

課題名 (タイトル) :

高性能生化学ネットワークシミュレーションの研究

利用者氏名 : ○高橋恒一, 海津一成, 小泉守義, サティア・アルジュナン, 宮内敦  
所属 : 神戸研究所 生命システム研究センター 生命モデリングコア  
生化学シミュレーション研究チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

近年の分子生物学や生化学の発展により、生化学ネットワークの構成分子や制御構造に関する多くの知見がもたらされ、計算機上に生命システムの再構成を行う「細胞シミュレーション」が可能となった。

これまでの細胞シミュレーションは細胞を試験管内のような攪拌系として捉え、その内部での生化学反応を酵素反応速度論に基づいて表現する方法が主流であった。

しかしながら、実際の細胞内は不均一であり、なおかつ数 mM の超高濃度で巨大分子がひしめきあう所謂「分子混雑」の状態にある。さらに細胞骨格や染色体などより大きな構造体も多数存在する。

こうした細胞内環境が生命システム、とりわけ信号伝達系に及ぼしている影響について理解することは、それを制御する上で非常に重要である。

我々は、分子の局在や細胞内環境を考慮した 1 分子粒度の生化学シミュレーションによって、より現実的な細胞内反応系の再現を行う。これにより、これまで指摘されなかった細胞内特有の新たな性質を導き出すことを目指している。

この粒子反応拡散計算は、従来の手法の延長では解決することの難しい多くの課題を含んでおり、その計算量も膨大である。従って、大規模な並列化を前提とした高精度・高性能な技術の開発が必要不可欠であり、その技術的な革新も重要な研究テーマである。

2. 具体的な利用内容、計算方法

空間や各分子のゆらぎを考慮した 1 分子粒度生化学シミュレーション手法として、本研究チームでは、大きくわけて「eGFRD 法」と「微視格子法」と呼ばれる二つの技法を用いている。

これらの手法は、どちらも反応を各分子の拡散運動から計算し、その粒子性をも考慮しており、前述の空間的な局在や細胞内環境を適切に再現することが可能である。

以下に、まず各手法の概要と大規模並列計算機との関係性について述べ、続いてこれらの手法を用いて扱う対象について記した。

2. 1. eGFRD 法

enhanced Greens Function Dynamics (以下、eGFRD) 法は、非常に正確 (高時空間解像度) かつ超高性能の粒子反応拡散計算手法である。

当手法は旧来用いられてきたブラウン動力学法と GFRD 法、Gillespie 法を組み合わせた手法であり、離散イベントシミュレーションとして逐次的に計算する。これにより各々の欠点を補いつつ、各分子の状況に応じて最も適切な手法を動的に適用し、その計算量を大幅に削減することが可能となった。

eGFRD 法では他の反応拡散手法ではおざなりにしがちな 0.1 ナノ秒から 1 マイクロ秒の時間スケールまで正確に計算することができ、実際にそれによって生じる新たな反応特性について明らかにされている (Takahashi ら、2010)。

一方でその正確さ故に、扱う分子数・濃度に応じて 1 度の計算に長時間を必要とし、さらに eGFRD 法は確率論的な手法であることから、統計的に十分な試行回数

を得るためには大規模並列計算機の利用が必要となる。

## 2. 2. 微視格子法

微視格子法は、本研究チームのサティア・アルジュナンによって開発された手法であり、eGFRD 法と比べ、短い時間スケールにおいては正確さにおいて劣るものの、多数の粒子や高濃度のシミュレーションにおいて高速に計算を行うことができる (Arjunan ら、2010)。

微視格子法では、空間を分子の直径スケール（一般に数ナノメートル）の六方最密格子に分割し、各分子の動きを格子上の移動として捉えることで計算を行う。空間の表現法や計算手法が直観的に理解しやすいこともあり、複雑な形状の膜や細胞内骨格などの細胞内環境を表現する場合にも柔軟に対応することができる。

ヒトなどの真核細胞の大きさは数十マイクロメートル程度に及び、信号伝達系に関わる分子の数も非常に多い。従って、こうした複雑かつ 1 千万粒子を超える超多体モデルを扱う上で、微視格子法は非常に強力な手法である。

とはいえ、多細胞や複雑な細胞構造をもつモデルを扱うためには膨大なメモリを確保することが必要であり、また計算量も増加する。微視格子法では、1つのモデルを空間的に分割し、各 CPU コアに割り振る並列計算アルゴリズムの開発を行うことで、従来困難であった 1 細胞まるごとの 1 分子粒度シミュレーションを大規模並列計算機上で実現することを目指している。

## 2. 3. 本研究の対象とする生命システム

1 分子粒度シミュレーションが対象とする生命システムは信号伝達系から遺伝子発現系まで多岐に及ぶが、本報告書では特に信号伝達系における「ゆらぎ」と「分子混雑」を対象とした。

信号伝達系は、細胞が細胞外の情報を自身に反映するための一連の反応であり、原核生物から真核細胞まで共通してみられ、癌や免疫、分化までその応用面においても非常に重要である。

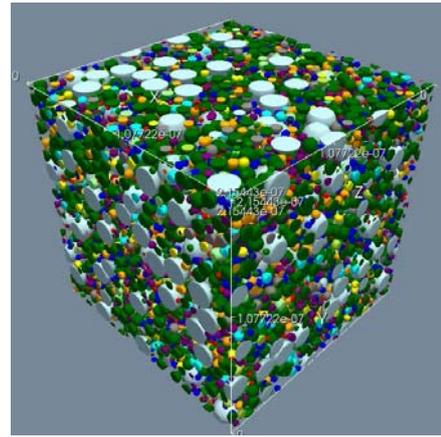


図 1. 原核細胞質中を模した分子混雑の様子

細胞の信号伝達過程において「ゆらぎ」は排除不可能な要素であり、近年では細胞がゆらぎを積極的に利用している例も挙げられるようになった。本研究では、1 分子粒度の反応拡散系におけるゆらぎの特性や影響について考察する。

さらに、分子混雑や異常拡散の生物学的な効果やその実態について明らかにすることを目的に細胞核内の分子拡散動態を再現した。これにより、細胞核内の染色体構造について示唆を与えると同時に、遺伝子発現制御という信号伝達系の最終的な出力につながる重要な要素についても明らかにできる。

## 3. 結果

### 3. 1. eGFRD 法の C++による実装 (小泉)

現在、eGFRD 法と微視格子法のプログラムは共に、スクリプト言語 Python 及び GNU C++によって実装されており、GNU Scientific Library、Boost などのライブラリを用いている。これにより、計算速度を犠牲にすることなくスクリプト言語の簡便さを活用できる。

eGFRD 法に関して、データ構造やグリーン関数などについては既に C++で実装されていたが、他の部分については Python で実装されていた。今回、計算をより高速化するために、これまで Python で書かれていたブラウン動力学法の処理を C++によって実装した。

ブラウン動力学法は eGFRD 法の中で繰り返し呼ばれ、実行速度のボトルネックとなっていることが予測されており、特に分子混雑下のシミュレーションを行う上

では他の処理と比べてことさら支配的になる。

ブラウン動力学法で特に高速化が見込まれる各時間ステップ処理を BDPropagator と呼ばれるクラスにまとめ、C++によって実装した。このクラスは eGFRD 法の演算中に利用されるだけでなく、ブラウン動力学法単体で用いる場合にも効果を発揮する。ブラウン動力学法は次項に示されるような全ての領域で高い濃度をもつ分子混雑のシミュレーションを行う際に用いられる。

現在は RICC 等を用いて正確なベンチマークやプロファイルをとるとともに、eGFRD 法の他の部分についても C++化によって利便性を損なわずに高速化できるよう検討している。

### 3. 2. 染色体構造と分子混雑による細胞核内分子の拡散動態の予測 (高橋・海津)

前述の通り、試験管内 (*in vitro*) と細胞内 (*in vivo*) の最も大きな違いのひとつが「分子混雑」である。細胞質中ではリボソームに代表される巨大分子がひしめきあっており、拡散や反応に試験管内とは異なる特性が表れることが比較的古くから指摘されてきた。しかしながらその実態には明らかになっていない点が多い。

中でも細胞核内、特に染色体領域は染色体を構造化していると言われるヌクレオソームによって分子混雑状態にあり、さらに細胞分裂期によってその状態は大きく変化すると考えられてきた。

また、染色体領域での分子の拡散動態は転写因子やヒストン・クロマチン構造による遺伝子発現制御について考える上で非常に重要な問題である。

そこで我々は近年の 1 分子観察技術による染色体領域における分子の拡散動態とシミュレーションによる予測を比較することで、1 分子観察では直接観察することの難しい短い時間スケール (数ナノ秒) での拡散動態やヌクレオソームや染色体の構造について明らかにすることを試みた。

具体的には、染色体領域について予想されている幾つかの条件下で様々な大きさの粒子 (追跡分子) を拡散させ、その動態を前項で実装されたブラウン動力学法を用いて計算した。この結果は FCS により示された既

存のデータと比較・検討される。

今回検討したヌクレオソーム (背景分子) の状態は以下の通りである。

- (1) ヌクレオソームが動いていない場合
- (2) ヌクレオソームが他の分子と同様自由に拡散している場合
- (3) ヌクレオソームも拡散するが、その動きが局所に制限されている場合

ヌクレオソームの染色体領域における濃度は、間期と分裂期で 0.1mM、0.5mM 程度とされる。また、追跡分子の大きさは、蛍光たんぱく質 GFP を基準とし、ヌクレオソームより大きい GFP 五量体 (数十ナノメートル) 程度までを検討の対象とした。

まず間期を示すヌクレオソーム 0.1mM 条件下では、(1)、

(2) 共に追跡分子の拡散動態には異常拡散が観察された (次頁、図 2.)。しかしながら、(1) では非常に長い時間スケールにわたって異常拡散が観察されたのに対し、(2) では数十から数百マイクロ秒より長い時間スケールでは通常拡散と同様の動態を示すことを明らかにした。このことは FCS により染色体領域内で拡散を観察したデータを判断する上で非常に有用である。

とりわけ、GFP 五量体に相当するヌクレオソーム程度の大きさを持つ分子については (1) 条件下で拡散が制限されることが予測された。このことは、分裂期においても分子が染色体領域中に浸潤 (拡散) することが *in vivo* で観察されていることと矛盾しており、ヌクレオソームにも何らかの移動 (位置ゆらぎ) が存在することを示唆している。(2) では 0.5mM 濃度下でも自由拡散が観察される。

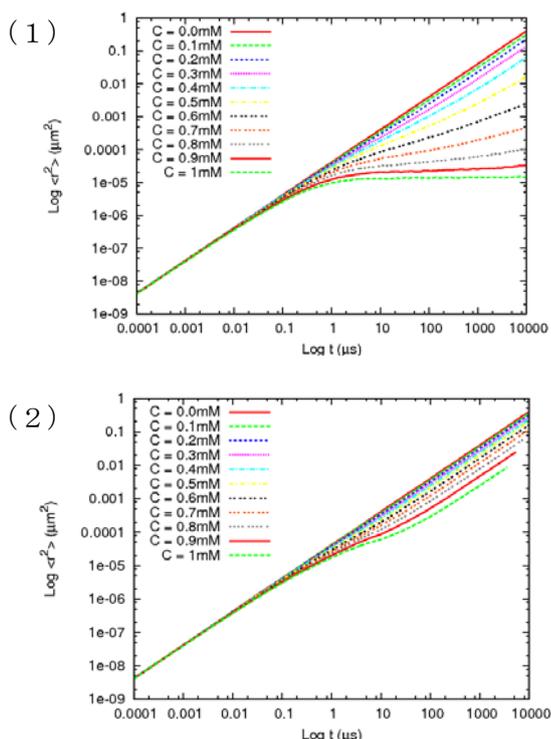


図 2. 各濃度における (1)、(2) 条件下での GFP 五量体分子の拡散動態

しかしながらヌクレオソームは DNA によって繋がれたポリマーのような状態にあり、(2) で想定したような自由な拡散を仮定することは難しい。また、これまでの知見でもヌクレオソーム自体の拡散はほとんど観察されていない。

従って我々は、これまでの観察技術では測定が困難な数マイクロ秒程度の短い時間スケールでのヌクレオソームの運動ゆらぎが染色体領域中の分子の拡散に及ぼす影響について調べるため、(3) の条件を検討した。

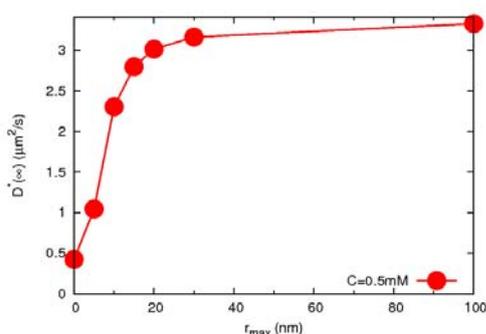


図 3. (3) 条件下ヌクレオソーム濃度 0.5mM における GFP 五量体分子の拡散動態

ヌクレオソームの拡散速度は (2) と同様とした。ヌクレオソームの移動可能な球状領域の半径を R とし場

合の五量体の終端拡散速度を図 3. に示した。半径 R が 0 のときが (1) に相当し、無限大としたとき (2) に等しい。

図 3. に示した結果により GFP 五量体程度の分子の場合、ヌクレオソームが数十ナノメートル程度動くことができれば (2) で示したのと同様の結果が得られることが明らかになった。

ここで仮定した条件の場合、ヌクレオソームの運動は数十マイクロ秒で減衰するため、FCS や 1 分子追跡などでは直接観察することは難しく、これまでの知見とも矛盾しないように思われる。

以上の成果は、国立遺伝学研究所の前島一博博士との共同研究で行われたものである。

### 3. 3. 微視格子法の並列化 (高橋・小泉・アルジュナン・宮内)

微視格子法は、空間 (時間) を 1 分子の粒度で分割することにより、分子混雑条件下でも高効率な 1 分子粒度計算を実現する。

これを拡張し、空間を大きく分割し、それぞれを各 CPU に割り振った並列計算により、単一 CPU での扱うことの可能なモデルよりもより大きなモデル (例えば分子混雑や高等真核細胞まるごと 1 つ) を扱うことを可能にする手法の開発を行っている。

現在、そのプロトタイプが完成しつつあり、数千コアまでの並列化によって実行性能が得られることを確認している。

### 4. まとめ

本研究では 1 分子粒度シミュレーション技法である eGFRD 法と微視格子法の開発を行うとともに、分子混雑条件下での分子の拡散動態について検討を行った。

eGFRD 法、微視格子法共にまだ開発途上の技術ではあるが、本年度において RICC を用いてその実行に必要な要件等についてまとめ、実際に利用する環境を整備できた。さらに前項に示したように高速化・並列化による

性能の向上をも実現しつつある。

染色体領域での分子混雑に関する研究では、分子の拡散動態を高精度・高解像度な手法を用いることで予測し、細胞内環境（染色体構造）に関して重要な裏付けを与えた。

また、本研究成果は、実験的に観察の困難なヌクレオソームの局所的なゆらぎが遺伝子発現制御などに重要な意味を持つことをも示唆している。

## 5. 今後の計画・展望

各シミュレーション技法の開発については、まず本年度行った実装や拡張による性能を確認するため、RICC 上での本格的な使用を行うとともに、その結果を基にさらなる拡張を行っていく予定である。

eGFRD 法については今回のブラウン動力学法と同様に C++化を推し進める。微視格子法についてはその並列化はまだまだ初期段階にあたり、これからの RICC 上での本格的な検討が期待される。また、eGFRD 法やブラウン動力学法についても、微視格子法と同様、大規模並列計算機上での高性能シミュレーションを目指した並列化について検討しており、次年度以降検討していく予定である。

生物学的な成果として、今回は特に拡散動態について特に研究を行ってきたが、同時に分子混雑が反応（信号伝達・遺伝子発現）に与える影響についても通常の計算機上で考察をはじめており、今後は RICC 上で本格的な運用へと移行する予定である。

## 6. RICC の継続利用を希望の場合は、これまで利用した状況（どの程度研究が進んだか、研究においてどこまで計算出来て、何が出来ていないか）や、継続して利用する際に行う具体的な内容

現在、eGFRD 法、微視格子法共に RICC 上での利用環境が整備された段階であり、今後の本格的な運用が期待される。

前項に記したように、次年度以降、今年度行った実装及び拡張に関するデータを RICC 上にて収集し、実際の成果としてまとめていくつもりである。

特に並列化については今後も、より大規模の並列環境においても性能ができるよう拡張していく。

また次年度以降では、同時に、当手法を用いて実際の大規模な細胞モデルを用いて生物学的な研究に着手することを予定している。

分子混雑シミュレーションについては、拡散動態についての詳細な結果が今回得られたため、それを成果としてまとめるとともに、今後は反応を含めたモデルについて拡張し、実際の生命システムの特性的について検討していく。

詳細に関しては前述 3. から 5. を参照のこと。

## 7. 一般利用で演算時間を使い切れなかった理由

研究をシミュレーション技法の開発と同時に進めており、主にシミュレーション環境の整備に時間を費やしていたため。

平成 22 年度 RICC 利用研究成果リスト

**【国際会議、学会などでの口頭発表】**

海津一成、E-Cell 4 を用いた分子粒度計算の実際、E-Cell Workshop 2010 summer、2010 年 8 月 23 日、慶應義塾大学先端生命科学研究所