

生成系人工知能による高速 AFM データの高解像度化技術の開発

京都大学大学院理学研究科 准教授 寺川 剛 *

1. はじめに

高速原子間力顕微鏡 (AFM) イメージングは、生体分子の動態をナノメートルスケールかつリアルタイムで観察可能な強力な手法であり、近年、タンパク質の構造変化や分子機械の動作原理の解明において重要な役割を果たしている。一方で、高速 AFM は測定速度と空間分解能の間にトレードオフが存在し、取得される画像はしばしばノイズや分解能の制約を含む。そのため、観察された構造から分子の詳細な立体構造や状態遷移を定量的に解釈することは依然として困難である。

このような課題に対し、近年発展が著しい生成系人工知能 (AI) を活用し、低分解能データから高分解能構造情報を再構成する試みが注目されている。本研究では、実験データと分子シミュレーションを統合することで、AFM 画像から分子構造を推定する枠組みを構築し、その基盤技術の確立を目指した。具体的には、溶液中 AFM 観察によって得られたコンデンシン分子の構造像に対し、粗視化分子動力学シミュレーションにより生成した構造アンサンブルを対応付けることで、画像情報から分子構造を再構成する手法を開発・適用した。さらに、機械学習に基づく粗視化ポテンシャルの構築にも取り組み、単量体の構造情報のみから多量体形成を再現可能なデータ駆動型モデルの開発を行った。これらの研究は、生成系 AI による構造推定の基盤となる「データ生成」と「構造空間の学習」の両面を補完するものである。

本報告では、まず AFM データと分子シミュレーションの統合による構造推定手法の実証として、コンデンシンの ATP 加水分解サイクルに伴

う構造変化機構の解明について述べる。続いて、機械学習を用いた粗視化モデル開発の成果を報告し、最後に本研究の意義と今後の展望について述べる。

2. 高速 AFM と分子シミュレーションの統合によるコンデンシン構造変化機構の解明

2.1. 研究背景と目的

コンデンシンは、染色体の凝縮や分配に必須な分子モーターであり、ATP 加水分解に伴う構造変化を通じて DNA ループ形成を駆動することが知られている^[1]。しかし、その分子機構、特に ATP 結合・加水分解の各段階においてどのような構造状態を取り、それらがどのように相互に遷移するのかについては、依然として不明な点が多い。このような構造ダイナミクスを直接観察する手法として高速 AFM は有力であるが、取得される画像は空間分解能やノイズの制約を受けるため、観察像から分子の詳細な三次元構造を一意に決定することは困難である。本研究では、この課題を克服するために、高速 AFM データと分子シミュレーションを統合することで、観察像から分子構造を推定する枠組みを構築することを目的とした。具体的には、溶液中 AFM によって取得したコンデンシンの画像に対し、粗視化分子動力学シミュレーションにより生成した構造アンサンブルを対応付けることで、各画像に対応する分子構造を推定し、ATP 加水分解サイクルに共役した構造変化機構の解明を目指した^[2]。

* (助成採択時) 京都大学大学院 理学研究科 生物科学専攻 助教

2.2. 高速 AFM と分子シミュレーションの統合

手法

本研究では、高速 AFM で取得された低分解能画像から分子構造を推定するために、分子シミュレーションと統合した解析手法を構築した。まず、出芽酵母由来のコンデンシン複合体について、ATP、ATP アナログ (AMP-PNP、ATPcS)、および ADP 存在下で溶液中 AFM 観察を行い、異なるヌクレオチド状態における構造像を取得した。次に、これらの AFM 画像に対応する分子構造を推定するため、Smc2/4 二量体に対する粗視化分子動力学シミュレーションを実施し、ヘッドの会合・解離およびヒンジの開閉状態を組み合わせた条件で構造アンサンブルを生成するとともに、AFM 画像へのフィッティングにより一致度 (coverage score) に基づいて最適構造を選択した。このようにして、実験で得られた AFM 画像とシミュレーションにより生成された構造アンサンブルを対応付けることで、「画像から構造を推定する」枠組みを実現した。本手法は、低分解能観察データに対して構造空間を探索し最適構造を同定する点で、生成系人工知能に類似したアプローチと位置づけられ、高速 AFM データの高解像度化に向けた基盤技術となる。

2.3. ヌクレオチド依存的構造変化の観察結果

高速 AFM 観察の結果、コンデンシン複合体はヌクレオチド状態に応じて顕著に異なる構造をとることが明らかとなった。まず、ヌクレオチド非存在下 (Apo 状態) では、ヒンジが閉じ、コイルドコイルアームが折れ曲がったコンパクトな構造 (One-form) がほぼ全ての分子で観察され、この構造は既報のクライオ電子顕微鏡構造とよく一致した。一方、ATP アナログである AMP-PNP や ATPcS 存在下では、アームが大きく開いた O-form が頻繁に観察され、ATP 結合に伴うヒ

ンジ開放が示唆された。また、B-form や I-form も一部で観察され、ATP 結合状態では複数の構造状態が共存することが明らかとなった。さらに、ADP 存在下ではヘッドが解離した V-form が多く観察されるとともに、ヘッド間距離が著しく拡大した新規構造 (Hat-form) が同定された。Hat-form は ATP 加水分解後の状態に特有であり、従来の研究では明確に捉えられていなかった構造である。以上の結果から、コンデンシンはヌクレオチドの結合および加水分解に応じて複数の構造状態を取り、それらの分布が大きく変化することが明らかとなった。これらの観察結果は、構造ダイナミクスが ATP 加水分解サイクルと強く結びついていることを示す直接的な証拠である。

2.4. 構造フィッティングと統計解析による機構解明

本研究では、高速 AFM 像から得られた分子形態を定量的に解釈するため、粗視化分子動力学シミュレーションによりヘッドの会合・解離およびヒンジの開閉状態を組み合わせた構造アンサンブルを生成し、AFM 画像へのフィッティングと統計解析により、アーム領域の一致度 (coverage score) に基づいて最適構造を同定した。

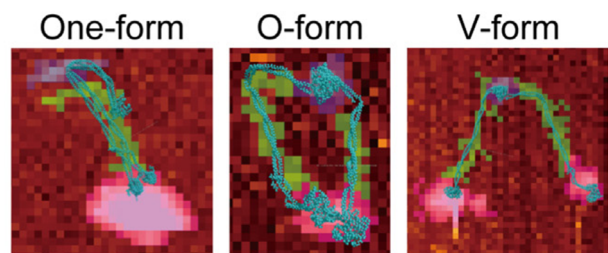


Fig. 1 コンデンシン分子構造の AFM 像へのフィッティングの代表例。粗視化分子動力学シミュレーションコンデンシン分子のコンフォメーションをシアンで示している。

その結果、ヌクレオチド存在下の AFM 像を再現するにはヒンジが開いた構造が不可欠であり、

ATP 存在下ではヘッド会合とヒンジ開放が協調的に誘起される一方、ADP 存在下ではヘッド解離状態が優勢となり、V-form および Hat-form が多くの画像を説明した。以上の結果は、ATP 結合に伴うヘッド会合とヒンジ開放の共役、および ATP 加水分解後のヘッド解離と構造再編成という一連の構造変化機構を示すものである。本研究は、AFM 観察像と分子シミュレーションを統合することで、コンデンシンの ATP 駆動型構造ダイナミクスを分子レベルで定量的に解明したものである。

2.5. コンデンシンの ATP 駆動構造変化モデル

以上の観察および統計解析の結果に基づき、コンデンシンの ATP 加水分解サイクルに共役した構造変化モデルを提案する。まず、ヌクレオチド非存在下では、ヒンジが閉じたコンパクトな One-form が安定状態として存在する。ATP が結合すると、ヘッドの会合が誘起されると同時にヒンジが開いた状態へと遷移し、O-form あるいは I-form が形成される。この状態ではアームが柔軟に開閉し、DNA の取り込みに適した構造をとると考えられる。その後、ATP の加水分解に伴いヘッドは解離し、V-form へと遷移する。さらに一部の分子では、ヘッド間距離が大きく拡大した Hat-form が形成され、この状態では Brn1 の一部が解離している可能性が示唆される。これらの結果は、ATP 結合によるヒンジ開放とヘッド会合、加水分解後のヘッド解離という一連の構造変化が段階的に進行することを示している。本モデルは、コンデンシンが ATP 駆動的に構造状態を切り替えることで DNA ループ形成を実現するという枠組みを支持するものであり、特にヒンジ開閉とヘッド状態の共役が DNA 取り込みおよび放出の制御に重要な役割を果たすと考えられる。

2.6. 本研究の意義

本研究は、高速 AFM 観察と分子シミュレーションを統合することで、分子構造ダイナミクスを定量的に解釈する新たな枠組みを提示した点に意義がある。特に、低分解能かつノイズを含む AFM 像に対し、構造アンサンブルとの対応付けによって分子構造を推定する手法は、高速 AFM データの高解像度化に向けた基盤技術となる。また、本研究により、コンデンシンの ATP 加水分解サイクルに共役した構造変化機構が明らかとなり、染色体構造制御の分子機構理解に新たな知見を与えた。さらに、本手法は他の動的タンパク質複合体にも適用可能であり、広範な生体分子の機能解明に寄与することが期待される。加えて、本研究のアプローチは生成系人工知能による構造推定と本質的に共通する枠組みを有しており、今後の AI 技術との融合により、より高精度な構造再構成手法へと発展する可能性を持つ。

3. 機械学習粗視化モデルによるタンパク質構造アンサンブルの再現と多量体形成の予測

3.1. 研究背景と目的

天然変性領域 (IDR) を含むタンパク質は多様な構造アンサンブルを取りながら機能を発現し、転写制御や液-液相分離 (LLPS) に重要な役割を果たす。これらの機構理解には構造分布とその遷移を記述できる分子動力学 (MD) シミュレーションが有効であるが、全原子 MD は計算コストの制約から大規模系や長時間ダイナミクスの解析に限界がある。本研究では、この課題を解決するため単量体の全原子 MD から得られた構造分布を教師データとして機械学習により粗視化力場を構築し、単量体情報のみから多量体形成に伴う構造変化の再現を検証するとともに、構造多様性と会合過程を統一的に記述する枠組みの確立を目指した。

3.2. 機械学習粗視化モデルの構築

本研究では、タンパク質の構造アンサンブルを再現する粗視化力場の構築に拡散モデルに基づく機械学習手法 (Two-for-one; TFO) を用いた^[3]。TFO モデルは全原子分子動力学シミュレーションから得られた構造分布を教師データとし、その確率分布の勾配 (スコア) を学習して力場を導出する。本研究では GCN4 ロイシンジッパー単量体の全原子 MD から得られた構造アンサンブルを 1 残基 1 ビーズに粗視化して学習に用い、各残基を 1 粒子として分子内相互作用をデータ駆動的に記述することで、参照構造に依存する Gō モデルと異なり特定構造を過度に安定化せず構造アンサンブル全体を再現できるモデルを構築した。さらに、多量体形成の再現に向けて分子内相互作用に TFO 力場、分子間相互作用に Gō ポテンシャルと排除体積ポテンシャルを組み合わせたハイブリッドモデルを用い、単量体情報に基づく力場によるタンパク質会合過程の再現を試みた。

3.3. 単量体構造アンサンブルの再現

構築した TFO 力場の妥当性を検証するため、GCN4 ロイシンジッパー単量体の粗視化分子動力学シミュレーションを行い全原子 MD で得られた構造アンサンブルと比較した結果、ヘリックスとコイルが混在する分布を概ね再現し単量体の構造多様性を表現できた。PCA および tICA 解析でも粗視化 MD の分布は全原子 MD とよく一致した。一方、AICG2+力場ではヘリックス構造が過度に安定化され部分的な構造ゆらぎを十分に再現できなかったが、TFO 力場は参照構造に依存せず構造アンサンブル全体を反映することでヘリックス-コイル間の多様性を適切に記述できた。以上の結果は機械学習に基づく粗視化力場が単量体の複雑な構造分布を再現できることを示している。

3.4. 多量体形成の予測

次に、単量体の構造アンサンブルのみを学習した TFO 力場を用いて GCN4 ロイシンジッパーの二量体形成を再現できるかを検証した。分子内相互作用には TFO 力場、分子間相互作用には Gō ポテンシャルと排除体積ポテンシャルを用いたハイブリッドモデルにより粗視化分子動力学シミュレーションを行った結果、初期状態で離れて配置した 2 本の単量体は自発的に接近・会合しコイルドコイル様の二量体構造を形成した。この過程では会合に伴いヘリックス構造が安定化し、単量体で部分的にほどけていた構造が秩序化されることが示され、「coupled folding and binding」を再現できた。さらに PCA および Q スコア解析からも二量体の構造分布は全原子 MD に近づくことが確認され、単量体情報のみから学習した力場が多量体形成を予測し構造遷移の物理的傾向を捉えていることが示唆された。

3.5. 学習データ依存性の検証

次に、構築した TFO 力場の学習データ依存性を検証した。単量体 MD から高ヘリックスフレームを除外して再学習した結果、70%以上除外では単量体分布と二量体でのヘリックス安定化は維持された一方、60%以上除外では単量体分布は再現されたものの二量体でのヘリックス安定化は失われた。これらは、二量体形成の再現に一定量のヘリックス情報が必要であり、TFO の表現能力がデータの量と質に依存することを示す。

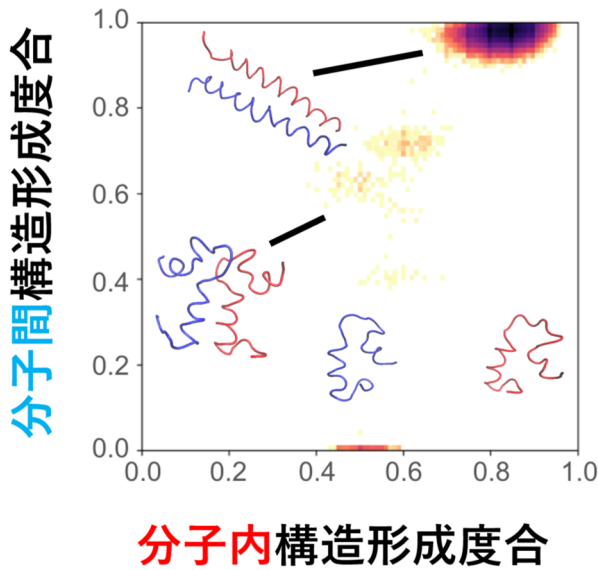


Fig. 2 分子内および分子間の構造形成度合いに基づく状態分類。横軸は分子内構造形成度合い、縦軸は分子間構造形成度合いを示す。ヒートマップは各状態の出現頻度(または確率密度)を表している。

3.6. 本研究の意義

本研究は、機械学習により構造アンサンブルから粗視化力場を構築し、単量体の情報のみから多量体形成を予測できることを示した点に意義がある。従来の粗視化モデルは参照構造に依存し構造多様性の表現に限界があったが、本手法は構造分布を学習することで IDR に特徴的な構造ゆらぎを含むアンサンブルを再現でき、学習データに含まれない多量体形成を再現したことからモデルが物理的傾向を捉えていることを示唆し、今後の汎用的な分子モデリング手法の発展につながると期待される。

4. 結論

本研究では、「生成系人工知能による高速 AFM データの高解像度化技術の開発」という目的のもと、実験と計算を統合した新たな分子構造解析手法の構築と、その応用を行った。まず、高速 AFM

観察と粗視化分子動力学シミュレーションを組み合わせることで、低分解能 AFM 像から分子構造を推定する枠組みを確立し、コンデンシンの ATP 加水分解サイクルに共役した構造変化機構を明らかにした。この結果、ヘッド会合とヒンジ開放の共役、および加水分解後のヘッド解離という一連の構造ダイナミクスが定量的に示された。さらに、機械学習に基づく粗視化モデルの開発により、単量体の構造アンサンブルのみから多量体形成を予測できることを示し、データ駆動型アプローチの有効性を実証した。これらの成果は、実験データとシミュレーションを融合することで情報の解像度を高める新しい方法論を提示するとともに、生成系人工知能と分子シミュレーションの統合による次世代の構造解析技術の基盤を築くものである。

5. 参考文献

- [1] Rutkauskas M, Kim E; In vitro dynamics of DNA loop extrusion by structural maintenance of chromosomes complexes, *Current Opinion in Genetics & Development*, 90, (2025), 102284.
- [2] Hiroki Koide, Noriyuki Koder, Shoji Takada, Tsuyoshi Terakawa; Solution AFM Imaging and Coarse-grained Molecular Modeling of Yeast Condensin Structural Variation Coupled to the ATP Hydrolysis Cycle, *Journal of Molecular Biology*, 437 -17, (2025), 169185.
- [3] Arts M., Garcia Satorras V., Huang C. W., Zügner D., Federici M., Clementi C., Noé F., Pinsler R., van den Berg R.; Two for One: Diffusion Models and Force Fields for Coarse-Grained Molecular Dynamics, *Journal of Chemical Theory and Computation*, 19 - 18, (2023), 6151–6159.

6. 本研究にかかわる成果発表等

- ・ Hiroki Koide, Noriyuki Kodera, Shoji Takada, Tsuyoshi Terakawa; Solution AFM Imaging and Coarse-grained Molecular Modeling of Yeast Condensin Structural Variation Coupled to the ATP Hydrolysis Cycle, Journal of Molecular Biology, 437 -17, (2025), 169185.

7. 謝辞

京都大学理論生物物理学研究室のメンバーの皆様には、本研究を通じて多くの議論とご支援を賜り、心より感謝申し上げます。また、貴重なご議論をいただいた Christian Haering 教授、ならびにタンパク質精製を行っていただいた寺川 まゆ博士に深く感謝いたします。本研究は、公益財団法人島津科学技術振興財団研究開発助成の支援を受けて実施されました。