

課題名 (タイトル) :

細胞内環境における全原子シミュレーション

利用者氏名 : ○原田隆平*, Jaewoon Jung*, 松永康佑*, 宮下尚之*,**, 森貴治*,**, 狩野康人*,**
Michael Feig**

所属 : *計算科学研究機構 粒子系生物物理研究チーム

**生命システム研究センター 分子機能シミュレーション研究チーム

報告内容

近年、X線結晶構造解析や核磁気共鳴分光法(NMR)、一分子測定、in-cell NMR および高密度 NMR などの実験技術の急速な発展により、細胞内における様々な生命現象が、核酸や蛋白質、糖鎖などの生命分子から構成されるネットワークにより制御されていることが明らかになってきた。これらの生命現象において、蛋白質はミリ秒から秒の時間スケールで、自らの構造揺らぎを巧妙に利用することにより、機能発現に関係する大規模な構造変化を引き起こす。一方で、周囲の生体分子との相互作用を通じて、酵素反応や分子認識、物質輸送といったピコ秒からナノ秒の時間スケールにおよぶ生化学反応が、厳密に制御されている。このような生体分子のマルチスケールな生命現象を、原子・分子レベルで解明することは非常に挑戦的で、興味のある研究課題である。

本研究グループでは、新しい分子動力学シミュレーション手法の開発と、その適用として蛋白質のダイナミクスや、機能発現のメカニズム解明を研究目標とする。本年度は特に、生体分子の集合体である細胞を強く意識した超大規模分子動力学シミュレーションを念頭に置き、分子混雑を露わに考慮した全原子分子動力学シミュレーションをはじめとして、レプリカ交換分子動力学法を用いた反応経路探索法および効率的構造空間探索法の開発、データ同化と分子動力学シミュレーションの融合によるタンパク質動態の解明や、マルチスケール分子動力学シミュレーション法の開発などを行ったので、それらの研究成果を中心に報告する。

課題 1 : 細胞内環境における全原子シミュレーション (Feig・原田)

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

細胞環境は混み合っており、ほとんどの生化学実験や生体分子シミュレーションが仮定してきた希薄溶液

環境とは異なる。この事実は、周囲の蛋白質および溶媒分子との相互作用(混み合い効果)について、これまで考慮されてこなかったことを意味しており、現在もなお細胞内の正確な描像を得ることは困難である。

混み合い効果は、当初より剛体球モデルを用いて調べられてきた。この理論モデルによると、混み合い環境では、排除体積効果によって蛋白質間にある小さな分子が排除され蛋白質同士が接近するため、折りたたまれた天然構造が安定となる。しかしながら近年、in-cell NMR や高密度 NMR によって細胞環境における蛋白質のダイナミクスや構造安定性を観察することが可能となり、天然構造が不安定化することが報告され始めた。この実験結果は、自由エネルギーにおけるエントロピー項の寄与に加えて、蛋白質間相互作用に関するエンタルピー項の寄与も重要であることを意味しており、エントロピー項が支配的であると仮定していた剛体球モデルの限界を示唆している。

このような背景から本研究では、蛋白質間相互作用を詳細に考慮すべきであると考え、蛋白質・溶媒を露わに含んだ分子混雑環境を全原子モデリングし、相互作用を分子力場で取り込んだ分子動力学シミュレーションを行った。これにより、分子混雑によって希薄溶液中における熱力学的平衡およびダイナミクスがどのように補正されるべきかを考察した。

2. 具体的な利用内容、計算方法

細胞環境を模したシステムとして、2種類の蛋白質(Protein G および Villin)を用いて、分子混雑をモデリングした。(図 1) また、水分子の数とボックスサイズを変化させることにより、混み合い度合の異なる複数のシステムを用意し、分子混雑環境における溶媒および蛋白質の物理化学的性質を希薄溶液と比較した。

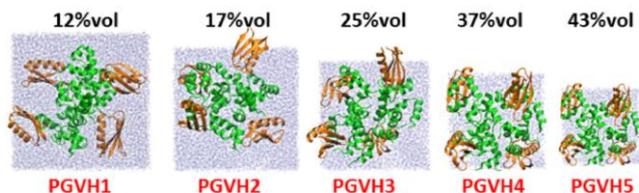


図 1. 2つの蛋白質により構成された分子混雑のシステム。(PGVH1-PGVH5) (Protein G:茶×4+Villin:緑×8) 混み合いの強さは、全体の体積に対する蛋白質の体積比(%vol)で表されている。

3. 結果

分子混雑が蛋白質の構造安定性に与える影響を調べるために、自由エネルギー地形を計算した。

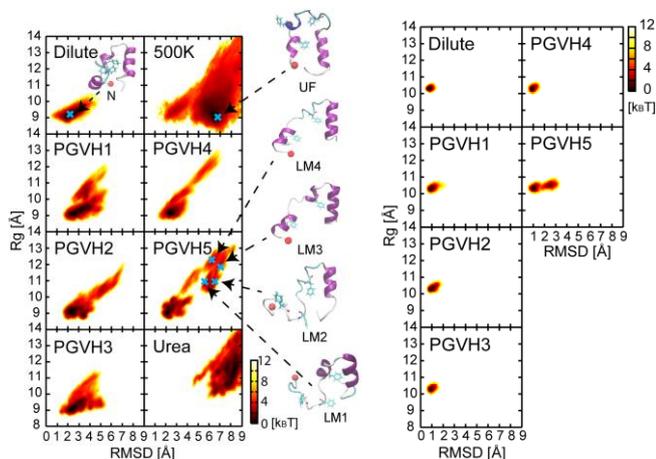


図 2. 分子混雑環境における Protein G(右)と Villin(左)の RMSD(天然構造から平均自乗距離)と Rg(慣性半径)の 2つの反応座標に射影された自由エネルギー地形。最も混み合いが強いシステム(PGVH5)の特徴的な 4つの準安定状態(LM1-LM4)とその構造を共に示す。

計算の結果、Protein Gと比較して Villin が分子混雑によって大きく天然構造が不安定化した。Villin は、全ての混み合い度合(PGVH1-PGVH5)において天然構造が不安定化していることがわかった。これとは対照的に、Protein G は混み合いが最も強いシステム(PGVH5)を除いては、天然構造の安定性が保たれていることがわかった。このように、蛋白質の種類に依存して構造安定性が変化することが観測され、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(500 K)や変性剤(Urea)によるアンフォールディングシミュレーションから計算した自由エネルギー地形とも異なっていることがわかった。この計算結果は、混み合いの強さと共に、分子混雑環境に存在する特有の相互作用が蛋白質の構造安定性を制御しているということを示唆している。図 3 に ProteinG-Villin 間の平均残基距離を示す。Protein G と Villin 間には、4つの特徴的な相互作用が存在しており、混み合いが弱い環境

では静電相互作用が支配的であり(A, B)、混み合いが強くなるにつれて疎水-極性残基間の相互作用(C)、極性-極性残基間の相互作用(D)が支配的になるように、蛋白質間相互作用がシフトしていく過程が観測された。これらの相互作用により、蛋白質の構造安定性が変化し、希薄溶液中では観測されることのない様々な準安定状態(LM1-LM4)が生み出されると考えられる。

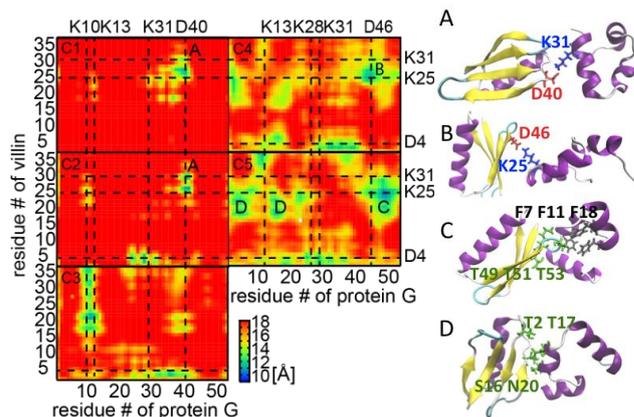


図 3. 分子混雑環境における Protein G - Villin 間の平均残基間距離(左)と 4つの特徴的な 2量体構造(右)。

4. まとめ

分子混雑環境における蛋白質間相互作用をどのように特徴づけていくかは、これからの重要な研究課題である。今後は、蛋白質の種類を考慮して分子混雑環境をモデリングしていく過程で、どのような相互作用が現れてくるのか、また混み合いの強さと相互作用がどのように相関しているのかについて、系統的に解析していく予定である。

5. 今後の計画・展望

今後の研究展望として、これらの分子混雑を露わに考慮した全原子分子動力学シミュレーションの計算結果に基づき、細胞環境で起こる様々な生命現象を予測可能な細胞内分子モデルの構築につなげていく予定である。

課題 2 : Development of improved reaction-path determination scheme in the QM/MM calculations (Jung)

1. Background and purpose of use project, relationship of the project with other projects

To understand enzyme reaction with protein environment, two methods are necessary. One is an accurate QM/MM calculation and the other is efficient reaction path determination scheme. In particular, we need an efficient way to find reaction path with QM/MM scheme which reduces the QM calculation as small as possible. Even though there are several ways for this method, many of them do not use proper target function and efficient parallelization is difficult. In this project therefore, we aimed to develop a new method that is more suitable to find the reaction path with QM/MM scheme.

2. Specific development of calculation method

For enzyme reactions of QM/MM, we introduce three modifications into the locally updated planes scheme (CO-LUP) to overcome these problems : (1) An improved tangent estimation of the reaction path, which is used in the NEB method, (2) Redistribution of images using an energy-weighted interpolation before updating local tangents, and (3) Reduction of the number of constraints, in particular translation/rotation constraints, for improved convergence. This method shows the same accuracy but improved speed over the existed NEB/string method.

3. Result

First, we test the method on the isomerization of alanine dipeptide without QM/MM calculation, showing that the method is more efficient than the existed string method both in accuracy and efficiency (Figure 1). We also apply the method for defining the reaction paths of the rearrangement reaction catalyzed by chorismate mutase (CM), and of the phosphoryl transfer reaction catalyzed by cAMP-dependent protein kinase (PKA) using generalized hybrid orbital QM/MM calculations. The reaction energy barrier of CM is in high agreement with the experimental value. The path of PKA reveals that the enzyme reaction is associative and

there is a late transfer of the substrate proton to Asp 166, which is in agreement with the recently published result using the NEB method (Figure 2).

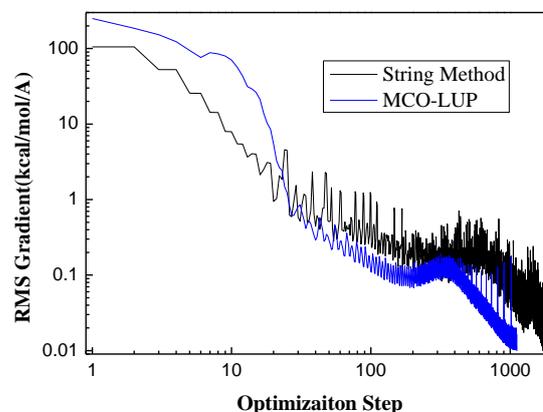


Figure 1

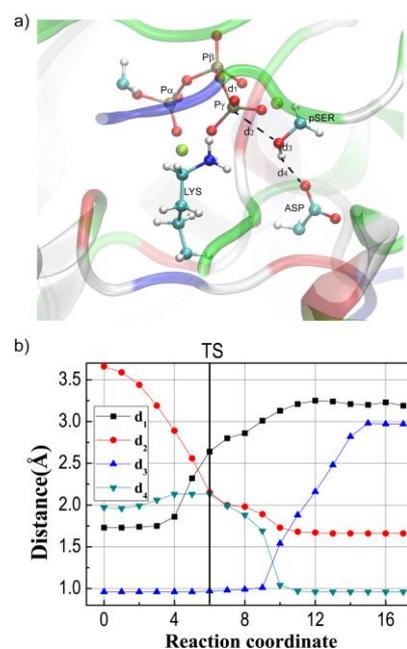


Figure 2

4. Conclusion

We have developed a method for calculating reaction paths. The new method features the advantage of a well-defined target function and enabling efficient optimization for QM/MM systems.

5. Schedule and prospect for the future

The project has focused on reaction paths at zero temperature. However, accurate energetics of enzyme reactions requires the free energy path at a finite temperature. We plan to investigate this using molecular dynamics simulation.

課題 3: パスサンプリング・データ同化によるタンパク質動態の解明 (松永)

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

タンパク質機能にとって重要な立体構造変化はミリ秒〜で起こる遅い過程であり、既存の **Brute force** シミュレーションでこれを再現するのは困難である。そこで我々は、単一のシミュレーションの代わりに、多数のコピー系を構造変化パス上に配置し、それらマルチコピー系を並列に計算することで粗視化空間における最適パスをサンプルする手法の開発に取り組んでいる。今期は、「京」コンピュータの一般利用課題に採択された多剤排出トランスポーター(**AcrB**)の構造変化パスサンプリング課題の事前計算として **RICC** で初期パス作成・平衡化計算などを行った。

一方で、近年の 1 分子計測技術の発達により、こうしたタンパク質構造変化揺らぎを実験で観察することが可能となってきたが、得られるデータは限られた構造部位に関する低次元情報であることが多く、そこから高次元状態空間ないしはネットワークの状態遷移モデルを構築することはデータ数の制限もあり困難であることが多い。そこで我々は、1 分子の **Förster resonance energy transfer (FRET)** 時系列データからシミュレーションによって低次元→多次元の「逆問題」をサンプリングし、限られた実験データから多次元情報に基づいた状態遷移モデルを構築する手法の開発している(逐次データ同化法)。具体的には、粗視化タンパク質モデルを使ったシミュレーションの最中に **FRET** データを逐次的に取り込み軌道を補正してやることで、シミュレーションの力場パラメータやランダムノイズ、時間積分に起因する誤差を修正しながら、実験データを再現する構造の時間発展アンサンブルを得ることを目指している。

2. 具体的な利用内容、計算方法

多剤排出トランスポーターの構造変化パスサンプリングに用いている有限温度ストリング法では、まず 1) 構造変化をよく記述する粗視化空間を定義し、2) 粗視化空間上で構造変化前後を結ぶ適当な初期パスを作成する、プロセスが必要となる。他のタンパク質(アデニル酸キナーゼ)に対する先行研究により、構造変化をよく

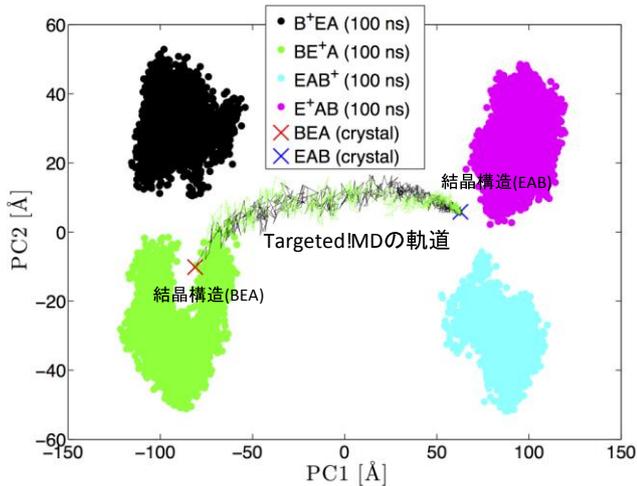
記述する粗視化空間としては、事前に構造変化前後周りの平衡サンプリングシミュレーションを行い、そこから統計解析(主成分解析)により抽出した集団座標を使うことが良いと分かっている。そこで、多剤排出トランスポーターに対して、構造変化前後で各 100 ns の平衡サンプリングシミュレーションを行った。また、初期パス作成には、構造変化方向へ拘束力をかけて構造変化を起こさせる方法(**Targeted MD**)を 10 ns 行った。

逐次データ同化法では、多数のコピー系を同時に流し、逐次的に実験データを取り入れながら尤度に基づいて復元抽出を行っていく。まず前提として、シミュレーションが時間解像度の粗い 1 分子 **FRET** 時系列を追跡しなければならないが、全原子モデルを用いての追跡は困難である。そこで、まず粗視化モデル(**Mixed-Elastic Network Model, Go-like Model**)を実装した。その後で、逐次データ同化のアルゴリズムを実装し、粗視化モデルシミュレーションから作成した疑似 **FRET** データを追跡できるか、他の自由度の振り舞いを推定できるか?を検証した。これには、およそ **RICC512** ノードを用いておよそ 1 千~1 万個のコピー系を使用した。その後、粗視化モデル自由度のみならず、モデルのパラメータも同時に推定できるかを検証した。パラメータ空間は時間発展がないために、復元抽出によるコピー系の退化問題が起こりやすい。そこで統数研の上野らによって提案された **Merging particle filter** 法を導入し、パラメータ推定が改善されるかを検証した。

3. 結果

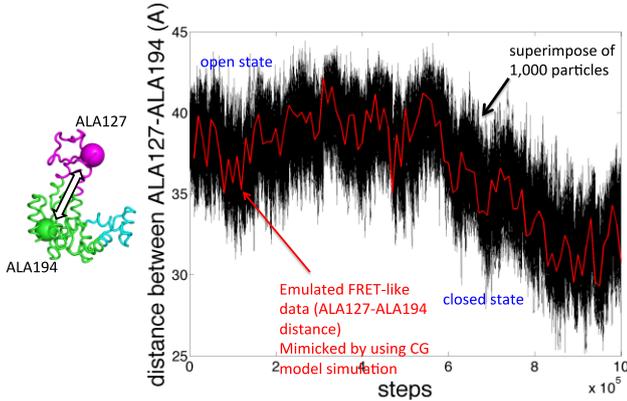
多剤排出トランスポーターの構造変化は、**proton transduction** がエネルギー源となっていると考えられている。そこで、粗視化空間抽出のための平衡サンプリングでは、特定の残基がプロトン化されている・されていない場合の二通りの系のシミュレーションを 100 ns 行った。その結果、プロトン化の有無によって抽出した粗視化空間において顕著な分布の違いがみられた(下図)。これは、粗視化空間がよく構造変化を記述するための証左と思われる。

集団座標上の平衡サンプリングデータと初期パス



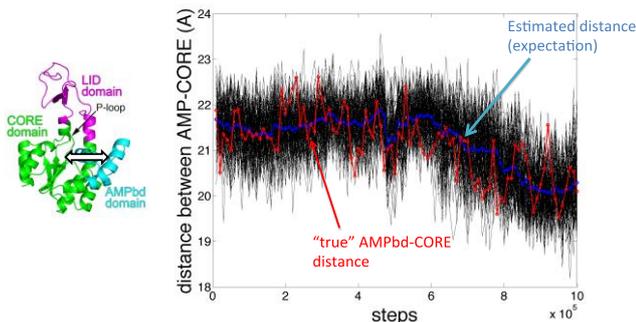
逐次データ同化手法では、実装したアルゴリズムで観測部位以外の自由度のふるまいを推定できるかを検証した。その結果、モデルパラメータによっては、観測部位以外の自由度をよく推定できることがわかった(下図)。一方で、コピー数が少ない際には、復元抽出による退化問題がおこることがわかった。これは特にパラメータ推定の際によく観察された。

Application to an emulated FRET data by using ENM model implemented in GENESIS (1,000 particles)



Inference of AMPbd-CORE domain distance

[模擬実験で見ている距離とは場所が異なるので、推定が難しい問題]



4. まとめ

多剤排出トランスポーターのパスサンプリングを行うために、事前の平衡サンプリング・初期パス作成を行った。途中結果は、抽出した粗視化空間が構造変化をよく記述することを示唆しており、今後「京」コンピュータでのプロダクションランの結果が期待できる。逐次データ同化手法では、粗視化モデルによるアルゴリズムのパフォーマンスを検証した。今後は実際の FRET データへの応用を目指す。

5. 今後の計画・展望

今後は、得られた成果・開発した方法論を積極的に学会誌・学会上で発表していく予定である。また逐次データ同化手法に関しては、今後開発した方法論を実際の FRET データを応用することを目指す。

課題 4：レプリカ交換分子動力学インターフェースプログラムの開発 (宮下)

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

生体分子の構造や構造変化機構、結合機構は細胞活動や疾患機構解明に繋がる重要な情報を与える。これらを予測する方法として、拡張アンサンブル法と呼ばれる方法がある。この方法では、エネルギー空間上のランダムなサンプリングを誘起し、結果として構造サンプリング効率を上げる。また自由エネルギー地形を得て、生体中での生体分子構造やその構造変化、揺らぎ方などの情報を効率的に得ることができる。

レプリカ交換分子動力学法は、拡張アンサンブル法の一つである。この方法は同じシステムでかつ、温度の違う複数のレプリカの分子動力学 (MD) 計算を行い、ある MD ステップ数毎にパラメータ (温度) の交換をメトロポリスクライテリアに従って行う方法である。この方法は温度だけでなく、他のパラメータ交換も可能であり、多次元パラメータ交換もできる (温度と距離パラメータなど)。しかし、多次元にすると非常に多くのレプリカが必要となる。即ち非常に多くのリソースを必要とする。しかし、理研で開発された京コンピュータには非常に多くの node がある為に、そのような計算が可能となる。

我々は、将来的に細胞環境中の膜タンパク質の構造

変化予測や生体膜中のラフトのタンパク質への影響などの研究を生命システム研究センターの研究として行う為に、レプリカ交換分子動力学プログラム:REINを開発した。(K用の開発は次世代計算科学研究プログラムのプロジェクトである)

しかし、京コンピュータは全ての人が使える物ではないし、テスト計算などは通常のスパコンやPCクラスターでできると便利である。REINは、PCクラスターでも動く様に設計されている。本課題では、このREINをRICCでも起動できる様に改良した。

2. 具体的な利用内容、計算方法

REINでは、MD部分は他の既存のプログラムを用いている。REINはそのMDを制御するプログラムとなっている。即ちREINから複数のMDを実行し、その結果から必要な情報を取り出し、レプリカ交換及び次のMDのインプットを作成し、再び複数のMDを実施するスケジューラのようなプログラムである。このMDの実行の制御方法として、京コンピュータではmpi-2の機能である動的プロセス生成の機能を使っている。また、その為に計算機ノード内の共有ディスクを用いるスタイルとなっている。

しかし、RICCでは動的プロセス生成の利用が許されていない。従って、そのままではREINはRICC上では実行できない。そこで、各々のMDはノード内利用に限定し、MPI/OpenMP hybrid形式でかつsystem関数を用いてMDの制御をする様に改良した。また、FTLでは各ノードのlocal discを利用するので、FTLでも使える様に改良した。その際、可搬性を保つ為に全体の構成の大きな変化が無い様に設計した。サポートソフトウェアもRICC用のオプションを作成した。

3. 結果

複数のテストを行い、RICCでも起動する事が確認できた。

4. まとめ

京コンピュータ用に作成しレプリカ交換インターフェースプログラムREINを改良し、RICCで起動する用に改良した。また、RICCで起動の確認を行った。

5. 今後の計画・展望

現在、REINを用いて京コンピュータで実施する計算の準備を行っている。また、そのサポート計算をRICCで実施する予定である。開発としては、GPGPUに対応させる予定である。

課題5: マルチスケール分子動力学シミュレーション法の開発(森)

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

細胞の周囲を取り囲む細胞膜表面では、細胞内外の物質濃度の調整や情報伝達など様々なプロセスが膜タンパク質を介して精密にコントロールされている。細胞膜はリン脂質やスフィンゴ脂質、コレステロールなどの生体分子から構成され、近年、脂質ラフトと呼ばれる細胞膜中のマイクロドメインが細胞プロセスの活性に重要であることも分かってきた。しかしながら、細胞膜は非常に複雑なシステムであるため、X線結晶構造解析などの実験的手法ではその構造を調べるのが難しい。また、シミュレーションを用いてその構造を予測しようとしても、通常の方法では限られた計算時間の中で十分にサンプリングすることは容易ではない。そこで本研究では、生体膜系に対する効率の良いシミュレーション法として「表面張力レプリカ交換法」の開発を目指した。

2. 具体的な利用内容、計算方法

シミュレーションソフトウェアには、本研究室で開発中のGENESISを用いた。GENESISは分子および細胞スケールを対象にした分子動力学シミュレーションソフトウェアで、ハイブリッドMPI/OpenMPで並列化されており、並列環境下においても効率よく計算を実行することができる。本研究では、開発した表面張力レプリカ交換法(図1)をGENESISに導入し、DPPC二重膜系を対象としてシミュレーションを実行した。脂質分子の力場にはCHARMM36(全原子モデル)を用い、温度、圧力コントロールにはLangevin dynamics, 結合拘束にはSHAKE, SETTLE, 非結合相互作用にはLookup table法および、Particle Mesh Ewald法を用いた。

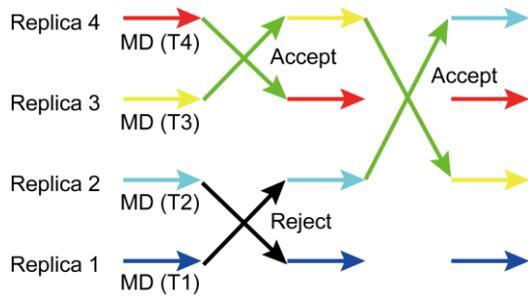


図 1 レプリカ交換法の概要

3. 結果

下図にシミュレーションにより得られたスナップショットを示す。脂質二重膜は壊れることなく急激に水平方向に構造変化していることがわかる。また、シミュレーションにおいて表面張力空間のランダムウォークを実現し、それにより表面積だけでなく膜厚もランダムに時間変化していた。

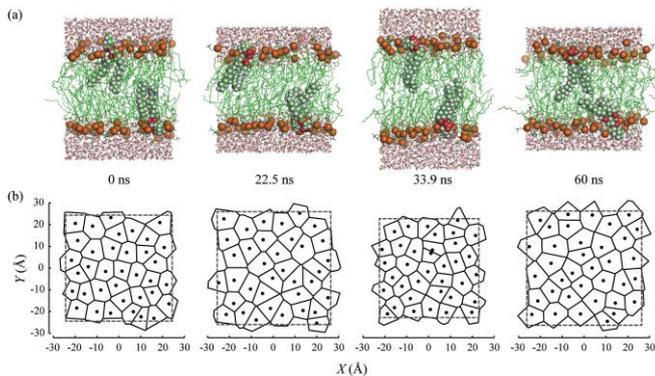


図 2 表面張力レプリカ交換シミュレーションにより得られた DPPC 二重膜のスナップショット

4. まとめ

本研究において表面張力レプリカ交換法の開発に成功した。本手法を DPPC 脂質二重膜系へ適用したところ、膜の急激な構造変化が観察され、生体膜系の効率の良い構造サンプリングが可能であることが実証された（本成果については、現在論文を執筆中である）。

5. 今後の計画・展望

本研究で開発した表面張力レプリカ交換法をより複雑なシステム、例えば膜タンパク質系や混合脂質二重膜系へと適用し、膜を隔てた物質輸送や情報伝達など生体膜中で起こる様々な現象の分子機構を明らかにしていく。

平成 24 年度 RICC 利用研究成果リスト

【論文、学会報告・雑誌などの論文発表】

1. Ryuhei Harada, Naoya Tochio, Takanori Kigawa, Yuji Sugita and Michael Feig, "Reduced native state stability in crowded cellular environment due to protein-protein interactions", *Journal of the American Chemical Society*, DOI: 10.1021/ja3126992
2. 原田隆平、"細胞環境における水和構造と蛋白質の構造安定性"、日本生物物理学会誌「生物物理」新進気鋭シリーズ (印刷中)
3. 原田隆平、博士論文紹介 "超並列マルチスケールシミュレーションMSFEL法で探る生体分子の自由エネルギー地形解析"、分子シミュレーション研究会会誌「アンサンブル」vol.15 No.1 January 2013 page61-65
4. Jaewoon Jung, Suyong Re, Yuji Sugita, and Seiichiro Ten-no, "Improved constrained optimization method for reaction-path determination in the generalized hybrid orbital quantum mechanical/molecular mechanical calculations", *Journal of Chemical Physics*, 138, 044106 (2013)
5. Yasuhiro Matsunaga, Ryotaro Koike, Motonori Ota, Jeremy R.H. Tame, and Akinori Kidera, "Influence of Structural Symmetry on Protein Dynamics", *PLoS ONE*, 7, e50011 (2012).
6. Yasuhiro Matsunaga, Hiroshi Fujisaki, Tohru Terada, Tadaomi Furuta, Kei Moritsugu, and Akinori Kidera, "Minimum Free Energy Path of Ligand-Induced Transition in Adenylate Kinase", *PLoS Computational Biology*, 8, e1002555 (2012).
7. Suyong Re, Wataru Nishima, Naoyuki Miyashita, Yuji Sugita, Conformational flexibility of N-glycans in solution studied by REMD simulations, *Biophysical Review*, 4, pp179-187 (2012)
8. Wataru Nishima, Naoyuki Miyashita, Yoshiki Yamaguchi, Yuji Sugita, Suyong Re, Effect of Bisecting GlcNAc and Core Fucosylation on Conformational Properties of Biantennary Complex-Type N-Glycans in Solution, *Journal of Physical Chemistry*, 116, pp8504-8512 (2012)

【国際会議、学会などでの口頭発表】

1. 原田隆平, 杉田有治, Michael Feig, "Hydration and Protein Stability under Crowded Environments", The 6th Mini-Symposium on Liquid, 2012年6月 九州大学
2. Ryuhei Harada, Yuji Sugita, Michael Feig, "Protein dynamics and stability under crowded environments", 2012年9月 名古屋大学, 第50回日本生物物理学会年会・若手奨励賞選考会
3. 松永康佑 "有限温度ストリング法によるアデニル酸キナーゼの最小自由エネルギー経路探索", シンポジウム「化学反応経路探索のニューフロンティア 2012」, 東京大学, 2012年9月
4. Naoyuki Miyashita, Suyong Re, Yuji Sugita, レプリカ交換インターフェースプログラム(REIN) 3E1110, 第50回生物物理学会年会, 名古屋大学, 2012年9月
5. 宮下尚之, 李秀榮, 杉田有治, Replica-Exchange Interface program (REIN) 20pAB-1, 日本物理学会 2012年 秋季大会, 横浜国立大学, 2012年9月

【その他】

1. Ryuhei Harada, Yuji Sugita, Michael Feig, “Protein stability under cellular environments”, Biophysical Society 57th Annual Meeting, Philadelphia (USA), 2013/2 (Poster presentation)
2. Jaewoon Jung, Suyong Re, Yuji Sugita, Seiichiro Ten-no, “Improved Constrained Optimization Method for Reaction-Path Determination in Quantum Mechanical / Molecular Mechanical Calculations”, Biophysical Society 57th Annual Meeting, Feb. 1-6, Philadelphia, USA (Poster presentation)