

## 2.2.1 当該年度(平成26年度)における研究成果

### ○ 研究開発課題

#### I 戦略課題 1：細胞内分子ダイナミクスのシミュレーション[統括：杉田 有治（理化学研究所）]

細胞質中の分子混雑、生体膜環境、膜を介した物質及び信号伝達など細胞環境を強く意識した分子および細胞スケールシミュレーションの実現を目指し、細胞内信号伝達経路の1分子粒度計算、膜タンパク質による物質輸送の解明、核内 DNA タンパク質複合体の構造予測と機能解明を行う。

##### I-1 杉田 有治（理化学研究所）

信号伝達経路上のリン酸化酵素への細胞質分子混雑の影響

##### I-1-1 実施計画

本研究では、「細胞内分子ダイナミクスのシミュレーション」研究の一環として、「京」を用いた大規模な分子動力学計算を実行する。細胞環境（分子混雑など）を露に考慮した環境での生体分子の大規模構造変化を計算機中に実現する。これにより、分子混雑環境の物理化学的パラメータを抽出して、粗視化モデルを構築する。

平成 26 年度は、(1) 平成 25 年度に行った細胞質環境のシミュレーションに関する詳細な解析を行い、タンパク質の拡散、安定性、基質認識などへの影響を検討する、(2) MEK-ERK 複合体のモデリングを行うとともに、分子混雑が酵素反応に及ぼす影響を検討する。また、「細胞内分子ダイナミクス」研究チーム間の連携をはかるための班会議や、実験研究者などと連携するために必要な研究セミナーを随時開催する。

##### I-1-2 実施内容（成果）

###### (1) バクテリア細胞の細胞質分子混み合い環境に関する大規模分子動力学計算

1) 細胞内環境は 70%が水で占められており、残りの 30%をタンパク質や RNA などの生体高分子や各種の代謝物などの低分子化合物が占めている。この分子混み合い環境におけるタンパク質などの動態はこれまで十分に理解されていなかったが、近年、In-cell NMR などの実験によって新しい知見が得られつつある。我々は、1170 万原子および 1 億原子を含む巨大な分子混雑系を構築し、「京」を用いた大規模分子動力学計算を行うことによりこの問題に取り組んだ。この分子混雑系は、プロテオーム解析等でタンパク質の濃度等が明らかになっているバクテリア細胞 (*Mycoplasma Genitalium*) の細胞質環境をモデル化したものであり、個々のタンパク質等は共同研究者の Michael Feig (Michigan State University) が Homology Modeling で構築した。水や代謝物も含む全原子分子動力学計算は、計算機構杉田チームで開発中の分子・細胞シミュレータ GENESIS を用いて行った。GENESIS は「京」などのスーパーコンピュータでの並列化を意識して開発されたソフトウェアであり、数万ノードの並列計算を行っても十分な性能を発揮できる。平成 26 年度は 25 年度に引き続き、8000 ノード程度を用いた計算を実施した。2 年間の「京」を用いた計算によって、1000 万原子系については 130ns のプロダクションを実施し、1 億原子系については 20ns のプロダクションが終了した。

得られたトラジェクトリから、タンパク質や RNA の立体構造安定性、並進や回転拡散に関するパラメータを記述、タンパク質と代謝物の相互作用、タンパク質間相互作用の解析を行った。それぞれの分子の周囲の分子混雑度合いを調べたところ、並進と回転の拡散定数と良い相関が

見られた。この計算と比較するための流体力学効果を含めた Brownian dynamics を平成 26 年度に実施したが、全原子分子動力学計算で得られた並進拡散定数は、流体力学効果を考慮した Brownian Dynamics や実験で得られた値と良く一致した。タンパク質と代謝物の相互作用については、ATP だけでなく異なる性質の代謝物の運動性とタンパク質との相互作用を検討したが、実験では見えない新たな知見が複数発見されつつある。今後はタンパク質間相互作用の解析に重点をおいて解析結果をまとめて論文化していく予定である。

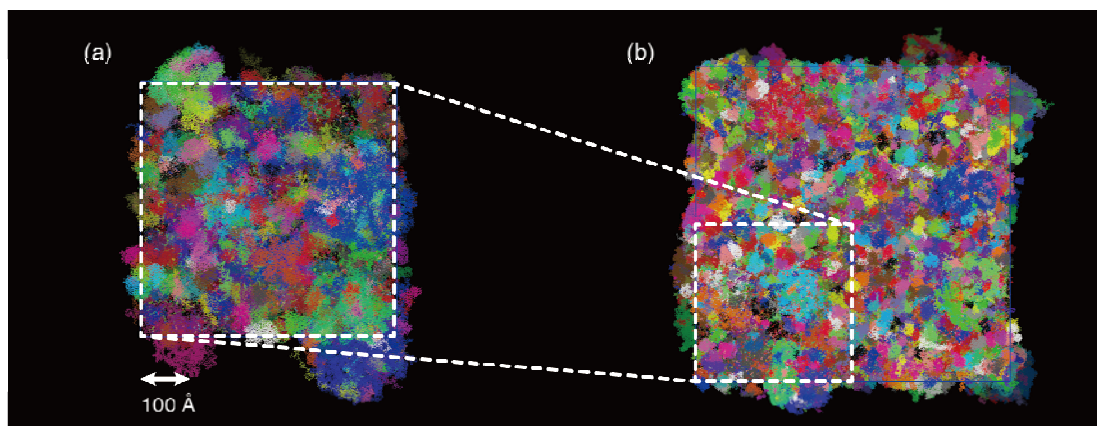


図 1. バクテリア細胞質の全原子モデルシミュレーションのスナップショット。(a) 1170 万原子系、(b) 1 億原子系。

2) タンパク質混雑系について、溶液 NMR で分子運動の検証を理研の木川と連携して試みた。緩和解析によって、個々のタンパク質の回転緩和は混雑環境に関わらず変化しないが、複数のタンパク質が協同的に動く緩和時間が混雑環境に影響していることが明らかになった。このような緩和解析は、希薄溶液中で用いられる通常モデルフリー解析では明らかにすることができず、より複雑なモデルを必要とした。タンパク質の回転緩和を高精度に計算するために、異なるモデルを用いた解析を行い分子動力学から信頼性のある回転拡散定数の導出を試みた。

## (2) 信号伝達経路上のリン酸化酵素の分子運動

1) EGF 信号伝達経路の下流に位置する MEK-ERK のリン酸化反応は、分子混雑の度合いに影響することが京大の青木によって実験的に示されており、この原因を分子モデルで説明することを目指している。この研究の困難は MEK-ERK 複合体の立体構造情報が未だに得られていないことである。昨年度までに、MEK と ERK それぞれの立体構造情報と生化学的な実験情報を用いて、タンパク質間ドッキングシミュレーションを行った。複数のドッキングサーバーを用いてシミュレーションを行い、構造を分類した後に、MM/PBSA を用いた自由エネルギー計算を実施し、エネルギー的に安定な候補をしばった。その結果 2 つの構造の候補が得られてきたが、粗視化分子モデルを用いて独立にシミュレーションを行った高田チーム(金田)の結果と比較した所、良い一致を得ていた。しかし、その後、新たなモデルが高田チームと林チームによって提案され、モデルの構築に関して実験家も含めた議論を複数回行った。

林チームと連携して、分子混雑環境が酵素反応自身に与える影響を調べるため、QM/MM 自由エネルギー計算に必要な初期構造モデルの構築に協力した。具体的には、酵素の構造安定性等に強い影響を与えないことが明らかになっている BSA を混雑物として選び、BSA が複数含まれ

る分子混雑環境の分子動力学計算を行った。得られた平衡構造を林チームに送り、ターゲットとする酵素を埋め込むことによって、QM/MM 自由エネルギー計算が進行している。

## I-2 高橋 恒一（理化学研究所）

### 細胞内環境下での信号伝達経路のモデリング基盤整備およびコード開発

#### I-2-1 実施計画

本研究では、「細胞内分子ダイナミクスのシミュレーション」研究の一環として、細胞内環境を考慮した細胞内信号伝達経路のモデリングとシミュレーション、またその実現に必要な計算手法の開発を行う。これまでの信号伝達経路のシミュレーションでは細胞内空間を陽に考慮しない反応ネットワークモデルや、巨視的な濃度勾配は考慮するが分子一つ一つの実体を解像しないメソスケールモデルが主だった。本研究課題では「京」の計算能力を活用し、1分子粒度かつ細胞まるごとスケールで、細胞表面の受容体が信号分子を感知してから核内で転写因子が遺伝子の発現のスイッチを入れるまでの一連の反応の計算を実現することを目指す。モデリングは、1分子計測とも連携しながら行う。

平成26年度の研究計画は、以下の通りである。(1) 反応拡散シミュレーションソフトウェア pSpatioocyte の開発に関しては、細胞内構造物等を再現するためのコンパートメント機能の追加およびプログラムのチューニングを実施する。(2) EGF 信号伝達経路モデルに関しては、細胞内環境（特に分子混雑）・細胞間のタンパク質量の違い等が、その経路の応答におよぼす影響を検証する。27年度にプロダクトランを予定しているため、その準備も行う。

#### I-2-2 実施内容（成果）

##### (1) pSpatioocyte の開発

平成26年度は、細胞形状および細胞内構造物を表現するためのコンパートメント機能の実装を主に行った。まず、実際の細胞の形状および細胞内構造物（細胞核）を、共焦点レーザー顕微鏡により観察し、その3次元スライス画像を取得した（図1A）。細胞種は、これまで実験に利用してきたPC12細胞を使用した。得られた画像データから、ImageJおよびV-CADを用いて、細胞形状および細胞核の構造を再構成した（図1B）。再構成したデータをpSpatioocyteに読み込むモジュールを実装し、pSpatioocyte上で可視化した結果を図1Cに示した。実験データより再構成した細胞形状および細胞核の構造を、pSpatioocyte上で精度よく再現することができた（図1）。今年度、pSpatioocyteに実装した読み込みモジュールは、3Dモデリングの分野でスタンダードなSTLフォーマット（Standard Triangulated Language）に対応しており、他の実験データも、このフォーマットで3次元データを作成することで、比較的容易に読み込ませることが可能である。

また、計算時間の大部分を占める拡散処理部に対する改良を行った。同じ拡散速度を持つ化学種を一括して処理することで化学種の数が増加した時に問題となる並列性能低下を改善した。上期から下期にかけてはOpenMPによるスレッド並列化の改良を行った。従来は一般的なcritical構文を用いたアクセス競合回避法を採っていたが、3スレッド以上では並列性能が劣化する問題があった。そこで制御の観点をスレッド側からオブジェクト側に移しlock変数を用いる独自の方法を考案した。これをpSpatioocyteに実装した結果、8スレッドまで速度向上が可能になった。さらに改良したプログラムを実際に「京」で32768ノードまで実行し、大規模並列時においても有効であることを確認した。

## (2) EGF モデリング・シミュレーション

平成 26 年度は、EGF 経路の細胞内反応部分である Raf/MEK/ERK 経路 (MAPK カスケード) のシミュレーションを実施し、分子混雑がその応答特性におよぼす影響を検証した。MAPK カスケードの反応スキームは図 2A に示すように、タンパク質の二重リン酸化モチーフが 2 つ連なった構造である。非混雑下でのシミュレーションの結果、入力/出力の応答曲線の非線形性は比較的高く、ヒル式でフィッティングしたところ、そのヒル係数およそ 3.6 であった (図 2B 紫)。一方、分子混雑条件では、入力/出力応答の非線形性が低下し、ヒル係数も 3.0 まで低下した (図 2B 緑)。特に低入力時での出力が上昇することにより、応答特性が変化した。このような分子混雑による応答の非線形性の低下は、実験でも確認されており (Aoki et al., 2011; 2013)、この結果と傾向は一致した。しかしながら、実際の実験では、分子混雑による反応系の応答特性の変化が今回のシミュレーション結果よりも大きく、その変化幅は一致しなかった。この原因として、我々のシミュレーションでは考慮できていない影響、特に分子構造の影響が大きいと考えられる。pSpatioocyte では、タンパク質分子を球状と仮定して、シミュレーションを行っている。しかしながら、実際のタンパク質の形状は様々で、その反応機構もまた複雑である。したがって、今回のシミュレーション結果は、分子混雑が詳細な反応過程にも大きな影響をおよぼしていること示唆した。

EGF 経路全体に関するシミュレーションでは、経路を構成する各タンパク質のばらつきと EGF 応答ばらつきとの関係を精査した。構築した EGF 経路の数理モデルには、EGF と受容体の結合から、ERK タンパク質の核内移行までが含まれる。EGF モデルの生物学的な妥当性を示すため、実験データとシミュレーション結果の比較を行った。そして、EGF 濃度に対するリン酸化 ERK および核内 ERK の応答の、実験値とシミュレーション値が一致することを確認した。次に、異なる EGF 濃度における核内 ERK 応答のばらつきを実験とシミュレーションで比較した。EGF 低濃度 (0.05 ng/ml) および高濃度 (50 ng/ml) において PC12 細胞で計測した核内 ERK 応答のばらつきを図 3A に示した。EGF 高濃度条件の方が、低濃度条件よりも分布の幅が狭くなっており、ばらつきが抑制されていることが分かる。同条件でのシミュレーション結果を図 3B に示した。実験とシミュレーション、両者は、非常によく似た分布を示しており、核内 ERK 応答のばらつきがよく再現できた。そして、各 EGF 濃度における核内 ERK タンパク質の分布が、パスウェイのどのタンパク質のばらつきに起因するかを調査した (図 3C)。その結果、EGF 低濃度条件においては、複数のタンパク質のばらつき (EGFR, Ras, Raf, MEK) により、核内 ERK 応答のばらつきが生み出されていることが分かった。一方、EGF 高濃度条件では、ERK タンパク質のみのばらつきにより、核内 ERK タンパク質のばらつきが生み出されていることが分かった。

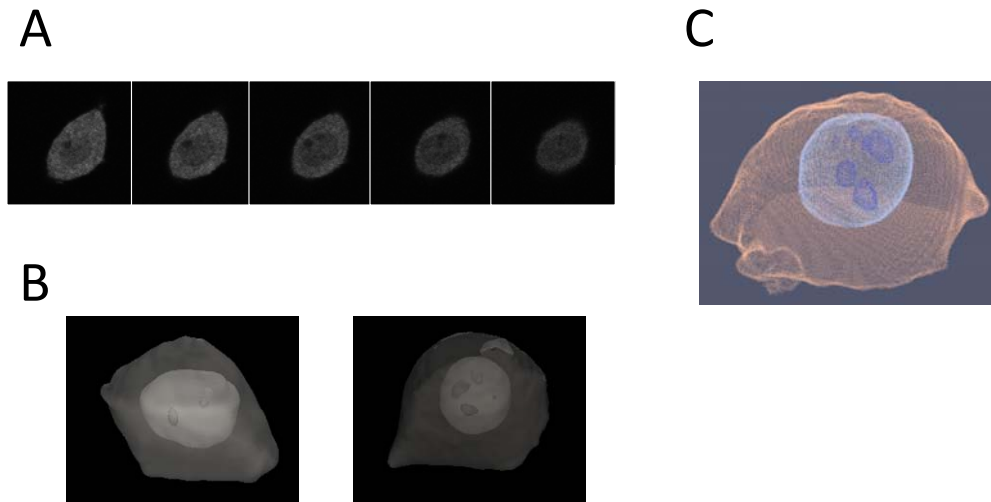


図1 細胞形状および細胞内構造物の再構成 A) 共焦点レーザー顕微鏡により観察したPC12細胞の3次元スライス画像。B) 3次元スライス画像から再構成した細胞形状および細胞核の構造。C) pSpatioocyte上で再現された細胞形状および細胞核。

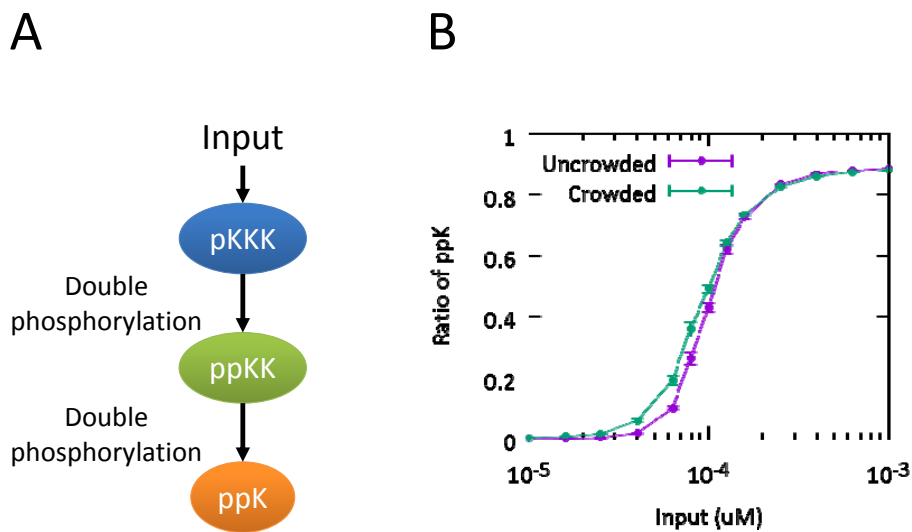


図2 分子混雑が応答特性におよぼす影響。A) MAPKカスケードの反応スキーム。二重リン酸化モチーフが2つ連なっている。B) MAPKカスケードの応答特性。非分子混雑下(紫)では、非線形性の強い応答特性を示す(ヒル係数は3.6)。分子混雑下(緑)では、非分子混雑下に比べて、応答特性の非線形性が低下した(ヒル係数は3.0)。

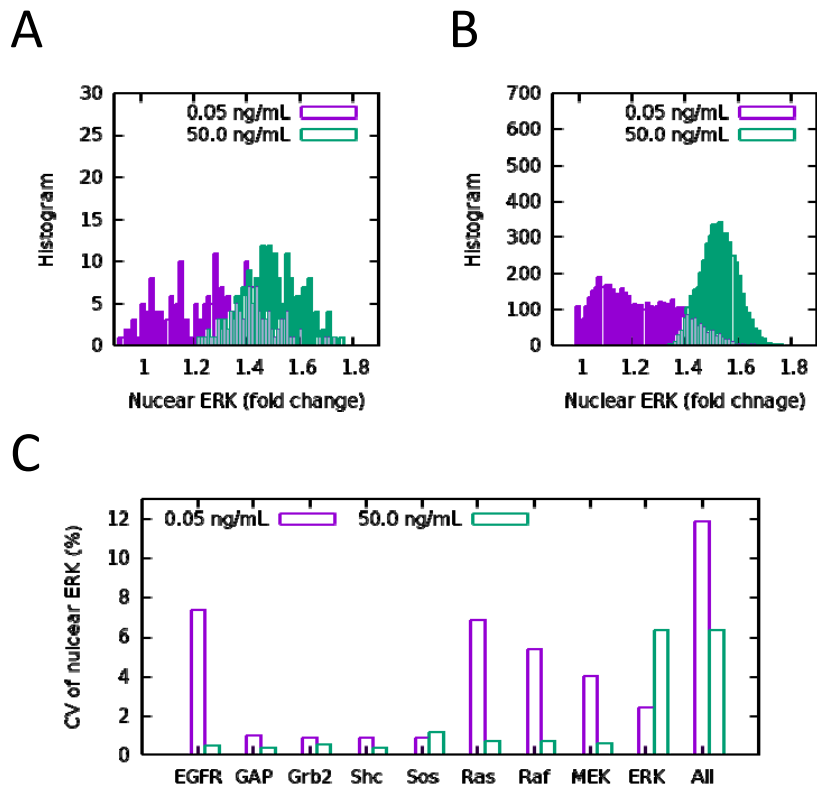


図3 EGF 応答ばらつきとタンパク質ばらつきの関係。A) PC12 細胞で観測された核内 ERK 応答のばらつき。B) 核内 ERK 応答のばらつきのシミュレーション。C) 各タンパク質ばらつきが核内 ERK 応答のばらつきにおよぼす影響。

### I-3 林 重彦 (京都大学)

#### 信号伝達経路上のリン酸化酵素の反応性解析

##### I-3-1 実施計画

本研究では、「細胞内分子ダイナミクスのシミュレーション」研究の一環として、信号伝達経路上のリン酸化酵素の反応性解析のための研究開発を実施する。細胞内信号伝達は、複数のリン酸化酵素のリン酸化反応により行われる。それらのリン酸化酵素は、その酵素自身が別のリン酸化酵素によりリン酸化を受けることにより、リン酸化活性が変化し、信号伝達が制御される。従って、リン酸化酵素の反応活性は、細胞内信号伝達の鍵となる化学的過程であり、主な創薬ターゲットの一つとなっている。そこで、細胞環境内におけるリン酸化酵素の反応活性の分子機構を明らかにし、細胞内信号伝達の理解と制御に向けた分子論的基盤を構築することを目指す。

平成 26 年度では、MAP リン酸化酵素である ERK と、MAPK リン酸化酵素である MEK のそれぞれに対して、リン酸化活性と相関すると考えられる ATPase 活性の化学反応性を、ハイブリッド QM/MM 自由エネルギー法を用いて解析する。昨年度までに行なってきたリン酸化による自由エネルギー極小構造変化の解析を更に推し進め、リン酸化による反応活性制御の分子機構を明らかにする。また次年度の細胞内環境におけるシミュレーションを目指し、大規模系に適用可能なハイブリッド QM/MM 自由エネルギー法の開発を行う。

##### I-3-2 実施内容 (成果)

###### (1) MEK 及び ERK の ATPase 活性の解析

MAPK リン酸化酵素である MEK1 に対して、分子動力学 (MD) シミュレーションを用いて、ATPase 活性に対するリン酸化の影響を解析した。MEK1 は、自身がリン酸化を受けることにより MAP リン酸化酵素 ERK をリン酸化し活性化する。また、MEK1 は、単体でも自身が受けたリン酸化により ATPase 活性が上昇することが知られている。そこで、MD シミュレーションにより、MEK1 のリン酸化前後の構造変化を解析することにより、リン酸化の ATPase 活性に及ぼす影響の分子機構を解析した。

MEK1 の初期構造は、重要なループ領域が見えている阻害剤を結合した X 線結晶構造を用いた (PDB-ID: 3PP1)。阻害剤を取り除き、周期境界条件の水中に入れた (図 1a)。基質は ATP 及び  $Mg^{2+}$  がタンパク質に結合している。タンパク質中の titratable な残基は、すべて標準プロトン化状態のままである。系を中性化するため、 $Na^+$  イオンを 7 個加えた。系の原子数は 67,916 原子である。分子力場は、タンパク質・イオンには Amber ff03、水には TIP3P を用い、ATP に対しては Meagher ら (J. Comput. Chem. 24, 1016 (2003)) が開発した分子力場を用いた。MD 計算は、AMBER を用いて GPU 上で行った。NPT アンサンブルで 300 K の平衡化を行い、さらに NVT アンサンブルで 100 ns の平衡化を行った。その後、リン酸化されたタンパク質 (MEK1-PP) モデルを、S218 及び S222 をリン酸化することにより構築した。リン酸はすべて脱プロトン化しているとした。リン酸化セリン残基の分子力場は Homeyer ら (J. Mol. Model. 12, 281 (2006)) の開発したものをを用いた。系の中性化のために  $Na^+$  イオンを 4 個加えた。これらのリン酸化を受けていないタンパク質 (MEK1) 及び MEK1-PP のシミュレーション系に対し、上記の平衡化後からそれぞれ 3 マイクロ秒の MD シミュレーションを行い、構造変化を解析した。



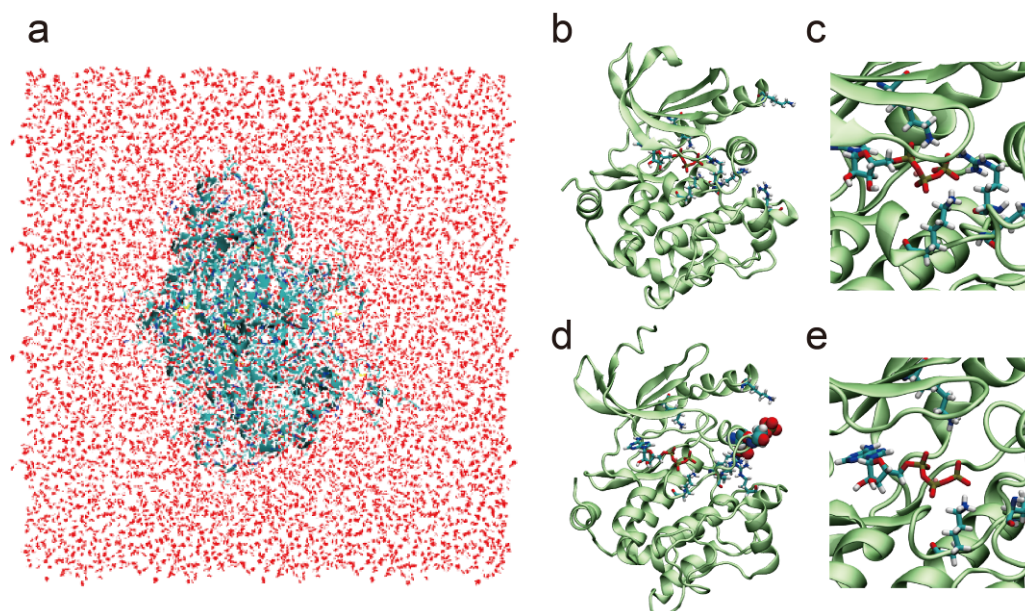


図 1: MEK1 の MD シミュレーション。a シミュレーション系。b,c リン酸化を受けていない MEK1 の 3 マイクロ秒後のスナップショット：全体構造 (b) と ATP 分子結合部位 (c)。d,e リン酸化した MEK1-PP の 3 マイクロ秒後のスナップショット：全体構造 (d) と ATP 分子結合部位 (e)

図 1b-d に 3 マイクロ秒後の MD シミュレーションのスナップショットを示す。リン酸化されていない MEK1 では Arg227 のグアニジニウム基が ATP の  $\gamma$  リン酸基に強く配位している構造が観測された。この配位により、ATP 加水分解を引き起こす攻撃水が排除されると共に、 $\gamma$  リン酸の強い安定化により、ATPase 活性が抑えられていると考えられる。一方、リン酸化を受けた MEK1-PP では、Arg227 のグアニジニウム基が、すぐ前の S218 及び S222 に付加したリン酸基に配位することにより、ATP の  $\gamma$  リン酸基周辺を大きく開け、ATP 加水分解が可能な構造となることを明らかにした。また、ERK の MD シミュレーション系の構築を開始した。

## (2) Ras-GAP 複合体の MEK 及び ERK の ATPase 活性の解析

MEK-ERK 複合体のモデルとして、信号伝達経路の上流にある GAP-Ras 複合体の GTP 加水分解反応に関する研究を行った。前年度までに QM/MM 自由エネルギー法を用いて決定した、大きな構造変化を伴う複合体構造の反応の始状態及び遷移状態に対して、その間の自由エネルギー差 (すなわち反応活性化自由エネルギー) を、Bennett acceptance ratio 法を用いた自由エネルギー摂動法により求めた。その際、遷移状態生成に伴う大きな構造変化を十分記述するために、1 マイクロ秒の自由エネルギー摂動計算を行った。その結果、反応活性化自由エネルギーは 16.3 kcal/mol と見積もられ、実験値 (15.9 kcal/mol) と良く一致した。そこで、酵素反応活性に対する遷移状態生成に伴う大きな複合体構造変化の影響を明らかにするために、大きな複合体構造変化が起きない 10 ナノ秒の自由エネルギー摂動計算を行った。その結果、反応活性化自由エネルギーは 22.2 kcal/mol となり、1 マイクロ秒の自由エネルギー摂動計算の結果に比べて非常に高い反応障壁となった。すなわち、本研究で見出された反応遷移状態生成に伴う大きな複合体構造変化は、反応障壁を大きく下げ、酵素反応活性に大きな役割を果たしていることを明らかにした。

上記の様に、酵素反応の反応遷移状態生成において、大きな複合体の構造変化が重要な役割を果たしている場合には、細胞内における混雑環境効果が酵素活性に重要となる可能性がある。

そこで、混雑環境内での Ras-GAP 複合体の酵素活性を調べるために、BSA 混雑環境下での反応始状態の自由エネルギー構造最適化を、QM/MM 自由エネルギー法を用いて行った。杉田グループより提供された BSA 混雑環境系に、上記の Ras-GAP 始状態の構造をはめ込み、226,414 原子のシミュレーション系を構築した (図 2)。また、QM/MM 自由エネルギー計算のために、MD シミュレーションによるサンプリングを行う GENESIS に改良を加えた。現在、QM/MM 自由エネルギー法による自由エネルギー構造最適化計算が進行中である。

### (3) 「京」における QM/MM 自由エネルギー法の計算コードの改良

QM/MM 自由エネルギー法の QM/MM 部分は、GAMESS を使用しているが、これまでの「京」における使用においては、GAMESS のデータサーバが 1 ノード内のコアの半分を使用しているため、演算の効率が低かった。そこで、通常の使用で用いる密度汎関数法による計算の場合に、データサーバを使用しないようにすることにより、すべてのコアを演算に使用することを可能にし、計算効率の 2 倍の向上を達成した。図 3 に FX10 におけるベンチマーク系 (QM 分子: trialanine 33 原子、MM 領域: 水溶媒 27 万原子、153,600 サンプル) の並列化効率を示す。1,000 コアを超えても非常に良い並列化効率が得られている。

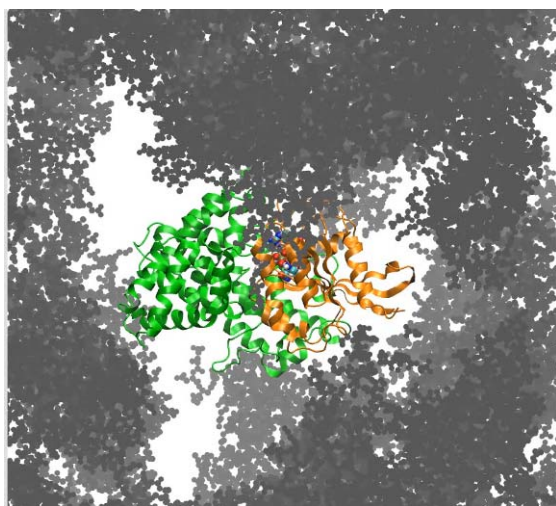


図 2: BSA 混雑環境中の Ras-GAP 複合体のシミュレーション系。

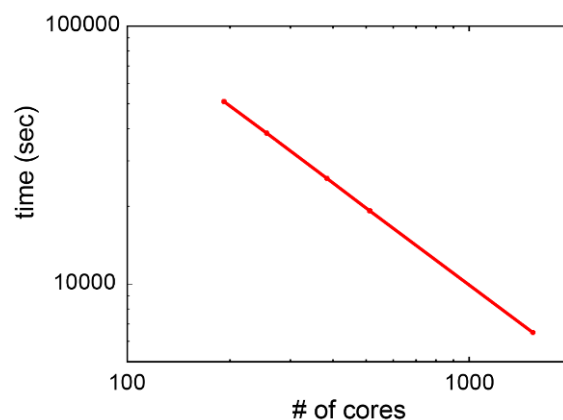


図 3: FX10 におけるベンチマーク系の並列化効率。

## I-4 高田 彰二 (京都大学)

### 粗視化分子モデルを用いた信号伝達経路上のリン酸化酵素複合体とクロマチンの動的モデリング

#### I-4-1 実施計画

本研究では、戦略課題1の中で、粗視化分子シミュレーション技法を適用して、1) 哺乳類における信号伝達経路の代表的モデルの一つである MAPK 系のリン酸化カスケードの構造動態に基づく理解と予測および、2) 核内混み合い環境におけるヌクレオソーム多量体からなるクロマチンモデルの構造動態と、動的なクロマチン上の遺伝子情報の検索過程の理解と予測を、それぞれ目指す。必要に応じて、密接に関連する実験研究者と連携し、シミュレーションに基づいた予測を実験研究者にフィードバックする。

平成26年度は、粗視化シミュレータ CafeMol を用いて、細胞内環境における信号伝達分子リン酸化酵素複合体 MEK-ERK の動的モデリングを行なうとともに、核内環境におけるモデルクロマチンの折れたたみ構造とその構造における転写因子の動態を解析する。

#### I-4-2 実施内容 (成果)

##### (1) 信号伝達経路上のリン酸化酵素複合体の動的モデリング

1) **研究の背景:** 細胞が周囲の環境に適応するためには、外部からの刺激を核内まで伝達させて遺伝子発現を変化させる機構が必要となる。具体的に、シグナル伝達経路の代表的モデルの一つである哺乳類の MAPK カスケードでは、外からの刺激を契機に、Ras, Raf (MAPKKK) が活性化し、さらに MEK1 (MAPKK)、ERK2 (MAPK) へと次々に上流のタンパク質が下流のタンパク質をリン酸化することで活性化する。この活性化(リン酸化)反応の連鎖によって最終的に信号が核内まで伝達される。MAPK カスケードは、細胞の増殖、分化、アポトーシスの決定のみならず、細胞のがん化とも強い関連を持っている事が知られているので、伝達における活性化機構の解明は医学的な見地からも強く望まれる。しかし、その活性化の分子機構の詳細は未だに不明である。また、ERK2 タンパクはリン酸化された後に核内に移行し転写因子と相互作用すると考えられる。しかし、クロマチン構造(ヌクレオソーム)が高濃度に存在している核内混み合い環境において、(ERK2 タンパクが) どの様に振る舞う(拡散運動する)のかその詳細は分かっていない。

**研究の目的:** 本研究の目的は哺乳類の MAPK カスケードにおける信号伝達の分子機構に迫る為に、(実験的手法では未だに明らかにされていない) MEK1 (MAPKK) と ERK2 (MAPK) の複合体の構造予測(モデリング)をすることである。

**方法:** MEK-ERK の複合体構造の予測の為に我々は、以下2つの方法(i, ii)を各々独立に遂行した。

(i) 我々が開発してきた粗視化シミュレーションソフトウェア: CafeMol を適用し、MEK-ERK ドッキングの粗視化シミュレーションを行う。その際、分子相互作用については AICG2 モデルを、分子間相互作用については、Debye-Hückel 型の静電相互作用、疎水性相互作用と斥力相互作用を考慮する。サンプリングの際は、室温(300K)で Underdamped の Langevin シミュレーションを遂行し、その中から分子間相互作用エネルギーの低い構造を有力(最適)構造として抽出した。

(ii) 由良敬氏（お茶の水女子大）と亀田倫史氏（産総研）の協力を得て、タンパク質ドッキングソフトウェア ZDOCK によりサンプリングされた MEK-ERK 複合体構造(3000 サンプル)から、KYG-PROTEIN 法を用いたスコアリングにより有力構造を抽出した。

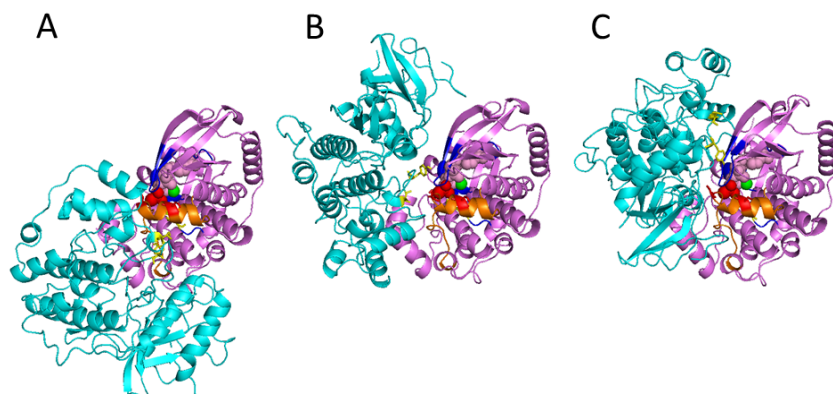


図 1: 得られた MEK-ERK 複合体のモデル構造(3つの有力候補)

我々は、(i)粗視化シミュレーションにより図1-A, Bの2つの有力構造を、(ii) KYG-PROTEIN 法[由良、亀田]により図1-Cの有力構造を1つ、それぞれ構築した。これらの3つの複合体では、MEKの触媒部位とERKの活性化ループ(リン酸化サイト)がある程度近くに存在する事が分かった。特に、粗視化シミュレーションにより構築した二つの構造(1-A, B)は、我々とは全く独立な方法[Wang等(杉田研)]により作成された代表的な複合体構造とよく似ている事が分かった。[Wang等はClusProによりサンプリングした構造をMMPBのフリーエネルギーによりスコアリングして構造抽出した。彼らの構造と我々の構造はMEK-ERK結合界面のコンタクトマップ等が似ている。] 今後は今回得られた3つの複合体構造の中から最も有力な構造を選出する為に、全原子MDによる調査を遂行する必要がある。

## (2) クロマチンの動的モデリング

### 1) ヌクレオソーム3量体の構造動態とセントロメア特異的構造

セントロメアは、動原体が結合するサイトを作るクロマチンの特別な領域であり、そのことによって染色体の分離を可能にしている。動原体の形成機構を詳細に知るために、セントロメア特異的なクロマチンの構造情報が不可欠であるが、これまでにその構造はほとんど明らかになっていなかった。我々は、合成されたヌクレオソーム3量体のX線溶液散乱データと粗視化分子シミュレーションによって、二つのヌクレオソーム3量体の構造を明らかにした。対象とした系は、H3-CA-H3とH3-H3-H3である。ただし、H3はカノニカルなヌクレオソーム、CAはセントロメア特異的なH3バリエーションであるCENP-Aを含むヌクレオソームである。両方のトリヌクレオソームについて、7つの塩濃度で粗視化分子動力学シミュレーションを行った(図2)。シミュレーションで得られるSAXSプロファイルと実験のSAXSプロファイルを比較することで、実験とつじつまがあう構造アンサンブルを得ることが出来た。実験と符合する計算条件において、H3-CA-H3系では、1つ目と3つ目のヌクレオソームがドッキングし、間のリンカーDNAはX字型に交差していなかった。対照的に、H3-H3-H3はリンカーDNAがクロスした構造をとっている。さらに、たくさんのH3ヌクレオソーム多量体のなかに、CAヌクレオソームを1分子だけ含めた構造のモデルを作成すると、CAのある場所でクロマチンファイバーが折れた。これは、セントロメアが動原体マシナリーの結合サイトを与える足場に

なる可能性のある構造であり、CAヌクレオソームの機能が示唆された。

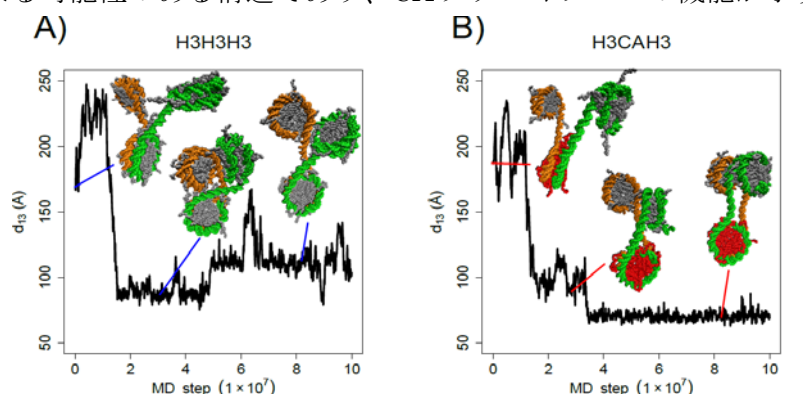


図2: トリヌクレオソーム H3-H3-H3 と H3-CA-H3 の分子シミュレーション。H3-H3-H3 では、二つのリンカーDNA がX字型にクロスすることが多いが、セントロメア特異的な H3-CA-H3 では、リンカーDNA はほぼ平行にならび、最初と最後のヌクレオソームがドッキングすることが多い。

## 2) アセチル化によるヌクレオソーム・クロマチン構造動態の変化

真核生物においては、さまざまなエピジェネティックな制御によって、クロマチン構造が多様に変化しており、その変化が遺伝子発現制御の主たる機構になっている。ヒストンアセチル化は、代表的なエピジェネティック制御であり、例外はあるものの、概ねクロマチン構造をほどくことによって遺伝子発現を活性化するように働くことが知られている。しかし、どのヒストンテールのアセチル化がどのような仕組みで発現を制御するのか、その詳細には不明な点が多い。本研究では、ヌクレオソーム3量体からなる系を例にとって、様々なヒストンテールアセチル化が、その構造をどのように変えるのかを研究した。計算には、我々が開発を進めてきた粗視化生体分子シミュレーションソフトウェア CafeMol を用いた。アセチル化はリジン残基への化学修飾であり、もとのリジンが持つ-1価の電荷がゼロになる。粗視化モデルの範囲内では、アセチル化を電荷の変化で表現した。

ヒストンテールアセチル化の影響を調べるために、次の6個のトリヌクレオソーム系の計算を行った: 非アセチル化、H3テールのみアセチル化、H4テールのみアセチル化、H2A テールのみアセチル化、H2B テールのみアセチル化、およびすべてのテールのアセチル化。各ヒストンのアセチル化サイトは、実験的に知られているものを使い、同じテールにある複数のアセチル化可能なリジンはすべて同時にアセチル化させることにした。各系において、 $10^8$  MD step のシミュレーションを独立に10回繰り返した。構造のコンパクトさを様々な指標で調べた。図3は、慣性半径  $R_g$  の分布を示したものである。図にあるように、H4テールアセチル化の影響がもっとも大きく、トリヌクレオソームをより広がった構造に変化させた。他のヒストンのアセチル化は、それよりは小さな影響であった。さらに、詳細な解析の結果、H4テールアセチル化は、そのヌクレオソームに隣接するヌクレオソームとの相互作用を弱くしていることが分かった。それは、H4テールの正電荷が減少し、それが隣接するヌクレオソームの DNA との相互作用を弱めているためであった。

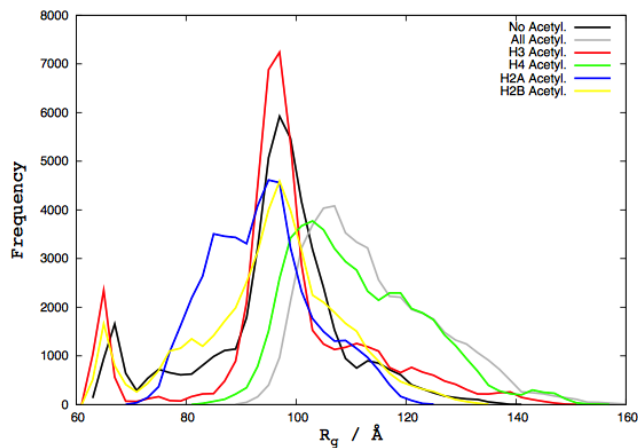


図 3 : 様々なアセチル化状態のヌクレオソーム 3 量体の慣性半径の分布図。

### 3) クロマチン構造と転写因子動態の連関

核内に移行した ERK2 蛋白質の核内混み合い環境における拡散運動を調査した。核内に移行した ERK2 タンパク質の拡散運動を調査する為に、我々が開発してきた粗視化分子動力学シミュレーター CafeMo1 を適用した。核内混み合い環境を実現する為に、20 ヌクレオソームからなるクロマチン構造を系に含めて計算を行った(ヌクレオソーム濃度が 0.1mM (図 4 左) と比較的低濃度のケースと高濃度[0.5mM]のケース (図 4 右) の 2 通りを調査した)。各ヌクレオソームのヒストン部位及び ERK2 タンパク質に対しては AICG モデルを適用し、2 本鎖 DNA に対しては Pablo グループ等(2009)により開発された DNA モデルを適用した。ERK2 タンパクとクロマチン間の分子間相互作用については Debye-Hückel 型の 静電相互作用と斥力相互作用を考慮した。我々はこのモデルに対して室温(300K)、イオン強度(0.2M)の条件で Underdamped の Langevin シミュレーションを遂行した。

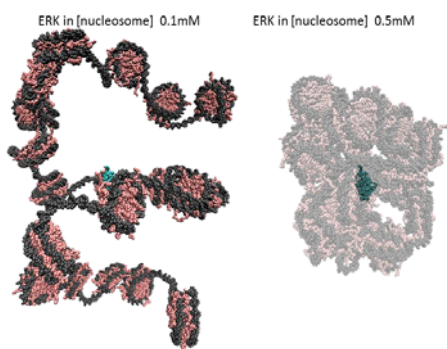


図 4 : 核内混み合い環境における ERK 蛋白質(スナップショット)

各ヌクレオソーム濃度の場合、およびヌクレオソーム・クロマチンがない溶液中の場合について、平均 2 乗変位の時間依存性から拡散係数を計算した (図 5)。比較的大型の転写因子 p53 は、ヌクレオソーム 0.1 mM 環境下では拡散可能であるが、ヌクレオソーム 0.5 mM になるとほとんど引っかかっており拡散することが出来なくなる。小型の転写因子 HMGB1 の場合、0.5 mM ヌクレオソーム環境下でも、遅いながらも拡散可能であり、拡散係数は溶液中の約 4 割程度になった。また、HMGB1 と同程度の非転写因子である ERK の場合、0.1 mM, 0.5 mM と徐々に拡散係数は減少するものの、0.5 mM ヌクレオソームでも溶液中の 6 割以上の拡散係数であった。

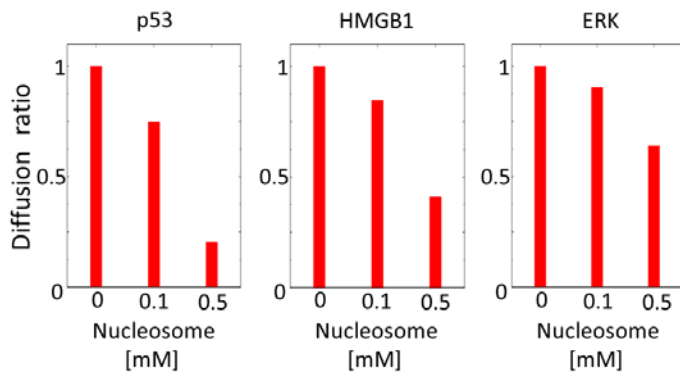


図 5：核内混み合い環境における ERK 及び転写因子(p53, HMGB1)の拡散係数(ヌクレオソームが無い環境における拡散係数との比)

## I-5 池口 満徳 (横浜市立大学)

### 核内DNA結合タンパク質の機能ダイナミクス研究

#### I-5-1 実施計画

本研究では、「細胞内分子ダイナミクスのシミュレーション」研究の一環として、核内DNA結合タンパク質の機能ダイナミクスの研究を実施する。核内DNAは、ヒストンタンパク質にDNAが巻き付いたヌクレオソーム構造を基本構造とし、複数のヌクレオソームが凝集した構造をとっている。近年、このヌクレオソームに代表される核内DNA結合タンパク質複合体の構造やダイナミクスが遺伝子の発現や維持に深く関わっていることが明らかになってきた。本研究では、そのような核内DNA結合タンパク質の機能ダイナミクスのあり方を、実験系と連携しつつ、分子シミュレーションを用いて解明することを目指す。

平成25年度に行った、ヌクレオソーム単量体の全原子分子動力学 (MD) シミュレーションとX線小角散乱 (SAXS) 実験の連携研究 (all-atom MD-SAXS) を踏まえ、平成26年度には、より大規模な対象へ適用可能な、粗視化分子動力学シミュレーションとX線小角散乱 (SAXS) 実験の連携研究 (CG MD-SAXS) 法の開発を行うとともに、それをヌクレオソーム構造に応用する。実験情報と合致するような、ヌクレオソーム構造の粗視化パラメータを導出し、ヌクレオソーム構造の機能ダイナミクスの解明を目指す。

#### I-5-2 実施内容 (成果)

##### (1) ソフトウェアの開発・高度化の状況

分子シミュレーションと実験を密に連携させるために、粗視化分子動力学シミュレーション (MD) 法とX線小角散乱 (SAXS) 実験法を組み合わせる粗視化MD-SAXS法の開発を行った (図1)。この粗視化MD-SAXS法のソフトウェアは、「京」において高並列のMPI並列計算が実行できるようにした。この手法を様々な系に対してテスト計算することで、粗視化MD-SAXS法の有用性を検討した。これにより、粗視化分子動力学シミュレーションとX線溶液散乱実験を密に連携させることが可能となった。

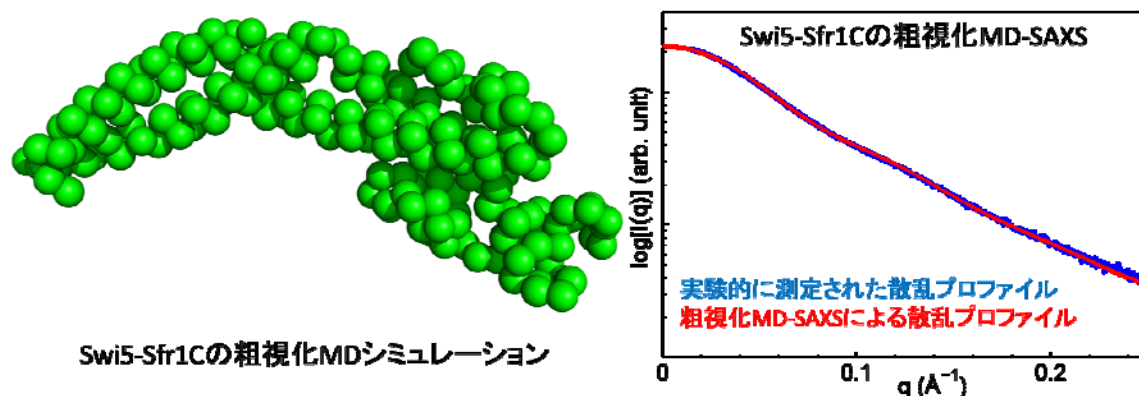


図1 粗視化MD-SAXS法によるシミュレーション結果と実験結果の比較



## (2) 研究開発の実施状況

### 1) 様々なヒストンバリエントを含んだヌクレオソームの粗視化 MD-SAXS 研究

今年度は、核内 DNA 結合タンパク質のうち、ヒストンバリエントの種類の異なる様々なモノヌクレオソームに着目して粗視化 MD-SAXS 解析を行った。ヌクレオソームに含まれるヒストンバリエントの違いにより、遺伝子発現制御に変化が生じることが報告されており、ヒストンバリエントによるヌクレオソームの構造変化は、エピジェネティクスの観点から重要である。粗視化 MD-SAXS の結果、カノニカルヌクレオソームは、メジャーな状態として、溶液中で両方の DNA 末端がコアダメインに巻き付いた状態をとり、また、マイナーな状態として、片方の DNA 末端がコアダメインから離れた状態をとっていることが分かった (図 2)。さらに、ヒストン H3 が CENP-A というヒストンバリエントに置き換わっているヌクレオソームでは、両方の DNA 末端がコアダメインから少し離れた構造をとっており、河野らの全原子シミュレーションとよい一致を示すことがわかった。さらに、ヒストン H2A が、H2A.B というヒストンバリエントに置き換わっているヌクレオソームでは、両方の DNA 末端がかなりコアダメインから離れた構造が多いことが分かった。これは、ヒストンバリエント H2A.B が、密に凝集しているヘテロクロマチンではなく、転写活性が高く、緩い状態にあるユークロマチンという領域で多く観察されるという実験事実とよく一致している。

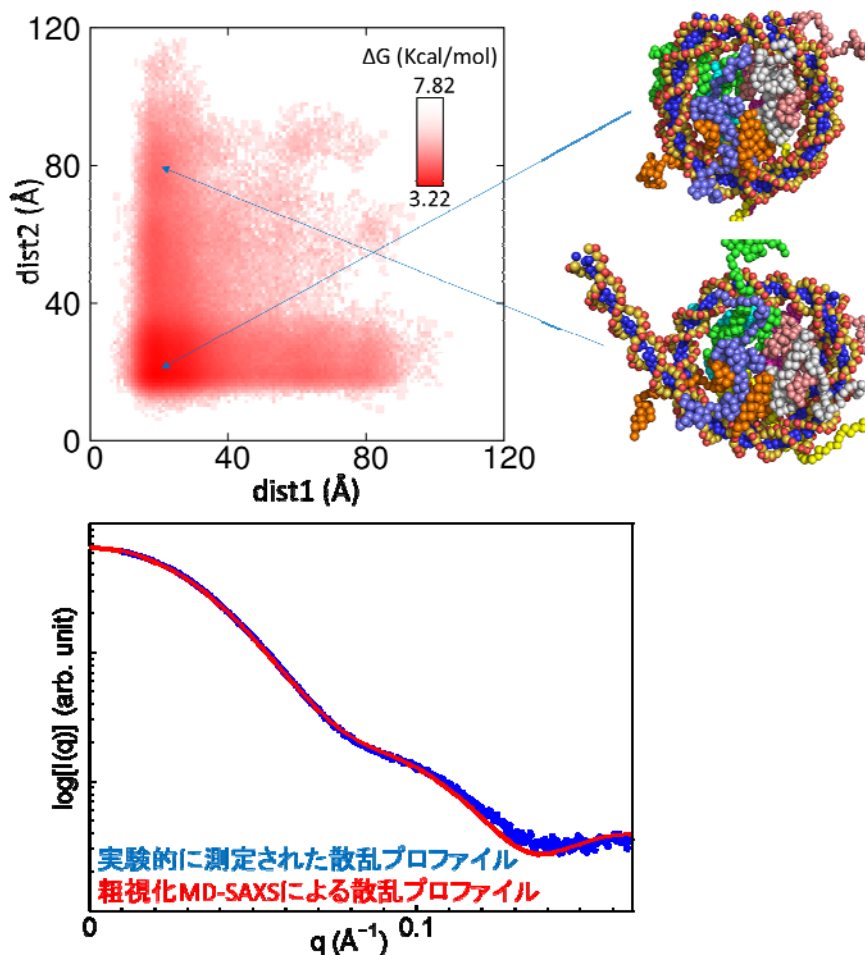


図 2 粗視化 MD-SAXS 法によるカノニカルヌクレオソームの自由エネルギー地形と X 線小角散乱のデータの比較。X 線小角散乱データは、横浜市立大学の佐藤衛研究室より提供された。

## I-6 河野 秀俊（日本原子力研究開発機構）

### 全原子モデルにもとづくヌクレオソームポジション変化の自由エネルギープロファイル計算

#### I-6-1 実施計画

本研究では、「細胞内分子ダイナミクスのシミュレーション」研究の一環として、核内DNAタンパク質複合体の構造予測と機能解明のための研究開発を実施する。核内DNAは、ヒストンタンパク質に約150塩基対のDNAがほぼ2回巻きついたヌクレオソーム構造を基本構造とし、このヌクレオソームがコンパクトに凝集した構造をとっている。近年、このヌクレオソーム構造のポジションやその変化が遺伝子制御に深く関わっていることが明らかになってきた。また、エピジェネティクスの観点からも、ヌクレオソームを構成するカノニカルヒストンとその変異体の違い、ヒストンの修飾及びDNAの修飾が遺伝子発現制御や細胞分化などにおいて重要な役割を果たしていることがわかってきた。そこで、これらの修飾や変異体とヌクレオソームポジション変化の関係を明らかにすることにより、遺伝子発現制御や細胞分化の分子メカニズムの一端を明らかにすることを目指す。

また、「全原子モデルにもとづくヌクレオソームポジション変化の自由エネルギープロファイル計算」の研究を行う上で、関連する研究者と必要な協議等を行うとともに、本格実施に必要な研究体制の整備を行う。

平成26年度は、昨年度に引き続き、ヒストンヴァリエントを含むヌクレオソームについて、巻き付いているDNAがほどける過程の自由エネルギープロファイル計算を行う。昨年度はヌクレオソームが約ひと巻まで解けた過程を計算した。今年度は、残りのひと巻が解ける過程を、カノニカルヌクレオソームとヒストンヴァリエントを含むヌクレオソームについて行う。また、ヒストンタンパク質が化学修飾(アセチル化)されたヌクレオソームでも同様な計算を行い、化学修飾の効果を定量的に調べる。これらすべての全原子モデルのシミュレーション結果を粗視化シミュレーションの力場決定に資することで、チーム内で連携して研究を進める。

#### I-6-2 実施内容（成果）

##### (1) ソフトウェアの開発・高度化の状況

自由エネルギープロファイルを計算するにあたり、2つの原子集団の重心間距離を反応座標とできるように、プログラムの改変を行った。また、少数のアミノ酸置換などわずかな違いをもつタンパク質の分子動力学計算は、お互いに非常に似通ったダイナミクスを示すことが多いため、ダイナミクスの違いがアミノ酸置換によるものか本来の揺らぎの範囲なのかを見極めることが非常に難しい。そこで、機械学習によく用いられている Linear Discriminant Analysis (LDA)法を導入し、有意な差を見つけ出す方法を開発した。LDAでは、同じ条件のシミュレーション結果を射影した分布はその平均の周りによくクラスター化するようにし、一方、違う条件のシミュレーション結果の射影は上記の平均から分離するように、繰り返し射影する空間を修正していく方法である。結果、これまで主成分解析では見つけることができなかった違いを発見することができた。この方法は、主成分解析と並び、新たなデータ解析方法を与える (Sakuraba and Kono, 論文準備中)。

##### (2) 研究開発の実施状況

###### 1) 2種類のモノヌクレオソームのDNA解離過程の自由エネルギー計算

ヒストン組成の違いがヌクレオソーム DNA の解離にどのような違いをもたらすか調べるために、カノニカル H3 ヒストンとそのヴァリエント CENP-A（染色体の均一な分配に重要でセントロメアに局在）をそれぞれもつヌクレオソームについて、DNA 解離の自由エネルギープロファイル計算を行っている。DNA の末端間距離もしくはヒストンタンパク質の重心と DNA 両端の塩基対の重心を反応座標にとり、adaptively biased MD 法及びアンブレラサンプリング法により、自由エネルギープロファイルを計算した。図 1 に示すように、今年度までに約半分の DNA が解離する過程の計算が終了した。結果、H3 と CENP-A が直接 DNA と相互作用する両末端の DNA の解離(図 1 の 1)では、自由エネルギーに大きな差が見られ、CENP-A の DNA は両末端が解けた状態が安定であることが分かった。また、その要因となるアミノ酸残基も特定された。

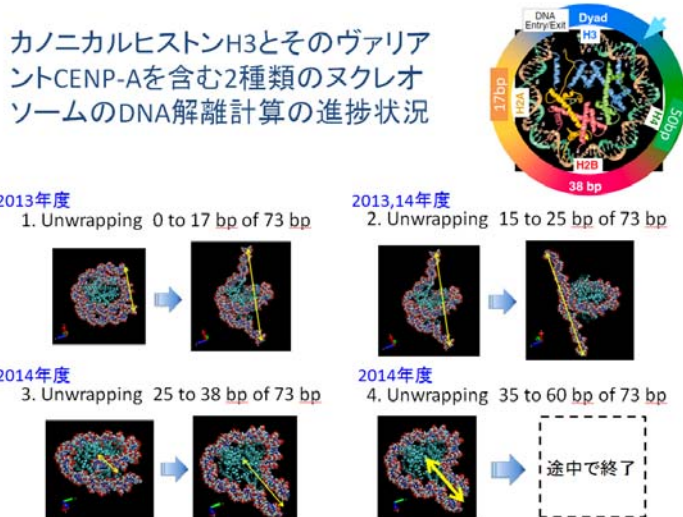


図 1. DNA 解離過程のヌクレオソームの構造 (H3 ヌクレオソーム)

図 1 の 2, 3 に対応する解離過程では、DNA は H2A, H2B ヒストンとの相互作用を失う。この部分は、2 つのヌクレオソームに共通であるため、自由エネルギープロファイルも類似すると期待された。実際、計算された自由エネルギープロファイルから、解離にかかるエネルギーコストは、2 つのヌクレオソームでほぼ同じ値を示し、それぞれ約 15、40kcal/mol であった(図 2)。図 1 の 2 の過程では、自由エネルギープロファイルのカーブに差が見られた(図 2 上段)。これは解離の過程で、H3 ヌクレオソームでは両末端がほぼ均等に解離したのに対し、CENP-A ヌクレオソームでは片側の DNA が先に解離し、あとで反対側の DNA が解離したことによる。図 2 の下段に示すように、Dyad の正反対の付近の DNA が解離する(図 1 の 3 に対応)には、大きなエネルギーコストがかかることが分かった。一定の力で DNA を引っ張ってヌクレオソーム DNA を解離させた 1 分子実験でも、DNA 末端から 23 から 33bp 付近の解離に時間がかかることが報告されている(Hall, M. et al. (2009). *Nat Struct Mol Biol* **16**, 124-9)。本シミュレーション結果は、その結果によく対応する。Dyad の正反対の付近まで DNA が解離すると、DNA はヒストンにひと巻の状態になる。その状態では、DNA 同士の静電的な反発力がなくなり、結果的に DNA はより強くヒストンに巻き付いた状態になる。従って、Dyad の正反対付近の DNA の解離には大きなコストがかかると思われる。

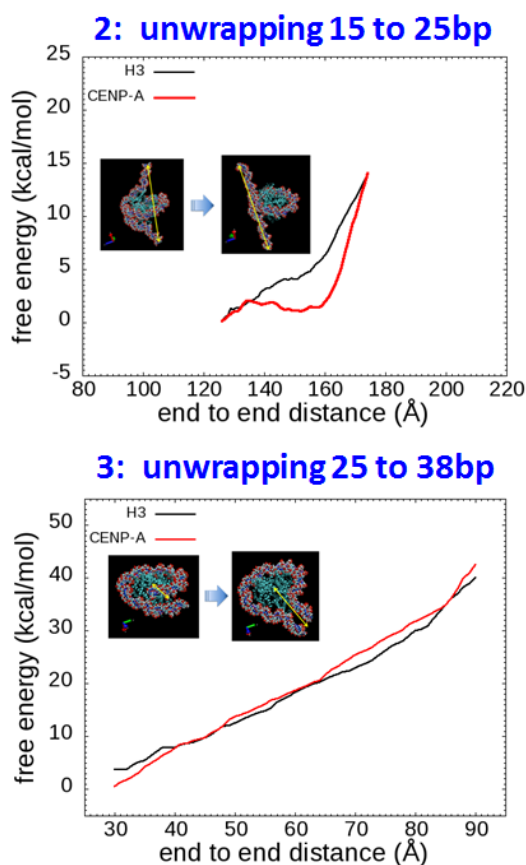


図 2. 自由エネルギープロファイル

## 2) メチル化 DNA が巻き付いたヌクレオソームの分子動力学計算

DNA のメチル化は、発生と細胞分化において非常に重要な役割を担っており、エピジェネティクスに深く関わっている。ヒトなどの場合、メチル化は CG 配列のシトシン(C)に対して特異的に起こる。そこで、DNA メチル化がヌクレオソームの立体構造に与える影響を調べるために、胡桃坂ら(早稲田大学)によって決定された、メチル化 DNA を持つヌクレオソームの結晶構造をもとにシミュレーション計算を行った。この構造は、12 のメチル化部位を持つ。H26 年度は、異なる初期速度を持つ独立な 100ns の計算を 2 本、さらにコントロールとしてメチル化されていないヌクレオソームの計算を 2 本、合計 4 本のシミュレーションを実行した。

構造アンサンブルを比較するため、主成分分析を行った。現在までの解析結果から、DNA の CG 配列はメチル化の有無にかかわらず、隣り合う二つの塩基対がねじれる (twist) 角の揺らぎが大きい、つまり、DNA のヘリックスが締まったり緩んだりする動きが大きいことが分かった (図 3)。これまでの研究から、DNA 単体での CG 配列は、TA 配列と並び特に揺らぎが大きく、多様な構造をとることが分かっていたが、その傾向はヌクレオソーム状態でも見られた。また、同じ計算条件のトラジェクトリから得られた構造を、第 1 と第 2 成分からなる分布の重なり具合で評価すると、メチル化なしの場合、3/4 の箇所がほぼ一致した分布になるのに対し、メチル化された場合、1/2 の箇所が異なる構造分布を示し、メチル化されることで CG 配列の構造多様性が増大する可能性が示唆された。H27 年度は、統計的な有意性を高めるためにさらなるシミュレーションを実行するとともに、メチル化の有無による DNA 構造や揺らぎの違いに関する詳細な解析を実行する予定である。

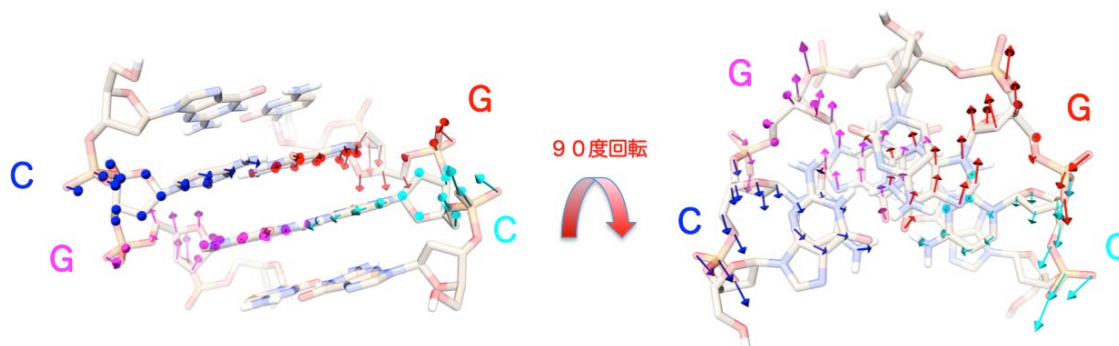


図 3. 主成分分析法によって決定された CG 配列重原子の固有ベクトル (図中の矢印)。塩基対を形成する C と G (青と赤、シアンとマゼンタ) がそれぞれ逆の方向 (左の図における上下方向) に動きやすく、2つの CG 塩基対 (青-赤、シアン-マゼンタ) が互いにねじれる (twist) 動きが生じることがわかる。本解析は 100 ns のシミュレーション 4 本 (メチル化なし 2 本、メチル化有り 2 本)、12 箇所の CG 配列の全ての構造を用いて行った。

## 3) ヒストンテールの構造サンプリング

ヒストンタンパク質の化学修飾は遺伝子発現制御や細胞分化などにおいて重要な役割を果たしている。とりわけヒストンの N 末端領域であるヒストンテールへの修飾が核内での DNA 収納状態をダイナミックに変えることが知られている。しかし、テールはいわゆる天然変性状態であることが知られており、実験的に詳細な構造情報を知ることは非常に難しい。そこで本研究では、修飾の有無によるテールやヌクレオソームの構造の違いを明らかにするために分子動力学計算 (MD) を行った。高い空間分解能を持つ MD シミュレーションは、テールの詳細な構造多形を

解析するのに有用な研究手段である。しかし、正に偏った電荷を持つテールは、負に偏った電荷を持つ DNA に強い静電相互作用によって張り付いてしまうため、従来の MD ではテールのとりうる様々な立体構造を十分に探索することができなかった。

我々は、構造的に比較的安定な DNA やヒストンコア領域の立体構造を維持しつつ、テールの様々な立体構造を効率よく探索することができる新たな MD 手法、adaptive lambda square dynamics (ALSD) 法を開発した(Ikebe et al., 2014, *J. Comp. Chem.*)。ALSD はテールの相互作用パラメータをスケールする変数、 $\lambda$  を新たに導入した MD である。 $\lambda = 1$  の時、ALSD は従来の MD と同様に DNA と接触したテールの立体構造を探索し、 $\lambda < 1$  の時は DNA-テール間の相互作用が弱められ、従来の MD よりも大きなテールの構造変化を促すことができる。ALSD ではシミュレーション中に  $\lambda$  がハミルトニアンに従って運動を行い、指定された  $\lambda$  値の範囲内を行き来することによって、テールは DNA から離れた状態と接触した状態を行き来し、様々なテールの立体構造を探索することが可能になる。

H26 年度は、ヒストンタンパク質が化学修飾された系 (H3 ヒストンテールの 14 番目のリジン残基がアセチル化された系 (K14ac)) について ALS D 計算を行い、これまで実現されてなかったアセチル化の有無での構造分布の比較を行った。アセチル化によってリジン残基の正電荷が中和されると、負電荷を持つ DNA との間の引力相互作用が弱まるため、テールは DNA から剥がれ外に突き出すようになると従来は考えられていた。しかし、本計算結果から、リジン残基のアセチル化だけではテールが DNA から剥がれるようなダイナミックな状態変化はほとんど起こらず、依然としてテールは DNA 周辺に存在することが分かった (図 4 (a), (b))。加えて K14ac 系では、アセチル化無し (WT) の系と比べてヌクレオソームに巻きついている DNA が剥がれやすくなり (図 4 (c))、テールはヘリックス構造を採る傾向がより強まることが分かった。H27 年度では本計算結果のさらなる解析を進め、修飾の有無によるテールやヌクレオソームの構造状態の変化に関する詳細を明らかにする。また、テールに複数のアセチル化が修飾された系について ALS D 計算を行い、アセチル化の数がその構造に与える影響について研究を行う予定である。

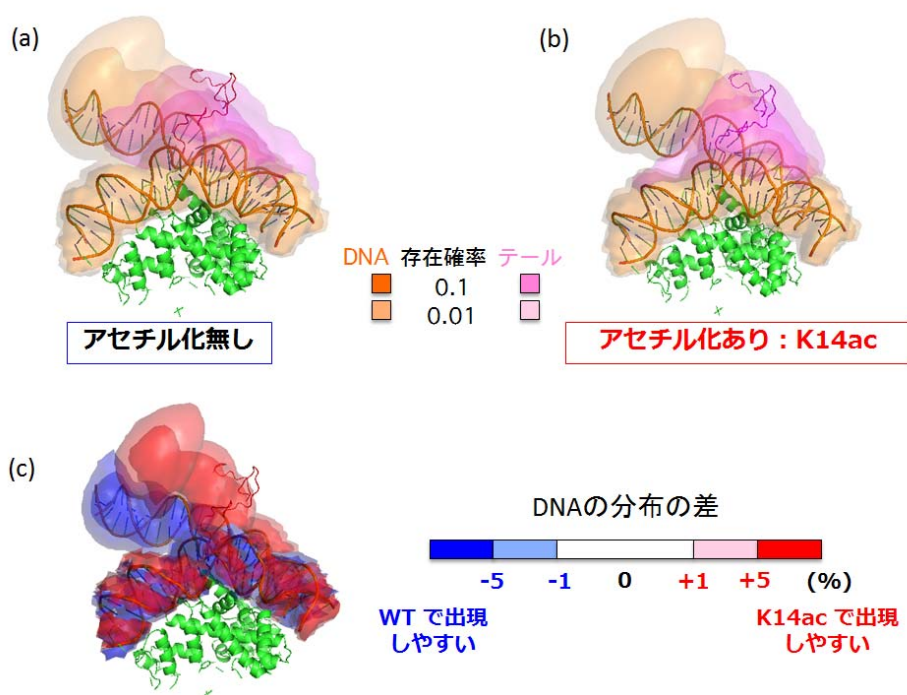


図 4. テールと DNA の空間分布。(a) アセチル化無し (b) アセチル化あり (c) アセチル化の有無による DNA の空間分布の差