

玉米黑粉菌 CYP51 稀有密码子和 mRNA 二级结构分析及与杀菌剂戊唑醇分子对接*

李书祥¹⁾ 韩睿^{1, 2)} 袁利玲¹⁾ 熊丽¹⁾ 袁永泽¹⁾ 杨江科³⁾ 闫云君³⁾ 刘德立^{1)**}

¹⁾华中师范大学生命科学学院遗传调控与整合生物学湖北省重点实验室, 武汉 430079; ²⁾青海大学农林科学院园艺研究所, 西宁 810016;

³⁾华中科技大学生命科学与技术学院, 武汉 430074)

摘要 通过截短玉米黑粉菌 CYP51(P450-14DM, UmCYP51)基因(去除编码跨膜区部分)和选取不同的表达载体, 构建了 9 种重组表达质粒, 在大肠杆菌中进行 UmCYP51 基因的表达, 发现只有 BL21(DE3)/pET32-Um-35 重组表达工程菌获得了表达. 对稀有密码子和 mRNA 翻译起始区二级结构进行分析, 结果表明稀有密码子和 mRNA 翻译起始区二级结构对 UmCYP51 蛋白的表达都有影响. 适用于稀有密码子表达的菌株 Rosetta(DE3)不利于 UmCYP51 蛋白的表达; 同时只有翻译起始区二级结构自由能值最低的重组载体 pET32-Um-35 可以表达. 为了设计以 UmCYP51 为靶标的新型抗真菌抑制剂, 基于最新解析的真核生物人类的 CYP51 晶体结构, 利用同源模建的方法构建了 UmCYP51 的三维结构并进行了分子动力学模拟优化. 通过与商品化杀菌剂戊唑醇进行分子对接获得了此类抑制剂与 UmCYP51 的理论结合方式, 阐述了戊唑醇分子的杀菌机理, 为开发新型的抗真菌抑制剂奠定了基础.

关键词 UmCYP51, 翻译起始区, 稀有密码子, 同源模建, 分子对接

学科分类号 Q617, Q786, Q811.4

DOI: 10.3724/SP.J.1206.2010.00617

玉米黑粉菌(*Ustilago maydis*)是引起玉米黑粉病的主要病原菌, 发病时直接造成农作物整穗无粒, 给农业生产造成巨大的经济损失. 细胞色素 P450 甾醇 14 α - 去甲基化酶(P450-14DM, CYP51)是真菌细胞膜甾醇生物合成过程中的一个关键酶, 该酶的缺乏会导致真菌细胞膜结构的破坏和功能的丧失, 最终造成真菌死亡^[1]. 农作物真菌 CYP51 已经被证明是许多重要的内吸性杀菌剂特别是三唑类杀菌剂的重要靶标^[2], 因此, 利用玉米黑粉菌 P450-14DM(UmCYP51)为靶标开发新型杀菌药物具有一定的现实意义.

目前所有已经测定的真菌 CYP51 蛋白均是膜结合蛋白, 解析其晶体结构仍然是细胞色素 P450 研究中的一个难题. 同源模建和分子对接的方法能很好地应用于生物大分子的理论模拟和药物设计筛选^[3], 但基于原核生物结核分支杆菌 CYP51 (*Mycobacterium tuberculosis* CYP51, MtCYP51)晶

体结构同源模建的 UmCYP51 三维结构在筛选开发新型高效抗真菌新药存在很大的局限^[4]. 2010 年 Lepesheva 等^[5]成功解析了第一个真核生物布鲁氏锥虫(*Trypanosoma brucei*)CYP51 的晶体结构; 随后 Strushkevich 等^[6]解析了真核生物人类羊毛甾醇 14 α - 脱甲基酶(CYP51)的晶体结构. 本文以最新解析的人 CYP51 晶体结构为模板, 通过同源模建构建了 UmCYP51 的三维结构, 采用分子对接等手段分析了杀菌剂戊唑醇与 UmCYP51 的具体结合方

* 国家自然科学基金(30771429, 31071653), 国家高技术研究发展计划(863)(2007AA05Z417)和教育部博士点基金(20060511002)资助项目.

** 通讯联系人.

Tel: 027-67865534, E-mail: deliliu2002@yahoo.com.cn

收稿日期: 2011-01-18, 接受日期: 2011-04-01

式,明确了UmCYP51与此类抑制剂结合时起重要作用的氨基酸残基,阐明了杀菌剂戊唑醇的抑菌作用机理,为进一步研发和筛选新型高效抗真菌新药奠定了良好基础.

1 材料和方法

1.1 UmCYP51 基因重组表达质粒的构建和表达

参照华中师范大学遗传调控与整合生物学湖北省重点实验室构建UmCYP51重组表达工程菌的条件和方法^[7],选择野生型、N端截短20个氨基酸残基(60个碱基)和N端截短35个氨基酸残基(105个碱基)的UmCYP51突变体,与3种表达载体pET28、pET32和pGEX-KG构建了9种重组表达质粒,分别为pET28-Um、pET28-Um-20、pET28-Um-35、pET32-Um、pET32-Um-20、pET32-Um-35、pGEXKG-Um、pGEXKG-Um-20和pGEXKG-Um-35.稀有密码子可能会影响到目的基因表达,实验中选择了稀有密码子表达效应不同的两种菌株*E. coli* BL21 (DE3)和Rossetta (DE3)进行重组质粒的表达.将得到的各种重组质粒分别转化大肠杆菌,30℃以0.5 mmol/L IPTG分别诱导2、4、6和8 h,4 000 r/min离心集菌,分别取样并进行SDS-PAGE分析.

1.2 稀有密码子分析和翻译起始区二级结构预测

利用codon usage database(<http://www.kazusa.or.jp/coden>)对UmCYP51基因ORF进行分析.通过计算机辅助预测,采用RNAStructure 4.5对构建的9种重组表达质粒翻译起始区(translation initiation region, TIR)的mRNA二级结构和自由能(ΔG)进行分析预测^[8].

1.3 UmCYP51 的同源模建和评价

以UmCYP51的蛋白质序列(来源于Swiss-Prot and TrEMBL数据库,收录号:A1XG20)为探针,在NCBI数据库中利用protein blast程序搜索Brookhaven蛋白质晶体数据库,得到模板蛋白人的细胞色素P450 甾醇14- α -脱甲基酶的三维结构,收录号为3LD6^[9].利用CLUSTERW程序^[9]进行序列比对,确定模板与序列之间的残基匹配.根据序列比对的结果,采用MODELLER 9v7^[10]自动建模程序搭建UmCYP51的三维结构并进行侧链异构化,对所得模型在pH为7.0条件下进行加氢、补键等化学修饰,得到UmCYP51的初始结构模型.

模型优化采用GROMACS 4.0.3^[11]程序包进行分子动力学模拟,将得到的UmCYP51初始结构模型置于立方体盒子中心,选择溶质原子到盒子壁的距离为1.0 nm,然后加入水分子,水分子选用SPC模型,添加3个钠离子取代体系中的水分子以平衡体系中多余的负电荷.采用GROMOS96 43a1力场^[12],将蛋白质-水体系用最陡下降法(steepest descents, SD)进行5 000步的能量优化,以避免模拟时原子间的不合理碰撞.然后在300K下用2 fs的积分步长进行了50 ps的限制性动力学优化.最后使用热浴耦合使系统温度保持在300K,用Maxwell分布产生初始原子速度,每隔2 ps记录一次轨迹,进行3 ns的MD模拟直到体系能量达到平衡.在MD模拟中,用PME (Particle-Mesh Ewald)方法处理静电作用^[13],范德华相互作用截断半径取0.9 nm.分子动力学模拟得到稳定的模型后再经简单的能量优化用于下步的分子对接,模型的合理性评估采用Procheck^[14]检查键长、键角及二面角的合理性.

1.4 分子对接

分子对接计算在Autodock 4.2.3^[15]程序上进行.受体蛋白UmCYP51采用AutoDockTools (ADT)加氢和添加Kollman电荷.商品化杀菌剂戊唑醇分子用化学图形软件Chemdraw构建,经sybyl 7.0软件用Powell方法Gasteiger Hückel力场优化,用Biopolymer模块加氢并为所有原子加上部分电荷,将其作为对接中配体的初始结构^[16].对接选用的受体格点盒子大小为7 nm×7 nm×7 nm,格点间隔为0.037 5 nm,设置盒子中心位于活性中心Heme的铁原子上.采用拉马克遗传算法(Lamarckian genetic algorithm, LGA),将局部能量搜索与遗传算法相结合,从100个任意的配体初始位置开始,对小分子构象和位置进行全局搜索,配体和受体间的能量匹配用半经验势函数打分^[17].遗传算法的种群数(Population size)选择默认的150,能量极大值(Maximum number of evals)设为(medium)2 500 000,最大变异数(Maximum number of generations)为27 000,遗传交叉模式(GA Crossover mode)为twopt模式.对配体分子进行100次独立的对接实验,最后依据对接能和成簇分析,筛选出合理的构象作为复合物的结构进行分析^[18].氢键作用的判据参数^[19-21]:氢键给体-受体间的最大距离为0.30 nm,氢键给体-氢-受体间的键角大于120°.

2 结果与分析

2.1 UmCYP51 重组质粒的表达

将构建的 9 种重组质粒分别转入 *E. coli* BL21 (DE3) 菌株和适合稀有密码子表达的 *E. coli* Rossetta (DE3) 菌株进行重组质粒的诱导表达, 通过添加 0.5 mmol/L IPTG 在 30℃ 条件下分别诱导 2、4、6 和 8 h, SDS-PAGE 检测结果表明, 只有 pET32-Um-35 重组质粒在 BL21(DE3) 菌株中得到表达, 在 70 ku 处发现明显的特异性表达带(图 1), 与预测的分子质量大小相符. 其他 8 种重组质粒在 *E. coli* BL21 (DE3) 和 Rossetta (DE3) 菌株中均未表达.

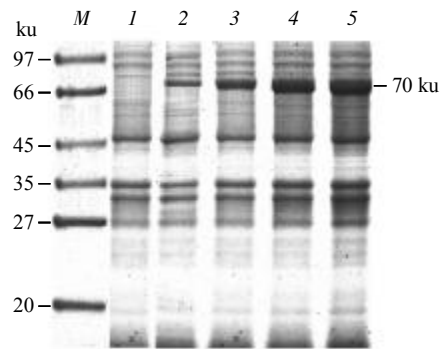


Fig. 1 SDS-PAGE analysis of recombinant protein

M: Protein molecular mass markers; 1: Uninduced; 2~5: Induced for 2, 4, 6 and 8 h.

2.2 UmCYP51 基因稀有密码子分析

UmCYP51 基因序列中显示的稀有密码子和统计结果分析见图 2 和表 1. 结果表明 UmCYP51 基因含有稀有密码子共 26 个, 占总基因序列比例只

有 4.63%, 这表明 UmCYP51 适合在 BL21(DE3) 菌株中表达, 适用于稀有密码子表达的菌株 Rosetta (DE3) 反而不利于 UmCYP51 蛋白的表达, 这与实验结果是一致的.

ATG GTG GCC TCC TCG TCT TCG GCG ACA GCG TCG TTG CTC GAC CAG CTG TTT GCG CTT ACG CCT CTC GCC GAT TCT TCG GCC TGG ATC AAA ACC
 ATC ACG GTG CTC GTT TTG CTT CCC TTG CTC GCG GTC GTT CTC AAC GTA GCA TCG CAG CTT CTG CTT GCG ACT CCC AAG AAT CAT CCT CCG GTG
 GTG TTT CAC TTT GTG CCT GTG ATT GGA TCG GCG ATC TAC TAT GGT ATC GAC CCT TAC AAG TTC TTT TTC GAA TGC CGC GAA AAG TAT GGC GAT
 GTG TTC ACG TTT GTT CTG CTC GGA CGC AAG ATC ACG GTG GCG TTG GGA CCC AAG GGA AGC AAT CTT GTC TTT AAC GCC AAG CAT CAG CAG GTG
 ACG GCA GAG GAC GCG TAT ACG CAC CTC ACC ACG CCC GTG TTT GGC AAA GAG GTG GTA TAC GAT GTG CCT AAT GCG GTG TTT ATG GAG CAG AAA
 AAG TTT GTC AAG GTG GGT CTG TCG ATC GAA AAC TTT CGC GTC TAC GTG CCG CAG ATC GTG GAT GAA GTG CGA GAG TAC ATC AAG AGC GAT GCG
 CGT TTC AGC GCA CTC AAG ACG CGC AAG ACG ATC ACA GTC GAT ATT TTT CAA GCC ATG TCG GAA CTC ATC ATC CTG ACC GCA TCC AGA ACG CTG
 CAG GGC AAG GAA GTC AGA CAG GGT CTC GAC AAG TCA TTT GCG CAA CTG TAT CAC GAT CTC GAC TCG GGC TTC ACT CCG ATC AAC TTT GTG ATT
CCC AAC CTG CCG CTG CCG AGC AAC TTT AAA GCG GAC AGG GCG CAG AAG AAA ATG TCG CAG TTC TAC CAG GAC ATT GTG GCG AAA CGA CGA GCT
 GCG GGT GCA TCC ACG TCT GCC GAC GAC GCC AGT GGC GAA AAC GAT ATG ATC GCA GCA CTG ATC GAG CAA AAG TAC AAG AAC GGA CGT GCA CTC
 AGC GGC GTC GAG ATT GCA CAT ATG ATG ATT GCA CTC TTG ATG GCG GGT CAG CAC ACG AGC AGC GCC ACG TCG TCG TGG GCG TTT CTG CGT CTG
 GCT AGT CGA CCC GAA ATC ATC GAG GAG CTA TAC GAG GAG CAG CTC AAC GTG TAC TCT GAC GGA CAC GGT GGT TTG AGG GAA CTC GAC TAC GAG
 ACG CAA AAG ACG TCG GTG CCT CTG CTG GAT GCT GTG GTC AAA GAG ACG TTG CGA CTG CAC CCG CCT CTG CAC AGC ATC ATG CGA TAT GTG AAA
 TCC GAC CTT GCC GTT CCA CCT ACG CTC TCG TCG CCG ACG TCG ACT AAA TCC GAG CCG GAT GCC CAC TAT GTG ATT CCC AAG GGC CAC TAC ATT
 ATG GCT GCA CCT GGC GTG TCG CAA GTG GAT CCT CAG ATT TGG AAG TCA TCG GAC CAG TTT GAT CCA CAC CGG TGG TTG GAC GCT ACG ACC GCT
 GCG GCG ATG CAG GAC TCG GCG GAG GAC AAG CAA GAC TTT GGC TTC GGA ATG ATC TCC ACC GGT GCC AAT AGC CCA TAC CTT CCC TTT GGC GCC
 GGA AGA CAT CGG TGC ATC GGA GAG CAG TTC GCC TAT CTC CAG ATC GGC GTC ATC CTA GCC ACC TTC GTC CGC ATC TTC AAA TGG CAC CTC GAC
 TCC AAA TTC CCC GAT CCC GAC TAC CAA AGC ATG GTC GTC CTT CCC AGC AAA AAC GGT TGC GCC ATC GTC CTC ACC CCC CGA GCC GAA TCC CTC
 CAC CTC GAC TAG

Fig. 2 Analysis of rare codons

Different rare codons are shown with underlines.

Table 1 The statistics of rare codons

Amino acid	Arg			Leu	Pro
Rare codon	AGG	AGA	CGA	CTA	CCC
Times	2	3	7	2	12
Frequency(Times/562)%	0.36	0.53	1.25	0.36	2.14
Total frequency/%	4.63				

2.3 翻译起始区二级结构预测结果

根据 GANOZA 等的方法分析翻译起始区, 选择翻译起始点上游-30 到下游+70 的 100 个碱基进行^[2]. 结果表明在构建的 9 种重组表达质粒中, pET32-Um-35 具有最低的二级结构自由能(图 3). mRNA 翻译起始区存在稳定的二级结构对翻译起始不利, 这种二级结构在翻译过程中会影响核糖体与其结合位点以及起始密码子的结合, 干扰蛋白质

表达. 如果二级结构过于稳定, 会阻碍翻译起始^[2], 而当翻译起始区结构稳定性降低时, 这种抑制作用会减弱并使目的蛋白的表达量提高^[2]. 实验中使用了同样的诱导表达条件, 而只有翻译起始区二级结构自由能最低的 pET32-Um-35 蛋白得到表达; 表明 UmCYP51 基因翻译起始区二级结构的稳定性对该蛋白质的原核表达也有影响, 选择不稳定的翻译起始区二级结构有利于 UmCYP51 的优化表达.

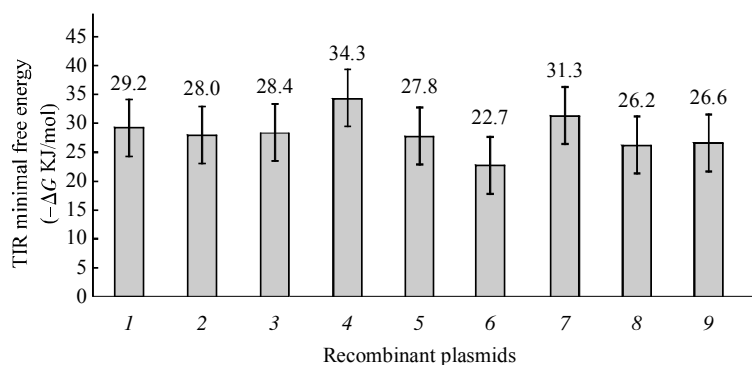


Fig. 3 The minimal free energy of ($\Delta G/(\text{kJ}\cdot\text{mol}^{-1})$) nine recombinant plasmids

1: pET28-Um; 2: pET28-Um-20; 3: pET28-Um-35; 4: pET32-Um; 5: pET32-Um-20; 6: pET32-Um-35; 7: pGEXKKG-Um; 8: pGEXKKG-Um-20; 9: pGEXKKG-Um-35.

2.4 UmCYP51 的同源建模

通过序列比对发现 UmCYP51 与模板蛋白人类 CYP51 氨基酸序列的同源性达到 39%, 而一直以来作为真核生物 CYP51 同源建模模板的原核生物 MtCYP51 与 UmCYP51 的同源性只有 27%^[4]. 由自动建模程序 modeler 构建的初始模型, 经过能量最

小化和分子动力学模拟后, 得到了 UmCYP51 三维结构的合理模型, 如图 4a 所示, UmCYP51 的模建结构含有 12 个 α 螺旋, 6 个 β 折叠和 11 个 β 转角, 与模板蛋白构象非常相似. 通过与模板蛋白叠合(图 4b)发现 UmCYP51 与 3LD6 主链碳原子的均方根偏差为 0.079 nm.

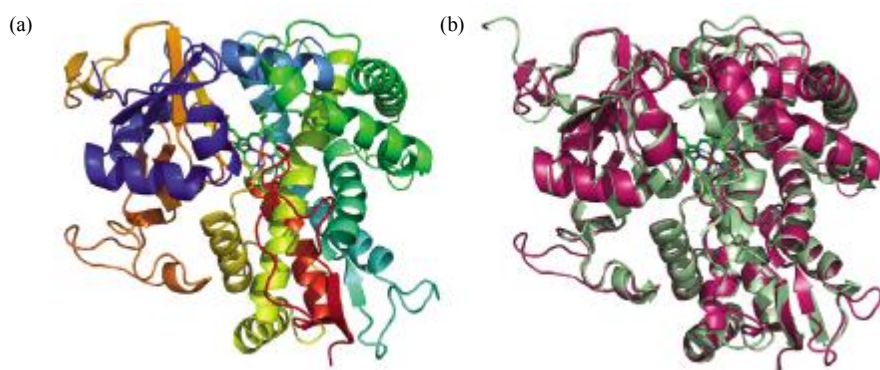


Fig. 4 Homology modeling of UmCYP51

(a) Ribbon representation of UmCYP51. (b) Superposition of the model of CYP51 (pink) and the template (green).

使用 Procheck 程序评价模型结构的立体化学参数, 此程序凭经验比较所给蛋白质结构与最合理的蛋白质结构之间立体化学性质的差异, 评价结果中主要考察所有氨基酸残基骨架的二面角分布, 通过程序生成的 Ramachandran 图来表示(图 5). 分析得到的模建蛋白发现 85.1%的残基落入最佳区域, 12.2%的残基落入其他许可区, 1.6%的残基落入勉强许可区, 处于合理区域的氨基酸比例为 98.9%, 上述结果说明模建结构氨基酸的二面角是合理的.

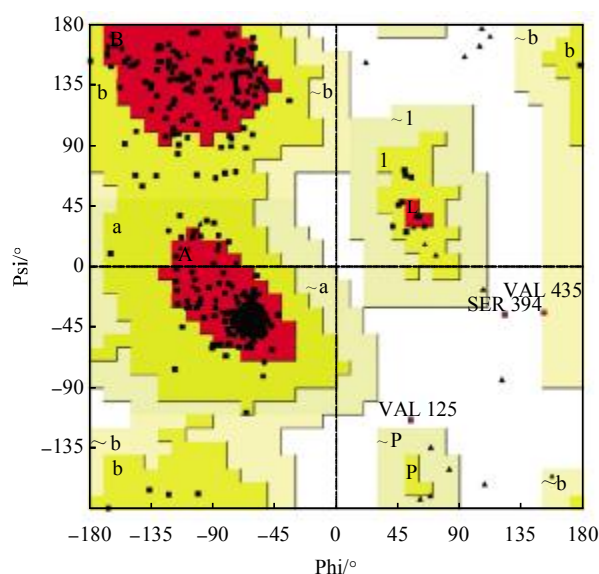


Fig. 5 Ramachandran plot of UmCYP51 modeling

The favored and most favored region is yellow and red respectively, pale yellow is the generally allowed and disallowed regions is white, the black foursquares represent the amino acids in the favored and most favored regions, the black triangles and red foursquares represent the amino acids in the generally allowed and disallowed regions respectively.

2.5 分子对接

在同源模建所得玉米黑粉菌 CYP51 三维结构基础上, 利用 Autodock 4.2.3 将戊唑醇对接到 MD 模拟优化后的 UmCYP51 受体蛋白质的活性中心(图 6), 受体与戊唑醇的最低相互作用自由能为 -59.75 kJ/mol. 分析对接结果可以看出, 戊唑醇分子三唑环上的氮原子通过与 UmCYP51 血红素卟啉环中铁原子形成配位键产生相互作用, N 原子

与 Fe 原子之间的距离为 0.245 nm. 戊唑醇羟基的氢原子与受体 Ala265 羰基氧原子之间形成了氢键^[19-21], 距离为 0.245 nm, 给体 - 氢 - 受体键角为 121.63° . 在受体活性中心形成的大疏水口袋中, 配体分子戊唑醇主要通过疏水作用与受体相互作用. 其中一部分疏水氨基酸如 Phe82、Val87、Phe95 和 Ala265 共同组成一个疏水空腔, 与戊唑醇的叔丁基发生作用. Tyr74、Leu77、Thr78 和 Phe185 四个氨基酸残基则形成了一个疏水狭缝, 与戊唑醇的单氯代苯基发生疏水相互作用. 这些分析表明配体戊唑醇通过配位键、氢键和疏水作用牢固地与受体 UmCYP51 结合在一起, 发挥抑制作用. 这一作用机理为我们设计新型的 UmCYP51 抑制剂分子提供了理论基础. 因为三唑环基团和单/双卤代苯基在唑类杀菌剂中具有特异性, 在设计新型抑制剂时可以保留这些基团. 在唑类抑制剂远离血红素一端引入疏水作用较强的基团, 或是引入能与受体氨基酸残基产生氢键作用的基团, 则能加强抑制剂分子与 UmCYP51 的相互作用, 进而设计新型的 UmCYP51 抑制剂.

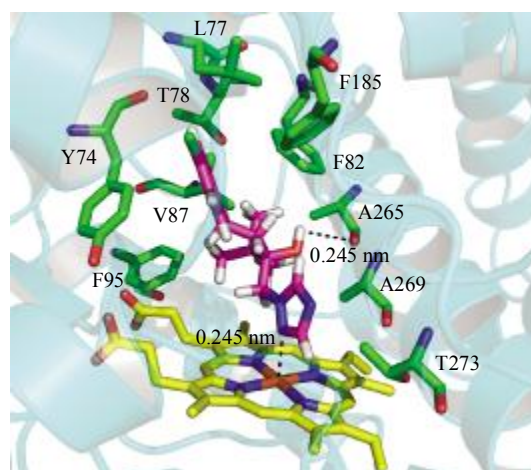


Fig. 6 Docking result of inhibitor tebuconazole with UmCYP51

3 讨论

玉米黑粉菌 P450-14DM(UmCYP51)的高效表达是进行体外抑菌活性测试, 进而阐述抑菌机理和开发新型、高效、广谱类抗真菌抑制剂的基础. 理

论上真核细胞偏爱的密码子和原核系统不同,用原核系统表达真核基因时,真核基因中的一些密码子对原核细胞来说可能是稀有密码子进而导致表达效率和表达水平很低.通过对UmCYP51基因的密码子分析表明其稀有密码子使用频率较低,采用适合稀有密码子表达的菌株Rosetta(DE3)反而不利于UmCYP51蛋白的表达.实验证实,在Rosetta(DE3)菌株中9种重组质粒均没有表达,而在优化后的BL21(DE3)菌株中重组质粒pET32-Um-35表达成功.说明外源基因的表达与mRNA密码子依赖性有直接的关系,真菌UmCYP51偏爱的密码子和原核系统基本一致.

RNA二级结构的稳定性同样是影响外源基因在大肠杆菌中表达水平的因素.本实验中选择了3种不同的UmCYP51基因型(野生型、60个碱基和105个碱基)和3种表达载体,构建了9种重组表达工程菌,分析了翻译起始区二级结构稳定性对UmCYP51基因的影响.在翻译过程中,不稳定的翻译起始区二级结构有利于核糖体与其结合位点以及起始密码子的结合,从而提高目的蛋白的表达量^[2].分析表明pET32-Um-35具有最低的二级结构自由能,这种不稳定的二级结构能有助于其蛋白质表达.实验结果证实选取BL21(DE3)菌株时只有pET32-Um-35可以表达.本实验对表达系统翻译起始区(TIR)二级结构进行了分析和优化,实现了UmCYP51基因在大肠杆菌中的高效表达.

本文采用同源建模的方法对玉米黑粉菌CYP51的三维结构进行了建模,分子对接分析得到了玉米黑粉菌CYP51与杀菌剂戊唑醇最佳作用模式的详细理论信息.从二者之间的相互作用能、氢键和疏水作用的形成来看,在三唑类抑制剂远离血红素的一端上引入疏水作用较强的基团可以用于开发新型的UmCYP51抑制剂.从UmCYP51与三唑类抑制剂的对接结果中发现,玉米黑粉菌CYP51的Tyr74、Leu77、Thr78、Phe82、Val87、Phe95、Phe185、Ala265和Ala269等残基在与抑制剂分子的相互作用过程中起到重要的结合作用,通过进一步定点突变分析,可以研究近年来出现三唑类抑制剂对UmCYP51的抗药性.分子对接结果表明建模的UmCYP51 3D结构适合于进行UmCYP51抑制剂的虚拟筛选,为开发设计新型的抗真菌抑制剂奠定了良好基础.

参 考 文 献

- [1] Lepesheva G, Hargrove T, Kleshchenko Y, *et al.* CYP51: A major drug target in the cytochrome P450 superfamily. *Lipids*, 2008, **43**(12): 1117-1125
- [2] Lamb D C, Kelly D E, Venkateswarlu K, *et al.* Generation of a complete, soluble, and catalytically active sterol 14 alpha-demethylase-reductase complex. *Biochemistry*, 1999, **38** (27): 8733-8738
- [3] Sheng C, Miao Z, Ji H, *et al.* Three-dimensional model of lanosterol 14 alpha-demethylase from cryptococcus neoformans: active-site characterization and insights into azole binding. *Antimicrob Agents Chemother*, 2009, **53**(8): 3487-3495
- [4] Zhao L, Liu D L, Zhang Q Y, *et al.* Expression and homology modeling of sterol 14 alpha-demethylase from *Penicillium digitatum*. *FEMS Microbiol Lett*, 2007, **277**(1): 37-43
- [5] Lepesheva G I, Park H W, Hargrove T Y, *et al.* Crystal structures of *Trypanosoma brucei* sterol 14alpha-demethylase and implications for selective treatment of human infections. *J Biol Chem*, 2010, **285**(3): 1773-1780
- [6] Strushkevich N, Usanov S A, Park H W. Structural basis of human CYP51 inhibition by antifungal azoles. *J Mol Biol*, 2010, **397**(4): 1067-1078
- [7] 韩睿, 邓灵福, 黎晨, 等. 玉米黑粉菌cyp51基因结构分析与克隆表达. *生物工程学报*, 2008, **24**(10): 1747-1753
Han R, Deng L F, Li C, *et al.* *Chin J Biotech*, 2008, **24** (10): 1747-1753
- [8] 张静, 石秀凡, 刘次全, 等. Yeast基因下游二级结构与多聚腺苷作用信号. *生物化学与生物物理进展*, 2001, **28**(4): 523-527
Zhang J, Shi X F, Liu C Q, *et al.* *Prog Biochem Biophys*, 2001, **28**(4): 523-527
- [9] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994, **22**(22): 4673-4680
- [10] Eswar N, Webb B, Marti-remom M A, *et al.* Comparative protein structure modeling using MODELLER. *Curr Protoc Protein Sci*, 2007, **2**(12): 15-32
- [11] Lindahl E, Hess B, Vander S D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J Mol Model*, 2001, **7**(8): 306-317
- [12] van Gunsteren W F, Billeter S R, Eising A A, *et al.* Biomolecular Simulation: The GROMOS96 Manual and User Guide. Groningen: vdf Hochschulverlag AG an der ETH Zürich and BIOMOS b.v., 1996: 21-25
- [13] YORK D, YANG W. The fast fourier poisson method for calculating ewald sums. *J Chem Phys*, 1994, **101**(4): 3298-3300

- [14] Laskowski R A, Macarthur M W, Moss D S, *et al.* PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystall*, 1993, **26**(2): 283–291
- [15] Morris G M, Goodsell D S, Halliday R S, *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J Comput Chem*, 1998, **19**(14): 1639–1662
- [16] 朱海梅, 陈慰祖, 王存新, 等. HIV-1 整合酶与抑制剂金精三羧酸复合物的分子动力学模拟. *化学学报*, 2004, **62**(8): 745–749
Zhu H M, Chen W Z, Wang C X, *et al.* *Acta Chim Sin*, 2004, **62**(8): 745–749
- [17] 李 博, 刘 明, 胡文祥. 芬太尼类化合物与阿片 μ 受体相互作用的分子对接与分子动力学模拟. *物理化学学报*, 2010, **26**(01): 206–214
Li B, Liu M, Hu W X. *Acta Phys-Chim Sin*, 2010, **26**(01): 206–214
- [18] 丛肖静, 谭建军, 王存新, 等. HIV-1 跨膜蛋白 gp41 与 N-取代吡咯衍生物的结合模式研究. *生物化学与生物物理进展*, 2010, **37**(8): 904–915
Cong X J, Tan J J, Wang C X, *et al.* *Prog Biochem Biophys*, 2010, **37**(8):904–915
- [19] Wallace A C, Laskowski R A, Thornton J M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*, 1995, **8**(2): 127–134
- [20] Azizian H, Bahrami H, Pasalar P, *et al.* Molecular modeling of *Helicobacter pylori* arginase and the inhibitor coordination interactions. *J Mol Graphi Mod*, 2010, **28**(7): 626–635
- [21] 马晓慧, 陈慰祖, 王存新, 等. 用从头药物设计方法合理设计 HIV-1 整合酶抑制剂. *军事医学科学院院刊*, 2005, **29**(4): 255–358
Ma X H, Chen W Z, Wang C X, *et al.* *Bull Acad Mili Med Sci*, 2005, **29**(4): 255–358
- [22] Ganoza M C, Kofoed E C, Marliere P, *et al.* Potential secondary structure at translation-initiation sites. *Nucleic Acids Research*, 1987, **15**(1): 345–360
- [23] Rex G, Surin B, Besse G, *et al.* The mechanism of translational coupling in *Escherichia coli*. Higher order structure in the atpHA mRNA acts as a conformational switch regulating the access of *de novo* initiating ribosomes. *J Biol Chem*, 1994, **269** (27): 18118–18127
- [24] De S M, Vanduin J. Control of translation by mRNA secondary structure in *Escherichia coli*: a quantitative analysis of literature data. *J Mol Biol*, 1994, **244**(2): 144–150

Analysis of Rare Codon and mRNA Structure About *Ustilago maydis* CYP51 and Molecular Docking With Fungicide Tebuconazole *

LI Shu-Xiang¹⁾, HAN Rui^{1,2)}, YUAN Li-Ling¹⁾, XIONG Li¹⁾, YUAN Yong-Ze¹⁾,
YANG Jiang-Ke³⁾, YAN Yun-Jun³⁾, LIU De-Li^{1)**}

¹⁾ Hubei Key Laboratory of Genetic Regulation and Integrative Biology, College of Life Science, Huazhong Normal University, Wuhan 430079, China;

²⁾ Institute of Horticulture Qinghai Academy of Agriculture & Forestry, Qinghai University, Xining 810016, China;

³⁾ College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Abstract To get a better optimization expression of the *Ustilago maydis* CYP51 (P450-14DM, UmCYP51) protein in *E. coli* BL21 (DE3), the different lengths of UmCYP51 gene that lacked the coding region for the putative membrane-spanning segment of the N-terminus were truncated. The first one is the wild type, the second one with 20 amino acids (60 base pairs) in N-terminus was truncated and the third one with 35 amino acids (105 base pairs) was truncated. Then these genes were incorporated into different expression vectors (pET28, pET32 and pGEX-KG) to construct nine recombinant expression plasmids (pET28-Um, pET28-Um-20, pET28-Um-35, pET32-Um, pET32-Um-20, pET32-Um-35, pGEXKG-Um, pGEXKG-Um-20 and pGEXKG-Um-35). The expression of recombinant plasmids were performed using 0.5 mmol/L of isopropyl β -D-thiogalactoside (IPTG) at 30°C. The culture harvested every 2 h up to 8 h. It was found that only recombinant plasmid pET32-Um-35 was expressed in *E. coli* BL21(DE3). Codon usage database (<http://www.kazusa.or.jp/coden>) was used for the analysis of rare codon and software RNAstructure 4.5 was employed to study the mRNA secondary structure of translation initiation region. The results showed that rare codons rate in UmCYP51 gene is only 4.63%, the Rosetta (DE3) strain expressing some rare codons is not suitable for the protein expression of UmCYP51. Only the lowest energy of mRNA structure for pET32-Um-35 can obtained protein expression. These results are compatible with the experiments. Moreover, to design novel antifungal compounds against UmCYP51, based on the recently determined X-ray crystal structure human CYP51, a three-dimensional structure model of UmCYP51 was built through homology modeling using MODELLER 9V7 program. After refinement of the energy minimization and MD simulation using GROMACS 4.0.3 package, the UmCYP51 model was evaluated by PROCHECK Ramachandran plot statistics that indicated the designed model was in good quality. Commercial fungicide tebuconazole was docked into the model protein using Autodock 4.2.3 program to form the binding pattern of inhibitor with UmCYP51. The docking conformation of tebuconazole in the active site of UmCYP51 showed that the N-4 of the triazole ring was bound to heme iron with a distance about 0.245 nm. The hydroxy group of tebuconazole formed hydrogen-bonding interaction with the oxygen atom of carbonyl group for Ala265 with a distance about 0.245 nm. The mechanism of inhibitory activity of tebuconazole against UmCYP51 obtained from this study could aid in designing new antifungal compounds targeting this enzyme.

Key words UmCYP51, translation initiation region (TIR), rare codon, homology modeling, molecular docking

DOI: 10.3724/SP.J.1206.2010.00617

*This work was supported by grants from The National Natural Science Foundation of China (30771429, 31071653), Hi-Tech Research and Development Program of China (2007AA05Z417) and The Specialized Research Fund for the Doctoral Program of Higher Education (20060511002).

**Corresponding author.

Tel: 86-27-67865534, E-mail: deliliu2002@yahoo.com.cn

Received: January 18, 2011 Accepted: April 1, 2011