

文章编号: 1000-5013(2009)04-0420-05

苏云金杆菌 I 型抗癌晶体蛋白的分子建模

袁宇熹, 张同武, 陈智山, 林毅

(华侨大学 化工学院, 福建 泉州 362021)

摘要: 通过同源建模, 得到 I 型抗癌晶体蛋白 Parasporein 1 的初始三维结构. 经分子动力学方法优化后, 利用 Ramachandran Plot 检测和结构匹配等方法对模型进行评价. 结果显示, 结构模型中的键长、键角及二面角的分布合理, 与模板蛋白的主链 α 碳原子的均方根差值为 0.537 592, 在合理范围之内. 此外, 通过分析 Parasporein 1 结构域 I 中的 β loop β 结构片段, 推测出第 2 个蛋白酶处理激活位点的位置在两个芳香族疏水性氨基酸 F167 与 Y168 之间.

关键词: I 型抗癌晶体蛋白; 同源建模; 分子动力学模拟; 蛋白酶; 激活位点

中图分类号: Q 811.4; Q 71

文献标识码: A

苏云金芽孢杆菌(*Bacillus thuringiensis*, Bt) 是一种革兰氏阳性菌, 在形成芽孢的同时, 菌体内会合成各种形状的伴孢晶体蛋白. 已有不少晶体蛋白被证实具有特异的杀虫活性, 但越来越多的研究发现, 自然界中分布更多的是没有杀虫活性的 Bt 菌株^[1]. 近 10 年来, 一些研究者对伴孢晶体蛋白的非杀虫活性进行了研究. 1999 年, Mizuki 等^[2] 首先发现没有杀虫活性的 Bt 菌株 A1190, 其所含的 81 ku 伴孢晶体蛋白经蛋白酶处理激活后对体外培养的人癌细胞(人白血病 T 细胞 MOLT-4、人宫颈癌细胞等)具有杀伤力, 而且不表现溶血作用. 目前, 已发现了 4 种不同类型的具有抗癌作用的伴孢晶体蛋白. 这类蛋白统称为 Parasporein, 具体定义为“Bt 及其相关细菌产生的无溶血作用但可优先杀死癌细胞的伴孢晶体蛋白”^[3]. 苏云金芽孢杆菌 I 型抗癌晶体蛋白(Parasporein 1, 又称 Cry31A a), 由 723 个氨基酸残基构成, 经蛋白酶作用后在 N-端发生两次酶切, 形成由 15 ku 和 56 ku 紧密结合的联合体才具有抗癌活性^[3]. 但是, 对于 Parasporein 1 酶切后的结构尚不清楚. 由于 Parasporein 1 的三维结构未见报道, 尤其是其中的两个酶切位点和其作用机理的阐明更有赖于三维结构的解析. 为此, 本文利用同源建模(Homology Modeling)、分子力学(Molecular Mechanics)、分子动力学模拟(Molecular Dynamics Simulation)等方法, 对 Parasporein 1 的三维结构进行预测及优化, 并推测蛋白酶激活的第 2 个位点.

1 数据库与分析方法

1.1 数据库

(1) 数据库. 国际互联网上的生物信息学数据库(National Center for Biotechnology Information, NCBI), 蛋白质结构数据库(Protein Data Bank, PDB) 等.

(2) 分析软件. Material Studio Modeling 4.0 软件(含 Discover 模块); 可视化软件, 如 Molsoft-ICM, Swiss PdbViewer 3.7, Discover Studio Visualizer 1.5; 比对软件 CLUSTAL W(1.81) Multiple Sequence Alignment.

(3) 服务器. 生物信息学在线服务器, 如同源建模的 Swiss Model(<http://swissmodel.expasy.org/>), 寻找保守序列的 Block Maker(http://blocks.fhcrc.org/blocks/make_blocks.html), 结构域分

收稿日期: 2008-05-17

通信作者: 林毅(1976-), 男, 副教授, 主要从事生物医药与生物农药的研究. E-mail: lyhxm@hqu.edu.cn.

基金项目: 国家自然科学基金资助项目(40601046); 细胞生物学与肿瘤细胞工程教育部重点实验室开放基金资助项目(B0510011)

析的 SMART (<http://smart.embl-heidelberg.de/>), 蛋白质结构预测的 Predict Protein(<http://www.predictprotein.org/>), 跨膜螺旋预测的 SeqFacts(<http://bip.weizmann.ac.il/sqfbin/seqfacts>).

1.2 分析方法

从 NCBI 数据库中搜索得到抗癌晶体蛋白 Parasporin 1 的氨基酸序列, 用 FASTA 程序在 PDB 数据库中搜索同源蛋白并进行 BLASTP 多序列对比, 从结果中寻找具有晶体结构的相似性最高的模板. 将 Parasporin 1 氨基酸序列与模板比对结果提交到 Swiss-Model 服务器进行比对或自动建模, 得到初始的模型. 对初始模型进行结构优化和动力学模拟, 分别用 Material Studio 软件中 Discover 模块的 minimizer 和 dynamics 程序执行.

Discover 是分子力学优化、分子动力学动态模拟程序之一, 是有机小分子、生物大分子结构优化、动力学模拟的应用模块, 可以进行包括最陡下降、共轭梯度、牛顿力学等多种分子力学极小化和分子动力学模拟. 同时, 它也可以进行高温模拟淬火搜索能量稳定点, 所得模型可以通过可视化软件进行观察.

对所得的空间结构进行分析. SMART 用于结构域分析, 可找出各结构域的位置; PredictProtein 对蛋白质结构、功能域(motif)预测; SeqFacts 用于跨膜螺旋预测. 最后, 利用 Ramachandran Plot 检测和结构匹配等方法对模型进行评价.

2 结果与分析

2.1 Parasporin 1 氨基酸序列及比对

Parasporin 1(Cry31Aa) 经蛋白酶作用后, 所形成的由 15 ku 和 56 ku 紧密结合的联合体才具有抗癌活性, 但酶切的位置尚不确定. 虽然酶切位置可以通过生物质谱或蛋白的 N 端测序来获得, 但通过分子建模也能预测其酶切位置, 而且更加简便快速. 因此, 以 Parasporin 1 全长的氨基酸一级结构进行结构同源建模.

从 NCBI 数据库中, 可搜索得到抗癌晶体蛋白 Parasporin 1 的氨基酸序列(登陆号为 AB031065). 然后, 将其提交到 NCBI 数据库, 用 FASTA 程序在 PDB 数据库中搜索同源蛋白, 并进行 BLASTP 多序列比对(比对时参考酶切位点来选取模板), 寻找具有晶体结构的模板. 比对结果发现, 其相似性最高为 22%, PDB 的登陆号为 1ciy(Cry1Aa). 将抗癌晶体蛋白 Parasporin 1 与模板 1ciy 的氨基酸序列进行比对, 结果如图 1 所示.

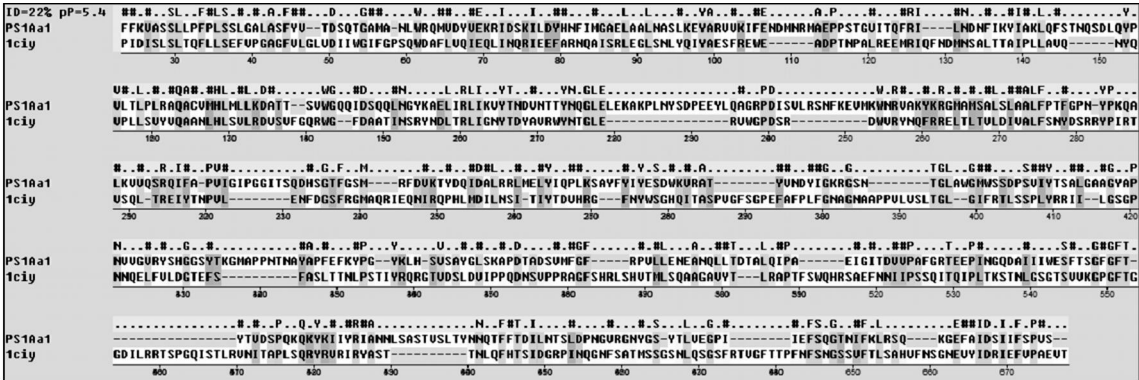


图 1 Parasporin 1 与氨基酸的序列比对

Fig.1 Alignment comparison between Parasporin 1 and the template 1ciy

2.2 Parasporin 1 的同源建模及模型的优化

目前, 由蛋白质的一级序列预测空间结构的诸多方法中, 同源蛋白结构建模技术的可靠性已被广泛接受, 是应用较多的结构预测技术. 三维结构建模是利用 Swiss-Model 服务器(Homology Modeling)进行的. 其主要步骤如下: 提交抗癌晶体蛋白 Parasporin 1, 并与模板 1ciy 的氨基酸序列比对(CLUSTAL W(1.81) Multiple Sequence Alignment); 将比对结果提交到 Swiss-Model 服务器, 进行比对建模; 同源建模返回的结构(pdb 格式) 在 ICM 中可视化, 即可得到 Parasporin 1 初始的三维结构. 但由于 Parasporin 1 与 Cry46Aa1 的相似性较低, 所得到的模型不够准确, 因此, 需要对 Parasporin 1 初始的三维结

构进行修正.

对于一个模型来说, 体系的势能面上有许多极小值点. 体系势函数对这些点的坐标的一阶微商等于零, 二阶微商大于零(或 Hessian 矩阵本征值大于零). 满足这样条件的一个点对应体系一个可能的平衡构型. 使体系具有最低能量的极小值点叫全局极小值, 其他的极小值点叫局部极小值.

分子力学中的优化, 就是确定这些极小值点所对应的体系的几何构型. 优化通常包括两个步骤: (1) 依据给定的体系的初始构型和所选力场, 建立描述该体系的势函数(目标函数); (2) 根据一定的优化算法, 逐步改变体系的坐标, 使体系的能量逐步降低, 直到得到具有最低能量的构型. 这一步骤常需要上千次的迭代, 迭代次数(n)的多少主要决定于优化算法的性质、目标函数的形式, 以及体系的大小等因素.

分别用 Material Studio 软件中 Discover 模块的 minimizer 和 dynamics 程序, 对同源建模的 Paraspurin 1 初始三维结构进行结构优化和动力学模拟. 在 CVFF(Consistent Valence Force Field) 力场下, 先用最陡下降法进行 100 步的结构松弛计算, 分子势能下降趋势已趋于缓和; 接着, 采用 200 步的共轭梯度法进行结构优化, 分子势能下降趋势趋于平稳, 如图 2 所示.

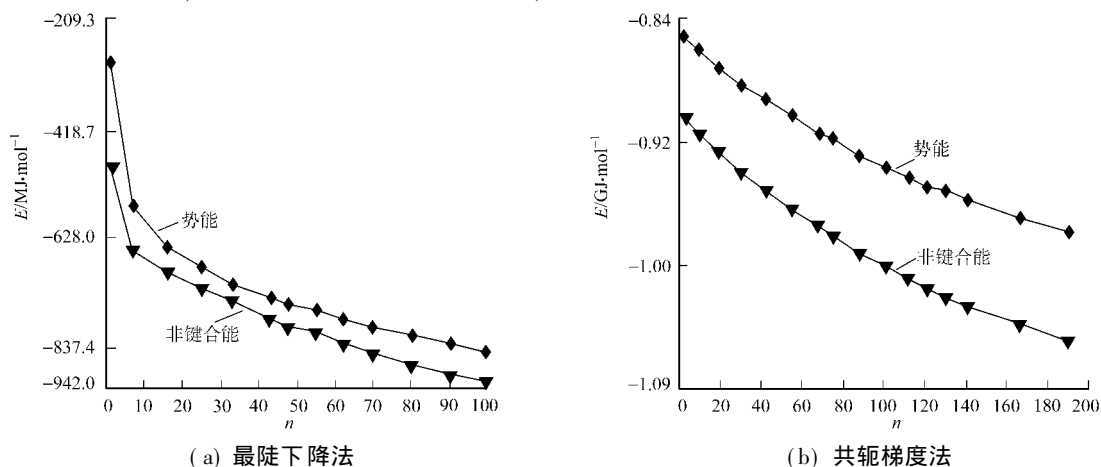


图 2 Paraspurin 1 初始三维结构的优化和动力学模拟

Fig. 2 Optimization and dynamic simulation for the initial structure of Paraspurin 1

由于分子动力学优化不能解决局部势垒问题, 因此, 为了克服局域势垒, 需对建模松弛后的结构进行了 50 ps 的常温(300 K, Andersen 控温, NVT 体系)分子动力学计算. 每隔 10 fs 收集一个构象, 选定能量最低的构象为最终的建模结果. 在分子动力学模拟过程中, 体系势能和键能随时间的变化, 如图 3 所示. 对于监控系统, 是否达到平衡是很重要的. 从图 3 可以看出, 在最初的 30 ps 后, 体系的能量达到平衡. 即在其余 20 ps 的模拟中, 体系已经达到了平衡.

图 4 为 Paraspurin 1 的稳定构型. 经分析, 三级结构含 14 个 α 螺旋, 28 个 β 折叠, 39 个 loop 环, 这与二级结构预测的结果基本一致. 修正后的总能量为 $-9.835\,271\text{ MJ}\cdot\text{mol}^{-1}$.

经跨膜螺旋预测, 该蛋白中可能存在两种跨膜螺旋: 方向由内向外(AA: 149~168)、方向由外向内(AA: 141~164). 与其他杀虫晶体蛋白的典型结构域相似, Paraspurin 1 的空间结构是由 Domain I(147~396), Domain II(397~587) 和 Domain III(588~723) 3 个结构域组成.

2.3 Paraspurin 1 模型的评价

在同源建模中, Ramachandran Plot 检测可评估模型结构的立体化学参数, 如分子中的键长、键角, 以及二面角的分布. 尤其是主链的构型中, φ, ψ 二面角分布的情形可用产生的 Ramachandran Plot 检测. 其中, Residues in most favoured regions 是结构性质中比较主要的参考指标.

一般预期模型二面角在该区域的分布大于 90% 为正常分布. 但是, 由于 Gly 没有侧链, 对二面角的形成大小没有空间上的阻碍作用, 且模板结构二面角在此区域的分布也可能低于 90%. 所以, 新模拟的结构的数值必须与模板结构的数值互比较分析, 然后, 才考虑其中最接近模板结构的模拟结构. Spd-view 可以用来评估主链构型中 φ, ψ 二面角的分布, 即 Ramachandran Plot 检测.

如图 5 所示, 模拟的 Paraspurin 1 结构中 91.2% 的 φ, ψ 二面角落在核心区, 而具有 X-ray 晶体结构 1civ 主链的 φ, ψ 二面角有 94.5% 落在核心区. 对 Paraspurin 1 和 1civ 进行结构匹配的检验, 两个蛋

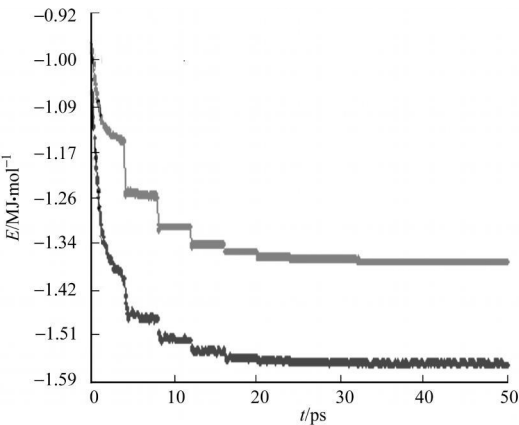


图 3 体系势能和非键能量随模拟时间的变化
Fig. 3 Potential energy and nonbond energy changes with simulation time

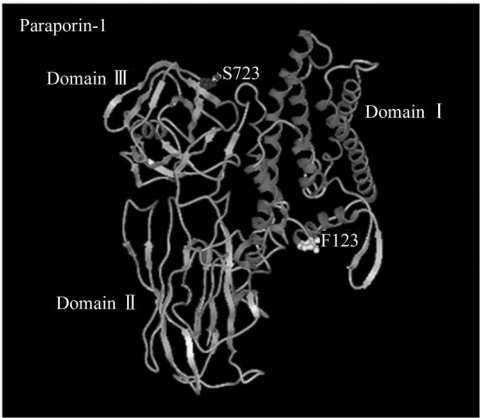


图 4 同源模建优化后 Paraspornir 1 三维结构
Fig. 4 Homology modeling optimization Paraspornir 1 three dimensional structure

白质的主链 α 碳原子的均方根差值为 0.537 592, 在合理范围之内, 二者的叠合模型如图 6 所示. 以上的评测结果表明, Paraspornir 1 的模建结构是合理可信的.

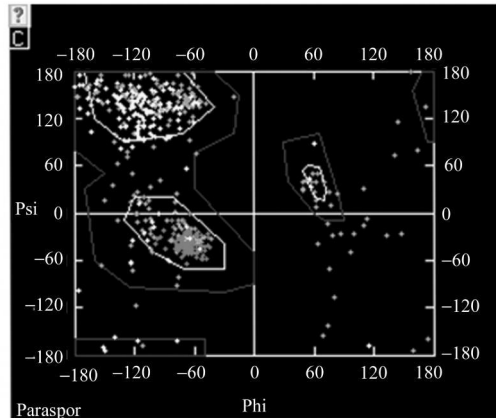


图 5 Paraspornir 1 主链的 Φ, Ψ 二面角于 Ramachandran Plot 上的分布
Fig. 5 Φ, Ψ dihedral angle of Paraspornir 1 main chain in the distribution Ramachandran plot

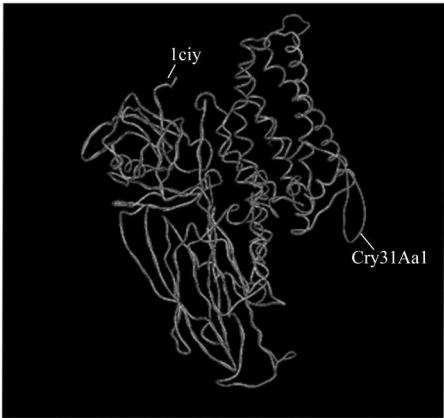


图 6 Paraspornir 1 和 Icy 主链 α 碳原子叠合的立体示意图
Fig. 6 Paraspornir 1 and Icy main chain of carbon atoms in a three dimensional matrix composite

3 讨论

文[4]已经通过同源模建, 成功获得了 Paraspornir 2 的空间结构模型, 而本研究模建了另一类抗癌晶体蛋白 Paraspornir 1 的空间结构, 两者在空间结构上存在很大差异. 研究表明, 其对癌细胞的作用机理也有很大的差异. 即 Paraspornir 1 能快速增加钙离子从细胞外进入 Hela 细胞的流入量^[5], 而 Paraspornir 2 以寡聚体形式镶嵌在细胞膜内, 与某种受体蛋白结合, 引起膜通透性增加, 使细胞质外泄而导致癌细胞死亡^[6].

在 Cry 晶体蛋白中, Domain I 为毒性区. 分析 Paraspornir 1 的结构(图 7), 发现在 Domain I 中除了与 Cry 蛋白类似的 7 个 α 螺旋外, 在该结构域的表面(靠近 N 端)还存在着一个不稳定的 β loop β 结构片段, 其氨基酸序列为 (167) FYVTDSQTGA (176). 该结构中部含有 5 个亲水氨基酸, 而两端含有较多的疏水性氨基

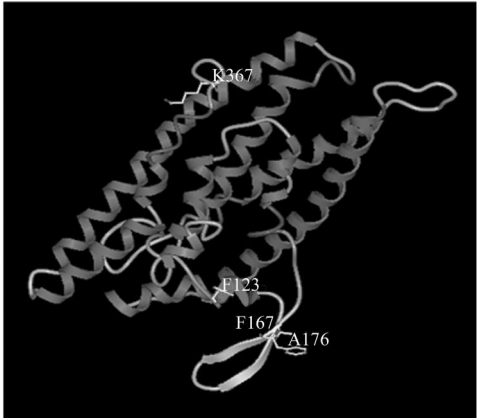


图 7 Paraspornir 1 的 Domain I 示意图
Fig. 7 Structure of Paraspornir 1 Domain I

酸,易受到蛋白酶 K 的作用.

由 Predict Protein 搜索得到 Parasporin 1 蛋白的结构和功能位点可知,在该结构片段的两端分别存在 1 个 N-十四烷基化位点,氨基酸序列分别为(162) GALASF(167), (175) GAMANL(180);而 N-十四烷基化位点是和膜结合的位点.由此推测,第 2 个酶切位点的位置在两个芳香族疏水性氨基酸 F167 与 Y168 之间.

与 Cry 蛋白相似, Domain I 的长疏水和亲水脂螺旋可能参与细胞膜上孔洞的形成, (175) GAMANL(180) 与 Domain I 的螺旋相连嵌入膜内并与膜结合,而(162) GALASF(167) 插入膜内与膜结合并与其 N 端形成 15 ku 的结构.该结构与在细胞膜上形成孔洞的 56 ku 的结构紧密结合形成联合体,所形成 15 ku 的结构可能增强了细胞膜孔洞的形成并赋予某一结构上的特性,最终导致 Parasporin 1 蛋白能快速增加钙离子从细胞外进入癌细胞的流入量.

参考文献:

- [1] ROH J Y, PARK H W, JIN B R, et al. Characterization of novel non toxic *Bacillus thuringiensis* isolates from Korea [J]. Lett Appl Microbiol, 1996, 23(4): 249-252.
- [2] MIZUKI E, OHBA M, AKAO T, et al. Unique activity associated with non insecticidal *Bacillus thuringiensis* parasporal inclusions: In vitro cell killing action on human cancer cells[J]. J Appl Microbiol, 1999, 86(3): 477-486.
- [3] KITADA S, ABE Y, ITO A, et al. Molecular identification and cytotoxic action of parasporin, a protein group of novel crystal toxins targeting human cancer cells[C] // Biotechnology of *Bacillus thuringiensis* and Its Environmental Impact Victoria B C: Minireviews, 2005: 1-8.
- [4] 林 毅, 张同武, 陈智山, 等. 抗癌晶体蛋白 Parasporin 2 空间结构的研究[J]. 激光生物学报, 2008, 17(2): 181-185
- [5] KATAYAMA H, YOKOTA H, AKAO T, et al. Parasporin 1, a novel cytotoxic protein to human cells from non insecticidal parasporal inclusions of *Bacillus thuringiensis* [J]. J Biochem, 2005, 137(1): 17-25.
- [6] KITADA S, ABE Y, SHIMADA H, et al. Cytotoxic actions of Parasporin 2, an anti tumor crystal toxin from *Bacillus thuringiensis* [J]. Journal of Biological Chemistry, 2006, 281(36): 26350-26360.

Molecular Modeling of Cancer Cell-Killing Crystal Protein Parasporin 1 from *Bacillus thuringiensis*

YUAN Yur-xi, ZHANG Tong-wu,
CHEN Zhi-shan, LIN Yi

(College of Chemical Engineering, Huaqiao University, Quanzhou 362021, China)

Abstract: The structure of Parasporin 1, the first cancer cell-killing crystal protein from *Bacillus thuringiensis*, was constructed using homology modeling, and was optimized by molecular mechanics method. The simulated structure was evaluated using Ramachandran Plot and structural matching. The distributions of bond length, bond angle, and dihedral angle were rational. The RMSD (root mean square deviation) value of α carbon atom on the main chain between Parasporin 1 and model protein was 0.537-0.592, which was within reasonable bounds. The analysis of β loop β motif in Domain I suggest that the second site digested by proteinase K were located between F167 and Y168.

Keywords: Parasporin 1; homology modeling; molecular dynamics simulation; proteinase K-digested sites

(责任编辑: 黄晓楠 英文审校: 陈国华)