

生物技术通报

· 研究报告 ·

BIOTECHNOLOGY BULLETIN

2012年第11期

东方肉座菌 EU7-22 纤维素酶基因的克隆及序列分析

龙传南¹ 成奕瑾¹ 邬小兵² 刘健¹ 龙敏南^{1,2}(¹厦门大学能源研究院, 厦门 361005 ; ²厦门大学生命科学学院, 厦门 361005)

摘要: 东方肉座菌 EU7-22 与 XC-9、里氏木霉、康宁木霉、黑曲霉、斜卧青霉进行产纤维素酶比较, 结果表明菌株 EU7-22 具有较高的产纤维素酶能力及完整的纤维素酶系。根据里氏木霉和绿色木霉的外切葡聚糖酶, 内切葡聚糖酶及 β -葡萄糖苷酶相关基因序列, 设计引物 PCR 扩增出菌株 EU7-22 *cbh*、*cbh*、*eg*、*eg* 及 *bgl*。基因序列经 NCBI Blast 分析表明, *cbh* 与绿色木霉 *cbh1* 基因 (FJ871063) 同源性最高达 99%; *cbh* 与康宁木霉 *cbh2* 基因 (DQ504304) 同源性最高达 99%; *eg* 与长枝木霉 *eg1* 基因 (GU144298) 同源性最高达 99%; *eg* 与绿色木霉 *eg2* 基因 (EF602036) 同源性最高达 99%; *bgl* 与菌株 *Trichoderma* sp. SSL *bgl* 基因 (FJ040193) 同源性最高达 100%。5 种纤维素酶基因编码的相应氨基酸序列与其他木霉纤维素酶的氨基酸序列相似性也非常高。对上述纤维素酶基因编码的相应蛋白的分子量、等电点、N-糖基化位点、信号肽序列进行分析; 对纤维素结合区及糖基水解酶家族特征结构区进行了定位; 用 SWISS-Model 模拟了酶蛋白的三级结构。

关键词: 东方肉座菌 纤维素酶 基因 同源性

Cloning and Sequence Analysis of Cellulase Genes from
Hypocrea orientalis EU7-22Long Chuannan¹ Cheng Yijin¹ Wu Xiaobing² Liu Jian¹ Long Minnan^{1,2}(¹School of Energy Research, Xiamen University, Xiamen 361005 ; ²School of Life Sciences, Xiamen University, Xiamen 361005)

Abstract: The fungi of *Hypocrea orientalis* EU7-22, *Hypocrea orientalis* XC-9, *Trichoderma reesei*, *Trichoderma koningii*, *Aspergillus niger* and *Penicillium decumbens* were investigated to produce cellulase. The results indicated that the strain EU7-22 had highly cellulase activities in comparison with other fungi, and possessed of integrated cellulase system. According to the reported cellobiohydrolase, endoglucanase and β -glucosidase gene sequences of *T. reesei* and *Trichoderma viride*, the primers were designed and five cellulase genes (*cbh*, *cbh*, *eg*, *eg*, *bgl*) were successfully cloned by PCR. By conducting sequence alignment analysis with NCBI Blast, it was found that the homology of same cellulase genes between strain EU7-22 and other *Trichoderma* were: 99% to *cbh1* (FJ871063) from *T. viride*, 99% to *cbh2* (DQ504304) from *T. koningii*; 99% to *eg1* (GU144298) from *Trichoderma longibrachiatum*; 99% to *eg2* (EF602036) from *T. viride*; 100% to *bgl* (FJ040193) from *Trichoderma* sp. SSL. Furthermore, the corresponding amino acid sequences were also quite similar. The molecular weight, isoelectric point, N-glycosylation sites and signal peptide sequence of these cellulase genes encoding the corresponding protein were analyzed. The cellulose-binding domain and conserved domains of glycosyl hydrolases family were confirmed. By using SWISS-Model, the tertiary structure of cellulase proteins were predicted and simulated.

Key words: *Hypocrea orientalis* Cellulase Gene Homology

木质纤维素生物质是地球上来源最丰富的生物质资源^[1,2], 若能将其有效降解成糖类(如葡萄糖、木糖、低聚木糖、阿拉伯糖等)物质, 以这些糖类

物质为平台进一步转化加工成人类所必需的生物基燃料、生物基化学品和生物基材料^[3,4], 对解决当前世界面临的能源危机、环境污染和粮食短缺等问

收稿日期: 2012-04-10

基金项目: 国家重点基础研究发展计划(“973”计划)(2010CB732201), 中央高校基本科研业务费专项资金项目(201112G026), 国家自然科学基金项目(31170067)

作者简介: 龙传南, 男, 博士研究生, 研究方向: 生物能源; E-mail: longcn2004@126.com

通讯作者: 龙敏南, 男, 教授, 研究方向: 生物质资源利用; E-mail: longmn@xmu.edu.cn

题具有重大的战略意义。

纤维素酶在天然纤维素生物质分解过程中发挥关键作用。它是一个复杂的多酶体系,由许多具有高效协同作用的水解酶组成,任何单一的纤维素酶组分都难以高效地水解纤维素。根据纤维素酶催化功能,可分为外切 β -1,4-葡聚糖酶(exo-1,4- β -D-glucanase, EC 3.2.1.91, 简称 CBH),内切 β -1,4-葡聚糖酶(endo-1,4- β -D-glucanase, EC 3.2.1.4, 简称 EG)及 β -1,4-葡萄糖苷酶(β -1,4-glucosidase, EC 3.2.1.21, 简称 BGL)等3种组分^[5]。

东方肉座菌(*Hypocrea orientalis*)属于木霉属,菌株 EU7-22(原命名为灰绿曲霉 EU7-22),在菌株 XC-9^[6]基础上诱变育种获得。本研究拟对菌株 EU7-22, XC-9 与里氏木霉(*Trichoderma reesei*; *Hypocrea jecorin*),康宁木霉(*Trichoderma koningii*),黑曲霉(*Aspergillus niger*),斜卧青霉(*Penicillium decumbens*)产纤维素酶能力进行比较;并且首次从菌株 EU7-22 中克隆外切葡聚糖酶基因,内切葡聚糖酶基因及 β -葡萄糖苷酶基因。采用生物信息学方法分析酶蛋白的性质,以期为后续开展纤维素酶的基因工程与蛋白质工程研究奠定基础。

1 材料与方法

1.1 材料

1.1.1 主要试剂 PCR 扩增试剂, DNA Marker 和 pMD19-T 载体均购于宝生物工程(大连)有限公司;胶回收试剂盒购于天根生化科技(北京)有限公司;对硝基酚- β -纤维二糖(pNPC),对硝基酚- β -葡萄糖苷(pNPG),对硝基酚(pNP)购于Sigma公司。其他试剂购于上海生工生物工程技术有限公司和国药集团化学试剂有限公司。

1.1.2 菌株及培养基 菌株:东方肉座菌(*H. orientalis*) EU7-22,东方肉座菌 XC-9,大肠杆菌(*Escherichia coli*) DH5 α 为本实验室保存。里氏木霉(*T. reesei* H. jecorin, 编号 40358),购于中国工业微生物菌种保藏管理中心 斜卧青霉(*P. decumbens*, 编号 3.686),黑曲霉(*A. niger*, 编号 3.316),康宁木霉(*T. koningii*, 编号 3.2774)购于中国微生物菌种保藏管理委员会普通微生物中心。

PDA培养基:马铃薯200g洗净,去皮,切成小块,

煮沸30min,6层纱布过滤,定容至1L,加葡萄糖20g,自然pH;固体培养基按2%浓度加入琼脂粉。

LB培养基:胰蛋白胨10g,酵母提取物5g,氯化钠10g,用去离子水定容至1L;固体培养基按2%浓度加入琼脂粉。

产纤维素酶培养基:2%(W/V)甘蔗渣(粉碎过40目筛网),1%(W/V)麸皮,0.5%蛋白胨,2.5g/L KH_2PO_4 , Mandels 营养液, pH5.2。

1.2 方法

1.2.1 粗酶液制备 将浓度约为 10^6 孢子/mL的不同丝状真菌菌液接入产纤维素酶培养基中,30,180 r/min 振荡培养4d后,将发酵液常温离心(6000 r/min, 10 min),上清液即为粗酶液。

1.2.2 纤维素酶活测定

1.2.2.1 滤纸酶活(FPA activity)测定^[7] 将Whatman NO.1 滤纸(50 mg, 1 cm \times 6 cm)卷成小卷,放进25 mL 具塞刻度试管中,加入1.5 mL 柠檬酸钠缓冲液(50 mmol/mL, pH5.0),再加入用相同缓冲液适当稀释(使吸光度值在最适量程范围内,即0.2-1.0)的0.5 mL 酶液,轻轻摇匀,使滤纸完全浸泡在液体中,50 保温1h后,按DNS法测定并换算出还原糖含量。以100 水浴灭活5min的粗酶液为空白对照(下同)。以每分钟催化底物水解生成1 μmol 还原糖定义为一个国际酶活单位(IU)。

在此酶活测定中,丝状真菌东方肉座菌 EU7-22、里氏木霉、东方肉座菌 XC-9 和康宁木霉所产酶液,分别稀释5倍、3倍、3倍、2倍;黑曲霉,斜卧青霉所产酶液不稀释。

1.2.2.2 内切葡聚糖酶活(CMC activity)测定^[7] 移取稀释酶液(东方肉座菌 EU7-22、里氏木霉、东方肉座菌 XC-9 和康宁木霉所产酶液,分别稀释10倍、6倍、6倍和4倍;黑曲霉,斜卧青霉所产酶液不稀释)0.5 mL 于25 mL 具塞刻度试管中,加入2% CMC-Na 的柠檬酸-柠檬酸钠缓冲液(pH5.0) 1.5 mL,然后50 水浴30min,按DNS法测定并换算出还原糖含量。以每分钟催化底物水解生成1 μmol 还原糖定义为一个国际酶活单位(IU)。

1.2.2.3 外切葡聚糖酶(CBH activity)与 β -葡萄糖苷酶(BGL activity)酶活测定^[8] 移取稀释酶液(东

方肉座菌 EU7-22、里氏木霉、东方肉座菌 XC-9、康宁木霉和黑曲霉所产酶液,分别稀释 10 倍、6 倍、3 倍、2 倍和 2 倍;斜卧青霉所产酶液不稀释) 50 μ L 于离心管中,加入 50 μ L 用柠檬酸钠缓冲液(50 mmol/mL, pH5.0)配制的 pNPC (5 mmol/L)(用于 CBH 酶活测定)和 pNPG (5 mmol/L)(用于 BGL 酶活测定),50 $^{\circ}$ C 水浴 30 min,再加入 1 mL Na_2CO_3 (0.5 mol/L) 终止反应。在 405 nm 波长下测定吸光值。一个酶活力单位 (IU) 定义为每分钟催化底物水解生产 1 μ mol pNP 所需的酶量。

不同丝状真菌产纤维素酶能力比较,根据 3 次重复产酶试验,所测定的酶活为依据。

1.2.3 菌株总 DNA 提取 东方肉座菌 EU7-22 经 PDA 液体培养基培养 36 h 后,按照 SDS 裂解法提取总 DNA^[9]。

1.2.4 引物设计与合成 根据 GenBank 中里氏木霉及绿色木霉外切葡聚糖酶 基因 (cbh :GenBank 登录号 GL985084 和 AY368686.1);外切葡聚糖酶 基因 (cbh :GenBank 登录号 GU724763.1 和 AY-368688.1);内切葡聚糖酶 基因 (eg :GenBank 登录号 AY928809.1 和 AY343986.1);内切葡聚糖酶 基因 (eg :GenBank 登录号 DQ178347.1 和 AY368-688.1); β -葡萄糖苷酶 基因 (bgl :GenBank 登录号 U09580.1 和 FJ882071.1) 的序列,分别设计 5 对引物(表 1),引物由上海生工生物工程有限公司合成。

表 1 东方肉座菌纤维素酶基因 PCR 扩增引物

引物名称	引物序列 (5'-3')	扩增基因
CBH -F	ATGTATCGGAAATTGGCCCTCATCTC	cbh
CBH -R	TTACAGGCACTGAGAGTAGAATTCGTTC	cbh
CBH -F	ATGATTGTCGGCATTCTCACC	cbh
CBH -R	TTACAGGAACGATGGGTTTGC	cbh
EG -F	ATGGCGCCCTCAGTTACA	eg
EG -R	CTAAAGGCATTGCGAGTAG	eg
EG -F	ATGAACAAGTCCGTGGCTCCAT	eg
EG -R	CTACTTCTTTCGCGAGACACG	eg
BGL -F	ATGCGTTACCGAACAGCAG	bgl
BGL -R	CTACGCTACCGACAGAGTGCT	bgl

1.2.5 纤维素酶基因的克隆 以提取的东方肉座菌 EU7-22 总 DNA 为模板,采用 TaKaRa 公司的高保真

Taq 酶,利用引物对 CBH -F&R、CBH -F&R、EG -F&R、EG -F&R 和 BGL -F&R 进行 PCR 扩增,分别获得 cbh、cbh、eg、eg 及 bgl 基因片段。扩增 cbh 基因的反应程序为:94 $^{\circ}$ C 预变性 5 min;94 $^{\circ}$ C 变性 30 s,57 $^{\circ}$ C 退火 45 s,72 $^{\circ}$ C 延伸 2 min,30 个循环;72 $^{\circ}$ C 充分延伸 10 min。扩增 cbh、eg 和 eg 基因的反应程序为:94 $^{\circ}$ C 预变性 5 min;94 $^{\circ}$ C 变性 30 s,55 $^{\circ}$ C 退火 45 s,72 $^{\circ}$ C 延伸 2 min,30 个循环;72 $^{\circ}$ C 充分延伸 10 min。扩增 bgl 基因的反应程序为:94 $^{\circ}$ C 预变性 5 min;94 $^{\circ}$ C 变性 30 s,54 $^{\circ}$ C 退火 45 s,72 $^{\circ}$ C 延伸 2.5 min,30 个循环;72 $^{\circ}$ C 充分延伸 10 min。

PCR 扩增产物经胶回收试剂盒纯化后,连接到 pMD19-T 载体上,转化 E. coli DH5 α 感受态细胞中,先进行蓝白斑筛选,然后用 M13 通用引物对阳性克隆子进行筛选和测序验证(北京六合华大基因科技股份有限公司)。

1.2.6 纤维素酶基因序列的生物信息学分析 将克隆基因序列在 GenBank 数据库中进行同源性比对 (<http://www.ncbi.nlm.nih.gov/BLAST>);根据 BLAST 比对的结果,分析基因内含子位置并去掉内含子,利用 Primer Premier 5.0 进行核酸及蛋白质序列分析;用 ProtParam 工具^[10]分析基因编码蛋白质的氨基酸序列组成、相对分子质量、等电点等理化性质 (<http://web.expasy.org/protparam>);利用 SignalP-4.0 软件^[11]分析 N-末端信号肽序列 (<http://www.cbs.dtu.dk/services/SignalP/>);利用 NetNGlyc 软件^[12]进行蛋白质 N-糖基化位点预测 (<http://www.cbs.dtu.dk/services/NetNGlyc>);利用 Prosite motif search^[13]对蛋白质的功能位点进行搜索 (<http://prosite.expasy.org/>);利用 SWISS-Model^[14]预测和模拟蛋白质的三级结构 (<http://swissmodel.expasy.org/workspace/index.php?>)。

2 结果

2.1 丝状真菌产纤维素酶能力比较

丝状真菌东方肉座菌 (H. orientalis) EU7-22,东方肉座菌 (H. orientalis) XC-9,里氏木霉 (H. jecorin),康宁木霉 (T. koningii),黑曲霉 (A. niger),斜卧青霉 (P. decumbens) 液态发酵产纤维素酶能力比较,结果如图 1 所示。产外切葡聚糖酶 (CBH)

的能力, 东方肉座菌 EU7-22 > 里氏木霉 > 东方肉座菌 XC-9 > 黑曲霉 > 康宁木霉 > 斜卧青霉; 产内切葡聚糖酶 (EG) 的能力, 东方肉座菌 EU7-22 > 里氏木霉 > 东方肉座菌 XC-9 > 康宁木霉 > 黑曲霉 > 斜卧青霉; 产 β -葡萄糖苷酶 (BGL) 的能力, 东方肉座菌 EU7-22 > 里氏木霉 > 东方肉座菌 XC-9 > 康宁木霉 > 黑曲霉 > 斜卧青霉; 产滤纸酶活 (FPA) 的能力, 里氏木霉略大于东方肉座菌 EU7-22 > 东方肉座菌 XC-9 > 康宁木霉 > 黑曲霉 > 斜卧青霉。结果表明, 出东方肉座菌 EU7-22 不但产酶能力强, 而且纤维素酶系齐全, 是一株比较有潜力的菌株, 被应用于生物能源领域。在后续试验中, 该菌株的主要纤维素酶基因也成功被克隆。

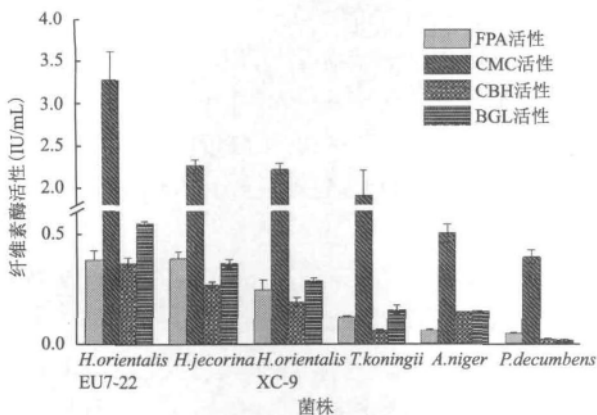
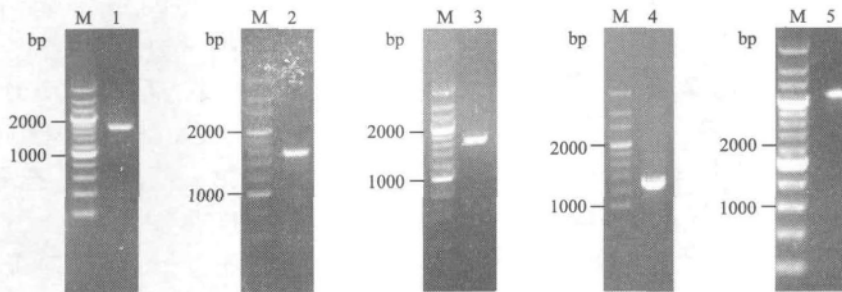


图 1 不同丝状真菌产纤维素酶能力比较



M. 200 bp DNA Marker; 1. *cbh* I; 2. *cbh* II; 3. *eg* I; 4. *eg* II; 5. *bgl* I

图 2 纤维素酶基因 PCR 扩增片段的琼脂糖凝胶电泳

外切葡聚糖酶 基因全长 1 583 bp (GenBank 登录号 JQ238605) , 与康宁木霉 *cbh2* 基因(DQ504304) 同源性最高达到 99%。根据康宁木霉, 里氏木霉及绿色木霉 *cbh2* 基因内含子位置, 分析菌株 EU7-

2.2 PCR 扩增纤维素酶基因片段

以东方肉座菌 EU7-22 总 DNA 为模板, 利用引物对 CBH -F&R、CBH -F&R、EG -F&R、EG -F&R 和 BGL -F&R 分别进行 PCR 扩增, 获得外切葡聚糖酶 基因 (*cbh* , 约 1 700 bp) 片段, 外切葡聚糖酶 基因 (*cbh* , 约 1 600 bp) 内切葡聚糖酶 基因 (*eg* , 约 1 600 bp) 片段, 内切葡聚糖酶 基因 (*eg* , 约 1 300 bp) 片段和 β -葡萄糖苷酶 基因 (*bgl* , 约 2 400 bp) 片段。PCR 产物进行琼脂糖凝胶 (1.0%) 电泳, 结果 (图 2) 显示, PCR 反应均扩增出单一的 DNA 条带, 基因片段的大小与预期一致。

2.3 纤维素酶基因序列 Blast 分析

序列测定结果经 NCBI Blast 分析 : 东方肉座菌 EU7-22 外切葡聚糖酶 基因全长 1 680 bp (GenBank 登录号 JQ238604) 与绿色木霉 *cbh1* 基因(FJ871063) 同源性最高达到 99%。根据绿色木霉、里氏木霉及康宁木霉 *cbh1* 基因内含子位置, 分析菌株 EU7-22 *cbh* 基因含有 2 个内含子, 位于 462-527 bp 与 1 225-1 293 bp 处。利用 Primer Premier 5.0 将 *cbh* 基因 (去掉内含子) 翻译成相应氨基酸序列 (编码 514 个氨基酸) 并进行 Protein blast, 结果显示, *cbh* 基因编码的蛋白质属于糖基水解酶家族 7A, 与绿色木霉外切葡聚糖酶 (ACZ06578) 氨基酸序列同源性最高达到 100%。

22 *cbh* 基因含有 3 个内含子位于 93-143 bp、528-583 bp 及 832-894 bp 处。利用 Primer Premier 5.0 将 *cbh* 基因(去掉内含子) 翻译成相应氨基酸序列 (编码 470 个氨基酸) 并进行 Protein blast, 结果显示

cbh 基因编码的蛋白质属于糖基水解酶家族 6, 与康宁木霉外切葡聚糖酶 (ABF56208) 氨基酸序列同源性最高达到 100%。

内切葡聚糖酶 基因全长 1 566 bp (GenBank 登录号 JQ238606), 与长枝木霉 eg1 基因 (GU144298) 同源性最高达 99%。根据长枝木霉、里氏木霉、绿色木霉 eg1 基因内含子位置, 分析菌株 EU7-22 eg 基因含有 2 个内含子位于 771-893 bp 和 1 490-1 546 bp 处。利用 Primer Premier 5.0 将 eg 基因(去掉内含子)翻译成相应氨基酸序列(编码 461 个氨基酸)并进行 Protein blast, 结果显示, eg 基因编码的蛋白质属于糖基水解酶家族 7B, 与长枝木霉内切葡聚糖酶 (ACZ34302) 氨基酸序列同源性最高达 100%。

内切葡聚糖酶 基因全长 1 257 bp (GenBank 登录号 JQ238607), 与绿色木霉 eg2 基因 (EF602036) 同源性最高达 99%。根据绿色木霉, 里氏木霉 eg2 基因内含子位置, 分析菌株 EU7-22 eg 基因不含内含子。利用 Primer Premier 5.0 将 eg 基因翻译成相应氨基酸序列(编码 418 个氨基酸)并进行 Protein blast, 结果显示, eg 基因编码的蛋白质属于糖基水解酶家族 5, 与绿色木霉内切木聚糖酶 (ADJ10267) 氨基酸序列同源性最高达 100%。

β -葡萄糖苷酶 基因全长 2 369 bp (GenBank 登录号 JQ904600), 与菌株 *Trichoderma sp.* SSL bgl 基因 (FJ040193) 同源性最高达 100%。根据菌株 *Trichoderma sp.* SSL, 绿色木霉, 里氏木霉 bgl 基因内含子位置, 分析菌株 EU7-22 bgl 基因含有 2 个内含子位于 66-135 bp 和 1 896-1 959 bp 处。利用 Primer Premier 5.0 将 bgl 基因(去掉内含子)翻译成相应氨基酸序列(编码 744 个氨基酸)并进行 Protein blast, 结果显示, bgl 基因编码的蛋白质属于糖基水解酶家族 3, 与菌株 *Trichoderma sp.* SSL 葡萄糖苷酶 (ACH92574) 氨基酸序列同源性最高达 100%。

2.4 纤维素酶基因编码蛋白质 ProtParam、NetNGlyc 及 SignalP 分析

ProtParam 分析表明, 东方肉座菌 EU7-22 CBH, CBH, EG, EG 及 BGL 酶蛋白相对分子质量

分别为 54.23、49.61、48.16、44.22 和 78.21 kD; 等电点分别为 4.58、4.91、4.72、4.96 和 5.96。

NetNGly 分析显示, 东方肉座菌 EU7-22 CBH 酶蛋白含有 4 个 N-糖基化位点 (N⁶²SST、N⁸¹ETC、N²⁸⁷TSF、N⁴⁰¹ETS); CBH 酶蛋白含有 3 个 N-糖基化位点 (N³⁸WSG、N³¹²ASS、N³³³ITS); EG 酶蛋白含有 5 个 N-糖基化位点 (N⁷⁸TTL、N¹⁶⁴GSL、N²⁰⁴GTL、N²⁰⁸TSG、N³⁶⁶PSN); EG 酶蛋白含有 1 个 N-糖基化位点 (N¹²⁴FTG); BGL 酶蛋白含有 1 个 N-糖基化位点 (N²³⁹TTW)。

SignalP 分析表明, 东方肉座菌 EU7-22 CBH 酶蛋白 N-末端前 1-17 个氨基酸 (MYRKLALISAFALAR) 为信号肽序列; CBH 酶蛋白 N-末端前 1-18 个氨基酸 (MIVGILTTLATLATAAS) 为信号肽序列; EG 酶蛋白 N-末端前 1-22 个氨基酸 (MAPSVTLPLTTAMLALVRLVAA) 为信号肽序列; EG 酶蛋白 N-末端前 1-21 个氨基酸 (MNKSVAPLLLAASLTVGGVAA) 为信号肽序列; BGL 酶蛋白 N-末端前 1-19 个氨基酸 (MRYRTAAALALAAAPFVTA) 为信号肽序列。

2.5 纤维素酶蛋白 Prosite motif search 及 SWISS-Model 三维结构模拟

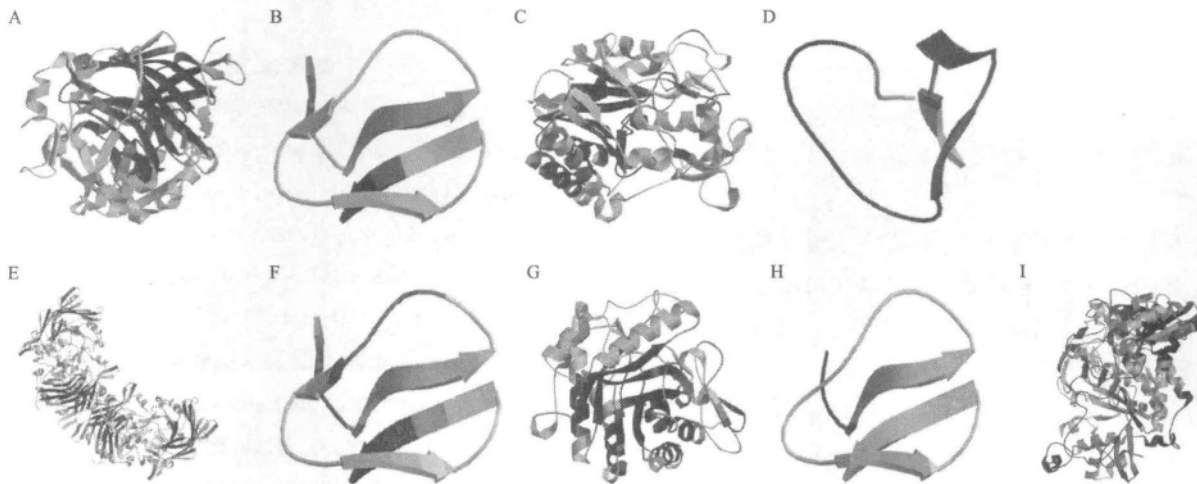
纤维素酶分子的结构大多具有独立活性的两个结构域, 其中一部分是只能水解可溶性纤维素, 但不能水解不溶性纤维素, 并且具有催化功能, 被称为催化结构域 (core or catalytic domain, CD); 另一部分能吸附于纤维素表面, 但对纤维素降解作用不大, 被称为结合结构域 (cellulose-binding domain, CBD), CBD 在纤维素酶中位于氨基端或羧基端; CD 与 CBD 之间通过一段高度糖基化的连接桥 (linker) 相连。

利用 Prosite motif search 对蛋白质的功能位点进行分析, CBH 酶蛋白含有 1 个纤维素结合域 (CBD): C⁴⁸⁶GGIGYSGPTVCASGTTTCQVLNEFYSC⁵¹³。CBH 酶蛋白含有 2 个糖基水解酶家族 6 的特征序列 (V¹⁹⁰VYDLPDRDCAALASNG²⁰⁶, L²³⁸LVIEPDSL²⁴⁷) 及 1 个纤维素结合域 (C³⁴GGQNWSPGCCASGSTCVYSNDYYSQC⁶¹)。EG 酶蛋白含有 1 个纤维素结合域 (CBD): C⁴³³GGIGYTGCKTCTSGTTCQYGN-

YYSQC⁴⁶⁰。EG 酶蛋白含有 1 个糖基水解酶家族 6 的特征序列 (V²³²WFGIMNEPH²⁴¹) 及 1 个纤维素结合域 (C²⁹GGIGWSGPTNCAPGSACSTLNSYYAQC⁵⁶)。

BGL 酶蛋白不含有任何特征序列。

利用 SWISS-Model 预测和模拟纤维素酶蛋白催化结构域与结合结构域的三级结构, 如图 3 所示。



A. CBH I 酶蛋白 CD 域; B. CBH I 酶蛋白 CBD 域; C. CBH II 酶蛋白 CD 域; D. CBH II 酶蛋白 CBD 域; E. EG I 酶蛋白 CD 域; F. EG I 酶蛋白 CBD 域; G. EG II 酶蛋白 CD 域; H. EG II 酶蛋白 CBD 域; I. BGL I 酶蛋白 CD 域

图 3 SWISS-Model 模拟纤维素酶 CD 与 CBD 域三级结构

3 讨论

天然纤维素生物质是地球上最丰富却未被有效利用的可再生资源, 高效地将纤维素生物质全组分转化为生物燃料对解决人类目前面临的能源短缺具有重要的战略意义。根据纤维素酶活测定结果, 东方肉座菌 EU7-22, 相比其他丝状真菌 (里氏木霉, 康宁木霉、黑曲霉和斜卧青霉), 表现较强的产纤维素酶能力, 并且拥有完整的纤维素酶酶系。

Goedegebuur 等^[15]总结, 在 *T. reesei* 中至少产生 7 种纤维素酶 (CBH I、CBH II、EG I、EG II、EG III、EG IV 和 EG V), 这些 CBHs 与 EGs 在菌株分泌的纤维素酶中所占含量分别为 >50%、20%、15%、10%、1%、<1% 和 <1%。而 CBH I、CBH II、EG I 和 EG II 4 种酶的含量总和超过 95%, 可见在纤维素降解过程中, 发挥巨大功能。EGs 与 CBHs 先分解纤维素产生纤维二糖, 进一步被 β -葡萄糖苷酶分解产生葡萄糖。因此, 东方肉座菌 EU7-22 的主要纤维素酶基因 (*cbh I*、*cbh II*、*eg I*、*eg II* 和 *bgl I*) 被克隆。这些基因序列及其内含子的位置、数量、长短以及编码氨基酸序列的保守性, 可为纤维素酶家族亚群的分类提供依据^[15]。

对纤维素酶蛋白进行糖基化及信号肽分析, 可为异源表达研究酶蛋白提供指导意义。如在酵母中表达研究酶基因, 需要去除信号肽序列^[16]。因此, 必须先了解清楚该酶蛋白有没有信号肽序列, 具体有多少个氨基酸为信号肽序列。表达的酶蛋白有没有活性, 活性的高低如何, 这涉及到酶蛋白糖基化问题^[17]。

SWISS-Model 预测和模拟蛋白质的三级结构。CBH 酶蛋白催化结构域的三级结构, 如图 3-A, 参考模板为 *H. jecorina* Cel7A (PDB:2v3i), 且该区域两者氨基酸序列相似性达 97.01%; CBD 结构域位于该蛋白质的羧基端, 其三级结构如图 3-B, 参考模板为 *H. jecorina* CBH I 的 C-末端区域 (PDB:1cbh)^[18], 且该区域两者氨基酸序列相似性达 91.67%。

CBH 酶蛋白催化结构域的三级结构如图 3-C, 参考模板为 *H. jecorina* Cel6A (PDB:1qk2)^[19], 且该区域两者氨基酸序列相似性达 98.90%; CBD 结构域位于该蛋白质的氨基端, 其三级结构如图 3-D, 参考模板为 *H. jecorina* CBH I 纤维素结合域 (PDB:1az6)^[20], 且该区域两者氨基酸序列相似性达 70.97%。

EG 酶蛋白催化结构域的三级结构如图 3-E, 参考模板为 *H. jecorina* Cel7B (PDB: 1eg1)^[21], 且该区域两者氨基酸序列相似性达 95.15%; CBD 结构域位于该蛋白质的 C 端, 其三级结构如图 3-F, 参考模板为 *H. jecorina* CBH I 的 C-末端区域 (PDB: 1cbh)^[18], 虽然模拟图与图 3-B (外切葡聚糖酶

CBD 结构域) 相同, 但是该区域与 *H. jecorina* 外切酶 I 纤维素结合域的氨基酸序列相似性只有 72.22%。

EG 酶蛋白的催化结构域的三级结构如图 3-G 所示, 参考模板为 *H. jecorina* Cel5A (PDB: 3qr3)^[22], 且该区域两者氨基酸序列相似性达 94.80%; CBD 结构域位于该蛋白质的 N 端, 其三级结构如图 3-H 所示, 参考模板为 *H. jecorina* CBH I 的 C-末端区域 (PDB: 1cbh)^[18], 虽然模拟图与图 3-B (外切葡聚糖酶 CBD 结构域) 相似, 但是该区域与 *H. jecorina* 外切酶 I 纤维素结合域的氨基酸序列相似性仅有 62.86%。

BGL 酶蛋白只含有催化结构域, 其三级结构如图 3-I, 参考模板为 *Kluyveromyces marxianus* Cel3 (PDB: 3abz)^[23], 且该区域两者氨基酸序列相似性达 23.75%。

Prosite motif search 分析的酶蛋白纤维素结合域氨基酸序列及位置与 SWISS-Model 预测及分析的序列完全符合。通过对纤维素酶蛋白各功能域分析及预测, 可为进一步构建具有高催化效力的酶蛋白突变体提供重要依据^[24, 25]。

4 结论

本研究证明东方肉座菌 EU7-22 相比其他丝状真菌 (里氏木霉、康宁木霉、黑曲霉和斜卧青霉) 具有较高的产纤维素酶能力, 并具有完整的纤维素酶系。菌株 EU7-22 主要纤维素酶基因 (cbh、cbh、eg、eg 和 bgl) 被成功克隆。序列分析表明, 与木霉属其他菌株相应的基因相似性均达 99%, 氨基酸序列相似性均为 100%, 并具有糖基水解酶家族特征序列。CBH、CBH、EG、EG 及 BGL 的分子量分别为 54.23、49.61、48.16、44.22 和 78.21 kD; 等电点分别为 4.58、4.91、4.72、4.96 和 5.96; 含有糖基化位点数分别为 4、3、5、1 和 1 个。

CBH N 端前 17 个氨基酸, CBH N 端前 18 个氨基酸, EG N 端前 22 个氨基酸, EG N 端前 21 个氨基酸及 BGL N 端前 19 个氨基酸为信号肽序列。这些纤维素酶蛋白 CD 及 CBD 域的三级结构被模拟。

参考文献

- [1] 曲音波. 木质纤维素降解酶系的基础和技术研究进展. 山东大学学报: 理学版, 2011, 46 (10): 160-170.
- [2] Sukumaran RK, Singhania RR, Mathew GM, et al. Cellulase production using biomass feed stock and its application in lignocellulose saccharification for bio-ethanol production. *Renew Energ*, 2009, 34 (2): 421-424.
- [3] Kamm B, Kamm M. Biorefineries-multi product processes. *Adv Biochem Engin Biotechnol*, 2007, 105: 175-204.
- [4] 方翎, 秦玉琪, 李雪芝, 等. 纤维素酶与木质纤维素生物降解转化的研究进展. *生物工程学报*, 2010, 26 (7): 864-869.
- [5] Chtistian P. The *Trichoderma* cellulase regulatory puzzle. *Enzyme Microb Tech*, 1993, 15 (2): 90-99.
- [6] Xu C, Long MN, Wu XB, et al. Screening and characterization of a high cellulase producing strain *Aspergillus glaucus* XC9. *Front Biol China*, 2006, 1: 35-40.
- [7] Ghose TK. Measurement of cellulase activities. *Pure Appl Chem*, 1987, 59: 257-268.
- [8] Saha BC. Production, purification and properties of endoglucanase from a newly isolated strain of *Mucor circinelloides*. *Process Biochem*, 2004, 39: 1871-1876.
- [9] 汪天虹. 分子生物学实验 [M]. 北京: 北京大学出版社, 2009: 21-22.
- [10] Wilkins MR, Gasteiger E, Bairoch A, et al. Protein identification and analysis tools on the ExPASy Server. *Methods Mol Biol*, 1999, 112: 531-552.
- [11] Petersen TN, Brunak S, Heijne GV, et al. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*, 2011, 8: 785-786.
- [12] Gupta R, Brunak S. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing*, 2002, 7: 310-322.
- [13] Sigrist CJA, Cerutti L, de Castro E, et al. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Res*, 2010, 38: 161-166.

- [14] Arnold K, Bordoli L, Kopp J, et al. The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. *Bioinformatics*, 2006, 22: 195-201.
- [15] Goedegebuur F, Fowler T, Phillips J, et al. Cloning and relational analysis of 15 novel fungal endoglucanases from family 12 glycosyl hydrolase. *Curr Genet*, 2002, 41: 89-98.
- [16] Song JZ, Liu BD, Liu ZH, et al. Cloning of two cellobiohydrolase genes from *Trichoderma viride* and heterogenous expression in yeast *Saccharomyces cerevisiae*. *Mol Biol Rep*, 2010, 37: 2135-2140.
- [17] Hui JPM, White TC, Thibault P. Identification of glycan structure and glycosylation sites in cellobiohydrolase II and endoglucanases I and II from *Trichoderma reesei*. *Glycobiology*, 2002, 12: 837-849.
- [18] Kraulis J, Clore GM, Nilges M, et al. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry*, 1989, 28: 7241-7257.
- [19] Zou J, Kleywegt GJ, Ståhlberg J, et al. Crystallographic evidence for substrate ring distortion and protein conformational changes during catalysis in cellobiohydrolase Ce16A from *Trichoderma reesei*. *Structure*, 1999, 7 (9) : 1035-1045.
- [20] Mattinen ML, Kontteli M, Kerovuo J, et al. Three-dimensional structures of three engineered cellulose-binding domains of cellobiohydrolase I from *Trichoderma reesei*. *Protein Sci*, 1997, 6 (2) : 294-303.
- [21] Kleywegt GJ, Zou JY, Divne C, et al. The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes. *J Mol Biol*, 1997, 272: 383-397.
- [22] Lee TM, Farrow MF, Arnold FH, et al. A structural study of *Hypocrea jecorina* Cel5A. *Protein Sci*, 2011, 20 (11) : 1935-1940.
- [23] Yoshida E, Hidaka M, Fushinobu S, et al. Role of a PA14 domain in determining substrate specificity of a glycoside hydrolase family 3 β-glucosidase from *Kluyveromyces marxianus*. *Biochem J*, 2010, 431 (1) : 39-49.
- [24] Carrard G, Koivula A, Soderlund H, et al. Cellulose-binding domains promote hydrolysis of different sites on crystalline cellulose. *Proc Natl Acad Sci USA*, 2000, 97 (19) : 10342-10347.
- [25] Lantz SE, Goedegebuur F, Hommes R, et al. *Hypocrea jecorina* CEL6A protein engineering. *Biotechnol Biofuels*, 2010, 3: 20.

(责任编辑 李楠)