

doi: 10.3969/j.issn.0490-6756.2020.06.025

Curcin 和 curcin C 与腺嘌呤互作位点的比较分析

吴朋, 邓聿杉, 陈放, 徐莺

(四川大学生命科学学院 生物资源与生态环境教育部重点实验室, 成都 610065)

摘要: 麻疯树核糖体失活蛋白 curcin 和 curcin C 均具有抗肿瘤活性,但后者的活性水平显著高于前者,为了探索造成这一差异的结构基础,本文采用在线的同源建模预测软件 SWISS-MODEL,对两种麻疯树核糖体失活蛋白 curcin、curcin C 的三维结构模型进行预测,采用 SYBYL 对预测模型进行能量最小化优化,采用 PROCHECK、VERIFY 3D 和 ERRAT 软件对优化前后的模型进行质量评估,随后采用 AutoDock 软件将预测的模型与腺嘌呤进行分子对接分析.结果显示,两种蛋白采用与 Ricin A 类似的方式与腺嘌呤相互作用,但在相互作用的氨基酸残基种类、数量以及形成的氢键和疏水相互作用上还是存在差异,其中,curcin C 与腺嘌呤具有最强的结合能力,curcin 则最低.

关键词: 麻疯树; 核糖体失活蛋白; 结构预测; 分子对接; 腺嘌呤

中图分类号: Q518.3 **文献标识码:** A **文章编号:** 0490-6756(2020)06-1186-07

Comparative analysis of the interaction sites of curcin, curcin C with adenine

WU Peng, DENG Yu-Shan, CHEN Fang, XU Ying

(Key Laboratory of Bio-Resources and Eco-Environment Ministry of Education,
College of Life Sciences, Sichuan University, Chengdu 610065, China)

Abstract: The ribosome-inactivating proteins of *Jatropha curcas* curcin and curcin C have antitumor activity, but the activity level of the latter is significantly higher than the former. In order to explore the structural basis for this difference, the online homology modeling prediction software SWISSMODEL was used to predict the three-dimensional structural models of the ribosomal inactivating proteins curcin and curcin C. SYBYL was used to optimize the energy of the prediction model, and PROCHECK, VERIFY 3D and ERRAT software were used to evaluate the quality of the model before and after the optimization, and then AutoDock software was used to analyze the predicted model and adenine for molecular docking. The results show that the two proteins interacted with adenine in a similar way to Ricin A, but there were differences in the types and number of amino acid residues that interact, as well as the hydrogen bonds and hydrophobic interactions formed. Among them, curcin C had the strongest binding ability with adenine, and curcin had the lowest.

Keywords: *Jatropha curcas*; Ribosome-inactivating protein; Structure prediction; Molecular docking; Adenine

收稿日期: 2020-03-09

基金项目: 国家自然科学基金(31870315)

作者简介: 吴朋(1994-), 男, 汉族, 贵州遵义人, 硕士研究生, 研究方向为植物蛋白抗肿瘤. E-mail: 1347357302@qq.com

通讯作者: 徐莺. E-mail: xuying@scu.edu.cn

1 引言

核糖体失活蛋白(Ribosome inactivating proteins, RIPs)能通过其 rRNA N-糖苷酶活性去除核糖体 28S rRNA 上特异位置腺嘌呤的糖苷键从而抑制蛋白质合成^[1]. 因其具有抗肿瘤、抗病毒和抗真菌等多种生物学活性而受到关注^[2].

Curcin 和 curcin C 均是从大戟科麻疯树属植物麻疯树(*Jatropha curcas*)中分离得到的核糖体失活蛋白. 1976 年, Stripe 首次从麻疯树胚乳中分离出 curcin^[3], Barbieri 等认为其属于 I 型 RIPs^[4]. 2002 年, 林娟^[5]等人确定其分子量为 28.2 kDa, 具有 rRNA N-糖苷酶活性. 2017 年, Zhang 等人^[6]从麻疯树完全伸展开的子叶中分离纯化出一种新型 I 型核糖体失活蛋白 curcin C. 研究表明, curcin 和 curcin C 的氨基酸序列同源性达到 81.88%, 但抗肿瘤活性存在较大的效率差异(例如后者对骨肉瘤细胞的细胞毒性可达前者的 10 倍)^[6-7], 这表明两者的催化效率或作用机制可能存在差异. 我们注意到两种核糖体失活蛋白的蛋白体外翻译抑制的 IC₅₀ 分别为 0.19 和 0.05 nM, 暗示它们的 rRNA N-糖苷酶活性可能同样也存在差异^[6-7]. 由于造成酶活性差异的主要原因是蛋白的底物结合和催化中心, RIPs 在完成其 rRNA N-糖苷酶活性时, 首先需要与底物分子特异位点的腺嘌呤结合. 两种蛋白是否因为在结合腺嘌呤上存在差异, 因而进一步导致活性的不同目前还不清楚. 为了回答这一问题, 本文利用生物信息学方法, 首先预测出 curcin、curcin C 蛋白的三维结构模型, 并对模型质量进行优化和评估, 之后采用分子对接方法探究 curcin、curcin C 与腺嘌呤相互作用位点的比较分析, 为阐明其活性差异的潜在机制提供一种的线索.

2 材料与方法

2.1 材料

麻疯树核糖体失活蛋白 curcin (GenBank 登记号: AAL58089.1) 以及 curcin C (GenBank 登记号: ABZ04128.1) 的氨基酸序列由 NCBI 蛋白质数据库 (<http://www.ncbi.nlm.nih.gov>) 检索获得.

2.2 方法

2.2.1 序列检索与三维结构预测 利用在线预测软件 SWISSMODEL^[8-12] (<https://swissmodel.expasy.org/>) 预测 curcin、curcin C 的三维模型.

2.2.2 模型优化 采用 SYBYL 软件, 选用 AMBER7 FF99 力场和 AMBER 电荷对 curcin 和 curcin C 模型进行能量最小化处理.

2.2.3 模型评估 为了进一步确定模型的有效性, 我们采用 PROCHECK^[13]、VERIFY 3D^[14-15]、ERRAT^[16] 对模型进行评估.

2.2.4 分子对接 分子对接实验使用 AutoDock 4.2^[17], 以鉴定 curcin、curcin C 与腺嘌呤可能的相互作用残基. 对蛋白质进行预处理: 加氢、计算电荷、分配原子 AD4 类型; 对腺嘌呤小分子做加氢以及计算电荷处理. 设置 GridBox 大小在 X、Y、Z 轴分别为 60、64、50 Å, 间距为 0.375 Å^[18]. 蛋白质为刚性结构, 小分子配体为柔性结构. 选择遗传算法 (Genetic Algorithm), 寻找蛋白质受体上配体合适的结合位置.

3 结果与讨论

3.1 Curcin 和 curcin C 的三维建模和初步评估

通过序列和结构比对, 寻找一个与 curcin、curcin C 同源的由实验测定的蛋白质结构, 本次建模选择的模型 PDB id 为 1IFS, 为蓖麻 ricinA 链和游离腺嘌呤形成的复合体^[19]. 其中 ricin A 链与 curcin、curcin C 的氨基酸序列相似度分别为 36.48%、34.69%, 均大于 30%, 满足了 SWISSMODEL 三维结构模型预测的需求. 在此基础上, 预测出 curcin、curcin C 的三维结构示意图(图 1A、B). 为了判断所预测结构的质量, 采取 QMEAN 软件对上述预测结果进行检验(图 1C、D). 分值分别是 -2.97、-2.87, 表明 curcin、curcin C 模型质量合格. 其中 curcin 模型具有 10 个 α -螺旋, 10 个 β -折叠; curcin C 模型具有 12 个 α -螺旋, 8 个 β -折叠. 导致 α -螺旋差异的两处位置分别是 curcin 的 Ala56-Ala57 对应 curcin C 中为 Thr56-Thr57、Pro134 对应为 Ser134. 造成 β -折叠差异的两处位置分别是 curcin 的 Ser78 对应 curcin C 的 Leu78、Glu81 对应为 Lys81.

3.2 Curcin 和 curcin C 模型的评估

蛋白质三维结构的稳定性主要取决于原子的最小能量排列, 通过能量最小化的优化可以使预测的蛋白质结构更接近自然状态^[18]. 将上述预测出的模型经 SYBYL 软件能量最小化优化后, curcin、curcin C 模型总能量分别从 -3 033.609 kcal/mol 和 -3 072.181 kcal/mol 降低至 -5 073.993 kcal/mol 和 -5 316.474 kcal/mol. 为了进一步

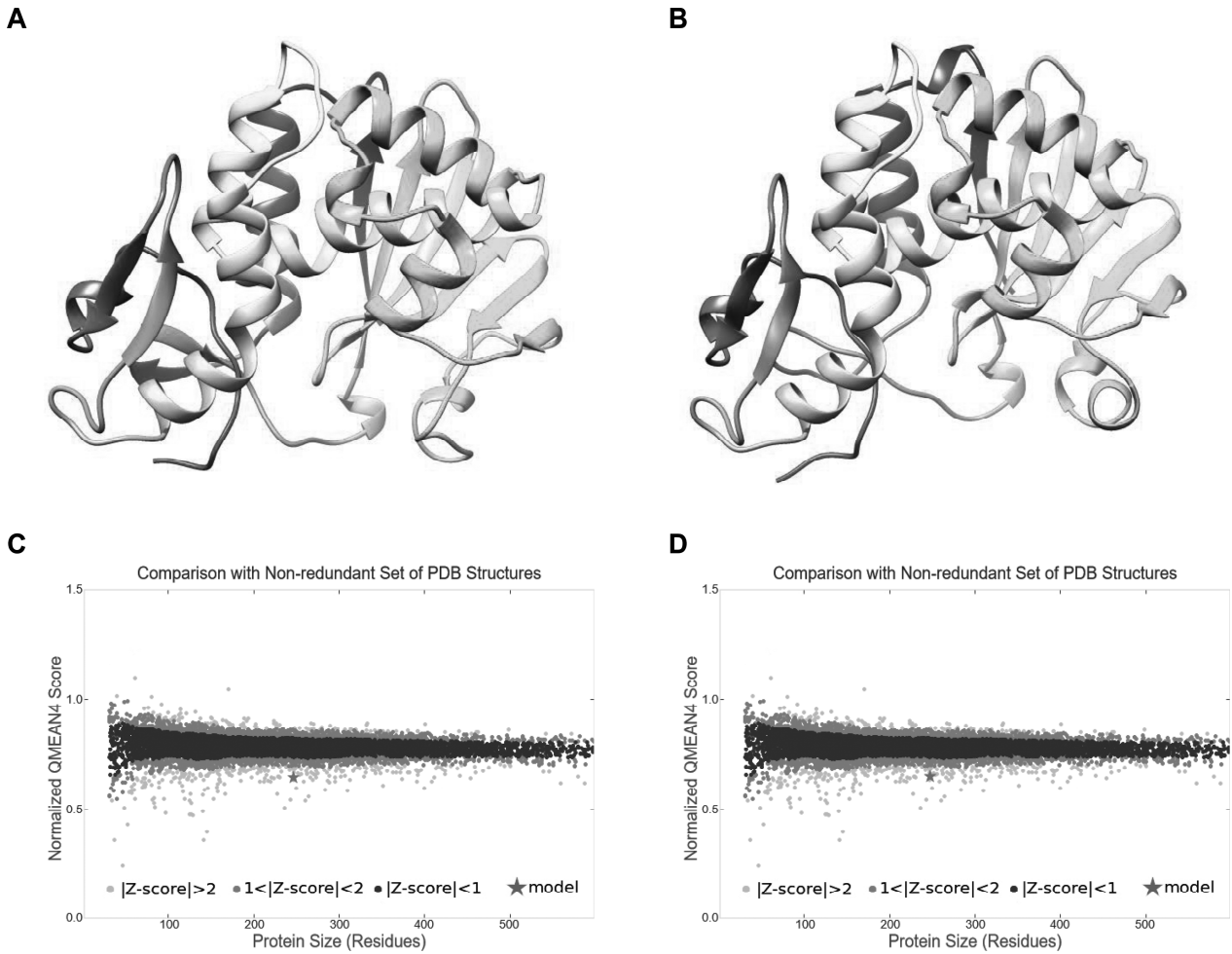


图 1 基于 SWISS MODEL 预测的 curcin 和 curcin C 三维结构示意图

(A, B) Curcin, Curcin C 三维模型; (C, D) QMEAN 分别评估 curcin, curcin C 模型.

Fig. 1 Schematic diagram of curcin and curcin C 3D structure predicted by SWISS MODEL

(A, B) Curcin, curcin C three-dimensional models; (C, D) Curcin and curcin C models were evaluated by QMEAN respectively.

验证优化前后 curcin、curcin C 模型的准确性, 采取 PROCHECK、VERIFY 3D、ERRAT 三个软件对模型质量进行评估. 结果如下: 应用 PROCHECK 的评估表明, 优化前 curcin、curcin C 的允许区域和额外允许区域百分比之和分别为 98.1% (图 2A) 和 97.8% (图 2B), 优化后分别为 97.7% (图 2C) 和 97.8% (图 2D). 可以发现优化对于模型的分值影响不大, 所有的分值均大于 90%. 对 curcin C 来说, 没有发生变化; 对于 curcin 来说, 分值甚至减小. VERIFY 3D 评估表明, 优化使得两个

模型的分值都得到提高. 前 curcin、curcin C 分别有 88.21% (图 2E)、80.97% (图 2F) 的残基拥有大于 0.2 的 3D/1D 值, 优化后分别为 94.31% (图 2G)、86.23% (图 2H). ERRAT 的评估结果类似于 VERIFY 3D, 优化前 curcin、curcin C 的得分分别是 96.581 (图 2I)、97.046 (图 2J), 优化后分别为 98.312 (图 2K)、90.377 (图 2L), 即优化使得模型的质量提高, 不过值得指出的是优化前后的分值均大于 85. 综上分析结果提示优化前后的 curcin 和 curcin C 模型均具有较好的质量, 都可用于下一步分析.

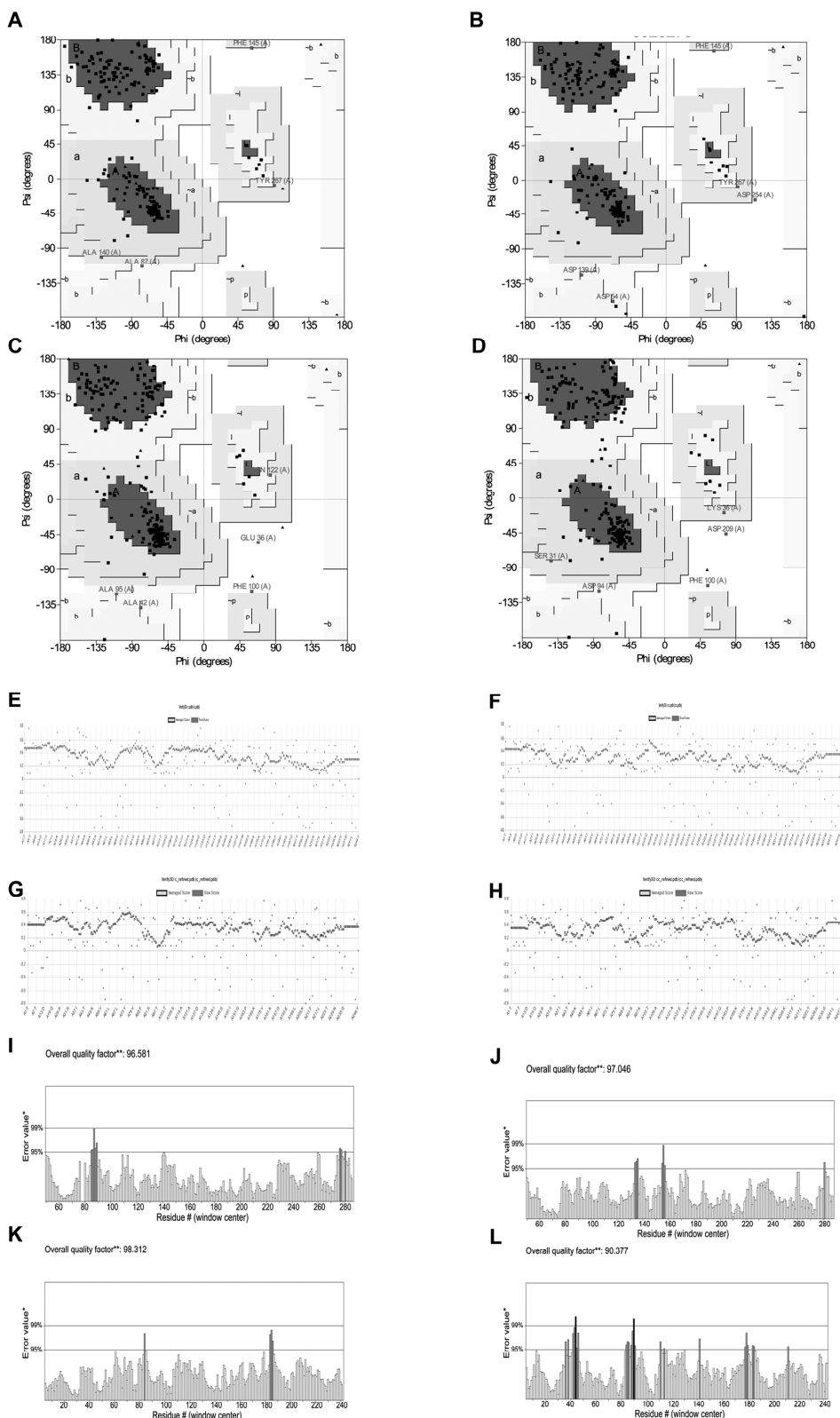


图 2 Curcin, curcin C 模型的质量评估

(A, B, C, D) PROCHECK 评估优化前后的 curcin, curcin C 模型; (E, F, G, H) VERIFY 3D 评估优化前后的 curcin, curcin C 模型; (I, J, K, L) ERRAT 评估优化前后的 curcin, curcin C 模型。

Fig. 2 Quality evaluation of curcin and curcin C models

(A, B, C, D) Curcin and curcin C models before and after optimization were evaluated by PROCHECK; (E, F, G, H) Curcin and curcin C models before and after optimization were evaluated by VERIFY 3D; (I, J, K, L) The curcin and curcin C models before and after optimization were evaluated by ERRAT.

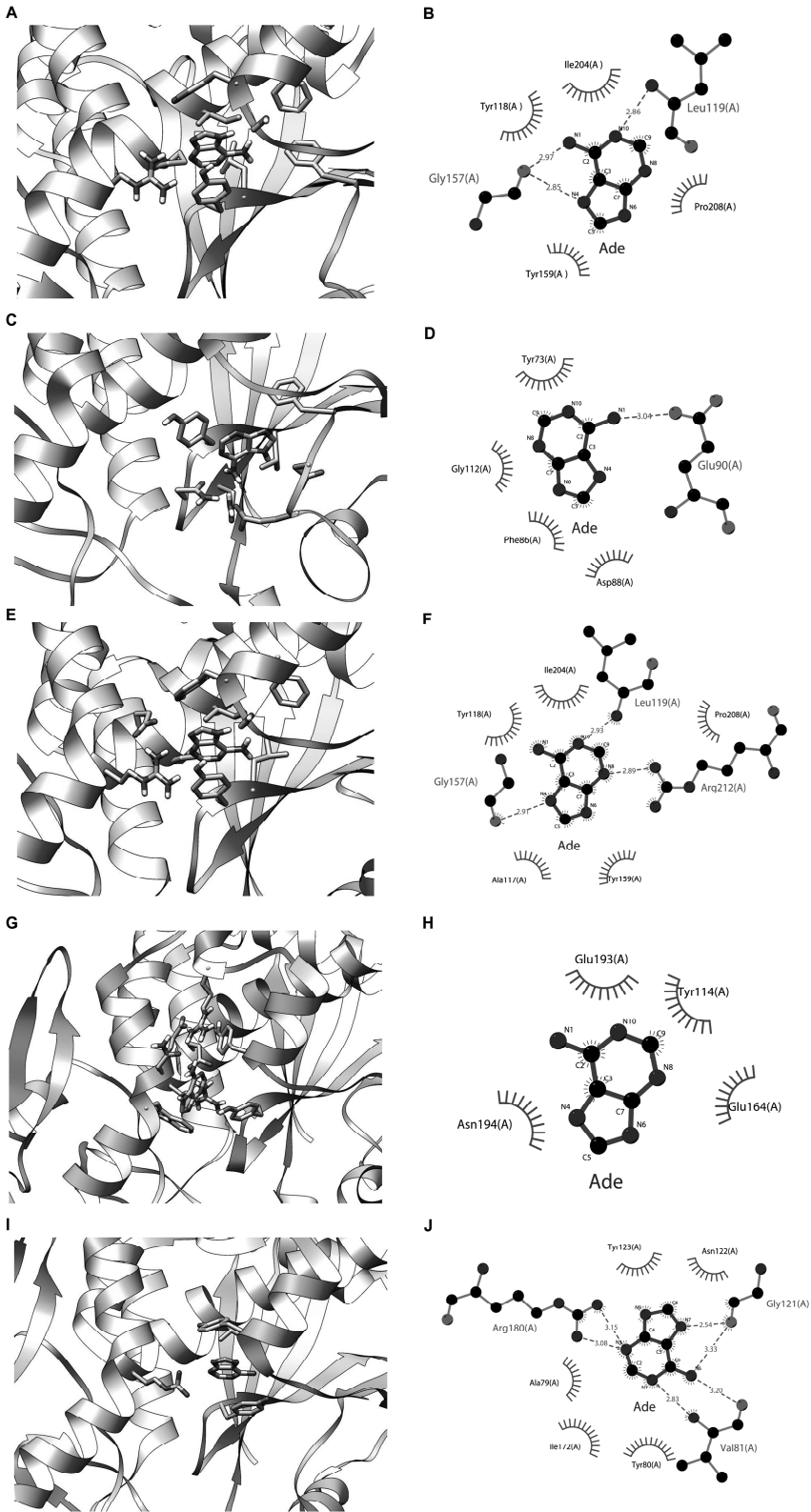


图 3 Curcins 和 curcins C 与腺嘌呤相互作用示意图

(A, B) 优化前 curcins 与腺嘌呤相互作用示意图; (C, D) 优化后 curcins 与腺嘌呤相互作用示意图; (E, F) 优化前 curcins C 与腺嘌呤相互作用示意图; (G, H) 优化后 curcins C 与腺嘌呤相互作用示意图; (I, J) Ricin A 与腺嘌呤相互作用示意图。

Fig. 3 Schematic of the interaction between curcins, curcins C and adenine

(A, B) Schematic diagram of interaction between curcins and adenine before optimization; (C, D) Schematic diagram of interaction of curcins and adenine after optimization; (E, F) Schematic diagram of interaction of curcins C and adenine before optimization; (G, H) Schematic diagram of the interaction between curcins C and adenine after optimization; (I, J) Schematic diagram of the interaction between Ricin A and adenine.

3.3 与腺嘌呤分子的对接研究

为了预测和比较 curcin 和 curcin C 与腺嘌呤的相互作用位点,采用 AutoDock 4.2 软件将腺嘌呤分别与优化前后的 curcin 和 curcin C 模型进行对接分析,选取结合能最低的对接结果并使用 UCSF Chimera v. 1. 12. 1^[20] 和 LigPlot +^[21] 软件进行可视化(图 3). 可以看见优化前后的 curcin 模型氨基酸残基与腺嘌呤相互作用的结合能分别为 -7.52kJ/Mol 和 -5.32kJ/Mol . 优化前的 curcin 模型与腺嘌呤分子共形成 3 个氢键,其中 Leu119 与腺嘌呤的 N10 形成氢键,Gly157 分别与腺嘌呤的 N1、N4 形成氢键,Tyr118、Tyr159、Ile204、Pro208 则与腺嘌呤形成疏水相互作用(图 3A, 3B). 优化后的 curcin 模型则是通过 Glu90 与腺嘌呤的 N1 形成一个氢键, Try73、Phe86、Asp88、Gly112 与腺嘌呤具有疏水相互作用(图 3C, 3D). 优化前后的 curcin C 模型氨基酸残基与腺嘌呤相互作用的结合能分别为 -8.2kJ/Mol 和 -5.29kJ/Mol . 优化前 curcin C 与腺嘌呤也形成了 3 个氢键,分别为 Leu119 与腺嘌呤的 N10、Gly157 与 N4、Arg212 与 N8 之间,Ala117、Tyr118、Tyr159、Ile204、Pro208 共 5 个氨基酸残基则参与了与腺嘌呤的疏水相互作用(图 3E, 3F). 优化后的 curcin C 模型的氨基酸残基与腺嘌呤未形成氢键,仅 Tyr114、Glu164、Glu193 以及 Asn194 共 4 个氨基酸残基参与了与腺嘌呤的疏水相互作用(图 3G, 3H).

在上述四个对接组合中,可以看出,结合能最低的组合是优化前的 curcin C 模型与腺嘌呤组合,该组合中分别有 3 个氨基酸残基参与了氢键的形成,5 个氨基酸残基参与了与腺嘌呤的疏水互作,从数量上也是最多的. 并且,我们还注意到,参与腺嘌呤互作的氨基酸残基及其位置与 ricin A 非常类似,后者也是有 3 个氨基酸残基(Val81、Gly121、Arg180)参与了氢键形成,5 个氨基酸残基(Ala79、Tyr80、Asn122、Tyr123、Ile172)参与了疏水互作(图 3I, 3J). 唯一的区别是 Pro208 替代了 Ser158 (ricin A 中为 Asn122)参与疏水作用. 同样地,我们发现优化前的 curcin 模型与腺嘌呤组合的结合能也低于另外两个优化后模型组合,参与的氨基酸数量虽然相对较少,但也非常类似于 ricin A. 前期的研究显示,在 Ricin A 链中 Tyr80、Tyr123、Glu177、Arg180、Trp211 是 rRNA N-糖苷酶活性的关键活性位点,而且这些位点在 RIPs 中高度保

守. 因此,优化前的模型结构虽然分值相对较低,但可能更接近真实状态. 造成这一现象的原因则可能是在能量最小化优化时,没有考虑到该模型是在有腺嘌呤参与时形成的. 当然这也意味着优化前的模型并非两个核糖体失活蛋白游离时的状态. 至于 curcin 和 curcin C 与腺嘌呤的相互作用存在差异,则提示它们的 rRNA N-糖苷酶活性可能存在差异,至于这种差异是导致它们的抗肿瘤活性差异的原因需要进一步的实验验证.

参考文献:

- [1] Fabbrini M S, Katayama M, Nakase I, *et al.* Plant ribosome-inactivating proteins: progresses, challenges and biotechnological applications (and a few digressions)[J]. *Toxins (Basel)*, 2017, 9: 314.
- [2] Puri M, Kaur I, Perugini M A, *et al.* Ribosome-inactivating proteins: current status and biomedical applications [J]. *Drug Discov Today*, 2012, 17: 774.
- [3] Stirpe F, Pession-Brizzi A, Lorenzoni E, *et al.* Studies on the proteins from the seeds of *Croton tiglium* and of *Jatropha curcas*. Toxic properties and inhibition of protein synthesis in vitro [J]. *Biochem J*, 1976, 156: 1.
- [4] Barbieri L, Battelli M G, Stirpe F. Ribosome-inactivating proteins from plants[J]. *Biochimica Et biophysica Acta*, 1993, 1154: 237.
- [5] 林娟, 颜钊, 唐琳, 等. 麻疯树核糖体失活蛋白的分离纯化和作用机制研究[J]. *高技术通讯*, 2002, 12: 36.
- [6] Zhang Y, Yang Q, Li C, *et al.* Curcin C, a novel type I ribosome-inactivating protein from the post-germinating cotyledons of *Jatropha curcas*[J]. *Amino Acids*, 2017, 49: 1619.
- [7] Lin J, Yan F, Tang L, *et al.* Antitumor effects of curcin from seeds of *Jatropha curcas* [J]. *Acta Pharmacologica Sinica*, 2003, 24: 241.
- [8] Waterhouse A, Bertoni M, Bienert S, *et al.* SWISS-MODEL: homology modelling of protein structures and complexes[J]. *Nucleic Acids Res*, 2018, 46: W296.
- [9] Bienert S, Waterhouse A, De Beer T A, *et al.* The SWISS-MODEL repository-new features and functionality [J]. *Nucleic Acids Res*, 2017, 45: D313.
- [10] Bertoni M, Kiefer F, Biasini M, *et al.* Modeling protein quaternary structure of homo-and hetero-oligomers beyond binary interactions by homology

- [J]. *Sci Rep*, 2017, 7: 10480.
- [11] Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models [J]. *Bioinformatics*, 2011, 27: 343.
- [12] Guex N, Peitsch M C, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer; a historical perspective [J]. *Electrophoresis*, 2009, 30 Suppl 1: S162.
- [13] Laskowski R A, Macarthur M W, Moss D S, *et al.* Procheck - a program to check the stereochemical quality of protein structures [J]. *J Appl Cryst*, 1993, 26: 283.
- [14] Luthy R, Bowie J U, Eisenberg D. Assessment of protein models with three-dimensional profiles [J]. *Nature*, 1992, 356: 83.
- [15] Bowie J U, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure [J]. *Science*, 1991, 253: 164.
- [16] Colovos C, Yeates T O. Verification of protein structures: patterns of nonbonded atomic interactions [J]. *Protein Sci*, 1993, 2: 1511.
- [17] 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]. *J Comput Chem*, 1998, 19: 1639.
- [18] Srivastava M, Gupta S K, Abhilash P C, *et al.* Structure prediction and binding sites analysis of curcin protein of *Jatropha curcas* using computational approaches [J]. *J Mol Model*, 2012, 18: 2971.
- [19] Weston S A, Tucker A D, Thatcher D R, *et al.* X-ray structure of recombinant ricin A-chain at 1.8 Å resolution [J]. *J Molecul Biolo*, 1994, 244: 410.
- [20] Pettersen E F, Goddard T D, Huang C C, *et al.* UCSF Chimera—a visualization system for exploratory research and analysis [J]. *J Comput Chem*, 2004, 25: 1605.
- [21] Wallace A C, Laskowski R A, Thornton J M. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions [J]. *Protein Eng*, 1995, 8: 127.

引用本文格式:

中文: 吴朋, 邓聿杉, 陈放, 等. Curcin 和 curcin C 与腺嘌呤互作位点的比较分析 [J]. *四川大学学报: 自然科学版*, 2020, 57: 1186.

英文: Wu P, Deng Y S, Chen F, *et al.* Comparative analysis of the interaction sites of curcin, curcin C with adenine [J]. *J Sichuan Univ: Nat Sci Ed*, 2020, 57: 1186.