

○ 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タを抽出して、粗視化モデルを構築する。

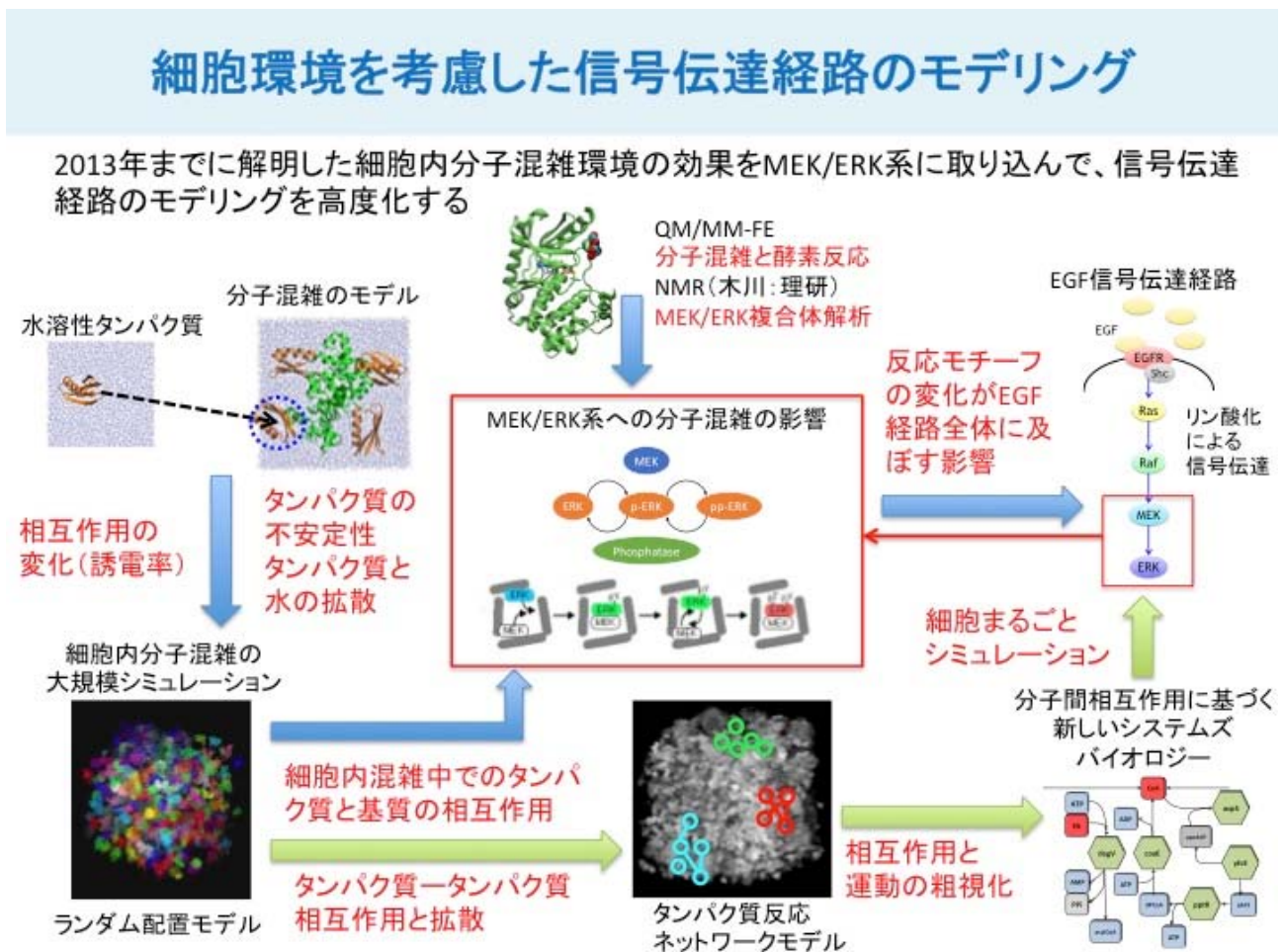


図 1. 研究計画全体を示す概念図

平成 25 年度は、(1) バクテリア細胞の細胞質環境を再現するシミュレーションを行い、拡散や粘性などの流体力学効果を導出する、(2) 信号伝達経路上のリン酸化酵素の分子運動を拡張アンサンブル法などで解析するとともに、ERK-MEK 複合体のモデリングを行なう、(3) 粗視化モデルを構築する手法を開発するとともに、時間ステップを延ばす工夫を行い、大規模粗視化シミュレーションの実現に備える。

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、1000 万以上の原子を含む巨大なシミュレーション系を構築し、「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。シミュレーションは、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。2000 ノードから 8000 ノード程度を用いた計算を実施したが、「京」のジョブの混み具合を検討して、比較的少ないノードで計算を実行せざるを得なかった。1000 万原子系については、100ns のプロダクションを実施し、1 億原子系については 5ns 程度のプロダクションが終了した。1 億原子系については平成 26 年度も引き続き計算を実施し、10ns 程度のプロダクションを完了する予定である。

得られたトラジェクトリから、タンパク質や RNA の並進拡散、回転拡散に関するパラメタを記述した。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 年度に実施する予定である。タンパク質と代謝物の相互作用についても検討を開始した。特に、ATP が多く含まれているため、ATP を基質とする ACKA というタンパク質との相互作用を調べたところ、混雑環境でも基質結合部位を認識できることを示す事ができた。代謝物の多くは、非特異的な相互作用によって生体高分子表面にトラップされており、その運動を解析中である。

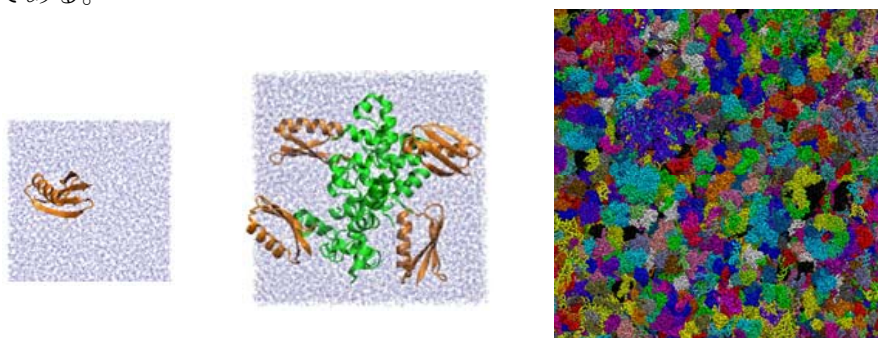


図 1. 細胞環境を考慮したシミュレーションの模式図。左から、希薄溶液中のタンパク質 (Protein G)、分子混雑を模したシステム (Protein G と Villin headpiece subdomain)、バクテリア細胞質を模した 1000 万原子系 (「京」で計算を実施中)

2) バクテリア細胞質中ではなく、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。

(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いる。この原因を分子モデルで説明することがこの課題の目標ではあるが、問題点として、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があげられる。そこで、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ぼった。その結果 2 つの構造の候補が得られてきた。この結果を、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、良い一致を得た。現在、これらの複合体の候補構造を立体構造決定に役立たせることができるかどうかを検討しているところである。

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25年度は、微視格子反応拡散コード pSpatiocyte の京コンピュータにおける最適化、24年度に構築した EGF 経路モデルの改良および pSpatiocyte によるテストランを行なう。具体的には、(1) 24年度にハイブリッド並列化した pSpatiocyte をさらにチューニングし、スケーリング性能を向上させる。(2) 細胞の核や膜のような構造を再現可能にするコンパートメントを pSpatiocyte に実装する。(3) 京コンピュータにて、pSpatiocyte を利用し、細胞環境（分子混雑）を含まないケースでの上皮成長因子（EGF）シグナル伝達経路モデルのテストランを実施する。

I-2-2 実施内容（成果）

(1) pSpatiocyte の開発および高度化

1) 京コンピュータ向けのチューニングとスケーリング結果

25年度上期において、拡散過程の並列化によって生じた計算結果の不一致を解消するため通信手順の改良を実施した。これによって分子数が保存しない不具合が解消された。また、「京」に特有の実行環境に適合させるために OpenMP を用いて主要ループを分割し、拡散過程にハイブリッド並列化を施し、化学反応過程には直接法による並列化を実装した。25年下期では上期の通信手順の変更によって生じた並列性能の低下を解消するため、通信の前処理と後処理のコーディングにおける Vector コンテナの処理方法を改善した。この処置によって数百プロセス程度で飽和していた並列スケーリングを数万プロセスまで向上させることができた。またプログラムの物理的な検証として、拡散および反応に関する理論値や既存の逐次プログラムの結果との比較を行い、妥当性を確認した。さらに実計算へ向けた取り組みとして、代表的なシグナル伝達経路である MAPK の簡略化モデルを対象に、実時間で30秒に相当する計算を実施した。猶、この計算には「京」の4096コアを用いて約16時間を要した。

2) コンパートメントの実装

25年度は、細胞内に存在するオルガネラや細胞膜・核膜などをモデル化するためのコンパートメントの実装を進めた。具体的には膜状の拡散速度や反応速度の合理的な設定に関し検討を行い、3次元過程を直接2次元に制限することにした。当初は膜上の分子数が少ないため、原理的には通信ネックによる性能低下の可能性も懸念されるが、物理的には膜上分子の拡散速度が細胞質内に比べて著しく小さいので、ほぼ問題無いものと予想される。現在開発は遅れ気味であるが、これによってコード開発の作業量を抑えることが可能となり、早期開発に目途を付けた。膜形状は初めに単純な立方体として作成する機能を用意し、十分な動作確認を行った後に任意形状を読み込む機能を付加する方向で進めている。

(2) EGF シグナル伝達経路のモデリング・シミュレーション状況

1) 非分子混雑下での EGF シグナル伝達経路モデルのシミュレーション

EGF 受容体 (EGFR) は、細胞膜に存在しており、細胞外の EGF と結合し、二量体化、リン酸化を経て、複数のアダプタータンパク質 (Grb2, Sos, Shc 等) と結合する。その結果、活性型の EGFR となり、細胞膜上の Ras タンパク質を活性化させる (RasGTP)。RasGTP は、細胞質中の Raf タンパク質を細胞膜へとリクルートし、細胞膜上で活性化させる。活性型の Raf タンパク質は、細胞質中の MEK タンパク質を二重リン酸化し (MEKPP)、さらに、MEKPP は、ERK を二重リン酸化する (ERKPP)。そして、ERKPP は細胞核内へと移行し、細胞の増殖・分化等に関与する転写因子を制御する。以上の一連の流れを含む数理モデルが、これまでに構築されており (図 1)、その数理モデルは、73 種類のタンパク質、144 本の反応、85 個の反応パラメータから構成される。本年度は、細胞間での各タンパク質分子の濃度ばらつきに着目し、その現象を数理モデルへと反映させ、シミュレーションを実施した。

細胞に含まれる各タンパク質濃度は、細胞ごとにばらついていることが知られており、定量的な計測から、その分布は対数正規分布に従うと報告されている。バラツキの程度は、変動係数 (標準偏差を平均値で割った値) で表現され、タンパク質の多くは、25%前後の値を示す。一方、タンパク質濃度に関しては、Ras, Raf, MEK, ERK の 1 細胞あたりの濃度が測定されているため、その値を数理モデルに利用した。その他のタンパク質に関しては濃度が測定されていないため、典型的なタンパク質濃度 (1 μ M 前後) を利用した。細胞間のタンパク質濃度のバラツキを再現するため、数理モデルに含まれるタンパク質の初期濃度を変動係数 25%の対数正規分布に従ってランダムに変更し、シミュレーションを実施した。pSpatioocyte のコンパートメントシステムの実装が完了していないため、前年度と同じく E-Cell3 Spatioocyte を利用して、この EGF モデルのシミュレーションを実施した。Spatioocyte は、シングルコアのみでの動作であるため、実際の細胞の大きさ (1pL) の 1/10 程度まで縮小し実行した。

図 2 に、EGF 0.05 ng/mL における、核内 ERK の実験データ (PC12 細胞) と初期値バラツキを考慮しない場合と考慮した場合のシミュレーション結果を示した。初期値のバラツキを考慮したシミュレーション (図 2C) の方が、バラツキを考慮しない場合 (図 2B) よりも、実験データ (図 2A) をよく再現出来ていることが分かる。シミュレーションデータのサンプル数が少ないため、定量的な比較は出来ないものの、タンパク質のバラツキが細胞の

応答のバラツキに大きく影響していることが示唆された。また、当研究室で開発している蛍光顕微鏡シミュレータにより、シミュレーション結果を可視化した結果を図 3 に示した。実際の蛍光顕微鏡により観測された画像 (図 3A) とシミュレーション結果 (図 3B) が非常に良く一致していることが分かる。

今後の予定としては、pSpatioocyte へのコンパートメントシステムの実装が完了次第、EGF モデルのシミュレーションを京コンピュータで実施する。さらに、細胞環境 (分子混雑) を含めたシミュレーションも実施する予定としている。

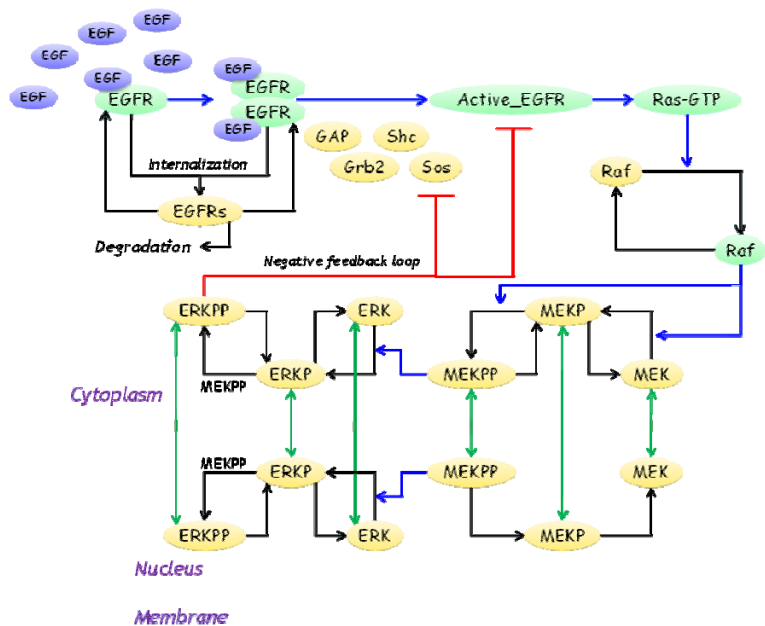


図 1 EGF シグナル伝達経路モデルの概略。細胞膜上の EGF 受容体と EGF リガンドの結合から、ERK タンパク質の核内移行までを含んでいる。

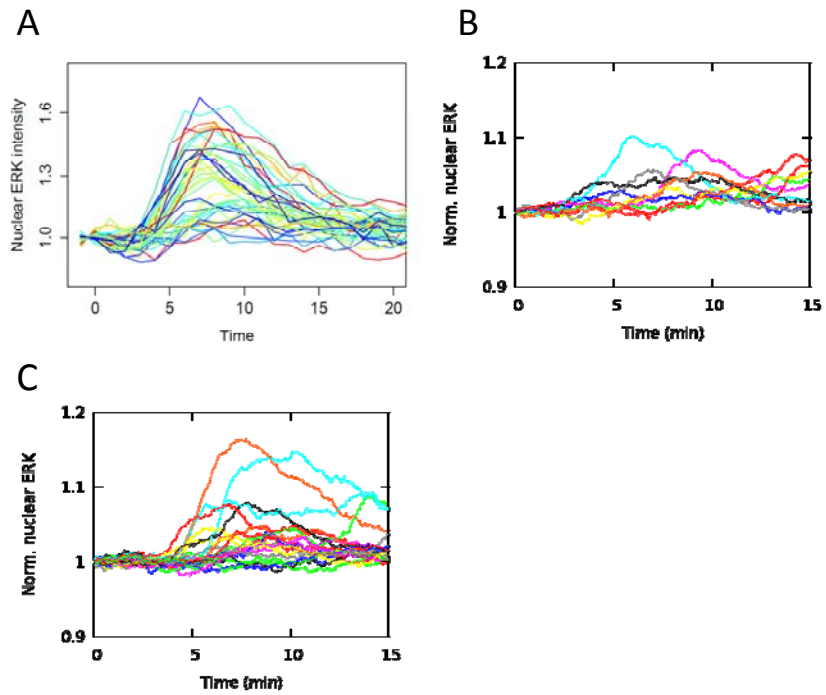


図2 核内 ERK の実験データとシミュレーション結果の比較。A) EGF 濃度 0.05 ng/mL における核内 ERK のタイムコースデータ。PC12 細胞を用いて測定した。B) 初期値のバラツキを考慮しない場合、C) バラツキを考慮した場合、それぞれの核内 ERK のシミュレーション結果。

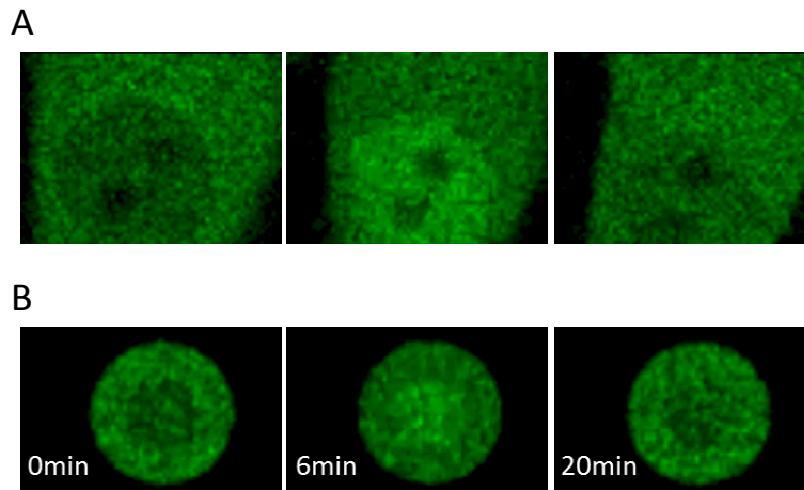


図3 ERK タンパク質分布の実験およびシミュレーションの比較。緑は、ERK タンパク質を示す。A) 蛍光顕微鏡により観察した EGF 添加後の ERK タンパク質の分布の変化。細胞核 (画像中央部) に一過的に ERK タンパク質が蓄積する。B) 蛍光顕微鏡シミュレータにより可視化したシミュレーション結果。実験データと同じく、画像中央部の細胞核部分に ERK タンパク質が一過的に蓄積する。

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 25 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。特に、リン酸化を受けている状態と受けていない状態についての自由エネルギー極小構造を探索することにより、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。またリン酸化活性を制御する薬剤結合について、ハイブリッド QM/MM 自由エネルギー法を用いた新規な解析手法の開発に着手する。

I-3-2 実施内容 (成果)

本年度は、MEK-ERK 複合体のモデルとして、信号伝達経路上の上流にある GAP-Ras 複合体の GTP 加水分解反応に関する研究を行った。まず、複合体構造に対して、反応の始状態、遷移状態、中間状態、及び終状態の 4 状態に対して、QM/MM 自由エネルギー法を用いた構造最適化を完了した。シミュレーション系は、MM 領域は 77,974 原子、QM 領域は 76 原子で構成され、周期境界条件のもと Ewald 法 (MD は PME 法) を用いてクーロン長距離力を記述している (図 1a)。QM 領域は B3LYP/6-31+G(d,p)、MM 領域は Amber ff03 force field を用いている。それぞれの構造最適化に要した時間は、それぞれ、始状態に 877 ns、遷移状態に 457 ns、中間状態に 673 ns、及び終状態に 525 ns である。また、それぞれの状態の最適化構造から、600 ns の分子動力学 (MD) 計算を行い、平衡分布を解析した。

その結果、反応遷移状態や終状態の生成に伴い、Ras と GAP の相対的な位置が大きく移動することが明らかになった (図 1b)。反応遷移状態生成における複合体の構造変化は、Ras に結合している反応基質の遷移状態構造を安定化する近傍の switch-I ループの構造変化が、ループに含まれる Pro34 のシフトを引き起こし、それが GAP の Leu902 と相互作用を通して GAP の相対位置の変化をもたらしていることが示唆された (図 1c)。また、反応始状態では、基質近傍のキャビティーに、攻撃水以外にも 3 個程度の水分子が結合しているが、遷移状態生成に伴う構造変化により、これらの水分子は排除されることが明らかになった。また、遷移状態や終状態の生成に伴い、Ras と GAP 間の塩橋の組み換えが観測された、例えば、終状態の生成に伴い、Ras の Glu37 と GAP の Arg928 の間に塩橋の生成が観測された (図 1d 及び 1e)。これらの塩橋の組み換えにより、構造変化が促進されていることが示唆される。このように、MEK-ERK 複合体にもみられるような、三リン酸からのリン酸移動を伴う信号伝達複合体の酵素活性には、大きなタンパク質の構造変化が関与することが明らかになった。細胞環境ではそのような大きな構造変

化が変調を受けることが予想されるため、酵素活性も変化する可能性が示唆された。

次に、QM/MM 自由エネルギー法の大規模系の拡張を目指し、およそ上記の 4 倍の原子数 (34 万原子) をもつ F_1 -ATPase の系の ATP 加水分解反応に対する適用を行った (図 2)。その結果、反応の始状態と遷移状態直前の中間状態に対して、自由エネルギー構造最適化に成功した。構造化に要した時間は、それぞれ 774 ns と 330 ns である。反応の遷移状態近傍における、熱ゆらぎの効果を含む自由エネルギー最適化構造は、それを無視したポテンシャルエネルギー最適化構造に比べて、リン酸の解離距離の増加や、触媒活性を与えるアルギニンフィンガーのより強い配位などが観察され、熱ゆらぎの効果により、反応活性化エネルギーが下がることが示唆された。

また、MEK 系に関しては、QM/MM 自由エネルギー計算に向けて、シミュレーション系の構築と、MD 計算による平衡化を現在行っている。

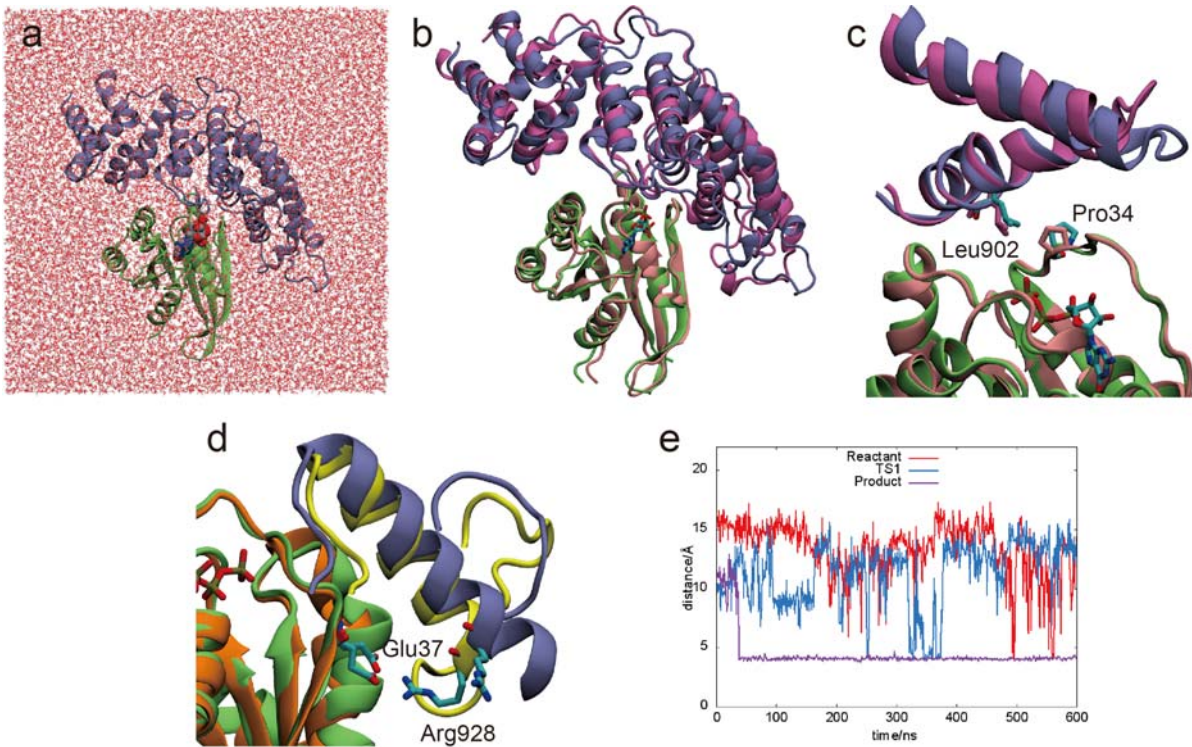


図 1: a Ras-GAP 複合体の QM/MM シミュレーション系。Ras (緑) と GAP (青) の複合体が周期境界条件の水環境の中に配置されている。GTP 基質分子はファンデルワールス表示で示されている。b 反応遷移状態生成に伴う構造変化。反応始状態 (Ras : 緑、GAP : 青) と終状態 (Ras : ピンク、GAP : マゼンタ) の Ras 部分が RMSD フィットされて示されている。反応遷移状態において、GAP 部分が相対的に移動していることがわかる。c Switch-I ループの Pro34 と GAP の Leu902 の相互作用と構造変化。d 反応終状態生成に伴う Ras-GAP 間の塩橋の生成。反応始状態 (Ras : 緑、GAP : 青) では、Ras の Glu37 と GAP の Arg928 は塩橋を作っていないが、終状態 ((Ras : オレンジ、GAP : 黄) では塩橋が生成している。e Glu37 と Arg928 の距離の変化。

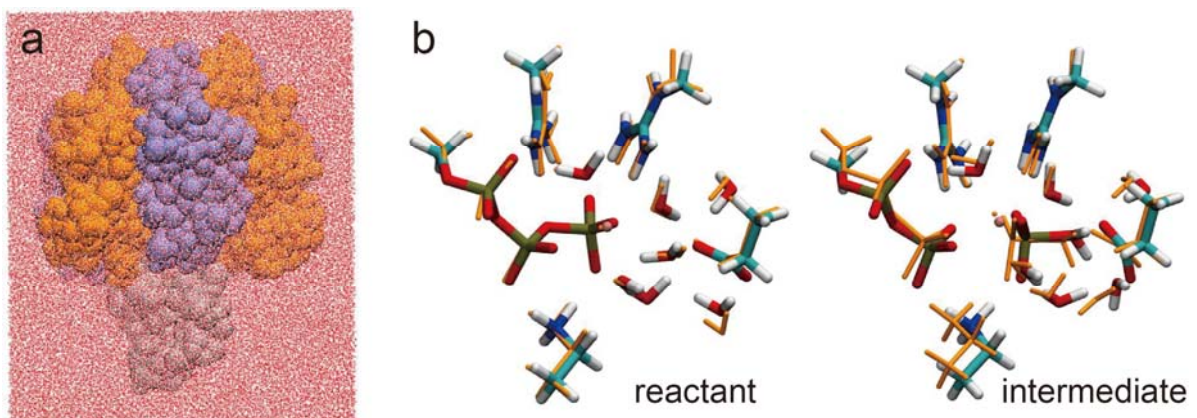


図 2: a F_1 -ATPase のシミュレーション系。b QM 領域の自由エネルギー最適化構造。オレンジは 0 K でのポテンシャルエネルギー最適化構造。

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25年度は、粗視化シミュレータ CafeMol を用いて、信号伝達における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、ヌクレオソーム多量体からなるクロマチンが核内混雑環境でどのような形態で超高密に折りたたまれているのかを解析する。

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) 細胞外から与えられた様々な刺激は、信号伝達経路を通して核内に伝達(増幅)され、最終的には蛋白の活性や遺伝子発現を変化させる。信号伝達経路の代表的モデルの一つとして、MAPK 系のリン酸化カスケードがある。この MAPK カスケードにおいては、上流のもの(MAPKK)は下流のもの(MAPK)をリン酸化する事で活性化する。しかし、一体どの様に MAPKK が MAPK をリン酸化するのか、その分子スケールにおける動的機構の詳細については良く分かっていない。その解明を阻む主たる原因は、MAPKK-MAPK 複合体の詳細構造が実験的に未だに明らかにされていない事にある。特に哺乳類における MAPK カスケードは、細胞の成長、分化、アポトーシスだけでなく、癌化等の様々な細胞運命に強い関連がある事が知られている。その為、リン酸化の動的機構(分子機構)について明らかにする事は医学的にも非常に重要である。

哺乳類の MAPK 系リン酸化カスケードにおけるリン酸化の(動的)分子機構を構造ベースの粗視化シミュレーションを適用する事で調査する。その際、主に MEK1 (MAP-kinase-kinase:MAPKK) がどの様に ERK2 (MAP-kinase:MAPK) をリン酸化するのか、その動的機構に焦点を当てる。

シミュレーションは、我々のグループで開発している生体分子粗視化シミュレータ CafeMol を用いて行った。その際、MEK (MAPKK)、ERK (MAPK) の分子内相互作用については、構造依拠型の AICG2 モデル(Li, Takada et al., 2012)を適用した。この AICG2 は、従来型の Go-like ポテンシャルよりも原子的な相互作用(効果)を考慮(反映)したポテンシャルになっている。PDB で構造が取られていないギャップ領域については、フレキシブルローカルポテンシャル[統計ポテンシャル](Terakawa and Takada, 2011)を適用した。一方、分子間相互作用については排除体積効果と Debye-Huckel 型[イオン濃度:0.15M]の静電相互作用を適用した。MD の時間発展については、Underdamped の Langevin ダイナミクス(温度:300K)を適用した。

実験(Garai et al., 2012)により、MEK (MAPKK) の N 末(D-domain) と ERK (MAPK) の CD-domain はある程度の親和性(平衡解離定数 $k_d=18.7 \mu\text{M}$) で相互作用している事が知られている。そこで、まず D-domain と CD-domain が Go-コンタクトにより非常に強く相互作用している条件下で MEK が ERK の周りをどの様に動くかそのダイナミクスを調査した。図 1 (上) は、ERK2 固定座標上(極座標)でみた MEK1 の重心の位置分布(確率密度分布)である。ここで、極座標(Θ 、 Φ)は図 1(下)の様においた: ERK2 の長軸方向を Z 軸方向にあわせ、リン酸化サイトを x 軸方向に合わせた。注目ベクトルの z 軸からの傾きを Θ とし、xy 平面に射影されたベクトル成分の x 軸からの傾きを Φ とした。図 1 上の分布図において、黄色の円で囲まれた領域が、ERK2 のリン酸化サイト(T185/Y187)の位置に相当する。本来、MEK1 が ERK2 をリン酸化する際には、MEK1 が結合している ATP(γ -ATP)が ERK2 のリン酸化サイト(T185/Y187)に十分近付く必要がある。しかしながら、D-domain と CD-domain の間の相互作用(アンカー相互作用)が存在する場合には、MEK は ERK のリン酸化サイト(T185/Y187)に殆んど近付く事ができない事が分かった(図 1 上: 確率密度分布より)。これは、アンカー相互作用が存在する場合、MEK1 の N 末のフレキシブルなループ領域が ERK2 に上手く巻き付かなくては、MEK1 の ATP は ERK2 のリン酸化サイト(T185/Y187)に到達できない事が原因となっていると考えられる。

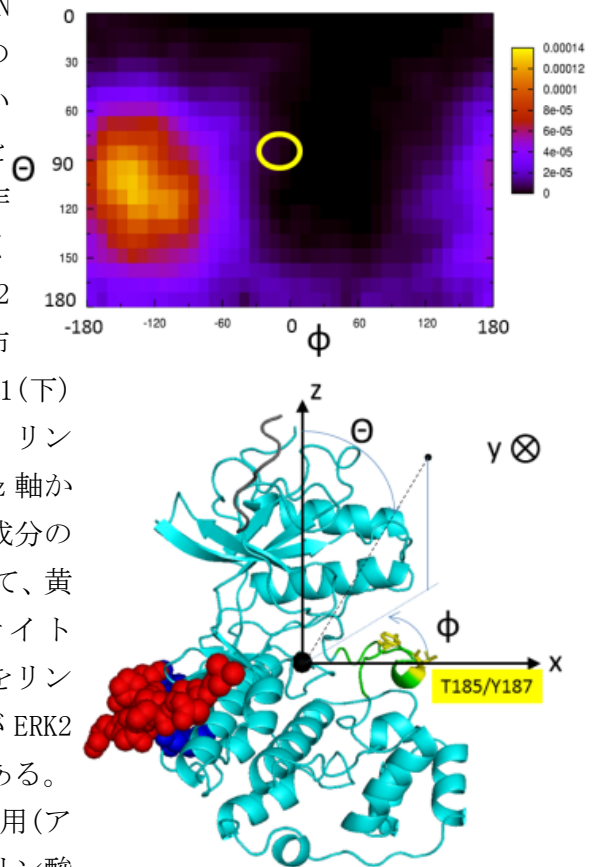


図 1 上図は ERK2 固定座標上(極座標)でみた MEK1 の重心の位置分布(確率密度分布)。下図は極座標系(Θ 、 Φ)を表す。

そこで、次に D-domain と CD-domain の間のアンカー相互作用が存在しない場合に同様シミュレーションを行った。但し、MEK と ERK が結合する前に互いの重心が拡散によって無限遠に離れてしまう事を防ぐ為に、MEK の重心と ERK の重心の間に距離拘束の調和振動子型のポテンシャル(自然長:70 Å、バネ定数:20.0kcal/mol)を印加した。その結果、アンカー相互作用が存在する場合(図 1 上)とは全く異なり、MEK1 の重心はある程度の高い確率で ERK2 のリン酸化サイト(T185/Y187)に近付く事が出来ると分かった。

以上の結果は、D-domain(MEK) と CD-domain(ERK)のアンカー相互作用が強過ぎず、ある程度の強度で存在する事が、MEK による ERK のリン酸化過程にとって重要であることを示唆している。我々の研究グループでは、酵母の MAPKK-MAPK 系のドッキングダイナミクスも調査しているが、定性的に同様の結果が得られている。

信号伝達経路(MAPK カスケード)における MEK (MAPKK) と ERK (MAPK) のドッキングダイナミクスを CafeMol による粗視化シミュレーションにより調査した。その結果、MEK の D-domain と ERK の CD-domain の間の相互作用の強度(有無)によって、ERK のリン酸化領域(T185/Y187)へのアプローチ過程(ダイナミクス)が定性的に大きく異なる事が分かった。

今後は、MEK の ATP(γ リン酸)と ERK のリン酸化サイト(T185/Y187)に Push-Pull-Release

(Ravikumar et al., 2012)のプロトコルに従い距離拘束のポテンシャルを適用する事で、妥当性のある MEK-ERK の複合体構造を効率的にサンプリングする。そうして得られた複合体構造から、リン酸化過程で重要となる MEK と ERK の間のコンタクト(相互作用)の情報の抽出を目指す。更に、細胞内環境を考慮に入れた MEK-ERK の複合体形成過程についても調査する。

(2) クロマチンの動的モデリング

1) ヌクレオソーム 3 量体および 20 量体の初期構造モデルを作成し、安定構造を探索するシミュレーションを行った。20 量体では、さまざまな混雑環境を設定して計算を行い、得られた構造の特徴を解析した。旧来言われてきたジグザグ型やソレノイド型に似た特徴は得られなかった。

さらに、20 量体がさまざまなヌクレオソーム密度に存在する状況における、転写因子の動態を、粗視化分子シミュレーションにより解析した。p53 (図 2) は、0.1 mM 程度のヌクレオソーム濃度下にいると、遅いもののある程度拡散が可能であるのに対し、0.5mM ヌクレオソーム下では、ゆらぎは見られるものの、拡散は大きく制限された。p53 より小型の転写因子では、より障害を受けない拡散が見られた。

2) 粗視化分子モデルにおける GBSA モデルの開発

一般に生体分子システムの分子内および分子間相互作用のなかで、静電相互作用と疎水性相互作用が主要なものであり、その有効な記述は重要である。我々グループが開発を進めている一般的粗視化分子シミュレーションソフトウェア CafeMol においては、現在までに Debye-Hückel 型の静電相互作用関数、経験的な粒子の埋もれ度によって表現された簡潔な疎水性相互作用関数が実装されているが、より高精度かつ系統だった相互作用関数の開発実装が望ましい。とくに、クロマチンヌクレオソーム系では、静電相互作用が主要な役割をもち、より高精度なモデルの必要性は高い。そこで、本研究開発では、より精度の高い静電相互作用と疎水性相互作用を表現するために、それぞれ、原子レベルの陰溶媒モデルとして著名な一般化ボルン(GB)と溶媒接触表面積(SA)モデルを、マルチスケールモデリングによって粗視化モデルに拡張し、それを CafeMol に実装した。

原子モデルで長年にわたり開発されてきた陰溶媒静電相互作用の GB 法には、近似の方法に数多くの表式があるが、そのなかで分解能を変えるときに同様に使えるものを検討した結果、Dominy & Brooks の公式 (Dominy & Brooks, J. Phys. Chem. B. 1999) を用いることとし、粗視化モデルに対してパラメータを最適化した。一群の生体分子に対して、原子モデルを用いて Poisson-Boltzmann (PB)方程式を解くことによって計算されるボルン半径を、よりよく再現できるように粗視化モデルにおける一般化ボルンのフィッティングパラメータを最適化した。一群の生

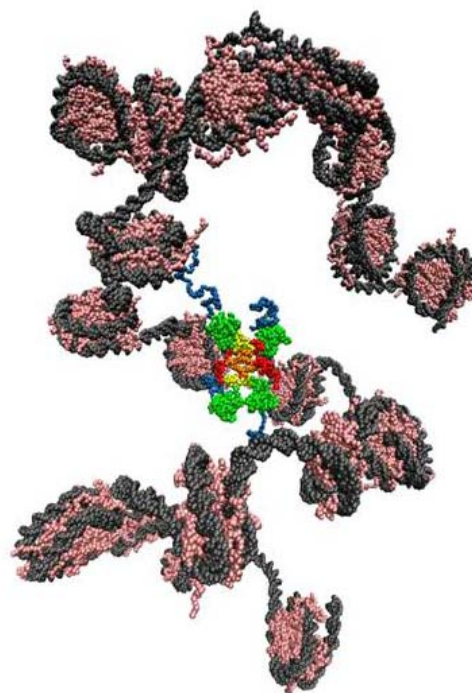


図 2. ヌクレオソーム 20 量体が、約 0.1 mM のヌクレオソーム密度にあるときの転写因子 p53 の動態を解析した。図は p53 をヌクレオソーム 20 量体の中心においたもの。

体分子セットとしては、21個のタンパク質と、20個の DNA、および14個の RNA を含めた。このパラメータを用いて、別の一群のテスト分子群に対して、粗視化 GB 法の溶媒和自由エネルギーを、原子モデルの Poisson-Boltzmann (PB) 方程式の解から求められる溶媒和自由エネルギーと比較した (図3)。図の通り、両者はきわめてよく相関し、粗視化モデルで溶媒和自由エネルギーをある程度精度よく再現することができることを示した。

陰溶媒疎水性相互作用の溶媒接触表面積

(SA) モデルについては、すでに $C\alpha$ のみの粗視化モデルによる近似表現が提案されている (the Parameter OPTimised Surfaces (POPS) 法、Fraternali & Cavallo, Nuc. Acid. Res. 2002)。我々はこの POPS モデルを適用することとし、それを CafeMol に実装した。

今後、GB 法と SA 法と、その他の部分の重みを決定する必要がある。マルチスケール法によってその重み因子を決定し、GBSA 法をヌクレオソーム・クロマチン系に適用し、従来よりも精度の高い粗視化モデルによるシミュレーション研究を進める。

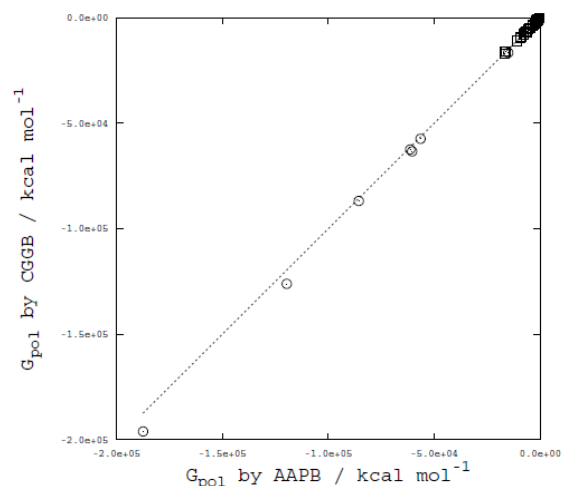


図 3. 一群のテスト分子群について、全原子モデルの PB 方程式から求められる溶媒和自由エネルギー (横軸) と、粗視化モデルの GB モデルから得られる溶媒和自由エネルギーの相関を表す。

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年度は、クロマチン関連因子やDNA修復関連タンパク質等の核内DNA結合タンパク質について、実験系と連携しつつ、その機能ダイナミクスを解明する。分子シミュレーションは計算機内のモデルを扱う研究であるため、実際の実験との連携・検証が欠かせない。原子レベルの立体構造情報を提供するX線結晶解析だけでなく、溶液状態の構造情報を提供するX線溶液散乱やNMRなど様々な実験情報と、シミュレーション結果を照らし合わせて検証することが必要である。しかし、実験情報とシミュレーション結果の比較は必ずしも簡単でなく、実験状況をシミュレーションで再現するなど、連携法を確立することが重要である。そこで、本年度は溶液系の実験情報とシミュレーションの連携法の検討を行い、クロマチン関連因子やDNA修復関連タンパク質等の核内DNA結合タンパク質に適用して、そ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

我々が開発している分子シミュレーションシステム MARBLE について、「京」へのチューニングを進めた。ウィークスケールング改良のために、計算対象の系のサイズを変えたデータに対しベンチマーク計算を行い、比較的大きな系で計算効率が悪いところを特定し、改良を試みた。また、以前、「京」用に最適化していた短距離力計算のカーネルコードを導入することで、SIMD化等の効率をアップさせた。長距離力についても、最適化を進めた。

分子シミュレーションの計算結果を、実験結果と比較検討するために、溶液状態の構造情報を提供するX線小角散乱(SAXS)法と分子動力学シミュレーション(MD)法を連携させるMD-SAXS法の開発を行った。既に、全原子MDをベースにしたall-atom MD-SAXS法の基本的な部分は完成していたが、開発者でないと使いづらいソフトウェアだったため、そのユーザー・インターフェースやスクリプトを開発することで、一般のユーザーにも使うことができるような環境を整えた。

しかし、ヌクレオソームのような巨大系では、原子をすべて露わに扱う全原子MDでは、構造サンプリングがどうしても不十分になってしまうため、粗視化MDの結果も統合的に扱うマルチスケールシミュレーションが必要になってくる。そこで、粗視化MDも、MD-SAXSに対応させるべく、粗視化MD-SAXS法の開発も進めた。

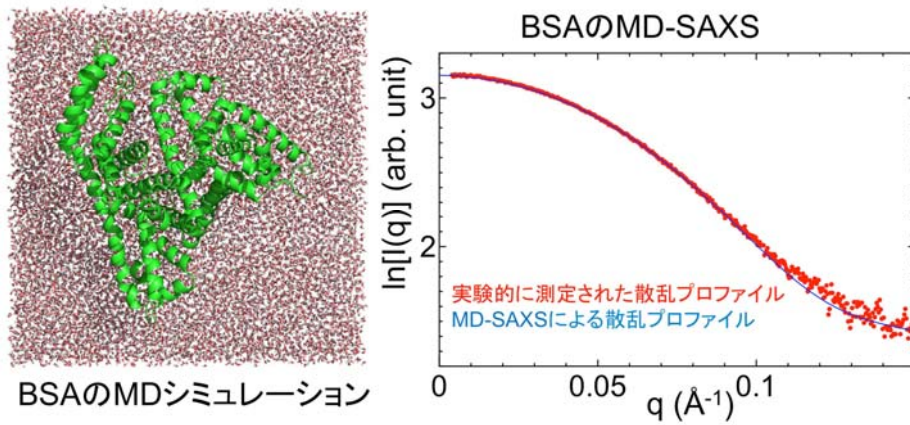


図. 全原子 MD-SAXS 法によるシミュレーション計算と実験結果の比較。

BSA の実験データは、横浜市立大学の佐藤衛研究室より提供された。

(2) 研究開発の実施状況

1) モノヌクレオソームの全原子 MD-SAXS 研究

今年度は、核内DNA結合タンパク質のうち、モノヌクレオソームに着目して、全原子 MD-SAXS 解析を行った。ヌクレオソームは、河野グループと高田グループと連携して、全原子モデルから粗視化モデルに至るマルチスケールシミュレーションを用いて、動的構造の研究を実施している対象である。ヌクレオソームにおいて、実験結果とコンシステントなシミュレーション条件を見出しておくことは、今後、大規模にヌクレオソームのシミュレーション研究を実施する上で、重要なことである。そこで、溶液状態のヌクレオソームの概形情報がわかる X 線小角散乱の実験データと MD の計算結果を比較し、溶液状態でヌクレオソームはどのような構造になっているかを調べた。その結果、ヌクレオソームの DNA 末端は、ヌクレオソームに巻き付いた構造よりも、多少離れた構造のほうが、SAXS 実験データに合うことが分かった。

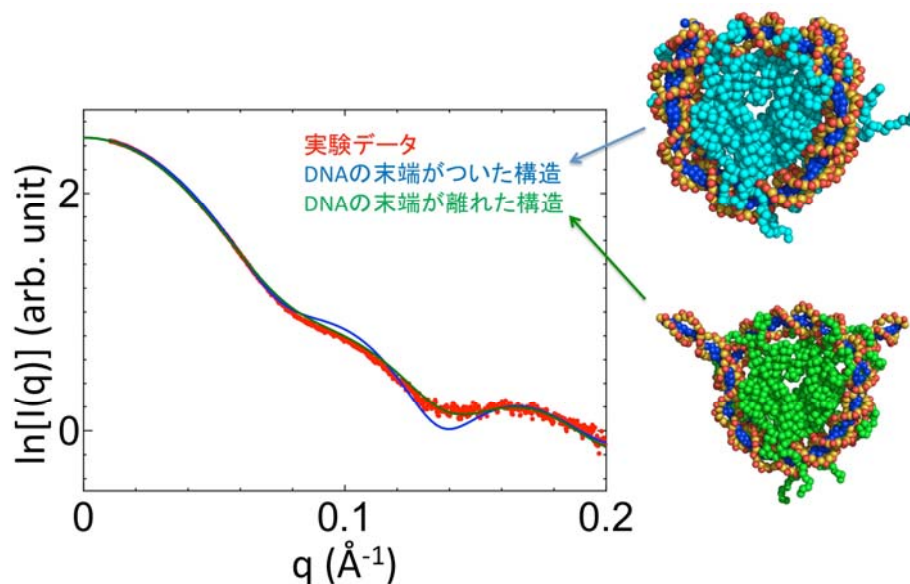


図. 全原子 MD-SAXS 法によるヌクレオソームの構造と X 線小角散乱のデータの比較。

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 石谷 隆一郎 (東京大学)

遺伝子の転写調節に関わる因子の分子動力学シミュレーションによる制御機構の解明研究

I-6-1 実施計画

本研究では、最終的には予測データを実験面からも検証することを視野に入れるが、研究課題自体は純粋に計算機シミュレーションを使用した理論研究のみを実施するものとする。平成25年度は、転写調節に関わる因子の中から、特に HHM を取り上げ、MD シミュレーションを行う。HHM は、Class II 型の塩基性 Helix-Loop-Helix (bHLH) 型転写因子に結合することで、Class II bHLH 型転写因子が DNA に結合することを阻害し、転写の抑制に関わる因子である。そして、最終的には TGF β による情報伝達を阻害するように働き、細胞のコンテキストに依存して細胞増殖あるいは停止へと細胞の状態を変化させる役割を持つことが知られている。本研究では、当研究室で得られている HHM の結晶構造をもとに、HHM 単独、あるいは Olig1 などの bHLH 型転写因子結合状態などの MD シミュレーションを実施し、動的な転写調節機構の解明を目指す。

I-6-2 実施内容 (成果)

(1)

本課題では、「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、遺伝子の転写調節因子による制御機構の解明に向けた研究開発を実施した。細胞の状態や周囲の環境に応じた遺伝子発現の調節は、様々な転写調節因子によって担われている。この調節の過程では、蛋白質間の相互作用と動的な構造変化が起こっていると考えられている。そのため、転写調節機構の解明には、原子レベルのスタティックな立体構造解明だけでなく、そのダイナミクスの解明が必要不可欠である。本課題では、転写調節因子あるいは調節因子とターゲット複合体の分子構造に分子動力学(MD)シミュレーションを適用し、転写調節の動的機構解明を目指した。

まず、転写調節に関わる因子の中から、特に HHM を取り上げ、MD シミュレーションを行った。HHM は、Class II 型の塩基性 Helix-Loop-Helix (bHLH) 型転写因子に結合することで、Class II bHLH 型転写因子が DNA に結合することを阻害し、転写の抑制に関わる因子である。そして、最終的には TGF β による情報伝達を阻害するように働き、細胞のコンテキストに依存して細胞増殖あるいは停止へと細胞の状態を変化させる役割を持つことが知られている。まず、当研究室で得られている HHM の結晶構造からは HHM は L 字型をした構造を取っているが、溶液中では自発的に中心部分の dnHLH 領域が解離して、線形の構造を取ることが示唆されていた。この変化の詳細を調べるために、dnHLH 領域が解離した L 字型構造から開始する MD シミュレーションを行ったところ、線形にはならず、L が折りたたまったような形状に変化することがわかっていった。本試験研究ではさらに長時間 (100 ns 程度)、あるいは多数のシミュレーションを行うことで線形に変化するかどうかを検討したが、実験結果に一致するような結果は得られなかった。今後さらに計算時間を延ばすあるいは試行回数を増やすなどして検討を行う必要が有る。

さらに、炎症応答や免疫制御、細胞の生存、がん細胞の接着・浸潤などに関与する多くの遺伝子発現を調節する転写調節因子 NF- κ B を制御する A20 の解析を行った。A20 は直鎖状ユビキチン鎖に特異的に結合することで、NF- κ B のシグナルを減衰させ細胞内シグナルの制御にかかわることがわかっている。我々のグループでは、A20 と直鎖ジユビキチン鎖複合体の結晶構造を解明したが、それを初期構造とし、直鎖

型, **K63** 結合型の 2 通りに対して MD シミュレーションを行った結果, **A20** は直鎖型に対してより安定に結合することが検証できた. さらにトラジェクトリーを詳細に解析することで, 直鎖型ユビキチンの場合, 主に水素結合からなる **A20** と近位ユビキチンとの相互作用が安定に保たれるのに対し, **K63** 結合型ユビキチンの場合, 2 つのユビキチンの微妙な配向の差に対応して, **A20** と近位ユビキチンとの相互作用が不安定となり, 40 ナノ秒程度で **A20** から解離することが判明した. 一方で, 遠位ユビキチンとの相互作用は主に疎水性相互作用から成るが, この相互作用に関しては, 直鎖型, **K63** 型両者のシミュレーションに於いて同様の相互作用が安定して保たれていた. 生化学実験からは **A20** は **K63** 結合型よりも直鎖型ユビキチンにより高親和性で結合することが示されているが, その機構が原子レベルのダイナミクスで明らかとなった.

I-7 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-7-1 実施計画

本研究では、「細胞内分子ダイナミクスシミュレーション」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25年度は、京コンピュータ用にチューニングした分子動力学計算プログラムSCUBAを用い、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。前年度に行ったカノニカルヒストンから構成される系との比較を行い、プロファイルの違いがどのような物理化学的なメカニズムによって生じるか明らかにする。また、メチル化したDNAが巻き付いたヌクレオソームでも同様な計算を行い、メチル化の効果を定量的に調べる。さらに、ヌクレオソーム構造下でのヒストンテールの構造探索を行い、テールの構造とクロマチン構造の関係を調べる。これらすべての全原子モデル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-7-2 実施内容（成果）

(1) ソフトウェアの開発・高度化の状況

大規模分子動力学計算プログラムSCUBAは、定期的な京コンピュータの計算環境アップデートに応じて再コンパイルを行った。それにより、別段不具合が生じることはなく、並列化効率99%以上を保持したままであった。今年度は、プロダクションランに専念した。

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結果を出すため、その差異を見つけるが一般的に非常に難しい。そこで、このわずかな構造の差を見つけ出して解析する手法についての研究に着手した。

(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

ヒストンの組成が異なるにもかかわらず、モノヌクレオソームのX線結晶構造は一見するとどれも同じ構造のように見える。しかし、細胞内では異なるタイミングでそれらのヌクレオソームが構築されることから、その構造やダイナミクスの特性を調べることは、転写、複製、組み換えなど生命の基本原理を理解するうえで非常に重要である。

本研究では、細胞周期に関わらず一定して発現されるカノニカルヒストンから構成されるヌクレオソーム（カノニカルヌクレオソーム）と染色体分裂期のセントロメアに特異的に構成される CENP-A ヒストンを含むヌクレオソーム（CENP-A ヌクレオソーム）の2種類について、それぞれ延べ2.5 マイクロ秒の分子動力学計算を行った。内訳は、反応座標全体をサンプリングする **adaptively biased** 分子動力学計算を1.5 マイクロ秒、その後、位相空間内で構造アンサンブルのオーバーラップを保証するための **umbrella sampling** を1.5 マイクロ秒実施し、自由エネルギープロファイルを得た。

図1に示すように、反応座標は DNA の両端間距離に取った。この反応座標上に、同時に100のヌクレオソームのレプリカを用意し、分子動力学計算を行った。各レプリカはサンプリングした位置に一定のポテンシャルエネルギーを落とすことで、サンプリングされたところが多数回サンプリングされるのを防ぎ、反応座標全体を舐めるようにサンプリングするようにした。1.5 マイクロ秒後、1 マクロ秒の **umbrella** サンプリングを行い、最終的な構造アンサンブルを取得した。得られた構造アンサンブルを重みつきヒストグラム法により解析し、自由エネルギープロファイルを得た。図2に自由エネルギープロファイルを示す。このプロファイルから、CENP-A ヌクレオソームの DNA は、H3 に比べて揺らぎが大きく、常に20~30 塩基対がヒストンから解離している状態が安定であること、カノニカルヌクレオソームは DNA を巻き付けている状態が最も安定であることなどがわかった。また、カノニカルヌクレオソームの自由エネルギープロファイルに見られるエネルギーのジャンプ（矢印部分）は、DNA とヒストンタンパク質の水素結合を失うことに対応していることがわかった。

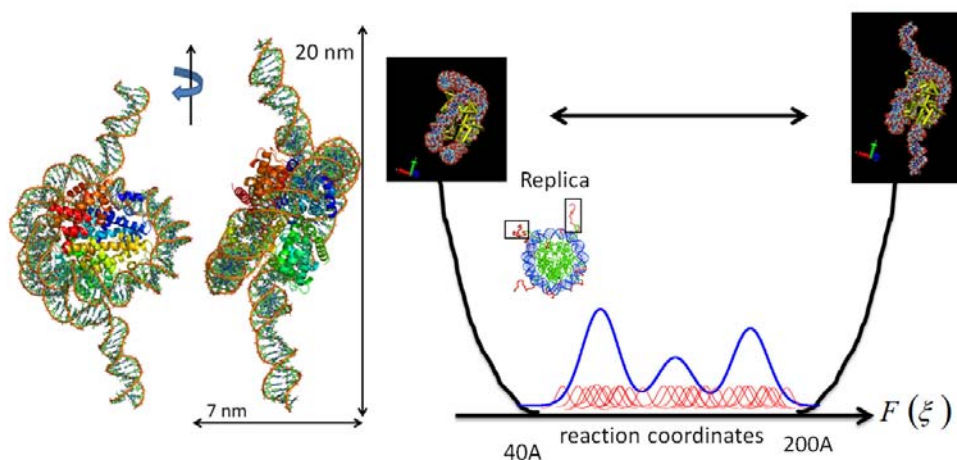


図1 設定した反応座標(左)と adaptively biased 分子動力学計算の概要

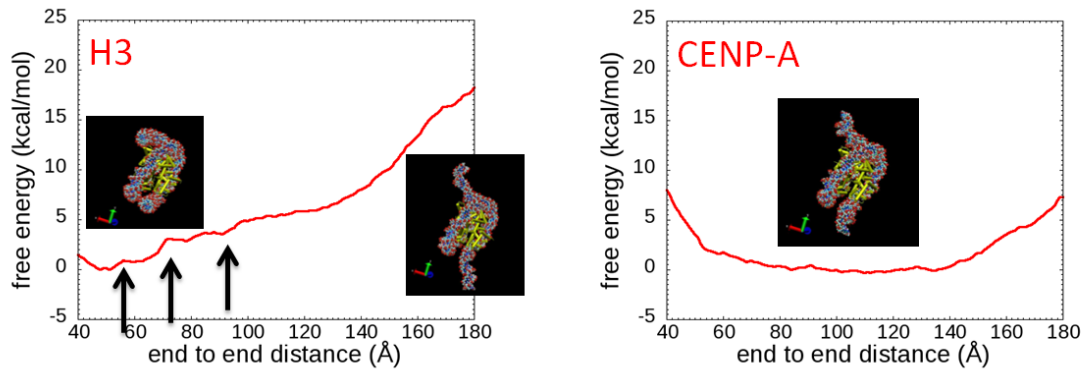


図2 カノニカル(左)と CENP-A(右)ヌクレオソームの自由エネルギープロファイル

また、自由エネルギープロファイルから、1塩基対のDNAを解離させるのに必要なエネルギーは、カノニカルヌクレオソームでは約0.4kcal/mol, CENP-Aヌクレオソームでは約0.1kcal/molであると見積もられた。一分子実験 (Mihardja et al. PNAS 2006, Kruihof et al. Biophys. J. 2009) では、対応するエネルギーコストは約0.1kcal/molと算出されており、実験結果と計算結果はよい一致を示した。

現時点の反応座標では、図1からもわかるように、部分的なDNAの解離しか計算できていない。H26年度は、完全にヒストンからDNAが解離する過程までの自由エネルギー計算を実施する予定である。

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、CG配列のシトシン(C)に対して特異的に起こる。シトシンの水素原子がメチル基に置き換わったとしても、その構造へのインパクトは小さいと考えられる。そこで、ヌクレオソームに巻き付いたDNAにおいて、すべてのCG配列のシトシン(合計8か所)をメチル化し、シミュレーションを実行した。独立に100nsの計算を4本実行した。100ns間の時間では、メチル化されたシトシンはグアニンと安定したベースペアを形成し、メチル化されたことによってベースがフリップアウトするようなことはどの部位でも観察されなかった。メチル化された部分のDNA構造や揺らぎの詳細な解析を現在実行中である。

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とそれに伴うテールの構造変化が、核内でのDNA収納状態をダイナミックに変える。このダイナミクスが転写制御に深く関わっていることが数多くの実験から示唆されている。テールの化学修飾による転写制御のメカニズムを理解するためには、修飾の有無によりテールの構造状態がどのように変化するかを明らかにすることが必要であるが、実験からその詳細な構造状態を知ることは非常に難しい。そこで、分子動力学計算により構造状態を調べた。高い空間分解能を持つ分子動力学(MD)シミュレーションは、テールの詳細な構造変化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ではテールの取り得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。

そこで我々は、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、 λ 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 λ がハミルトニアンに従って運動を行い、指定された λ 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。H25年度は、修飾の無いテールの構造状態をALSDによって明らかにした。この計算により、テールの存在領域は従来考えられていたよりもDNA周辺に限定されていること(図3)や、テール中に存在する二種類の正電荷アミノ酸であるリジンとアルギニンのDNA接触率が著しく異なることなど、これまで明らかとなっていなかったテールの性質を明らかにした。H26年度では修飾(アセチル化)されたテールで同様の計算を行い、両者の比較を行うことによって、修飾の有無によるテールの構造状態の変化を明らかにする。また、得られた構造アンサンブルは、粗視化シミュレーション結果の妥当性を示す指標となる。そこで、同じ修飾を持たせるなど共通分子を計算対象とするなどの研究チーム内での調整をおこなった。

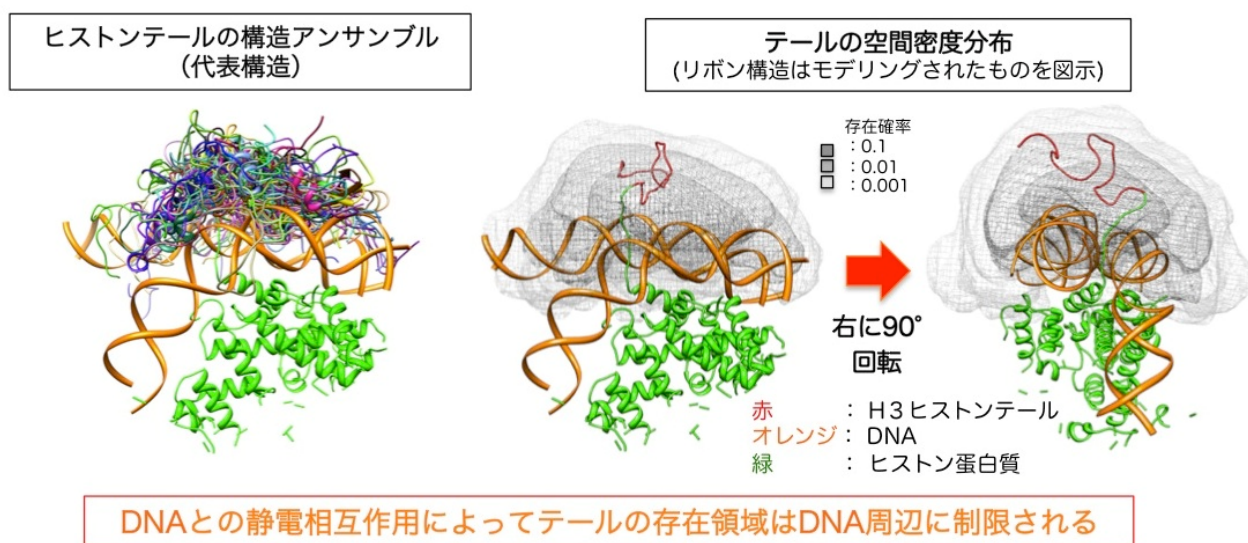


図3 H3ヒストンテールの構造分布。右図は、テールの空間密度分布を表す。