

単粒子解析における新画像分類法の確立

開発 タンパク質電顕画像の大規模・高精度分類法の

タンパク質の構造解析法の一つとして、結晶サンプルを必要としない「単粒子構造解析法」が注目されている。この方法は、電子顕微鏡画像より切り出した、さまざまな角度を向いている数千から数万のタンパク質の単粒子画像を、位置や角度毎に分類し、それぞれの分類毎に加算平均処理をすることで、バックノイズを減少させタンパク質の3次元構造を決定することができる。当研究部門では、この生成データをもとに3次元構造を高精度に解析する有効な解析法の開発を行ってきた。この方法により、生物の発生や記憶に深く関与しているIP3レセプターの解析を行い¹⁾、以前本誌で報告した方法で10万枚の粒子画像を二週間で拾い上げ、解析期間の飛躍的な短縮に成功している²⁾。

タンパク質の構造を高精度に解析するためには、粒子画像の角度毎に画像を分類する精度が極めて重要となる。従来の方法では、位置や角度の分類精度が低いため分解能が制限されていた。さらに、粒子画像の主成分分析による情報圧縮をした後に階層型分類を行っているため、あまりにも多くの計算時間とメモリを消費していた。そのため、数十万枚規模の画像分類を行うことは極めて困難であった。

我々は、自己組織化法の一つであるGrowing Neural Gas Network (GNG)を改良することで、

膨大な単粒子画像の分類が極めて高精度にかつ短時間に行えることを見出した。GNG法では、粒子画像を超多次元空間内に存在する分布として捉え、この分布に従い適切に分類するため、従来の画像分類法に比べて精度が高くノイズに強い特徴がある(図1)。さらに、各画像グループのネットワークが自動的に形成され、このネットワーク構造が実際の粒子の投射角度にほぼ対応している。図2では、左側に粒子上部の投射画像があり、右側に行くに従い側面の投射像へと徐々に変化する。さらに、不純物は、粒子のネットワークから突き出すように右上に配置される。こうしたネットワークの情報を用いることで、3次元構造の再構成が容易になり高速化できるものと思われる³⁾。こうして得られた3次元構造から、我々は機械としてのタンパク質の仕組みを知ることが可能となる。

GNGを用いた本手法では、メモリの消費も極めて少ないため100万枚規模の膨大な粒子画像を高速かつ高精度に分類することが可能であり、単粒子構造解析の分解能と解析速度の飛躍的な向上に寄与し、タンパク質の機能解析に貢献するものと思われる。また、この方法は結晶化が困難な膜タンパク質やタンパク質複合体の解析にも応用が可能であるため、今後その重要性はますます高くなるものと期待される。

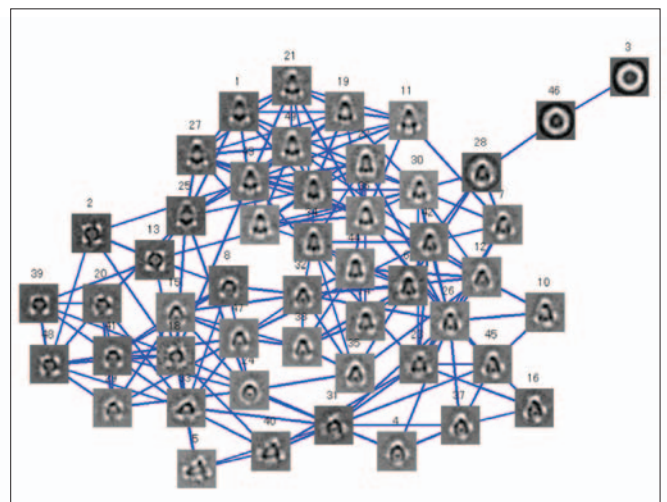
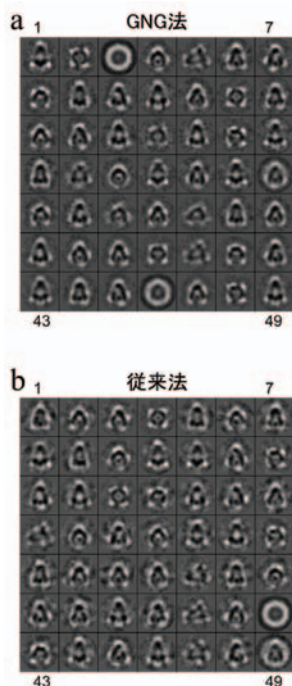


図1(左) (a) GNG法によるNa channelタンパク質の画像分類 (b) 従来法による画像分類

図2(上) GNG法による分類画像のネットワーク構造、タンパク質粒子の投射角度を反映したネットワークが自動的に形成される



おぐらとしひこ
小椋俊彦
t-ogura@aist.go.jp
脳神経情報研究部門

関連情報

- 共同研究者：佐藤主税(脳神経情報研究部門)。
- 1) T. Ogura, C. Sato : J. Struct. Biol., Vol. 145, 63-75 (2004).
- 2) C. Sato, K. Hamada, T. Ogura, A. Miyazawa, K. Iwasaki, Y. Hiroaki, K. Tani, A. Terauchi, Y. Fujiyoshi, K. Mikoshiba : J. Mol. Biol., Vol. 336, 155-164 (2004).
- 3) T. Ogura, K. Iwasaki, C. Sato : J. Struct. Biol., Vol. 143, 185-200 (2003).