

科学研究費助成事業 研究成果報告書

令和元年6月5日現在

機関番号：82401

研究種目：新学術領域研究(研究領域提案型)

研究期間：2014～2018

課題番号：26119006

研究課題名(和文) マルチレゾリューション法を用いたタンパク質複合体の高解像度動的解析

研究課題名(英文) Molecular dynamics simulations of protein complexes based on multi-resolution models

研究代表者

杉田 有治 (Sugita, Yuji)

国立研究開発法人理化学研究所・開拓研究本部・主任研究員

研究者番号：80311190

交付決定額(研究期間全体)：(直接経費) 115,250,000円

研究成果の概要(和文)：実験等で得られる低解像度構造情報と分子動力学などで得られる高解像度シミュレーション情報を組み合わせた新しい解析手法を開発した。具体的には、クライオ電子顕微鏡と分子動力学を組み合わせたフレキシブルフィッティングにおけるアンサンブル計算や並列化手法、一分子解析と分子動力学を機械学習で組み合わせたモデリング手法などで、新規性の高い開発である。これらは分子動力学ソフトウェアGENESISの機能として導入し、効率の良い構造探索手法やマルチスケール分子モデルと組み合わせて計算を行うことができる。開発した手法を膜タンパク質やタンパク質・核酸複合体などの動的構造の解析に応用した。

研究成果の学術的意義や社会的意義

本研究は、シミュレーションなどの理論的な手法だけでは解決できない問題を実験データをうまく組み合わせることで克服する新しいアプローチである。また、クライオ電子顕微鏡の利用が今後ますます発展することを考慮すると、そのモデリングを分子動力学で高解像度化できることは大きな意義がある。実際、製薬企業などでクライオ電子顕微鏡を用いた構造解析を薬物との相互作用を理解するために導入し始めており、それらの解析に貢献することができる。開発した手法はGENESISに導入し、今後一般公開していくため、広いユーザーが利用することができる。

研究成果の概要(英文)：We developed new methods combining experimental information with simulation data to obtain reliable structure-dynamics relationship of proteins. Generally, experimental data tend to be low-resolution and low-dimension, whereas simulation can provide high-resolution structural information. In our methods, we combined cryo-EM with MD simulation in flexible fitting calculation and utilized machine learning for combining single-molecule analysis with MD simulation. The developed methods were installed into GENESIS software package and can be used together with enhanced sampling methods and multi-scale/multi-resolution molecular models. We applied these methods to understand structure-dynamics relationships of membrane proteins or protein/nucleic acid complexes.

研究分野：生物物理学

キーワード：分子動力学 クライオ電子顕微鏡 一分子構造解析 フレキシブルフィッティング マルチスケール・マルチレゾリューション 膜タンパク質

様式 C - 19, F - 19 - 1, Z - 19, CK - 19 (共通)

1. 研究開始当初の背景

蛋白質や核酸などの生体分子は生命における機能素子であり、その機能を十分に理解するためには立体構造を明らかにすることが重要である。生体分子のX線結晶構造解析は飛躍的に進歩し、多くの蛋白質等の立体構造が原子解像度で明らかになっている。その中にはリボゾームなどの蛋白質核酸複合体や膜輸送体などの膜蛋白質が含まれ、立体構造から機能の理解が進んだものもある。しかし、このような構造解析で解明されたのは結晶場の中での静的な構造情報であり、蛋白質や核酸が機能する細胞という環境の中での動的な構造は未だに未解明の事が多い。そのため、蛋白質などの動的構造を解明するために様々な計測手法が開発されてきた。高速AFMやin-cell NMR、一分子蛍光測定などのように世界に先駆けて国内の実験家が新しい手法を開発してきた例も少なくない。しかし、動的な構造を計測しようとするると必然的に結晶中の静的な構造解析と比べて低解像度のデータになるため、方法の妥当性の検証を行うことが困難であった。

2. 研究の目的

実験で得られた低解像度データを、計算機シミュレーションなどを通して動的な立体構造情報に変換することが理論的には可能である。例えば、X線小角散乱(SAXS)では散乱プロファイルが実験データとして得られるが、結晶構造を出発構造とした分子動力学計算を行い、その結果を用いて散乱プロファイルを計算し実験データと直接比較することで溶液中の構造情報を得ることができる。本研究課題ではこの例のように、実験データと計算機シミュレーションを組み合わせることで高解像度の動的構造解析を実現することを目指す。

本研究課題で主に用いるシミュレーション手法は分子動力学計算である。郷ポテンシャルに代表される粗視化分子モデルを用いた長時間ダイナミクスや、レプリカ交換法などの拡張アンサンブル法に代表される方法は、効率の良い構造探索と自由エネルギー解析を可能にした。本研究では、複数の構造間を遷移する大規模構造変化をシミュレーション可能な計算手法の開発を行う。そのために、並列化効率の良い分子動力学ソフトウェアを利用するだけでなく、異なるモデルと手法を組み合わせたマルチレゾリューション法を導入する。機能発現において大規模な構造変化を行う蛋白質は必然的に分子量が大きくなるため、単ドメインからなる蛋白質というよりは生体超複合体や膜蛋白質などの柔らかい構造を持つ蛋白質を研究対象とし、実験・計測データを評価しうる長時間ダイナミクスを解析する。

3. 研究の方法

実験から得られる構造情報には、低解像度データ(例えば電子顕微鏡やSAXS, FRETなどの蛍光一分子測定)と高解像度データ(例えばX線結晶構造解析やNMR)が存在する。低解像度データは巨大な生体超分子や複合体などの構造の概観を得るのに用いられる事が多く、また、動的な構造情報を与える。一方で高解像度データは原子構造を与えるが結晶場の影響を強く受けた静的な情報である。分担者のTama Florence と連携研究者の宮下治は異なる解像度を持つ様々な構造情報を組み合わせた構造最適化に豊富な経験を有するため、その手法を杉田チームで開発中の分子動力学シミュレータGENESIS に導入する。研究代表者の杉田は分子動力学ソフトウェアのみならず、レプリカ交換分子動力学法(REMD)などの効率の良い構造探索アルゴリズムや自由エネルギー計算法などの開発と応用計算を行ってきた。本研究ではスーパーコンピュータ「京」などの利用に適した超並列分子動力学ソフトウェアに効率の良い構造探索アルゴリズム等を導入し、複数の構造間を遷移できるアルゴリズムの開発を行う。また、分子を記述するモデルとして、全原子モデルのみならず、複数の原子を束ねた粗視化分子モデルを組み合わせる。さらには、X線結晶構造解析では困難な水素イオンの位置を予測できる方法等も組み合わせ、長時間の分子運動を正確に

(すなわち高解像度に)解析できる手法を開発する.このような手法を「マルチレゾリューション法」と呼び,複数の分子モデルとアルゴリズムを統合的に利用できるソフトウェア環境を GENESIS 上で構築する.これらの開発は,連携研究者の森貴治, Jaewoon Jung や小林千草などと共同で行う.

4. 研究成果

(1) 分子動力学ソフトウェア GENESIS の開発と GPU 計算機への最適化

我々は,スーパーコンピュータ「京」などの多くの CPU を並列化することで計算を加速するソフトウェア GENESIS を開発し,巨大な生体分子系(100 万を超える原子を含む分子系)のシミュレーションを実現した.そのために, OpenMP と MPI を組み合わせた Hybrid 並列化を実現することで多くの CPU を用いた場合にも並列化効率が下がらない工夫をした.その一方,中規模(数十万原子を含む)分子系を計算するために, GPU を有効に活用できるアルゴリズムの開発に取り組んだ.我々の開発において,最も計算時間のかかる非共有結合相互作用(短距離力)を GPU で計算し,それ以外の演算は CPU で行うモデルを考案し, GENESIS に導入した.この計算モデルの利点は, AMBER などの GPU プログラムと異なり,複数の GPU 計算機を並列化して用いることができることである.そのため,従来の GPU プログラムでは難しかった巨大な生体分子系についても並列 GPU 計算機を用いてシミュレーションすることが可能になった.

(2) タンパク質の大規模構造変化を記述する最小自由エネルギー経路の探索

タンパク質が機能を発現する際に, Open form と Close form など異なる立体構造へと変化する例が多く見られる. X 線結晶構造で明らかになっているのは少数の状態に限定されるので,それらを接続する計算手法が必要である.様々な手法が提案されているが, 2 つの異なる立体構造を接続する最小自由エネルギー経路を探索するアルゴリズム(String method)が信頼性の高い解を与えることがわかってきた.しかし,従来の String method はいくつかの解決すべき問題を含んでおり,様々な分子系に応用するためにはさらなる改良が必要であった.我々は GENESIS に String method を導入し, 2 つの立体構造変化を記述する反応座標の次元依存性について検討した. Adenylate Kinase の Open-Close forms の構造変化に注目し,粗視化分子モデルを用いた String method を実行した.ここでは主成分解析を行い,少数の主成分を反応座標として用いた.その結果,主成分の数が少なすぎる場合には Open-Close forms の間に存在する Semi-Close form を記述することができないことが明らかになった.構造変化の遷移状態を記述するためには,少なくとも 10-20 程度の主成分を用いる必要があった.この手法とプログラムを用いて,膜タンパク質であるカルシウムイオンポンプやヘムトランスポーターの大規模構造変化のシミュレーションに取り組み,その自由エネルギー変化を予測することに成功した.

(3) 分子動力学を用いたクライオ電子顕微鏡電子密度へのフレキシブル・フィッティング

クライオ電子顕微鏡の利用は生体超分子の解析に必須なものとなってきており,その構造とダイナミクスを解析するためにますます活用されるようになってきている. Direct electron detector の改良によって,原子解像度の電子密度が得られるようになってきたが, X 線結晶構造解析と比較するとまだ解像度が低い場合が多い.そのため,一つのタンパク質分子について高解像度の立体構造(X 線結晶構造解析によって得られる)を用いて低解像度の電子密度マップ(クライオ電子顕微鏡によって得られる)を満たす原子モデルを予測する手法が必要である. F. Tama らはその予測を行うために基準振動解析や分子動力学シミュレーションを用いたハイブリッド解析法を考案し,様々な系に適用してきた.しかし,従来の手法では, A) 非常に構造変化が大きな場合には答えが収束しない場合があること, B) 電子密度マップと計算で得られるモデ

ルの相関項(cross correlation: CC) の計算に多くの時間がかかることなどの問題が生じていた。そこで、我々は第一の問題を克服するためにアンサンブル計算法を、第二の問題を解決するためにCCの並列化計算法を新たに開発し、分子動力学ソフトウェアGENESISに導入した。アンサンブル計算法では、複数の分子動力学を並行に実施し、電子密度マップを満たす構造をアンサンブルとして予測する。その際、レプリカ交換分子動力学法などの効率の良いアルゴリズムを用いることで、その予測精度を上げることができた。CCの並列化を行うためにツリー法や空間分割法を考案し、分子サイズやモデル(粗視化あるいは全原子)に応じて最適な計算スキームを選べるようにした。これらの開発によってリボゾームなどの巨大分子のフレキシブルフィッティング計算を溶媒を含む全原子分子動力学法を用いて実行することができるようになった。現在、このモデルを用いてタンパク質・DNA複合体などさらに複雑な分子系の解析に取り組んでおり、高解像度のダイナミクス情報を得ることが可能となってきた。

(4) 機械学習を用いた構造・ダイナミクス情報の統合

一分子計測を行うことでタンパク質につけたDye間の距離情報に関する時系列データを観測することができる。これらは正確なダイナミクスの情報を与える一方で、低次元の物理量であり、そこから立体構造を再現することは難しい。その一方で、分子動力学はトラジェクトリの解析から三次元立体構造情報を詳細に得ることが可能である。しかし、計算で用いる分子力場に結果が依存すること、さらに、比較的短いダイナミクス情報しか得られない、など問題点も多い。我々は、一分子計測で得られるダイナミクス情報と分子動力学で得られる構造情報を機械学習で組み合わせることで、正確な(すなわち、実験データを再現する)構造ダイナミクス情報を得ることを試みた。我々は、Hidden Markov Modelを用いて、立体構造情報をHidden State、一分子計測で得られるPhoton CountをObservableとしてこれらをマルコフ状態モデルで記述した。ここで実験データとしてえられるPhoton Countをダイナミクス情報として用いて、マルコフ状態モデルに必要な遷移確率マトリックスを最適化していくことで、正確な構造ダイナミクス情報の記述を目指した。具体的には、NIHのEatonらによって行われたWW-domainの折れ畳みに関する実験データを用いて、分子動力学シミュレーションで得られたマルコフ状態モデルを最適化した。最適化されたマルコフ状態モデルが記述する結果は、Eatonらの実験だけでなく、独立に行われたFershtらのPHI値解析をも再現するものであり、WW-domainのダイナミクスを正確に記述できていると考えられた。また、この手法は一般的なものであり、今後様々なケースに用いることで、実験データとシミュレーションを融合する新しい手法として注目されている。

(5) 構造生物学との連携によるタンパク質動的構造の解析

本新学術領域内で、Sec DF(塚崎智也:奈良先端科学技術大学)、Heme transporter(杉本:理研)、Tom20(神田大輔:九州大学)などにかんする構造生物学と分子動力学を用いた共同研究を実施した。それらの結果はいずれも論文として報告している(あるいは論文投稿準備中)。特に、実験だけでは見えないプロトン化による変化や、構造変化の中間体、結晶におけるコンタクトフリーなスペースの影響などの詳細を分子動力学が解明することに成功した。

5. 主な発表論文等

[雑誌論文](計36件)

1. T Mori, M Kulik, O Miyashita, J Jung, F.Tama, Y.Sugita: Acceleration of cryo-EM Flexible Fitting for Large Biomolecular Systems by Efficient Space Partitioning, Structure 27, 2019, 161~174. 査読あり. 10.1016/j.str.2018.09.004
2. T Nagai, F.Tama, O Miyashita: Cryo-Cooling Effect on DHFR Crystal Studied by

- Replica-Exchange Molecular Dynamics Simulations, *Biophys. J.* 116, 2019, 395~405. 査読あり. 10.1016/j.bpj.2018.11.3139
3. A Srivastava, F Tama, D Kohda, O Miyashita: Computational investigation of the conformational dynamics in Tom20-mitochondrial presequence tethered complexes, *Proteins* 87, 2018, 81~90. 査読あり.10.1002/prot.25625
 4. Y Matsunaga, Y Sugita: Linking time-series of single-molecule experiments with molecular dynamics simulations by machine learning, *eLife* 7, 2018, 32668. 査読あり.10.7554/eLife.32668
 5. A Furukawa, K Yoshikaie, T Mori, H Mori, Y Morimoto, Y Sugano, S Iwaki, T Minamino, Y Sugita, Y Tanaka, T Tsukazaki: Tunnel Formation Inferred from the I-Form Structures of the Proton-Driven Protein Secretion Motor SecDF, *Cell Reports* 19, 2017, 895-901. 査読あり. doi: 10.1016/j.celrep.2017.04.030
 6. O Miyashita, C Kobayashi, T Mori, Y Sugita, F Tama: Flexible fitting to cryo-EM density map using ensemble molecular dynamics Simulations, *J. Comp. Chem.* 38, 2017, 1447-1461. 10.1002/jcc.24785
 7. J Jung, A Naruse, C Kobayashi, Y Sugita: Graphics Processing Unit Acceleration and Parallelization of GENESIS for Large-Scale Molecular Dynamics Simulations, *J. Chem. Theory Comput.* 12, 2016, 4947-4958. 査読あり. 10.1021/acs.jctc.6b00241
 8. Y Matsunaga, Y Komuro, C Kobayashi, J Jung, T Mori, Y Sugita: Dimensionality of Collective Variables for Describing Conformational Changes of a Multi-Domain Protein, *J. Phys. Chem. Lett.* 7, 2016, 1446-1451. 査読あり. 10.1021/acs.jpcllett.6b00317
 9. R Matsuoka, A Shimada, Y Komuro, Y Sugita, D Kohda: Rational design of crystal contact-free space in protein crystals for analyzing spatial distribution of motions within protein molecules, *Protein Science*, 25, 2016, 754-768. 査読あり. 10.1002/pro.2867
 10. J Jung, T Mori, C Kobayashi, Y Matsunaga, T Yoda, M Feig, Y Sugita: GENESIS: A hybrid-parallel and multi-scale molecular dynamics simulator with enhanced sampling algorithms for biomolecular and cellular simulations, *WIREs Comp. Mol. Sci.* 5 2015, 310-323. 査読あり.10.1002/wcms.1220

[学会発表](計48件)

1. Tama Florence, Computational tools to characterize structure and dynamics of biomolecular systems from single molecule experiments, American Chemical Society National Meeting, Boston, USA, 2018.
2. Tama Florence, Hybrid modeling approaches to study structures and dynamics of biological systems, 63rd Annual Meeting of the Biophysical Society, Baltimore, USA, 2018.
3. Yuji Sugita, Mechanisms for Protein-Ligand Binding in Solution and in Cellular Environments, The 18th KIAS conference on Protein Structure and Function, Seoul, Korea, 2018.
4. Yuji Sugita, Conformational changes between E1P and E2P states of SERCA by MD simulations based on string method and free-energy calculations, Symposium on

- “ Membrane Protein Simulations and Free Energy Approaches ” in the 256th ACS National Meeting, Boston, USA, 2018.
5. Yuji Sugita, Development of the flexible-fitting MD simulation method for cryo-EM images of large macromolecular structures and dynamics, Symposium on “ COMP Meets CRYO: New Frontiers in Flexible Fitting, Image Processing & Refinement of Cryo-EM Data ” in in the 256th ACS National Meeting, Boston, USA, 2018
 6. Yuji Sugita, Machine Learning Approach for Protein Dynamics by combining smFRET and MD simulations, Telluride Summer Conference on Free Energy Calculations: Three decades of adventure in chemistry and biophysics, Telluride, USA, 2017.
 7. Tama Florence, Hybrid approaches to characterize structure and dynamics of biomolecular systems from single molecule experiments, Telluride Workshop, Telluride, USA, 2016.
 8. Yuji Sugita, Recent development of replica-exchange molecular dynamics simulation methods, 251st ACS national meeting, San Diego, USA, 2016.
 9. Yuji Sugita, Domain motion analysis of membrane proteins by bioinformatics and molecular dynamics simulations, The international workshop on computational science and engineering, Hong Kong City University, Hong Kong, 2014.
 10. Osamu Miyashita, Atsushi Tokuhisa, Florence Tama, Coarse-Grained Modeling of Structure and Dynamics of Biomacromolecules 3., Telluride Workshop on Coarse-grained modeling of structure and dynamics of biomolecules, Telluride Science Research Center, Telluride, USA, 2014.

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

6 . 研究組織

(1)研究分担者

研究分担者氏名 : TAMA, Florence

ローマ字氏名 : TAMA, Florence

所属研究機関名 : 名古屋大学

部局名 : 理学研究科

職名 : 教授

研究者番号 (8 桁) : 20648191

科研費による研究は,研究者の自覚と責任において実施するものです.そのため,研究の実施や研究成果の公表等については,国の要請等に基づくものではなく,その研究成果に関する見解や責任は,研究者個人に帰属されます.