

Oxford, 1983.

- [6] Ehinger, B.: *Retina*, 2, 305, 1983.
[7] Lam, D. M. K. et al.: *Exp. Eye Res.*, 31, 729, 1980.
[8] Lam, D. M. K. et al.: *Proc. Natl. Acad. Sci., USA*, 75, 6310, 1978.
[9] Pourcho, R. G.: *Brain Res.*, 198, 333, 1980.
[10] 杨雄里:《生理科学进展》, 16, 79, 1985.

- [11] Ehinger, B.: *Vision Res.*, 23, 1281, 1983.
[12] Osborne, N. N.: In *Biology of Serotonergic Transmission* (Ed. Osborne, N. N.), Wiley, New York, 1982.
[13] Karten, H. J. et al.: In *Neurotransmitter Interaction and Compartmentation* (Ed. Bradford, H. F.), Plenum Press, New York, 719, 1983.

[本文于 1984 年 8 月 7 日收到]

蛋白质和核酸动力学的计算机模拟

施蕴渝 贡汝槐 王存新

(中国科学技术大学理论分子生物物理研究组, 合肥)

近几年, 核酸和蛋白质内部的动力学已成为广泛引起人们兴趣的课题。为了说明生物大分子结构和功能的关系, 不仅需要了解原子的平均位置, 而且也需要知道原子位移随时间的涨落。大量实验表明生物大分子结构不仅有刚性而且存在柔性, 而且这种柔性 with 生物大分子作用机制有密切关系。最近 X 射线晶体衍射关于蛋白质温度因子的研究, 在各向同性和谐振的假设下, 已能估计出这种涨落的数值^[1], 核磁共振, 荧光去偏振, 激光拉曼光谱等实验也说明了生物大分子柔性的存在^[2-4]。

在大量实验工作的基础上, 这方面的理论工作也有了很快的发展。分子动力学模拟曾成功地应用于含有大量原子的气体和液体, 从中可以得到系统热力学的信息。第一个将分子动力学模拟用于蛋白质的是 M. Karplus^[5], 他们对包含 48 个氨基酸残基的小蛋白质牛胰蛋白酶抑制剂进行了真空环境下的分子动力学研究。随后, 肌红蛋白, 细胞色素 c, 铁硫蛋白, DNA, tRNA 等一系列生物大分子在真空环境下的分子动力学模拟工作也完成了。最近 Van Gunsteren 等人^[6]考虑了溶剂和晶体环境, 对上述工作作了改进, 使理论计算值和实验值取得了比较好的一致性。现在分子动力学模拟已成为研究生物大分子动力学的一个重要方面。1981 年 5 月欧洲分子生物学组织^[7]与 1982 年 8 月国际理论物理中心举办的暑期讨论班^[8]都对上

述内容作了专题介绍。还举行过多次国际会议^[9,10], 并有专题评论。

本文是根据文献及近年来我们自己工作的体会, 对生物大分子动力学计算机模拟的基本原理和方法, 目前取得的主要成果及存在的问题, 作一介绍。

一、原理和方法

分子动力学方法是一种重要的统计力学方法。在准各态历经假设下, 任何力学量对系综的平均等于该力学量对时间的平均, 而力学量对时间的平均可以从经典运动方程所决定的原子运动轨迹得到。具体来说, 该方法是:

在一定的初始条件 $\{\mathbf{r}_i(0)\}$, $\{\mathbf{v}_i(0)\}$ 下解蛋白质中所有原子的运动方程组

$$\frac{d^2\mathbf{r}_i}{dt^2} = \mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)/m_i \\ i = 1, 2, \dots, n$$

其中 $\mathbf{r}_i$ 是第 i 个原子的位置矢量, m_i 是第 i 个原子的质量, $\mathbf{F}_i$ 是第 i 个原子所受到的作用力, 由位能函数 V 导出

$$\mathbf{F}_i(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = -\nabla_i V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \\ i = 1, 2, \dots, n$$

分子的能量照理应用量子力学的方法来计算, 但是由于生物大分子是一个大而复杂的系统, 原子数很多。国际量子生物协会主席 Clementi 估计^[12]: 量子力学计算的上限是 150—

200 个原子 (不包括有大量对称性的情况)。迄今为止所发表的从头计算还没有超过 60—80 个原子的, 而且从头计算所得的结果并不比简单的多且花钱少得多的半经验位能函数法所得的结果准确多少^[13], 因此目前国际上主要采用半经验位能函数法。

一个典型的位能函数的形式如下:

$$V = \sum_{\text{键长}} \frac{1}{2} K_b (b - b_0)^2 + \sum_{\text{键角}} \frac{1}{2} K_\theta (\theta - \theta_0)^2 \\ + \sum_{\text{非正常二面角}} \frac{1}{2} K_\xi (\xi - \xi_0)^2 + \sum_{\text{正常二面角}} K_\phi [1 \\ + \cos (n\phi - \delta)] + \sum_{\text{非键}} \left[\frac{C_{12}}{r_{ij}^{12}} - \frac{C_6}{r_{ij}^6} \right. \\ \left. + \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \right]$$

其中第一项是共价键振动能, 第二项是键角弯曲能, 第三、第四项分别是非正常二面角及正常二面角扭转能, 最后一项是范德华相互作用和静电相互作用能。参数 C_6 由 Slater-Kirkwood 公式得:

$$C_6 = (4\pi\epsilon_0)^{-2} (3/2 c \hbar m_e^{-1/2}) \alpha_i \alpha_j [(\alpha_i/N_i)^{1/2} \\ + (\alpha_j/N_j)^{1/2}]^{-1} \\ C_{12} = \frac{1}{2} C_6 (r_i + r_j)^6$$

其中 α_i 是原子极化率; N_i, N_j 是原子有效外壳层电子数; r_i, r_j 是所考虑的一对原子的范德华半径; ϵ_0 是真空介电常数; e 是电子电荷; $\hbar = h/2\pi$, h 是普朗克常数。 b, θ, ξ, ϕ 分别是键长, 键角, 非正常二面角, 正常二面角。 r_{ij} 是一对原子间距离, q_i, q_j 是原子部分电荷。 $b_0, \theta_0, \xi_0, K_b, K_\theta, K_\xi, K_\phi, n, \delta, \epsilon_r$ 为参数。参数是通过实验数据的拟合而确定。

模拟从已知的蛋白质 X 射线晶体坐标出发, 首先用能量极小化方法对坐标进行修正, 得到原子的初位置, 再指定一个温度, 按照此温度下的麦克斯威速率分布, 给原子赋予一定的速度, 然后进行分子动力学模拟。在模拟过程中系统要逐渐趋于平衡, 因为蛋白质起始构象所相应的能量一般不是最低能量, 所以在模拟初期势能会逐渐减少, 温度要逐渐升高。为了使

系统达到某一指定的温度 T_0 , 所有速度都要乘上一个修正因子 $[2T_0/\langle T \rangle - 1]^{1/2}$ 。在第一个微微秒中每 0.1 微微秒修正一次, 以后每当 $|\langle T \rangle - T_0| > \Delta T$ 时才修正, 直到温度在一个较长时间范围内稳定在 $T_0 \pm \Delta T$ 范围内。 $\langle T \rangle$ 是平均温度, 由系统的平均动能按下式算出来

$$\frac{1}{2} \sum_{i=1}^N m_i \langle v_i^2 \rangle = \frac{3}{2} N K_B \langle T \rangle$$

其中 m_i 是第 i 个原子的质量, $\langle v_i^2 \rangle$ 是第 i 个原子的速度平方的平均值, K_B 是玻尔兹曼常数。模拟过程中还要消除分子的整体运动。

达到平衡所需要的时间主要取决于相互作用的范围, 系统的大小, 初始构象和速度。

微分方程用 Gear 预估校正算子或 Verlet 算子求解^[14], 积分时间步长一般为 $1 \times 10^{-3} - 2 \times 10^{-3}$ 微微秒。

为了避免慢的温度漂移, 可以将系统与一个温度为 T_0 的热浴相耦联, 即在分子动力学模拟的每一步按下述关系:

$$\frac{dT(t)}{dt} = \tau^{-1} [T_0 - T(t)]$$

对温度重新标定, 耦合强度由温度弛豫时间 τ 决定, τ 一般取 0.01 微微秒。

蛋白质在室温下共价键的键长变化很少, 其涨落的方差为 $0.003 \text{ \AA}$, 因此可以对键长加以约束, 这可以通过 SHAKE^[15] 的方法来实现。

为了更真实地反映生物大分子的实际情况需要模拟晶体环境或溶液的情况。为模拟晶体环境, 需要将中心分子坐标按照空间群进行对称操作, 从而得到一个晶胞中其它分子的坐标。然后加周期性边界条件开拓出去。为模拟溶液中的情况亦须要按照水的密度为 1 克/厘米^3 在生物大分子周围加上事前经过分子动力学模拟的水分子, 然后再加上周期性边界条件开拓出去。

通过分子动力学模拟可以得到在某一温度下, 球蛋白中每一个原子在几十到一百微微秒范围内运动的轨迹, 并由此得到每个原子距平

衡位置的均方根涨落。这可通过下述关系与由X射线晶体衍射测温度因子推算出的原子距平均位置的均方根涨落对比。在各向同性和谐振假设下,原子位置的均方根涨落 $\langle \Delta r^2 \rangle_{\text{dyn}}$ 与温度因子 (Dybye-Waller 因子) B 之间有下列关系:

$$\langle \Delta r^2 \rangle_{\text{dyn}} = \frac{3B}{8\pi^2} - \langle \Delta r^2 \rangle_{\text{dis}}$$

其中 $\langle \Delta r^2 \rangle_{\text{dis}}$ 是晶格无序或其它作用对 B 的贡献。分子动力学模拟的结果还可以与核磁共振测出的弛豫参数对比。

一旦得到了原子的运动轨迹,还可作相应的谱分析和相关分析,从而得到频谱和功率谱,人们即可对相应的运动特性进行分析。归一化的时间相关函数按下述表达式计算:

$$C_A(t_n) = \frac{\sum_{m=1}^{N-n} A(t_m)A(t_m + t_n)}{\frac{N-m}{N} \sum_{m=1}^N [A(t_m)]^2}$$

其中 A 是位移矢量 $\Delta r = (\Delta x, \Delta y, \Delta z)$ 或其分量。

二、主要成果

胰蛋白酶抑制剂是由 58 个氨基酸残基(包

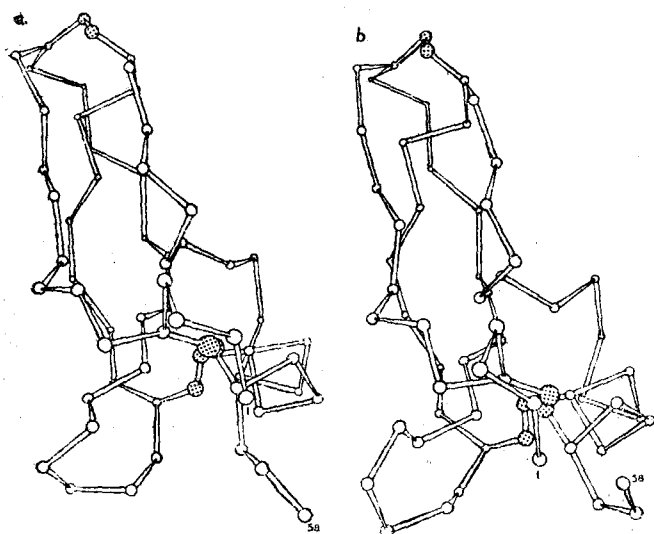


图1 胰蛋白酶抑制剂的分子结构

(a) X射线衍射所得的分子结构

(b) 3.2微微秒分子动力学模拟后所得的分子结构

括 454 个重原子)组成的蛋白质(在计算时氢原子已被并入它们所结合的重原子),有四个以内部氢键相结合的水分子。M. Karplus 于 1977 年^[5]首先完成对该蛋白质的分子动力学模拟。图1是牛胰蛋白酶抑制剂X射线衍射所得平均分子结构(a)及分子动力学模拟 3.2 微微秒后的分子结构(b)。二者十分相似,但也有一些小的差别。这个结果说明分子柔性的存在,而且所有的残基柔性并不相等,而是某些区域柔性显然更大,最大的偏移在C末端以及在图左下方。

最近 W. F. Van Gunsteren 等人^[6]对上述工作进一步改进,完成了单位晶胞中 4 个胰蛋白酶抑制剂分子及 560 个水分子的分子动力学模拟,总自由度为 11844。分子动力学模拟所得的平均结构与X射线晶体衍射所得的平均结构的均方差

$$= \left(\frac{[\langle r^2 \rangle - \langle r \rangle^2]}{2} \right)^{1/2}$$

对 C_α 原子是 0.8 Å, 对所有原子是 1.2 Å。一般来说分子动力学模拟在真空中的精确度是 0.1—0.4 nm, 包括了水和晶体环境精确度可达 0.1—0.15 nm。

实验和理论工作表明,被研究的大多数蛋白质主链原子位置的均方根涨落为 ~ 0.4 Å, 长的侧链末端这种涨落可达 1.5 Å。对于有氢键维系的二级结构部分 (α 螺旋, β 折叠), 其涨落小于无规卷曲部分。

分析胰蛋白酶抑制剂, 细胞色素 c 等蛋白质中原子位置随时间的变化, 结果表明原子涨落可视为高频局部振动与低频集合性振动二者的叠加。前者的时间范围是亚微微秒, 平均频率约 $\bar{\nu} = 200 \text{ cm}^{-1}$, 而后的时间范围是 1—10 微微秒, 甚至更长, 频率为 $\bar{\nu} \approx 3\text{--}30 \text{ cm}^{-1}$ 。前者对主链原子的均方根涨落大约贡献 40%, 对蛋白质内部所有原子都是均一的, 而后者仅涉及到几个近邻原子, 一个残基或蛋白质内部某一区域许多原子形成

的集团,反映蛋白质内部不同区域的不同特征。后者与生物大分子的功能有更密切的联系。

自 1977 年第一篇关于蛋白质的分子动力学模拟的工作发表以来,这方面的论文已发表了几十篇。所研究的对象包括胰蛋白酶抑制剂,肌红蛋白,细胞色素 c,铁硫蛋白, DNA 和 tRNA。理论工作紧密与 X 射线晶体衍射,核磁共振,中子衍射等实验工作相结合,已成为研究生物大分子动力学的一个重要方面。

在此基础上又发展了随机动力学,活化动力学等方法。并将这些方法用于酶反应机理的研究及药物设计。

三、分子动力学方法的局限性

1. 分子动力学使用的是经典力学,解的是系统中所有原子的运动方程,没有考虑量子效应。不过对大分子来说,在 300°K 下这并不是一个严重的问题。

2. 原子运动轨迹依赖于势函数,而势函数是半经验性质的;势函数越可靠,分子动力学模拟的结果将越好。

3. 体系的大小有限制,仅仅有限数目的粒子可以被模拟;目前已经模拟的最大系统含 7000 个原子。

4. 模拟的时间范围也受限制。使用现代的计算机,模拟的时间从几十微微秒到毫微微秒。

后两条限制随着计算机速度的提高,以及理论或技术的发展将会改变。

生物大分子的分子动力学模拟工作还刚刚开始,可以期望此方法在今后几年内将更广泛地被应用于各种生化问题,从而提供更多的有关生物大分子内部动力学及生化反应方面的信息。

参 考 文 献

- [1] Frauenfelder, H. et al.: *Nature* (London), **280**, 558, 1979.
- [2] Wagner, G. et al.: *Nature* (London), **275**, 274, 1978.
- [3] Holowka, D. A. et al.: *Biochemistry*, **15**, 3379, 1976.
- [4] Genzel, L. et al.: *Biopolymers*, **18**, 507, 1976.
- [5] McCammon, J. A. et al.: *Nature* (London), **267**, 585, 1977.
- [6] van (Gunsteren, W. F. et al.: *Proc. Natl. Acad. Sci. USA*, **80**, 4315, 1983.
- [7] Karplus, M.: *Structural Molecular Biology* (Davies, D. B. Edited), Plenum, New York, 1981.
- [8] Karplus, M.: Summer college in biophysics, International Centre for Theoretical Physics. 1982.
- [9] Karplus, M.: *Mobility and function in protein and nucleic acid*. Pitman, London (Ciba Foundation symposium 93), 1982.
- [10] Workshop: *Molecular Dynamics Methods and Protein Structure*, Chapel Hill, USA, 1984.
- [11] Karplus, M. et al.: *CRC Crit. Rev. Biochem.*, **9**, 293, 1981.
- [12] Clementi, E.: *Computational aspects for large chemical system*, Sping-Verlag 1980.
- [13] Olson, W. K.: «访华在吉林大学演讲稿», 1981.
- [14] Gear, C. W.: *Numerical initial value problems in ordinary differential equation*.
- [15] van Gunsteren, W. F.: *Molecular Physics*, **40**, 1015, 1980.

[本文于 1984 年 12 月 5 日收到]

短梗霉多糖的性质和应用

梁 金 虎

(中国科学院生物物理研究所,北京)

天然多糖多取自植物的根、茎和果实以及海藻,也有许多则利用微生物来生产。它们的特点是无毒无害,食用性好,不会引起三致(致畸,致癌和致突变)。美,日利用微生物发酵生产多糖,广泛应用于工业,农业,医药卫生和食

品工业等领域。目前,天然多糖的世界年产量约 50 万吨^[1]。

本文主要介绍一种较重要的微生物多糖——短梗霉多糖(Pullulan)的性质及其应用。