

生体高分子表面積の高速計算アルゴリズム*

3G-5

升屋正人 岡本祐幸

分子科学研究所 理論研究系

1. はじめに

タンパク質や核酸など生体高分子の立体構造予測に関する研究では、溶媒（主として水）の効果を誘電率として暗に取り入れているに過ぎず、水和自由エネルギーを計算しているわけではなかった。このため、最小エネルギー構造とネイティブ構造は必ずしも一致していない。また、RISM理論やスケール粒子理論など explicit に溶媒を扱う水和自由エネルギーの計算方法では、長い計算時間を必要とするため、シミュレーションへの導入は現実的ではないとされている。そこで、われわれは水和自由エネルギーを高速に計算することでこれらの問題を回避し、分子量の大きな生体高分子を対象として、水和自由エネルギーをも含んだ立体構造予測シミュレーションを行うことを目指している。

水和自由エネルギーは、分子の溶媒接触表面積 (solvent accessible surface area) に比例するものとし、それぞれの原子団ごとに経験的方法により求められているパラメータを用いて計算することで、従来の方法に比べて高速に計算できる。原子団 i の溶媒接触表面積を A_i 、パラメータを σ_i とすると、分子の水和自由エネルギー E_s は以下の式で表される。

$$E_s = \sum_i \sigma_i A_i \quad (1)$$

この方法では分子の表面積を高速に計算することにより水和自由エネルギーを高速に計算することが

可能となる。そこで、本研究では分子表面積の高速計算アルゴリズムを提案し、生体高分子の表面積と水和自由エネルギー計算に応用した例を示す。

2. 点の分布

分子の表面積の計算方法には解析的なものと数値的なものの2つが提唱されている。解析的なものは正確ではあるが、計算に時間がかかる、場合によっては発散し表面積を求めることができない、といった問題がある。このため、われわれは数値的なものを採用することとした。数値的な方法では球の表面に点を分布させ、点の数を数えることにより表面積を求める。このため点の分布の方法が精度に大きく影響することが知られている。

球の表面に均等に分布させることのできる点の数は最大12個である。これ以上の点を分布させる場合には、均等に点を分布させるアルゴリズムが必要となる。正20面体の各面を構成する3つの辺に囲まれた球面を三角形パッチで分割するもの (icosahedral tessellation)[1] などいくつかのアルゴリズムが提唱されているが、われわれはモンテカルロシミュレーションを用いて最適な点の分布を決定することで、より精度の高い表面積の計算を可能とした。

球の表面に分布している点が均等であるかどうかは、以下のポテンシャル関数 U を評価することで明らかにすることができる。

$$U = \sum_{x < y} (180 - \beta_{xy})^2 \quad (2)$$

ここで、 β_{xy} は表面上の点 x と y の立体角である。 U

*Fast calculation of biomolecule surface areas
Masato Masuya and Yuko Okamoto
Institute for Molecular Science
Myodaiji, Okazaki, Aichi 444, Japan

が最小となるような点の分布を求めることで精度の高い表面積計算が可能となる。通常のエネルギー最小化アルゴリズムではエネルギー極小値に留まるとい問題があるが、われわれが開発した新しい方法を用いることで、エネルギー極小値に留まらずに、最小値を求めることが可能となった。

3. 表面積の計算

分子の表面積は以下の手順で求める。

1. 分子に含まれている各原子 i の中心の座標 (x_i, y_i, z_i) と半径 r_i を与える。
2. すべての原子 i について、原子 $j (i \neq j)$ によって覆われている表面上の点をマークする。
3. マークされていない点をカウントし、以下の式で表面積 A を算出する。

$$A = 4\pi \sum_i r_i^2 \frac{n_i}{N} \quad (3)$$

ここで、 n_i は、原子 i の表面上の点のうち、マークされていない点の数、 N は一つの球上に分布させた点の数である。

本研究では計算の初期段階で空間を立方体に分割することで高速化を図った。単位立方体の一辺の長さは、原子半径のうち最大のものの2倍とする。各立方体についてそれ自身とその近傍の26の立方体、あわせて27の立方体の中だけを考慮することにより計算の効率化が可能となる。これは、その他の立方体に含まれている原子が注目している立方体中の原子と交わることはないためである。

4. 表面積と水和自由エネルギーの算出

水和自由エネルギーの算出には、溶媒分子の平均半径 (1.4 Å) を原子の van der Waals 半径に加えた半径を r_i に用いた溶媒接触表面積を用いる。Human parathyroid hormone fragment(残基数:37, PDB

ID: 1HPH) と Protein G(残基数: 56, PDB ID: 1PGA) について、表面積と水和自由エネルギーを求めた結果を表1に示す。

表1: 表面積と水和自由エネルギー

分子	A (Å ²)	E_s (kcal/mol)	Time (s)
1HPH	3705.9	-98.1	0.142
1PGA	3644.1	-117.6	0.193

パラメータ σ_i としては、Ooi ら [2] のものを用いた。計算時間は DEC AlphaStation 500/500(Alpha21164A 500MHz) での所用時間を示している。立体構造予測シミュレーションでは数十万回の再計算が必要となるが、われわれの方法を用いることにより、現実的な時間で水和自由エネルギーを考慮したシミュレーションが可能であることが示された。

参考文献

- [1] Frank Eisenhaber, Philip Lijnzaad, Patrick Argos, Chris Sander, and Michael Scharf. The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of Computational Chemistry*, 16(3):273-284, 1995.
- [2] Tatsuo Ooi, Motohisa Oobatake, George Némethy, and Harold A. Scheraga. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proceedings of the National Academy of Science of the United States of America*, 84:3086-3090, 1987.