

溶媒接触表面積 (SASA: Solvent Accessible Surface Area)

山内 仁喬

2021 年 1 月 31 日

1 溶媒接触面積について [1]

タンパク質と直接接触している水分子の数や、タンパク質の特定の部分と接触している水分子の数に興味があることがある。このような場合に、どの程度溶質が溶媒に露出しているかを測定する手段の一つが、溶媒接触表面積 (Solvent Accessible Surface Area) を計算することである。

ここでは、溶質 (例えば、タンパク質) が水中に存在している状況を考える。つまり、溶質分子の周りには多くの水分子が隙間なくギッチリと詰まっている。プローブ球を溶質分子のファンデルワールス表面に接するように、満遍なく転がした時、プローブ球の中心点の軌跡が描く表面を溶媒接触表面と定義する。その面積が、溶媒接触表面積である。水分子の半径を 1.4 \AA とすれば、溶媒接触表面積は溶質分子に含まれる、各原子のファンデルワールス半径に 1.4 \AA を加えた半径を持つ球の集合体の表面積と言い換えられる。

1.1 溶媒接触表面積の計算方法

溶媒接触表面積の具体的な計算方法として、

- 球面上にランダムに点を打ち、露出した点を数えることで表面積を推定する、モンテカルロ的な方法
- 球面上に均一に点を割り振り、球面を多面体で近似し (例えば、正 20 面体)、多面体を構成する幾何構造の面積から近似する方法 (DSSP プログラム)[2]
- ガウス・ボンネの定理を用いた厳密解による計算法
- 円盤状に切断して近似計算する Lee-Richards の方法 (NACCESS プログラム)[3]

が挙げられる。

1.1.1 モンテカルロ法による溶媒接触表面積の計算方法

溶媒分子に含まれる原子が球であると仮定して、原子 A, 原子 B, 原子 C, ... が配置している系を考える。

■Step.1 原子 A の座標を原点に持ち、原子 A のファンデルワールス半径にプローブ球の半径を足した、半径 $r_{A,\text{probe}} = r_{\text{LJ}}^A + r_{\text{LJ}}^{\text{probe}}$ を持つ球の球面上にランダムな点を複数打つ。原子 A のファンデルワールス半径にプローブ球の半径を足した半径を、便利のために溶媒接触半径と呼ぶこととし、その球表面を原子 A の溶媒接触表面と呼ぶことにする。プローブ球として水分子を用いる場合、 $r_{\text{LJ}}^{\text{probe}} = 1.4 \text{ \AA}$ と設定することが多い。球面座標は

$$x = r \sin \theta \cos \phi \quad (1)$$

$$y = r \sin \theta \sin \phi \quad (2)$$

$$z = r \cos \theta \quad (3)$$

であるから、半径 r の球面上に一様に分布する点の座標 (x, y, z) は、 z 軸の $\cos \theta (\equiv u)$ と角度 ϕ の値をそれぞれ -1 から 1 、 0 から 2π の一様分布から取り出すことで決定できる。つまり、2 つの一様乱数

$$u \in [-1, 1] \quad (4)$$

$$\phi \in [0, 2\pi) \quad (5)$$

を用いると、

$$\sin \theta = \sqrt{1 - \cos^2 \theta} = \sqrt{1 - u^2} \quad (6)$$

であるから、

$$x = r \sin \theta \cos \phi = r \sqrt{1 - u^2} \cos \phi \quad (7)$$

$$y = r \sin \theta \sin \phi = r \sqrt{1 - u^2} \sin \phi \quad (8)$$

$$z = r \cos \theta = ru \quad (9)$$

のように決定できる。

■Step.2 ランダムに打った点が溶媒に対して露出しているかどうかを判定する。判定には、原子 A と重なり合っている原子のみを考えればよい。つまり、原子 X ($X = B, C, \dots$) に対して、 $r_{A,X} \leq r_{A,\text{probe}} + r_{X,\text{probe}}$ のとき、原子 A と原子 X の溶媒接触半径は重なりを持ち、重なっている領域は溶媒に露出していない。したがって、ランダムに打った点が溶媒接触半径の重なり合う領域に存在しているかどうかを判定すればよい。原子 A の溶媒接触表面上に打った点と原子 X の中心との距離を計算する。この距離が $r_{X,\text{probe}}$ よりも小さい時、原子 A と原子 X との重なり合っている領域に存在しているため、溶媒に露出していない点と定義する。そうでない時、溶媒に露出している点と定義する。なお、 $r_{A,X} > r_{A,\text{probe}} + r_{X,\text{probe}}$ である時は、原子 A と原子 X の溶媒接触半径は重なり合っていないので、原子 X は考える必要がない。原子 A に関する溶媒接触表面積は

$$\text{SASA}(\text{原子 A}) = \frac{\text{原子 A の球面上に打った点の内、露出している点の数}}{\text{原子 A の球面上に打った点の総数}} \times 4\pi r_{A,\text{probe}}^2 \quad (10)$$

と計算できる。

■Step.3 ある溶質分子の溶媒接触面積は、溶質分子内の原子の溶媒接触面積の和

$$\text{SASA} = \sum_{i=1}^N \text{SASA}(i) \quad (11)$$

で計算される。

1.2 疎水性相互作用エネルギーとの関係性

実験的に、炭化水素など疎水的な水分子の疎水性相互作用エネルギーが溶媒接触表面積により近似で比例することが知られている [4, 1]。

1.3 溶媒露出度 (Solvent Accessibility)

溶媒接触面積を用いて、分子がどれだけ溶媒に露出しているかを定量化する場合、大きいアミノ酸残基ほど溶媒接触面積への寄与が大きくなってしまふ。そこで、各アミノ酸の溶媒への露出度を評価するときには、溶媒露出度 (Solvent Accessibility) あるいは相対溶媒露出度 (Relative Solvent Accessibility) と呼ばれる規格化

された値が使用されることがある。溶媒露出度は、

$$\text{Solvent Accessibility}(i_{\text{res}}) = \frac{\text{SASA}(i_{\text{res}})}{\text{SASA}(i_{\text{res_max}})} \in [0, 1] \quad (12)$$

と計算される。ここで、 $\text{SASA}(i_{\text{res_max}})$ はアミノ酸の最大溶媒接触面積である。通常、アミノ酸の最大溶媒接触面積は、二面角 ϕ と ψ を β シート領域になるように設定したペプチド Gly-X-Gly(X は目的のアミノ酸残基) から見積もる。溶媒露出度が大きい値 (1 に近い) ほど、その残基が溶媒に露出していることを意味し、小さい値 (0 に近い) ほど他の残基と接触していることを意味する。また、溶媒露出度の減少は他の残基との接触が増加していることを意味する。

参考文献

- [1] 藤博幸. タンパク質の立体構造入門—基礎から構造バイオインフォマティクスへ—. 講談社, 2010.
- [2] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, Vol. 22, No. 12, pp. 2577–2637, dec 1983.
- [3] B. Lee and F. M. Richards. The interpretation of protein structures: Estimation of static accessibility. *Journal of Molecular Biology*, Vol. 55, No. 3, pp. 379–IN4, feb 1971.
- [4] Y Nozaki and Charles Tanford. The solubility of amino acids, diglycine, and triglycine in aqueous guanidine hydrochloride solutions. *The Journal of biological chemistry*, Vol. 245, No. 7, pp. 1648–52, apr 1970.