

フラグメント間相互作用エネルギー行列の非計量多次元尺度構成法を用いた解析 Analysis of Inter Fragment Interaction Energy by Nonmetric Multi-Dimensional Scaling Method

阿部優生¹, 山中雅則²,
Yuki Abe¹, Masanori Yamanaka²

Abstract: Cluster analysis was performed on the interaction energy matrix obtained by the fragment molecular orbital method using the non-metric multidimensional construction method. We performed the analysis about one of the 1TG0-A (PDB ID:1TG0, FMO DB ID: XR38P). The result of inter-fragment interaction energy matrix was used as the distance matrix, and the multidimensional scaling method was executed. The embedding targets we chose are 2D, 3D, 4D, 5D, 6D, 7D. From the result, we observed that some amino acids form an interaction cluster.

1. フラグメント分子軌道法とフラグメント間相互作用エネルギー

フラグメント分子軌道(Fragment Molecular Orbital)法は、巨大高分子の量子化学計算法の1つである。この方法の特徴は、分子を比較的小さなフラグメントに分割し、個々のフラグメントを元にモノマーとダイマーの計算から全エネルギーや全電子密度の算出が可能なことである。

フラグメント間相互作用エネルギー(Inter Fragment Interaction Energy)は、フラグメント分子軌道法から得られるフラグメント間の相互作用エネルギーである。これを元に蛋白質を構成するアミノ酸同士の相互作用を解析する事ができる。このフラグメント間相互作用エネルギーは数値で得られる。数値ではなく、図として可視化したものが図1の左に示した IFIE Map である。図中の縦軸と軸横は共にアミノ酸番号であり、青と赤のプロットはフラグメント間相互作用エネルギーの正負、色の濃さは相互作用の強さを示す。

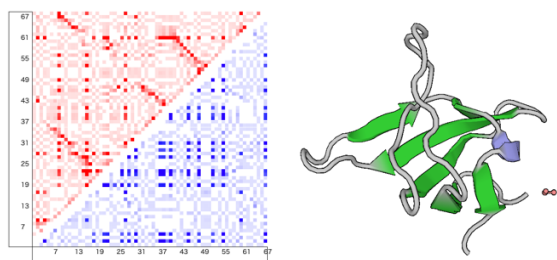


Figure.1 IFIE Map and 3D structure of 1TG0

2. 非計量多次元尺度構成法

非計量多次元尺度構成(Nonmetric Multi-Dimensional Scaling)法は、多変量解析の1つである。2点間の相対距離だけが与えられた点の集合について、その相対距離を反映させながら空間に点を配置する方法である。特徴は、データの相関性を空間的に把握可能なことと、

情報縮約能が大きいことから、数十次元の高次元データを2,3次元に縮約することが可能な事である。この方法は、線形計量多次元尺度構成(Multi-Dimensional Scaling)法の結果に満足しない場合や、データが非線形な関係である場合に他の解析方法よりも有用な場合がある。またここでの非計量とは、データにおいて順位相関のみを考慮するという意味である。

3. 本研究の目的

本研究の目的は、フラグメント間相互作用エネルギーの情報からアミノ酸の相互作用の情報を得ることに対する非計量多次元尺度構成法の有用性について調べることである。標的蛋白質のフラグメント間相互作用エネルギー行列から得られた距離行列に対し、非計量多次元尺度構成法を用いた様々な次元の解析結果のクラスター化の様子を分析し、有用性の判断を行なった。

本研究では具体例として、図1の右に示した標的蛋白質 1TG0-A-Xray13 (FMO DB: XR38P)のフラグメント間相互作用エネルギー行列の解析を行った。データはFMO データベースのものを使用した。解析の手順は以下のとおりである。[6]

1. フラグメント分子軌道法の計算からフラグメント間相互作用エネルギー行列を求める。
2. 得られたフラグメント間相互作用エネルギー行列を変換し距離行列 $D(n, n')$ とする。
3. 距離行列に対し非計量多次元尺度構成法を用い、収束するまで計算を行う。
 - 3-1. 乱数を用い初期位置 $\{\vec{r}_n\}$ とし、 $\sqrt{\sum_n |\vec{r}_n|^2} = N$ となる様に規格化する。 $\{\vec{r}_n\}$ の距離行列を $d(n, n')$ とする。
 - 3-2. 距離行列 $D(n, n')$ と $d(n, n')$ の間の中間に $M(n, n')$ を想定し、 RS を最小にする $M(n, n')$ を求める。

$$RS = \sqrt{\sum_{n,n'} (M(n_1, n_2) - d(n_1, n_2))^2}$$

3-3. $M(n, n')$ と $d(n, n')$ を昇べきの順にソートする.

3-4. $D(n, n')$ が k 番目に大きい時, $D(n, n')$ が何番目に大きいか, という順序を表す S を求める.

3-5. $V = S - k$ を求め, $V \neq 0$ の場合, 空間 R において n を n' に近づく様に以下の計算をし, 動かす. ここで, d は適当な小さな数である.

$$d \sum_{n'} V \frac{\vec{r}_n - \vec{r}_{n'}}{|\vec{r}_n - \vec{r}_{n'}|}$$

3-6. 解が収束するまで繰り返す.

4. 得られたアミノ酸座標を立体空間と Dendrogram に表示させる.

立体空間への配置の自然さは, 配置する空間次元に依存するため, 様々な次元を試した.

4. 結果

2-1 2D, 3D, 4D の配置

2次元及び3次元の解析結果を図2と図3に示した. 4次元の結果は3次元以下の空間に表示できないため, 成分を各々1つ落とした4種類の3次元表示を図4に示す.

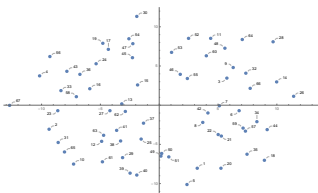


Figure.2 Result in 2D

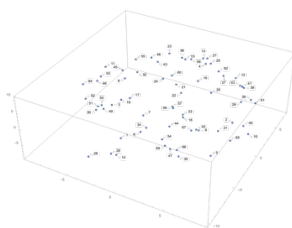


Figure.3 Result in 3D

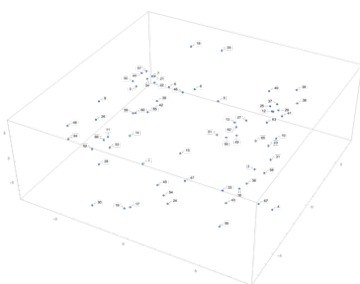


Figure.4 Result in 4D

3-2. 4, 5次元の Dendrogram での結果

4次元以上では空間に図示出来ないために Dendrogram を用いて結果を表示した. 結果を図4に示す. 4次元の立体空間から解析したクラスターは, Dendrogram から得られたクラスター解析結果と一致していた.

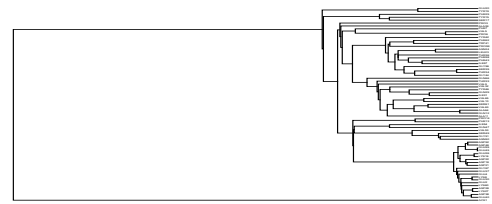


Figure 5 Result in 4D

5. まとめと考察

蛋白質のフラグメント間相互作用エネルギー行列を非計量多次元尺度構成法により調べた. 具体例として PDB ID: 1TG0 に対して行った. フラグメント間相互作用エネルギーは大半が同じ様な値であり, 低次元ではクラスターに分離されにくい結果が得られた. 埋め込み次元を変化させる事でクラスターの傾向を調べた. 2次元では円盤状に分布し, 3次元では球面状に分布した. 4次元以上の表示は不可能であるため, Dendrogram を用いた. Dendrogram の結果は高次元に埋め込むことで結果が改善されていた. 4~7次元では同様の結果が得られた. 図6に Dendrogram で解析したクラスターを立体空間に示した. 以上より, フラグメント間相互作用エネルギー行列の非計量多次元尺度構成法の解析には高次元に埋め込む必要があると考えられる.



Figure.6 Cluster illustration in 3D space

- [1] K, Kitaura et al., Chem. Phys. Lett. 313, 701, (1999).
- [2] C. Watanabe, et. al., Chem-Bio Informatics Journal, 19, 5, (2019).
- [3] Modern Multidimensional Scaling, Springer. New York. EFG(2000)
- [4] A part of this research was done in activities of the FMO drug design consortium (FMODD)
- [5] https://drugdesign.riken.jp/FMODB/detail.php?FMODB_ID=XR38P
- [6] 田口善弘(2001). 「非計量多次元尺度構成法への期待と新しい視点」『数理統計』 Vol.49, No. 1, pp133-153.