• 研究论文 •

分子动力学模拟极端嗜热核糖结合蛋白的热力学稳定性

冯献礼 赵 熹 于 辉 王乙博 孙铁东 黄旭日*

(吉林大学理论化学研究所理论化学计算国家重点实验室 长春 130023)

摘要 采用分子动力学模拟方法研究极端嗜热性核糖结合蛋白(tteRBP)的嗜热机理. 在常温(300 K)和最佳活性温度(375 K)时, 分别对 tteRBP 分子进行动力学模拟, 结果表明, 整体分子均保持结构稳定, 但分子内部的协调运动不同. 在 375 K 时蛋白整体柔性显著提高, 使分子能够局部调整构象以适应极端高温. 蛋白结构变化的分析也确认了高温时构象局部微调对蛋白极端高温稳定性的关键作用.

关键词 tteRBP(Thermoanaerobacter tengcongensis ribose binding protein); 分子动力学; 分子内部协调运动; 柔性; 极端嗜热

Molecular Dynamics of an Extremely Thermophilic Ribose Binding Protein

Feng, Xianli Zhao, Xi Yu, Hui Wang, Yibo Sun, Tiedong Huang, Xuri*

(State Key Laboratory of Theoretical and Computational Chemistry, Institute of Theoretical Chemistry, Jilin University, Changchun 130023)

Abstract Molecular dynamics simulations for extremely thermophilic protein thermoanaerobacter tengcongensis ribose binding protein (tteRBP) were performed to investigate the thermophilic mechanism of the protein. The comparative analysis of molecular trajectories of room temperature (300 K) and optimal activity temperature (375 K) shows that the protein conformations are stably maintained, but the concerted motions are different. The flexibility of the protein at 375 K significantly increases, so the protein can adjust the local conformation to adapt to extreme temperature. The analysis of the changes in protein structure confirmed that the local conformation adjustment at 375 K plays a key role on the extreme high temperature stability.

Keywords tteRBP (Thermoanaerobacter tengcongensis ribose binding protein); molecular dynamics; intramolecular concerted motions; flexibility; extremely hemophilic

在生命进化过程中, 一些生物采取多种多样的、复杂的方式使蛋白质性能最大优化以适应环境^[1~4]. 近年来得到越来越多的关于蛋白质的环境、遗传和结构等方面的信息, 使人们可以更多地了解影响蛋白质高温热稳定性的因素^[5~7].

tteRBP 核糖结合蛋白来源于一种极端嗜热细菌(最佳生长温度为 353 K), 这种细菌最先从我国云南的温泉中发现^[8]. 这种蛋白属于外周胞质结合蛋白超级家族(periplasmic binding protein: PBP). 这个家族的成员在原

核物质运输^[9]、趋药性^[10]和细胞内信息传递系统发挥作用^[11]. PBP 的空间结构是由两个或三个 β 链组成一个灵活的铰链, 连接两个形成劈型结构的结构域组成, 每个结构域由一个三层的 $\alpha/\beta/\alpha$ 三明治结构组成, 两个结构域之间的界面是配体结合位置^[12~14]. tteRBP 的空间结构如图 1 所示(取自晶体结构).

tteRBP 的最佳活性温度为 375 K, 属于极端嗜热核糖结合蛋白. Matthew 等^[15]最先在 2008 年测定其晶体结构, 并通过与另一种主体结构极其相似的嗜热核糖结合

* E-mail: huangxuri200278123@yahoo.com.cn; Tel.: 0431-88499721.

Received July 3, 2011; revised November 20, 2011; accepted December 6, 2011.

图 1 蛋白 tteRBP(用带状表示)和底物 Ribose (Pyranose form)(用球棍表示)的结构

Figure 1 Overview of the tteRBP (in cartoon representation) and Ribose (Pyranose form) (in ball-and-stick representation)

蛋白比较分析, 推测出两种核糖结合蛋白侧链的作用不同, 进而导致其耐高温的机理. 本文选择 tteRBP 晶体结构作为分子模拟的初始结构, 对常温(300 K)和最佳活性温度(375 K)分别进行 20 ns 的分子动力学模拟, 比较 tteRBP 在两种温度下结构方面的差异, 找出其耐高温的原因.

1 计算方法

模拟的初始结构使用 RCSB 库中的晶体结构(PDB 编号: 2IOY). 模拟采用 GROMACS3.3.1^[16~18]程序包, TIP3P 水模型, 使用 AMBER03^[19]力场. 用 LINCS^[20]算法约束所有键长, 体系原子的初速度由 Maxwellian 分布给出, 系统温度通过由溶剂和溶质分别与参考温度的热浴耦合来保持, 温度耦合常数为 0.1 ps. 系统压强通过与外部压强浴耦合实现, 压强耦合常数为 0.5 ps, 压强为常压^[21]. 范德华和库仑相互作用的截断半径均为 1.4 nm, 短程邻近截断距离(rlist)为 0.8 nm.

系统模拟过程: 首先建立具有平衡 TIP3P 水分子的矩形盒子, 然后 tteRBP 放入盒子中, 体系与盒子边的最小距离为 0.9 nm, 盒子大小为 4.828 nm×4.358 nm×6.346 nm, 其中含有 20930 个水分子. 为保持整个体系的总电荷为零, 盒子中加入 7 个钠离子作为抗衡离子, 并除去相应位置的水分子. 准备好的体系都使用了周期化条件, 用最速下降法对体系进行能量最小化, 直到最大作用力小于 100 kJ/(mol·nm)为止, 然后运行 200 ps 的平衡动力学模拟. 之后分别在常温(300 K)和最佳活性温度(375 K)时进行 20 ns 自由分子动力学模拟, 步长为 2.0 fs.

熵是衡量分子柔性大小的有效参数之一, 我们采用了 Schafer 等^[22,23]提出的熵变计算方法来计算熵变.

2 结果与讨论

2.1 动力学体系的平衡

两个体系的动力学模拟轨迹中蛋白 α 碳原子的均方根偏差(RMSD)随时间变化的趋势如图 2 所示, 从图 2 中可以很明显地看出体系在 4 ns 左右达到平衡状态并且 tteRBP 在 300 和 375 K 的平均 RMSD 分别为 0.07 和 0.12 nm, 波动幅度变化比较小, 接着对 4~20 ns 的数据进行分析.

图 2 蛋白在两种温度 300 K (A)和 375 K (B)下模拟轨迹的 RMSD

Figure 2 Time evolution of RMSD of MD structure of the tteRBP300 K (A) and tteRBP375 K (B) respectively

2.2 蛋白结构变化的分析

盐桥、侧链相互作用、氢键及溶剂可及表面积等因素, 对嗜热酶结构有重要影响, 这里对其分析以寻找影响 tteRBP 极端高温稳定性的原因. 首先对两种温度下蛋白的侧链相互作用进行分析, 同时观察蛋白的盐桥数目变化情况. 蛋白的侧链相互作用及强度如图 3 所示(图 3 中亮点表示横纵坐标代表的残基有相互作用, 亮度表示相互作用强度), 结果表明没有明显变化. 同时通过测量有盐桥相互作用的残基间距离变化如表 1, 发现两种温度下盐桥数目没有明显变化. 这说明蛋白的侧链相互作用和盐桥不是影响蛋白极端高温稳定性的关键因素.

接着对蛋白与配体、蛋白内部、蛋白与溶剂间的氢键相互作用进行分析, 300 和 375 K 下蛋白与配体、蛋白内部、蛋白与溶剂间的氢键数目分别为 11.179, 215.295, 477.611 和 11.187, 221.122, 550.068, 由此可以看出在两种温度下, 蛋白与配体和蛋白内部的氢键相互作用基本保持稳定, 这表明了蛋白的构象保持整体稳定, 而 375 K 时蛋白与溶剂间的氢键相互作用强度显著增强, 表明

图3 蛋白在两种温度 300 K(上)和 375 K(下)的侧链相互作用图

Figure 3 Side chain interaction maps of the protein in the 300 K (up) and 375 K (down). Vertical and horizontal axis represent the residue

蛋白构象发生了局部调整。

在温度分别为 300 和 375 K 时, 蛋白的疏水性、亲水性和总的溶剂可及表面积分别为 51.907, 85.643, 137.550 和 51.130, 86.404, 137.704 nm², 由此得出两种温度下蛋白疏水性、亲水性和总的溶剂可及表面积几乎没有变化, 这说明蛋白表面的空间结构没有发生大的变化。

2.3 蛋白柔性分析

骨架 α 碳原子均方根浮动(RMSF: root-mean-square fluctuations)是显示蛋白柔性的重要参数。蛋白 tteRBP 分别在两种温度下的 RMSF 如图 4 所示, 同时对比二者的空间结构信息, 可以明显发现蛋白 tteRBP 在两种温度下其整体结构保持稳定。在 375 K 时蛋白的柔性波动, 除因残基 112~114 进入 α 螺旋导致邻近的 loop 区残基

119~122 比 300 K 时明显下降外, 整体有明显的提高, 这是高温下个别残基进出 α 螺旋和 β 折叠从而局部结构调整造成的。如残基 13~16, 55, 194~196, 253 进入 α 螺旋, 而残基 28~29, 42, 145~147, 178~179, 226 则从螺旋中解离出来。残基 98 进入 β 折叠, 同样残基 88, 153~154, 184, 232~236, 264~268 从二级结构解离。由此可以看出 tteRBP 在保持总体结构稳定的情况下, 高温时有较多的残基从二级结构中解离出来, 这就大大增加了蛋白的柔性, 使蛋白能够进行构象调整, 以有利的空间结构来适应极端高温, 有利于维持极端高温下蛋白的结构稳定。

图4 蛋白在两种温度 300 K (A)和 375 K (B)下模拟轨迹的 RMSF

Figure 4 Ca root mean square fluctuations as a function of residue number for tteRBP300 K (A) and tteRBP375 K (B) respectively. Abscissa represents the residue

2.4 本质动力学分析

本质动力学分析是研究蛋白质内部协调运动的重要手段, 它建立在蛋白 Ca 原子协方差矩阵对角化的基础上的, 把与蛋白性能有关主要的结构重排及运动情况的信息从局部的、任意的波动信息重提取出来, 以有利于我们更加详细的进行比较^[24,25]。表 2 列出了两个分子动力学模拟轨迹对角化协方差矩阵的前五个特征值, 由此可知前三个特征向量表征的分子内部协调运动在分子内部运动中有重要作用, 因此这里主要对前三个特征向量表示的运动进行具体分析。如图 5 所示, 在两种温度下第一个特征向量表示蛋白的整体运动强度高, 整体柔性没有明显区别。而第二个特征向量表示的运动强度明显降低, 整体柔性区别明显。第三个特征向量表示的运动强度弱, 整体柔性区别也不是很明显。

从图 5a 和图 5b 可看出第一特征向量表示的运动在 375 K 时残基 24~30 运动强度比 300 K 时稍弱, 而残基 45~57, 167~171, 199~207 明显增强, 高温下蛋白的配体结合口袋开口处残基 132~139 运动强烈, 在 300 K 时相同区域运动强度轻微。高温下蛋白部分区域运动强度明显增强, 特别是蛋白的配体结合口袋开口处的显著增强, 有利于蛋白与配体的结合。

从图 5c 和图 5d 可以看出整体柔性有十分明显的差

异, 在 300 K 时 N 端残基 269~272, 121 运动强度大, 残基 202~207, 216~219 运动强度较强, 而蛋白的其它部分运动微弱, 特别是 C 端结构域一侧运动不明显. 在 375 K 时不仅 N 端残基 268~274, 还有残基 53~56, 132~136(蛋白的配体结合口袋开口处), 174~182 运动强度大, 残基 27~34, 42~44, 71~79 运动强度较强, 蛋白的其它大部分区域也有较为明显的运动. 由此可以看出第二个特征向量所表示的运动在高温时不但部分区域增强, 整体也比 300 K 时增强明显, 特别是蛋白的配体结合口袋附近区域有显著增强, 这就充分说明高温下分子内部协调运动大大提升. 为了更进一步对此说明, 这里接着对第三个特征向量表示的运动进行分析. 由图 5e 和图 5f 可明显看出, 二者整体运动强度都大幅度减弱, 在 300 K 时除残基 72~83, 121~122, 176~183, 204~208, 255~256 有明显运动, 蛋白其他部分特别是 C 端结构域一侧运动不明显. 而在 375 K 时除残基 167~171, 199~207 运动明显外, C 端结构域中的残基

24~30, 47~57 运动明显, 特别是 300 K 时没有运动的蛋白的配体结合口袋开口处残基 132~139 运动强烈. 通过以上对比可知蛋白在高温时内部协调运动除个别区域外整体明显增强, 特别是蛋白的配体结合口袋开口处的柔性显著提高, 这显然有利于蛋白分子进行构象的调整.

分子内部熵值是分子体系整体柔性大小的直接体现, 这里对其进行计算以进行简单验证上述分析. 计算结果表明, 蛋白在 300 和 375 K 时的蛋白骨架(Protein Backbone)的熵值分别为 4349.43 和 5282.09 J/mol. 显示了高温下蛋白整体柔性明显增强, 与前面的分析一致.

结合以上比较分析, 可以证明在高温时 tteRBP 蛋白分子内部协调运动显著增强, 使蛋白分子结构进行了一系列局部微调, 在保持了蛋白分子整体结构的稳定的前提下, 也极大地提高了蛋白的整体柔性, 使蛋白能调整结构来适应极端高温.

图 5 两种温度 300 K (a, c, e)和 375 K (b, d, f)下前 3 个特征向量表示的运动

Figure 5 Motions described by the first three eigenvectors of tteRBP300 K (a, c, e), tteRBP375 K (b, d, f) respectively. The direction of arrow (next to Ca) indicated the direction of movement. The length of arrow indicated the intensity of movement.

表 1 两种温度 300 和 375 K 下有盐桥相互作用的残基间距离**Table 1** Distance of residues which have salt bridge interaction at 300 and 375 K respectively

Residues	tteRBP300 K	tteRBP375 K
88~89 ^a	0.42871 ^b	0.43285
122~153	0.47683	0.48803
127~159	0.42965	0.44010
165~168	0.45360	0.46984
166~192	0.37911	0.38389
250~255	0.45521	0.46684

^a Digital represent the number of residue. ^b The unit is nm.**表 2** 两种温度 300 和 375 K 下前 5 个特征向量的特征值**Table 2** Eigenvalues (nm) of the first five eigenvectors for tteRBP 300 and tteRBP 375 K respectively

Species	teRBP 300 K	tteRBP 375 K
Eigenvector 1	0.136742	0.23976
Eigenvector 2	0.105782	0.156875
Eigenvector 3	0.0439468	0.128062
Eigenvector 4	0.0425549	0.0960423
Eigenvector 5	0.0362693	0.0637953

References

- Kumar, S.; Nussinov, R. *Cell. Mol. Life Sci.* **2001**, *58*, 1216.
- Berezovsky, N.; akhnovich, E. I. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*(36), 12742.
- Vieille, C.; Zeikus, G. J. *Microbiol. Mol. Biol. Rev.* **2001**, *65*(1), 1.
- Jaenicke, R.; Bohm, G. *Curr. Opin. Struct. Biol.* **1998**, *8*(6), 738.
- Querol, E.; Perez-Pons, J. A.; Mozo-Villarias, A. *Protein Eng.* **1996**, *9*, 256.
- Kumar, S.; Tsai, C. J.; Nussinov, R. *Protein Eng.* **2000**, *13*, 179.
- Haney, P.; Konisky, J.; Koretke, K. K.; Luthey-Schulten, Z.; Wolynes, P. G. *Proteins* **1997**, *28*, 117.
- Xue, Y. F.; Xu, Y.; Liu, Y.; Ma, Y. H.; Zhou, P. *Int. J. Syst. Evol. Microbiol.* **2001**, *51*, 1335.
- Boos, W.; Shuman, H. *Microbiol. Mol. Biol. Rev.* **1998**, *62*(1), 204.
- Davidson, A. L.; Shuman, H. A.; Nikaido, H. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*(6), 2360.
- Neiditch, M. B.; Federle, M. J.; Pompeani, A. J.; Kelly, R. C.; Swem, D. L.; Jeffrey, P. D.; Bassler, B. L.; Hughson, F. M. *Cell* **2006**, *126*(6), 1095.
- Bjorkman, A. J.; Mowbray, S. L. *J. Mol. Biol.* **1998**, *279*(3), 651.
- Magnusson, U.; Chaudhuri, B. N.; Ko, J.; Park, C.; Jones, T. A.; Mowbray, S. L. *J. Biol. Chem.* **2002**, *277*(16), 14077.
- Sharff, A. J.; Rodseth, L. E.; Spurlino, J. C.; Quiocho, F. A. *Biochemistry* **1992**, *31*(44), 10657.
- Matthew, J. C.; Yaji, T.; Malin, A.; Homme, W. H. *BMC Struct. Biol.* **2008**, *8*, 20.
- Berendsen, H. J. C.; Van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43.
- Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 6.
- Spoel, D. V. D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 701.
- Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A. *J. Comput. Chem.* **2003**, *24*, 999.
- Hess, B.; Bekker, H.; Berendsen, H.; Fraaije, J. *J. Comput. Chem.* **1997**, *18*, 1463.
- Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
- Schafer, H.; Mark, A. E.; van Gunsteren, W. *J. Chem. Phys.* **2000**, *113*(18), 7809.
- Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115*(14), 6289.
- Amadei, A.; Linssen, A. B.; de Groot, B. L.; van Aalten, D. M.; Berendsen, H. J. *J. Biomol. Struct. Dynm.* **1996**, *13*, 615.
- Groot, B. L. d.; Amadei, A.; Scheek, R. M.; Nuland, N. A. J. V.; Berendsen, H. J. C. *Struct. Funct. Genetics* **1996**, *26*, 314.

(A1107031 Qin, X.)