

蛋白质与极性配体复合物的绝对结合自由能计算*

王存新** 宋伟 陈慰祖

(北京工业大学生命科学与生物工程研究院, 北京 100022)

摘要 介绍了用分子动力学模拟与热力学积分法相结合, 模拟蛋白质与配体的绝对结合自由能的方法. 通过分子转换法, 使蛋白质分子 (包括水分子) 与配体小分子之间的相互作用逐渐减弱 (或增强) 至完全消失 (或完全出现). 运用体约束方法, 计算了配体与受体结合后平动、转动自由度的丧失即熵效应所引起的自由能变化. 以胰蛋白酶双突变体 (D189G/G226D) 与极性配体苯甲脒为例, 研究了蛋白质活性部位与极性配体的相互作用对结合自由能的影响, 该复合物绝对结合自由能的模拟结果 (-15.5 kJ/mol) 与实验值 (-10.5 kJ/mol) 相近.

关键词 分子转换法, 热力学积分法, 蛋白质-配体复合物, 结合自由能

学科分类号 Q61

在分子模拟中, 体系的结构和动力学性质可以通过对分子动力学 (MD) 模拟所得轨迹进行系综平均得到. 但是有两个与生化功能密切相关的热力学量——熵与自由能则不能用以上方法直接求得, 它们应包括对整个体系相空间的充分采样, 因此理论上需要一个无限长的模拟时间. 10 年前建立起来的自由能微扰和热力学积分方法, 已成功地应用于生物大分子体系两种不同状态的相对自由能变化的计算, 取得了重要的进展^[1~4]. 但是如何用 MD 模拟方法计算蛋白质与配体复合物的绝对结合自由能, 一直是人们十分关心的研究问题.

5 年前, Hermans 等^[2]提出了分子转换法, 较好地计算了蛋白质与非极性小分子配体的绝对结合自由能. 该方法主要是引入一种称为体约束 (body restraint) 的势, 考虑蛋白质与配体结合时, 分子从自由状态变为复合物状态过程中, 平动、转动自由度的丧失所带来的自由能变化. 这部分自由能的变化主要是熵的贡献, 如配体分子从自由态到结合态, 是一个熵减的过程. 因静电相互作用对结合自由能的贡献较大, 对配体分子为极性的复合物体系, 如何正确处理活性部位极性带电基团之间的相互作用, 就显得特别重要. 因此, 如何用 MD 模拟与热力学积分法相结合, 正确计算和预测蛋白质与极性配体复合物的绝对结合自由能, 是一个至今尚未很好解决的问题. 本工作的主要目的是, 在方法学上弄清蛋白质与配体结合时由于自由度消失所引起的熵效应对结合自由能的影响, 选择结构与功能都比较清楚的胰蛋白酶双突变体与极性配体苯甲脒复合物为研究对象, 弄清蛋白质活性部位与极性配体的相互作用对结合自由能的影响.

1 体系与方法

选择胰蛋白酶双突变体 (D189G/G226D) 与极性配体苯甲脒的复合物体系为研究对象, 蛋白质数据库 (PDB) 代码为 1BRA. 该体系结合自由能的实验值为 10.5 kJ/mol (结合常数 $K_i = 1.5 \times 10^4 \text{ L} \cdot \text{mol}^{-1}$)^[5]. 苯甲脒分子是一刚性小分子, 所有原子基本位于同一平面, 带有一个单位正电荷. 胰蛋白酶突变体的活性口袋内可容纳 6 个水分子, 酶与配体结合后, 苯甲脒取代了其中 2 个水分子的位置.

图 1 为计算绝对结合自由能的热力学循环图, 其中 P 为受体蛋白, L 为配体分子, P-L 表示蛋白质-配体复合物, P- $n\text{H}_2\text{O}$ 表示配体分子被 n 个水分子取代, 模拟中水分子采用简单点电荷 (SPC) 模型^[6]. 下标 w 表示水相, g 表示气相, s 表示标准态, r 表示约束态. ΔG_1 为在标准态时 1 个水分子的去溶剂化自由能; ΔG_2 为气相水分子由标准态变为约束态的自由能之差; ΔG_3 为液相下的复合物分解为液相下的受体与气相下的配体所引起的自由能变化, 同时还考虑了配体进入气相后 2 个水分子取代苯甲脒配体的过程, 这是本工作的关键步骤, 称为分子转换法 (molecular transformation); ΔG_4 是气相下配体分子由约束态变为标准态所引起的自由能变化; ΔG_5 为在标准态时配体分子的溶剂化自由能.

* 国家自然科学基金 (29992590-2, 30170230, 10174005) 和北京市自然科学基金 (5032002) 资助项目.

** 通讯联系人.

Tel: 010-67392724, E-mail: cxwang@bjpu.edu.cn

收稿日期: 2002-08-12, 接受日期: 2002-11-21

受体与配体在标准情况下的结合自由能的实验值应为图 1 中 ΔG 的负值, 因此绝对结合自由能为: $-\Delta G = -(n\Delta G_1 + n\Delta G_2 + \Delta G_3 + \Delta G_4 + \Delta G_5)$ (1) 因为苯甲脒在活性口袋中取代了两个水分子的位置, 所以在模拟时, 将图 1 热力学循环中水分子的数目 n 取为 2.

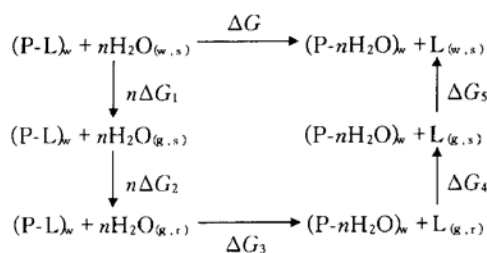


Fig. 1 The scheme of thermodynamic cycle for calculating the absolute free energy

关于约束对自由能的贡献, 这里只作简单介绍, 详见文献 [2]. 在气相下, 配体由约束态变为标准态所引起的自由能变化 ΔG_4 可表示为:

$$\Delta G_4 = -(\Delta G_{r,x}^0 + \Delta G_{r,\theta}^0 + \Delta G_{r,\phi}^0 + \Delta G_s^0) \quad (2)$$

其中 $\Delta G_{r,x}^0$ 、 $\Delta G_{r,\theta}^0$ 和 $\Delta G_{r,\phi}^0$ 分别为质心位置约束、角约束和二面角约束所引起的体系自由能的变化, ΔG_s^0 为配体分子对称性所引起的自由能的变化. 对于配体苯甲脒分子, 分别做质心位置约束、角约束和二面角约束. 对于 SPC 水分子, 仅做质心位置约束, 因此气相下水分子由标准态变为约束态所引起的自由能变化为:

$$\Delta G_2 = \Delta G_{r,x}^0 \quad (3)$$

其中 $\Delta G_{r,x}^0$ 为质心位置约束所引起的自由能变化.

MD 模拟和自由能计算采用 SIGMA 程序库 (见: Hermans, *et al.* SIGMA Documentation, 1999). 复合物体系被置于 $6 \text{ nm} \times 6 \text{ nm} \times 6 \text{ nm}$ 的立方体盒子中, 周围用 5 515 个水分子填充. MD 模拟采用周期

性边界条件. 蛋白质与配体结合时体系自由能的变化 ΔG_3 , 通过 MD 与热力学积分方法相结合来实现 (热力学缓慢生长法), ΔG_3 可写为:

$$\Delta G_3 = \int_0^1 \langle \partial U / \partial \lambda \rangle_\lambda d\lambda \quad (4)$$

其中 U 为体系势函数, λ 为耦合参数, $\langle \dots \rangle_\lambda$ 表示系综平均. 势函数 U 是 λ 的函数, $U(\lambda=1)$ 对应于复合物状态, $U(\lambda=0)$ 对应于非结合自由状态. MD 积分步长取为 2 fs, 体系温度为 300 K, 压力为 $1.01 \times 10^5 \text{ Pa}$. 在 $\lambda=1$ 时做 20 ps 的 MD 平衡, 然后慢慢在 100 ps 内将 λ 由 1 变为 0, 随后在 $\lambda=0$ 时再做 20 ps 的 MD 平衡, 然后慢慢增加 λ , 在 100 ps 内将 λ 由 0 变为 1. 因此, 以上模拟 ΔG_3 过程总的模拟时间为 240 ps.

2 结果与讨论

2.1 约束引起的自由能变化 ΔG_2 和 ΔG_4

对于水分子, 模拟时只做了质心位置约束, ΔG_2 是对水分子加约束后引起的体系自由能的变化, 所以方程 (3) 中 ΔG_2 为正值. 对于配体分子, 分别做了质心位置约束、角约束和二面角约束, 其目的是使配体分子的平动和转动自由度消失. 因为 ΔG_4 为释放配体的约束而引起体系自由能的变化, 所以由方程 (2) 计算出的 ΔG_4 应为负值. 因为约束所引起体系自由能的变化与约束力常数有关, 所以在模拟时, 我们把质心位置约束、角约束和二面角约束的约束力常数 K 分别取为 5×4.186 、 10×4.186 、 50×4.186 三种情况进行计算, 其中位置约束力常数 (K_x) 的单位为 $100 \text{ kJ}/(\text{mol} \cdot \text{nm}^2)$, 角约束力常数 (K_θ) 和二面角约束力常数 (K_ϕ) 的单位为 $\text{kJ}/(\text{mol} \cdot \text{rad}^2)$. 约束力常数 K 取不同值时 ΔG_2 和 ΔG_4 的值列于表 1.

Table 1 The free energy changes corresponding to every processes in Fig. 1 obtained from simulation data for different restraint parameters K^1

K	ΔG_1	ΔG_2	ΔG_3	ΔG_4	ΔG_5	ΔG
5×4.186	+ 26.4	+ 19.6	+ 182.1	- 28.1	- 230.2	- 15.5
10×4.186	+ 26.4	+ 22.2	+ 195.1	- 33.3	- 230.2	- 28.5
50×4.186	+ 26.4	+ 28.3	+ 185.4	- 45.4	- 230.2	- 18.8

¹⁾ The unit for free energy changes is in kJ/mol. The unit for restraint parameters K : K_x in $100 \text{ kJ}/(\text{mol} \cdot \text{nm}^2)$, K_θ and K_ϕ in $\text{kJ}/(\text{mol} \cdot \text{rad}^2)$.

2.2 水化自由能 ΔG_1 和 ΔG_5

水分子的去水化自由能 ΔG_1 取自文献 [7], $\Delta G_1 = 26.4 \text{ kJ/mol}$. 苯甲脒分子的水化自由能 ΔG_5 , 通过 MD 模拟得到. 应用热力学缓慢生长

法, 将处于水相中的苯甲脒分子与水分子的相互作用慢慢减弱至完全消失. 体系自由能随 λ 的变化情况如图 2 所示. 由图 2 可以看出, 体系水化自由能的变化具有比较好的可逆性. 模拟得出的 ΔG_5

的平均值为 -230.2 kJ/mol 。

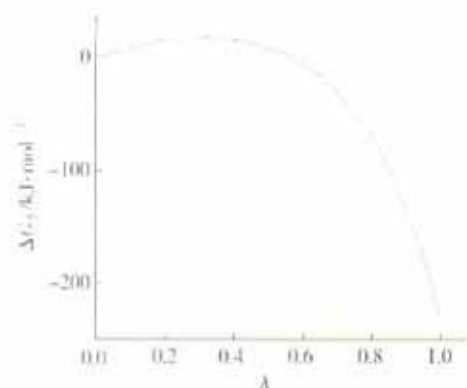


Fig. 2 The hydration free energy ΔG_3 of benzamidine versus coupling parameter λ

The solid curve (red) represents the forward simulation (λ from 0 to 1).

The dotted curve (blue) shows the reverse simulation (λ from 1 to 0).

2.3 用分子转换法模拟得到的自由能变化 ΔG_3

分别取 3 种不同的约束力常数 ($K = 5 \times 4.186$, 10×4.186 , 50×4.186), 在温度 $T = 300 \text{ K}$, 压力为 $1.01 \times 10^5 \text{ Pa}$, 应用热力学缓慢生长法模拟了胰蛋白酶与苯甲脒从结合态 ($\lambda = 1$) 到非结合态 ($\lambda = 0$) 的整个过程, 并考虑到苯甲脒消失后水分子占据其留下的空穴。 ΔG_3 的计算结果列于表 1。由模拟结果看出, 约束力常数 K 值的选取对模拟结果有一定的影响。由表 1 可见, 当约束力常数选为 5×4.186 时, 总的模拟结果为 -15.5 kJ/mol , 与实验结果 -10.5 kJ/mol 比较接近。

图 3a 给出了 20 ps MD 平衡后胰蛋白酶与苯甲脒复合物 (PDB 代码 1BRA) 结合位点附近的结构, 虚线表示氢键网络。图 3 中的折线表示胰蛋白酶结合位点附近的 C_α 主链, BEN 表示苯甲脒分子, WAT 表示水分子。苯甲脒分子、水分子和有关氨基酸残基均用球棍模型表示, 只显示重原子, 省略了氢原子。由图 3a 可以看出, 苯甲脒与胰蛋白酶结合时, 复合物活性口袋里有 WAT3、WAT4、WAT5 和 WAT6 4 个水分子, 其中 WAT5 埋藏在口袋深处, 苯甲脒的两个胺基 N 原子, 一个 N1 原子与 WAT6 氧原子和 Gly216、Gly219 主链羰基氧原子形成氢键, 另一个 N2 原子与 WAT3、WAT4 的氧原子和 Ser190、Asp226 的侧链氧原子形成氢键, 如图 3a 中的虚线所示。图 3b 显示出苯甲脒分子从结合位点附近消失后结合位点附近的结构示意图。由于 WAT3、WAT4、WAT5 在口袋深处, 它们的相对位置大致保持稳定, 图 3a 中的 WAT6 水分子在口袋顶部附近, MD 模拟过程中, 随着苯甲

脒消失, WAT 6 水分子的氧原子与苯甲脒的 N1 原子的氢键随之消失, WAT 6 的位置随之发生了变化, 由另一水分子 WAT 1000 取代。由图 3b 可以看到, 两个水分子 WAT 1' 和 WAT 2' 占据了苯甲脒分子消失后留下的空穴。当苯甲脒分子转换为两个水分子后, WAT 1' 与 Gly216、Gly219 的主链羰基氧原子、

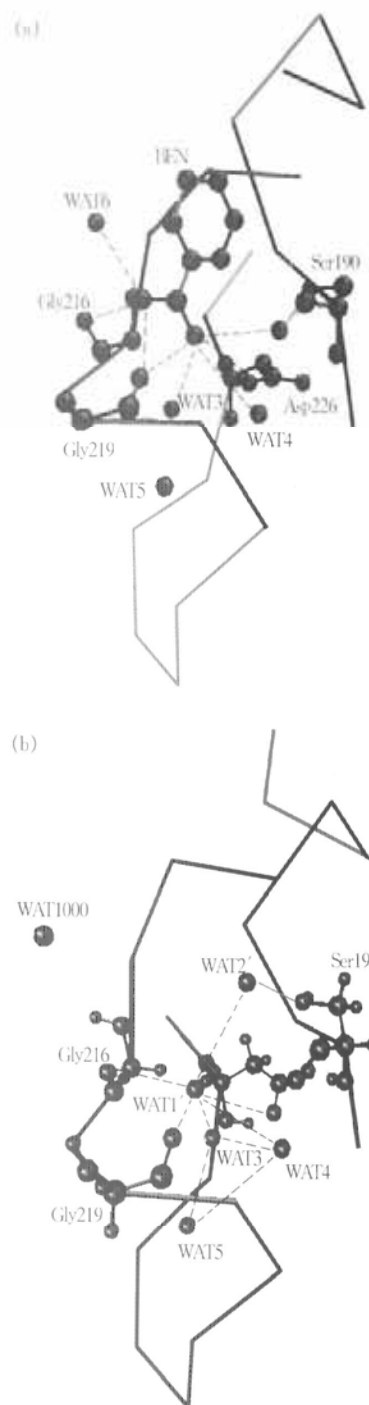


Fig. 3 Hydrogen binding network around the binding site
(a) The snapshot structure around the binding site of the trypsin and benzamidine complex; (b) The snapshot around the binding site when benzamidine disappears from binding site. BEN: benzamidine, WAT: H_2O .

Asp226 的侧链氧原子以及 WAT 3、WAT 4 形成氢键。WAT 2' 与 Ser190 的侧链羟基和 WAT 1' 形成氢键。

需要指出的是, 苯甲脒转化成的水分子 WAT 1' 和 WAT 2' 的空间位置对模拟结果有直接的影响。本工作应用体约束方法使水分子 WAT 1' 和 WAT 2' 的氧原子分别约束在苯甲脒的适当位置, 并使 WAT 1'、WAT 2' 与苯甲脒无任何相互作用。经过多次试验, 发现约束位置选在苯甲脒的 N2 和 C6 原子上结果较好。当用分子转换法模拟时, 随着苯甲脒分子的慢慢消失, 水分子 WAT 1' 和 WAT 2' 慢慢从合适的约束位置上生长出来, 并与周围环境发生相互作用, 形成合理的氢键网络, 从而得到了比较合理的结构。

3 结 论

用 MD 模拟和分子转换法模拟了蛋白质分子在亲水的结合位点与配体分子的相互作用及绝对结合自由能, 并将该方法应用于胰蛋白酶与极性配体苯甲脒的复合物体系。模拟结果表明, 约束力常数的选取对模拟结果有一定影响, 对于本工作所研究的体系, 当约束力常数选为 5×4.186 时, 总的模拟结果 (-15.5 kJ/mol) 与实验结果 (-10.5 kJ/mol) 比较接近。从 MD 模拟轨迹分析可知, 蛋白质活性位点附近的氢键网络对复合物的稳定结构有重要作用。另外, 对于如何用体约束和分子转换法正确处

理配体进入气相后水分子取代配体的过程, 该方法及相关程序还有进一步改进的地方, 比如约束水分子的个数选取和约束方法的改进, 将在下一步工作中加以解决。

致谢 感谢 Hermans 教授提供的 SIGMA 程序包和对本工作提出的指导意见。

参 考 文 献

- 1 Beveridge D L, Dicapua F M. Free energy via molecular simulation: applications to chemical and biomolecular systems. *Annu Rev Biophys Biomol Struct*, 1989, **18**: 431~492
- 2 Hermans J, Wang L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding. Application to a complex of benzene and mutant T4 lysozyme. *J Am Chem Soc*, 1997, **119** (11): 2707~2714
- 3 Wang C X, Shi Y Y, Zhou F, *et al.* Thermodynamic integration calculations of binding free energy difference for Gly169 mutation in subtilisin BPN'. *Proteins: Struct Funct Genet*, 1993, **15** (1): 5~9
- 4 Jorgensen W L, Buckner J K, Boudon S, *et al.* Efficient computation of absolute free energies of binding by computer simulation: application to the methane dimer in water. *J Chem Phys*, 1988, **89** (6): 3742~3746
- 5 Perona J J, Tsu C A, McGrath M E, *et al.* Relocating a negative charge in the binding pocket of trypsin. *J Mol Biol*, 1993, **230** (3): 934~949
- 6 Berendsen H J C, Postma J P M, van Gunsteren W F, *et al.* *Intermolecular Forces*. Dordrecht: Reidel, 1981. 331
- 7 Hermans J, Pathiaseril A, Anderson A. Excess free energy of liquids from molecular dynamics simulation. application to water models. *J Am Chem Soc*, 1988, **110** (18): 5982~5986

The Absolute Free Energy Calculation of The Protein and Polar Ligand Complex*

WANG Curr Xin**, SONG Wei, CHEN Wei-Zu

(Institute of Life Science and Bioengineering, Beijing Polytechnic University, Beijing 100022, China)

Abstract Molecular dynamics simulation and thermodynamic integration method were used to calculate the absolute binding free energy of the protein-ligand complex. By the molecular transformation method, the interactions between protein (plus solvent) and its ligand are gradually decreased (or increased) into a non-interacting (or full interacting) state. A potential of a body restraint was used to calculate the free energy changes caused by the loss of translation and rotation freedom of the ligand molecule, that is called as the entropy effect. A mutant trypsin (D189G/G226D) and its polar ligand (benzamidine) were selected as a model to study the influence of the interactions between protein and polar ligand on the binding free energy. The calculation result of the absolute free energy for the model complex (-15.5 kJ/mol) is close to the experimental data (-10.5 kJ/mol).

Key words molecular transformation method, thermodynamic integration method, protein-ligand complex, binding free energy

* This work was supported by grants from The National Natural Sciences Foundation of China (29992590-2, 30170230, 10174005) and Beijing Natural Science Foundation (5032002).

** Corresponding author. Tel: 86-10-67392724, E-mail: cxwang@solaris.bjpu.edu.cn

Received: August 12, 2002 Accepted: November 21, 2002