

文章编号:1000-2367(2020)02-0094-06

DOI:10.16366/j.cnki.1000-2367.2020.02.014

鞘氨醇单胞菌中威兰胶合成基因 *welS* 和 *welG* 的信息学研究

李丽¹,常爱平¹,钱锦²,杨衍钰¹,何侨妹¹,沈亚领²,王颖璐¹,朱虎¹

(1.福建师范大学 工业生物催化福建省高校工程研究中心;福建省先进材料化工基础重点实验室;
化学与材料学院,福州 350007;2.华东理工大学 生物反应器工程国家重点实验室,上海 200237)

摘要:威兰胶是一种由微生物合成的天然高分子多聚糖,具有优异的流变性能和良好的稳定性,广泛应用于石油开采、建筑材料、日用化学品、食品、医药等领域.本文运用生物信息学手段,对威兰胶合成过程中涉及到的关键基因 *welS* 和 *welG* 进行了研究,主要包括 *welS* 和 *welG* 基因序列及其编码蛋白 WELS 和 WELG 的性质及结构.结果表明 WELS 和 WELG 含有 412 和 545 个氨基酸,分子量分别为 43.1 kDa 和 59.5 kDa,等电点分别为 11.12 和 10.39,均为碱性疏水性蛋白.WELS 和 WELG 在结构域组成类型上具有相似性,左右结构域形成大致平行的角度,并且在 β -sheet 部分具有相似的“环形”结构.这些研究结果为建立威兰胶代谢合成调控新策略,提高威兰胶的发酵产量奠定了基础.

关键词:威兰胶;*welS*;*welG*;结构特性;生物信息学

中图分类号:Q503

文献标志码:A

威兰胶(Welan gum),又名韦兰胶或威伦胶,是由海洋来源鞘氨醇单胞菌(*Sphingomonas* sp.WG)和产碱杆菌(*Alcaligenes* sp.)在生长代谢过程中分泌的一种水溶性胞外多聚糖^[1-2].威兰胶与结冷胶、黄原胶等同属于鞘氨醇类^[3],由 D-葡萄糖、D-葡糖醛酸、D-葡萄糖和 L-鼠李糖共同组成威兰胶的结构骨架^[4],侧链一般连接单链 L-鼠李糖(占 2/3),也可能会连接单链 L-甘露糖(占 1/3),除此以外,一半左右的四糖片段上还带有乙酰基和甘油基团^[5].作为一种新型材料,威兰胶水溶液因其较高的热稳定性,并且对酸、碱稳定,在 pH 为 2~12 时,仍能保持黏度不变的独特性质^[6,7],因此在水泥、混泥土等建筑材料^[8]、日用化学品^[9]、食品^[10]、医药^[11,12]和石油开采领域^[13-15]中具有广泛的应用价值和潜在的发展前景.虽然目前鞘氨醇单胞菌发酵生产威兰胶的产量达到 39.95 g/L,葡萄糖的转化率也提高到 56.48%^[16,17],但相较于广泛使用的黄原胶(糖转化率达 70%~90%)^[18],糖的转化率还是相对较低.因此,提高威兰胶的产量仍然是威兰胶生产发展必不可少的一步.

要进一步提高威兰胶产量和降低其生产成本,深入解析其生物合成途径及调控机制有着至关重要的意义.对威兰胶发酵生产的合成路径已多次被报道^[5,7,19-20],然而有关威兰胶生物合成中涉及的遗传信息和基因组尚未得到很好的研究.通过对鞘氨醇单胞菌(*Sphingomonas* sp.ATCC 31555)的全基因组序列^[21]分析,明确了合成威兰胶的基因簇序列^[22],*wel* 簇的核心区域约为 24 kb,含有约 24 个基因,主要有四基因 *rml* 簇 *rmlABCD*、编码四糖重复单元组装的糖基转移酶的基因(*welB*,*welK*,*welL* 和 *welQ*),以及编码参与聚合、链长度调节和威兰胶的输出的一些酶的基因(如 *welS*,*welG*,*welC* 和 *welE*)^[23].

多糖生物合成的最后一步,是将多糖从细胞膜上分泌出去.这个转运过程需要能量,并且可能通过一个特殊的转运系统,该系统确保将高聚物从脂类载体上释放出来并且进行跨膜转运.推测的威兰胶四糖重复单元聚合和分泌的途径是 Wzy 型途径^[19,20],首先在同 Wzx 相似的转位酶 WelS 的作用下,重复单元从细胞质

收稿日期:2019-09-14;修回日期:2020-02-09.

基金项目:国家 863 计划(2015AA020925);国家自然科学基金(U1805234 and 31800075);福建省自然科学基金(2019J01264).

作者简介:李丽(1996—),女,四川南充人,福建师范大学在读硕士研究生,研究方向为环肽及多糖自组装.

通信作者:朱虎,E-mail:zhuhu@fjnu.edu.cn;王颖璐,E-mail:wangyl@fjnu.edu.cn.

中翻转经过内膜进入周质空间,然后是在 Wzy 相似的聚合酶 WelG 作用下,重复单元聚合在一起.但是目前对 *welS* 和 *welG* 基因的报道较少,对该基因编码蛋白的结构、理化性质的研究更是寥寥无几.因此,通过生物信息学手段对威兰胶合成过程中的四糖基本结构单元的聚合和运输步骤中涉及的 *welS* 和 *welG* 基因进行深入研究,对提高威兰胶的产量具有重要的意义.本研究基于 *welS* 和 *welG* 基因序列,利用在线软件分析它们的理化性质和结构特点,为后期深入研究 *welS* 和 *welG* 的功能并构建威兰胶发酵调控新策略奠定了基础.

1 材料与方法

welS 和 *welG* 基因序列均从课题组前期测定的鞘氨醇单胞菌 (*Sphingomonas* sp. WG) 全基因组序列 (NCBI ACCESSION NO: NZ_LNOS000000000)^[23] 中推导获得.应用 NovoPro (<http://www.novopro.cn/tools/translate.html>) 在线软件,将 *welS* 和 *welG* 基因序列翻译为氨基酸序列;ProtParam (<http://web.expasy.org/protparam/>) 预测演绎蛋白质 WELS 和 WELG 的理论等电点和相对分子质量和稳定系数;使用 ProtScale (<http://web.expasy.org/protscale/>) 预测演绎蛋白的亲疏水性;利用 ExPASy 在线服务器上的 SOPMA (<https://www.expasy.org/>) 预测演绎蛋白质的二级结构;SWISS-MODEL (<https://www.swissmodel.expasy.org/interactive>) 进行三级结构同源建模并预测蛋白质空间构象;在 NCBI Conserved Domain Search 平台 (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 预测演绎蛋白质的结构域;使用 SignalP 4.0 Server (<http://www.cbs.dtu.dk/services/SignalP-4.0>) 在线软件预测蛋白是否含有信号肽;通过 TMHMM (<http://www.cbs.dtu.dk/services/TMHMM/>) 预测演绎蛋白质的跨膜区.

2 结果分析

2.1 WELS 和 WELG 蛋白的一级结构预测分析

利用软件 NovoPro 将 *welS* 和 *welG* 基因序列翻译成氨基酸序列,并利用 ProtParam 得到相应的编码蛋白质 WELS 和 WELG 的一级结构 (NCBI ACCESSION NO: NZ_LNOS000000000). 分析结果表明,这两个蛋白分别含有 412 和 545 个氨基酸残基,分子量分别为 43.1 kDa 和 59.5 kDa. WELS 和 WELG 的氨基酸组成中疏水性氨基酸分别为 Ala, Cys, Ile, Leu, Met, Phe 和 Val, 与常温和中温的蛋白酶相比,含量较高,并且 WELS 的疏水性比 WELG 的疏水性强. WELS 和 WELG 的带电的氨基酸 (Asp, Arg, Lys 和 Glu) 组成比例高于常温和高温蛋白酶,这些疏水性氨基酸和带电氨基酸对于维持蛋白质较高的热稳定性是非常有利的. WELS 和 WELG 的等电点分别为 11.12 和 10.39, 表明这两个蛋白质均为碱性蛋白质;不稳定系数 (Instability index) 分别为 24.79 和 46.33;脂肪系数 (Aliphatic index) 分别为 129.34 和 126.65.

2.2 WELS 和 WELG 蛋白的二级结构预测分析

用 SOPMA 软件预测 WELS 和 WELG 蛋白的二级结构 (图 1), 结果发现,预测的 WELS 蛋白含约 63.11% 的 α -螺旋, 4.37% 的 β -转角, 14.32% 的延伸链和 18.20% 的不规则卷曲; WELG 蛋白含约 54.13% 的 α -螺旋, 3.67% 的 β -转角, 13.58% 的延伸链和 28.62% 的不规则卷曲.

2.3 WELS 和 WELG 蛋白结构域预测分析

利用 NCBI 中 Conserved Domains 对 WELS 和 WELG 蛋白序列进行结构域分析. 结果表明,对于 WELS, 多糖生物合成蛋白的膜是从第 1 位到 277 位的氨基酸, 氨基酸序列对比是从第 8 位到 274 位的氨基酸; 多糖生物合成 C-末端 d 的膜是从第 328 位到 411 位的氨基酸, 氨基酸序列对比是从第 328 位到 405 位的氨基酸. 对于 WELG 而言, O-抗原连接酶的膜是从第 318 位到 466 位氨基酸, 氨基酸序列对比是从第 324 位到 465 位的氨基酸. 因此, 两个蛋白酶在结构域组成上具有相似性, 这有可能是引起它们具有相似疏水性的主要原因.



图1 WELS (上) 和 WELG (下) 的二级结构预测

Fig.1 Secondary Structure Prediction of WELS (top) and WELG (bottom)

2.4 WELS 和 WELG 蛋白三级结构预测分析

利用在线软件 SWISS-MODEL, 根据同源建模法, 预测出 WELS 和 WELG 的三级结构(如图 2 所示), 分析结果表明, WELS 和 WELG 具有相似的立体三维结构, 但 WELS 的三级结构比 WELG 的三级结构更加紧凑, 左右两个结构域形成大致平行的角度, 并且在 β -sheet 部分具有相似的“环形”结构. 值得注意的是, 因为这两种酶相似的空间结构, 使

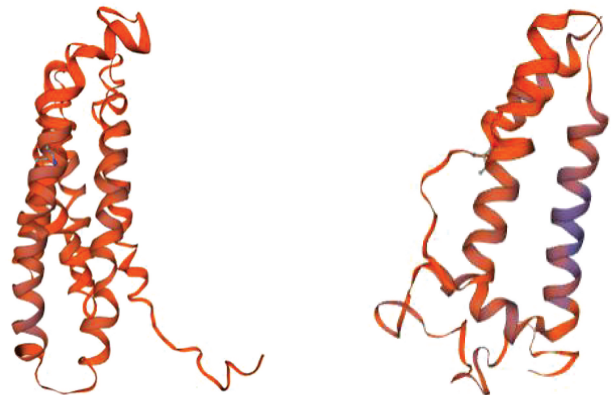


图2 WELS (左) 和 WELG (右) 的三级结构预测

Figure 2 Tertiary Structure Prediction of WELS (left) and WELG (right)

得构建具有高催化活性与功能的杂交酶成为可能.当然,这与目前报道的细菌脂多糖转运组装膜蛋白复合体有着巨大的差异,来自福氏志贺菌的 LptD-LptE 膜蛋白复合体是由两个蛋白所形成的“塞子-桶”结构模式,并含有两对非连续的、跨结构域的二硫键.LptD 的跨膜区是由最多 β 链所组成,尺寸最大的 β 桶状整合膜蛋白.在组成 β 桶的第 3 和第 26 条 β 链间创造了一个宽度为 16 埃的脂多糖出口,其大小可以允许脂多糖从桶壁上的出口侧向输入到脂双层中,成功完成脂多糖的转运^[24].结构更复杂的是,Lpt 系统在细胞内膜转运上涉及一个由 LptB,LptF 和 LptG 构成的 ABC 转运体复合物,其中 lptB 位于大肠杆菌全基因组的 72.0min 处,结构上呈完全折叠构型,具有典型的“L”构型,并包含高度保守的 Q-loop^[25].

2.5 WELS 和 WELG 蛋白信号肽预测

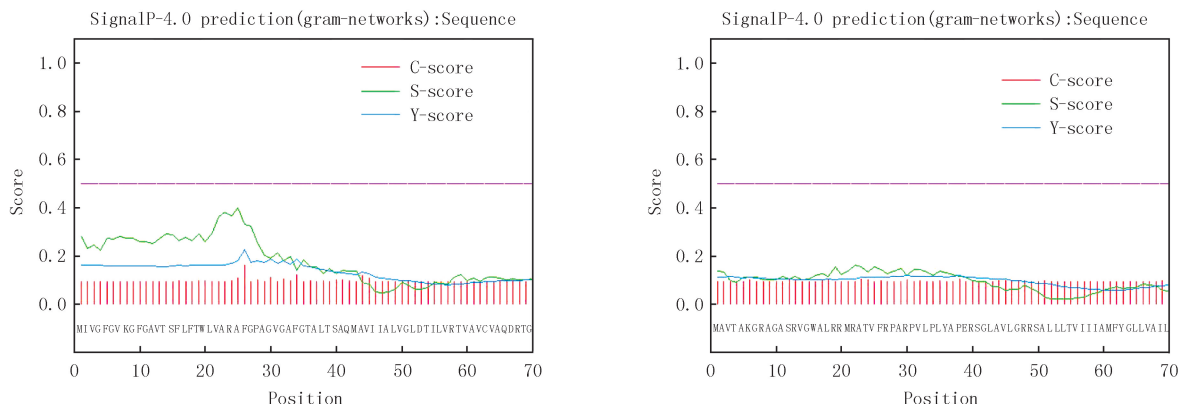


图 3 WELS (左) 和 WELG (右) 信号肽预测结果

Fig.3 Signal Peptide Prediction of WELS (left) and WELG (right)

利用 SignalP 4.0 Server 对 WelS 和 WelG 蛋白进行信号肽预测,结果如图 3 所示,可以看出两种蛋白的 C-score,S-score 和 Y-score 均未达到阈值(0.5),说明这两种蛋白均不含信号肽,为非分泌蛋白,合成后不需要运输,直接在合成处发挥作用.

2.6 WELS 和 WELG 蛋白的跨膜区分析

利用 TMHMM Server v. 2.0 在线软件分析蛋白质 WELS 和 WELG 的跨膜区,分析结果如图 4 所示.WELS 和 WELG 均具有较多的跨膜结构,因此它们均属于跨膜蛋白,这也印证了 WELS 和 WELG 在威兰胶分泌和聚合中发挥着重要功能.

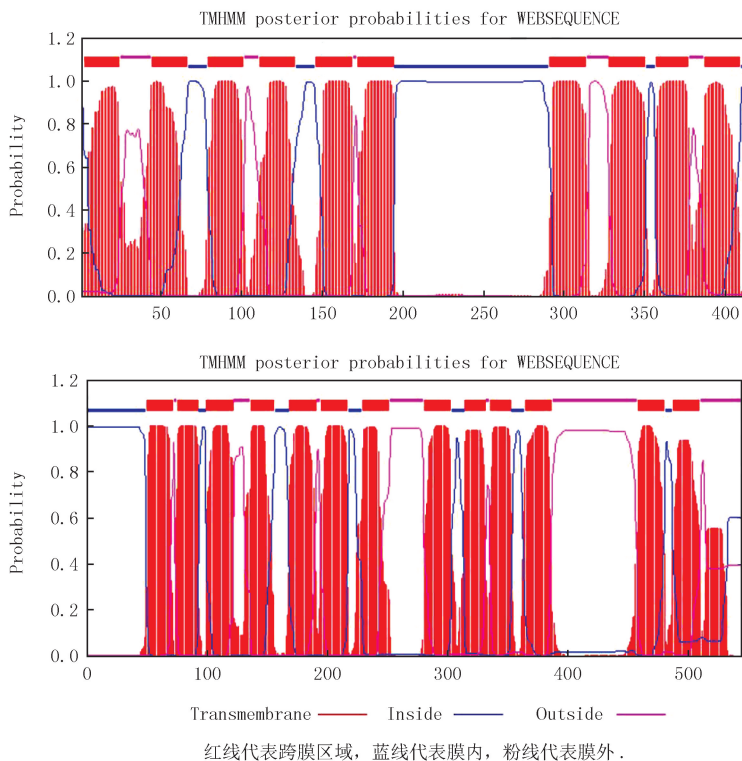


图 4 WELS (上) 和 WELG (下) 跨膜区分析

Fig.4 Transmembrane Region Analysis of WELS (top) and WELG (bottom)

3 结 语

本研究在前期完成威兰胶

合成基因簇测序工作的基础上,利用 PROTPARAM, SOPMA, SWISS-MODEL 等多种生物信息学软件,对威兰胶合成关键基因 *welS* 和 *welG* 进行了分析,分析结果表明他们编码的蛋白质 WELS 和 WELG 分别含有 412 和 545 个氨基酸,其中疏水性氨基酸分别为 Ala, Cys, Ile, Leu, Met, Phe 和 Val, 与常温和中温的蛋白酶相比,含量较高.WELS 和 WELG 分子量分别为 43.1 ku 和 59.5 ku, 等电点分别为 11.12 和 10.39, 均为碱性疏水性蛋白.在对蛋白 WELS 和 WELG 高级结构预测中发现两种蛋白均含有较高比例的 α -螺旋结构,较多 α -螺旋结构的存在有利于蛋白质结构的稳定性,并有利于提高蛋白质分子的热稳定性;同时 WELS 和 WELG 在结构域组成类型上具有相似性,通过对其进行同源建模分析,发现这两种蛋白具有相似的三维结构,左右结构域形成大致平行的角度,并且在 β -sheet 部分具有相似的“环形”结构;WELS 比 WELG 的三级结构更加的紧凑.通过软件预测,发现这两种蛋白均不存在信号肽,说明此类蛋白质为非分泌蛋白;两种蛋白均属跨膜蛋白,提示其可能在威兰胶分泌和聚合中发挥重要功能.这些结果将为未来提高威兰胶产量等研究提供重要的信息,奠定良好的基础.但是,该研究成果仅预测了可能参与核苷酸糖前体合成及四糖重复单元组装的部分蛋白的编码基因,对于威兰胶合成及代谢调控的具体机制仍不清楚.

参 考 文 献

- [1] 周万龙,焦学,孙亚杰,等.一株海洋来源鞘氨醇单胞菌胞外多糖的含量测定[J].中国海洋药物,2017,36(2):27-34.
ZHOU W L, JIAO X, SUN Y J, et al. Determination of extracellular polysaccharides produced by marine derived *Sphingomonas* sp. WG [J]. Chinese Journal of Marine Drugs, 2017, 36(2): 27-34.
- [2] LI H, LI J, ZHOU W L, et al. An efficient production of a novel carbohydrate polymer Sphingan WL [J]. Journal of Chemical Technology & Biotechnology, 2018, 93(12): 3472-3482.
- [3] WU M M, HUANG H D, LI G Q, et al. The evolutionary life cycle of the polysaccharide biosynthetic gene cluster based on the Sphingomonadaceae [J]. Scientific Reports, 2017, 7(46484): 1-14.
- [4] JANSSON P E, WIDMALM G. Welan gum (S-130) contains repeating units with randomly distributed L-mannosyl and L-rhamnosyl terminal groups as determined by FABMS [J]. Carbohydrate Research, 1994, 256(2): 327-330.
- [5] KAUR V, BERA M B, PANESAR P S, et al. Welan gum: microbial production, characterization, and applications [J]. International Journal of Biological Macromolecules, 2014, 65(5): 454-461.
- [6] TAKO M, KIRIYAKI M. Rheological properties of Welan Gum in aqueous media [J]. Journal of the Agricultural Chemical Society of Japan, 2006, 54(12): 3079-3084.
- [7] LI H, XU H, XU H, et al. Biosynthetic pathway of sugar nucleotides essential for welan gum production in *Alcaligenes* sp. CGMCC2428 [J]. Applied Microbiology and Biotechnology, 2010, 86(1): 295-303.
- [8] PLANK J. Applications of biopolymers and other biotechnological products in building materials [J]. Applied Microbiology and Biotechnology, 2004, 66(1): 1-9.
- [9] LI H, JIAO X, SUN Y J, et al. The preparation and characterization of a novel sphingan WL from marine *Sphingomonas* sp. WG [J]. Scientific Reports, 2016, 6(37899): 1-9.
- [10] 郭建军,李建科,陈芳,等.韦兰胶的特性、生产和应用研究进展[J].中国食品添加剂,2008(2):87-91.
GUO J J, LI J K, CHEN F, et al. The research progress of properties, production and application on Welan Gum [J]. China Food Additives, 2008(2): 87-91.
- [11] XU L, QIU Z, GONG H, et al. Rheological behaviors of microbial polysaccharides with different substituents in aqueous solutions: Effects of concentration, temperature, inorganic salt and surfactant [J]. Carbohydrate Polymers, 2019, 219: 162-171.
- [12] Martín-Piñero MJ, García MC, Muñoz J, Alfaro-Rodríguez MC. Influence of the welan gum biopolymer concentration on the rheological properties, droplet size distribution and physical stability of thyme oil/W emulsions [J]. International Journal of Biological Macromolecules, 2019, 133: 270-277.
- [13] XU X, NIE Z, ZHENG Z, et al. Production and rheological properties of Welan gum produced by *Sphingomonas* sp. ATCC 31555 with different nitrogen sources [J]. Journal of Molecular Microbiology and Biotechnology, 2017, 27(1): 55-63.
- [14] HAN D K, YANG C Z, ZHANG Z Q, et al. Recent development of enhanced oil recovery in China [J]. Journal of Petroleum Science and Engineering, 1999, 22(1/3): 181-188.
- [15] 张宏方,王德民,岳湘安,等.不同类型聚合物溶液对采油残余油的作用机理研究[J].高分子学报,2003(3):321-325.
ZHANG H F, WANG D M, YUE X A, et al. Various types of polymer solution's mechanism of action on residual oil in porous media [J]. Acta Polymerica Sinica, 2003(3): 321-325.
- [16] LIU X, LIN L, XU X, et al. Two-step economical welan gum production by *Sphingomonas* sp. HT-1 from sugar industrial by-products [J]. Carbohydrate Polymers, 2018, 181: 412-418.

- [17] ZHU H, SUN S W, LI H, et al. Significantly improved production of Welan gum by *Sphingomonas* sp. WG through a novel quorum-sensing-interfering dipeptide cyclo(L-Pro-L-Phe)[J]. International Journal of Biological Macromolecules, 2019, 126: 118-122.
- [18] LIU X L, LIN L J, XU X Y, et al. Two-step economical welan gum production by *Sphingomonas* sp. HT-1 from sugar industrial by-products[J]. Carbohydrate Polymers, 2018, 181: 412-418.
- [19] ROSALAM S, ENGLAND R. Review of xanthan gum production from unmodified starches by *Xanthomonas com prestris* sp.[J]. Enzyme & Microbial Technology, 2006, 39(2): 197-207.
- [20] LI H, LI J, XUE J, et al. Characterization of the biosynthetic pathway of nucleotide sugar precursor UDP-glucose during sphingane WL gum production in *Sphingomonas* sp. WG[J]. Journal of Biotechnology, 2019, 302(1): 1-9.
- [21] WANG X Y, TAO F, GAI Z H, et al. Genome sequence of the welan gum-producing strain *Sphingomonas* sp. ATCC 31555[J]. Journal of Bacteriology, 2012, 194(21): 5989-5990.
- [22] SCHMID J, SPERL N, SIEBER V. A comparison of genes involved in sphingane biosynthesis brought up to date[J]. Applied Microbiology and Biotechnology, 2014, 98(18): 7719-7733.
- [23] LI H, FENG Z M, SUN Y J, et al. Draft genome sequence of *Sphingomonas* sp. WG, a welan gum-producing strain[J]. Genome Announcements, 2016, 4(1): e01709-15.
- [24] RUSSELL E B. Structural biology: Lipopolysaccharide rolls out the barrel[J]. Nature, 2014, 511: 37-38.
- [25] 莫婷, 刘马峰, 程安春. 部分革兰氏阴性菌脂多糖运输蛋白 LptD 的结构及功能研究进展[J]. 微生物学通报, 2017, 44(1): 217-224.
- MO T, LIU M F, CHENG A C. Structure and function of LptD in some Gram-negative bacteria[J]. Microbiology China, 2017, 44(1): 217-224.

Bioinformatics study on *welS* and *welG* key Genes involved in Welan Gum Synthesis in *Sphingomonas* sp. WG

Li Li¹, Chang Aiping¹, Qian Jin², Yang Yanyu¹, He Qiaomei¹, Shen Yaling², Wang Yinglu¹, Zhu Hu¹

(1. Fujian Provincial University Engineering Research Center of Industrial Biocatalysis, Fujian Provincial Key Laboratory of Advanced Materials Oriented Chemical Engineering, College of Chemistry and Materials Science, Fujian Normal University, Fuzhou 350007, China; 2. State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: Welan gum is a kind of natural polysaccharide produced by microbial metabolism. Due to its excellent rheological properties and stability, it has been widely used in petroleum exploitation, building materials, daily chemicals, food and medicine industry, etc. This study focusses on the bioinformatics analysis for the key genes *welS* and *welG* involved in the biosynthesis of welan gum, mainly including gene sequences and properties and structures of their coding proteins WELS and WELG. The results showed that the encoding proteins WELS and WELG contain 412 and 545 amino acids with molecular weight 43.1 kDa and 59.5 kDa, respectively. The isoelectric point of WELS is 11.12, and of WELG is 10.39, combining with the hydrophilicity studies, suggesting that both of these proteins are basic hydrophobic proteins. The domain composition of WELS and WELG are similar, which form a proximal parallel angle between the left and right domains, and construct a similar "annular" structure in the beta-sheet moiety. These analyzed results will provide a new regulation strategy for biosynthesis of welan gum, and theoretical support for improving its fermentation yield.

Keywords: Welan gum; *welS*; *welG*; structure characteristic; bioinformatics

[责任编辑 王凤产 杨浦]