

# 1 分子時系列情報に基づく反応ネットワーク構造の構成論とその複雑さ解析

●小松崎 民樹 ◆李 振風 ◆馬場 昭典  
北海道大学電子科学研究所分子生命数理研究分野

## <研究の目的と進め方>

近年の計測技術の飛躍的な進展により、これまで集団平均に埋もれていた生体分子、それらの複合体ならびに機能システムの動的な挙動を1分子レベルで追跡することが可能となり、分子の個別性に由来する構造多型性、長時間記憶などの新たな動態現象が発見されてきた。しかしながら、それらの動的発現原理はまだ解明されていない。本研究課題では、システムを化学反応の複雑ネットワークとして捉え、実際に観測される1分子観察データから背後に存在する化学反応ネットワークを構築する新しい「構成論」を開発する。マクロレベルでは、化学反応ネットワークは各化学成分の濃度を変数とする連立微分方程式として表現されるが、ミクロレベルでは濃度として分子の個性を単純化できる保証はなく、化学反応ネットワークは一般に多次元状態空間上の「アトラクター残骸」間の遷移のネットワークとして表現される。情報理論、不変多様体理論ならびに計算力学に立脚して、ミクロおよびマクロにおける化学反応ネットワーク構造がもつ複雑さおよびランダムネスを定量化し、ミクロとマクロをつなぐ「階層間のつながり」を考察する。また、化学反応ネットワークのトポロジーの動的変化も調べて、ネットワークの可塑性を考察する。

隷属原理の前提そのものが問われるメソスケールで生起する生命現象に対して、(観測される)1分子時系列情報のみから「状態」を規定し、遷移ネットワーク構造を構成するとともに、マクロ(=濃度)からミクロ(=状態空間上のアトラクター残骸)における「状態」概念の関係を考察し、ミクロとマクロを架橋するための新たな概念・研究手法の創出を目指す。

## <2007年度の研究の当初計画>

①分子動力学計算や実験データから背後に潜む多次元有効自由エネルギー地形を構成する手法の開発：前年度に引き続き、各準安定状態における局所平衡が達成される時間スケール、準安定状態から別の準安定状態に遷移していく時間スケールの両者を1分子時系列情報から同定し、多次元自由エネルギー地形が正当化される時間スケールを考察しつつ、構成・可視化する方法論を開発する。今年度は、時系列から局所平衡が達成する時間スケールを直接評価する手法の可能性を考察する。蛋白質の計算機シミュレーション、1分子時系列情報に適用し、多次元自由エネルギー地形の概念の再考、時間スケールに依存した地形変化を調べることを通して、経路に依存した分子記憶を解析する。

②1分子時系列情報から背後に潜む状態空間の遷移ネットワーク構造を再構成する手法の開発：前年度に引き続き、カオス理論、情報理論などに立脚し、局所平衡を予め規定しないで、異なる時空間スケールの(状態空間上の)反応ネットワーク構造を構成する「第一原理的」時系列解析理論を開発する。ウェーブレット多重分解された1分子時系列集団からシンボル化された状態間遷移列を評価し、(動的に揺らぐ)遷移ネットワーク構造を時空間スケールの異なる階層毎に再構成する。今年度は、ネットワークのトポロジー解析方法の開発に主眼を移して、複雑に張り巡らされたネットワーク構造における定常的経路、階層間の情報流の検出手法を開発する。また、状態遷移ネットワークのもつ幾何学的構造から詳細釣り合いならびに分子記憶を評価する新しい解析手法を考察する。具体的に開発した手法を種々の1分子時系列データなどに適用し、1分子時系列情報に立脚した状態遷移ネットワークの多様性の原理を考察する。

## <2007年度の成果>

1) 昨年度開発した、局所平衡を予め仮定しないで、異なる階層

の時空間スケールにおける「状態 (Causal State)」および「状態間遷移ネットワーク」を構成する方法論を発展・拡張させた。この方法では、1分子観察によって得られた時系列を異なる時間スケール毎にウェーブレット多重時間分解し、階層毎にある長さのシンボル時系列を抽出する。そして、その時系列断片間の遷移確率を算出し、遷移確率分布が等しいシンボル時系列の集まりを一つの「状態」として規定する。このとき、各状態は同じ固有の遷移確率分布を持つため、異なる状態間を繋ぐ反応ネットワーク構造を階層毎に構成することができる。蛍光寿命揺らぎの自己相関関数において異常拡散が観測されているフラビン還元酵素の1分子時系列データ (Yangら *Science* 302, 262 (2003)) に対し適用した。時間スケールと共に変化する状態遷移トポロジーを自然な形で(=トッピング的に状態数や状態のつながりを仮定しないで時系列から抽出する形で)再構成し、ネットワーク上のマルコフキネティクスから蛍光寿命揺らぎの自己相関関数を解析的に導出することに成功した。約0.1s以下の時間スケールにおいてブラウン運動を仮定した理論値は実験結果を再現しないが、(1分子時系列から導出した)状態遷移ネットワークに基づく理論値は時間スケールに応じて変化する拡散現象の多様性を理論的に評価することができた(図1参照)。

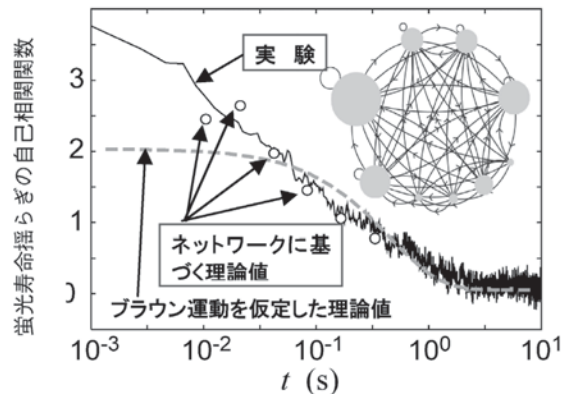


図1. 蛍光寿命揺らぎの自己相関関数(実験)と階層的状态遷移ネットワークから見積もった自己相関関数(理論)

2) 複雑に張り巡らされた状態ネットワークの次数(=ノード(状態)あたりのリンク数)や状態遷移の多様性が時間スケールとともに、どのように変化するかを調べた。図2にフラビン還元酵素の1分子時系列データから評価された32ms、120msおよび480msにおけるネットワーク構造を可視化した。ここで、横軸は各状態に対応する蛍光寿命から推定した色素分子間距離を表し、縦軸は各ノードからそのほかのノードへの相対距離の期待値(値が大きいほど、ネットワークの重心から遠い)を表す。ここでノード間の相対距離は遷移確率分布の類似度(Hellinger距離)を採用した。これはノードの分布が(縦軸方向に対して)コンパクトであればあるほど、各ノードからの遷移パターンが一様になっていくことに相当する。図2より、ブラウン拡散の時間領域に近づくにつれて、ネットワークがコンパクトになり、ノード(状態)ひとつ当たりのリンク数が増大し、状態遷移ネットワークのもつ多様性が減少することが明らかとなった。

さらに、社会学、生態学、WWW、代謝ネットワークなどにおいて重要な概念である“スケールフリー”(Barabasiら *Science*

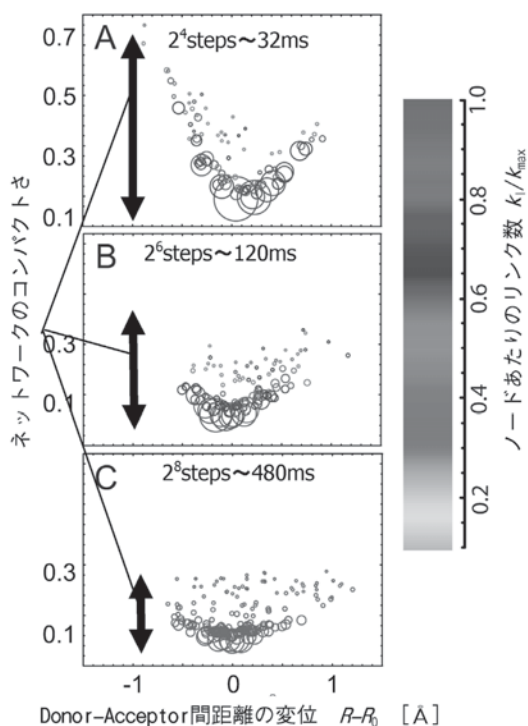


図 2 各時間スケールに応じて変化する状態遷移ネットワーク (理論的には時間無限大の極限で、状態はひとつに収束する)

286, 509 (1999) の存在を解析した。その結果、ブラウン運動で表すことが困難な異常拡散を示す時間領域のネットワークではスケールフリー性を有しているが、通常のブラウン運動として表現される480ms以上の時間領域の時間成分から構成されたネットワークはスケールフリー性を有していないことなどが示された。

3) 生体分子、細胞、組織、そして個体に至る生命システムは常に外界に晒されながら、ミクロレベルでの“刺激”がマクロレベルまで伝達し頑健な機能を作り出している。階層を越えて情報が伝達していくメカニズムを一分子時系列情報に立脚して評価できれば、生体機能の原理を究明することに繋がる。我々は状態遷移ネットワークの情報から、時間スケールの異なる複数の時系列情報のあいだの情報伝達、情報伝達の方向性、強さを移動エントロピー (Schreiber, *Phys. Rev. Lett.* 85, 461 (2000)) を用いて定量化する方法論を開発した。46ビーズモデル蛋白質の折れ畳み温度におけるポテンシャルエネルギー揺らぎ時系列に対して適用し、異なる時間スケールにおける状態間の移動エントロピーを解析した。その結果、時間スケールの近接したネットワーク間の相互相関が一般的に大きいが、情報伝達の大きさは必ずしも双方向的でなく、遅い時間スケール上のネットワーク上の遷移情報が早い時間スケールにおけるネットワーク上の遷移を支配する傾向があること、などが初めて明らかとなった。

4) 計算力学では、手順として (連続量である) 時系列データを離散的な時系列に変換して、離散化 (=シンボル化) された時系列に対して条件付き確率分布が等しいシンボル時系列の集まりを一つの“状態”として定義する。1分子観察においては、時系列データは常に観測ノイズを含むため、シンボル化は必須であることは自明であるが、時系列をシンボル化する際に損失する情報を最小にするためにはどのような手続きが最適であるかは未解決な難問として残っている。我々はHenon-Heiles系と呼ばれる2自由度力学系を例に、得られるネットワークがもつ多様性 (すべての状態の滞在確率に対するシャノン情報量) を最大にするように座標の時系列データをシンボル化すると、時系列情報の背後に存在する相空間構造の非一様性を反映した (ノードの遷移の非等方性/唯一性と局所リヤプノフ指数が良く対応した) ネットワーク構造が再構成できることを見出した。

#### <国内外での成果の位置づけ>

階層を越えた「状態変化」のつながりの産物である生命システ

ムを理解するための理論的なアプローチには、大別して、背後に存在する数理構造を提唱するトップダウン的構成論的手法と微視的な立場からマクロな現象の再現を試みるボトムアップ的還元論的手法が存在する。しかしながら、前者は大胆な仮定や粗視化のために自然と乖離したモデルに陥る可能性が存在する一方で、後者は個々の微視的事象を枚挙するだけでシステム全体を捉えることは困難である。本研究成果は両アプローチを橋渡しする方法論ならびに新規概念を提出するもので、国内外に類を見ない独創的なものである。現在、国内外の1分子分光の実験家から、多くのサンプルが寄せられている。我々の研究手法は1分子観察研究に対する新たなダイナミクス解析基盤を創出するものと期待される。

#### <達成できなかったこと、予想外の困難、その理由>

当初計画①において、(計算機実験で求められた) 残基間距離の時系列情報から得られる有効自由エネルギー地形がどれくらい実際の自由エネルギー地形の形態を再現しているかを理論的に検証した。全体的には良好な対応を見出したが、両者のあいだのズレが、高次元情報を一次元に縮約して見ることによるのか (つまり、観測上の問題)、または、蛋白質の運動が完全に統計的に振舞っていないことによるのか (つまり、系のダイナミクスの問題) を区別して評価することができなかった。自由エネルギー地形概念は、各準安定状態において局所平衡が成り立ち、かつ溶媒からの摩擦効果が無視できる状況において成立する概念であるため、本質的に区別して議論することが困難であるが、なんらかの検証方法を考える必要がある。

#### <今後の課題>

状態空間の階層的遷移ネットワーク構造の可視化技術を確立するとともに、生体ネットワークにおけるスケールフリー性の起源を考察する。環境に適応しながら、時々刻々変化するネットワーク構造の遍歴現象を追跡し、一分子計測に基づいた生体システムの“適応”、“創発”のメカニズムを探る。また、各時間スケールの異なる複数の時系列情報のあいだの情報伝達、情報伝達の方向性、強さおよび階層間の情報伝達のカスケード性を評価する新しい指標を開発し、生体機能におけるミクロとマクロを橋渡しする情報伝達メカニズムを考察する。

#### <成果公表リスト>

1) 論文/プロシーディング (査読付きのものに限る)

- 0708011832 (論文)  
Kinoshita, M., Kamagata, K., Maeda, A., Goto, Y., Komatsuzaki, T., and Takahashi, S., Development of a technique for the investigation of folding dynamics of single proteins for extended time periods, Proceedings of National Academy of Sciences USA, 104(25), 10453-10458 (2007)
- 0801152318 (論文)  
Baba, T., Komatsuzaki, T., “Construction of effective free energy landscape from single molecule time series”, Proceedings of National Academy of Sciences USA, 104(49), 19297-19302 (2007)
- 0801152327 (論文)  
Matsunaga, Y., Li, C.-B., and Komatsuzaki, T., “Anomalous Diffusion in Folding Dynamics on Minimalist Protein Landscape”, Phys. Rev. Lett. 99, 238103(4pages) (2007)
- 0801152323 (論文)  
Li, C.-B., Yang, H., and Komatsuzaki, T., Multiscale Complex Network of Protein Conformational Fluctuation Buried in Single Molecule Time Series, Proceedings of National Academy of Sciences USA, 105, 536-541 (2008)
- 0801152335 (その他)  
松永康佑, 小松崎民樹, “高分子とカオス—異常拡散と階層的規則性—” 高分子 57 (2008) (invited) 印刷中
- 0801241228 (その他)  
白川智弘・郡司ベギオ幸夫・小松崎民樹「惑星科学と非線形科学の接点: 化学反応から生物計算まで」遊星人 16(4) 322-329 (2007) (invited)