

2024年9月20日

YYYY/MM/DD

独立行政法人日本学術振興会理事長 殿

To: President, Japan Society for the Promotion of Science

## 研究活動報告書

### Research Report

#### 1. 受入研究者/ Host researcher

受入研究機関・部局・職

Name of Host Institution, Department and Title

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

受入研究者氏名

Host Researcher's Name

杉田有治

#### 2. 外国人招へい研究者/ Fellow

所属研究機関・部局・職

Name of Institution, Department and Title

University of Florida, Department of Chemistry, Assistant Professor

外国人招へい研究者氏名

Fellow's Name

Alberto PEREZ ANTONANZAS

#### 3. 採用期間/ Fellowship Period

2024年

4月

15日

～

2024年

7月

14日

#### 4. 研究課題/ Research Theme

マルチスケールシミュレーションとベイズ推定による蛋白質核酸複合体の構造予測

#### 5. 研究活動報告/ Research Report

##### (1) 研究活動の概要・成果/ Summary of Research Results

##### [概要]

フロリダ大学化学科の Alberto Perez Antonanzas 教授 (以下、Perez 教授) は、JSPS 外国人招聘研究者として、理化学研究所開拓研究本部杉田理論分子科学研究室に約3ヶ月間滞在し、Perez 教授らが米国で開発してきたタンパク質・DNA 複合体構造の予測を行うために分子動力学シミュレーションとベイズ統計を組み合わせた計算手法 MELD 法と、理化学研究所杉田グループを中心に開発している分子動力学ソフトウェア GENESIS を組み合わせた計算手法の開発に取り組んだ。Perez 教授が理研に滞在した3ヶ月という短い期間で、新しい計算手法・ソフトウェアの開発、タンパク質・DNA 複合体の予測などの応用研究の全てを完了することは困難であったが、理研の若手研究員などとも密接な共同研究を行うことで MELD 法についての深く理解することができた。また、マルチスケール分子モデルを用いた MELD 法の高速度化と精密化についての議論を行い、GENESIS に組み込む将来的な計画を具体化することができた。

##### [目的と計画]

タンパク質や核酸、さらにそれらの複合体についての立体構造情報は、X 線結晶構造解析や核磁気共鳴、低温電子顕微鏡などによって実験的に得ることができる。さらに近年、実験的に決定された非常に多くの立体構造情報をデータベース (Protein

Data Bank: PDB) から利用し、機械学習を用いることで理論的に未知のタンパク質の立体構造を予測することが可能になってきた。AlphaFold に代表される機械学習の手法は、従来提案されてきた立体構造予測の手法と比較して正確な立体構造を短時間で予測することができる。予測された立体構造情報を用いて、機能の詳細を理解し、創薬に応用した研究例も存在する。しかし、この手法だけでは実験的に得られた構造情報が不十分な問題、例えば、天然変性状態 (Intrinsically Disordered Protein) の動的な構造情報や、RNA や DNA とタンパク質との複合体の立体構造など、を高い精度で予測することは困難である。このような問題に取り組むためには、実験から得られた構造情報と分子動力学シミュレーションを、ベイズ統計を用いて組み合わせた計算手法が有効であると考えられる。Perez 教授らの研究グループはそのような目的のために、MELD 法を開発し、タンパク質の立体構造予測、タンパク質と基質分子の複合体構造予測、タンパク質と核酸の複合体立体構造予測などを精力的に行ってきた。さらにこの手法を発展させ、従来は困難であって研究課題に挑戦するため、Perez 教授と理化学研究所開拓研究本部杉田理論分子科学研究室および同計算科学研究センター粒子系生物物理研究チームの共同研究を行った。

## [成果]

### 1. 分子動力学ソフトウェア OpenMM に導入された MELD 法の理解

MELD 法の詳細を理解するため、Perez 教授らが分子動力学ソフトウェア OpenMM に導入した MELD 法のプログラムを理研のサーバーで実行し、その動作を検証した。分子動力学ソフトウェア OpenMM は、GPU 計算機を用いて高速な分子動力学を行うことができるとともに、Python と C++ の階層を分離し、ユーザーが Python で新たな機能やモデルを簡単に導入できる利点を有する。この利点を活かして、Perez 教授らは MELD 法を OpenMM に導入して、様々な応用計算を行ってきた。その詳細を理解するため、我々は共同で小タンパク質 (Protein-G) の折れ畳み計算を OpenMM に導入された MELD 法を用いて計算した。ここで、計算時間を短縮するために、水溶媒を露わに扱うことはせずその静電相互作用等を、一般化 Bom モデルを用いて近似した隠溶媒モデルを用いた。これにより図に示すように短い計算時間で Protein-G の折れ畳み過程を計算することができた。MELD 法では様々な拘束条件を用いているが、この計算ではタンパク質の二次構造・Strand Pair・疎水性コンタクトに関する拘束条件を組み合わせている。その詳細を理解することで、GENESIS への MELD 法の導入への知見を得ることができた (図 1)。

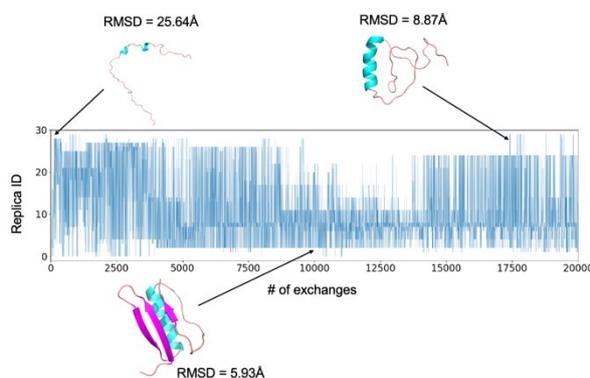


図 1. OpenMM/MELD を用いた Protein-G の折れ畳み計算のトラジェクトリ

### 2. GENESIS に導入された残基レベルの粗視化モデルを用いたタンパク質・核酸複合体の計算

今回の共同研究のターゲットとして、Perez 教授らによる先行研究 (Nucleic Acid Research 2023) で予測したタンパク質・DNA 複合体の構造予測を高速化・精密化することを目指した。図 2 に示すように、陰溶媒を用いた MELD 法では多くの構造の予測に成功しているものの、正確な予測ができなかった例も多い。その理由として、静電相互作用を正確に見積もれていないこと、あるいは、構造探索が十分ではなかったことが原因であると考えられる。後者を改善するために、アミノ酸残基を 1 つの粒子で、塩基を 3 つの粒子で近似する粗視化モデルを用いて構造探索をどの程度高速に実現することができるかを確認した。ここでタンパク質と DNA の間に存在する分子間相互作用は、PwMcos モデル (Tan et al., J. Chem. Theory. Comp. 2018) を仮定したが、相互作用が強すぎるとタンパク質と DNA が結合したまま構造変化ができないため、適度な強さで構造変化ができるような調整を行なった (図 3)。

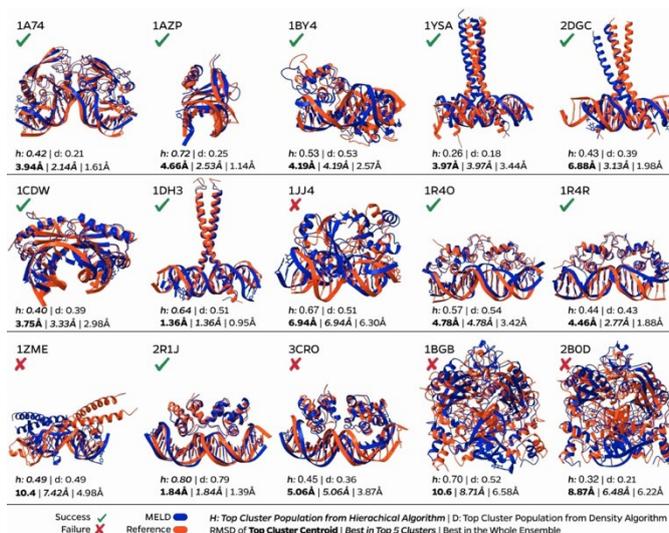


図 2. Perez 教授らによるタンパク質・核酸複合体予測構造

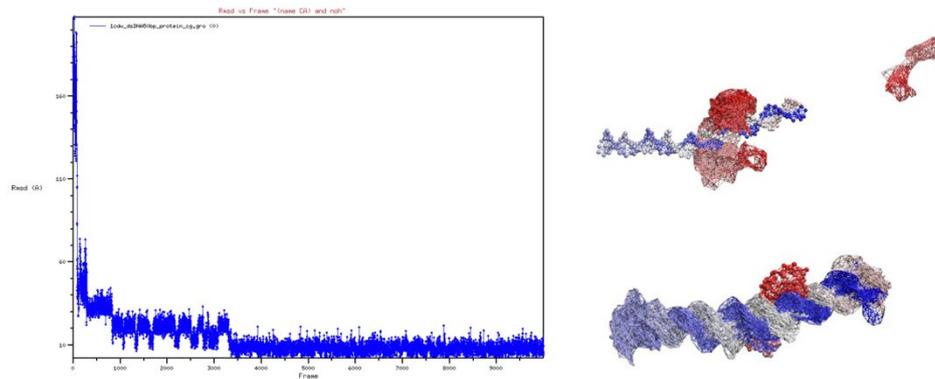


図3. PWMcos のパラメタ調整により、タンパク質 (赤) と核酸 (青) の結合様式の前測を確認した例。このパラメタでは、結晶構造を正確に前測しているが、Unbinding は実現できていない。一方で、Unbinding を強調するようなパラメタでは、タンパク質と核酸の複合体構造が再現できなかった。

### 3. 粗視化モデルを用いた MELD 法の開発

Perez 教授らによる先行研究では、DNA の二重らせん構造の中心とタンパク質の間に拘束力を仮定し、その拘束条件をバランスよく組み合わせることで、DNA 二重らせん構造上でのタンパク質の構造探索を実現している。共同研究を開始した初期には、PWMcos モデルと MELD 法の拘束条件の両者を同時に用いることを計画していたが、結果的にこの2つが同じような効果をするため、PWMcos モデルを利用せずに、MELD 法の拘束条件のみを用いる方法を検討した。MELD 法では様々な距離拘束条件を仮定した分子動力学計算を実施するが、その全ての条件を採用するのではなく、エネルギーの高い拘束条件を無視することで実行的にベイズ統計を満たす確率分布を再現する。そのため、どのような距離条件を採用することが、タンパク質と DNA の間のダイナミクスを維持しつつ最も安定な立体構造を前測することができるかが検討課題である。Perez 教授の滞在期間にはその最適な条件検討までは完了することができなかったため、今後、理研の研究者が継続して取り組むことになった。

### 4. GENESIS への MELD 法の導入による陽溶媒モデルを用いた MELD 法の計算

分子動力学ソフトウェア GENESIS は、並列化による高速化を行うことで溶媒を露わに含む巨大な分子系のシミュレーションを効率的に実行することができる。MELD 法をこのプログラムに導入するためには、多数の距離拘束条件を含む分子動力学計算を行う必要がある。しかし、非常に多くの距離拘束条件を含む分子動力学計算を行う場合には、並列化効率が極端に悪くなる可能性が高い。この問題を回避するための距離拘束条件の並列化アルゴリズムについて議論を行った。

#### (2) 主な研究発表 (雑誌論文、学会、集会、知的財産権等) / Main Research Publications

[雑誌論文]

なし

[学会発表]

- Alberto Perez, “Searching for the BET interactome through AI and Molecular dynamics simulations”, IUPAB2024, Kyoto, 2024/06/24~06/28.
- Alberto Perez, “Synergizing simulations and experiments: faster exploration of biologically relevant states”, IUPAB2024 Hands-on Training Program, CHARMM-GUI/GENESIS MD Tutorial, Kobe, 2024/06/30~07/02.

#### (3) その他/Remarks

今回の共同研究を行うにあたり、受け入れ研究者以外にも以下の理研の研究者が Perez 教授との共同研究に参画した。

- 水谷淳生 (開拓研究本部特別研究員)
- Cheng Tan (計算科学研究センター研究員)
- Jaewoon Jung (計算科学研究センター研究員・開拓研究本部専任技師)

今後は、水谷研究員を中心に Perez 教授との共同研究を継続し、Zoom での議論やフロリダ大学への派遣を行う計画である。