

• 研究论文 •

## 人顶体酶三维结构的同源模建及其与 KF950 的分子对接研究

吕加国\* 盛春泉 张珉 季海涛  
张万年 周有骏 朱驹 蒋俊航

(第二军医大学药学院 上海 200433)

**摘要** 采用同源模建方法首次构建了人顶体酶的三维结构模型, 模型的可靠性经 Ramachandran 图和 Profile\_3D 图验证. 采用 InsightII/Binding site 方法准确定位了人顶体酶的活性位点, 并研究了顶体酶重要功能残基在活性位点的立体分布. 在此基础上, 通过柔性分子对接方法首次阐明了顶体酶高效抑制剂 KF950 与靶酶活性位点的相互作用模式, 发现特异性的氢键相互作用是 KF950 产生高抑制活性的重要分子基础. 其研究结果将为合理设计新型顶体酶抑制剂, 寻找男性口服避孕药奠定坚实基础.

**关键词** 人顶体酶; 同源模建; 活性位点; 4-胍基苯甲酸(4-甲氧甲酰基)苯酯甲烷磺酸盐; 分子对接

## Homology Modeling of Human Acrosin and Its Molecular Docking Study with KF950

LÜ, Jia-Guo\* SHENG, Chun-Quan ZHANG, Min JI, Hai-Tao  
ZHANG, Wan-Nian ZHOU, You-Jun ZHU, Ju JIANG, Jun-Hang

(School of Pharmacy, Second Military Medical University, Shanghai 200433)

**Abstract** Homologous 3D model of human acrosin was built on the basis of the crystal coordinates of ram and boar acrosin, while the reliability of the model was assessed by Ramachandran plot and Profile-3D analysis. The active site of human acrosin was searched by Insight II/binding site analysis and important functional residues were located at the active site. To explore the binding mode of the 4-guanidinobenzoates with the active site of human acrosin, KF950 was docked into the active site and specific hydrogen-bonding interaction was found to play an important role in inhibitor recognition and orientation. These results provided a basis for the rational design of specific inhibitors to the target enzyme and the discovery of novel contraceptive agents with high potency.

**Keywords** human acrosin; homology modeling; active site; 4-guanidinozoic acid (4-methoxycarbonyl)-phenyl ester monomesylate; molecular docking

受精是精子与卵子相互结合启动新生命发育的过程. 当精子接近或进入卵丘细胞与卵透明带结合时, 其头部发生顶体反应. 顶体反应使精子顶体内所含的顶体酶(acrosin)释放出来, 顶体酶水解透明带糖蛋白, 支持精子穿过卵透明带, 从而实现精卵融合. 顶体酶位于精子头顶体内膜上, 在顶体反应中起着关键性作用, 是受

精过程中一种不可缺少的重要蛋白水解酶.

顶体酶最初是以酶原形式存在. 在受精的顶体反应过程中, 顶体内容物释放后不久, 顶体酶原(proacrosin)自动激活成  $\beta$ -顶体酶( $\beta$ -acrosin). 其激活过程包含对分子两端的处理, 如小鼠顶体酶原转化为  $\beta$ -顶体酶最初是残基 R23 和邻近的缬氨酸之间的肽键断裂, 形成一条 23

\* E-mail: KF950LDS@yahoo.com.cn

Received September 13, 2005; revised November 17, 2005; accepted January 19, 2006.

上海市计生委局管科技发展基金(No. 05JG05011)资助项目.

个氨基酸的“轻链”和一条 394 个氨基酸的“重链”，它们通过 2 对二硫键交联形成  $\alpha$ -顶体酶( $\alpha$ -acrosin). 此时  $\alpha$ -顶体酶就已具有了蛋白水解酶的活性. 通过 C 末端三次连续蛋白内切分解, 共除去 76 个残基, 其中包括特征聚脯氨酸序列, 最终产生高活性的  $\beta$ -顶体酶( $\beta$ -acrosin)<sup>[1]</sup>.

顶体酶属丝氨酸蛋白水解酶家族中胰蛋白酶族<sup>[2,3]</sup>, 具有丝氨酸水解酶家族特征的 His, Asp, Ser 催化三联体, 并具有潜在的糖基化位点 N(2)和 N(169). 有研究表明顶体酶上活性位点为两个, 彼此分开, 其中一个用于分解蛋白质的催化活性位点; 另一个是与卵细胞透明带糖蛋白受体结合的结合位点<sup>[4,5]</sup>.

尽管在受精过程中顶体酶的精确作用仍不十分清楚, 但体内或体外实验表明, 天然或合成的胰蛋白酶抑制剂可以抑制顶体酶的活性, 从而阻止受精过程的发生. 由于顶体酶的重要功能和显著的高细胞特异性, 因此寻找其专属性抑制剂对于研究新型避孕药有非常重要的意义. 目前人顶体酶的晶体结构仍未测得, 但是已经测得雄羊和雄猪顶体酶的晶体结构. 由于雄羊和雄猪顶体酶与人顶体酶的氨基酸序列具有较高的同源性, 因此利用雄羊和雄猪顶体酶的晶体结构来模建人顶体酶的三维结构具有非常高的可靠性.

我们以雄羊和雄猪顶体酶的晶体结构模建了人顶体酶的三维结构, 并搜寻了可能的药物结合空腔, 同时阐明了人顶体酶高效抑制剂 KF950 与蛋白的相互作用模式. 上述研究为基于人顶体酶开展合理药物设计, 寻找男性口服的避孕药提供思路.

## 1 材料与方法

雄羊顶体酶(分辨率 0.21 nm, Rcryst=0.200)和雄猪顶体酶(分辨率 0.29 nm, Rcryst=0.212)的晶体结构坐标来源于 Brookhaven 蛋白质晶体结构数据库, 记录号分别为 1FIW 和 1FIZ. 人顶体酶的蛋白序列来源于 NCBI 数据库. 所有分子模建、能量优化和柔性对接计算均在 Origin 300 服务器上, 采用 InsightII 2000 和 SYBYL 6.9 商用软件包完成. 计算中选用的各项参数除特别指明均采用缺省值.

### 1.1 序列联配

采用 InsightII/Homology 中的 Align123 程序对以上 3 条序列进行匹配, 参数设置: Blosom-62 矩阵, Gap penalty 为 10, Gap extension penalty 为 1. 然后基于晶体结构知识, 对匹配结果进行微调, 尽量避免在二级结构保守区引入 gap (图 1).

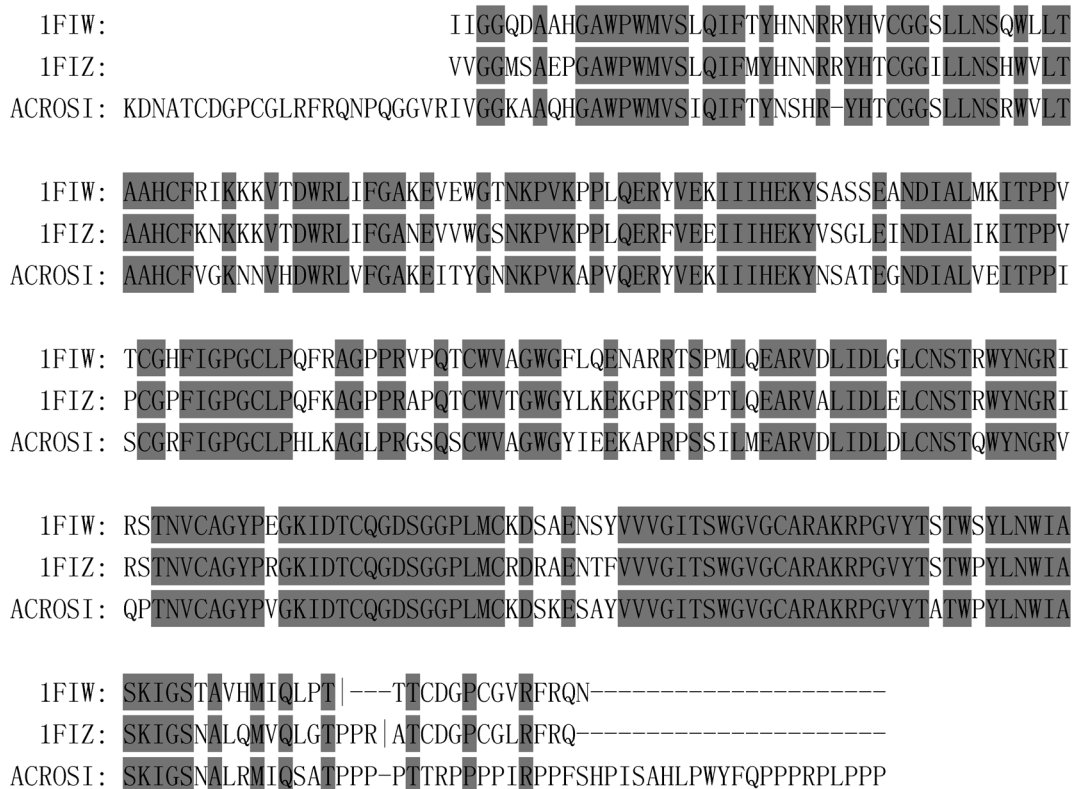


图 1 人、雄羊和雄猪顶体酶的序列联配

Figure 1 Sequence alignment between human acrosin, ram acrosin and bore acrosin

## 1.2 人顶体酶的分子结构模建

采用 InsightII 2000 软件包 Homology 模块中的 Modeler 程序模建蛋白分子. 参数设置: 每次 Modeler 计算构建 5 个模型, 每个模型产生 2 个不同的 loop 区, 优化水平为中度, 模型构建不考虑氢原子, 共计算 5 个循环. 对各模型以 Profiles\_3D 图检验其得分值, 选取得分最高的模型, 进一步确定模建蛋白结构变化区即 loop 区的三维结构. 采用 Modeler/Refine\_Loop 程序, 对每一个 loop 区计算其 5 种构象, 优化水平为中度, 选取 Profiles\_3D 图检验得分最高的构象. 最终所得的模型进行进一步能量优化.

## 1.3 模建蛋白溶剂化模型的构建和能量优化

以上得到的模建蛋白在 pH 为 7 条件下加氢, 得到初始模型. 将整个体系加入 0.5 nm TIP3P 水分子, 采用 InsightII/Discover 95.0 模块进行能量优化, 力场选用 CVFF 力场. 首先采用最陡下降法(Steepest descent method), 当能量梯度低于  $42 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$  时, 采用共轭梯度法(Conjugate gradient method)再优化至能量梯度小于  $4.2 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$ . 能量优化时采用分级方法, 首先对溶剂水和蛋白的氢原子进行能量优化, 进而优化环区侧链、环区主链、结构保守区侧链和结构保守区主链, 直至全部放开优化, 优化过程中采用 Ramachandran 图和 Profiles\_3D 图随时检验优化所得蛋白质的结构合理性, 得到最终的溶剂化模型.

## 1.4 人顶体酶上活性位点的搜寻

采用 InsightII 2000 软件包的 Binding site analysis 模块来搜寻和分析人顶体酶上空腔的大小和分布情况. 参数设置如下: 探针分子(grid)大小定义为  $0.1 \text{ nm}\times 0.1 \text{ nm}\times 0.1 \text{ nm}$ , Site\_Opensize 为 7.00, Site\_Cutoffsize 为 50.

## 1.5 KF950 与人顶体酶的柔性分子对接

采用 SYBYL 分子片段库中片段构建 KF950 分子结构. SYBYL/MINIMIZE, Tripos 标准分子力场, Gasteiger-Hückel 电荷, Powell 能量优化方法, 能量收敛标准为  $0.042 \text{ kJ}\cdot\text{mol}^{-1}\cdot\text{nm}^{-1}$ . 首先对化合物进行分子动力学模拟退火, 参数设置: 升温到 1000 K, 平台处理时间为 1.0 ps, 退火至 250 K, 退火时间为 1.5 ps, 指数型退火函数. Tripos 标准分子力场, Gasteiger-Hückel 电荷, 每次 50 个循环, 得 50 个构象并能量优化, 找到能量最低的构象作为下一步的起始构象. 反复操作, 以重复出现 5 次以上的最低能量构象作为该分子的最低能量构象, 用于下一步的柔性对接.

柔性对接采用了网格对接(Grid docking)策略, 网格点分辨率为 0.05 nm. 对接过程中, 活性位点以外部分以网格代替, 作用位点内原子可在能量许可范围内移

动, 因此能更好反映配体、受体诱导契合的过程. 首先将底物手工对接至两个模建蛋白的活性位点, 并实时评价相互作用能量, 寻找较低能量的初始复合物结构. 采用 InsightII/Affinity 模块中蒙特卡罗方法搜寻化合物在作用位点内的取向空间, 寻找 50 个复合物结构, 彼此间均方根偏差值(RMSD)大于 1 以排除类似结构, 并采用共轭梯度法优化 300 步. 为避免能量值跳跃过大, 采用 Quartic-Vdw-non-Coub 选项仅计算立体相互作用能量, 并设 Scale Vdw 系数为 0.1. 对蒙特卡罗搜寻得到的最优构象, 采取分子动力学模拟退火技术, 高温 500 K, 低温 300 K, 降温过程中共分为 100 步, 每步下作 100 fs 的分子动力学模拟, 最终优化 5000 步至收敛. 计算过程中, 采用 Cell multipole 模型计算非键相互作用.

## 2 结果与讨论

### 2.1 人顶体酶三维结构的总体特征

模建所得的人顶体酶的三维结构见图 2, 它与雄猪和雄羊顶体酶的晶体结构具有相似的总体拓扑结构特征<sup>[4]</sup>. 人顶体酶的三维结构呈“银杏”状, 其蛋白质的折叠特征与胰蛋白酶家族非常相似, 三维结构核心区域存在有多个反平行的  $\beta$ -桶状结构. 由于人顶体酶的 1~23 残基区域在模板晶体结构中没有对等的三维坐标, 因此该区域没有模建. 对顶体酶保守的催化三联体(His, Asp 和 Ser)以骨架碳原子作为叠合点进行叠合, 人与雄猪和雄羊顶体酶主链原子的均方根值(root mean square, rms)值分别为 0.04469 和 0.03628 nm, 全部原子的 rms 分别为 0.11642 和 0.09181 nm, 其中核心区域变化比较小, 大的偏差主要发生在蛋白质表面的 loop 区. 以 Ramachandran 图(图 3)分析发现模建蛋白 87.4%残基的  $\phi$ 和  $\psi$ 角落在最佳区域, 9.5%的残基落入其他许可区,

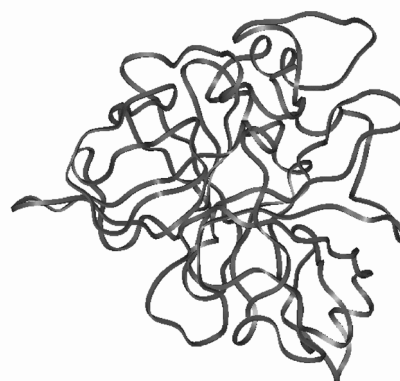


图 2 人顶体酶三维结构的带状示意图  
Figure 2 Ribbon representation of human acrosin

2.7%残基落在勉强许可区, 0.5%残基落在不允许区, 上述结果说明模建结构氨基酸的二面角是非常合理的. 采用 Profile\_3D 图(图 4)检测模建蛋白氨基酸序列与其三维结构的兼容性, 发现所有残基的分值均在 0.098 以上, 无不兼容残基.

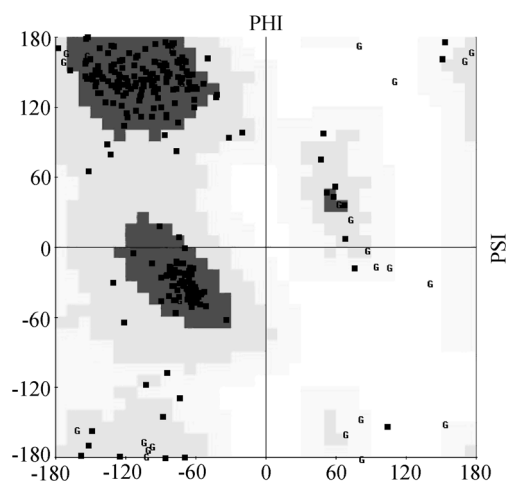


图 3 人顶体酶模建结构的 Ramachandran 图

Figure 3 Main-chain dihedral angles (Ramachandran plot) of modeling human acrosin

## 2.2 人顶体酶上活性位点的分析与定位

Binding site analysis 计算结果显示, 在人顶体酶三维结构上共发现有 6 个较大的空腔, 分别可以容纳 135, 120, 104, 68, 57 和 54 个探针分子, 定义每一个探针分子 0.3 nm 范围内的残基作为该空穴的组成残基(见表 1). 根据雄羊和雄猪顶体酶的晶体结构特征, 顶体酶序列联配结果和已有的顶体酶生物实验结论来准确定位人顶体酶的活性位点.

从一些有限的定点突变实验发现, 公羊顶体酶的 Arg 37, Arg 221, Lys 223 和 Arg 224 均为功能性残基<sup>[6,7]</sup>,

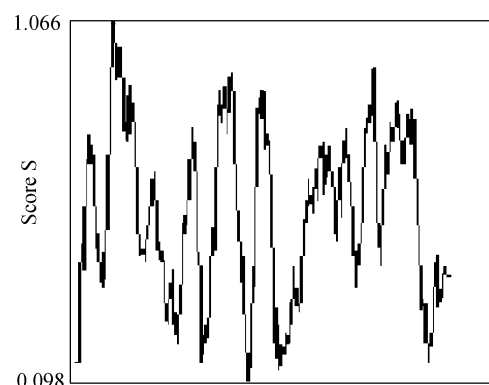


图 4 人顶体酶模建结构的 Profile\_3D 图

Figure 4 Profile\_3D plots of the modeling human acrosin 3D-structures

它们分别对应于人顶体酶的 Arg 50, Arg 249, Lys 251 和 Arg 252. 而雄羊顶体酶上的催化三联体 His 57, Asp 102 和 Ser 195 是顶体酶发挥催化活性的关键残基<sup>[4]</sup>, 它们分别对应于人顶体酶的 His 69, Asp 123 和 Ser 221. 上述这些残基在已知顶体酶序列中都是非常保守的. 上述结果结合对雄羊和雄猪顶体酶晶体结构知识可以判定 Site1 是人顶体酶的活性位点. 首先, 经定点突变实验证实的 4 个功能性残基 Arg 50, Arg 249, Lys 251 和 Arg 252 中有 3 个残基位于 Site 1, Arg 50 位于蛋白质的 N 端, 并不属于其它搜寻得到的空腔; 其次, 对顶体酶发挥催化功能至关重要的催化三联体 His 69, Asp 123 和 Ser 221 位于 Site 1; 再次, 将人顶体酶三维结构与雄羊和雄猪顶体酶晶体结构叠合比较后发现, Site 1 的位置与雄羊和雄猪顶体酶活性位点相一致, 在该区域结合有顶体酶的小分子抑制剂苯甲脒<sup>[4]</sup>. 对于搜寻得到的其它 5 个空腔, Site 2, Site 3, Site 4 和 Site 5 距离 Site 1 较远, Site 6 虽与 Site 1 邻近, 但从其残基分布, 空腔形状和大小来看, Site 6 无法成为活性位点.

表 1 人顶体酶模建结构上搜寻得到的 6 个空腔<sup>a</sup>

Table 1 Six cavities discovered on the modeled structure of human acrosin

Site	Size	Residue
Site 1	135	Ala 67, Ala 68, <u>His 69</u> , Ala 118, Thr 119, Glu 120, Asn 122, <u>Asp 123</u> , Trp 195, Thr 196, Asn 197, Gly 198, Arg 199, Val 200, Gly 212, Thr 216, Cys 217, Gln 218, Gly 219, Asp 220, <u>Ser 221</u> , Trp 243, Gly 244, Val 245, Gly 246, Cys 247, Ala 248, <u>Arg 249</u> , Ala 250, <u>Lys 251</u> , <u>Arg 252</u>
Site 2	120	Arg 153, Gly 154, Ser 155, Ile 185, Asp 186, Leu 187, Pro 202, Thr 203, Asn 204, Val 205
Site 3	104	Ile 24, Val 25, Gly 26, Gly 27, Lys 28, Ala 29, Gly 164, Tyr 165, Ile 166, Arg 172, Pro 173, Ser 174, Ile 176, Leu 177, Met 178, Glu 179, Ala 180, Thr 216, Asp 215, Cys 217
Site 4	68	Ala 149, Gly 150, Leu 151, Pro 152, Cys 227, Lys 228, Asp 229, Ser 230,
Site 5	57	Tyr 46, Asn 47, Ser 48, His 49, Asn 93, Asn 94, Lys 95, Glu 103, Arg 104, Tyr 105
Site 6	54	Thr 53, Cys 54, Cys 70, Gly 73, Lys 74

<sup>a</sup> Identified functional residues are under lined.

图5显示了人顶体酶活性位点在蛋白质三维结构中的位置,它位于蛋白质表面较为狭长的凹槽,周围分布有顶体酶的重要功能残基.基于该空腔可开展合理药物设计,发现能特异性作用于人顶体酶的小分子抑制剂.

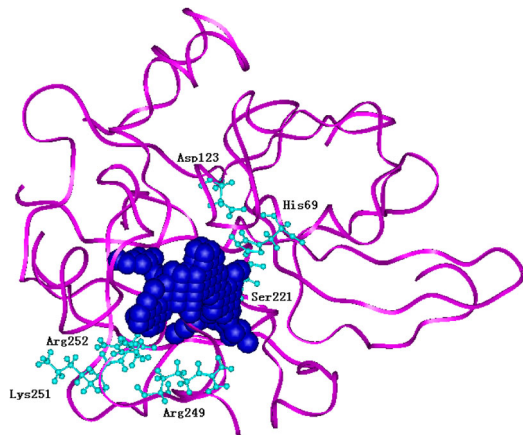


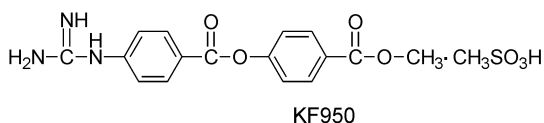
图5 人顶体酶活性位点在其三维结构中的位置

图中蓝色区域为活性位点的范德华表面,用球棒状模式显示的残基为人顶体酶的重要功能残基

Figure 5 The active site of human acrosin in 3D structure

### 2.3 KF950 与人顶体酶活性位点的相互作用模式

4-胍基苯甲酸酯类化合物 KF950 是顶体酶的高效抑制剂,它对人顶体酶的  $IC_{50}$  值为  $6.7 \times 10^{-10} \text{ mol} \cdot \text{L}^{-1}$ ,并且具有很好的体内活性,目前我们正对 KF950 进行临床前研究<sup>[8]</sup>.研究 KF950 与人顶体酶活性位点的相互作用模式有助于阐明 4-胍基苯甲酸酯类顶体酶抑制剂的分子作用机理,为进一步的结构优化指明方向,也可以为开展全新药物设计提供思路.



由于配体小分子与受体蛋白质的结合是一个“动态契合”的过程,配体和受体的构象在结合前后会发生一定程度的改变,而目前大多数对接算法仅考虑配体的柔性而忽略受体柔性,这往往会降低对接计算的准确性.此外,我们对接所采用的人顶体酶的三维结构为同源模建结构,其精度要低于晶体结构,考虑受体活性位点的柔性将有利于得到更为合理的结果.因此我们选用 InsightII/Affinity 模块来开展柔性分子对接研究,该对接策略能同时考虑受体活性位点和配体的柔性,在蒙特卡罗技术能详尽搜寻药物在结合位点的各种取向的基础上,采用分子动力学模拟退火方法对初始复合物进行

优化而得到复合物合理的低能构象<sup>[9]</sup>.

柔性对接结果显示, KF950 在人顶体酶活性位点处于“伸展”的构象,两者的相互作用能为  $-413.6756 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ ,其中范德华相互作用能和静电相互作用能分别为  $-240.45$  和  $-173.2256 \text{ kJ} \cdot \text{mol}^{-1} \cdot \text{nm}^{-1}$ .图6显示了 KF950 与人顶体酶活性位点的相互作用模式,氢键相互作用是 KF950 与人顶体酶活性位点的基本作用力. KF950 可以和人顶体酶活性位点形成 5 对氢键,这也可能是该类化合物产生高抑制活性的重要分子基础.胍基是 4-胍基苯甲酸酯类顶体酶抑制剂的必需基团,将该基团用其它基团替代后会大大降低的抑制活性.对接结果显示, KF950 的胍基可以和 Glu 120 的羧基负离子形成 2 对氢键( $\text{NH}_2 \cdots \text{OCO}$  键长为  $0.221 \text{ nm}$ ,  $\text{NH} \cdots \text{OCO}$  为  $0.170 \text{ nm}$ ),而且 Glu 120 也是顶体酶家族中非常保守的残基.因此可以推测,在顶体酶活性位点识别和准确定位化合物构象时,胍基与 Glu 120 的氢键结合起着十分重要的作用. 4-胍基苯甲酸酯类顶体酶抑制剂中的酯基部分也是非常重要的基团,它可以分别和 Tyr 196 侧链酚羟基( $\text{CO} \cdots \text{HO}$  为  $0.155 \text{ nm}$ )和 Arg 252 侧链胍基( $\text{CO} \cdots \text{H}_2\text{N}$  为  $0.177 \text{ nm}$ )形成 2 对氢键,其中 Arg 252 已经定点突变实验证实是顶体酶的重要功能残基.构效关系研究表明, KF950 分子结构末端苯环的对位取代基中以酯基取代活性最强,对接结果表明该酯基的羰基部分可以和 Gly 254 的主链氨基形成 1 对氢键( $\text{CO} \cdots \text{HN}$  为  $0.248 \text{ nm}$ ),该氢键有助于进一步增强化合物与顶

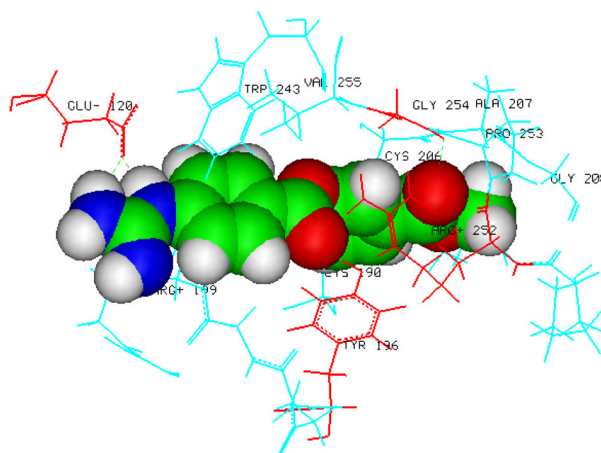


图6 KF950 与人顶体酶活性位点的对接结果

图中显示了 KF950 分子周围  $0.4 \text{ nm}$  范围内的残基,红色残基为能与 KF950 形成氢键相互作用的残基,氢键用点状虚线表示, KF950 分子用 CPK 模式显示

Figure 6 Docking result of KF950 with human acrosin

The residues within  $0.4 \text{ nm}$  from the ligand are shown, the residues involved in the hydrogen bonding with KF950 are colored red, hydrogen bonds are indicated with dotted lines and KF950 is displayed in CPK mode

体酶活性位点的结合. 基于KF950与顶体酶的作用模式可以发现, KF950还没有和顶体酶的催化三联体形成直接的相互作用, 这提示可以在KF950分子结构中进一步引入可以和催化三联体相作用的基团, 这样可以进一步提高分子与顶体酶的亲和力和特异性, 相关的药物设计与合成工作正在进行之中.

### 3 结论

以雄羊和雄猪顶体酶的晶体结构为模板, 采用同源模建方法首次模建了人顶体酶的三维结构, 模建结构的精度经 Ramachandran 图和 Profile\_3D 图检验. 采用 InsightII/Binding site 方法搜寻得到了人顶体酶的活性位点, 并定位了活性位点的重要功能残基. 在此基础上, 首次阐明了 4-胍基苯甲酸酯类顶体酶抑制剂 KF950 与人顶体酶活性位点的相互作用模式, 发现氢键相互作用是 KF950 产生高抑制活性的重要分子基础. 以上研究结果为进一步研究顶体酶的结构与功能关系提供了理论依据, 同时为设计新型顶体酶抑制剂, 寻找男性口服避孕药奠定坚实基础.

### References

- 1 Kashiwabara, S.; Baba, T.; Takada, M.; Watanabe, K.; Yano, Y.; Arai, Y. *J. Biochem.* **1990**, *108*(5), 785.
- 2 Rawlings, N. D.; Tolle, D. P. *Biochem. J.* **2004**, *378*(3), 705.
- 3 Barrett, A. J.; Rawlings, N. D. *Arch. Biochem. Biophys.* **1995**, *318*(2), 247.
- 4 Tranter, R.; Read, J. A.; Jones, R.; Brady, R. L. *Structure* **2000**, *8*(11), 1179.
- 5 Crosby, J. A.; Jones, R.; Barros, C.; Carvallo, P. *Mol. Reprod. Dev.* **1998**, *49*(4), 426.
- 6 Jansen, S.; Jones, R.; Jenneckens, I.; Marschall, B.; Kriegesmann, B.; Coadwell, J.; Brenig, B. *Mol. Reprod. Dev.* **1998**, *51*(2), 184.
- 7 Richardson, R. T.; O'Rand, M. G. *J. Biol. Chem.* **1996**, *271*(39), 24069.
- 8 Lü, J.-G. *CN 1438217A*, **2003** [*Chem. Abstr.* **2005**, *142*, 463459].
- 9 Song, Y.-L.; Zhang, W.-N.; Ji, H.-T.; Sheng, C.-Q.; Zhang, M.; Yao, J.-Z.; Yu, J.-X.; Zhou, Y.-J.; Zhu, J.; Lü, J.-G. *Acta Chim. Sinica* **2003**, *61*(11), 1860 (in Chinese).  
(宋云龙, 张万年, 季海涛, 盛春泉, 张珉, 姚建忠, 余建鑫, 周有骏, 朱驹, 吕加国, 化学学报, **2003**, *61*(11), 1860.)

(A0509133 ZHAO, C. H.; LING, J.)