

技术与方法

分子表面的计算机表示*

蔡文生 邵学广¹⁾ 潘忠孝 张懋森

(中国科学技术大学应用化学系, 合肥 230026)

摘要 提出了一种用于生成分子光滑表面的新算法。该算法从分布在一个包含整个分子表面的椭球上的三角网络开始, 逐步收缩网络直到所有的三角形最佳贴近分子表面。所使用的收缩包络椭球的技术只要稍加修改就可用于蛋白质空腔的表示。

关键词 分子表面, 表面积, 体积, 分子图形, 分子识别

学科分类号 TP399, Q71

分子表面的表示和计算对于分子的稳定性、分子间的相互识别以及分子三维结构的预测都起着十分重要的作用。溶剂化疏水自由能以及溶剂化自由能和焓都可以通过计算溶剂可及表面积而近似得到^[1~4]。锁和钥匙的表面互补原理是已知蛋白质结合部位, 确定其可能配体的重要依据。配体和受体间的互补性研究表明溶剂可及表面积减少、形状和电荷分布的互补性是决定其亲合性和专一性的主要决定因素^[5]。与此类似, 疏水性互补也能很好地用于建立蛋白质螺旋间的结合模型^[5,6]。蛋白质-蛋白质间的静电和形状互补是蛋白质识别的关键特性^[7~10]。用界面元方法(BEM)计算大分子周围的静电场, 意味着分子表面的正确估计^[11]。在蛋白质工程和蛋白质三维结构预测中, 二级结构单元或折叠结构域如何结合在一起, 对于建立具有蛋白质天然构象的合理模型是至关重要的^[12,13]。

在这些研究中, 人们提出了不同的模型和算法用来表示、计算和分析分子形状^[14~17], 并广泛地应用于对接(docking)问题和蛋白质折叠(folding)问题。

我们提出了一种新的分子表面生成方法: 将分布在椭球上的三角网络映射到所包含的分

子表面上。这种映射过程是通过逐步收缩包络椭球直到所有三角形顶点到达分子表面来实现的。另外, 算法稍加改动: 收缩技术改为扩张技术, 起始包络椭球改为空腔内含椭球, 就可方便地将该算法应用于蛋白质空腔的表示。

1 实现方法

分子表面的概念由 Lee 和 Richards^[18,19]首先提出。它包括三种不同类型(图1): 范德华表面(van der Waals surface), 由一组表示原子的范德华球堆积形成的表面(图1a); Lee 和 Richards^[18]最早提出的溶剂可及表面(solvent-accessible surface), 该表面为当一表示溶剂分子的探针球沿分子表面滚动时, 其球心的运动轨迹所得到的表面(图1b); 以及 Richards^[19]后来提出的分子表面的概念(molecular surface), 该表面为: 当一溶剂分子探针球沿分子表面滚动时所产生的不包含溶剂分子的表面(图1c)。这种分子表面由部分与溶剂探针球相切的原子的范德华表面(接触表面或凸球面)和将这些接触表面光滑地联结

* 中国科学院和国家教委留学回国人员科研资助费资助项目。

¹⁾ 中国科学技术大学化学系, 合肥 230026。

收稿日期: 1997-05-27, 修回日期: 1997-09-15

在一起的一组凹球面(球面三角)和马鞍形面(凹角面)组成. Connolly^[20~22]最先提出了分子点表面和三角划分表面的计算方法.

我们的方法是将分布在椭球上的三角网络,逐步收缩到分子表面上,从而得到三角分割的分子表面.为此要实现以下几步:a.分析分子表面的所有表面元素(凸球面、凹球面

和马鞍形面);b.构造一个包含整个分子表面的起始椭球,并在椭球表面产生一个均匀分布的三角包络网络;c.逐步向内收缩包络表面,直到所有网格点到达分子表面;d.用 Gouraud 阴影算法在图形终端上显示所得到的光滑分子表面.

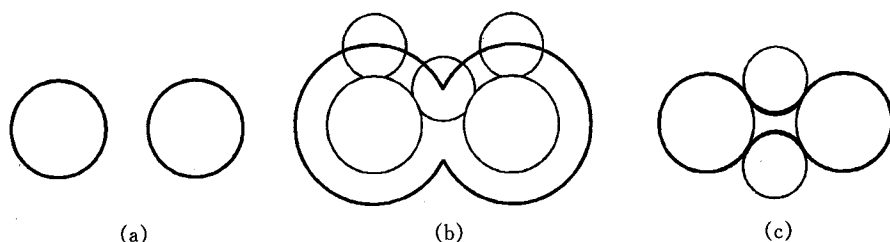


图1 分子表面的三种定义

(a) 范德华表面; (b) 溶剂可及表面; (c) 分子表面.

1.1 分子表面元素

当表示溶剂分子的探针球(半径通常为0.15 nm)沿分子表面滚动时,它的位置有以下三种情况:a.与一个原子相切,在该原子上所经历的表面为其范德华球的一部分,称为接触表面或凸球面;b.与两个原子相切,探

针球上连接两切点的弧线所经历的面为马鞍形面(图2a);c.与三个原子相切,三切点在其面向内球面上形成一凹球面三角(图2b).这三部分表面元素组成分子表面.至此,只要给定探针球半径,就可以确定分子表面的组成.

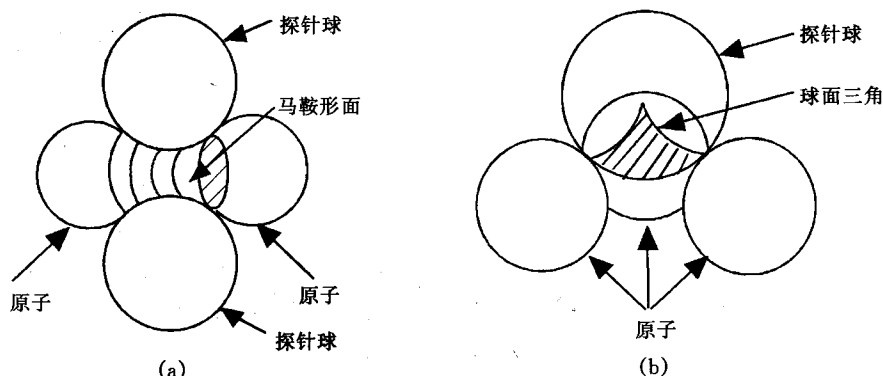


图2 表面元素的定义

(a) 马鞍形面; (b) 凹球面三角.

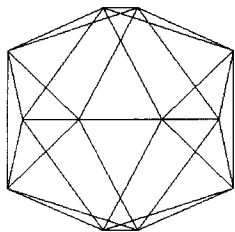
1.2 包络分子表面的椭球

我们所定义的包络分子表面的起始椭球,球心为离分子重心最近的原子范德华球球心,

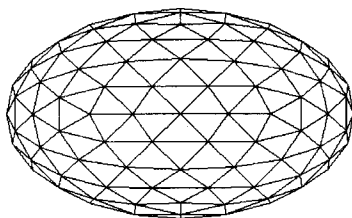
并包括所有原子范德华球的最小椭球.为了得到一个分布在椭球面上的三角网络,我们从一个正二十面体分割球(半径为单位长度,球心

在坐标原点)开始^[23](图3a). 等边三角形的分割是通过连接各边的中点, 从而使顶点分布高度均匀. n 次分割后在球表面上的顶点数和三角形数分别是 $12 + 10 \times (4^n - 1)$ 和 20×4^n . 然后根据椭球球心重新定位并按三个主轴

方向和半轴长做相应的扩展, 最终得到所有三角形顶点都在包络椭球面上的三角网络, (图3b). 分割次数 n 可根据分子的大小确定, 它影响计算的复杂性和所产生表面的光滑程度.



(a)



(b)

图3 分子包络椭球

(a) 起始正二十面体分割球, 顶点数=12, 三角形数=20; (b) 分布在包络椭球面上的三角网络, 顶点数=162, 三角形数=320.

1.3 逐步收缩包络椭球

1.3.1 收缩方向: 起始椭球经过逐步收缩得到一尽可能贴近分子表面的包络面. 每一步中网格点的收缩方向影响到所得到的下一包络表面的光滑性. 我们所采用的收缩方向为当前包络表面在该点的向内法向量, 即该点所有相邻三角形的向内法向量的平均向量. 这一平均向量将作为这一步的收缩方向. 通常情况相邻两点在下一步收缩过程中不会相交. 当一网格点到达分子表面后, 该点在之后的收缩过程中固定在这一位置上, 所有这些固定点将使得每一次收缩朝着最佳分子包络面进行.

1.3.2 与分子表面相交: 在收缩迭代的过程中要计算收缩表面上的三角形与分子表面元素间所有可能的相交.

为了减少计算时间, 我们引进表面元素局部包络球的概念: a. 对于凸球面, 局部包络球即为对应原子的范德华球; b. 对于凹球面三角, 局部包络球球心即为假想的滚动球的球心在形成该凹面的三个原子中心所定义平面上

的正交投影, 半径则为从局部包络球球心到球面三角各顶点的最大距离; c. 对于环形面(马鞍形面), 局部包络球的球心定义为: 与定义该环面的两原子球相切的假想滚动球球心所经历的圆弧的中心, 局部包络球半径为其球心到两原子的最大距离.

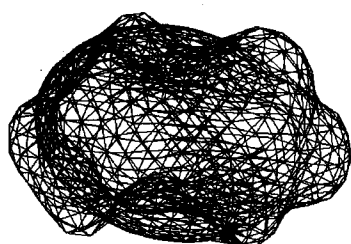
在计算移动后的网格点与移动前的位置构成的线段是否与表面元素相交之前, 首先判断该线段是否与该表面元素的局部包络球相交, 因为只有与表面元素的局部包络球相交才可能与该表面元素相交. 这样就大大减少了空相交的计算时间.

2 结 果

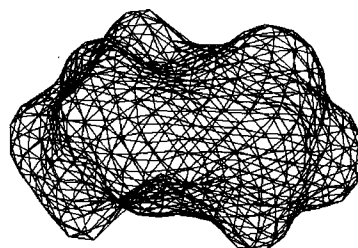
图4给出了用该算法计算一小分子(26个原子)分子表面的逐步收缩过程. 其中, 包络表面上顶点数为642, 三角形数为1280. 收缩步长为0.05 nm, 共需迭代8步. 图4a为第3步所得到的包络表面, 图4b为第8步得到的最终分子表面.

表 1 给出了计算分子表面 (MS) 和溶剂可及表面 (SAS) 所需要的 CPU 时间 (SGI 工作站 IRIS Indy), 分别为表中 MS-CPU 和

SAS-CPU. 三角形数为分布在椭球面上三角网络中的小三角形的个数.



(a)



(b)

图 4 计算分子表面的收缩过程

原子数=26, 顶点数=642, 三角形数=1 280, 收缩步长=0.05 nm, (a) 第 3 步所得到的包络表面;
(b) 最终 (第 8 步) 得到的分子表面.

表 1 从原子坐标生成分子表面(MS)、溶剂可及表面(SAS)
所需的 CPU 时间

原子数	三角形数	CPU 时间(MS)/s	CPU 时间(SAS)/s
33	1280	1.334	0.375
53	1280	2.330	0.539
243	5120	24.440	13.363
327	5120	69.446	33.502

注: 其中探针球半径=0.15 nm, 收缩步长=0.14 nm.

作用.

参 考 文 献

- 1 Chothia C. Structural invariants in protein folding. *Nature*, 1975, **254**: 304~308
- 2 Smith K C, Honig B. Evaluation of the conformational free energies of loops in proteins. *Proteins: structures, functions, and genetics*, 1994, **18**: 119~132
- 3 Perrot G, Cheng B, Gibson K D *et al.* MSEED: a program for the rapid analytical determination of accessible surface areas and their derivatives. *J Comput Chem*, 1992, **13**: 1~11
- 4 Kurochkina N, Lee B. Hydrophobic potential by pairwise surface area sum. *Protein Engineering*, 1995, **8** (5): 437~442
- 5 Miller J, Janin J, Lesk A M *et al.* Interior and surface of monomeric proteins. *J Mol Biol*, 1987, **196**: 641~656
- 6 Chothia C. The nature of the accessible and buried surfaces in proteins. *J Mol Biol*, 1976, **105**: 1~12
- 7 Norel R, Lin S L, Wolfson H J *et al.* Molecular surface complementarity at protein-protein interfaces, The critical sparse points in docking. *J Mol Biol*, 1996, **252** (2): 263~273
- 8 Jones S, Thornton J M. Principles of protein-protein interactions. *Proceedings of the United States of America*, 1996, **93** (1): 13~20
- 9 Meyer M, Wilson P, Schomburg D. Hydrogen bonding and molecular surface shape complementarity as a basis for protein docking. *J Mol Biol*, 1996, **264** (1): 199~210
- 10 Lawrence M C, Colman P M. Shape complementarity at protein/protein interfaces. *J Mol Biol*, 1993, **234** (4): 946~950
- 11 Xiang Z, Shi Y, Xu Y. Calculating the electric potential of macromolecules: a simple method for molecular surface trian-

3 结 论

本文所介绍的方法包括分子表面元素的分析、分子包络椭球的生成以及包络面的逐步收缩过程. 分子表面元素分为凸面、凹面 (球面三角) 和环面 (马鞍形面), 这些面之间是连续的. 分布在包络椭球面上的三角网络从分子的原子坐标计算而得. 包络面上的网格点逐步向内收缩最终得到最佳贴近的分子表面. 收缩方向影响所得到的表面. 我们将这种技术成功地应用于蛋白质分子空腔 (cavity) 的表示. 首先在感兴趣的空腔内定义一起始椭球, 然后沿向外法向量 (指向空腔表面) 逐步向外扩张分布在椭球上的三角网络. 这种方法的最终应用是计算分子的球面调和表面 (harmonic surface), 用以研究分子表面运动和分子间相互

- gulation. J Comp Chem, 1995, **16** (4): 512~ 516
- 12 Zehfus M. Improved calculations of compactness and a reevaluation of continuous compact units. Proteins: structures, functions, and genetics, 1993, **16**: 293~ 300
 - 13 Eisenstein M, Shariv I, Koren G *et al.* Modeling supramolecular helices: Extension of the molecular surface recognition algorithm and application to the protein coat of the tobacco mosaic virus. J Mol Biol, 1997, **266** (1): 135~ 143
 - 14 Perrot G, Maigret B. New determinations and simplified representations of macromolecular surfaces. J Mol Graph, 1990, **8**: 141~ 144
 - 15 Sanner M F, Olson A J, Spehner J-C. Reduced surface: an efficient way to compute molecular surfaces. Biopolymers, 1996, **38** (3): 305~ 320
 - 16 Cossi M, Mennucci B. Analytical first derivatives of molecular surfaces with respect to nuclear coordinates. J Comput Chem, 1996, **17** (1): 51~ 73
 - 17 You T, Bashford D. An analytical algorithms for the rapid determination of the solvent accessibility of points in a three-dimensional lattice around a solute molecule. J Comput Chem, 1995, **16** (6): 743~ 757
 - 18 Lee B, Richards F M. The interpretation of protein structures: estimations of static accessibility. J Mol Biol, 1971, **55**: 379~ 400
 - 19 Richards F M. Areas, Volumes, Packing and protein structure. Annu Rev Biophys Bioeng, 1977, **6**: 151~ 176
 - 20 Connolly M L. Analytical molecules surface calculation. J Appl Cryst, 1983, **16**: 548~ 558
 - 21 Connolly M L. Solvent-accessible surfaces of proteins and nucleic acids. Science, 1983, **221** (4612): 709~ 713

22 Connolly M L. Molecular surface triangulation. J Appl Cryst, 1985, **18**: 499~ 505

23 Dietrich A, Maigret B. Program for the visualization of inorganic crystals. J Mol Graph, 1991, **9**: 85~ 90

Representation for the Molecular Surface. CAI Wen-sheng, SHAO Xue-guang¹⁾, PAN Zhong-xiao, ZHANG Mao-Sen (*Department of Applied Chemistry, University of Science and Technology of China, Hefei 230026, China;* ¹⁾ *Department of Chemistry, University of Science and Technology of China, Hefei 230026, China*).

Abstract A new method for the representation of smooth molecular surface was proposed. The embracing surface, which started from a triangular mesh built on an ellipsoid embracing the whole molecular surface, is deflated stepwise until all the triangles will stick the surface of the molecule. This method can be easily used to get the cavity of the protein.

Key words molecular surface, surface area, volume, molecular graphics, molecular recognition

一种改进的分离嗜中性白细胞的方法^{*}

贝 玲 胡天惠 沈 恂

(中国科学院生物物理研究所, 北京 100101)

摘要 报道一种从全血中分离嗜中性白细胞的方法. 人血经过葡聚糖沉降, 淋巴细胞分离液梯度分离, 特殊分离液的洗涤和红细胞的溶胀后, 得到了嗜中性白细胞. 经瑞特氏染色和台盼蓝染色证明细胞纯度与存活率均在 95% 以上. 以化学发光和细胞色素 c 还原法证明了细胞的高活力与膜受体的完整性. 与国外的分离方法相比较, 该方法简便易行, 且分离效果好, 是一种高效、经济的分离方法.

关键词 嗜中性白细胞, 分离提纯, 呼吸爆发

学科分类号 Q2-33

嗜中性白细胞属于小吞噬细胞, 约占体内白细胞总量的 70%, 在人体内起着吞噬外来细菌及异物, 抵御疾病的重要作用. 因此, 对嗜中性白细胞的研究在基础及临床上都有着重

要意义. 然而, 从全血中分离嗜中性白细胞却存在一定困难: 由于嗜中性白细胞的密度与红

^{*} 国家自然科学基金资助项目 (39670205).

收稿日期: 1997-05-28, 修回日期: 1997-09-15