键数目随模拟时间的变化;e:基于主成分分析(PCA)的自由能形貌图;f:木聚糖小分子在模拟过程中的轨迹构象分布。

图 3 V6428-木聚糖复合物的分子动力学模拟分析

Figure 3 Molecular dynamics simulation analysis of the V6428-xyloglucan complex

2.4　候选蛋白的理化性质

为系统评估 V6428 蛋白的理化性质及翻译后修饰潜能,我们使用 ProtParam 进行的生物信息学分析显示,V6428 基因编码一个由 459 个氨基酸组成的蛋白质,理论分子量为 51.73 kDa(化学式:$C_{2337}H_{3497}N_{573}O_{731}S_{13}$)。该蛋白呈酸性,理论等电点(pI)为 4.42。它含有 70 个带负电的残基(天冬氨酸+谷氨酸)和 31 个带正电的残基(精氨酸+赖氨酸)(图 4a)。不稳定指数为 37.09,将 V6428 归类为稳定蛋白;脂肪族指数为 77.49,总平均亲水性(GRAVY)为 −0.270,分别表明其具有良好的热稳定性和整体亲水性特征。ProtScale 分析显示其具有亲水性,平均亲水指数为 −0.270,且未出现显著的高疏水峰区(图 4b)。使用 DeepTMHMM 进行的生物信息学分析显示,V6428 没有跨膜结构域,证实其是一种非跨膜蛋白(图 4c)。NetPhos 3.1 预测了 52 个高置信度磷酸化位点,其中包括 30 个丝氨酸、13 个酪氨酸和 9 个苏氨酸残基(图 4d),提示磷酸化介导的信号传导具有广泛的调控潜力。SignalP 6.0 信号肽预测,结果显示 V6428 最可能属于脂蛋白信号肽(Sec/SPII)类型,其切割位点位于第 20~21 位氨基酸之间(概率 0.751 078),各类信号肽似然值中 Sec/SPII 最高(0.774 8),而普通分泌信号肽(Sec/SPI)仅为 0.224 8,TAT 信号肽(Tat/SPI、Tat/SPII)及菌毛样信号肽(Sec/SPIII)均接近 0(图 4e),提示该蛋白可能经 Sec 途径转运并以脂蛋白形式锚定于细胞膜。这些结果表明 V6428 为酸性、稳定、亲水的非跨膜蛋白,富含磷酸化位点且携带典型脂蛋白信号肽,这些特征为其功能表达、亚细胞定位及调控机制研究提供了重要依据。



a:V6428 蛋白的氨基酸组成及带电残基分布;b:V6428 蛋白的亲水性分析;c:V6428 蛋白的跨膜结构域预测;d:V6428 蛋白的磷酸化位点预测;e:V6428 蛋白的信号肽类型预测。

图 4　V6428 蛋白的理化性质及翻译后修饰预测分析

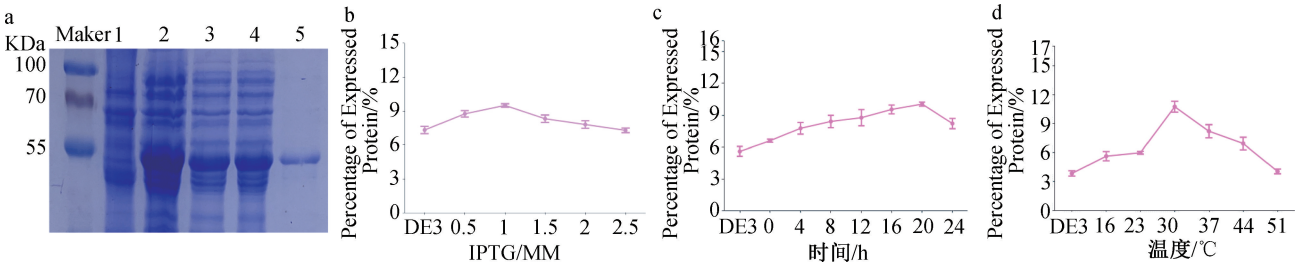
Figure 4　Prediction analysis of the physicochemical properties and post-translational modifications of V6428 protein

2.5　重组蛋白 V6428 的表达、纯化及酶学特性

为进一步验证 V6428 的活性,我们合成了 V6428 基因并将其连接至 pET28a(+)质粒载体。将重组质粒 pET28a-V6428 导入大肠杆菌 BL21(DE3)中,并通过限制性内切酶 NcoI 和 XhoI 进行双酶切鉴定。通过加入 IPTG 至终浓度 1 mmol·L^{-1},在 37

℃下诱导重组蛋白表达。重组 V6428 蛋白的 SDS-PAGE 分析显示在预期大小(52 kDa)处有明显条带,证实了重组 V6428 蛋白的成功表达。随后使用预装 HisTrap HP(GE Healthcare Life Sciences, Logan, UT)镍柱纯化重组 V6428 蛋白,获得了分子量约为 52 kDa 的单一条带(图 5a)。

为确定 V6428 的最佳表达条件,我们系统评估了 IPTG 浓度、诱导温度和诱导时间的影响。通过系统优化诱导参数,我们确立了表达重组 V6428 蛋白的最佳条件。IPTG 浓度的初步筛选(16 h,37℃)显示,1.0 mmol·L^{-1}为最佳诱导浓度,经 ImageJ 灰度分析定量,目标蛋白表达量占细胞总蛋白的 9.51%(图 5b)。然而,更高的 IPTG 浓度(2.5 mmol·L^{-1})导致表达水平下降,提示诱导剂毒性产生的代谢负担效应。随后使用确定的 1.0 mM IPTG 进行温度优化,结果显示在 30℃诱导 16 h 条件下 V6428 产量最高为(10.73%),在更高温度下观察到产量下降(图 5c)。在这些优化条件(30℃下 1.0 mmol·L^{-1} IPTG)下进行的最终时间进程实验确定最佳诱导时间为 20 h,此时 V6428 占细胞总蛋白 10.05%(图 5d)。超过此时间后表达量下降,可能反映了蛋白质不稳定性或培养基中营养耗竭。这些优化参数(1.0 mmol·L^{-1} IPTG,30 ℃,诱导 20 h)为 V6428 的大规模生产和后续功能表征奠定了基础。



a:V6428 的表达与纯化,M 代表蛋白质分子量标准图,第 1 行是未诱导细胞的总蛋白含量,第 2 行是诱导细胞的总蛋白含量,第 3 行是总蛋白在匀浆细胞的上清液中被诱导产生,第 4 行是细胞匀浆后的沉淀物中的总蛋白也被诱导产生,第 5 行是用镍柱亲和层析纯化所得蛋白质;b:不同诱导剂浓度下目标蛋白在总蛋白中的比例;c:不同诱导时间下目标蛋白在总蛋白中的比例;d:不同诱导温度下目标蛋白在总蛋白中的比例。

图 5 V6428 的原核表达及酶学特性分析

Figure 5 Analysis of the prokaryotic expression and enzymatic properties of V6428

底物特异性分析显示,V6428 能显著合成多糖。其对木糖的活性最高(4 927.54 units·g^{-1})(表 2),此外,V6428 还对纤维二糖、麦芽三糖、α 乳糖、麦芽糖、葡萄糖和木糖具有合成作用,并且在终浓度为 5 mmol·L^{-1}的 Zn^{2+}溶液中,V6428 对纤维二糖合成活性从 2 072.46 U 显著上升至 3 601.45 U,提示 Zn^{2+}对该酶的糖链合成功能具有促进作用,酶活性测定结果详见表 2。

表 2 酶活性测定结果表

Table 2 Results of enzyme activity measurement

U/g	纤维二糖	麦芽三糖	α-乳糖	麦芽糖	葡萄糖	纤维二糖+Zn^{2+}	木糖
酶活	2 072.46±132.83	2 833.33±527.59	2 775.36±925.98	3 478.26±200.16	4 398.55±836.98	3 601.45±570.21	4 927.54±119.73

3 讨论

碳水化合物活性酶(CAZymes)是微生物参与多糖降解、转化及合成代谢的核心酶系,主要包括糖苷水解酶(GHs)与糖基转移酶(GTs)两大功能家族,在肠道菌群的碳源利用、微生态互作以及功能性多糖的生物合成中发挥关键作用[26]。肠道共生菌的基因组通常富集大量 CAZymes 编码基因,其中拟杆菌门微生物尤其典型:它们通过多糖利用位点(PULs)协同调控 GH 与 GT 家族基因的表达,既能够利用外源多糖维持自身生长与定植,又可合成结构多糖,以参与宿主免疫互作与肠道微生态稳态的维持。本研究从卵形拟杆菌(*B. ovatus*)MAGs 中鉴定出大量碳水化合物活性酶,其中糖基转移酶(GT)家族占比达 14.3%,是仅次于糖苷水解酶(GH)的第二大类 CAZymes。这一高丰度的 GT 家族成员(如 GT2、GT4 等)为多糖链的延伸与合成提供了酶学基础,提示卵形拟杆菌在降解多糖获取能量的同时,也具备可观的多糖合成能力,可能参与胞外多糖(EPS)、荚膜多糖(CPS)或脂多糖(LPS)等结构多糖的生物合成。已有研究表明,拟杆菌属的荚膜多糖不仅参与宿主-微生物互作,还在生物被膜形成和肠道定植中发挥关键作用[27]。因此,卵形拟杆菌中丰富的 GT 基因资源可能与其适应肠道环境、逃

避免疫识别及维持肠道稳态密切相关。

B. ovatus 是人类肠道中重要的潜在益生菌,但是本研究还发现了大量抗生素抗性基因(ARG),其中 ugd(UDP-葡萄糖脱氢酶)和 bacA 不仅参与抗生素耐药,还在多糖合成中发挥双重功能。ugd 催化 UDP-葡萄糖转化为 UDP-葡萄糖醛酸,是荚膜多糖和胞外多糖合成途径中的关键酶。这一发现进一步强化了多糖代谢与抗性调控之间的功能联系,暗示卵形拟杆菌可能通过多糖合成相关基因的水平转移获得抗性优势。此外,抗性基因往往位于可移动元件上,这意味着卵形拟杆菌不仅自己能抵抗威胁,还可能将这些"武器基因"转移给其他肠道细菌,从而在更广的范围内重塑整个肠道微生态的平衡,存在着一些潜在的风险。因此,挖掘出 *B. ovatus* 编码的高效的功能蛋白以替代 *B. ovatus* 势在必行。

利用分子对接,我们发现 V6428 对多种底物都具有较强的碳水化合物结合能力,表明其可能是一个多功能的碳水化合物酶,对木聚糖具有最强的结合亲和力(-7.269 kcal·mol^{-1}),其次是麦芽三糖和麦芽四糖。值得注意的是,木聚糖是一种由木糖单元通过 β-1,4 糖苷键连接形成的半纤维素多糖,而 V6428 对木聚糖的高亲和力提示该酶可能参与木聚糖链的延伸或修饰。100 ns 分子动力学模拟进一步证实了蛋白-木聚糖复合物的结构稳定性:RMSD 平稳波动,平均形成 10~15 个氢键,自由能景观呈现单一低能谷,表明木聚糖以高度收敛的模式结合于 V6428 的活性口袋。由于木聚糖难溶于水,因此,在 V6428 的活性测定中,我们使用了木糖,并且发现了 V6428 具有将木糖合成为多聚的木糖的能力。

其中对木聚糖的合成活性最高(4 927.54 units/g),而对其他多种糖底物也表现出不同程度的合成能力。这一结果明确表明 V6428 是一种具有底物偏好的多糖合成酶,而非水解酶。尽管 CAZymes 注释中 GH 家族占主导,但 V6428 作为低相似度序列,其功能更接近于糖基转移酶,可能参与木聚糖骨架的延伸或支链修饰。在 V6428 的结构预测中,我们发现它具有一个 Zn^{2+} 结合结构域。因此,我们添加了 5 mmol·L^{-1} Zn^{2+},发现其能够增强 V6428 活性,这可能与 Zn^{2+} 作为糖基转移酶的辅助因子稳定催化构象有关[28]。

综合上述结果,本研究利用分子对接及分子动力学模拟技术从卵形拟杆菌中鉴定出一个新型高效的 GT 家族基因 V6428,并成功表征,该酶以木聚糖为最适底物,是一种金属酶。这一发现不仅丰富了卵形拟杆菌多糖合成酶的功能图谱,也为微生物胞外多糖的酶法合成提供了新的候选酶资源,具有潜在的生物技术应用价值。

4　结论

本研究从卵形拟杆菌 MAGs 中获得高质量基因组序列,解析了其碳水化合物活性酶(CAZymes)与抗生素抗性基因(ARG)组成。CAZymes 中糖苷水解酶(GH)占 69.9%,糖基转移酶(GT)占 14.3%,高丰度 GT 家族及双功能 ARG(ugd、bacA)提示多糖合成与抗性调控存在功能联系。利用分子对接技术筛选出了新型蛋白 V6428,并利用分子动力学模拟进行了进一步的验证。最后,利用原核表达技术对其进行了表征,并测定其活性,发现 V6428 是一个新型的高效的多糖合成酶。V6428 的发现为多糖合成相关产业提供了新候选酶资源。

参考文献(References)

[1] Yue B J, Zong G F, Tao R Z, et al. Crosstalk between traditional chinese medicine-derived polysaccharides and the gut microbiota: a new perspective to understand traditional chinese medicine[J]. Phytotherapy Research, 2022, 36(11): 4125-4138.

[2] Megur A, Daliri E B M, Baltriukienė D, et al. Prebiotics as a tool for the prevention and treatment of obesity and diabetes: Classification and ability to modulate the gut microbiota[J]. International Journal of Molecular Sciences, 2022, 23(11): 6097.

[3] Liu S G, Zhang Z S, Gu P, et al. The effect of extracellular polymeric substances on the distribution and transmission of antibiotic resistance genes treating antibiotic wastewater via microbial electrolysis cells[J]. Chemosphere, 2024, 364(10): 143284.

[4] Li W, Zeng J Y, Li Y Y, et al. A fascinating finding: The application of traditional Chinese medicine in aquaculture prevents the spread and diffusion of antibiotic resistance genes among gut microbes[J]. Aquaculture, 2024, 583(7): 740573.

[5] Ekstrom A, Taujale R, McGinn N, et al. Plant CAZyme: A database for plant carbohydrate-active enzymes[J]. Database, 2014, 20(14): 1-8.

[6] Nauom S, Da Silva Neto B R, Ribeiro M S, et al. Biochemical and molecular study of Trichoderma harzianum enriched secretome protein profiles using lectin affinity chromatography[J]. Applied Biochemistry and Biotechnology, 2019, 187(1): 1-13.

[7] Hayase E, Hayase T, Mukherjee A, et al. Bacteroides ovatus alleviates dysbiotic microbiota-induced graft-ver-

sus-host disease[J]. Cell Host & Microbe, 2024, 32(9):1621-1636.

[8] Martens E C, Roth R, Heuser J E, et al. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont [J]. Journal of Biological Chemistry, 2009, 284(27): 18445-18457.

[9] Chklovski A, Parks D H, Woodcroft B J, et al. CheckM2: A rapid, scalable and accurate tool for assessing microbial genome quality using machine learning [J]. Nature Methods, 2023, 20(8):1203-1212.

[10] Olm M R, Brown C T, Brooks B, et al. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication[J]. The ISME Journal, 2017, 11(12): 2864-2868.

[11] Chaumeil P A, Mussig A J, Hugenholtz P, et al. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database[J]. Bioinformatics, 2020, 36(6): 1925-1927.

[12] Zhang H, Yo H T, Huang L, et al. dbCAN2: A meta server for automated carbohydrate-active enzyme annotation[J]. Nucleic Acids Research, 2018, 46(1):95-101.

[13] Li Y C, Lu C L. BLASTP-ACC: Parallel architecture and hardware accelerator design for blast-based protein sequence alignment[J]. IEEE Transactions on Biomedical Circuits and Systems, 2019, 13(6):1771-1782.

[14] Trott O, Olson A J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading[J]. Journal of Computational Chemistry, 2010, 31(2):455-461.

[15] Wilkins E M R, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools in the ExPASy server [J]. Methods in Molecular Biology, 1999(112):531-552.

[16] Terrapon N, Lombard V, Gilbert H J, et al. Automatic prediction of polysaccharide utilization loci in bacteroidetes species[J]. Bioinformatics, 2015, 31(5):647-655.

[17] Dyrløv Bendtsen J, Nielsen H, Von Heijne G, et al. Improved prediction of signal peptides: Signal P 3.0[J]. Journal of Molecular Biology, 2004, 340(4):783-795.

[18] Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function[J]. Pacific Symposium on Biocomputing, 2002, 31:310-322.

[19] Abraham M J, Murtola T, Schulz R, et al. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers[J]. SoftwareX, 2015,27(9):19-25.

[20] Huang J, Rauscher S, Nawrocki G, et al. CHARMM36m: An improved force field for folded and intrinsically disordered proteins[J]. Nature Methods, 2017, 14(1):71-73.

[21] Hu L L, Li X Y, Li C Y, et al. Characterization of a novel multifunctional glycoside hydrolase family in the metagenome-assembled genomes of horse gut[J]. Gene, 2024, 927(5):148758.

[22] Mohammadi S, Tarrahimofrad H, Arjmand S, et al. Expression, characterization, and activity optimization of a novel cellulase from the thermophilic bacteria Cohnella sp. A01[J]. Scientific Reports, 2022, 12(1):10301-10317.

[23] Ito D, Nakano E, Karita S, et al. Characterization of a GH family 43 β-xylosidase having a novel carbohydrate-binding module from Paenibacillus xylaniclasticus strain TW1[J]. Journal of Applied Glycoscience, 2022, 69(3):65-71.

[24] Deshavath N N, Mukherjee G, Goud V V, et al. Pitfalls in the 3, 5-dinitrosalicylic acid (DNS) assay for the reducing sugars: Interference of furfural and 5-hydroxymethylfurfural[J]. International Journal of Biological Macromolecules, 2020, 156(15):180-185.

[25] Huang H T, Lyu W Q, Xu F Y, et al. Mechanism of yiqi huoxue huatan recipe in the treatment of coronary atherosclerotic disease through network pharmacology and experiments[J]. Medicine, 2023, 102(26):e34178.

[26] Berlemont R. The supragenic organization of glycoside hydrolase encoding genes reveals distinct strategies for carbohydrate utilization in bacteria[J]. Frontiers in Microbiology, 2023, 14(7):1179206.

[27] Béchon N, Ghigo J M. Gut biofilms: Bacteroides as model symbionts to study biofilm formation by intestinal anaerobes[J]. FEMS Microbiology Reviews, 2022, 46(2):54-68.

[28] Javaid T, Bhattarai M, Venkataraghavan A, et al. Specific protein interactions between rice members of the GT43 and GT47 families form various central cores of putative xylan synthase complexes[J]. The Plant Journal, 2024, 118(3):856-878.

(责任编辑:唐慧)