

平成23年度における研究成果

I 戦略課題 1：細胞内分子ダイナミクスシミュレーション[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

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研究を行う。「細胞内分子ダイナミクス」研究では、これまでの水中、脂質二重膜中、真空中などの環境ではなく、細胞環境（細胞質、生体膜、核など）を露に考慮した環境での生体分子ダイナミクスを研究する。そのためには、大胆な粗視化と精度の高い自由エネルギー計算が必要であり、本研究課題ではこれらを開発し、「京」での大規模自由エネルギー計算を実行し、最先端計測などと直接比較する。

平成23年度は、(1)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の開発と「京」への高度化を行うこと、(2)「細胞内分子ダイナミクス」の解析に必要なツール群を開発すること、(3)イオンポンプのダイナミクスに関して分子動力学計算を用いて解析することなどを実施する。

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

1) 分子・細胞シミュレータ GENESIS の開発

計算科学研究機構粒子系生物物理研究チームが主に開発している分子・細胞シミュレータ GENESIS の高度化や多次元レプリカ交換分子動力学計算法の開発などを行った。「京」の機能を効率よく使うための SIMD の最適化やソフトウェアパイプラインの利用の促進を行い、計算効率が向上した。また、レプリカ交換分子動力学計算法を導入することで並列に複数の分子動力学計算を実施することができるようになり、2,000 ノード以上の利用が実現した。また、GENESIS に搭載している分子動力学のハイブリッド並列化を実施したことにより、今後は 10,000 以上のノードでのレプリカ交換分子動力学計算の実施が可能になる。

マルチスケール・マルチレゾリューション分子動力学計算を実施するために、全原子モデルとしては CHARMM、AMBER の力場パラメタを導入した。また、粗視化分子モデルとしては MARTINI 分子力場と Go model を導入することに成功した。これらのモデルをレプリカ交換分子動力学計算などの enhanced conformational sampling 法と組み合わせて計算することが可能になったため、構造空間探索効率が格段に向上した。

2) MotionTree 法を用いたトラジェクトリ解析の実施

名古屋大学の太田らによって開発された MotionTree 法を、複数のドメインを持つ蛋白質の構造変化の解析に応用した。特に、本研究課題の重要なターゲットの一つでもあるイオンポンプのドメイン運動の解析に適用し、この手法を用いることで複雑なドメイン運動を効率よく解析できることを確認した。

(2) 研究開発の実施状況

1) 分子混み合いのモデル系のシミュレーションの研究

細胞環境を考慮したシミュレーションの第一弾として、細胞質内分子混雑環境を考慮したシミュレーションを2つ実施した。まずは、NMRの実験と比較するために、BSAあるいはLysozymeを混雑物として含む系におけるCI2の安定性変化を評価した。水を露に含む全原子分子動力学計算を行うことにより、BSA存在下では不安定化しないCI2がLysozymeとの静電相互作用によって不安定化することが分かった。この結果は、PielackらのNMRの測定結果と良く一致した。

次に、混雑濃度を変化させ(図1)、VillinとProtein-Gを含む系での水和の変化を検討した。全原子分子動力学計算を図1に示したそれぞれの系について300ns実行することによって、蛋白質濃度が変化した場合の周囲の水の拡散係数や誘電率の変化を見積もることに成功した(図2)。これらの物理パラメータを実験的に測定することは困難であるため、今回の計算結果は非常に有用であると考えられる。また、「京」の利用が実現した場合にはさらに大規模な系での分子混雑の解析が実現すると期待できる。

図1. 蛋白質濃度を变化させたシミュレーションモデル

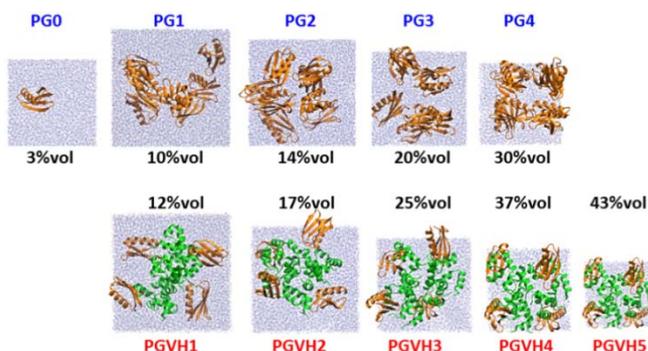
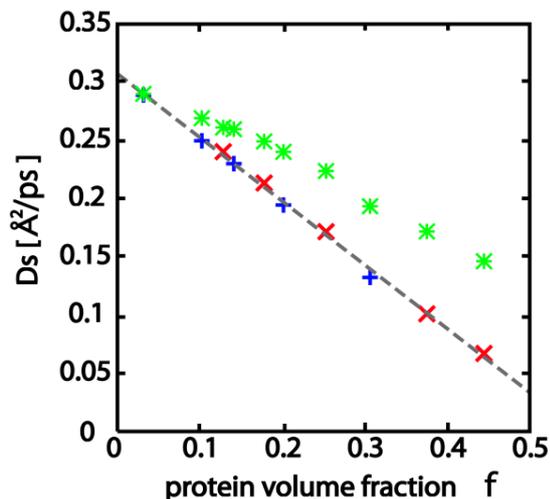


図2. 水の拡散係数の変化



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

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。

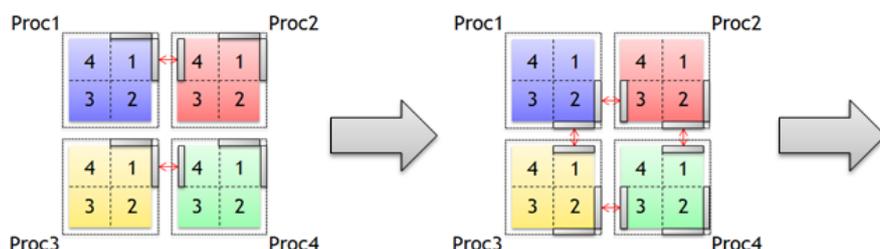
平成23年度は、複雑な反応ネットワークを1分子粒度で表現するために必要な基盤技術の開発を行い、次年度以降での具体的なモデル作成およびシミュレーション実行へのはずみを付ける。具体的には、(1)微視格子反応拡散コードの「京」へ向けた高度化、および(2)複雑な反応ネットワークをモデル化する際に不可避な組み合わせ爆発を解決するためのルールベース技術の導入、の二点に重点を置いてコード開発を行う。また、(3)真核細胞の信号伝達系の大規模シミュレーションで一番乗りを目指すため、EGF経路モデルのプロトタイプ作成を通じて準備研究期間で策定した研究戦略をさらに洗練させ具体化する。

I-2-2 実施内容（成果）

(1) 微視格子反応拡散コードの「京」へ向けた高度化

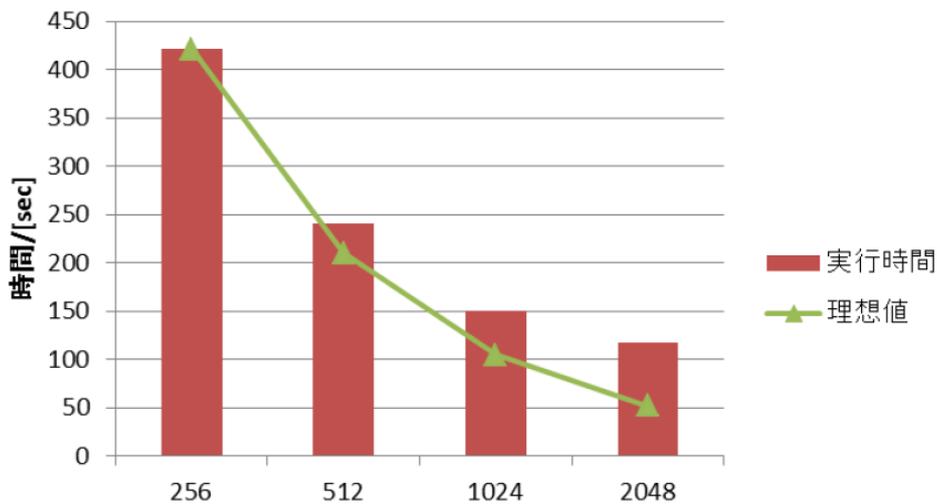
1) Parallel Spatiocyte (pSpatiocyte)の開発

本年度の開発で、本プロジェクトで用いる微視格子反応拡散コード pSpatiocyte の「京」への移植が完了した。また、反応過程の実装も完了し、反応=拡散の連成計算が可能になった。pSpatiocyte では6方細密格子で離散化された空間上の格子点の状態として分子を表現し、これをさらに空間分割により各プロセッサに分担させ並列化を行う（図参照）。このとき、袖領域にある分子の運動により隣接するプロセッサが担当する部分空間にある分子との衝突が起こる場合があり、この衝突解決を矛盾なく行うために、各プロセッサが担当する部分空間をさらに小さな部分空間に分割し、これらを毎ステップ違う順番でアップデートする。この方式により、逐次版コードと同等の精度を持った計算を矛盾なく並列計算出来るようになった。「京」への移植は、ベースとしたコード Spatiocyte が E-Cell 3 プラットフォーム上の C++ 言語で Boost テンプレートライブラリを利用した比較的高度なプログラミング技法を利用したものであったのに対し、pSpatiocyte では基本的な C++ 文法に限定したコーディングを心がけたため、最小限のコードの変更で移植が完了した。ただし、細胞シミュレーション分野では高度なソフトウェア工学が要求される場面もあり、今後のプロジェクト進展に従って利用する C++ 文法や依存ライブラリも高度化する事が予想されるため、この問題への対処を引き続き考慮する必要がある。



2) パフォーマンス測定

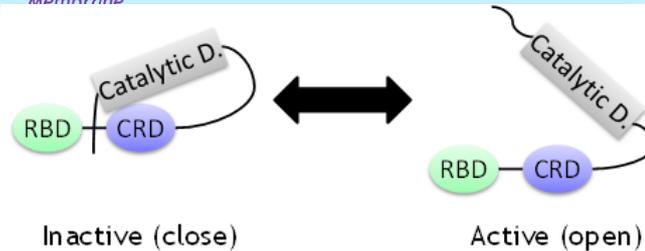
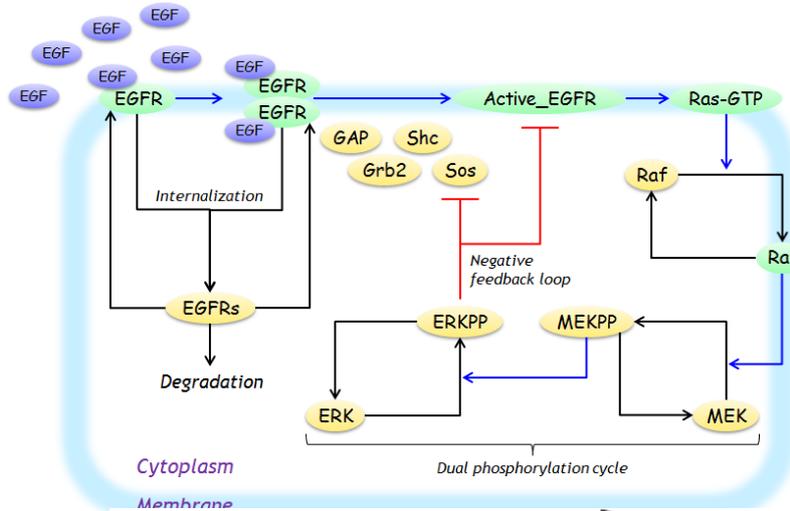
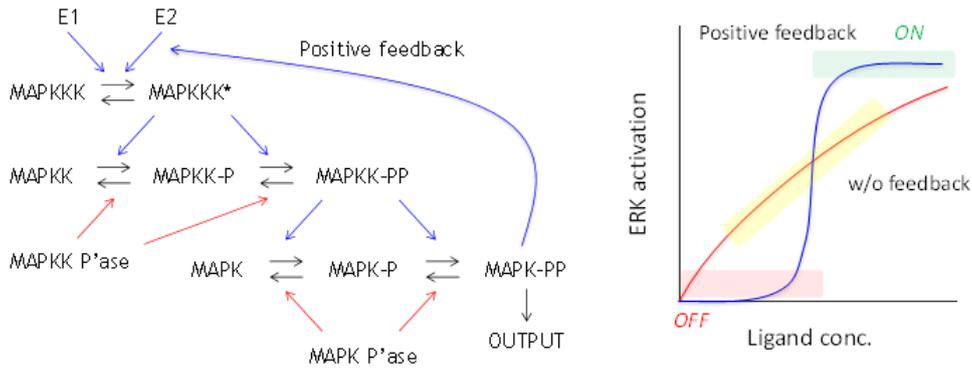
「京」実機を用いて pSpatioocyte の性能測定を行い、フラット MPI で 2,048 コアまでの強スケーリングを確認した。用いた条件は、格子数約 2.6 億 (640*640*640 の直方体)、分子数は約 8 千万 (細胞内の分子混雑と同等の 30% 占有率とした) である (図)。同じく、格子数 80 億、分子数 24 億の条件でも測定を行い、若干の効率低下を確認した (データ示さず)。これら 2 つの条件でのプロファイル比較により、マスタープロセスによるブロードキャストと C++ STL 利用箇所におけるボトルネックを同定し、改善策を策定した。本報告書作成時点では改善策の実装まで至っていないが、この実装が完了すれば、現在目指している 80 億格子において 10,000 コア以上までのスケーリングの達成に向け、理論上は大きな障害は見当たらない。なお、利用コア数が少ない場合に発生するメモリー量の制限により今回は 1 ノードあたり 4 コアのみ利用としたが、詳細プロファイルからメモリーバンド幅には十分余裕がある事がわかっており、高並列度において 8 コア利用する事に問題はないと考えられる。次回以降の性能測定はノードの全コア利用とする予定である。ハイブリッド並列の実装は 24 年度内で完了する予定である。



(2) 研究開発の実施状況

1) 研究目標

本プロジェクトでは真核細胞の 1 分子粒度まるごとシミュレーションの実現を目標としている。その具体的なターゲットとして、上皮成長因子 (EGF) シグナル伝達経路に狙いを定めた。ヒト細胞は EGF 刺激に対して不均一に応答する事が実験的に知られている [1, 2]。この現象は細胞種、刺激因子によらず広く見られる事が知られており、細胞のがん化、発生過程、薬剤耐性等に関与すると考えられている。応答不均一性の要因としては、ネットワークノイズ [3] や分子の構造多型 [4] が言われるが、議論は決着していない。このような反応ネットワークや分子状態の動態は分子混雑、細胞骨格、分子のクラスタ化などの細胞環境が大きな影響を及ぼす事が知られており [5, 6]、細胞環境をあらわに表現した 1 分子粒度モデルを用いたシミュレーションが必要である。本プロジェクトでは EGF 刺激に対する細胞の応答をモデル化するため、細胞に与えられる EGF 分子の濃度を入力、核内の活性型 ERK (Extracellular signal Regulated Kinase) 分子のレベルを出力と定義し、定義された入力に対する出力のばらつきや動態を計算する。



ネットワーク構造と応答特性[3]

Raf 分子の構造多型[4]

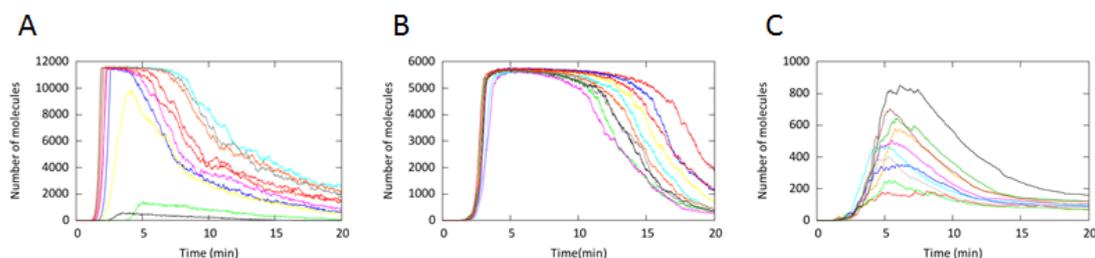
2) プロトタイプ EGF モデルの構築

本年度は EGF 応答経路の基本モデル構築を完了した。このモデルは 59 個の化学種と 109 個の反応から構成される。パラメータ条件は今後協力研究者による実験により詰めてゆくが、本年度作成したモデルでは既存モデル[7]の値 (モデル 1)、文献[8]からの推定値 (モデル 2)、モデル 2 の値を一部改変 (モデル 3)、の 3 パターン作成した。分子の拡散速度は $1 \mu\text{m}^2/\text{s}$ (細胞質)、および $0.01 \mu\text{m}^2/\text{s}$ (細胞膜)、プロトタイプ計算に用いた細胞サイズは実際の 1/100 程度とした。

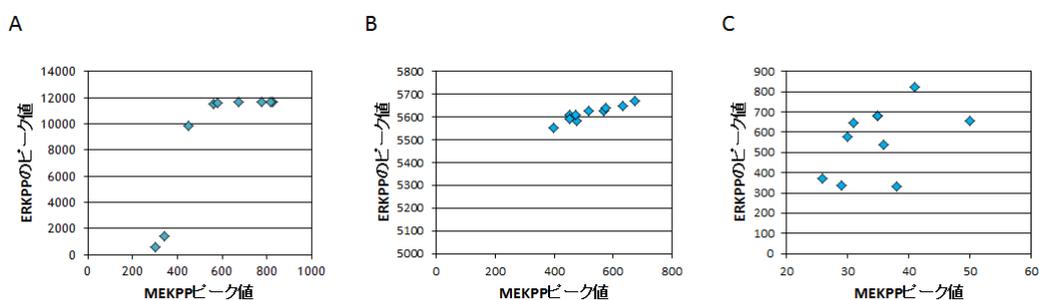
3) プロトタイプモデルのシミュレーション結果

モデル 1(A)、モデル 2(B)、モデル 3(C) の ERK の時間変化を図に示した。モデル 1 と 3 では、ERK のピーク値の変化が、モデル 2 では、ERK の減少する時刻の変化がそれぞれ見られた。実験的には、ERK のピーク値の不均一性が観察されており[2]、これを定性的に再現する結果と

なった。



次に、モデル 1(A)、モデル 2(B)、モデル 3(C)の ERK の応答特性と、反応経路直上の活性型 MEK のピーク値との関係を図に示す。反応経路最終段にあたるこの 2 重リン酸化反応においては、モデル 1 は超感受性を、モデル 2 と 3 は線形性を示し、この部分の応答特性が系全体の応答を決めていることがわかる。



ERK の 2 重リン酸化サイクルは、従来の数理解析では双安定性や超感受性を示すとされてきた [3, 9] が、近年の実験による線形的な応答の報告と矛盾していた [6]。線形的応答の理由の一つの候補としては分子混雑による反応機構の分配的 2 重リン酸化から連続的 2 重リン酸化への変化が提唱されている [6]。今回の計算は、受容体から ERK までの一連の反応経路全体を通じてどのようにシグナルが伝達されるかを 1 分子粒度で計算した初めてのものと思われるが、計算規模を限定したものであった。今後のプロジェクト進展により、より現実に近い規模での計算を実現したい。

参考文献

- [1] Santos *et al.*, *Nat. Cell Biol.* 9, 324, 2007
- [2] Saidon *et al.*, *Molecular Cell* 36, 885, 2009
- [3] Ferrell, Jr. *et al.*, *Science* 280, 895, 1998
- [4] Hibino *et al.*, *Biophys. J.*, 2009
- [5] Takahashi *et al.*, *PNAS*, 107, 2413, 2010
- [6] Aoki *et al.*, *PNAS* 108, 12675, 2011
- [7] Hornberg *et al.*, *Oncogene* 24 5533, 2005
- [8] Fujioka *et al.*, *JBC* 281, 8917, 2005
- [9] Markevich *et al.*, *JCB* 164, 353, 2004

I-3 太田 元規 (名古屋大学)

分子混雑によるタンパク質の安定性変化などのシミュレーションを行うための解析手法の開発

I-3-1 実施計画

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のシミュレーション」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、「分子混雑によるタンパク質の安定性変化などのシミュレーションを行うための解析手法の開発」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 23 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に着手する。具体的には 2 つの立体構造を比較し、その違いを階層的に表現するモーションツリー法を軌道の比較に利用することを検討する。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。

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) 配布用モーションツリー (MT) 描画ソフトウェアの開発

モーションツリー法 (MT 法) は同じタンパク質の異なる立体構造を 2 つ読み込み、2 つの構造間で発生する立体構造変化をツリーとして記述する。ツリーの葉の部分には各アミノ酸部位を表している。ツリーは階層的クラスタリングの結果となっており、どのようなアミノ酸部位のグループが、一緒に固まりとして運動しているのかが、ツリーから理解される。MT 法は当初、立体構造比較の方法として考案されたが、これを軌道解析にも応用するためには、広くユーザに利用してもらうことも大切である。そこで配布版 MT 描画ソフトウェアを開発することにした。このソフトウェアは C 言語でかかれており、Linux、MacOS でコンパイルが可能である。

利用法

```
% mtntr PDBcode1 ChainID1 PDBcode2 ChainID2 (return)
```

ここで、mtntr がプログラム、PDBcode1 は構造 1 の PDB 形式のファイル、ChainID1 は構造 1 のチェーン ID、PDBcode2 は構造 2 の PDB 形式のファイル、ChainID2 は構造 2 のチェーン ID。

出力として MT*.pdb、MT*.ras、MT.info、MT.ps が得られる。MT*.pdb は構造 1 と構造 2 を重ね合わせた PDB ファイル。*には数字が入る。数字は大きな構造変化の数を表す有効ノードの数にあわせて設定される。有効ノードの数が増えると出力される PDB ファイルの数も増える。MT*.ras は MT*.pdb を rasmol で描画するための情報ファイルで、ras ファイルを利用することで構造変化が発生している部分構造が色づけされる。MT.info には有効ノードとノードの高さの情報が出力されている。MT.ps にはツリーの絵がポストスクリプト形式で出力されている。図 1 にモーションツリーの例を示す。

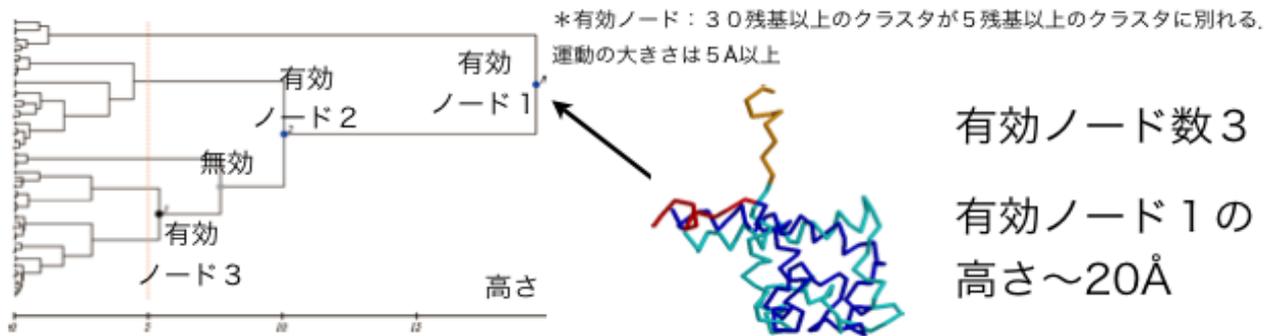


図1: 配布用 MT の出力例。この構造比較では3つの有効ノードが確認された。左図は MT. ps の結果。右の構造はノード1に対応する構造変化で、MT01.pdb と MT01.ras の出力結果を利用し描画した。赤とオレンジの部分が構造変化を表している。

2) MT 法のカルシウムポンプへの適用の実施

作成したソフトウェアを杉田グループに評価していただき、何度か改良を重ねた結果、あるていど満足のいくものになった(当初は計算速度が遅かったので高速化をおこなった)。これを杉田グループで保持しているカルシウムポンプの構造変化解析に適用した。カルシウムポンプについてはこれまでにいくつかの異なる状態で立体構造が決定されている。今回は E2 状態(リン酸とも Ca とも結合していない中間体)から分子動力学計算を行って求めた軌道をクラスタリングし、おおよそ4状態に分離されたスナップショットの代表構造

を利用した構造比較を行った。まずカルシウムポンプの立体構造の概略を図2に示す。カルシウムポンプは N、P、A と名付けられた3つの水溶性ドメインと膜貫通ヘリックスから形成されている。この構造(E2)から開始した分子動力学計算の軌道を図3に示す。この図は、得られた軌道のスナップショットを利用して主成分解析を行い、その第1主成分に軌道を射影したものである。各々のクラスタの中心付近の構造と、X線結晶構造解析で求められた E2 構造を比較したところ、図3下図の

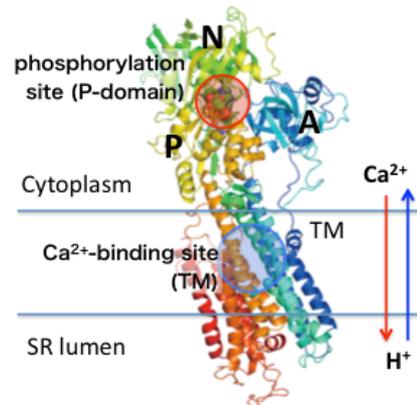


図2: カルシウムポンプのドメイン構造と膜貫通ヘリックス。図上部に水溶性の N、P、A ドメインが配置されている。

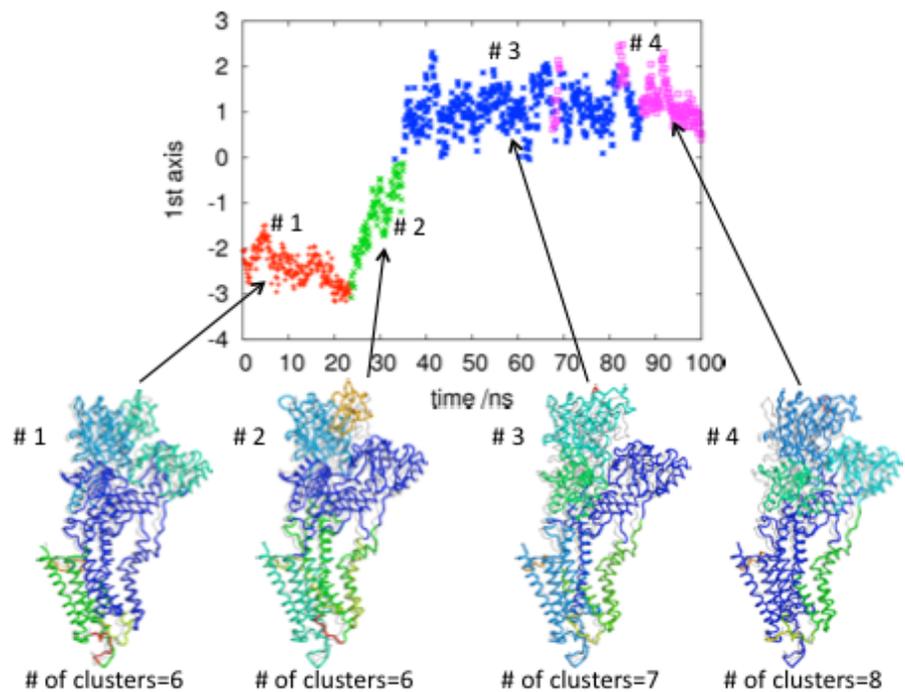


図3: カルシウムポンプの分子動力学計算の軌道を、スナップショットから求めた第1主成分に射影した時間発展プロフィール(上図)。軌道はおおよそ4つのクラスタに分割された。

ようなリジッドドメインが同定された。これらはおよそ、N、P、Aドメインや膜貫通ヘリックス部位、もしくはそれらの融合体に対応しており、分子の運動が既知ドメインを単位として行っていることが確認された。また、水溶性ドメインなどの同定は研究者の目視に基づくものであったことを考えると、MT法のドメイン分割法は人の感覚に沿ったものであることも確認された。

(2) 研究開発の実施状況

1) MT法を用いた軌道解析の研究

MT法を分子動力学から生成される軌道解析に利用するにあたり、対象とするサンプルとして、改変されたリプレッサのフォールディング軌道を用いることとした。以前、東工大TSUBAMEにおける伸びた状態からのフォールディング計算で、 $C\alpha$ 間のRMSD値が4.7Åほどまでフォールドする軌道を得ている(図4)。これを解析することとした。フォールディングにかかわるイベントが時間発展のどの段階に現れるのかを簡便に見つけることが目的となる。

2) MT法を用いた軌道解析の実施

時間発展は20psごとに記録されている。つまり、軌道は20psごとの連続した構造スナップショットになる。構造変化は1つ前の記録、つまり20ps前の構造と比較することにした。MTでは様々な構造情報を得ることができるが、構造変化が激しくおきているのか、あまりおきていないのかを見るための非常に初歩的なインデクスとして、有効

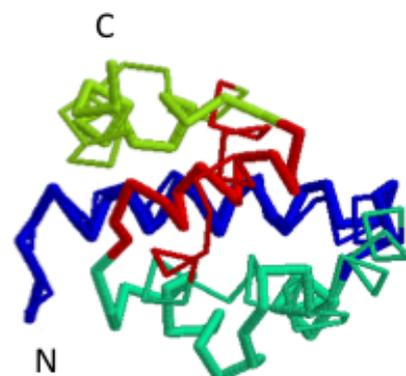


図4: 改変されたリプレッサの天然構造(太線)とシミュレーションで得られた構造(細線)。MT法で構造比較すると4ドメインに分割された。

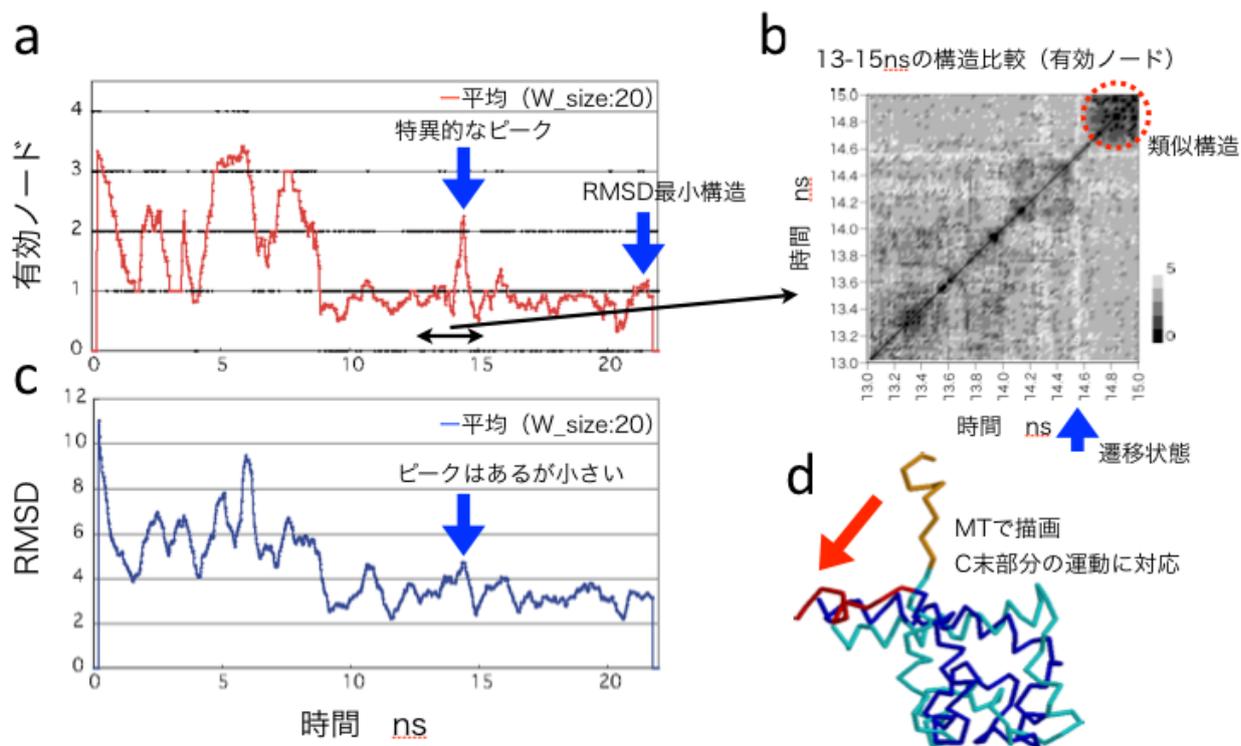


図5: MT法によるリプレッサの軌道解析。(a)平均有効ノードの時間変化。(b)ピーク前後の構造比較。(c)平均RMSDの時間変化。(d)ピーク前後で見られた構造変化。

ノードの数を利用することにした。有効ノードとは、30 残基以上のクラスタが双方 5 残基以上の 2 つのクラスタに分割するノード、ツリーの高さが 5Å 以上のノード、という 2 つの条件が満たされた時に採用される。有効ノードが少なければ単純な構造変化がおきていることを示し、有効ノードが多ければ複雑な構造変化がおきていることを示す。有効ノード数の時間変化は激しく変化しており見にくかったので、前後総計 20 の数値 (400ps に相当) で平均することにした。図 5a が平均有効ノードの時間変化である。これを見ると 14.5ns のあたりで特徴的なピークが観測されていた。つまり、このあたりで複雑な構造変化がおきていると推測された。13ns から 15ns の構造を利用して、all vs all の構造比較を実施した結果を図 5b に示す。図 5b は有効ノードの数がプロットされており、黒いセルは構造変化がないこと (有効ノード 0)、白いセルは構造変化が激しいこと (有効ノード 5) を表している。丁度青い矢印で示したあたりがピークに対応するが、その後に生成された構造に対応するセルは黒くなっており、これらは類似構造であることがわかる。青い矢印あたりに該当する箇所は縦横に白いセルが並んでおり、この構造はどの構造とも類似していない。この構造は遷移状態に相当するものと考えられる。同様の解析を RMSD で行った結果を図 5c に示す。14.5ns あたりにピークは観測されるが、図 5a より小さいピークである。実際にどういう構造変化がこのピークに該当するのかを調べたところ、C 末部分の運動に対応することがわかった。この C 末の構造変化により、この軌道が保持する天然状態と類似の構造 (図 5a の右の青矢印に相当) に転移していた。つまり、MT 法で得られたピークはフォールディングの最終課程でおきる構造変化を捉えていたと考えられる。このピークは図 5c では顕著でなかったことを考えると、MT 法は RMSD よりも構造変化として重要な運動に敏感であることが示唆された。

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

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のシミュレーション」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23年度は、膜を介した物質輸送に関わる輸送体の中から、特にマグネシウムイオン輸送体 MgtE を取り上げ、分子動力学シミュレーションを行う。現在までに、結晶構造から得られた(イオンを透過しない)閉状態構造を基にしたシミュレーションをすでに開始している。まず、そこから人為的に力がかかることでモデリングした開状態のシミュレーションを平成23年度内に完了する。そしてさらに、力をかけずにノンバイアスで開状態を得るための、マイクロ秒オーダーのシミュレーションを年度内に着手することを目指す。他の輸送体に関しては、現在、委託先研究室で解析を完了しつつある医薬学から見ても有用性が高い輸送体について、平成23年度内にシミュレーションを開始し、物質輸送の動的メカニズム解明を目指す。

I-4-2 実施内容 (成果)

(1) 研究開発の実施状況

マグネシウムイオン輸送体 MgtE の輸送メカニズムに関する研究の実施

実施計画にあるとおり、マグネシウムイオン輸送体 MgtE を取り上げ、分子動力学シミュレーションを行っている。平成23年度試験研究では、バイアスがない状態下での MD シミュレーションを行い、自発的な開状態の形成を目指した。予備的な計算として RICC を用い、野生型 MgtE について数百ナノ秒のシミュレーションを行った。今後、さらに長時間のシミュレーションを、京コンピュータを使用して行うことを予定している。

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プロトコルの確立

I-5-1 実施計画

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ヒストンの修飾、ヒストン変異体との置換及び DNA の修飾が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、遺伝子制御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プロトコルの確立」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 23 年度は、ヒストン修飾、ヒストン変異体および DNA 修飾がヌクレオソームポジション変化にどのような影響を与えるのか、それを定量的に明らかにするための計算基盤を構築するために、反応座標の取り方や効率的なサンプリング条件の検討を実施し、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プロトコルの確立を行う。

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

1) SCUBA の開発

これまで SCUBA は MPI 並列化を行っていたが、京コンピュータのマルチコア及び超並列化に対応するために、ハイブリッド並列化を行った。問題の系のサイズを固定して使用するノード数を増やすと、図に示すように計算時間が短縮され、3,792 ノードまで高い並列化効率を得ることができた。その結果、3,840 ノードまでノード拡大が許可された。

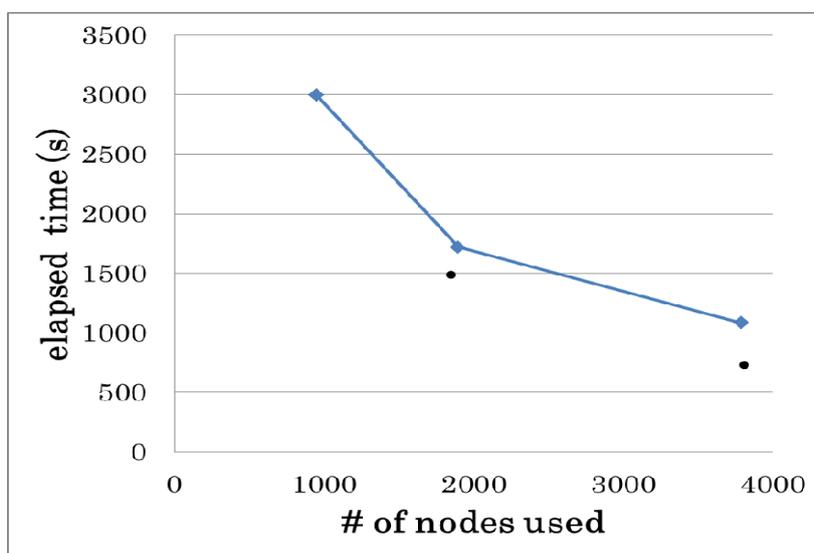


図 1. ノー

ド数) と計算時間。黒丸が理想的な計算時間。

ド数 (x 8 = core

した。結果、各 bin 内に少なくとも 4×10^5 の構造サンプル数がないと平均力が収束しないことがわかった。図 4 に収束の様子を示す。色は異なる bin を示し、横軸の構造サンプル数に応じて平均力が収束していく(図ではフラットになっていく)のがわかる。時間刻みを 1fs とすると、各 bin に最低でも 400ps の計算が必要になる。つまり、1 Å 刻みの window あたり 4ns である。仮に、反応座標の範囲が 100Å であるとすると、総計で 400ns のシミュレーションをすることになる。しかし、各 window での計算は完全に独立に行えるために、「京」では各 window の計算を異なるノードに割り当てて同時に処理する。従って、4ns のシミュレーションの計算時間ですべての計算を終えることができる。このようなシミュレーションは「京」に適した計算である。

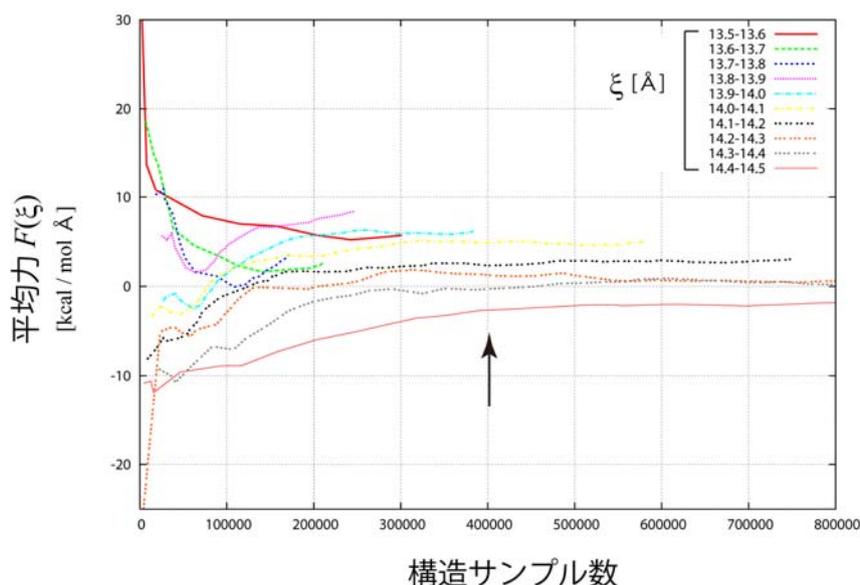


図 4. 構造サンプル数に対する平均力の収束の様子。色(線)は異なる bin を表す。矢印のところから平均力が収束しているのがわかる。

2) ヒストン H3 テールの構造探索

ヒストンタンパク質の N 末端の 30 アミノ酸程度の領域(ヒストンテールと呼ばれる)は、化学修飾を受け状態が変化することにより、ヌクレオソーム機能の発現制御に重要な役割を果たしている。しかしながら、この領域は非常に構造的な揺らぎが大きくその立体構造が明らかになっていない(図 5)。これまで、ヒストンテール部分のみのシミュレーション計算は行われているが[3]、本研究では DNA 存在下、つまり、ヌクレオソーム構造中における非修飾のヒストンテールの構造探索を開始した。それにより、DNA がテール構造に及ぼす影響、テール構造の揺らぎなどを調べることができる。複数あるヒストンテールのうち、H3 のテールの化学修飾が最もよく実験的に研究されているため、まず H3 テール自体の構造探索計算を開始した。

計算では、図 6(a)に示すようにピンクで示した球内の分子を陽に扱い、それ以外の水色部分は陰に扱うこととし、大幅な計算時間の短縮を図っている。陰に扱う部分が陽の領域に及ぼす平均的な静電相互作用を見積もるため、陰領域の計算をあらかじめ行っておく。その時、図 6(b)に示すようにセル多重極法を用いて、各セルの 4 重極子まで計算しておく。テールの構造探索の際には、各セルからの寄与を加味してシミュレーション計算を行うことで、ヌクレオソーム構造下でのテールの構造探索を行っている。次年度は、非修飾、修飾でのテールの構造空間の違いを調べる。

参考文献

- [3] Potoyan, D. A. and Papoian, G. A. (2011), J. Am. Chem. Soc. **133**, 7405.

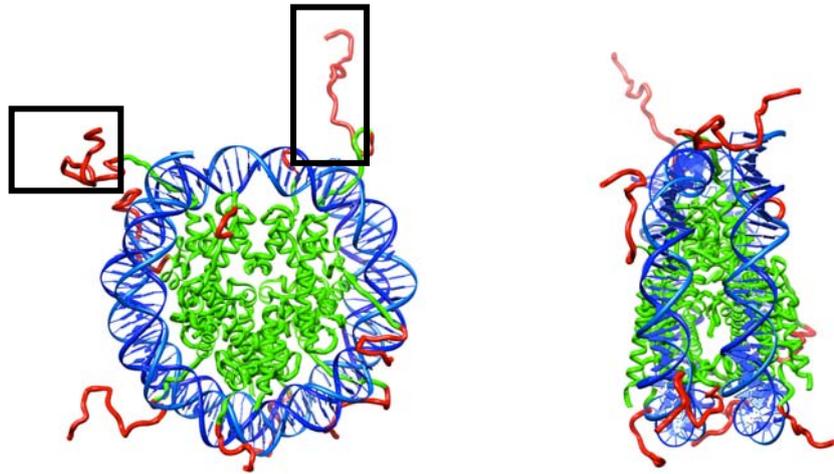


図 5. ヌクレオソームの構造 (PDB code: 1kx5) 青が DNA、緑がヒストン、赤がヒストンテール。赤の部分のテール構造はモデリングされた構造。太線で囲ったところが H3 のテール。

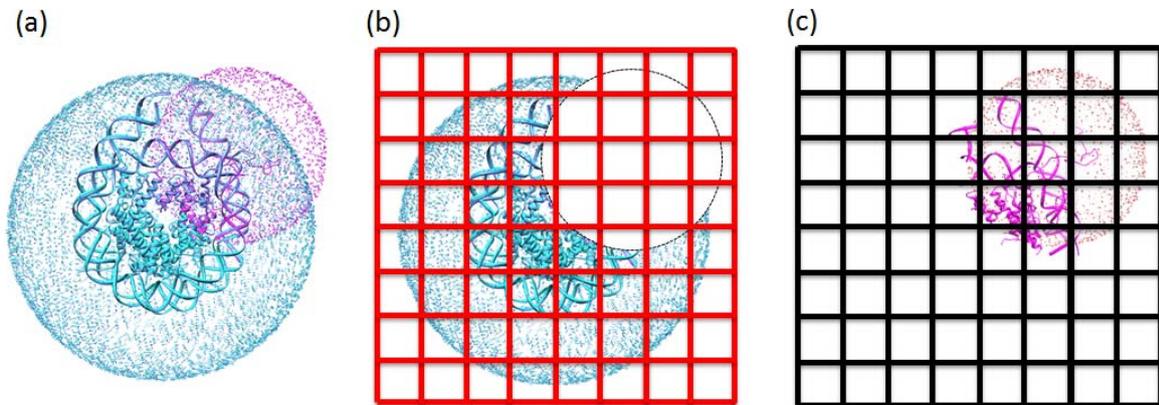


図 6. ヌクレオソーム計算の系。(a) H3 テールを含むピンク色部分は陽に扱って計算を行う。(b) セル多重極法を用いて、水色部分で示された陰に扱う部分の平均的な静電相互作用をあらかじめ計算しておく。図には示していないが、各セルはさらに分割し、近距離の相互作用を精度よく計算している。(c) (b) の結果を陰に取り込み、(c) の系で H3 テールの構造探索を行う。