

# アミノ酸配列空間上の蛋白質の折れ畳みダイナミクスにおける規則的階層性の探索

●小松崎 民樹

神戸大学理学部地球惑星科学科

## ＜研究の目的と進め方＞

1分子計測技術は、集団平均に埋もれている生体分子、それらの複合体および機能システムにおける生体系のダイナミクスならびにパスウェイに対する数多くの知見を呈示するものと期待されている。近年の1分子計測技術の進展により、エネルギー地形上の経路に関する「静的」な分布の情報に加え、生体分子が構造転移過程のあいだに運動の履歴を覚えているなどの「動的」情報に関する数多くの新知見が得られるようになった。しかしながら、系を構成する全自由度情報を得ることが不可能である以上、(得られる)ひとつないし複数の時系列データを加工・解析し、時系列データの背後に潜む生体系のエネルギー地形およびダイナミクスの構造(=状態空間の構造トポロジー)を抽出する解析手法を確立する必要がある。本研究では、

1. アミノ酸配列及び構造トポロジーと折れ畳みダイナミクスにおける規則性の関係を解析し、規則性に関連する重要なモードの抽出、階層性の探索を可能にする解析手法を開発する。天然構造に折れ畳み容易さとその“拡散的”運動に秘められた規則性の関係を解明する。
2. 「エネルギー地形の違いがダイナミクス・パスウェイにどのような影響を与えるか」を解明し、1分子計測により得られた分子記憶のメカニズムの原理を構造転移パスウェイの多重性・構造多型性の動的構造の観点から考察する。遷移ダイナミクスの規則性ならびにパスウェイ構造(複数の経路を通過する多重性、準安定状態の情報を構造転移の過程で保持するか否かを考察する構造多型性)の視点から、分子レベルで発現する記憶のメカニズムの原理を考察する。
3. 観測される生体分子時系列情報からその背後に潜む系のエネルギー地形およびダイナミクス・パスウェイの動的構造(状態空間の構造・次元性、経路の多重性・構造多型性)を抽出する汎用な解析手法を、物理化学、カオス時系列解析、パターン認識論および情報理論に立脚し開発する。

具体的には、ランダムさを含んだ接触相互作用を加えたビーズモデルに基づく折れ畳みシミュレーションおよび1分子計測などで得られた1分子時系列情報に対して、開発する解析手法を適用し、「エネルギー地形とダイナミクスの相関関係」、「構造多型性、折れ畳み経路の多重性と1分子レベルの記憶の関係」を情報論と分子論の両観点から解析し、生体分子の構造遷移ダイナミクスにおけるパスウェイの動的構造を解明する。また、タンパク質のアミノ酸配列、構造トポロジーとその折れ畳みの規則性がどのように関わり合っているのかを考察し、タンパク質設計に対する新しい指針・概念を創出する。

## ＜研究開始時の研究計画＞

- ①時系列情報からの状態空間再構成手法の開発および規則的階層性の状態空間構造トポロジーの解明

“ランダム”な時系列から大域的な規則構造を取り出す多次元埋め込み論、有限サイズリヤブノフ解析などの手法とタンパク質の動的構造解析の分野で発展してきた主成分解析を融合させて、ランダムさを含んだ接触相互作用を加えたビーズモデルに基づく折れ畳みダイナミクスの次元性・規則性とエネルギー地形との関係を解析する。タンパク質の“拡散的”な運動のなかから規則構造を抽出し状態空間の構造トポロジーの観点から考察する。

- ②多次元エネルギー地形可視化技術の開発と構造多様性指数を導入したエネルギー地形の特徴づけ

主成分、有効次元の抽出だけでは折れ畳みの“拡散的”な運動のなかに隠れた規則性を議論したことにはならない。そこで、重要な数～数十個の主成分で張られる空間に対して、折れ畳みの各段階に対して、その主成分に沿った運動を記述する規則構造を探索し、タンパク質の折れ畳み効率の違いをエネルギー地形、エルゴード性などの観点から解析する。

- ③タンパク質近傍の水の協奏的運動・熱運動の動的効果

タンパク質近傍の水分子配向は決してランダムではなく、ある規則的な場を形成していることが最近指摘されている。この場合、折れ畳みダイナミクスは必ずしも(水中のタンパク質の)ランジュバン型ダイナミクスという描像が正しくなく、周囲の水がタンパク質の折れ畳み易さを支援している可能性が存在する。水の動態場としてのゆっくりした協奏的ダイナミクスと機能、折れ畳みとの関連性を解析する。

- ④1分子時系列情報から背後に潜む多次元状態空間を評価する手法の開発

1分子計測技術は10分の1ミリ秒程度の時間分解能でタンパク質の単分子過程としての折れ畳みダイナミクスを測定することを可能にしてきた。更に数桁時間分解能が向上し、変性状態 $\leftrightarrow$ 天然状態の転移のほかに変性状態および天然状態におけるダイナミクス情報も数値化(例えば、FRET蛍光の輝度)することができる。実験家と密接な意見交換を行いながら、1分子計測データへ開発するカオス時系列解析手法の適用可能性を探る。

## ＜研究期間の成果＞

- ①時系列情報からの状態空間再構成手法の開発および規則的階層性の状態空間構造トポロジーの解明

1分子時系列情報のみから、背後に存在する多次元自由エネルギー地形および状態空間地形を再構成する手法をアラン分散、平均相互情報量ならびに埋め込み次元と主成分解析を融合させた時系列解析理論を開発した。

ガラス状態に成りやすい凸凹な地形を有している46個、3種のアミノ酸残基から構成されるビーズモデル(以下BLNモデルと呼ぶ)と二状態的転移を与える理想的なファネル型地形をもつGo-like BLNモデルの折れ畳みダイナミクスに対し適用し、「エネルギー地形の違いが

ダイナミクスにどのような影響を与えるか」、更には、「1分子計測により得られる（であろう）時系列からそのタンパク質のエネルギー地形が推測可能であるか」を考察した。

その結果、フラストレーションが「より小さい」フェネル型エネルギー地形をもつタンパク質はその転移温度において「より広範囲な」時間領域に渡って「より強い」非定常性を示し、“白色”ノイズ→1/fノイズへの「より先鋭な」転移を伴うこと、すなわち、フラストレーション最小のフェネル型エネルギー地形をもつタンパク質はダイナミクスの記憶をより保持しやすい傾向があること；主成分モードの殆どは、ただか100MDステップで「位置の記憶」を完全に忘れること；転移温度では大振幅主成分モードに沿って記憶が保持される傾向が強く、揺らぎの大きい数十成分の埋め込み次元が顕著に小さいことを見出した。

また、「多数のサドルを通過するダイナミクスにおいて、個々のベイスン間遷移は、多くの場合、系に依らず、“熱”運動のなかに潜む（運動方向の情報を含めた）規則的な経路を辿っている」ことを原理的に証明し、ダイナミクスの情報が頑健に生き残るための状態空間構造トポロジーを解明した。

## ②多次元エネルギー地形可視化技術の開発と構造多様性指数を導入したエネルギー地形の特徴づけ

複雑なタンパク質のキネテックスおよびダイナミクスを議論する上で、3N個の座標（Nは原子数）によって規定される高次元ポテンシャルエネルギー地形の性質を議論することは極めて重要である。系が感じる大域的な多次元エネルギー地形を俯瞰する方法論としては、極小エネルギー構造および第一ランクサドル構造（以後両者を総称して固定点と呼ぶ）の配列情報から構成されるdisconnectivity graphがひろく用いられている。我々は固定点間の計量関係をできるだけ保存するように多次元エネルギー地形を少数自由度空間に投影する新しいdisconnectivity graphを提案し、タンパク質が持つエネルギー地形のトポロジックな特徴を評価する構造多様性という概念を新規に導出した。ここで、構造多様性は、端的に表現すると、エネルギー地形を構成する極小エネルギー構造、鞍部点構造が3N次元空間においてどれくらい広範囲の多次元配位空間に分布しているかを評価するものである。我々は進化の所産として獲得された（とされる）フェネル型エネルギー地形をもつGo-like BLNモデルでは、階層的なエネルギー地形を有するBLNモデルに比べて、少数自由度に縮

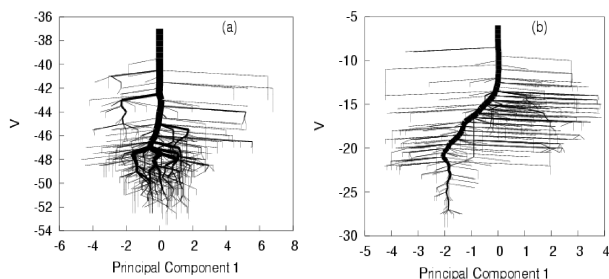


図3 固定点のあいだの距離関係を保存するdisconnectivity graph (a) BLNモデル (b) Go-like BLNモデル  
横軸は固定点間の距離関係を近似的にあらわす最も分散の大きい主成分座標で、縦軸はポテンシャルエネルギー値である。図中、根っここの幅はスーパーベイスンの構造多様性（＝スーパーベイスンに属している固定点の平均構造からのずれの総分散の99%を再現するのに必要な主成分の最小数）を表している。系は根っここの太い経路を辿って構造緩和する確率が高い（と期待される）ため、BLNモデルでは最安定構造に到る以前に複数の分岐路を正しく経由する必要がある（他の同程度のエネルギーをもつ極小エネルギー構造に捕獲されやすい）のに対し、Go-like BLNモデルでは大きな分岐路を経由することなく、最安定構造へ到る確率が高いことがよくわかる。

約可能であるエネルギー領域が最安定構造近傍に存在することを見出し、そのような領域における構造多様性が顕著に低いことを明らかにした。構造多様性が最安定構造付近で小さいという事実は「なぜエネルギー軸上の固定点密度が小さい最安定構造近傍のエネルギー地形が少数の主成分でよく近似され得るのか？」に対する解答を与えるものである。

また、「どの程度の時間経過すれば、運動量空間におけるエルゴード性が成り立っているのか」を調べるために、主成分座標に共役な運動量空間のエルゴード測度を解析した。その結果、温度が低いほど、（他の主成分モードと十分混合せず）1モードあたりの平均運動エネルギーに漸近しない主成分モード数が多いこと、また低温下ではそれらの主成分モードの多くが非ガウスの軌道を示す傾向にあること、などを新規に見出した。

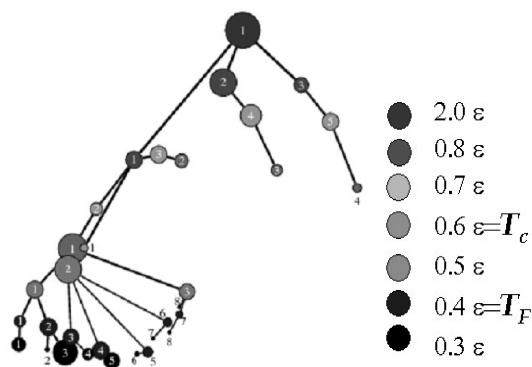
## ③タンパク質近傍の水の協奏的運動・熱運動の動的効果

従来、タンパク質以外の（特に近傍の水和している水分子を含めて）すべての水分子のダイナミクスを「ランダム力」＋「分子摩擦」として「タンパク質＋水」系は論じられてきた。これはタンパク質と水の微視的なダイナミクスの時間スケールが「分離」していることに依拠する。我々は「タンパク質＋水」を力学系の立場から「時間スケールの分離→ランジュバン方程式」の正当性を再評価し「タンパク質にとっての水とはなにか？」を解明することを目的として、ポリアラニンのヘリックスコイル転移ダイナミクスにおける水場の集団運動を解析した。ここで「水場」は、空間固定の立方格子、“細胞”、を（ポリアラニンの構造転移に比べて）高速に”通過する水が残す双極子モーメントベクトルの短時間平均された（遅い）計時変化として評価した。その結果、個々の水和水は微視的には“高速なランダム運動”であっても、粗視化されたメソスケールでは協奏的な運動を呈していること、協調的な集団運動を伴う水場の空間分布がポリアラニンの構造転移の過程で交互に交替していることなどを明らかにした。

## ④1分子時系列情報から背後に潜む多次元状態空間を評価する手法の開発

1分子観測実験では、通常、色素分子の退色効果により状態間遷移が数回しか起こらない短時間の1分子時系列の“集団”のみが得られる。第0近似として各準安定状態において局所平衡が達成されることを仮定し、退色効果などにより十分なサンプルが得られないような“短時間の”1分子時系列の“アンサンブル”から多次元自由エネルギー地形を構成する方法を開発した。この方法は細胞内化学反応ネットワークのもつ普遍的性質を考察する上でよく用いられている「(マクロ量である)濃度を変数とするマスター方程式」を（局所平衡仮定に立脚して）1分子レベルに還元することに相当する。ここで重要なことは（時系列から同定される）1分子構造空間における準安定状態のもつ潜在確率をマクロな濃度の代わりに変数として用いること、ならびに、その準安定状態は“観測される時間幅に依存して“状態数および各状態の（観測量に対する）分布関数の形”も変化することである。また、分布間距離から状態間距離を新たに定義し、計量多次元尺度法によりその距離関係を最大限保存するように多次元自由エネルギー地形を可視化する方法論も開発した。この方法によって、分子動力学シミュレーションを行わなくても1元的な1分子時系列情報のみから、多次元自由エネルギー地形上でのスーパーベイスン間遷移を抽出することができる。現在、1分子計測の実験研究者達との

国内外の共同研究に発展している。



非計量多次元尺度法により2次元平面に投影された(1分子時系列情報だけから構成された)多次元自由エネルギー地形非フェネル型エネルギー地形をもつタンパク質モデルの各温度領域(0.4εは折れ畳み温度、0.6εはCollapse温度)におけるFRET時系列情報から、準安定状態の数、各状態における自由エネルギー値および状態間距離を構成したもの。図中、円*i*の半径は各温度における準安定状態*i*の(相対)自由エネルギーに対応し、半径が大きいほど自由エネルギーが低く、その配置は状態間距離の関係をできるだけ保存している。各温度間における準安定状態の「距離」が一番近いもの同士を実線で結んでいる。たとえば、温度がCollapse温度(0.6ε)から0.7εと上昇すると(その温度における)最安定状態が(0.6ε①)→(0.7ε④)と大きく、自由エネルギー地形上のスーパーベイスンを遷移していることが時系列情報のみから表現できることがわかる(同様の説明が折れ畳み温度においても成り立つ)

#### 〈国内外での成果の位置づけ〉

2001年から現在に掛けて、国内8件(日本物理学会、日本生物物理学会、電子情報通信学会ソサイエティ大会、生理研研究会、分子研研究会、統数研研究会等)、海外5件(Telluride Workshops(コロラド, 米), CECAM Workshop(リヨン, 仏), Max-Planck Institut fur Physik Komplexer Systeme(ドレスデン, 独)等)の招待講演を当該研究課題に関連して行った。また、John-Wiley & Sons, Inc.から化学反応・生体分子ダイナミクスにおける非統計性・選択性に関する特集号Adv. Chem. Phys. 130(2005)をシカゴ大学R.S. BerryおよびS.A. Rice両教授らと共に編・出版した(1224頁)。本研究課題は1分子時系列情報から「多次元エネルギー地形上における構造多型性、折れ畳み経路の多重性とダイナミクスの関係」を解明するための新たな「メソスケールの理論化学」の枠組みを提供することに成功した。現在、当該研究で開発した解析手法は平成16年から始まったJST/CRESTプロジェクトにおいて更に改良を重ね、状態間遷移の非統計性、遷移状態概念の再考、熱揺らぎとタンパク質ダイナミクスのあいだの競合・協同性などを解明する研究に進展している。

#### 〈達成できなかったこと、予想外の困難、その理由〉

Steven Chu教授から長時間分子記憶を保持するヘアピンリボザイムの1分子時系列データ(Zhuang et al., Science, 296, 1473 (2002))を入手することができたが、当該研究課題で開発した時系列解析手法を適用し、状態空間を再構成するに至らなかった。これは、実験的に得られる1分子時系列は色素分子の退色効果により、埋め込み解析を行う上で十分な長さをもつ時系列情報を入手することが本質的に困難であるためである。そのため、短時間の1分子計測データのアンサンブルから背後に潜む状態空間を構成する必要がある。このような制約下においても威力を発揮する多次元自由エネルギー地形を構成する方法論を、近年、開発したので、今後の発展が期待さ

れている。

#### 〈今後の課題〉

短時間の1分子時系列のアンサンブルに対して、局所エルゴード性を予め規定せずに状態空間構造を再構成する動的解析理論を開発し、多次元自由エネルギー地形を構成する静的理論とともに、実験により得られる1分子時系列のアンサンブルに適用し、「エネルギー地形とダイナミクスの相関関係」、「構造多型性、折れ畳み経路の多重性と1分子レベルの記憶の関係」を考察する。

#### 〈研究期間の全成果公表リスト〉

##### 1) 論文/プロシーディング(査読付きのものに限る)

- 0602031252  
Tamiki Komatsuzaki, Kyoko Hoshino, Yasuhiro Matsunaga, Gareth J. Rylance, Roy L. Johnston and David J. Wales 'How Many Dimension is Required to Approximate Potential Energy Landscape of A Model Protein?' Journal of Chemical Physics 122, 084714 (2005).
- 0602031255  
Tamiki Komatsuzaki, Kyoko Hoshino and Yasuhiro Matsunaga, 'Regularity in Chaotic Transitions on Multi-Basin Landscapes' Advances in Chemical Physics 130B, 257-313 (2005) (invited)
- 0602031303  
Tamiki Komatsuzaki and R. Stephen Berry, 'Regularity in Chaotic Transitions on Two-Basin Landscapes' Advances in Chemical Physics 130A, 143-170 (2005) (invited)
- 0602031306  
Y. Matsunaga, K.S. Kostov and Tamiki Komatsuzaki, 'Hierarchical Regularity in Multi-Basin Dynamics on Protein Landscapes' AIP conference series for Slow Dynamics in Complex Systems 708, 302-305 (2004)
- 0602031322  
Y. Matsunaga and Tamiki Komatsuzaki, 'Protein Folding Dynamics: Ergodic Behavior in Principal Component Space' AIP conference series for Slow Dynamics in Complex Systems 708, 342-343(2004)
- 0602031325  
K. Hoshino, Y. Matsunaga, M. Miller, D.J. Wales and Tamiki Komatsuzaki, 'A Coarse-Graining of Energy Landscapes of Proteins -Structural Stability of the Most Stable States-' AIP conference series for Slow Dynamics in Complex Systems 708, 344-345(2004)
- 0303251128  
Yasuhiro Matsunaga, Konstantin S. Kostov and Tamiki Komatsuzaki, 'Multi-Basin Dynamics in Off-Lattice Minimalist Protein Landscapes' Journal of Physical Chemistry A 106, 10898-10907 (2002) (invited)
- 0303251140  
Tamiki Komatsuzaki and R. Stephen Berry, 'Chemical Reaction Dynamics: Many-Body Chaos and Regularity' Advances in Chemical Physics 123, 79-152(2002) (invited)
- 0303251148  
Tamiki Komatsuzaki and R. Stephen Berry, 'A Dynamical Propensity Rule of Transitions in Chemical Reactions' Journal of Physical Chemistry A 106, 10945-10950(2002) (invited)

10. 0303251124

小松崎民樹、松永康佑「タンパク質フォールディングのダイナミクス—異常拡散と階層的規則性」生物物理42,285-289(2002) (invited)

11. 0303251157

笹井理生・小松崎民樹・戸田幹人「ファネル談義—なぜ蛋白質は一義的に、かつ効率的に折れ畳むのでしょうかね—」物性研究78,475-509(2002)