

HPCと実験データの融合による 動的構造生物学の革新

令和7年度 研究課題報告会

研究代表者：Florence Tama / 杉田 有治

実施期間：令和7年度 ～ 令和11年度

生体高分子の構造から動的構造へ

研究目標

HPC と実験データの融合により、タンパク質の動的構造を解明する。

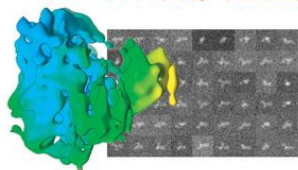
- 近年、実験手法の進歩や構造予測AIの開発により、タンパク質の静的構造を知ることとは比較的容易になった。
- しかし、タンパク質の機能発現の本質は、構造の「動き（ダイナミクス）」
- 一方で、タンパク質を「静止」させて観測する構造生物学的実験手法や、AI による構造予測では、このような動的構造を直接解析することは依然として困難である。
- 分子シミュレーションはタンパク質の動きに関する詳細な情報を与えることが可能であるが、計算時間の制約やモデルの精度といった限界も存在する。
- 本研究では、大規模分子シミュレーションと実験データを統合することで、タンパク質の動的構造を解析し、その生物学的機能との関係を明らかにする。

HPC × 大型実験施設の融合

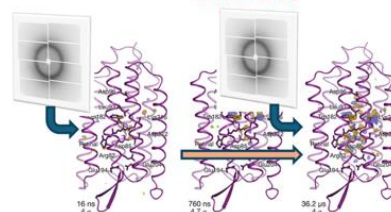
「富岳」の計算能力と、SPring-8/SACLAの最先端計測を融合活用

静的・断片的な実験データ
低解像度、ノイズ、
平均化された情報

クライオ電子顕微鏡
単粒子観測
クライオ環境, 中解像度



XFEL 時分割結晶構造解析
時分割, 常温
結晶環境

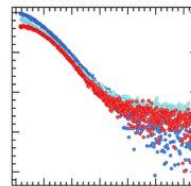


メリット
デメリット

実験, AI予測による
静的構造

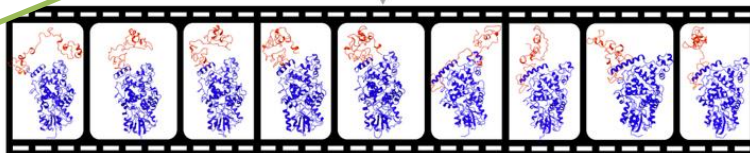


X線小角散乱
溶液環境
低情報量



HPC 富岳

大規模シミュレーション
で実験データの隙間を
物理学的に補完



データと大規模計算の
融合により高精度な
分子動画を生成

期待される成果と地域貢献

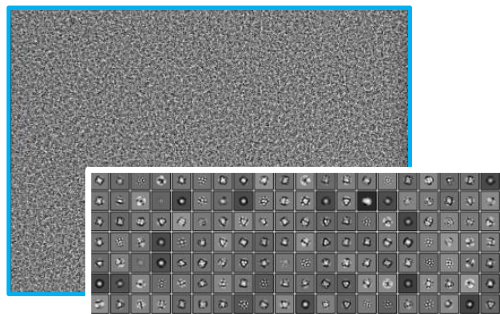
- 計算技術と多様な実験データを融合するアルゴリズム・ソフトウェアの開発
- MD シミュレーションを用いたハイブリッドモデリングの高度化
 - Cryo-EM 解析ソフトの強化による構造多様性研究の推進
 - XFEL TR-SFX の自動データ処理・シミュレーションによる解析効率化
 - SAXS データを用いたタンパク質ダイナミクス研究への応用
 - メソスケールシミュレーションの手法の高度化
- 開発する手法を GENESIS の機能とて公開し、地元大学・企業での利用を促進
- SPring-8 / SACLA と富岳を活用した生体分子ダイナミクス研究の新しい展開
- 教育活動、研究シンポジウム開催による人材育成と研究連携の促進

進捗状況：Cryo-EMデータからのタンパク質構造多様性の抽出

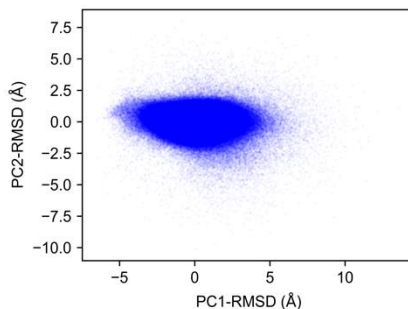
- 理研SPring-8で取得されたタンパク質GDHのcryo-EM観測データを解析。
- 富岳上でGENESIS/MDSPACEを実行し、120万画像を対象としたフィッティング計算を行うことで、サンプル内のGDHの構造分布の再構築を行なった。
- 得られた構造分布を実験による既知の構造と比較したところ、知られている離散的な構造だけでなく、より連続的かつ広範な構造分布が存在することが示された。
- 構造分布をもとにGDHの機能に関連する構造多様性について定量的評価を行なった。



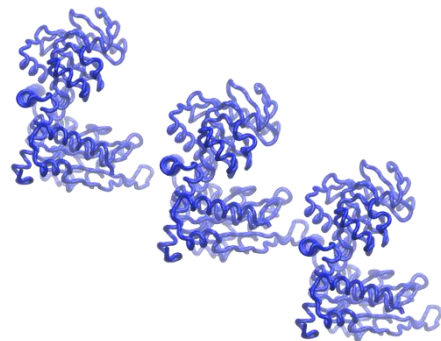
クライオ電顕



120万画像へのフィッティング



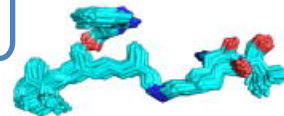
データが示す構造分布



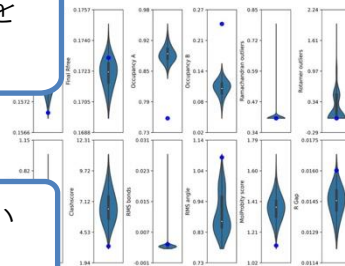
進捗状況：SACLA XFELによる時分割構造解析

- 分子動力学シミュレーションを用いた時分割シリアルフェムト秒結晶構造解析データからの構造モデリング自動化への取り組み。
 - 分子動力学シミュレーションを行い、観測対象が取りうる構造のサンプルを大量に生成。
 - それぞれのサンプルを初期構造として観測データをターゲットとした構造精密化を並列実行。
 - 作られたモデルの構造の立体化学的妥当性とデータ整合性を評価し、最適なモデルを選択。
- SACLAで取得された実験データにおいて実証。

MDシミュレーションにより
タンパク質の
取りうる構造を多量に生成



X線実験データをターゲットと
した構造精密化モデリングを
並列実行



得られたモデルの評価を行い
最も適切なモデルを抽出

構築モデルの一例

