

食用菌纤维素内切酶的分子建模与分析

黄婷婷, 易卓林, 茆灿泉

(西南交通大学 生命科学与工程学院/分子进化与应用生物学实验室, 四川 成都 610031)

摘要:采用 SWISS-MODEL、Discovery Studio 等生物信息方法预测和分析食用菌纤维素内切酶(EG)分子的三级结构。结果显示:草菇(*V. voluacea*)和云芝(*T. hirsuta*)EG的分子外形呈“桶”状,由外层的 α -螺旋和内层的 β -折叠组成,无规卷曲位于其2个底面;草菇EG的R129作为盐键形成中心,参与形成6个盐键,其催化残基E215和E248参与形成5个氢键,而里氏木霉(*T. reesei*)EGII的催化残基参与形成3个氢键。基于EG连接区和催化相关序列氢键分析,认为草菇和云芝的该区域相关碱基或氨基酸可进行突变和优化。目前未见对草菇和云芝EG的分子建模的研究报道。此研究对于食用菌纤维素内切酶结构的深入了解和结构优化及应用具有重要的意义和价值。

关键词:分子建模;纤维素内切酶;食用菌

中图分类号:Q71 文献标志码:A 文章编号:1000-2286(2011)03-0584-05

Molecular Modeling and Analyses of the Endoglucanases from Edible Fungi

HUANG Ting-ting, YI Zhuo-lin, MAO Can-quan

(School of Life Science and Engineering Southwest Jiaotong University, Chengdu 610031, China)

Abstract: SWISS-MODEL and Discovery Studio were used to predict and analyze the tertiary structure of edible fungi endoglucanases. It was found that the endoglucanases of *V. voluacea* and *T. hirsuta* appear to like a barrel with α -helices in their inner layer and β -folding sheets in its outer layer, with random coils in their two bottoms; as a salt-bond formation center in *V. voluacea* endoglucanase, arginine129 participates in the formation of six salt-bonds; catalytic residues glutamic acid215 and glutamic acid248 in *V. voluacea* endoglucanase participate in the formation of five hydrogen bonds, while the number for *T. reesei* endoglucanase II is three. Based on the analyses of the linker sequences and catalytic related hydrogen-bond, the authors suggest that amino acids in these sites might be mutable optimized for *V. voluacea* and *T. hirsuta* endoglucanases. The work would be helpful for structure optimization and utilization of edible fungi endoglucanases.

Key words: molecular modeling, endoglucanase, edible fungi

纤维素内切酶(endo-1,4- β -D-glucanase E.C.3.2.1.4,来自真菌简称EG,来自细菌简称Len)是纤维素复合酶中重要的一种,作用于纤维素分子内部的非结晶区,随机水解 β -1,4-糖苷键,将长链纤维素分子截短,产生大量带非还原性末端的小分子短链低聚糖,这些产物容易被 β -葡萄糖苷酶降解为葡萄糖分子,而葡萄糖分子是很重要的生物资源,可应用于食品、能源等产业。

收稿日期:2010-11-30 修回日期:2011-03-30

基金项目:中央高校基本科研专项资金(SWJTU09ZT28)

作者简介:黄婷婷(1986—),女,硕士生,主要从事分子进化工程研究,E-mail:huang77wanygi@163.com;* 通讯作者:茆灿泉,教授,博士生导师,E-mail:maocq@home.swjtu.edu.cn。

纤维素内切酶分子一般由两个功能域,即催化结构域(catalytic domain, CD)、纤维素结合结构域(cellulose binding domain, CBD)以及一个高度糖基化的连接子(linker)三部分组成^[1]。不同的纤维素酶分子有不同的功能域组成,有的可能不止两个功能域或者仅有一个功能域。CD是纤维素酶序列中最长的一段肽链,约有250~300个以上氨基酸。不同来源纤维素酶的CD基本一致。内切酶的活性位点位于一个开放的裂口(Cleft)中,它可能与纤维素链的任何部位结合并切断纤维素链。

食用菌是纤维素内切酶优良的生产者和利用者。深入了解和分析食用菌纤维素内切酶结构与功能具有重要的意义和价值。本文通过分子建模等生物信息学手段对属于糖苷水解酶家族5的草菇(*Volvariella volvacea* (Bull. ex Fr.) Sing)和云芝(*Trametes hirsuta* (Wulf.) Pilát.)纤维素内切酶结构进行了预测和分析,而对于其它不属于糖苷水解酶家族5来源于食用菌的纤维素内切酶(如双胞蘑菇EG属于糖苷水解酶家族7),由于同源建模对序列相似度的要求,这里未进行分析。本文还对同属于糖苷水解酶家族5^[2-3]的纤维素重要降解真菌——里氏木霉(*Trichoderma reesei*)的内切酶重要组分EG II进行了比较和分析。期望在分子结构层面对其差异进行解释,同时也对食用菌纤维素内切酶的人工分子进化改造提出可能的途径和方法。

1 材料与方法

1.1 序列来源

文中序列来自Genbank。*T. hirsuta* EG 登录号: BAD01163; *V. volvacea* EG 登录号: AAG59832; *T. reesei* EG II 登录号: ABA64553。

1.2 MEGA、Clustal X 比对 EGs 蛋白质序列

将收集到的fasta格式序列装载到Clustal X上。装载完成后,用鼠标拖至选择待比对的多条序列,然后选择“alignment”,在弹出的下拉列表中选择“do complete alignment”即可。

1.3 SWISS-MODEL 蛋白质建模

食用菌纤维素内切酶的分子建模采用SWISS-MODEL^[4-7]等生物信息学的方法对其进行预测。进入SWISS-MODEL(<http://www.expasy.org/swissmod/SWISS-MODEL.html>),选择“first approach mode”粘贴序列,点击“submit modeling request”按钮,等待并保存PDB文件。模板序列来自黄色嗜热子囊菌(*Thermoascus aurantiacus*)EG的B链(PDB ID: 1gzjB)。

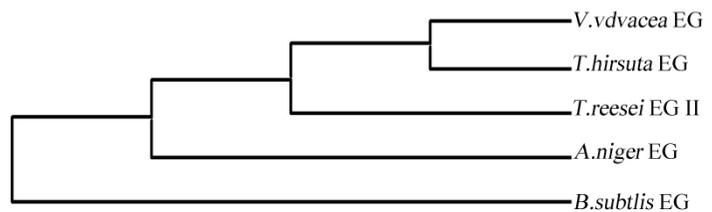
1.4 分子结构的观察比较

采用软件Discovery studio对通过分子建模构建的分子三级结构进行展示和分析比较。进入discovery软件。点击菜单中“new→3D structure”命令,将PDB格式的文件拖放到新建的窗口中。

2 结果

2.1 Clustal X、MEGA 同源性分析

通过系统发育软件MEGA对5条EG蛋白质序列(分别是里氏木霉EG II、草菇EG、云芝EG、黑曲霉*Aspergillus niger* EG和枯草杆菌*Bacillus subtilis* EG)进行同源性分析。结果表明,来源于食用菌的2条EG蛋白质序列之间同源性最高,它们同里氏木霉EG II序列也能较好地聚类,而来源于黑曲霉的EG及枯草杆菌EG同源性更低一些,如图1所示。



草菇EG、云芝EG、里氏木霉EG II、黑曲霉EG和枯草杆菌EG。

图1 5条EGs蛋白质序列的同源性分析(MEGA4.0)

Fig. 1 Evolutionary tree of five EGs protein sequences

通过Clustal X比对里氏木霉EG II和草菇、云芝EG的蛋白质序列发现,里氏木霉EG II序列连接区域中大量存在T和S残基,而云芝EG和草菇EG分子中的连接区域有许多非极性R基氨基酸代替了里氏木霉EG II的不带电荷极性R基氨基酸或带电荷残基,替换率为8.96%。例如,里氏木霉EG II的

I63、P69、G71、T109、D110、K122; 云芝 EG 为 P59、A65、P68、V94、P96、P108; 草菇 EG 为 P64、A70、P73、V100、P102、P114。

2.2 SWISS - MODLE 蛋白质建模

2.2.1 里氏木霉 EG II 由于模板的原因并没有获得全长范围的预测结果,而是主要集中在酶分子的 CD 催化区。里氏木霉 EG II 三级结构建模的结果包含里氏木霉 EG II 蛋白序列的 146 位到 398 位氨基酸残基。该区域属于催化区,后者从 125 位残基开始。这一段序列的三级结构呈桶状(图 2)。“桶”的“缺口”是 146 位上游的残基。“桶”中为活性裂隙,催化纤维素 1 β -糖苷键的断裂。“桶”是由外层的 α -螺旋和内层的 β -折叠组成 α -螺旋和 β -折叠是由无规卷曲相连。所有无规卷曲的大部分残基位于“桶”的上下两面。

图 3 显示了里氏木霉 EG II 的催化残基 E239、E350 在分子中的位置和它们的侧链羧基所在位置。2 个催化残基处于整个桶状结构凹处的中上部。其侧链上的羧基靠近。2 个残基与其他 3 个残基形成氢键。E239 侧链羧基的 1 个氧原子与 H309 咪唑基上的 -N-H 形成氢键。E350 侧链羧基的 2 个氧原子分别与 Y311 的酚羟基和 R151 的侧链胍基之间形成氢键。

2.2.2 草菇 EG 草菇 EG 分子建模的起止位置分别是 F84 和 V389。而由分子的功能域分析知道,草菇 EG 分子的 60 ~ 105 位氨基酸残基肽段为连接子,106 ~ 365 位肽段为催化区域。说明建模部分包含了连接子的部分和整个催化区域。建模结果可见整个结构呈较完整的桶状(图 4)。结合文献和多序列比对的结果,认定草菇 EG 的催化残基为 E215 和 E324。

稳定草菇 EG 2 个活性位点的氢键共有 5 个。E215 侧链羧基的 1 个氧原子和 H282 的酰胺氢原子形成氢键,键长为 2.704 Å。E215 的主链酰胺氢与 E248 的侧链羧基上氧原子形成氢键,键长为 2.728 Å。E324 的侧链羧基的 2 个氧原子分别与 Y284 的酚羟基和 R129 的侧链上胍基形成氢键,键长分别为 2.655 Å、2.846 Å。E324 的主链羧基氧与 Q283 的主链酰胺氢原子形成氢键,键长 2.954 Å。在草菇 EG 桶状结构的凹处比活性区域内更深的位置,发现一处盐键密集的地方,这些键对稳定活性区域的结构有重要的作用。R129 是一处盐键形成中心,参与形成了 6 个盐键。另外, R236 参与形成 2 个盐键。

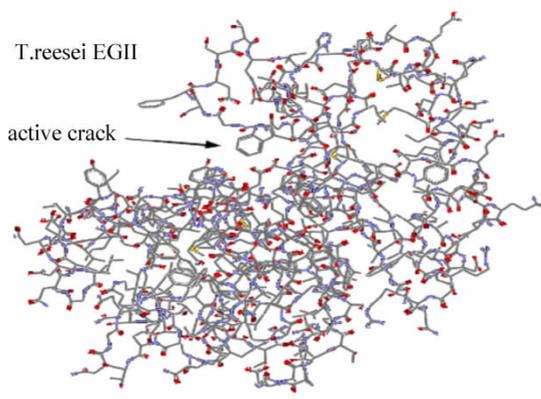
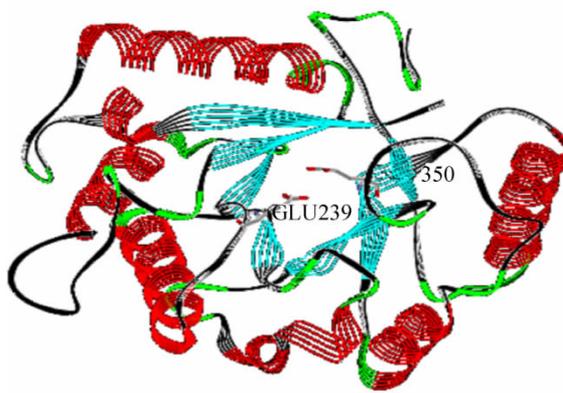
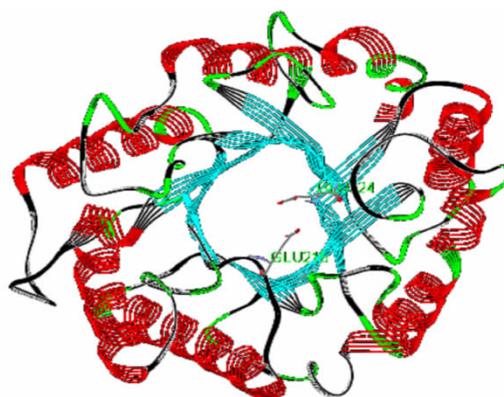


图 2 里氏木霉 EG II 活性裂隙
Fig. 2 Active crack in *T. reesei* EG II



图中所示 E239、E350 为里氏木霉 EG II 的催化残基。
E239 and E350 (as indicated) are the two catalytic residues of *T. reesei* EG II.

图 3 里氏木霉 EG II 的活性中心及 2 个催化残基
Fig. 3 Active center in *T. reesei* EG II and two catalytic residues



图中所示 E215、E324 为草菇 EG 的催化残基。
E215 and E324 (as indicated) are the two catalytic residues of *V. volvacea* EG.

图 4 草菇 EG 的活性中心和 2 个催化残基
Fig. 4 Active center in *V. volvacea* EG and two catalytic residues

	*:****: . . . * . . . : * :**** * : * . * . * : :		::*:****,*** . *. ** * : : . : * * * .
reeseiEGII	WFGIMN ^E PHD ^E VNIN ^I WAATVQ ^E EVVTA ^I RNAGATSQ ^F ISLPGNDWQSAGAFIS	T. reeseiEGI	IFDVH ^K YLDSDNSGTHAECT ^N NI ^D GA ^F SPLAT-WLRQNNRQAIL ^T ETGGGNV
virideEGIII	WFGIMN ^E PHD ^E VNIN ^I WAATVQ ^E EVVTA ^I RNAGATSQ ^F ISLPGNDWQSAGAFIS	T. virideEGI	IFDVH ^K YLDSDNSGTHAECT ^N NI ^D GA ^F SPLAT-WLRQNNRQAIL ^T ETGGGNV
virideEG	WFGIMN ^E PHD ^E VNIN ^I WAATVQ ^E EVVTA ^I RNAGATSQ ^F ISLPGNDWQSA ^A AFIS	T. virideEG	IFDLH ^K YLDSDNSGTHAECT ^N NI ^D GA ^F APLAT-WLRQNNRQAIL ^T ETGGGNV
hirsutaEG	IFDLMN ^E PHD ^I PAQTVFQ ^L MQAAVGNVRASGATSQ ^L ILVEGTSW ^T GAWTWT	T. hirsutaEG	AIQM ^H QYLDSDSGTS ^F TCV ^S STI ^G AERLQAATQ ^L WQKGLK ^G FLG ^E ITG ^T GNN
volvaceaEG	IFDVMN ^E PYQ ^I DASVFN ^L NQA ^A INGIRASGATSQ ^L ILVEGTAW ^T GAWSWES	V. volvaceaE	AIEM ^H QYLDSDSGTS ^F ATCV ^S STV ^G VERL ^R VATD ^L WLRNNL ^K GFLG ^E MGAGSN
ler	240.....250.....260.....270.....280.....29	ruler	0.....320.....330.....340.....350.....360..

图中矩形标出了各 EG 蛋白质序列中的催化残基 均为 E。

The rectangular marked the catalytic residues in the EGs which belong to glycoside hydrolase family 5. for each EG ,both of the two catalytic residues are Glu.

图 5 属于 GHF5 的食用菌 EG 和木霉 EG 的 2 个催化残基

Fig.5 Two catalytic residues in *V. volvacea* ,*T. hirsuta* and *T. reesei* EGs which belong to glycoside hydrolase family 5

2.2.3 云芝 EG 云芝 EG 分子建模包含了其连接子(54 ~ 87) 的部分和整个催化区域(88 ~ 341) 。云芝 EG 的催化残基是 E210 和 E319 ,与里氏木霉和草菇 EG 一样 ,云芝 EG 的 2 个催化残基位于其桶形结构的凹处中。稳定这 2 个催化残基的氢键共有 5 个。E210 的侧链羧基的一个氧原子和 H277 的酰胺氢原子形成氢键 ,键长为 2.72 Å。E210 的主链酰胺氢与 E243 的侧链羧基上氧原子形成氢键 ,键长为 2.742 Å。E324 的侧链羧基的 2 个氧原子分别与 Y279 的酚羟基和 R123 的侧链上胍基之间形成氢键 ,键长分别为 2.657 Å 2.843 Å。E319 的主链羧基氧与 Q278 的主链酰胺氢原子形成氢键 ,键长 2.915 Å。

3 讨 论

从 NCBI 数据库中我们仅收集到 2 种属于 GHF5 的食用菌 EG 蛋白序列(图 5) 。考虑里氏木霉 EGII 与草菇、云芝 EG 一样同属于 GHF5 ,具有相似的序列和结构 ,为此我们仅对该 3 种真菌 EG 进行了分子建模和结构分析。2 种食用菌 EG 建模的部分包含了连接子的部分和整个催化区域。整个结构呈较完整的“桶”状。“桶”是由外层的 α - 螺旋和内层的 β - 折叠组成 ,而 α - 螺旋和 β - 折叠是由无规卷曲相连。这一结构与糖苷水解酶家族 5 成员之催化区的典型结构类似^[8]。文中预测的里氏木霉 EGII 催化区的结构以及在催化反应中分别作为质子供体和亲核基团的 2 个关键残基在酶分子中的位置与 Yuqi Qin 等^[9]的预测结果一致。

比较草菇 EG 和里氏木霉 EGII 酶活性位点可以发现 ,稳定各自活性位点的氢键在数目和键长上都有差异。首先 ,稳定里氏木霉 EGII 的活性位点 E239 和 E309 的氢键数目一共只有 3 个 ,而草菇 EG 有 5 个。多序列比对、三级结构以及氢键分析共同说明存在于里氏木霉 EGII 中的 3 个氢键与草菇 EG 的其中 3 个氢键相似 ,这 3 对相似的氢键中 ,除 G - R 的键长相似以外 ,另 2 对都是草菇 EG 的较长。另外 ,草菇 EG 中与活性位点相关的其余 2 个氢键 ,即 E215 的主链酰胺氢原子与 E248 的侧链羧基上氧原子形成的氢键以及 E324 的主链羧基氧原子与 Q283 的主链酰胺氢原子形成的氢键 ,使得草菇 EG 的 2 个活性位点都较里氏木霉 EGII 的稳定。这使得 2 个酶在酶促反应中的行为出现差别 ,即出现酶催化特性的差别。可能是这 2 个氢键约束活性残基 ,使之间接约束活性残基的催化基团 ,纯化与底物的反应。与草菇 EG 类似 ,云芝 EG 分子中的催化残基 E210 和 E319 参与形成 5 个氢键。通过同里氏木霉 EGII 的比较 ,我们发现云芝 EG 除具有 3 对与前者相似的活性中心相关氢键以外 ,也有另外 2 对氢键 ,即: E210 的主链酰胺氢原子与 E243 的侧链羧基上氧原子形成的氢键以及 E319 的主链羧基氧原子与 Q278 的主链酰胺氢原子形成的氢键。由于草菇 EG 的催化残基相关的氢键较里氏木霉 EGII 多 2 个 ,即 E215 的主链酰胺氢与 E248 的侧链羧基上氧原子形成的氢键和 E324 的主链羧基氧与 Q283 的主链酰胺氢原子形成的氢键。这 2 个氢键很有可能影响活性中心发挥其作用。所以可以考虑将草菇 EG 的 E248 和 Q283 进行突变。类似地 ,我们提出对云芝 EG 的 E243 和 Q278 进行突变。蛋白质的结构影响功能 ,分析云芝、草菇 EG 与里氏木霉 EGII 之间结构上的相似与不同 ,可能为人工分子进化改造以提高食用菌纤维素内切酶的比活力提供思路^[11-14]。

参考文献:

[1] Gilkes N R , Henrissat B , Kilburn D G , et al. Domains in microbial beta - 1,4 - glycanases: Sequence conservation , function , and enzyme families [J]. Microbiol Rev , 1991 , 55(2) : 303-315.

- [2]胡莹,杨凌,杨胜利. 糖苷酶序列分类法和作用机理的研究进展[J]. 药物生物技术, 2006, 13(1): 66-70.
- [3]Henrissat B. A classification and mechanisms of glycosidaese[J]. Pharmaceutical Biotechnology, 2006, 13(1): 66-70.
- [4]Arnold K, Bor doli L, Kopp J, et al. The SWISS – MODEL Workspace: A web – based environment for protein structure homology modelling[J]. Bioinformatics, 2006, 22(1): 195-201
- [5]Kopp J, Schwede T. The SWISS – MODEL repository of annotated three – dimensional protein structure homology models [J]. Nucleic Acids Research, 2004, 32(4): 230-234.
- [6]Schwede T, Kopp J, Guex N, et al. SWISS – MODEL: an automated protein homology – modeling server[J]. Nucleic Acids Research, 2003, 31(3): 3381-3385.
- [7]Guex N, Peitsch M C. SWISS – MODEL and the Swiss – PdbViewer: An environment for comparative protein modelling[J]. Electrophoresis, 1997, 18(2): 2714-2723.
- [8]Lo Leggio L, Larsen S. The 1.62 Å structure of *Thermoascus aurantiacus* endoglucanase: completing the structural picture of subfamilies in glycoside hydrolase family 5[J]. FEBS Lett, 2002, 523(1-3): 103-108.
- [9]Qin Y Q, Wei X M, Song X, et al. Engineering endoglucanase II from *Trichoderma reesei* to improve the catalytic efficiency at a higher pH optimum[J]. Journal of Biotechnology, 2008, 135(2): 190-195.
- [10]Michael B, Howard, Nathan A, et al. Identification and analysis of polyserine linker domains in prokaryotic proteins with emphasis on the marine bacterium *Microbulbifer degradans* [J]. Protein Sci, 2004, 13(5): 1422-1425.
- [11]江华,于兆海. 里氏木霉纤维素酶系的分离及其酶学性质[J]. 南京林业大学学报, 2007, 31(6): 48-52.
- [12]Wu S F, Ding S J, Zhou R, et al. Comparative characterization of a recombinant *Volvariella volvacea* endoglucanase I (EG1) with its truncated catalytic core (EG1 – CM), and their impact on the bio – treatment of cellulose – based fabrics [J]. J Biotechnol, 2007, 130(4): 364-369.
- [13]Nozaki K, Seki T, Matsui K, et al. Structure and characteristics of an endo – beta – 1,4 – glucanase, isolated from *Trametes hirsuta*, with high degradation to crystalline cellulose [J]. Biosci Biotechnol Biochem, 2007, 71(10): 2375-2382.
- [14]Dashtban M, Maki M, Leung K T, et al. Cellulase activities in biomass conversion: measurement methods and comparison [J]. Crit Rev Biotechnol, 2010.

(上接第 567 页)

- [5]Livak K J, Schmittgen T D. Analysis of relative gene expression data using real – time quantitative PCR and the 2 – Delta Delta C (T) method [J]. Methods, 2001, 25: 402-408.
- [6]Higuchir, Fockler C, Dollinger G, et al. Kinetic PCR analysis: real – time monitoring of DNA amplification reactions [J]. Biotechnology, 1993, 11(9): 1026-1030.
- [7]Kel D, Chen Z, Yung W K. A reliability test of standard based quantitative PCR: exogenous standards [J]. Mol Cell Probes, 2000, 14(2): 27-135.
- [8]Becker K, Pan D, Whitel Y C B. Real – time quantitative polymerase chain reaction to assess gene transfer [J]. Hum Gene Ther, 1999, 10: 2559-2566.
- [9]Shim W S, Dlirenzo J, Decaprio J A, et al. Segregation of steroid receptor coactivator – 1 from steroid receptors in mammary epithelium [J]. Proc Natl Acad Sci USA, 1999, 96: 208-213.
- [10]Meijer O C, Steenbergen P J, Kloet E R. Differential expression and regional distribution of steroid receptor coactivators SRC – 1 and SRC – 2 in brain and pituitary [J]. Endocrinology, 2000, 141: 2192-2192.
- [11]Ogawa H, Nishi M, Kawata M. Localization of nuclear coactivators p300 and steroid receptor coactivator 1 in the rat hippocampus [J]. Brain Res, 2001, 890: 197-202.
- [12]黎寿丰, 赵振华, 黄华云, 等. *NCOA1* 基因多态性与邵伯鸡母系产蛋性状的相关性分析 [J]. 云南农业大学学报, 2011, 26(2): 215-219.