

分子体积及表面积的 Monte Carlo 模拟计算

商志才 俞庆森 林瑞森

(浙江大学化学系, 杭州 310027)

摘要 建立了用 Monte Carlo 方法模拟计算 van der Waals 分子体积和表面积的算法. 在一定的置信度条件下, 可获得在指定置信限内的期望值. 与 Bodor 算法比较, 此算法有更优的精度.

关键词: 分子体积和表面积, Monte Carlo 模拟, Bodor 算法

分子体积及表面积与许多理化参数有较好的相关性, 如溶解度, 正辛醇 / 水分配系数, 无限稀释活度系数等. 近来, 分子体积及表面积与药物的生物活性、有机物的毒性、多环芳烃的致癌活性、蛋白质的结构与功能等之间的关系更是得到了比较多的关注.

分子的体积和表面积可由基团加和法得到^[6], 但在有些场合精度显然不够. 严格的方法应采用 Wong 算法^[2], 将由“0.001 au 电子密度封套”占有的体积定义为分子体积, 但由于计算速度的原因, 对略大的分子如氯丙烷却不得不采用加和片段体积常数来计算. 故实际计算中, 常假定分子是由以原子核为中心, van der Waals 原子半径为半径的球形原子组成. 显然, 分子表面积 S 就是分子中 van der Waals 球的相交面, 分子表面积包围的体积就是分子体积 V .

文献大多采用网格 (grid) 技术. 如 Bodor 开发了一种基于均匀正交网格的数值积分技术 (numeric integration technique) 用于计算分子的体积及表面积, 并将之作为主要参数用于关联正辛醇 / 水分配系数, 取得了较好的结果^[3]. Eisenhaber^[4]、Silla^[5] 等则引进了科学计算可视化技术中常用的 Delaunay 三角面片法生成原子表面上均匀的球面三角网格, 从而计算分子的表面积, 进而用 Gauss-Ostrogradskii 理论等方法计算分子体积.

本文采用 Monte Carlo 模拟方法计算 van der Waals 分子体积和表面积, 在一定的置信度条件下, 理论上可获得任意置信限内的期望值.

1 计算方法

量子化学或分子力学计算可得到分子中各原子核的坐标, 设为 $x_0(L)$, $y_0(L)$, $z_0(L)$. 对原子 L , 采用 Monte Carlo 方法产生 N_t 个符合球均匀分布规律的伪随机点 $x(i)$, $y(i)$, $z(i)$, 对每个伪随机点, 测试检查是否满足

$$d = \sqrt{(x_0(p) - x(i))^2 + (y_0(p) - y(i))^2 + (z_0(p) - z(i))^2} > R(p)$$
$$i = 1 \cdots N_t; p = 1 \cdots L - 1 \quad (1)$$

这里, $R(p)$ 是 p 原子的 van der Waals 半径. 条件 (1) 保证测试点不属于先前已考察过的任何原子.

设满足 (1) 的伪随机点数为 N , 则原子 L 对分子体积的贡献可由下式估算:

$$V = \frac{4\pi R^3}{3} \left(\frac{N}{N_t} \right) \quad (2)$$

加和所有原子的体积贡献可计算得到分子体积. 重复模拟计算可得到分子体积的均值 \bar{V} , 并在一定的置信度下对均值的置信限作出估计:

$$\mu_V = \bar{V} \pm t(s/\sqrt{n}) \quad (3)$$

类似的算法可用于计算分子的表面积. 首先, 对原子 L , 用 Monte Carlo 方法产生 N_t 个符合球面均匀分布规律的伪随机点 $x(i)$, $y(i)$, $z(i)$, 对每个伪随机点, 测试检查是否满足

$$d = \sqrt{(x_0(p) - x(i))^2 + (y_0(p) - y(i))^2 + (z_0(p) - z(i))^2} > R(p) \quad (4)$$

$i = 1 \cdots N_t, \quad p \text{ 为不等于 } L \text{ 的所有原子}$

条件 (4) 保证测试点不在任何其他原子内.

设满足 (4) 的伪随机点数为 N , 则原子 L 对分子表面积贡献可由下式估算:

$$S = 4\pi R^2 \left(\frac{N}{N_t} \right) \quad (5)$$

加和所有原子的表面积贡献可计算得到分子表面积. 重复模拟计算可得到分子表面积的均值, 并在一定的置信度下对均值的置信限作出估计.

2 球均匀分布和球面均匀分布伪随机点的产生

2.1 球均匀分布

在笛卡尔坐标系中, 均匀分布在半径为 R 的球内的分布密度函数为

$$f(x, y, z) = \begin{cases} \frac{3}{4\pi R^3} & \text{当 } x^2 + y^2 + z^2 \leq R^2 \\ 0 & \text{其他} \end{cases}$$

在球坐标系 (r, ϕ, θ) 中, $0 \leq r \leq R$, $0 \leq \phi \leq \pi$, $0 \leq \theta \leq 2\pi$. 因为 $\frac{\partial(x, y, z)}{\partial(r, \phi, \theta)} = r^2 \sin \phi$, 容易求得其分布密度函数为

$$\begin{aligned} f(r, \phi, \theta) &= \frac{3}{4\pi R^3} \cdot r^2 \sin \phi = \frac{3r^2}{R^3} \cdot \frac{\sin \phi}{2} \cdot \frac{1}{2\pi} \\ &= f_r(r) \cdot f_\phi(\phi) \cdot f_\theta(\theta) \end{aligned}$$

其中

$$f_r(r) = \frac{3r^2}{R^3}, f_\phi(\phi) = \frac{\sin \phi}{2}, f_\theta(\theta) = \frac{1}{2\pi}$$

为了从 $f_r = \frac{3r^2}{R^3}$ 中抽样, 先作变换 $r = R\eta$, 于是有

$$f_\eta(y) = f_r(Ry)R = 3y^2 \quad y \in (0, 1)$$

可采用挑选抽样法^[1]： $Z_\eta = \max(\eta_1, \eta_2, \eta_3)$ ，则 r 的伪随机数序列为

$$Z_r = RZ_\eta$$

为了从 $f_\phi(\phi) = \frac{\sin\phi}{2}$ 中抽样，可令

$$y = \sin^2 \frac{\phi}{2} \text{ 即 } \phi = 2\arcsin\sqrt{y} \quad \text{当 } 0 \leq \phi \leq \pi, 0 \leq y \leq 1$$

因为 $\frac{dy}{d\phi} = \frac{\sin\phi}{2}$ ，容易求得 y 的分布密度函数为 $f(y)=1$ ，即为 $(0, 1)$ 内的均匀分布。故 y 从 $(0, 1)$ 取均匀伪随机数序列 ξ 时，用直接抽样法可得 ϕ 的伪随机数序列

$$Z_\phi = 2\arcsin\sqrt{\xi}$$

很显然， θ 取 $(0, 2\pi)$ 内的均匀分布伪随机数序列。

2.2 球面均匀分布

在笛卡尔坐标系中，均匀分布在半径为 R 的球面上的分布密度函数为

$$f(x, y, z) = \begin{cases} \frac{1}{4\pi R^2} & \text{当 } x^2 + y^2 + z^2 = R^2 \\ 0 & \text{其他} \end{cases}$$

在球坐标系 (r, ϕ, θ) 中， $r=R, 0 \leq \phi \leq \pi, 0 \leq \theta \leq 2\pi$ ，容易求得其分布密度函数为

$$f(r, \phi, \theta) = \frac{\sin\phi}{2} \cdot \frac{1}{2\pi} = f_\phi(\phi) \cdot f_\theta(\theta)$$

抽样可如前述情形进行。根据球坐标系的伪随机点坐标，可很方便地得到伪随机点的笛卡尔坐标系坐标： $x=r\sin\phi\cos\theta, y=r\sin\phi\sin\theta, z=r\cos\phi$ 。

3 计算示例

本文计算体积时，在原子内取 4000 个均匀分布伪随机点，计算表面积时，在原子表面取 1000 个均匀分布伪随机点。原子坐标由 AM1 半经验方法确定。在 99% 置信度条件下，以均值的 0.2% 或 0.5% 作为置信限，计算了 N_2 、 H_2O 分子的体积和表面积，结果如表 1 所示。

表 1 计算结果

Table 1 Calculated results

	$V^a/\text{\AA}^3$	$V^b/\text{\AA}^3$	$V^c/\text{\AA}^3$	$V^d/\text{\AA}^3$	$S^a/\text{\AA}^2$	$S^b/\text{\AA}^2$	$S^c/\text{\AA}^2$	$S^d/\text{\AA}^2$
N_2	23.59	23.61	23.59	23.63	40.96	42.77	40.97	40.87
H_2O	19.34	19.36	19.33	19.38	36.35	37.83	36.31	36.30

^aGeometry method; ^bBodor algorithm, $20 \times 20 \times 20$ grid; ^cThis algorithm, 0.2 % confidence limit; ^dThis algorithm, 0.5 % confidence limit

比较表 1 中同时给出的 Bodor 算法计算结果可以看出，本算法的计算精度特别是对表面积的计算结果明显优于 Bodor 算法。本文试算了文献 (3a) 所列 118 个化合物，最大的分子由 56 个原子组成。Bodor 算法结果与本算法 (置信限为 0.5%) 结果比较，体积的最大偏差为 0.93%，RMS 为 0.41；面积的最大偏差为 9.0%，RMS 为 3.64。分析 Bodor 算法不难看出，该方法固

有的缺陷决定了该方法不可能对表面积作精确估算, 虽然可对体积作较精确的估算, 但无法对计算结果作置信限估计。

显然, 本文建立的算法克服了其它方法的一些缺陷, 能成功地用于精确计算分子的体积和表面积。

参 考 文 献

- 1 Zhu Benren(朱本仁). Introduction to Monte Carlo Method(蒙特卡罗方法引论), Jinan: Shandong University Press(济南: 山东大学出版社), 1987
- 2 Wong M W, Wiberg K B, Frisch M J. *J. Comp. Chem.*, 1995, 16(3): 385
- 3 a) Bodor N, Gabanyi Z, Wong C K. *J. Am. Chem. Soc.*, 1989, 111:3783;
b) Bodor N, Huang M J. *J. Pharm. Sci.*, 1992, 81:272
- 4 Eisenhaber F, Lijunzaad P, Argos P, et al. *J. Comp. Chem.*, 1995, 16(3): 273
- 5 Silla E, Tunon I, Pascual-Ahuir J L. *J. Comp. Chem.*, 1991, 12(9): 1077
- 6 Reid R C, Prausnitz J M, Poling B E. The properties of Gases and Liquids, 4th ed., New York: McGraw-Hill Book Co., 1987

Monte Carlo Simulation Calculation of Volume and Surface Area of Molecule

Shang Zhicai Yu Qingsen Lin Ruisen

(Department of Chemistry, Zhejiang University, Hangzhou 310027)

Abstract The Monte Carlo simulation technique was developed to calculate the volume and surface area of molecules. In the fixed confidence degree, expected values within the specified confidence limit could be obtained. The result of this method was better than that of the Bodor algorithm method.

Keywords: Volume and surface area of molecule, Monte Carlo simulation technique, Bodor algorithm

[书 讯]

《固体表面与界面》

孙大明 席光康编著

本书是研究固体表面与界面的专著。它涉及到固体表面与界面的原子排列、电子结构、电磁声子、吸附与脱附、粒子与表面的相互作用以及表面与界面的分析方法等。对金属和非金属的表面与界面, 以及晶体和非晶体的表面与界面都进行了论述。

本书侧重于我国学者和作者本人近年来在表面科学的某些方面的研究进行概括; 同时也引用了国外一些学者的研究成果。

本书对于基础理论采用有简有繁的描述方式; 对于一些最新的研究成果, 则直接引用。对界面现象作了重点描写; 对表面电子态和表面吸附进行了专题讨论。注重理论联系实际, 基础理论与高新技术的结合。

本书由安徽教育出版社出版, 定价为 19.50 元。购书可与安徽合肥市跃进路 1 号该出版社联系。

Received 1997-04-08, revised 1997-05-30. Correspondent: Shang Zhicai.