

情報科学の生命分子科学への応用：

アミノ酸物性に基づくタンパク質特徴の可視化の試み

(広島市大院) ○津崎祐二、藤原久志、石渡孝

Application of information sciences to biomolecular sciences: An attempt to visualize characteristics of a protein on the basis of physical properties of amino acids

(Hiroshima City University) ○Y. Tsuzaki, H. Fujiwara, and T. Ishiwata

**【序】**我々は、情報科学の生命分子科学研究への応用を目指して、種々の試みに取り組んでいる。本研究では、タンパク質内の物性分布の特徴を可視化し、タンパク質科学研究を支援するソフトウェア開発を行う。

**【物性に基づくアミノ酸のグラフとコドン表】**宮田らは、物性（極性と体積）に基づくアミノ酸同士の「距離」と進化における「置換頻度」に相関があることを示した<sup>1)</sup>。この研究に着想を得て、横軸を相対極性、縦軸を相対体積として、アミノ酸 20 種類の物性のグラフを作成した（図 1）。アミノ酸の物性（図 1）とコドン表<sup>2)</sup>（図 2）には興味深い相関が見られた。すなわち、一文字目の塩基が C、A、G の場合、二文字目の塩基 U（赤）、C（水色）、A（緑）に応じて、アミノ酸の極性が小、中、大と変化することが分かった。また、一文字目が U または二文字目が G のコドン表には、色々な個性を持つアミノ酸が現れている。コドン表における位置とアミノ酸の物性（疎水性、電荷、残基量など）に興味深い相関があることは、これまでも報告されている<sup>2-4)</sup>。本研究では、図 1 のグラフに基づき、タンパク質アミノ酸配列の物性分布を可視化することとした。

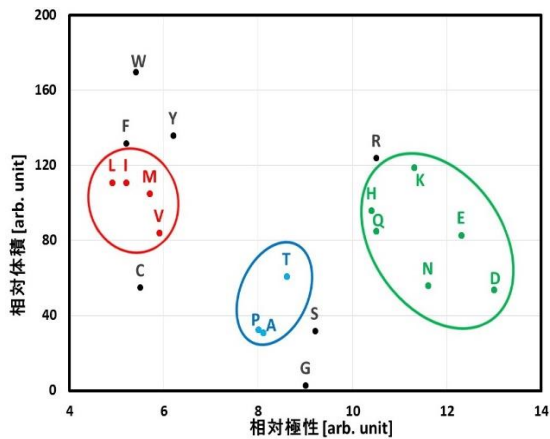


図 1：アミノ酸 20 種類の物性  
(相対極性と相対体積)

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

図 2：コドン表

**【アミノ酸配列物性分布の可視化：移動平均の計算】**図 1 に基づき、物性（極性と体積）に着目して、7 残基ごとに移動平均を計算することとした。すなわち、アミノ酸の相対極性および体積の位置  $i$  における移動平均  $\bar{P}(i)$  および  $\bar{V}(i)$  は、次式で計算できる。

$$\bar{P}(i) = \left( \sum_{j=i-3}^{i+3} P(j) \right) / 7, \quad \bar{V}(i) = \left( \sum_{j=i-3}^{i+3} V(j) \right) / 7$$

また、算出した移動平均 $\bar{P}(i)$ および $\bar{V}(i)$ の分布の特徴を表すため、以下の式を用いて重心( $\bar{P}, \bar{V}$ )を求めた。

$$\bar{P} = \frac{1}{n} \sum_{i=1}^n \bar{P}(i) \quad , \quad \bar{V} = \frac{1}{n} \sum_{i=1}^n \bar{V}(i)$$

**【ソフトウェア開発】** アミノ酸配列の移動平均と重心を計算し、可視化を行うソフトウェアを、Win32 Application Programming Interface (API)関数を主体として作成した。また、プログラム作成環境として、Visual Studio 2010 Professional (Microsoft)を用いた。移動平均と重心の解析結果 (図3の手前) と図1のグラフ (図3の後方) は仮想画面を利用して重ね合わせて表示するようにした。

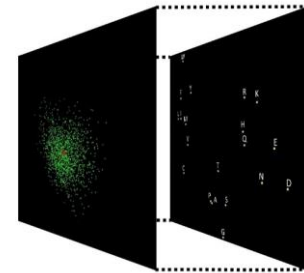


図3: 仮想画面の重ね合わせ

**【膜タンパク質のデータの取得】** 両親媒性インデックス確立に用いられたデータセットから18種類の膜タンパク質<sup>4)</sup>を選択し、これらのアミノ酸配列をUniProt<sup>6)</sup>より取得した。得られたデータは、膜貫通領域配列とループ領域配列とに分別した。

**【膜貫通領域とループ領域の解析】** 図4および5に膜貫通領域およびループ領域のアミノ酸配列の解析結果を示す。膜貫通領域配列の移動平均( $\bar{P}, \bar{V}$ )は、図1の赤色グループと水色グループの間に出現しており、疎水的な性質を示唆している。一方、ループ領域配列の移動平均( $\bar{P}, \bar{V}$ )は図中央付近に出現し、より親水的であること示唆している。タンパク質内の構造および物性 (図3) とコドン表 (図2) が、アミノ酸の物性 (図1) を仲立ちに関連付けられおり、興味深い。

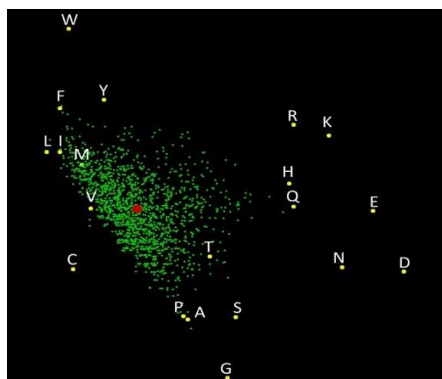


図4: 膜貫通領域アミノ酸配列の相対極性および体積の移動平均 (緑) と重心 (赤)

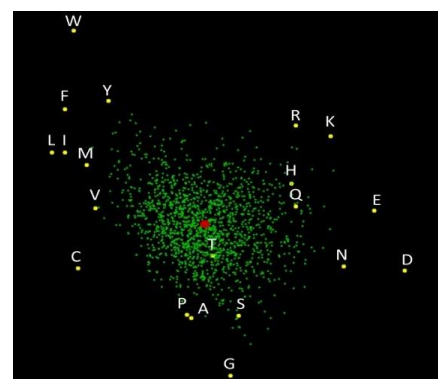


図5: ループ領域アミノ酸配列の相対極性および体積の移動平均 (緑) と重心 (赤)

#### 【参考文献】

- 1) T. Miyata, S. Miyazawa, and T. Yasunaga. *Journal of Molecular Evolution*, **12**, 219-236(1979)
- 2) ヴォート 生化学 (上) 第4版、田宮信雄、村松正實、他訳 (東京化学同人、東京、2012)
- 3) 美宅成樹、生物とは何か?-ゲノムが語る生物進化・多様性・病気-、pp.12-14,46 (共立出版、東京、2013)
- 4) 美宅成樹編、ゲノム系計算科学-バイオインフォマティクスを越え、ゲノムの実像に迫るアプローチ-、金田行雄、笹井理生監修 (共立出版、東京、2013)
- 5) S. Mitaku, T. Hirokawa, and T. Tsuji. *Bioinformatics*, **18**, 608-616 (2002)
- 6) <http://www.uniprot.org/>