

令和 5 年 6 月 11 日現在

機関番号：82118

研究種目：若手研究

研究期間：2020～2022

課題番号：20K15735

研究課題名（和文）クライオ電子顕微鏡・単粒子解析への量子コンピュータ技術の応用

研究課題名（英文）Application of Quantum Computer Technology to Cryo-EM Single Particle Analysis

研究代表者

守屋 俊夫（Moriya, Toshio）

大学共同利用機関法人高エネルギー加速器研究機構・物質構造科学研究所・特任准教授

研究者番号：20565014

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：当初は巡回セールスマン問題として捉えた3次元角度推定への量子アニーリングの応用研究に取り組んだ。シミュレーションデータセットでは良好な結果が得られたが、実験データセットを用いた場合、高い再現性と精度の3次元角度推定が得ることができなかつたため断念した。代替案としてタンパク質ダイナミクス解析への量子アニーリングの応用を検討した。基質無し条件下で取得したV型ATPase全体構造のデータセットの試験的なダイナミクス解析を行い、Vo部分のCリング構造と回転軸DFd複合体が連動して連続的な回転運動をする一部の様子を動画で可視化することに成功した。

研究成果の学術的意義や社会的意義

膜タンパク質や超分子複合体等の立体構造を原子分解能で可視化する技術として、クライオ電子顕微鏡を道いた単粒子解析が急激に成長している。それに伴い本手法に要求されるデータ量と計算量も急増し続けている。本研究の成果として得られた、莫大な計算量を必要とするクライオ電子顕微鏡・単粒子解析、特にダイナミクス解析のために構築した生体高分子構造解析向けのクラウド計算環境であるGoToCloudは、本課題の当初の目的であった本手法への量子コンピュータ技術の応用のための技術基盤となる。

研究成果の概要（英文）：Initially, we tried to apply quantum annealing to 3D angle estimation of each particle image by viewing the estimation as a traveling salesman problem. Although reasonable results were obtained with a simulation dataset, high reproducibility and accuracy of 3D angle estimation could not be obtained with an experimental dataset. As an alternative research theme, the application of quantum annealing to protein dynamics analysis was investigated. We performed the dynamics analysis of the whole structure of V-ATPase obtained under substrate-free conditions, and succeeded in visualizing a part of the continuous rotational motion of the C-ring structure of the Vo domain in conjunction with the rotational axis DFd complex domain as a movie.

研究分野：構造生物学

キーワード：クライオ電子顕微鏡 単粒子解析 量子コンピュータ技術 画像処理 機械学習 構造生物学 タンパク質ダイナミクス解析 クラウドコンピューティング

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、研究者個人に帰属します。

1. 研究開始当初の背景

分子レベルの生命活動において主要な役割を果たすタンパク質およびその複合体の立体構造決定手法は、生命科学の発展や創薬プロセスにおいて欠かせない技術である。創薬などへの産業応用のためには、2.0-3.5 Å の近原子分解能立体構造をより多くのより多様なタンパク質で安定して決定できることが要求される。近年の技術革新により、クライオ電顕・単粒子解析で生体高分子の近原子分解能構造の決定が可能になった。そのため、他の手法では構造決定が難しいタンパク質複合体等を中心にして本手法が広く用いられるようになっていた。

クライオ電顕・単粒子解析で問題になるのが、多くの計算処理ステップで複数の未知パラメータの最適化が必要なことである。つまり、数多くの未知パラメータが存在するため、多くの最適化問題を解決しなければならない。同時に単粒子解析ではより大量な画像データをより高速に処理するニーズが常にある。つまり、様々なアルゴリズムを駆使して分解能の向上を行いつつも、常に計算時間を実用的なレベルまで短縮しなければならない。そのため、これまでも様々な高速化が行われてきた。主な高速化は並列化実装によって実現されてきた。しかし、日常的に原子座標モデルの構築が可能なレベルの分解能の立体構造をより多くのより多様な生体高分子で安定して決定するためには、より小さいピクセルサイズで、より多くの、より大きいサイズの粒子画像が必要になる。さらに問題となるのが複数構造状態の解析である。クライオ電顕の急激な分解能向上と普及により、より多くの構造状態を分類できるようになる期待は高まっていた。しかし、必要な粒子数は、データセットに含まれる構造状態の数に比例して増加し、必要なメモリ量と計算量も急激に増える。そのため、現状普及している実装では、近い将来に処理時間を実用的な範囲に維持できなくなるのは明白であった。

2. 研究の目的

このような背景を踏まえ、近年になり急速に実用化が進められている量子コンピュータ技術をクライオ電顕・単粒子解析に導入する有効性と必要性は明らかであった。量子コンピュータはスーパーコンピュータの上位互換ではなく、得意な特定の問題を解くアクセラレーターとして位置付けられているが、その実用範囲は広い。特に「組合せ最適化問題」は従来型コンピュータにとって手に負えない難問の代表例である。これは無数にある組合せのうち条件を満たすものを探しあてる問題で、サイズが大きい場合には答えの候補の全件チェックに天文学的な時間が必要となる。量子コンピュータは量子重ね合わせによる並列計算で総当たりすることができるため、大局解かそれに近い局所解を高速に得ることが可能である。このことから、複数未知パラメータの最適化が複数のステップで必要なクライオ電顕・単粒子解析は、量子コンピュータの優位性を検討するには非常に適した分野であることがわかる。それにも関わらず、未だ量子コンピュータ導入の動きは本分野では見られなかった。そこで本研究では単粒子解析に量子コンピュータ技術を応用することで、その最終出力である生体高分子立体構造の分解能を向上しつつも実用的計算時間に収められるソフトウェアの開発を将来に渡っても停滞させないための技術基盤を確立することを目指した。

3. 研究の方法

まず、ボトルネックとなっている最適化問題を抽出、その従来解法を変更し量子コンピュータの既存技術をそのままアクセラレーターとして単粒子解析に適用することを試みた。本研究の初めの取り組みとして、「初期 3 次元構造の決定」ステップ中の「3 次元角度推定」プロセスに、量子アニーリングという巡回セールスマン問題(TSP)を解ける既存量子アルゴリズムの適用を行った。3 次元角度推定の最適化問題での粒子像間 3 次元一貫性を TSP の都市間距離に置き換えて問題が同一になるように工夫した。シュミレーションデータを用いたテストを 2000 量子ビット搭載の「D-Wave 2000Q」で実行した。n 都市 TSP 問題では n の二乗個の全結合量子ビットを必要だが、現行 D-Wave は疎結合ビットであるため最大 64 個の擬似的な全結合ビットしか使用できない。そこで、多い画像枚数の場合での提案アルゴリズムの正当性は従来コンピュータ用の TSP 解法である LKH3 アルゴリズムで検証し、実用性は D-Wave で検討した。その結果、高い再現性と精度の 3 次元角度推定を行うことができた。一方、実際の実験データセットを用いたテストでは高い再現性と精度の 3 次元角度推定が得ることができなかった。これはクライオ電顕像のノイズレベルが高いことに起因する計算誤差が原因だと考えて、計算誤差を小さくする工夫を行った。しかし、電顕画像ノイズレベルが高すぎる、電子顕微鏡特有のコントラスト伝達関数(CTF)による情報の変調を完全に補正できない等の問題があり、3 次元角度推定への量子アニーリングの応用を行うことは難しいと判断した。

そこで、代替案としてタンパク質ダイナミクス解析(一つのデータセットに含まれる複数の構造状態を決定し、タンパク質の動きを解明すること)への量子アニーリングを応用する検討を行った。ダイナミクス解析においては、深層学習をベースにした全く新しいアプローチのアルゴリズムの開発が活発化してきている。その一つである CryoDRGN (Zhong et al. Nat Methods. 2021, 18, 176) はデータセットに含まれる粒子画像と同じ莫大な数の状態の構造をクライオ

電顕マップとして出力する(図1)。しかし、各構造状態同士の物理化学的な関連が不明なため、これだけではタンパク質がどのように動くか(生物学的に意味のある構造変化)までは解明することができない。そこで、複数のCryoDRGN 構造同士の自由エネルギー的な近さを計算し、物理化学的に尤もらしい構造変化の最適なパスを見つけることに量子アニーリングを応用する研究に着手した。

4. 研究成果

ダイナミクス解析のケーススタディとして、イオン輸送を行う回転分子モーターであるV型ATPase(図2)のエネルギー変換機構を原子のレベルで解明を目指して、その全体構造の解析を行った。まず基質無し、ATP入り、ADP入り、二種類の阻害剤入りのV型ATPase全体構造の5つのデータセットを300kVクライオ電顕で取得した。これまでに、基質無し条件下で取得したデータセットの試験的な解析を行い、従来法である3次元クラス分けで7構造状態まで分離した。さらに、CryoDRGN を用いてダイナミクス解析も行い、Vo部分のCリング構造と回転軸DFd複合体が連動して連続的な回転運動をする様子の一部を動画で可視化することに成功した(図3)。

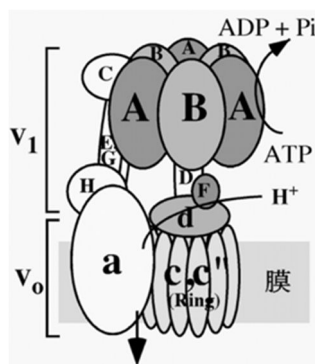


図2 V型ATPaseの構造

この解析結果をもとに、マルチスケール分割統治型MD法を用いて、CryoDRGNで得られた20状態のV型ATPase全体構造のクライオ電顕マップそれぞれの自由エネルギーを計算するアルゴリズムの研究開発を進めた。しかし、CryoDRGNで得られる構造状態の多くが、全体もしくは局所的に中低分解能(4-8Å程度)のマップとなることが分かった。これは、メモリ容量等の計算機ハードウェアの制限に応じて、計算可能な解像度も制限されてしまうことが主な原因である。この分解能レンジでの原子モデリングはハードルが高いため、特殊なアルゴリズムを必要とする。そこで、深層学習ベースの原子モデリングを行うDeepMAINMASTを利用して、中低分解能の分解能のクライオ電顕マップに含まれる二次構造(ヘリックス、シート、コイル/ターン)を正確に同定し、最小全域木と呼ばれるグラフ構造を用いて主鎖モデルのトレースを行い、ツリーグラフ構造を用いて主鎖の接続とCの位置を直接トレースすることで、原子モデルを決定する試みを開始した。また、同時に本ダイナミクス解析には莫大な計算量を必要であることが想定されていたため、この解析計算のために生体高分子構造解析向けのクラウド計算環境であるGoToCloudを構築した。現在、GoToCloudにDeepMAINMASTを導入する準備も行っている。

今後、上記の成果を全ての研究者が簡単に利用できるパブリックなクラウド計算環境上で統合し、従来法より遥かに適用範囲の広い実用的なダイナミクス解析の研究基盤を世界に先駆けて確立することを目指す。そして、その構造揺らぎと構造変化の複雑さのために未だ謎であるV型ATPaseの回転分子モーターのエネルギー変換機構を原子レベルで解明する。さらに、この結果をベースとして複数の構造状態同士の自由エネルギー的な近さを計算し、物理化学的に尤もらしい構造変化の最適なパスを見つけることに量子アニーリングを応用する予定である。

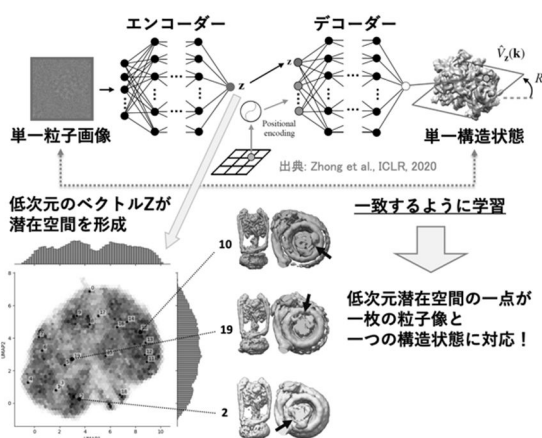


図1 CryoDRGNのダイナミクス解析

図1の下部には、低次元潜在空間の一点が一枚の粒子像と一つの構造状態に対応していることが示されています。また、異なる構造状態間の遷移パスも示されています。

この解析結果をもとに、マルチスケール分割統治型MD法を用いて、CryoDRGNで得られた20状態のV型ATPase全体構造のクライオ電顕マップそれぞれの自由エネルギーを計算するアルゴリズムの研究開発を進めた。しかし、CryoDRGNで得られる構造状態の多くが、全体もしくは局所的に中低分解能(4-8Å程度)のマップとなることが分かった。これは、メモリ容量等の計算機ハードウェアの制限に応じて、計算可能な解像度も制限されてしまうことが主な原因である。この分解能レンジでの原子モデリングはハードルが高いため、特殊なアルゴリズムを必要とする。そこで、深層学習ベースの原子モデリングを行うDeepMAINMASTを利用して、中低分解能の分解能のクライオ電顕マップに含まれる二次構造(ヘリックス、シート、コイル/ターン)を正確に同定し、最小全域木と呼ばれるグラフ構造を用いて主鎖モデルのトレースを行い、ツリーグラフ構造を用いて主鎖の接続とCの位置を直接トレースすることで、原子モデルを決定する試みを開始した。また、同時に

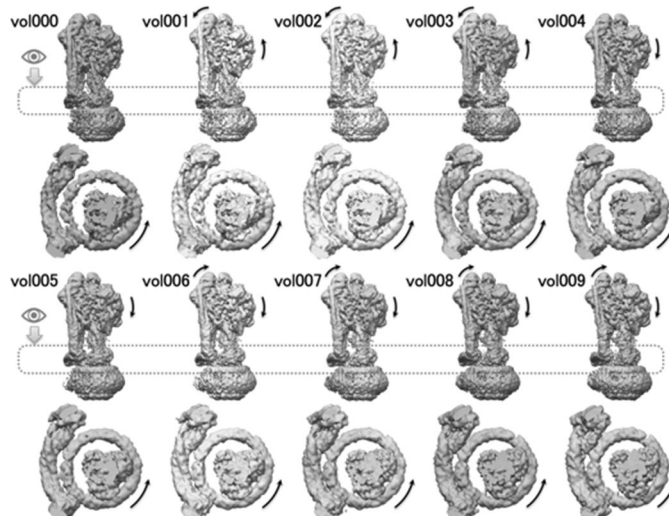


図3 V型ATPase全体構造のCryoDRGNマップ動画のコマ

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計7件（うち招待講演 5件 / うち国際学会 2件）

1. 発表者名 守屋俊夫
2. 発表標題 AWS ParallelClusterをハブとしたクライオ電子顕微鏡単粒子解析現場のIoT化
3. 学会等名 第22回日本蛋白質科学会年会
4. 発表年 2022年

1. 発表者名 Toshio Moriya
2. 発表標題 Application of Public Cloud for Cryo-EM and Improvement of User Support for X-ray Crystallography and MicroED at KEK
3. 学会等名 The 2023 Structural Biology Summit (招待講演) (国際学会)
4. 発表年 2023年

1. 発表者名 守屋俊夫
2. 発表標題 クラウド計算環境をハブとしたクライオ電顕ネットワークのIoT化で進化する単粒子解析
3. 学会等名 日本顕微鏡学会 第47回関東支部講演会 (招待講演)
4. 発表年 2023年

1. 発表者名 守屋俊夫
2. 発表標題 クライオ電子顕微鏡単粒子解析における分子ダイナミクス解析アルゴリズム研究の最前線
3. 学会等名 日本顕微鏡学会第77回学術講演会 (招待講演)
4. 発表年 2021年

1. 発表者名 守屋俊夫
2. 発表標題 クライオ電子顕微鏡単粒子解析計算の高度化の取組み
3. 学会等名 計算を中心としたバイオ分野の筑波大-KEK連携セミナー
4. 発表年 2021年

1. 発表者名 守屋俊夫
2. 発表標題 クライオ電子顕微鏡法を用いたプロテインダイナミクス研究の最前線
3. 学会等名 iBIX-JAXA-KEK物構研-QST合同タンパク質研究会（招待講演）
4. 発表年 2022年

1. 発表者名 Toshio Moriya
2. 発表標題 Recent advancements of protein dynamic analysis in cryogenic electron microscopy
3. 学会等名 The 6th International Symposium Of Quantum Beam Science（招待講演）（国際学会）
4. 発表年 2022年

〔図書〕 計0件

〔出願〕 計2件

産業財産権の名称 電子顕微鏡データ解析システム、及び電子顕微鏡データ解析方法	発明者 守屋俊夫、山田悠介	権利者 高エネルギー加速器研究機構
産業財産権の種類、番号 特許、特願2022-114077	出願年 2022年	国内・外国の別 国内

産業財産権の名称 解析システム、及び解析方法	発明者 山田悠介、守屋俊夫	権利者 高エネルギー加速器研究機構
産業財産権の種類、番号 特許、特願2022-211645	出願年 2022年	国内・外国の別 国内

〔取得〕 計0件

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------