

## 糖リガンドから見た糖結合サイトにおける分子間相互作用の幾何学的モデル化の研究

(東工大院・総理工<sup>1</sup>, 国際医福大・薬<sup>2</sup>, 東京理大・DDS 研セ<sup>3</sup>)○ 増田 修<sup>1</sup>, 後藤 了<sup>2,3</sup>, 山村 雅幸<sup>1</sup>

【背景・目的】リガンド分子がタンパク質の結合サイトにおいてどのように認識されるかについて研究をおこなうことはドラッグデザインの要となるため、現在でも様々なアプローチによる研究がおこなわれている。このような研究のうち、これまでの *in silico* ドラッグデザインは定量的構造活性相関 QSAR を基礎にしている。この QSAR に基づいた *in silico* ドラッグデザインではリガンド分子の基本骨格を活性の基礎に据え、それに対する化学修飾が活性に対し加成的（線形的）に影響を及ぼすということを前提とする。そのため、基本骨格を重ね合わせることで目指すリガンド分子の活性を予測する。しかしながら、近年ではこのような加成性が成り立たない事象や、様々な骨格を持つリガンド分子が同一のタンパク質結合サイトに結合する事象などが存在することも知られてきている。[1]

このような加成性が成り立たない事象を扱う場合、単純にリガンド分子の骨格構造だけではなく、化合物の様々な性質などを含めた総合的な情報を考慮する必要があると考えられる。そのため近年では、Fragment-Based ドラッグデザインや機械学習に基づいた性質の予測 [2] など、基本骨格のみに依存しないアプローチによる研究がおこなわれている。本研究では分子間相互作用の幾何学的な分布に着目し、単に基本骨格を重ね合わせるのではなく、リガンド分子からタンパク質結合サイトを見た場合に、分子間相互作用の分布や結合サイトの 3 次元構造がどう見え、それがどう抽象化可能か？ ということを検討し、さらにこのようなアプローチをおこなった場合に、同一のリガンド分子が結合する種々のタンパク質の結合サイトには共通性が見られるかどうかを検討することを目的とした。

【方法】解析対象には単一のリガンド分子が様々なタンパク質結合サイトに結合すること、および第 3 の生命鎖として生命を理解および薬学的に非常に重要だと考えられていること、という 2 つの点から単糖を用いた。結合サイトに糖鎖が結合しているものは、糖鎖を各々の単糖に分解して考えることとした。また、対象とするタンパク質-単糖複合体のデータセットには、文献 [3] で用いられているもののうちの Glucose (Glc) が結合しているものを用いた。なおこのデータには、生体高分子の 3 次元立体構造データベースである Protein Data Bank (PDB) のデータが用いられている。

結合サイトの共通性を考えるためには、各々の結合サイトの相同性を定量化する必要があると考え、結合サイトの形状および大きさに着目した相同性スコアを以下のように定義した。

$$\text{相同性スコア} = \frac{2V_{\text{common}}}{V_A + V_B}$$

$V_{\text{common}}$ : キャビティの共通部分の体積  
 $V_{A,B}$ : 各々の結合サイトのキャビティの体積

相同性スコアの計算には、結合サイトとリガンド分子の間で形成される分子間相互作用をすりあわせた状態での、結合サイトの空隙（キャビティ）およびその共通部分の大きさを用いる。そのためまず、上記で示した使用する糖結合タンパク質の 3 次元立体構造データを PDB から取得した。PDB ファイルに含まれる各原子の 3 次元座標からタンパク質の 3 次元モデルを構築し、リガンド分子の各原子から 10 Å 以内にあるアミノ酸、金属イオンなどを取得した。各原子に電荷を割り当て、リガンド分子の原子-タンパク質の原子間でイオン結合ベクトル、水素結合ベクトルを算出した。

分子間相互作用があると判定されたイオン結合ベクトルおよび水素結合ベクトルを各単糖リガンドの重心を中心とした半径 15 Å の球面に投影した。球面投影をおこなうと図 1 のような図が各結合サイトにおける各リガンド分子ごとに得られた。

2つの球面投影図が最も重なり合うような条件(回転軸)を固有値分解(主成分分析)によって求めた。固有値分解で求めた固有値を用いてリガンド分子周囲の原子を回転させた状態で、原子の3次元座標からメッシュを作成し結合サイトの体積を算出した。このメッシュの作成には計算幾何学で用いられるドロネー分割を利用した。

【結果・考察】 相同性スコアの算出を各リガンド分子についておこなうと、相同性スコアの行列が作成できる。ここで相同性スコアは各結合サイト間の相同性を定量化するものと定義したため、結合サイト間の距離と見なすことが可能であり、以下では相同性スコアの行列を距離行列とする。

この距離行列を、多次元尺度構成法の一つである主座標分析[4]を用いて分析をおこなった。主座標分析は与えられた距離行列から、元のデータ構造を再現させる場合に用いられる手法である。例として都市間の距離から元の地図を再現するということが挙げられる。これにより各結合サイトが空間的にどのような状態で付置しているかを明らかにすることが可能となる。

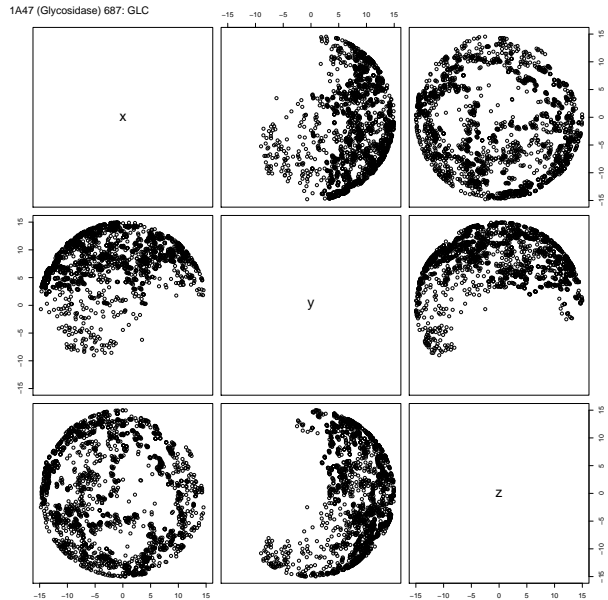


図 1: 球面投影の例 (単位は Å)

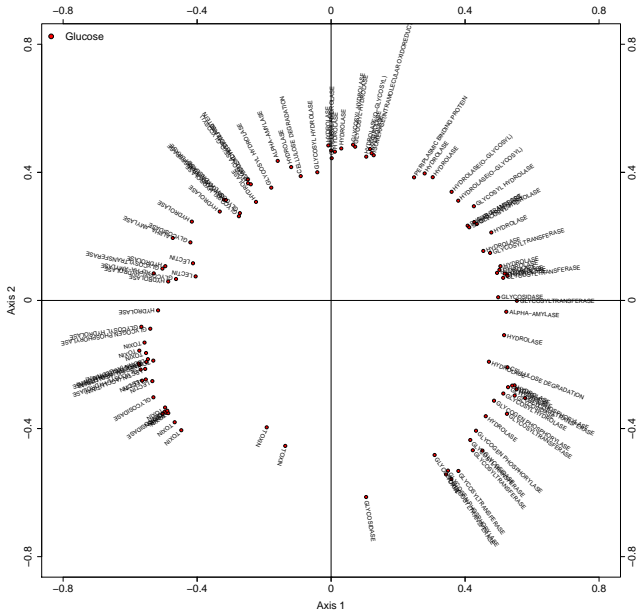


図 2: 主座標分析による Glc の解析結果

を結びつけるだけでなく、結合サイトが持つ生物学的な意義を知ることにつながるということを期待している。

#### 参考文献

- [1] S. J. Teague, *Nature Reviews Drug Discovery* **2**, 527-541 (2003).
- [2] C. M. Dobson, *Nature* **432**, 824-828 (2004).
- [3] C. Shionyu-Mitsuyama, T. Shirai, H. Ishida, T. Yamane, *Protein Eng.* **16**, 467-478 (2003).
- [4] Gower J. C., *Biometrika* **53**, 325-328 (1966).

主座標分析の結果、図2のような特徴的な円環構造が得られた。この円環構造のうちレクチンおよびトキシンは一部分に集まり、それ以外は酵素が並ぶという状況が見られた。酵素にはリガンド分子の構造変化を強いるものが多く、レクチンやトキシンはただ単に分子認識だけをおこなうことから、そのような構造変化があまり必要ない可能性が高いということを考えると、この円環構造はリガンド分子の構造変化をどのように強いるかということを表しているという可能性がある。そのため、もう一度結合サイトの構造およびリガンド分子の結合状態を解析する必要があると考えている。この円環構造のさらなる詳細については現在解析中であるが、このような手法によりこれまでのようなリガンド分子や結合サイトの構造と活性