

細胞内環境でのタンパク質や核酸の 分子ダイナミクスの解析

2015年9月30日

HPCI戦略プログラム 分野1 課題1

理化学研究所 杉田理論分子科学研究室

杉田 有治



自己紹介

■ 研究チーム紹介

■ グループリーダー: 杉田有治

- 理化学研究所杉田理論分子科学研究室主任研究員
- 理研計算科学研究機構TLと理研生命システム研究センターTLを兼務

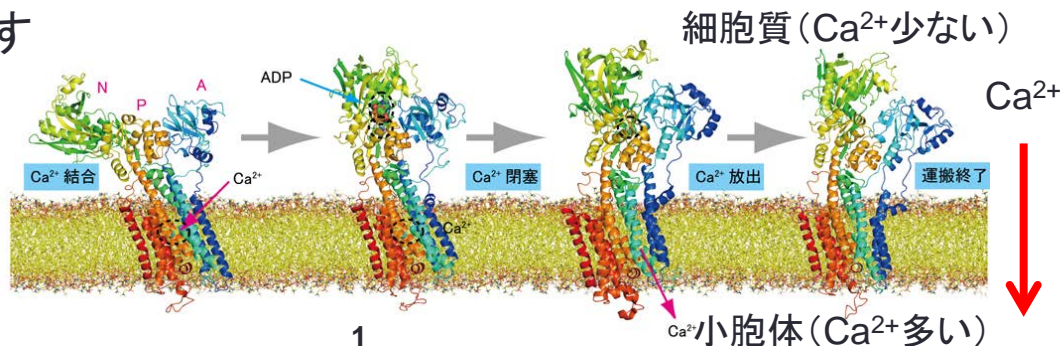
■ 課題参加者(チームリーダー)

- 河野秀俊(原子力機構)、高橋恒一(理研)、石谷隆一郎(東大)、太田元規(名大)、高田彰二(京大)、林重彦(京大)、池口満徳(横浜市大)

■ グループリーダーの自己紹介

- 中学生になったときに両親に初めてパソコンを購入してもらって以来、いろんなコンピュータを使ってきました
- 理学部化学科を卒業していることもあり、原子や分子のレベルで生命現象を理解したいとずっと思ってきました
- タンパク質の機能の複雑さや精密さに魅せられて、理論や計算でもっと深く理解したいと考えています

東大豊島近教授がX線結晶構造解析で決定したカルシウムイオンポンプの複数の構造。この構造変化は1万倍のイオン濃度差に逆らってCa²⁺を輸送するために必須



分子動力学って何？

■ 分子動力学はニュートンの運動方程式を解くことで、分子運動を予測する計算手法

- 利点: 実験では観れない原子レベルの詳細な分子運動や構造が調べられる
- 弱点: 現在の計算機では(原子モデルを利用した場合)百万分の1秒(マイクロ秒)程度の運動しか計算できない
 - それでも、約十億回の繰り返し計算を実行している
- 応用: 薬剤の結合状態の予測など

左の図は教科書に描かれている有名な絵だが、必ずしも正しくない(課題2で説明)



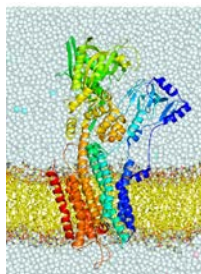
■ 計算機の進歩に伴って、急速に発展してきた。

- その基盤を構築した研究者(Karplus, Levitt, Warshel)に2013年度のノーベル化学賞が与えられた

最初のMD計算(1977)



最近のMD計算 (2005)



イオンポンプ(994アミノ酸)
100 ns (10^{-7} 秒)
生体膜中の計算

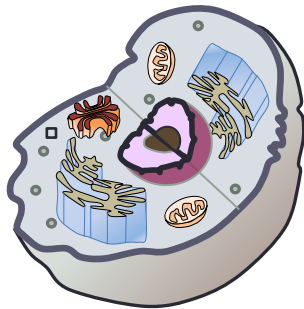
「京」を用いたMD計算

?

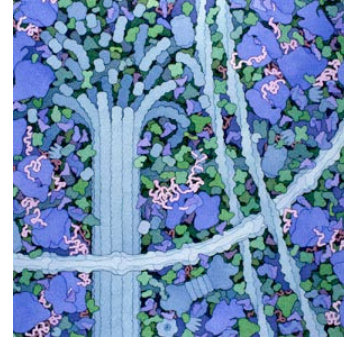
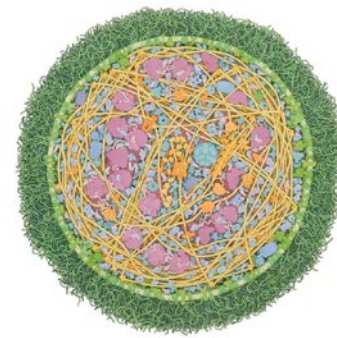
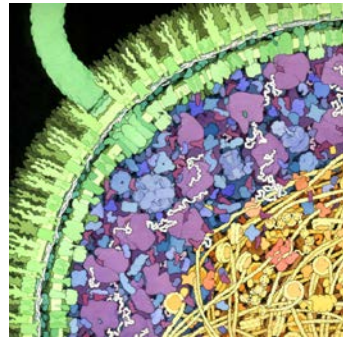
1つのタンパク質のリアルな短時間計算

細胞内環境(分子混雑)

- 試験管内(*in vitro*)のタンパク質反応と細胞内(*in vivo*)の反応では異なることが多い。なぜだろうか？
 - 一つの原因は、細胞質内は非常に多くの分子(タンパク質、RNA、低分子化合物)が含まれた濃厚溶液であり、この分子混雑環境が試験管内の希薄溶液とは大きく異なるからである



教科書に書かれた細胞の絵。核、ミトコンドリア、オルガネラを含む。細胞質は灰色で表示



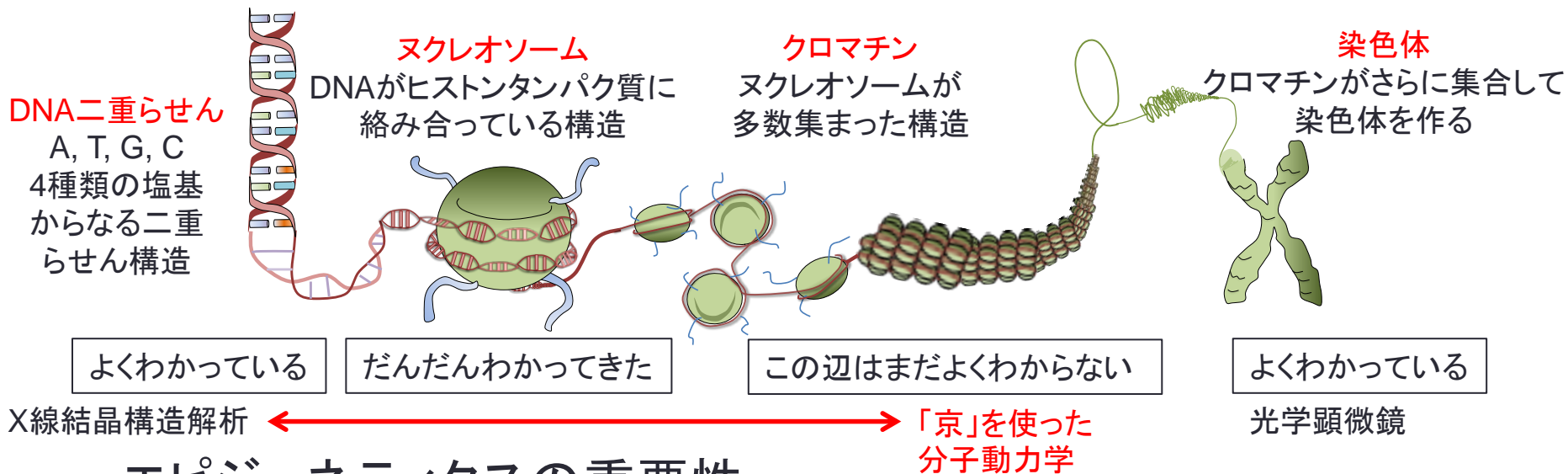
最新の生物学的知見を考慮した細胞内分子のイラストレーション。
Illustration by David S. Goodsell, the Scripps Research Institute

- このような複雑な環境の中で、複数のタンパク質が反応しながら、細胞膜から核内へと信号を伝えていく
 - このような濃厚溶液の中でどのような制御が行われているのか理解したい
 - そのために、細胞内環境を分子レベルで表現したシミュレーションを行い、そこでのタンパク質構造や分子運動を理解する

細胞内環境(核内DNA)

■ DNAは核の中で遺伝情報を保持している

- 近年、DNAが核の中で収納されている仕組みがわかってきた(真核細胞)



■ エピジェネティクスの重要性

- 同じDNAから、皮膚の細胞や神経細胞など様々な細胞ができるのはなぜか？
- 一卵性双生児でも生育環境が異なると違いがでるのはなぜか？
- 加齢や生活習慣(飲酒、喫煙)によって癌化の可能性が高まるのはなぜか？
- 近年の実験によって、DNAメチル化やヒストン化学修飾とそれに伴うヌクレオソームやクロマチンの変化が明らかになってきた
- それらの構造や分子運動の変化は実験的にはよく理解されていない

「京」を最大限利用するために

■「京」の利点と弱点

- 利点: 非常に多くのCPUが一度に利用できる。さらに、それらは高速なネットワークで接続されている
- 弱点: 一つ一つのCPUの速度はそれほど速くない。分子動力学専用計算機(MDGRAPE(RIKEN)やANTON(D.E.Shaw Research)の方が分子動力学を行うためにはずっと早い

■「京」を最大限に利用する戦略

- 数千から数万CPUを用いた超並列計算を実行
 - 超並列分子動力学ソフトウェアGENESISの開発
 - レプリカ(ターゲット分子のコピー)を用いた複数の分子動力学計算の連成(SCUBA)
- 数千の異なる条件の分子動力学を一斉に実行し、実験データを満たす構造モデルを予測(MD-SAXS)
 - SAXS(X線溶液散乱)実験プロファイルを満たす条件の検討

数千から数万CPUを用いた超並列計算

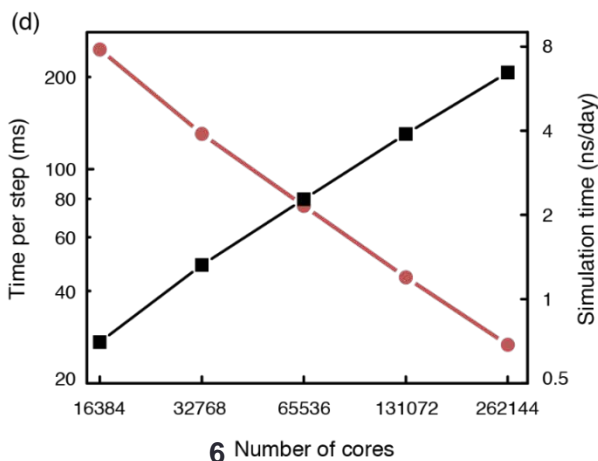
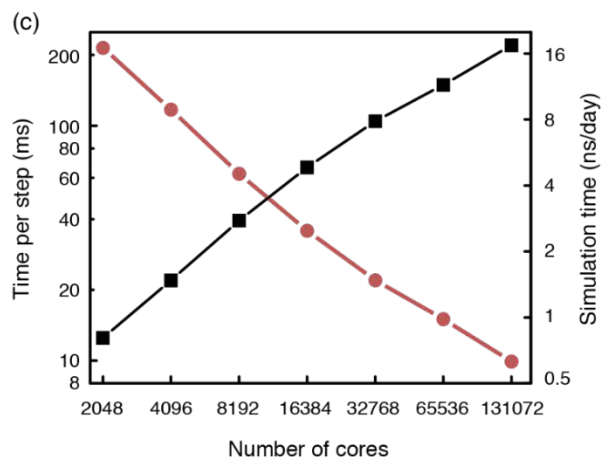
■ 超並列分子動力学ソフトウェアGENESISの開発

■ 「京」の性能を十分に発揮出来る分子動力学ソフトウェアをゼロから理研計算科学研究機構を中心に開発

■ 並列化とは？

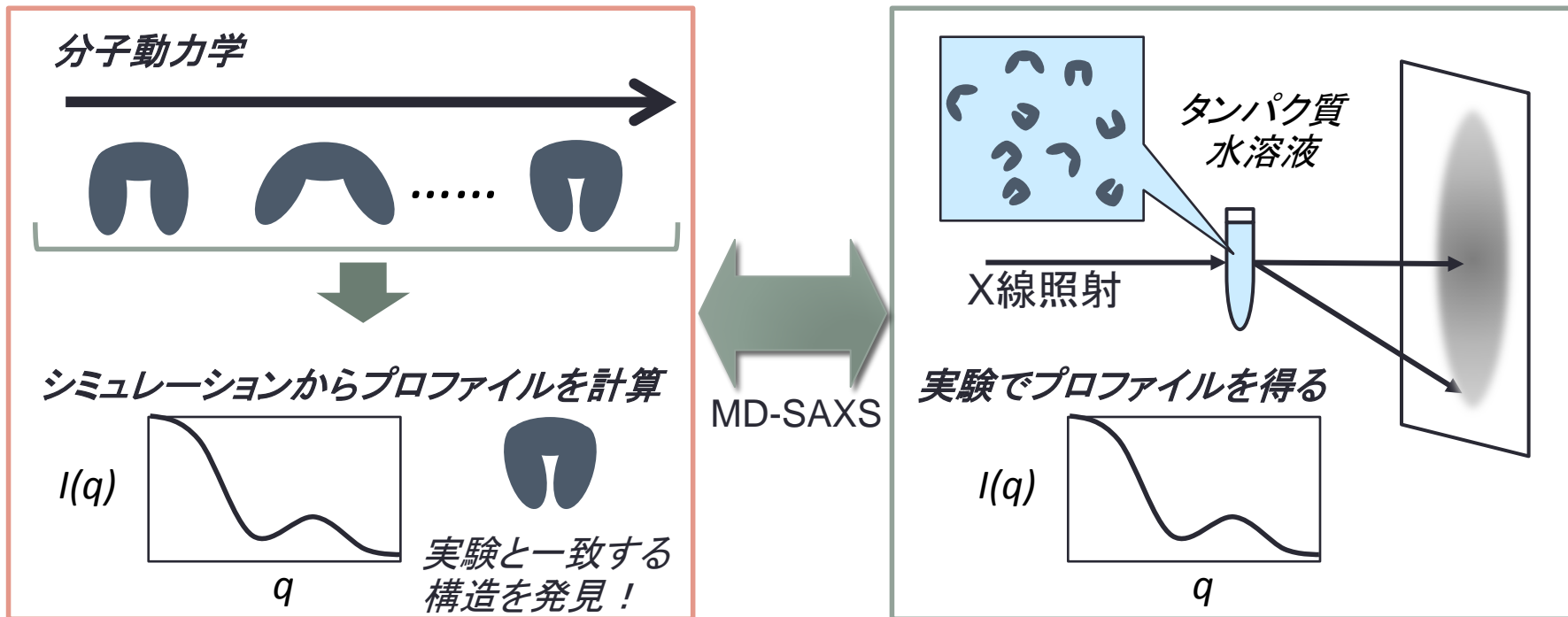
- 利点: 多数のCPU(N個と仮定)にデータと演算を分割することで、理想的には $1/N$ の計算時間に短縮できる可能性がある
- 弱点: CPU間の通信が必要となり、 $1/N$ には短縮できない。特に多くのCPUを使った場合にCPU間通信がボトルネックとなり、むしろ遅くなる場合もある

■ GENESISでは巨大分子系(数千万から1億原子を含む)の分子動力学を「京」を用いて実現した



数千の異なる条件の分子動力学を一斉に実行

- 分子動力学 (MD) と X線溶液散乱 (SAXS) のタイトな連携: MD-SAXS
 - 分子動力学で得られた各々の構造からプロフィールを計算し、**実験と比較して一致した構造**を正解と考える
 - 「京」の多くのCPUを用いて、イオン濃度や相互作用エネルギーを変えた**数千もの異なる条件の分子動力学を一斉に実行**することで高速な探索を実現



今後の応用や展開

- 分子動力学計算の過去、現在、未来
 - 過去: 水溶液中や脂質二重膜中のタンパク質の短い分子動力学(100ナノ秒: 10^{-7} 秒)
 - 現在(京): 細胞内環境を考慮した巨大生体分子システムの長時間分子動力学(数マイクロ秒: 10^{-6} 秒)
 - 未来: より長時間(数十マイクロ秒からミリ秒)の分子ダイナミクスを実現して、タンパク質機能の制御を実現。新しい創薬ターゲットの開拓など
- 長時間の分子動力学シミュレーションの結果を、最先端の実験計測で検証する必要がある。一分子計測実験、溶液X散乱実験、In-cell NMRなどと組み合わせる
- 開発したソフトウェアや計算手法をアカデミアのみならず製薬業界などでも広く利用できるように公開し、利用の普及をはかる
 - GENESISをはじめ課題1で開発したソフトウェアの多くはフリーソフトとして公開中

スケジュール

- 今後の予定(短期・長期的計画)
 - 「京」やポスト「京」を用いた大規模計算を継続するだけでなく、理研 QBiCをはじめとした実験グループと密接な連携を行う。
 - 本研究で得られた知見を生かしてタンパク質や核酸の機能制御を目指した研究をポスト「京」を用いて行う(重点課題1)。
- 研究の進捗、分析時期、サイエンス成果(論文等)の出る時期
 - マイコプラズマ細胞に含まれるほとんどのタンパク質やRNA、代謝物を含む超巨大システムの全原子分子動力学計算(世界最大規模の計算)
 - 現状: 計算はほとんど終了。もうすぐ論文投稿予定。
 - 今後: **細胞環境を考慮した創薬応用MD計算**をポスト京等で実施予定。
 - ゲノムDNAの収納状況から活性状態への移行時のエネルギー変化の計算
 - 現状: 計算はほとんど終了。年内に論文投稿予定。
 - 遺伝子発現、細胞分化、**細胞リプログラミング(ES細胞、iPS細胞)などの仕組みの解明**へ、さらには**エピゲノム創薬**へ。
 - ゲノムDNA(ヌクレオソーム多量体)の溶液構造に関するシミュレーションと放射光実験との連携
 - 現状: 計算はほとんど終了。もうすぐ論文投稿予定。
 - 「京」と、「SACLA」や「SPRING8」などを連携することで、**次世代の構造生物学(結晶構造に依らない)**を構築する。

まとめ

- 「京」を用いた大規模な分子動力学シミュレーションを行うことで、細胞内環境を考慮した分子ダイナミクス解析を実現しました。このような計算は従来の計算機・計算手法では不可能でした。
- 「京」の演算性能を十分に発揮するため、新しいソフトウェアや計算手法の開発を行いました。
- 数千から数万CPUを並列に利用することで、最大で1億原子を含む細胞内分子混雑システムの全原子分子動力学計算を実現しました。これは、マイコプラズマ細胞というバクテリアの細胞質の計算に対応します。
- ヌクレオソームやクロマチンなどの構造を理解・予測するために様々なモデル（全原子、粗視化モデルなど）や計算手法を組み合わせたマルチスケール計算を行いました。
- 実験研究とタイトな連携を行うとともに、今後は、ポスト「京」重点課題の中で分子動力学シミュレーションを創薬応用に役立てる研究開発も実施します。