

課題名 (タイトル): タンパク質・核酸など生体高分子の分子シミュレーション

利用者氏名: ○森次 圭

理研での所属研究室名: 情報基盤センター・計算工学応用開発ユニット

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

本研究チームは平成 24 年度までの次世代生命体統合シミュレーション研究開発プログラム・分子スケール研究において、タンパク質を中心とした生体分子のシミュレーション法とそのソフトウェアの研究開発 (コード名:  $\mu^2\text{lib}$ , <http://www.mu2lib.org> のサイトで公開) を行った。特に、全原子シミュレーションと疎視化モデルシミュレーションとの連成手法を新規開発することを目指した。プロジェクトでの研究目的は以下の 2 点である:

- ・次世代スーパーコンピュータ「京」の全計算機資源を用いて高効率で計算することができる
- ・それによって従来の分子シミュレーションの方法ではできなかったレベルの計算をすることができる

生命活動をタンパク質や核酸などの生体分子のレベルからシミュレーションによって解こうという分野における問題は、その巨大な系の大きさと生命現象の時間スケールの大きさである。その大きさのために、全原子シミュレーション法には巨大な計算機資源を用いても多くの場合、生命現象の解明が可能な系の大きさと計算時間の長さを実現するシミュレーションは不可能である。そこで不可避免的に疎視化モデルの利用が求められるが、そこには精度の制約が生まれる。従って、その両者の利点を併せ持つ連成計算 (全原子シミュレーション法の精度と疎視化モデルの効率) が必要となる。また、数十万コアという並列計算を実現するためには、不可避免的に弱連成のアルゴリズムであることが要請される。これらを実現するため、一つの応用研究として、利用者らは新規アルゴリズムである **MultiScale Enhanced Sampling (MSES)**法を開発した。

MSES 法は、全原子モデルと低自由度の疎視化モデルによる連成シミュレーションである。疎視化モデルのポテンシャルが規定する運動空間に全原子モデルをドライブし、全原子モデルと疎視化モデルとを接続するバネ強度を 0 に外挿することで、全原子モデルの空間

での分布関数を得ることができる。バネ強度の 0 への外挿は、バネ強度を変数としたハミルトニアンレプリカ交換法によって行う。従って、MSES 法はバネ強度の異なる多数のコピーを用いた弱連成のシミュレーションであり、高度の並列計算が可能である。レプリカ交換が疎視化モデルの自由度により決まることから、通常の温度レプリカ交換法と異なり全原子モデルの自由度の制限なく巨大系のサンプリングが実行可能となる画期的な方法である。昨年度までの研究において、 $\mu^2\text{lib}$  への MSES 法の実装および高度化は完了している。また応用研究として、ミニタンパク質シニョリンのフォールディング過程だけではなく、天然変性タンパク質 sortase の disorder-to-order 転移の構造探索や barnase-barster 複合体、HPr-EIN 複合体の自由エネルギー地形計算による複合体形成過程の理解、また、グルタミン結合タンパク質へのリガンド結合過程の解析といった、従来の拡張アンサンブル法では難しかった大規模系 (溶媒を陽に含んだタンパク質系という意味で) への適用も進めてきた。

2. 具体的な利用内容、計算方法

本研究チームでは、次世代生命体統合シミュレーション研究開発プログラムの一環としてマルチコピー・マルチスケール分子動力学シミュレーションのためのクラスライブラリ ( $\mu^2\text{lib}$ ) の開発を継続している。 $\mu^2\text{lib}$  ではマルチコピーシミュレーションを行っており、異なるパラメータを与えた数十の系のコピー (レプリカ) を発生させ、それらの間の相互作用を考慮しながら並行してシミュレーション実行する。各コピーについて数十のコア、合計数百のコアを用いた並列計算を OpenMP と組み合わせたハイブリッド並列により行なう。今年度は下記のような応用研究を行い、これらの方法の妥当性と物理化学的意味づけを評価した。

3. 結果

(1) グルコキナーゼの全原子構造サンプリング

グルコキナーゼはグルコースへのリン酸化転移を行う酵素であり、糖代謝を制御し全身のグルコース恒常性を保つうえで重要な役割を果たしていることから、糖尿病や低血糖症に向けた創薬ターゲットとなっている。グルコース濃度に対するリン酸化転移活性は **positive cooperativity** を示すことが知られており、これにはグルコース濃度に応じた構造分布の大きな変化が関与すると考えられている。本研究では、MSES 法によりグルコース結合・非結合状態での網羅的構造探索を行い、その自由エネルギー地形を直接計算・比較することで、**positive cooperativity** を示す分子機構を明らかにすることを旨とした。

結晶構造はドメインの大きな開閉運動が界面でのグルコース結合を制御することを示唆しているが、グルコース結合状態ではドメイン closed/open 状態で 2 つの basin が存在し高いエネルギー障壁を持つのに対し、グルコース非結合状態では open 状態を含む幅広い単一の basin 内を運動することが示された (図 1)。つまり、高グルコース濃度では closed/open 状態の構造変化が起こる前に closed basin でグルコースが再結合することで活性サイクルが高速化されるのに対し、低グルコース濃度では closed/open 間の遅い構造遷移が律速になり活性サイクルが遅くなることが分かった。また、結晶構造では見えていない、ドメイン界面にある  $\alpha 13$  ヘリックスの折れ畳みがエネルギー障壁に関わることを示した (図 2)。さらに、ドメイン運動により  $\alpha 13$  ヘリックスの折れ畳み、また、グルコースとの結合安定性が高度に制御されることを明らかにした。

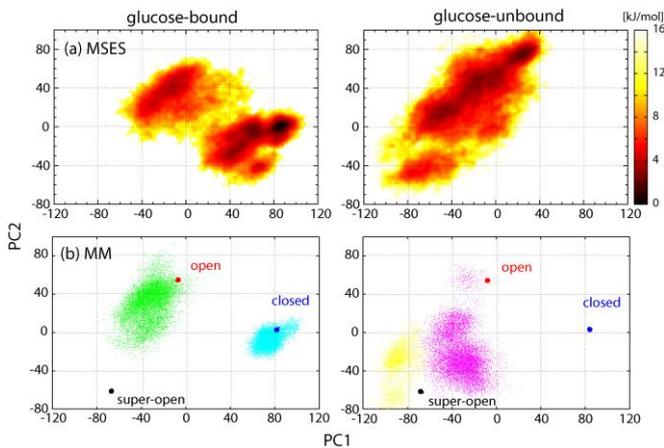


図 1 : (a)MSES、および(b)brute-force MD で得られたグルコキナーゼのドメイン運動に関する自由エネルギー地形。

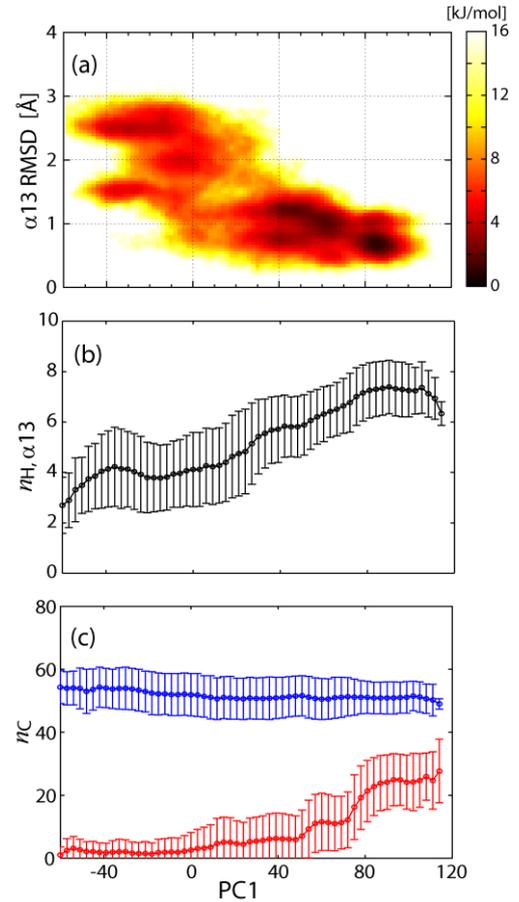


図 2 : (a)  $\alpha 13$  ヘリックスの折れ畳みに関する自由エネルギー地形。(b)ドメイン運動と  $\alpha 13$  ヘリックスの折れ畳み。(c) large(赤)、large(青)ドメインとグルコースとの原子コンタクト。

次に、weighted ensemble (WE)法の適用により、basin間の構造変化の時定数を計算し、**positive cooperativity** を示す **kinetic** な機構について調べた。得られた自由エネルギー地形が、2つの basin 間の構造変化がドメイン運動と  $\alpha 13$  ヘリックスの折れ畳みの 2 つの過程に分割可能であることを示唆するため、両方の過程の WE 計算を別々に行い時定数を計算した (図 3a)。WE 計算ではそれぞれ 22/54 の state に分割したのち、各 state すべてで 16 本の 50-ps MD を実行することを 100 回繰り返し state 間の確率の変化を短い時間で(50 ps x 100 = 5 ns) 計算することで、それぞれの過程での時定数を近似的に見積もった。図 3b のような kinetic なモデルを設定し、WE 計算および MSES 法で得られた自由エネルギー地形が示す状態間の自由エネルギー差により 2 状態間の時定数を計算した ( $1.1 \text{ ms}^{-1}$ ) ところ、グルコースリン酸化転移の反応レートの実験結果 ( $0.22 \text{ ms}^{-1}$ ) と同等であることを示し、グルコキナーゼで **positive cooperativity** が実現される kinetics を実証することができた。

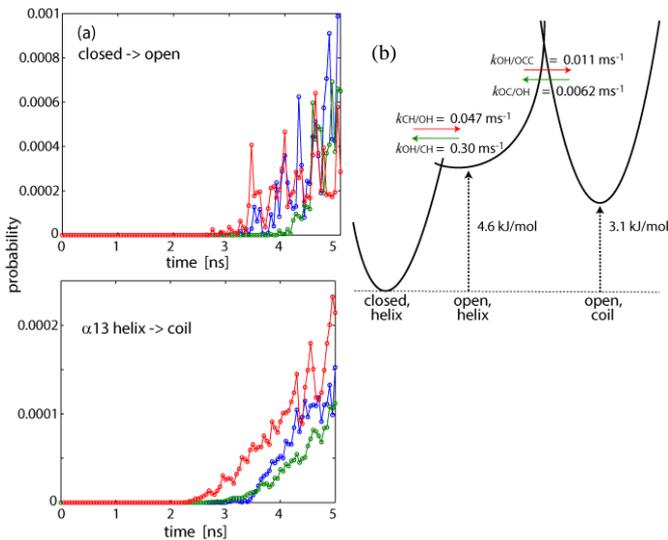


図 3 : (a) weighted ensemble 法で計算されたドメイン運動 (上図) および  $\alpha 13$  ヘリックスの折れ畳みの時定数計算。(b) weighted ensemble 法および自由エネルギー地形から計算した kinetic 定数。

(2) Ras による SOS のアロステリック活性制御の分子基盤のシミュレーション解析

MAPK シグナリングパスウェイにおいて、Ras のヌクレオチド交換反応 (GDP 結合型  $\rightarrow$  GTP 結合型) は重要な要素であり、その反応はグアニンヌクレオチド交換因子 SOS が触媒する。活性化された SOS は、allosteric site に RasGTP (GTP 結合型 Ras) を結合することでさらに活性化し、catalytic site に RasGDP (GDP 結合型 Ras) を結合しそれを RasGTP にする交換反応を起こす。それを実証する in vitro 実験として、allosteric site に Ras がいないとき、および、RasGDP の場合には、SOS が活性化されないという結果がある。そこで本研究では、allosteric site に Ras がいないとき、および、RasGDP/RasGTP の場合の SOS の構造ダイナミクスを 300 ns の brute-force MD シミュレーションにより計算・比較を行うことで、SOS アロステリック活性制御の分子基盤を明らかにすることをめざした。

まず SOS 単体のシミュレーションに対して Motion Tree を計算し、SOS の動的構造を解析したところ、allosteric site での Ras 結合を制御する Rem ドメインと catalytic site での Ras 結合を制御する helical hairpin が協同して動く様子が示され、そのダイナミックな協同性によりアロステリックな SOS の活性制御が実現されることを示した (図 4)。

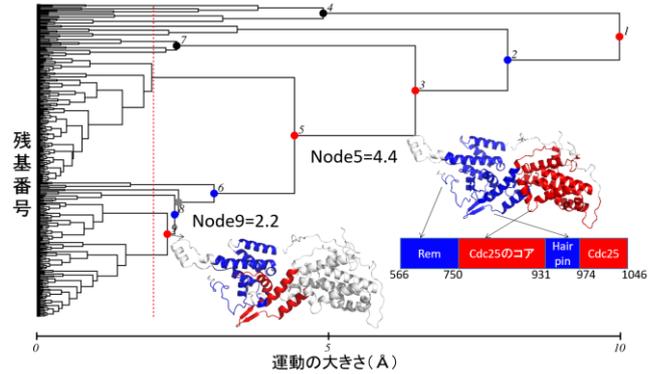


図 4 : SOS 単体シミュレーションから計算された Motion Tree。

次に、Ras のヌクレオチド状態変化に応じて SOS 構造ダイナミクスがどのように応答するかを比較した。Ras 結合がない時には、Cdc25 と Rem のドメイン運動が自由に起こるために open から close まで広範な構造をとった。それに対して、活性型になった SOS は、ゆらぎが close 型に限定されていた (図 5)。catalytic site 非結合型の RasGDP/noRas、RasGTP/noRas を比較すると、実験では allosteric site に結合した Ras が GTP 型の方が GDP 型よりも catalytic site での高い活性を与えるが、RasGTP がより活性型である open 型に寄った構造を取っていることを示しており、実験結果を説明したシミュレーション結果となっている。ここで、RasGTP/RasFree と RasGTP/noRas を比較すると、それらの間には大きな違いはなく、catalytic site の構造は allosteric site の結合によってのみ制御され、catalytic site への Ras の結合による induce はないものと思われる。さらに、allosteric site に結合した Ras の構造を Ras 単体の PC1 の軸方向で見たとき、RasGTP がより強く switch I を閉じていることが分かり、そのわずかな違いが SOS の Rem domain の位置を制御しているものと推測される。

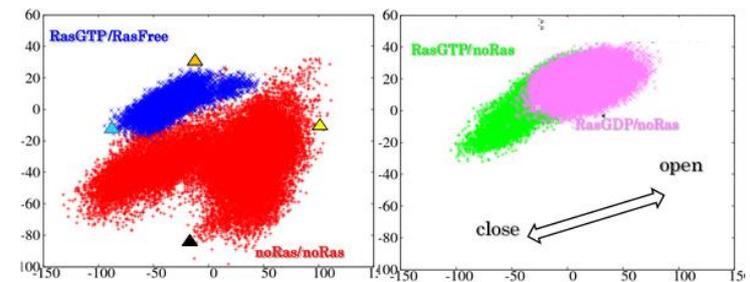


図 5 : Ras のヌクレオチド状態変化による SOS 構造ダイナミクス変化。縦軸横軸は、SOS 主成分解析での PC1/PC2。

## (3) ユビキチン化ヒストン H3 結合による DNMT1 活性化の分子シミュレーション

DNA methyltransferase 1 (DNMT1)は DNA 複製の際にメチル化 DNA の情報を正確に伝達する上で重要な役割を果たす。メチル化された DNA は DNA 複製にともない片方の DNA 鎖のみがメチル化されたヘミメチル化 DNA となるが、DNMT1 はヘミメチル化 DNA を基質として新生 DNA 鎖にメチル基を付加する。ヘミメチル化 DNA 結合タンパク質 UHFR1 により DNMT1 がヘミメチル化部位にリクルートされるが、それには DNMT1 の RFTS ドメインによるユビキチン化ヒストン H3 の認識が必要である。DNMT1 の RFTS ドメインと 2 つのサイトでユビキチン化されたヒストン H3 との複合体結晶構造が解かれ、その特異的な認識機構が示された。本研究では、DNMT1 全長 (RFTS + CXCC + BAH1 + BAH2 + catalytic ドメイン ; N 末端天然変性領域は除いた) とユビキチン化ヒストン H3 (Ub18+Ub23+ヒストン H3 テイル) との複合体モデルを構築し、その分子シミュレーションを行うことで、DNMT1 がユビキチン化ヒストン H3 との結合によってどのように自己阻害型から活性型に構造を変化させるかを明らかにすることをめざした。

まず、自己阻害型である human DNMT1 の結晶構造をもとに、ユビキチン化ヒストン H3 結合を阻害している RFTS 末端 C-linker を Targeted MD により動かすことで、ユビキチン化ヒストン H3 との複合体を構築した。活性型構造が未知のため、Targeted MD のシミュレーション時間(1-5 ns)、また、力の定数(0.5-5 kcal/mol/Å<sup>2</sup>)を変えながら自己阻害型からの RMSD が大きくなるような Targeted MD 計算を行ったところ、ユビキチン化ヒストン H3 を配置することが可能なランでは図 6 と同等の構造モデルが得られることから、モデルの妥当性を確認できた。

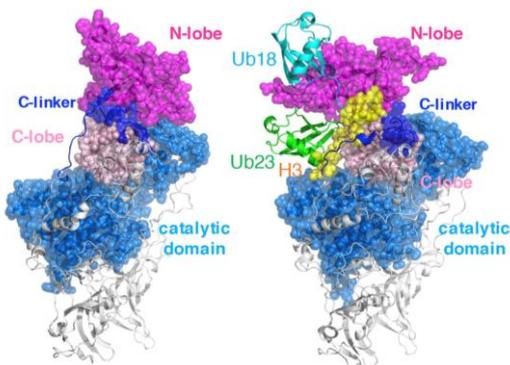


図 6 : DNMT1 の活性型複合体モデル構造

さらに、モデリングした 3 つの複合体について 100 ns の分子動力学シミュレーションを実行し、自己阻害型 DNMT1 単体の結果と比較したところ、ユビキチン化ヒストン H3 の結合により、RFTS と catalytic domain の相互作用を弱くし、界面にある触媒部位とヘミメチル化 DNA の結合を促進することが示された (図 7)。

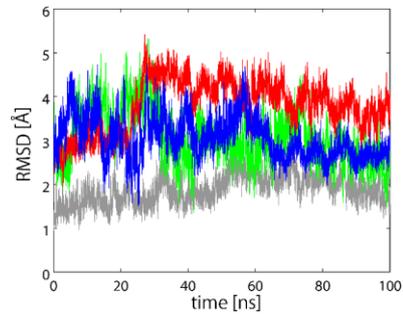


図 7 : catalytic domain から見た RFTS C-lobe の RMSD。自己阻害型 (灰色)、および活性型モデル (赤、緑、赤)。

## 4. まとめと今後の計画・展望

次世代生命体統合シミュレーション研究開発プログラムの一環としてマルチコピー・マルチスケール分子動力学シミュレーションのためのクラスライブラリ  $\mu^2\text{lib}$  を用いた MSES 法の応用研究を行った。創薬ターゲットとなっているグルコキナーゼ活性化の分子機構を明らかにし、MSES 法による全原子構造探索が創薬研究においても有用になる可能性を示した。系の自由度に応じて構造探索に必要なシミュレーション時間は指数関数的に増大するが、新規アルゴリズムとマルチコピー・マルチスケール手法の組み合わせにより初めて可能になったシミュレーション成果であるといえる。

来年度は、SOS の更なる研究として MSES 法による全原子構造探索を進め、今年度の brute-force MD 結果に確証を与えるようなシミュレーション結果をまとめたうえで、論文化に取り組む。また、新規の研究テーマとして、BW-MPC の大規模計算資源を効率的に用いることにより、MSES 法による CYP の全原子構造と ABC トランスポーターのストリング法による最小自由エネルギー経路探索を行う予定である。

平成 29 年度 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. Satoshi Ishiyama, Atsuya Nishiyama, Yasushi Saeki, Kei Moritsugu, Daichi Morimoto, Luna Yamaguchi, Naoko Arai, Rumie Matsumura, Toru Kawakami, Yuichi Mishima, Hironobu Hojo, Shintaro Shimamura, Fuyuki Ishikawa, Shoji Tajima, Keiji Tanaka, Mariko Ariyoshi, Masahiro Shirakawa, Mitsunori Ikeguchi, Akinori Kidera, Isao Suetake, Kyohei Arita, and Makoto Nakanishi, "Structure of the H3 ubiquitin code/Dnmt1 reader module complex reveals the basis for DNA methylation maintenance", *Molecular Cell*, (2017) 68: 350-360.
2. Kei Moritsugu, Tohru Terada, Hironori Kokubo, Satoshi Endo, Toshimasa Tanaka and Akinori Kidera, " Multiscale Enhanced Sampling for Glucokinase", *Journal of Chemical Physics*, *submitted*.