

平成24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）

マルチスケール・マルチレゾリューション自由エネルギー計算法による分子混雑と物質輸送に関する計算

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、生体膜の寄与など）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、構造変化に必要な自由エネルギー変化、基質結合に必要な自由エネルギー変化、酵素活性に必要な自由エネルギー変化などを計算し、膜蛋白質などの機能を定量的に解析する。

平成24年度は、(1) QM/MM自由エネルギー計算による酵素活性と、拡張アンサンブル計算に基づく構造変化の自由エネルギー変化を中規模の蛋白質について求めること、(2) 大規模構造変化を実現するマルチスケール法の開発、(3) 細胞環境を取り込む手法の開発と分子混雑による蛋白質の揺らぎや拡散運動の変化を解析する。

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、「分子混雑」や「物質輸送」に関わる実験研究者、「細胞スケールシミュレーション」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）

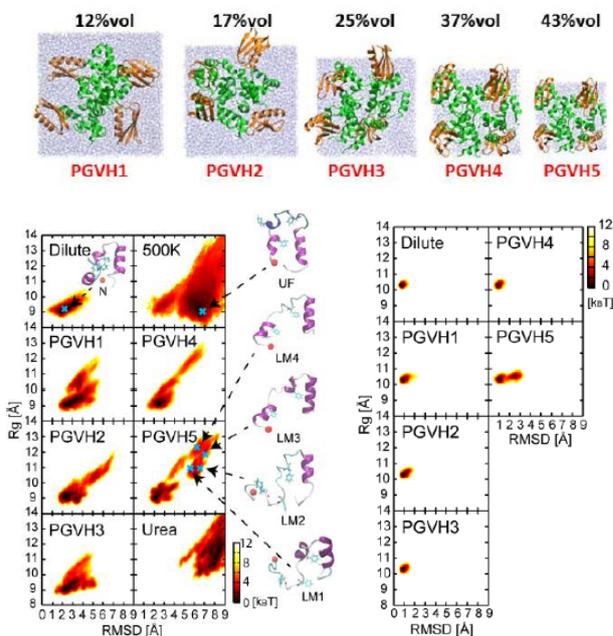
(1) ソフトウェアの開発・高度化の状況

細胞環境のような巨大システムの分子動力学計算を効率良く実行することを目的として、現在使用しているソフトウェア GENESIS（計算科学研究機構・粒子系生物物理研究チーム開発）に、並列ファイル入出力機能を新規に導入した。通常、小さな分子システムに対して分子動力学シミュレーションを行う場合は、構造ファイルやパラメータファイル、トポロジーファイルなどのインプットファイルはそれぞれ1個用意しておけば良い。しかしながら、細胞環境のような巨大分子システムの場合は、原子数が1億個以上あればインプットファイルの容量は10GB以上になり、ファイル転送や読み込み、変数のセットアップなどの前処理だけに数時間も要してしまう。新規に導入した並列ファイル入出力機能はこのような問題を改善するためのもので、本機能はインプットファイルをドメイン単位に分割し、分割されたファイルを各計算ノードに分散させて転送し、ファイルの入出力を各ノードで行う。この機能により、我々は1億原子系に対して数時間かかっていた計算の前処理を数十秒に短縮することに成功し、「京」において8192ノード以上でも効率よく計算を実行できるようになった。

(2) 研究開発の実施状況

1) 細胞内環境における全原子シミュレーション

細胞環境がタンパク質に対してどのような影響を与えるかを調べるために、異なる2種類のタンパク質を様々な濃度で含むシステムの分子シミュレーションを行い、タンパク質の物理化学的性質を分子混雑下と希薄溶液中の場合とで比較した。本研究では、動的・熱力学的性質が実験・理論的にもよく調べられている Protein G および Villin を使い、これらに対する分子動力学計算から、天然構造との RMSD および慣性半径を反応座標として自由エネルギー地形を解析した。その結果、Villin は分子混雑によって大きく天然構造が不安定化し、このような効果はすべての濃度条件下で現れることが分かった。一方、Protein G は混雑度が最も強いシステム (PGVH5) を除いては、どの濃度においても天然構造の安定性が保たれることがわかった (右図参照)。即ち、タンパク質の種類によっては構造安定性が異なり、常に希薄溶液中における天然構造が安定ではないことが示唆された。また、高温条件下 (500 K) や変性剤 (Urea) 存在下におけるタンパク質のアンフォールディングシミュレーションから得られた自由エネルギー地形と比較したところ、これらとも異なることが分かった。本研究により、分子混雑下特有の相互作用がタンパク質の構造安定性に影響を与えることを初めて明らかにした。



2) カルシウムイオンポンプのプロトン状態

カルシウムイオンポンプ (SERCA) は筋小胞体膜中に存在する代表的な膜輸送タンパク質である。 10^4 倍の濃度差に逆らって、筋細胞中から小胞体内部へと Ca^{2+} を輸送する。SERCA の反応サイクルでは、プロトンの対抗輸送、ATP の結合、加水分解などいくつかの反応が行われており、輸送に重要な役割を果たすことが知られている。本研究は、各反応機構を明らかにし、膜輸送タンパク質に共通なモデルを作成することを目的としている。今年度は特に SERCA のプロトン対抗輸送が、輸送機能にどのような役割を与えるのかについて研究を行った。その結果、プロトン状態の変化によってイオンの輸送経路である水分子のチャネルの構造に違いが現れることを示した。また、更に実験などで示唆された結合部位に存在する酸性残基の構造変化に注目した。プロトン移動による結合部位内の水素結合ネットワークの組み換えが、酸性残基の構造変化を引き起こすことを示した。これらの結果からプロトン対抗輸送が膜輸送に果たす役割について明らかにした。

3) 多剤排出トランスポーターの薬剤排出機構解明に関する研究

生体膜を介した物質輸送メカニズムの解明の一環として多剤排出トランスポーターの薬剤排出過程のシミュレーションを実施した。X線構造解析の結果と比較するため、タンパク質の立体構造、アミノ酸のプロトン化状態、空洞における脂質の有無などの条件を変えて、複数の平衡状態のシミュレーションを行った。その結果、プロトン結合部位の結合状態により、静電相互作用が異なり (図1)、周辺領域の水和構造、脂質、薬剤の結合位置 (図2) などが変化することが分かった。これら平衡状態のシミュレーションの結果により、プロトンの移動経路や構造変化を予測することに成功した。さらに、予測されたプロトンの移動、力を加えた構造変化を連続的に行うことで、平衡過程のシミュレーションと矛盾のない挙動を示すことも確認できた。

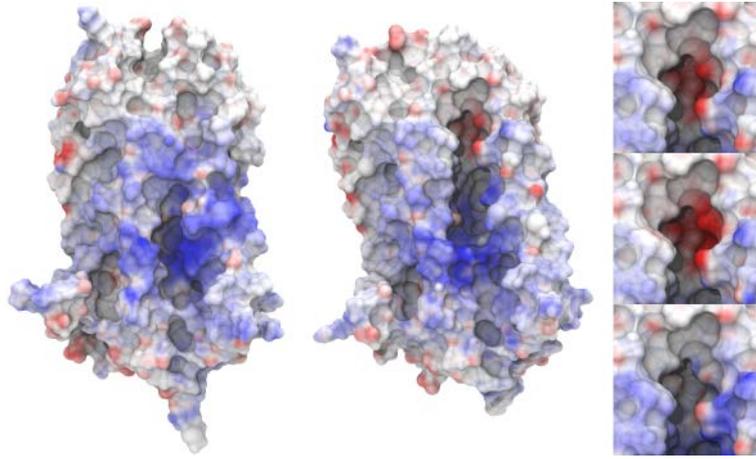


図1. 各構造・各プロトン化状態における静電ポテンシャル分布

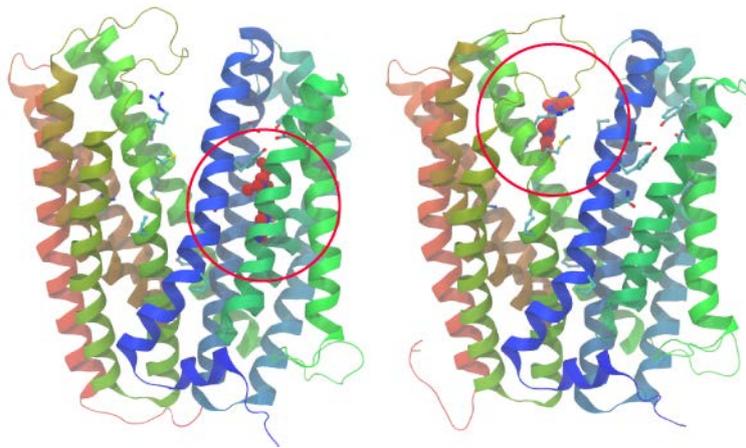


図2. 異なるアミノ酸のプロトン化による薬剤の結合・遊離状態

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24年度は、微視格子反応拡散コード pSpatioocyte の「京」での実行に向けた高度化および23年度構築したEGF経路モデルの改良を行なう。具体的には、(1)23年度に「京」へ移植した pSpatioocyte にハイブリッド並列化を施しさらにチューニングする。これを用い、(2)細胞内環境を考慮した反応モチーフレベルでの1分子粒度シミュレーションを実行し、動作を検証する。さらに、(3)タンパク質の細胞質から細胞膜および核内への移行といった動的な輸送ダイナミクスをEGF経路モデルに実装する。

I-2-2 実施内容 (成果)

(1) ソフトウェアの開発・高度化の状況

1) 微視格子反応拡散コード (pSpatioocyte) のハイブリッド並列化

前年度に、京コンピュータへと移植した細胞シミュレーションソフトウェア pSpatioocyte のハイブリッド並列化およびパフォーマンス測定を行った。現行の pSpatioocyte には、MPI ライブラリを利用したプロセス間通信のみが実装されており、このソフトウェアに OpenMP を利用したスレッド並列を実装した。具体的には、pSpatioocyte で、最も処理に時間を要する部分、タンパク質分子の拡散過程の処理をスレッドに分割し、処理の高速化を図った。ハイブリッド化した pSpatioocyte のパフォーマンスを、京コンピュータにて測定した (図1)。図1左には、512コア (64ノード) から、16384コア (2048ノード) を用いた場合の、ストロングスケールリングを示している。これによると、8192コア以降の高並列化では、若干低下するものの、良好にスケールアップしていることが分かる。また、図1右には、プロセス並列のみの場合 (Flat MPI) およびハイブリッド並列の場合 (Hybrid) の並列化効率を示した。低並列化 (~2048コア) で

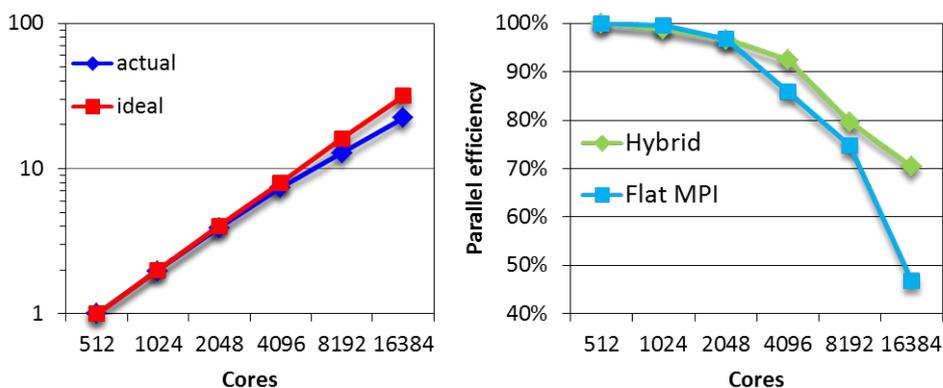


図1 pSpatioocyteのパフォーマンス測定結果

は、両方式に差は見られないが、高並列化（4096 コア～）すると、明らかに Hybrid の方が、高い並列化効率を示した。本年度の pSpatioocyte のハイブリッド並列化により、16000 コアまでの利用が実現し、今後は、更なるチューニングを施し、利用可能なコア数の上限を引き上げる予定である。

2) pSpatioocyte の動作検証

pSpatioocyte では、2 種類の反応様式、つまり、分子の拡散に依存する反応（拡散依存反応）と依存しない反応（拡散非依存反応）が想定されており、それぞれの反応が正確にシミュレーションできているか確認した。

まず、拡散非依存反応について、pSpatioocyte では、この反応の解法に、Gillespie の直接法と呼ばれる手法を利用している。表 1 に示す単純な反応モチーフを、pSpatioocyte、Gillespie の直接法および Spatioocyte (pSpatioocyte の元となったソフトウェア) で、シミュレーションし、その結果を比較した (図 2)。図 2 より、反応モチーフに含まれる 4 種類の化学種それぞれの計算結果が、完全に一致しており、pSpatioocyte にて、拡散非依存反応が正確にシミュレートできていることが確認できた。

一方、pSpatioocyte にて、拡散依存反応は、Collins-Kimball 法により計算されている。非拡散依存反応と同様に、2 種類の分子が結合する反応 (表 2) を pSpatioocyte および Spatioocyte を用いて計算し、その結果を比較した (図 3)。図 3 から、それぞれの化学種の計算結果は、完全に一致した。Spatioocyte では、Collins-Kimball 法による計算が正確に行われていることが確認されているため、pSpatioocyte も正確に計算できたこととなる。今後は、Michaelis-Menten 型の反応や MAPK モチーフ等のシミュレーションを実施し、その結果を検証する。

表 1 拡散非依存反応の動作検証に用いた反応モチーフ

反応 1	$\text{SpeciesA} \rightarrow \text{SpeciesB} + \text{SpeciesC}$
反応 2	$\text{SpeciesA} \rightarrow \text{SpeciesC} + \text{SpeciesD}$

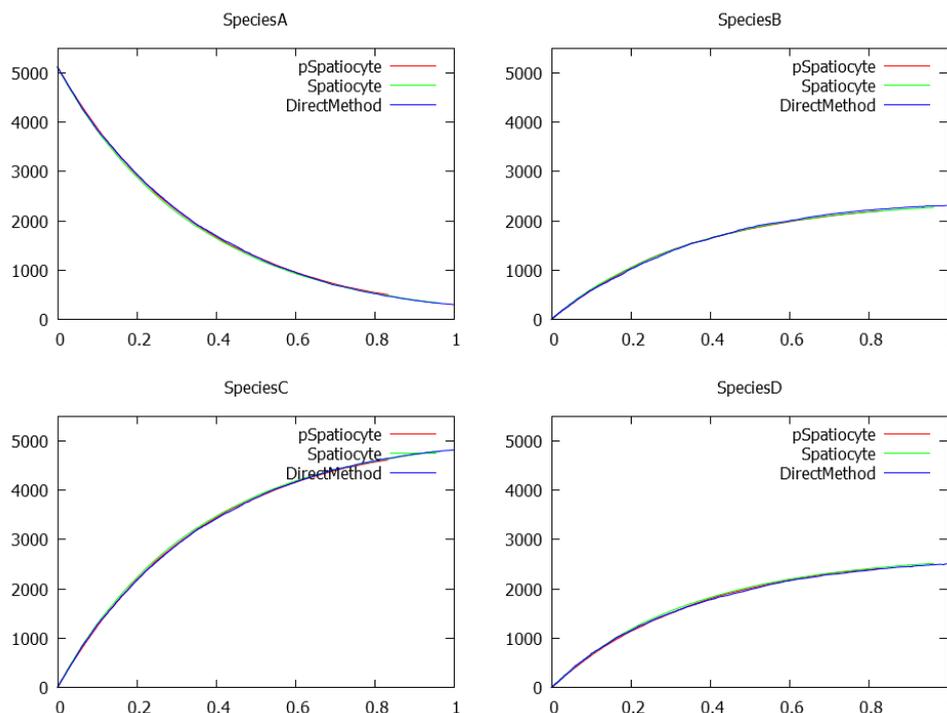


図 2 非拡散依存反応の計算結果の比較。反応モチーフに含まれる 4 種類の Species の結果が完全に一致している。

表 2 拡散依存反応の動作検証に用いた反応

反応 1	SpeciesB + SpeciesC -> SpeciesA
------	---------------------------------

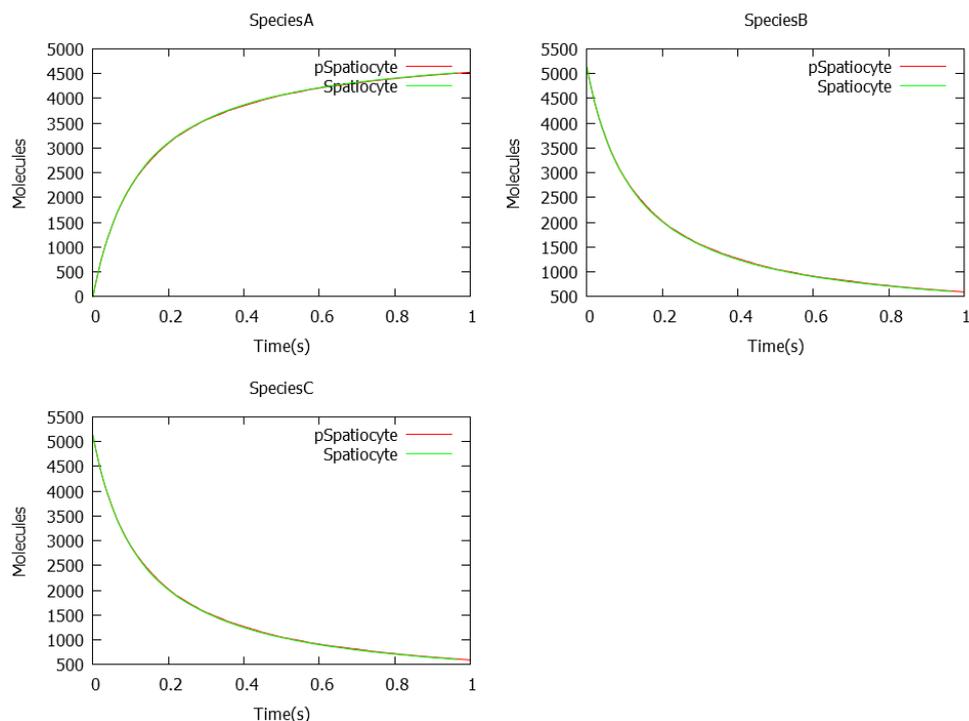


図 3 拡散依存反応の計算結果の比較。pSpatioocyteとSpatioocyteによる計算結果が完全に一致している。

(2) 研究開発の実施状況

1) EGF シグナル伝達経路モデルへの細胞膜および細胞核の追加

前年度に構築した、EGF シグナル伝達経路のプロトタイプモデルに、細胞核および細胞膜を追加し、細胞膜上の EGF 受容体の反応から、ERK タンパク質の細胞核への移行までをモデル化し、シミュレーションを実行した。

EGF 受容体 (EGFR) は、細胞膜に存在しており、そこで、細胞外の EGF と結合する。そして、二量体化、リン酸化、複数のアダプタータンパク質 (Grb2, Sos, Shc 等) との結合を経て、活性化型の EGFR となり、Ras タンパク質を GTP 型へと変化させる (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、Raf タンパク質は、細胞膜上で活性化する。活性化型の 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)。ERKPP は細胞核内へと移行し、様々な転写因子を制御する。また、MEK-ERK の二重リン酸化および脱リン酸化は、細胞質および細胞核の両方で行われることがよく知られている。以上の一連の反応を前年度のプロトタイプ EGF モデルに追加し、その概略図を図 4 に示した。この EGF モデルには、73 種類のタンパク質、144 本の反応、85 個のキネティックパラメータが含まれており、各タンパク質分子の個数や各キネティックパラメータには、実験値および文献値を利用した。また、細胞質および細胞核内のタンパク質分子の拡散係数は、 $1 \text{ um}^2/\text{s}$ に、細胞膜上のタンパク質分子は、 $0.01 \text{ um}^2/\text{s}$ にそれぞれ設定した。pSpatioocyte は未完成のため、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が、Spatioocyte は、シングルコアのみでの動作であるため、シミュレーション時には、実際の細胞の大きさの 1/500 程度まで縮小し、実行した。

図 5 に、EGF モデルのシミュレーション結果を示した。図 5 左上のグラフは、細胞核内の ERK タンパク質の動的挙動を示しており、シミュレーション毎 (細胞毎) にその挙動が異なってい

ることが分かる。このような核内 ERK の応答のバラツキは、細胞の応答不均一性と呼ばれており、実験的に観察されたデータと定性的に一致した。また、細胞核内の ERK のピーク値の分布を図 5 右上のグラフに示しており、これによると、二峰性の分布を示していることが分かる。ある EGF 濃度においては、核内の ERK が二峰性の分布を示すことが報告されており、こちらの結果も、実験データを定性的に再現した。さらに、図 5 下部には、当研究チームで開発中の蛍光顕微鏡シミュレータを利用して可視化したシミュレーション結果を示しており、上部が核内の ERK レベルが上昇しない細胞、下部が上昇した細胞の結果である。これを利用することで、シミュレーション結果と実験結果を直接、比較することが可能となる。今後の予定としては、(1) pSpatioCyet を利用しての実際の細胞スケールでのシミュレーション、(2) 細胞環境下(細胞内の構造物の配置や分子混雑など)でのシミュレーションを実施する。

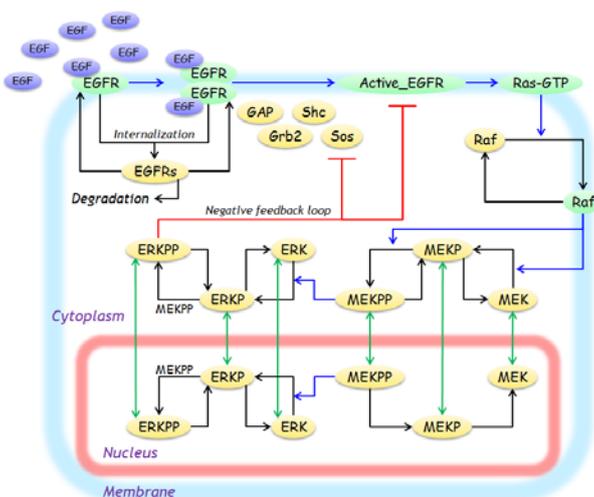


図 4 EGFシグナル伝達経路モデル

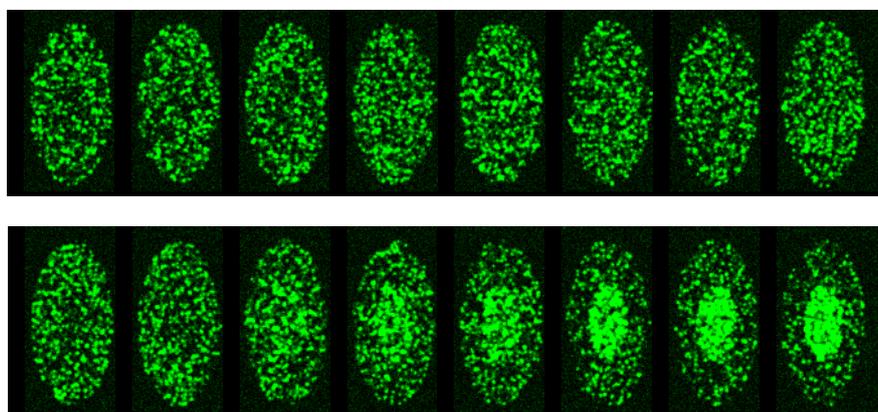
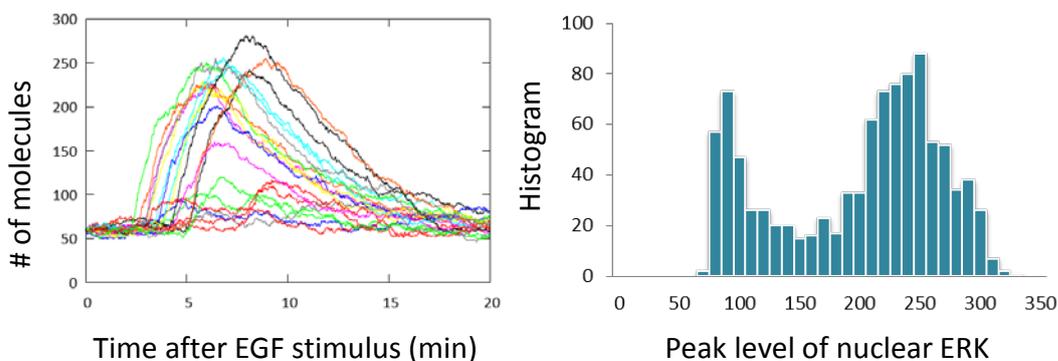


図 5 シミュレーション結果。(左上) 核内ERKの動的挙動。(右上) 核内ERKのピーク値の分布。(下) 蛍光顕微鏡シミュレータによるERKタンパク質の可視化。

I-3 太田 元規 (名古屋大学)

大量分子動力学軌道データに有効な初期解析法の開発

I-3-1 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、シミュレーション結果を解析する手法の研究開発を実施する。「細胞内分子ダイナミクスシミュレーション」研究では、高速な計算機を利用した分子シミュレーションが様々な階層、および環境で実施される。計算機が高速なため、シミュレーション結果は膨大なものとなり、解析を効率的に実施するには大量情報処理が求められる。本研究開発では、分子動力学シミュレーションが生成する大量の軌道データを比較し、効率的に意味のあるイベント（構造変化の軌跡）を自動で抽出する方法論の開発を目的とする。これにより、シミュレーション結果の事後処理が加速されることが期待される。

また、「大量分子動力学軌道データに有効な初期解析法の開発」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24年度は分子動力学法で生成された多数の軌道を分類するために、軌道を比較する新しい方法論の開発を継続して実施する。具体的にはカルシウムチャネルなどの分子動力学計算結果を対象に、その立体構造をモーションツリー (MT) 法で階層的に比較する。モーションツリー法の出力から軌道比較に必要な構造の違いの情報を効率よく取得する方法を調査する。また、軌道情報の圧縮表現についても検討を開始する。

I-3-2 実施内容 (成果)

(1) ソフトウェアの開発・高度化の状況

1) 24年度に開発したものは特になし。

(2) 研究開発の実施状況

1) カルシウムチャネルの構造比較

カルシウムチャネルの構造変化について、MT法を利用して研究を行っている。MT法を利用すると、構造決定者などのタンパク質のエキスパートが認識するドメイン構造が、自然と同定されることが、これまでの研究でわかっている。今年度は特に、E2・Pi状態 (E2状態にリン酸が付加した状態) の構造からE2状態の構造に至る構造変化について、分子動力学計算の結果を詳細に解析した。図1にE2・Pi状態のX線構造とE2状態のX線構造からMT法によってえた、ドメインを示す。この結果からE2・Pi状態からE2状態の間でおこる構造変化は、Piが放出されることでPドメインが緩和し、M1とM2の間が閉じる、というものであると解釈できる。分子動力学計算はE2・Pi状態、E2・Piの構造か

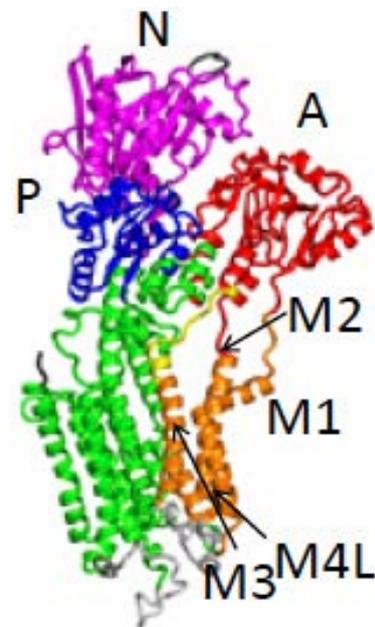


図1 : E2・Pi状態の構造とE2状態の構造間でおこる構造変化をMT法で解析した結果得られたドメイン構造

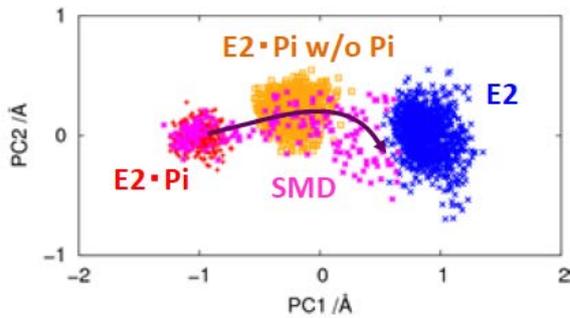


図 2 : P ドメインの構造変化の主成分解析結果

らリン酸をはずしたもの, E2 状態について行った. また, E2・Pi 状態からリン酸を序々にはずす計算を行った. 主成分解析の結果から, P ドメインの動きは, E2・Pi 状態から Pi を外した状態をへて, E2 状態に至るという経路をとることがわかった (図 2).

2) モーションツリー法を利用した分子動力学軌道の解析法

アクチンキャッピング

タンパク質 (CP) を題材に, MT 法を利用した分子動力学軌道の解析にとりくんでいる. CP はアクチンフィラメントの反矢じり端に結合し, その伸長を止める. CARMIL タンパク質

(CARMIL) は CP に結合し, キャッピング機能を阻害する. CARMIL は CP のアクチン結合面と違う部分に結合するので, アロステリック阻害を行っている. 我々はこのアロステリック阻害は, CP の構造変化を抑制する

ことにより行われるという仮説をたてている.

CP, CP/CARMIL 複合体について, 水を露わに含んだ分子動力学計算を実施した. それぞれ, 100ns の計算を 3 本行った. 使用プログラムは Amber,

図 4 : Free の CP と CP/CARMIL

で出現するリジッドドメインの数. 1-40,268-308 (左上) などはキノコの柄に相当する. X 軸近傍にある 445-450 が, V-1 結合ループの隣のループに該当する.

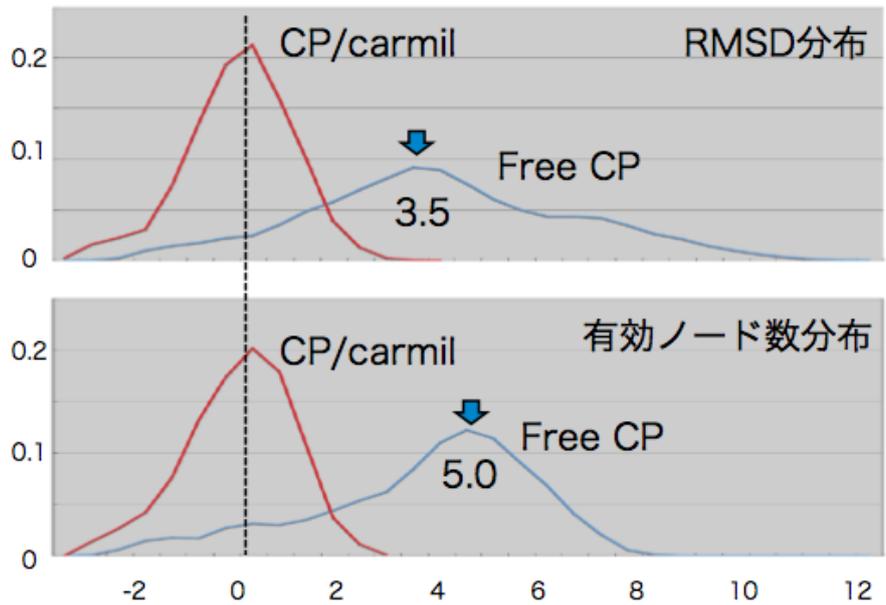
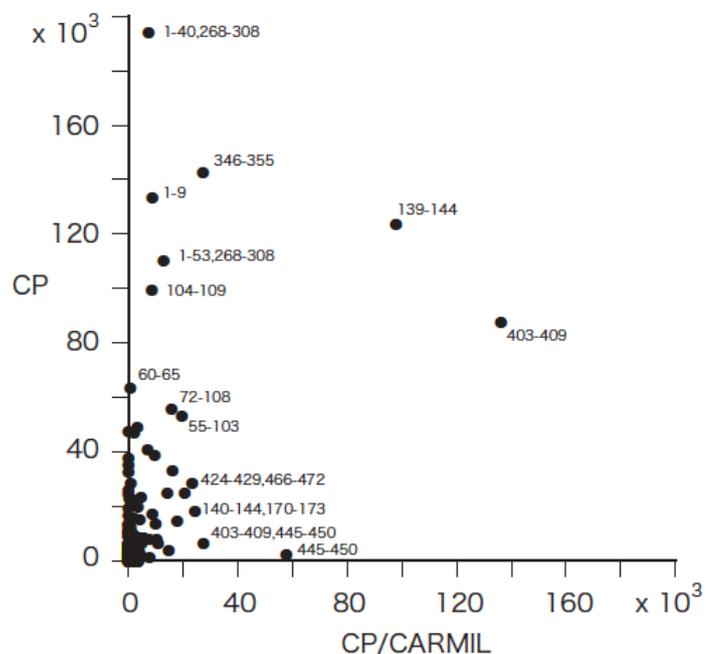


図 3 : CP 構造の比較結果. 軌道から 2 つの構造を取り出し, RMSD と MT の有効ノード数の分布を求めた. CP/CARMIL 複合体の分布を利用し SD スコア (X 軸) の分布とした.



マシンは情報科学研究科が有する Xeon (8 コア, 72 ノード) を利用した。

CP, CP/CARMIL 複合体の分子動力学計算の結果得られた軌道について, 構造比較を実施した。

CP, CP/CARMIL 複合体の軌道毎に解析を行った。軌道から 2 つの構造をとりだし, RMSD を計算した。一方, 2 つの構造から MT 法によりツリーを作成した。MT 法では構造変化の数 (有効ノードの数) が出力される。複雑な構造変化がおきる場合, 有効ノードの数は増加する。図 3 に RMSD と構造変化の数の分布を示す。CP が CARMIL の結合により運動を抑制される様子 (CARMIL が結合していない free の CP が揺らぐ様子) は, RMSD よりも MT の方が顕著に捉えていることがわかった (分布の差が顕著である)。

次に, CARMIL 結合により引き起こされる構造揺らぎの違いを表現するために, MT 法が出力するドメイン (正確にはリジッドドメインで, 構造変化の単位) について統計をとった。ある種の基準を導入し, 出力されるドメインをクラスタリングした。各ドメインについて, CP と CP/CARMIL の出現数の相関を見た (図 4)。Free の CP でよく出現するが CP/CARMIL では出現しないドメインは, CP をキノコに見立てた時の柄の部分に該当するところであった。一方, CP/CARMIL では良く出現するが CP では出現しないドメインは, V-1 という CP の阻害剤に結合するループの, 隣のループであった。CP は CARMIL 結合により, キノコの傘と柄の部分の運動が阻害される。その一方で, V-1 結合ループの隣のループ運動が活性化される。ということがわかった。CARMIL によるアロステリック阻害機構とここで得られた発見との関係性について現在調査を行っている。

3) 軌道情報の圧縮表現の検討

MT 法を利用した軌道情報圧縮法について検討した。以下のような方針を考えた。① Ca だけを考え, 他の原子は無視する。②構造変化のない構造は考慮しない。③構造変化は MT 法で求めたドメインの座標変換で表現する。④③のように考えた場合, 軌道は 1 つのリファレンス構造と座標変換の列になる。座標変換は 1 構造変化について, 3×3 の回転行列と 3 次元の移動ベクトルで表現されるので 12 個の数字になる。

実際にどの程度圧縮されるのかを調べるために, CP の軌道, 100ns, 50,000 スナップショットについて, 構造変化の数を計算した。スナップショット 1 とスナップショット 2 を比較し, 有効ノードを持つ MT が得られない場合, スナップショット 2 を捨てる。次にスナップショット 1 とスナップショット 3 を比較する。もしも構造変化があれば, 構造変化数としてカウントし, スナップショット 3 をリファレンス構造として更新する。こういう作業を行ったところ, 構造変化は 606 個であった。よって, ②の工程により構造情報は約 12% (606/50000) に圧縮された。

I-4 石谷 隆一郎 (東京大学)

膜輸送体の分子動力学シミュレーションによる膜を介した物質輸送メカニズムの解明 研究

I-4-1 実施計画

本課題では、「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、生体膜を介した物質輸送メカニズムの解明に向けた研究開発を実施する。この生体膜を介した物質輸送は、膜に埋め込まれた輸送体蛋白質によって担われているが、従来の研究から、輸送体蛋白質は動的な構造変化を伴って、膜を介した物質輸送を行うと考えられている。そのため、輸送機構の解明には、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、輸送体蛋白質に分子動力学(MD)シミュレーションを適用し、物質輸送の動的機構解明を目指す。

また、「細胞内分子ダイナミクスシミュレーション」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。

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24年度は、膜を介した物質輸送に関わる輸送体の中から、特にジペプチド・薬剤輸送体 POT を取り上げ、MDシミュレーションを行う。POT は細胞膜を介したプロトン濃度勾配をエネルギー源としてジペプチドや薬剤の取り込みを行う共輸送体である。当研究室で得られている POT の高分解能結晶構造をもとに、プロトン結合状態、基質結合状態、輸送体単独状態の MDシミュレーションを実施し、プロトン濃度勾配をエネルギー源とした輸送機構の解明を目指す。一方、昨年度から継続してマグネシウムイオン輸送体 MgtE の Mg²⁺イオン選択性機構解明を目指したシミュレーションも実施する。

I-4-2 実施内容 (成果)

(1) 研究開発の実施状況

まず、細胞膜を介したプロトン濃度勾配を用いてオリゴペプチドを輸送する、ペプチド輸送体 POT を取り上げた (図1)。平成23年度試験研究では、脂質二重膜を含む系で MDシミュレーションを行った (図2)。数百ナノ秒程度のシミュレーションを行ったところ、細胞内向き開状態の POT について、プロトン・基質オリゴペプチドともに非結合状態で、細胞内向き開状態から閉状態への自発的な構造変化を見出すことに成功した (図3)。

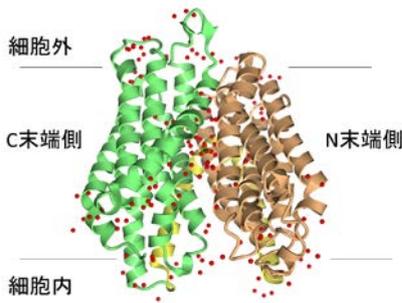


図1) オリゴペプチド輸送体 POT の結晶構造

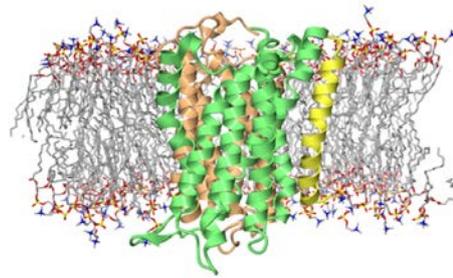


図2) POT の MD シミュレーションに用いた系

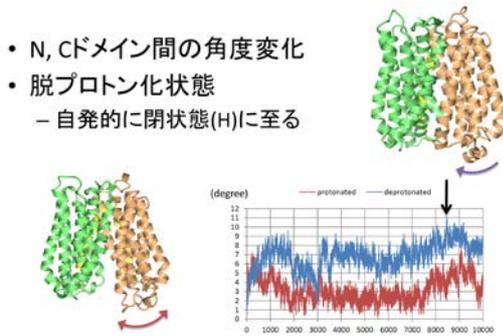


図3) 脱プロトン化状態のシミュレーション結果

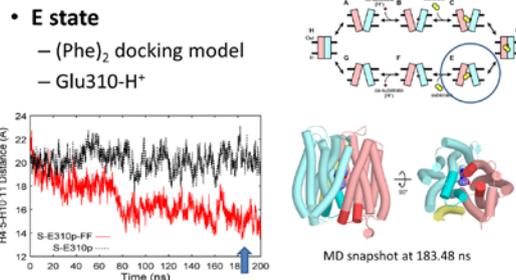


図4) 基質・プロトン結合状態のシミュレーション

この構造変化は、初期状態が異なる複数の計算でも見出されることから、POT はプロトン・基質オリゴペプチドともに非結合状態では閉状態に移行しやすいと考えられる。一方数種類の異なるプロトン化状態についてもシミュレーションを行ったが、こちらでは構造変化は見いだされなかった。これらの結果は実験結果とも一致するものである。さらに、ペプチド複合体についても長時間のシミュレーションを行った。その結果、非結合状態の場合同様に、閉構造への構造変化が起こることを見出した(図4)。現在、生化学的な実験データと併せて、細胞膜を介したプロトン濃度勾配を用いてオリゴペプチドを濃度勾配に逆らって輸送できる動的な分子機構のモデルを構築し、論文を執筆・投稿中である。

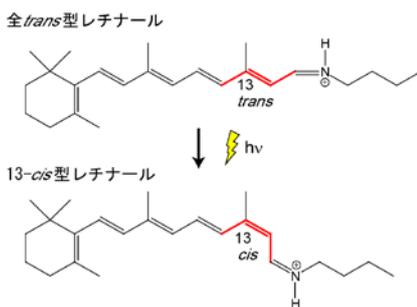


図5) レチナールの構造式

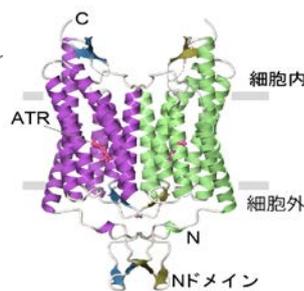


図6) ChR2 の結晶構造

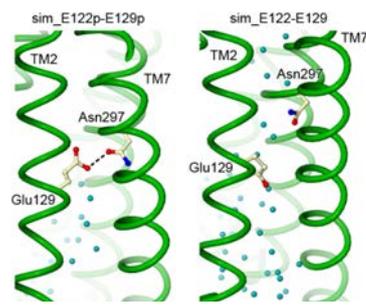


図7) チャンネル孔への水分子の流入

本研究ではさらに、光によって開閉が制御されるチャンネルである、チャンネルロドプシン(図6)に関して、脂質二重膜を含む系でMDシミュレーションを行った。光を受容するレチナールに関して、プロトン化状態、異性化状態(図5)について何通りか初期構造を作成し、各々数百ナノ秒程度のシミュレーションを行った。その結果、光を受容した後の状態に相当するシス型レチナールにおいて、計算機内でもイオンチャンネルが閉状態から部分的開状態へ変化することを見出した(図7)。現在、さらに長時間のシミュレーションを行うことで、より開いた状態に至るかどうか検討している。

I-5 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-5-1 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24年度は、前年度に確立した計算プロトコル及びチューニングした分子動力学計算プログラムを用いて、ヒストンタンパク質に巻きついているDNAがほどける過程について自由エネルギープロファイル計算を実行する。ヌクレオソームを構成しているヒストンタンパク質の組み合わせの違いによる自由エネルギープロファイルの違いを調べ、ヒストンタンパク質の組成が異なるヌクレオソームそれぞれのダイナミクス（物理的な性質）の違いを明らかにする。また、化学修飾を受けるヒストンテールについて、その分子ダイナミクスを調べる。

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大規模分子動力学計算プログラム SCUBA のチューニングを行い、4801 ノードで現在計算を進めている。原子間相互作用のエネルギー計算に必要な誤差関数の計算部分をテーブル参照に変えたことと原子間相互作用リストの並び替えによってメモリアクセスを最適化したことにより、約20%の計算スピードの向上に成功した。ただし、これによりFlops値は下がった。必ずしも、実際にかかる計算時間とFlops値で測る実行効率が対応しない。また、プログラムのSIMD化を行い、SIMD化率を約5%から約30%に向上させることに成功した。しかし、計算スピードの向上は見られなかった。ここでも、SIMD化率と実際の計算速度の関係は必ずしも比例するとは限らないことを示す結果となった。今年度でSCUBAのチューニングは一区切りをつけ、次年度以降はプロダクション計算に集中する。

（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。

そこで、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 CENPA ヒストンを含むヌクレオソーム (CENPA ヌクレオソーム) の2つについて、延べ1.5 マイクロ秒の分子動力学計算を行った。本計算では、構造のわかっていないヒストンテール領域は除いてある。これまでに報告されているヌクレオソームの計算時間が100 ナノ秒であることを考えると、京コンピュータを使うことによって延べ時間にして約15倍の計算を実行することができた。

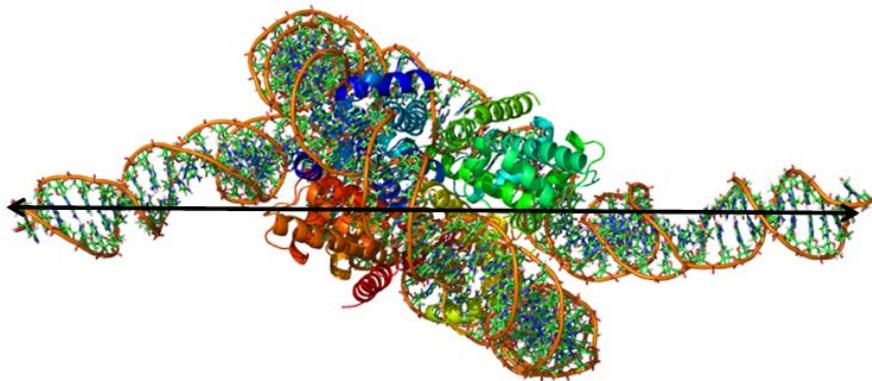


図1. 反応座標をDNAの両端距離にとる

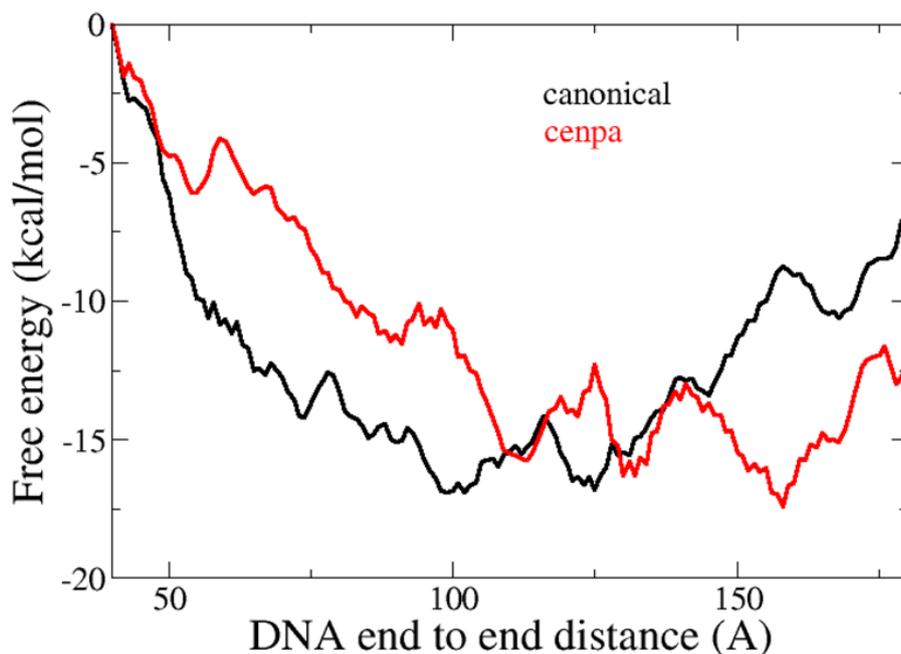


図2. 自由エネルギープロファイル

自由エネルギーの計算結果はまだ収束しておらず、自由エネルギーの細かな凸凹など定量的なことを議論する段階にはない。積極的に大規模計算を行うように指示があったにも関わらず、そのような計算ジョブはキュー待ち状態になってしまい、思うように計算が進まなかった。次年度は、前半から計算ジョブを投入していく予定である。

図1のように、DNAの両端の距離を反応座標とし、ヌクレオソームからDNAがほどけた状態の自由エネルギー計算を行った。図2に自由エネルギーを示す。反応座標は、40~180Åとし、その間で最もエネルギーが高い状態をゼロとして表示した。注目すべきは、どちらも結晶状態の構造(約65Å)からDNAがほどけた状態の方が安定であるという点である。図からわかるように、カノニカルヌクレオソーム(黒線)に比べてCENPAヌクレオソーム(赤線)の自由エネルギーの底は、DNAの末端間距離が大きいところにある(反応座標にして、110から160Åあたり)。つまり、CENPAヌクレオソームはよりほどけた状態の方が安定なことがわかる。しかし、自由エネ

2) ヒストンテールの構造サンプリング

効率よく分子の構造をサンプリングする方法自体、計算機シミュレーションの世界では大きな研究課題である。これまでいろいろな方法が提案されてきたが、partial multi-canonical (Okumura, 2008, JCP) は、一部の相互作用エネルギーのみをスケールすることによって効率的に構造空間を探索することができるので、ヒストンテールの構造探索に最適であると考えてプログラムに実装した。結果、この手法は理論的には欠陥がないが、静電場の強い系においてはある特定の構造に捕捉されて効率的な構造探索ができないことがわかった。そこで、新たに adaptive lambda square dynamics (ALSD) という方法を開発した。この方法は、原子間の相互作用パラメータをスケールする lambda 変数を取り入れ、それをダイナミックに変えながら、lambda 空間で一様に構造をサンプルすることで、先の方法の問題点を解決した方法である。得られたアンサンブルに適切な重みをかけることにより、常温(300K)でのアンサンブルが得られる。この方法が正しく動作しているかを示すために、ヒストンテールを想定したポリリジンの構造探索を行った。比較として、これまでの研究でさまざまな生体分子系に適用され、実験結果と良い一致を示すアンサンブルを得ることができる multi-canonical MD 計算を用いた。図3に示すように、両者でほぼ同じ自由エネルギー地形が得られた。また、自由エネルギー地形を得るための計算量を比較すると、ALSDは multi-canonical MD と比べて約20倍計算効率が良いことがわかった(図4)。今後は、この方法を用いてヒストンテールの構造サンプリング及びヒストン修飾のありなしの構造変化を調べる。

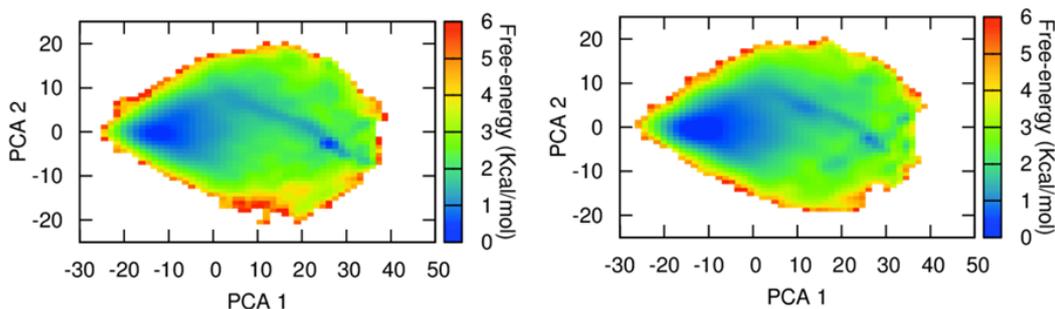


図3. ポリリジンの自由エネルギー面(2次元に射影)
Multi-canonical MD (左)と adaptive lambda square dynamics (右)

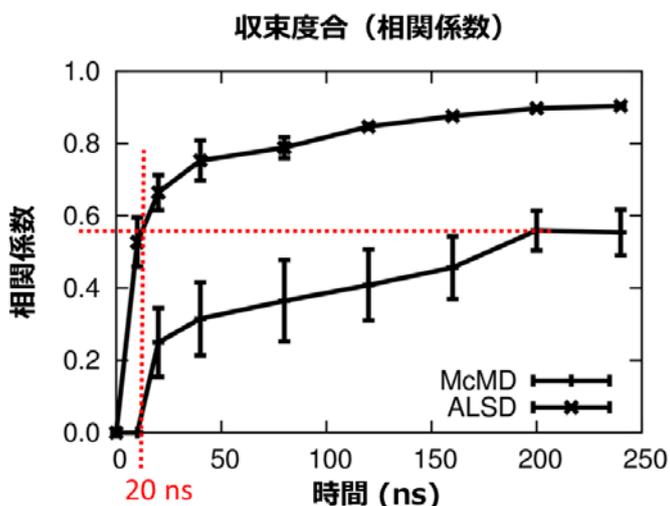


図4. 計算の収束度合
長時間の計算結果に対する、Multi-canonical MD (McMD)と
adaptive lambda square dynamics (ALSD) 収束具合。