

令和元年度 文部科学省

ポスト「京」で重点的に取り組むべき社会的・科学的課題に関する
アプリケーション開発・研究開発

令和元年度

「生体分子システムの機能制御による革新的創薬基盤の構築」

成果報告書

令和2年5月29日

国立研究開発法人 理化学研究所

奥野 恭史

本報告書は、文部科学省の科学技術試験研究委託事業による委託業務として、国立研究開発法人 理化学研究所が実施した令和元年度「生体分子システムの機能制御による革新的創薬基盤の構築」の成果を取りまとめたものです。

目次

1. 委託業務の題目	1
2. 実施機関（代表機関）	1
3. 委託業務の目的	1
4. 令和元年度（報告年度）の実施内容	2
4-1. 実施計画	2
4-2. 実施内容（成果）	5
4-3. 活動（研究会等）	90
4-4. 実施体制	92
別添1 学会等発表実績	
別添2 実施計画	

1. 委託業務の題目

「生体分子システムの機能制御による革新的創薬基盤の構築」

2. 実施機関（代表機関）

代表 機 関	機関名		国立研究開発法人理化学研究所			
	所在地		〒351-0198 埼玉県和光市広沢 2-1			
課題 責任者	ふりがな	おくの やすし	生年	西暦 1970 年 6 月 17 日 (49 歳)		
	氏名	奥野 恭史	月日	※2020 年 4 月 1 日現在		
	所属部署名	生命機能科学研究センター		役職	客員主管研究員	
	連絡先	Tel. 06-6155-0111 Fax. 06-6155-0112 E-mail yasushi.okuno@riken.jp				
事務 連絡 担当者	ふりがな	きら みずほ				
	氏名	吉良 みずほ				
	所属部署名	外部資金室	役職	主査		
	連絡先	Tel. 048-467-9642 Fax. 048-467-2634 E-mail mizuho.kira@riken.jp				

3. 委託業務の目的

この十数年、製薬業界では、世界的規模で、新薬の承認数が横ばい状態（20 品目程度／年）であるのに対し、研究開発費が増え続けているという深刻な問題に直面している。従って、開発費を抑えながら、効率的に新薬を開発することは、製薬業界にとっての重要課題の一つであり、計算機シミュレーションによる創薬に大きな期待が寄せられている。しかしながら、現状の創薬計算技術は、予測精度の低さや評価できる化合物数・標的タンパク質数の限界などの問題から、実際の実験に置き換わるほどの革新的技術に至っていない。本研究では、ポスト「京」の演算能力を最大限に活かす分子シミュレーション技術を開発することで、生体分子システムの時間的空間的機能解析に資する新たな計算生命科学を開拓するとともに、これらを用いた創薬計算フローと次世代創薬計算技術を開発することで、高精度かつ超高速の革新的な創薬計算基盤の構築を目指すものである。ポスト「京」を基軸とした創薬計算基盤を製薬会社に提供することで、医薬品開発が加速され、画期的新薬や希少疾患薬の創出が期待される。

このため、国立研究開発法人理化学研究所を中核機関として、分担機関である国立大学法人京都大学大学院医学研究科、国立大学法人京都大学大学院理学研究科、国立大学法人東京工業大学生命理工学院、国立大学法人東京大学大学院新領域創成科学研究科、国立大学法人東京大学大学院情報学環、国立大学法人東京大学先端科学技術研究センター、国立大学法人名古屋大学、公立大学法人横浜市立大学、国立研究開発法人産業技術総合研究所、及び国立研究開発法人量子科学技術研究開発機構と密接に連携し、研究開発を実施する。

4. 令和元年度（報告年度）の実施内容

4-1. 実施計画

令和元年度は、本格実施フェーズとして、実施計画（研究開発内容、目標・期待される成果、実施体制、必要計算資源、工程表、所要経費等）に基づいた研究開発を、分担機関の国立大学法人京都大学大学院医学研究科、国立大学法人京都大学大学院理学研究科、国立大学法人東京工業大学生命理工学院、国立大学法人東京大学大学院新領域創成科学研究科、国立大学法人東京大学大学院情報学環、国立大学法人東京大学先端科学技術研究センター、国立大学法人名古屋大学、公立大学法人横浜市立大学、国立研究開発法人産業技術総合研究所、及び国立研究開発法人量子科学技術研究開発機構と共同で実施する。本格実施フェーズ最終の本年度は、超大規模計算手法確立、次世代創薬計算手法確立、統合システム完成を実施する。

具体的には、サブ課題 A「ポスト「京」での MD 高度化とアルゴリズム深化」では、コデザインと新規手法を結合することを目標とし、長時間ダイナミクスを創薬応用計算において実現し、超大規模な生体分子系シミュレーション法を確立することで国際的優位性をもつ計算手法を確立する。サブ課題 B「次世代創薬計算技術の開発」では、次世代創薬計算手法を完成させ、DA 法、PPI、NPI、ウイルス標的、細胞標的について国際的優位性のある計算手法を確立する。サブ課題 C「創薬ビッグデータ統合システムの開発」では、創薬計算フローのテスト、およびゲノム医療・ゲノム創薬への応用を通じて、統合システムを改良・性能評価を行うことで、統合システムの完成を目指す。

(1) ポスト「京」での MD 高度化とアルゴリズム深化（サブ課題 A）

①ポスト「京」での MD 高度化

令和元年度は、長時間にわたる MD 計算を効率的に実施するために、数値積分の時間刻み幅を大きくする手法などを導入し、MD 計算のさらなる効率化を進める。さらに様々な分子系で安定な MD 計算を実行できるように評価と改良を行い、時間刻み幅の大きな手法での MD 計算手法を確立する。

②拡張アンサンブル法

令和元年度は、自由エネルギー摂動法（FEP）を GENESIS 公開版に移植、網羅的なテスト計算を実施した後に公開する。公開に伴い、MD 計算のセットアップ及び解析用ツールも整備する。これまでに、gREST/REUS 法を用いてキナーゼ・阻害剤ドッキングの効率化に成功したことをうけ、異なるサイズの阻害剤との結合状態や準安定状態などの予測を行う。また、反応座標を用いない gREST 法によるドッキング構造の予測も含めて gREST 法を用いた創薬応用計算の手法を確立する。

③長時間ダイナミクス法（再委託先：国立大学法人東京工業大学生命理工学院 北尾彰朗）

令和元年度は、PaCS-MD 法を用いた解離シミュレーションまたは結合・解離シミュレーションとマルコフ状態モデルを組み合わせた解析を用いた結合自由エネルギー・キネティクス計算を、タンパク質-ペプチドまたは低分子複合体の長時間ダイナミクス研究に応用する。さらに PaCS-MD 法を用いた解離シミュレーションとマルコフ状態モデルを組み合わせた結合自由エネルギー計算を用いてタンパク質-タンパク質複合体モデル構造の評価を行う。

④ハイブリッド QM/MM 自由エネルギー法（再委託先：国立大学法人京都大学大学院理学研究科化学専攻 林重彦）

令和元年度は、タンパク質の薬剤結合に対して、非経験的 QM/MM 自由エネルギー法に基づく高精

度な結合エネルギー計算法の開発を行う。これまでの HIV プロテアーゼに対する計算の進展を受け、QM/MM 自由エネルギー構造最適化及び自由エネルギー摂動法計算を行うことにより、天然タンパク質における天然基質分子の反応活性化状態や薬剤分子結合状態のプロトン化状態を同定する。さらに、薬剤耐性変異体にも同様の計算を行うことにより、薬剤耐性の分子科学的解析手法の開発を行う。また、リガンド分子に対する分子力場での記述が困難な光受容体タンパク質の活性化状態に対して QM/MM 自由エネルギー構造最適化及び自由エネルギー摂動法計算を行い、性能評価を行う。またリガンド分子の構造揺らぎの効果の計算における自由エネルギー外装法で必要となる MM/MM RWFE 法を開発する。

- ⑤粗視化モデリング（再委託先：国立大学法人京都大学大学院理学研究科生物科学専攻 高田彰二、国立大学法人名古屋大学 篠田渉）

令和元年度は、これまでに開発した大規模なタンパク質・核酸・リン脂質複合系に対する高精度な粗視化モデルを適用して、クロマチン動構造、転写制御因子とクロマチン構造の共役、生体膜とタンパク質の動態についての計算を実施する。具体的には、酵母の全遺伝子の転写開始サイトの前後に存在する -1、及び +1 ヌクレオソームの動的安定性と遺伝子発現特性の関係、山中因子 Oct4, Sox2 のヌクレオソーム結合動態の解析、膜タンパク質による生体膜の動構造変化等についての、細胞生物学的実証研究を行う。

- ⑥自由エネルギー計算法（再委託先：公立大学法人横浜市立大学 池口満徳）

令和元年度は、リガンドとタンパク質間の高効率・高精度な結合自由エネルギー計算法の開発を継続する。平成 30 年度に確立したタンパク質・リガンドの静電相互作用に関わるサイズ効果の影響のない方法を拡張し、より汎用的な結合自由エネルギーの計算法を開発する。そこで開発された方法を、タンパク質とリガンドの結合自由エネルギー計算に適用し、実験値との比較等の性能検証を行う。

(2) 次世代創薬計算技術の開発（サブ課題 B）

- ①動的分子機能制御（再委託先：公立大学法人横浜市立大学 池口満徳）

令和元年度は、リガンド結合に際して立体構造変化が機能に重要な役割を果たすタンパク質について、分子の動的機能の解析を継続する。スーパーコンピュータを用いて多数の全原子分子動力学シミュレーション計算を実施し、リガンドの種類（アゴニスト、アンタゴニストなど）によって、タンパク質に誘起される立体構造変化の相違についての分子メカニズムの研究を行う。さらに、近年の創薬で注目されている中分子もターゲットにし、その動的構造の解析によって膜透過などの機能との関係を明らかにする。

- ②タンパク質間相互作用制御（再委託先：国立大学法人東京大学 藤谷秀章・山下雄史）

令和元年度は、これまでに開発を進めてきた新しい FUJI 脂質力場を用いて、抗生物質が効かない多剤耐性菌などが持つ多剤排出ポンプの分子動力学計算を行う。特に、現実の多剤耐性菌細胞の膜成分比を再現した多成分膜モデルを構築し、その環境下で超巨大タンパク質複合体のタンパク質間相互作用が排出ポンプ制御にどの様に関わっているかを解析する。また、これまでの抗原-抗体系の分子動力学計算の結果にさらに詳細な解析を行い、タンパク質間相互作用の分子論的観点から抗体医薬品開発に応用する方法を検討する。

③ウイルス標的創薬計算技術（再委託先：国立大学法人名古屋大学 篠田渉）

令和元年度は、B型肝炎ウイルスカプシドを対象として、薬剤のカプシド吸収における自由エネルギー解析技術の開発・改良を進め、複数の逆転写酵素阻害剤や核酸アナログのカプシド吸収効率の評価を実施する。また、粗視化モデルの高精度化を行い、カプシド全体系での阻害剤吸収の分子メカニズムを長時間分子動力学シミュレーションによって解析する。形成阻害剤によるカプシドタンパク複合体の結合構造への影響を系統的に調べ、複数の阻害剤について構造安定化への寄与を評価する。

④核酸-タンパク質相互作用制御（再委託先：国立研究開発法人量子科学技術研究開発機構 河野秀俊）

令和元年度は、引き続き核内の DNA 収納状態の動構造を調べる。クロマチン状態の DNA は、タンパク質の結合やクロマチンの構造変化によって、引っ張りや振じれといった外力を受ける。この外力に対する応答をシミュレーションし、詳細な構造変化のメカニズム解析や応答の自由エネルギー解析を行う。また、その応答によって、ゲノム DNA がどの程度転写因子等のタンパク質からアクセスしやすくなるかを明らかにする。

⑤細胞内環境シミュレーション

令和元年度は、分子混雑環境におけるタンパク質・基質結合の分子機構を解析し、溶液中におけるタンパク質・基質結合との分子機構の違いを明らかにする。細胞内環境におけるタンパク質は、その周囲に複数の off-target タンパク質が存在する。分子混雑環境を用いて、基質あるいは薬剤が細胞内で target タンパク質に到達し、結合部位に強く結合するプロセスを解析する。

(3) 創薬ビッグデータ統合システムの開発（サブ課題 C）

①高精度薬剤デザイン（再委託先：国立大学法人京都大学大学院医学研究科 奥野恭史）

令和元年度は、サブ課題 A-C で開発される MD 計算法・各種創薬計算法・標的タンパク質データベースに加えて平成 30 年度に選定した AI 創薬手法を統合し、創薬ビッグデータ統合システムを完成させる。また、協力機関と連携し、創薬ビッグデータ統合システムを具体的な創薬ターゲットに適用することで性能評価を引き続き行う。特に、平成 30 年度に構築したゲノム医療分子シミュレーション基盤を実践的に活用する事で、計算フローのブラッシュアップを行う。

②機能制御部位データベース（再委託先：国立研究開発法人産業技術総合研究所 広川貴次）

令和元年度は、機能部位予測データベースのシステム改良とサブ課題 C①と連携した創薬ビッグデータ統合システムへの組み込みを行う。アンサンブル構造を活用したバーチャルスクリーニングやドッキングによる機能部位予測については、システムとして実装を行い、共同研究を通じて創薬標的タンパク質に適用し、性能評価と手法の高精度化を可能な限り続けて行う。

③創薬関連ビッグデータ（再委託先：国立大学法人東京大学大学院新領域創成科学研究科 津田宏治）

令和元年度は、強化学習と分子動力学計算を用いた構造変化パスウェイ探索アルゴリズムを開発する。現在ポスト「京」プロジェクトで使用されている PaCS-MD 法は、短い時間の MD を直列に組み合わせることで、初期構造から目標構造への軌跡を短時間で得ることができるが、反面、局所最適構造にトラップされやすい傾向がある。本研究では、木探索を用いることで、一旦トラップされても、後戻りして抜け出すことを可能にする。実際のタンパク質構造に適用し、性能を評価する。

④標的分子ネットワーク（再委託先：国立大学法人東京工業大学生命理工学院 北尾彰朗）

令和元年度は、これまで開発してきた分子動力学計算とエネルギー表示の理論を組み合わせたタ

ンパク質—タンパク質複合体モデル構造評価法 evERdock の効率化を完了する。具体的には機械学習などを導入することで計算時間を軽減する方法の改良を完成させる。さらに evERdock でモデル評価の対象とする複合体モデル構造を粗視化レベルで絞り込んでおくためのクラスタリング法の高精度化を加えて、標的分子ネットワーク評価法の応用を行う。

⑤生体系マルチスケールモデリング（再委託先：国立大学法人東京大学大学院情報学環 寺田透）

令和元年度は、心筋イオンチャネル hERG の電位センサードメインについて、前年度までに構築した立体構造変化の推定経路上のモデルを初期構造として、分子動力学シミュレーションを多数実施する。このトラジェクトリを解析・統合し、hERG 電位センサードメインの電位依存的な立体構造変化を記述するマルコフ状態遷移モデルを構築する。また、心筋イオンチャネルのポア領域の立体構造モデルに対して、不整脈を誘発することが知られている薬剤のドッキングを行い、分子動力学シミュレーションを用いて結合自由エネルギーを求める。ここでは複数の条件でドッキングを行い、実験値（IC50 値）を最も良く再現する条件を決定する。

（4）プロジェクトの総合的推進

プロジェクト全体の連携を密としつつ円滑に運営していくため、執行部会議を定期的実施し、プロジェクト運営全般の検討・決裁・指揮を行う。重点課題1 ワークショップ（全体会議）を開催し、プロジェクト全体の進捗状況を確認しつつ、プロジェクトで得られた成果については、学会やシンポジウムなどの発表にて成果を取りまとめる。

さらに、ポスト「京」創薬推進委員会、実験技術連携推進検討会、バイオ計算分野振興 WG において、外部有識者を招聘して意見を聞くとともに、協力機関との連携・調整にあたる。

令和元年度も引き続き、日本生物物理学会などバイオ計算に関心の高い研究者を数多く集める学会との共催により本課題に関連した発表枠を設けることで、効果的な討論、意見交換を行うほか、ポスト「京」プロジェクトで同じく健康長寿社会の実現をターゲットとする重点課題2との連携によるシンポジウムを開催し、プロジェクトで得られた成果についての討論や意見交換、さらに知見を広めるとともに本プロジェクトのプレゼンスを高める。

また、課題推進にあたり、第三者の視点で進捗状況の把握・評価・改善提言・指導等を行う諮問委員会を適宜開催する。

4-2. 実施内容（成果）

令和元年度は、本格実施フェーズとして、実施計画（研究開発内容、目標・期待される成果、実施体制、必要計算資源、工程表、所要経費等）に基づいた研究開発を、分担機関の国立大学法人京都大学大学院医学研究科、国立大学法人京都大学大学院理学研究科、国立大学法人東京工業大学生命理工学院、国立大学法人東京大学大学院新領域創成科学研究科、国立大学法人東京大学大学院情報学環、国立大学法人東京大学先端科学技術研究センター、国立大学法人名古屋大学、公立大学法人横浜市立大学、国立研究開発法人産業技術総合研究所、及び国立研究開発法人量子科学技術研究開発機構と共同で実施した。本格実施フェーズ最終の本年度は、超大規模計算手法確立、次世代創薬計算手法確立、統合システム完成を行った。

具体的には、サブ課題 A「ポスト「京」での MD 高度化とアルゴリズム深化」では、コデザインと

新規手法を結合することを目標とし、長時間ダイナミクスを創薬応用計算において実現し、超大規模な生体分子系シミュレーション法を確立することで国際的優位性をもつ計算手法を確立した。サブ課題 B「次世代創薬計算技術の開発」では、次世代創薬計算手法を完成させ、DA 法、PPI、NPI、ウイルス標的、細胞標的について国際的優位性のある計算手法を確立した。サブ課題 C「創薬ビッグデータ統合システムの開発」では、創薬計算フローのテスト、およびゲノム医療・ゲノム創薬への応用を通じて、統合システムを改良・性能評価を行うことで、統合システムを完成した。

(1) ポスト「京」での MD 高度化とアルゴリズム深化 (サブ課題 A)

①ポスト「京」での MD 高度化

本課題では、大規模生体システムの計算を実施するために、ターゲットアプリである超高並列分子動力学計算プログラム GENESIS のポスト「京」での高度化を行ってきた。分子動力学法は原子・分子間に働く相互作用力を用いて運動方程式を数値的に解くことで、粒子の運動を求める方法である。この計算の中で、運動方程式を短い時間ステップ幅で繰り返して解くため、各ステップで相互作用力を計算する必要がある。長時間のタンパク質の構造揺らぎや構造変化を観測するためには、非常に多くのステップを繰り返し計算することが必要となるため、膨大な演算量が必要である。実施には、ほとんどの演算時間は相互作用力を求める部分で費やされる。相互作用力の計算は、非結合相互作用と結合性相互作用に分けられるが、原子対の数が多い、すなわち、最大で原子の数の二乗に比例する非結合性相互作用計算が演算時間のボトルネックとなる。ポスト「京」(富岳)での MD の高度化は、主にこのボトルネックとなる非結合相互作用を中心に行ってきた。昨年度までの研究において、富岳の SIMD 幅を活かすアルゴリズム修正やチューニングによる高速化を進めてきた。令和元年度には、スレッド並列時の負荷分散を修正した。また、コンパイルオプションに対する性能を確認し、適切なコンパイルオプションの選定を行った。さらに、ボトルネック以外の区間に対してもスレッド並列の負荷分散やコンパイルオプションの最適化を通じて、高速化を行った。高度化された GENESIS を用いて、スーパーコンピュータ富岳の試作機、更には開発中の富岳において、MD 計算全体の性能評価を行った。アプリケーション実行性能として、「京」全体を利用した場合と比較して、「富岳」全ノードでは 125 倍の性能向上を実現することが示された。

また、長時間にわたる MD 計算を効率的に実施するために、さらに長いステップ幅でも安定した MD 計算を可能にする高精度な時間発展スキームを導入した。具体的には、新たな温度定義を導入することで、精度を損なうことなく、長い時間ステップ幅を用いた計算を可能にした。通常の MD 計算では、各時間ステップでの温度は、時間ステップ幅もしくは $1/2$ 時間ステップ幅での運動エネルギー (K_{full} もしくは K_{half}) を用いてエネルギー等分配則に基づき決められている。しかし、そのように決めた温度 (T_{full} もしくは T_{half}) は、時間ステップ幅の 1 次までの精度しかなく、時間ステップを大きくすると MD 計算の精度が著しく損なわれてしまう。そこで、 K_{full} と K_{half} を組み合わせることで、時間ステップ幅の 3 次までの精度を保つ新たな温度定義を導出し GENESIS に実装した。この新しい温度定義に対して、水分子や膜分子などに適用する事で、長いステップ幅を使用しても物性・性質を適切に計算可能であることを確認した。同結果は論文として国際誌に掲載された (J. Jung et al. J. Chem. Theory Comput. (2019))。

更に、タンパク質や核酸などの実際の生体内高分子に適用するために、温度・圧力計算手法の改善、

水素原子の質量を増加させる Hydrogen Mass Repartitioning (HMR)法の安定なパラメータ設定、数値的に安定な multiple time step (MTS)手法の開発を行った。これらの手法を富岳用にチューニング、高速化されたコードへ導入し、富岳での更なる性能向上を目指した。また、レプリカ交換法など生体分子の解析に重要な機能の導入も併せて行い、水溶性タンパク質、脂質二重膜、膜タンパク質、RNA など様々な生体分子に応用して、実際に長いステップ幅でも安定な計算を実現できることを確認した。現在利用できる「富岳」のノードを用いて、細胞質に含まれる多数のタンパク質、核酸、代謝物を含む超大規模な分子混雑環境のシミュレーションを長時間ステップ幅。この結果については国際的な会議に論文投稿中である。

このように、富岳の性能を活かすべく高度化、高速化されたコードと信頼性の高い温度定義と長いステップ幅の実現は、分子動力学法が観測できる空間的・時間的限界を大きく広げるものである。これらの手法が導入される GENESIS と富岳が広く利用されることで、生物物理の基礎的な問題だけでなく、医薬分野での応用などにも大きな飛躍が生まれることができると確信する。

②拡張アンサンブル法

GENESIS による高精度かつ超高速の創薬計算を実現するために、分子シミュレーションの要素計算技術を開発し評価している。令和元年度は、自由エネルギー摂動法 (Free Energy Perturbation 法、FEP 法) を GENESIS 公開版に移植しつつ、多様な系に対応するために、拘束条件を用いる手法と用いない手法 (MP-CAFEE 法) の両方を実装した。両手法は結合自由エネルギー計算に用いられる代表的なものであり、今回の実装により、絶対結合自由エネルギー、相対結合自由エネルギー、タンパク質の変異による結合自由エネルギー変化など、多様な結合自由エネルギー計算が可能となった。特に MP-CAFEE 法は国内の創薬計算で広く用いられており、今回の実装により創薬分野での実践的な利用が期待される。代表的なテスト系として FKBP と T4 リゾチームのテスト計算を実施し、実装及び手法の精度を確認した。図 1 に FKBP に対する 2 種のリガンド (L8 と L9) の相対結合自由エネルギーを計算した結果を示す。実験的には両リガンドの結合自由エネルギー差は 0.2 kcal/mol であることが知られているが、GENESIS を用いた計算はそのことを良く再現する (計算値は 0.0 ± 0.8 kcal/mol)。

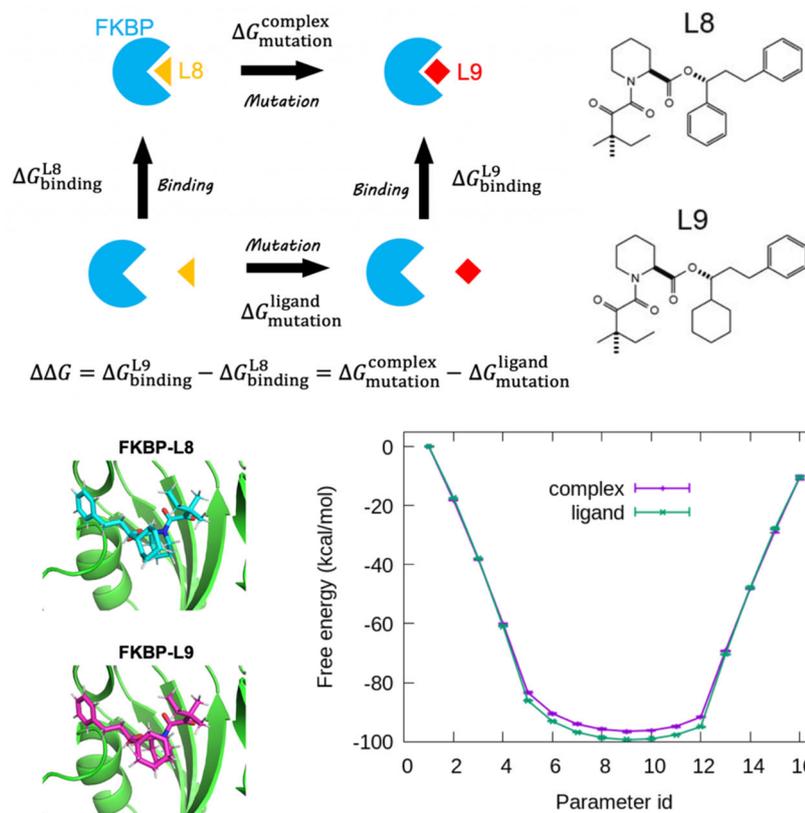


図 1 自由エネルギー摂動法 (FEP 法) を用いた相対結合自由エネルギーの計算スキーム (左) と、対象とした FK506 結合タンパク質 (FKBP) と 2 種のリガンド (L8 と L9)。FEP 計算では、L8 リガンドが FKBP に結合した状態を初期状態、L9 リガンドが FKBP に結合した状態を最終状態とした。同様に、L8 を含む水ボックスを初期状態、L9 を含む水ボックスを最終状態とした。タンパク質に結合した状態 (complex) と水に囲まれた状態 (ligand) の各々に対して、初期状態から最終状態への自由エネルギー変化を求めた (下)。

FEP 法は結合自由エネルギー計算の有用な手法であるが、分子の置換や相互作用のスケールリングを制御のために複雑な入力ファイルを作成する必要がある。FEP 計算のセットアップを自動化・簡素化することは、人為的なミスが減らすとともに、利用者の拡大にもつながる。令和元年度は、MD 計算のセットアップ及び解析用ツールとして、現在広く用いられている MD 計算のセットアップツール CHARMM-GUI の開発者と協力し、GENESIS の FEP 計算用に入力ファイルを生成するツールのプロトタイプを作成した。これら GENESIS の FEP 機能は、バージョン 1.5.0 として公開するとともに、GENESIS の公開ウェブサイトに FEP 計算のセットアップから計算、解析までの詳細をチュートリアルとして公開し、利用者が用意に実施できる環境を整備した。

また、これまでに、gREST/REUS 法を用いてキナーゼ・PP1 阻害剤ドッキングの効率化に成功したことをうけ、実際の分子標的薬であるダサチニブ (分子量 500 程度) とキナーゼとの結合状態や準安定状態を予測した。分子量が大きき場合でも X 線結晶構造を良く再現し、準安定状態やそれらへ至る経路も予測できた。分子量が大きくなったことで、PP1 阻害剤の計算に比べ、結合・脱離の観測回

数は半減するものの、結晶構造を高精度に予測出来ることを確認した (図 2)。結晶構造以外にも、準安定な結合状態があることも予測できた。さらに過渡的な初期複合体 (会合体) の構造と相互作用を調べ、PP1 阻害剤の結合と同様に、G-loop と呼ばれる保存性の高いアミノ酸残基群がリガンドを最初に補足し結合へと導く重要な役割を果たしていることを確認し、今後の薬剤設計に役立つ新たな視点を与えた。一方で、阻害剤の分子量が大きくなったことで、分子の柔軟性が増し、多様な結合状態を特定する新たな解析が必要であることもわかった。

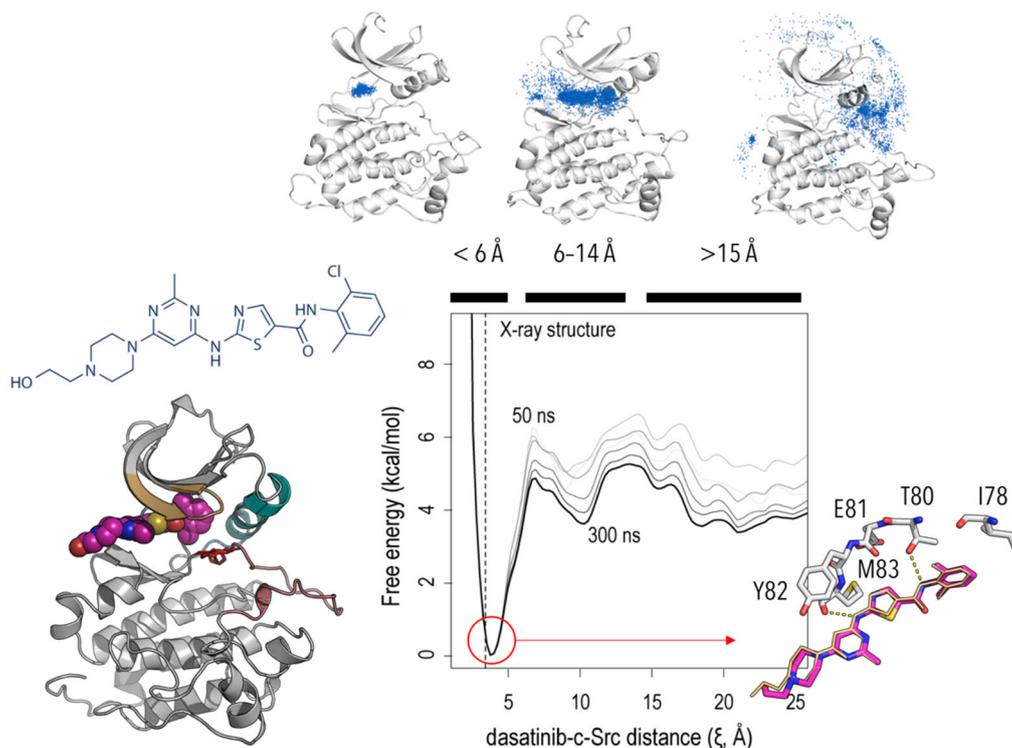


図 2 二次元レプリカ交換法 (gREST/REUS 法) の計算対象とした Src キナーゼとダサチニブ (リガンド) の構造 (左) と、計算から求めた結合の自由エネルギープロファイル (右)。自由エネルギーが最も低い点が X 線結晶構造 (黄) に対応する。タンパク質からの距離に沿ったリガンド分布の変化

また、ダサチニブの結合について、反応座標を用いない gREST 法によるドッキング構造予測にも成功した。gREST/REUS 法は、高い精度で結合予測が可能な反面、反応座標を用いる上、計算コストも非常に高いという課題がある。前年度の研究で、結合ポーズをより安価に求める手法として、gREST 法 (REST 法を改良した手法) に基づいた効率的なドッキング計算スキームを考案し、ベンチマーク計算に広く用いられている T4 リゾチームのリガンド結合に対して、X 線結晶構造を高い精度で再現することを示した。新たなスキームは、反応座標は用いずに、gREST 法のみを用いてタンパク質結合サイトの構造揺らぎを十分に取り込むことで探索効率を向上させる。さらに、Flat-bottom ポテンシャルを導入し、リガンドが結合サイト付近を十分探索するよう工夫している。令和元年度は、同手法を Src キナーゼとダサチニブの結合予測に適用した。gREST (溶質温度) には 10 レプリカを用いた。リガンドをタンパク質結合サイトの外側に配置した構造を初期構造とした。計算結果は良好で、タン

パク質の揺らぎを許すことにより、分子サイズの大きなダサチニブの結合も観測することに成功した (図 3)。

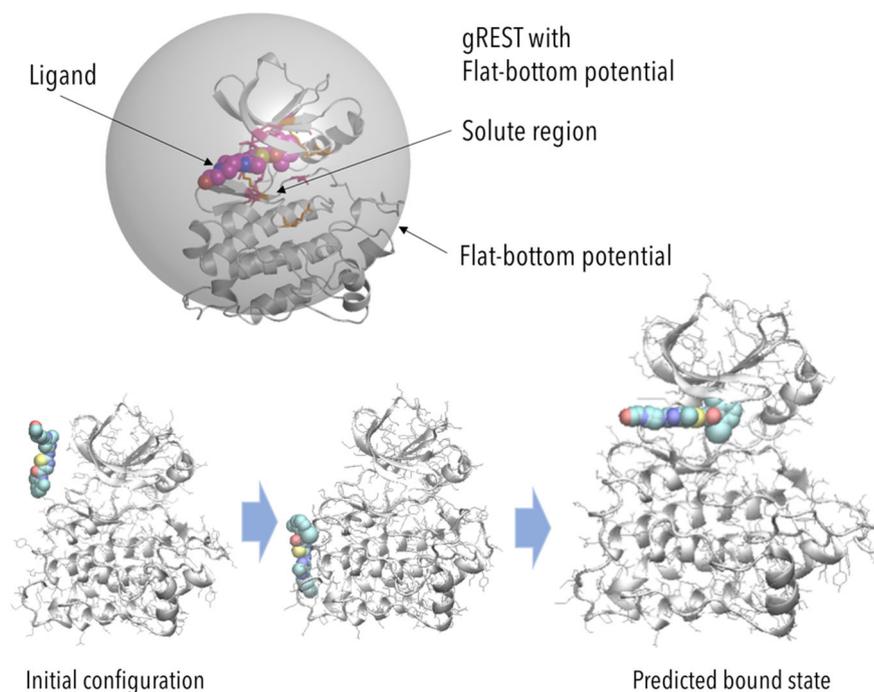


図 3 gREST 法と Flat-bottom ポテンシャルを用いた Src キナーゼとダサチニブの結合計算。リガンドと結合サイト付近のアミノ酸残基を gREST 法の solute (溶質部分) とした。Flat-bottom ポテンシャルはグレー球で示しており、リガンド分子は球の中を自由に動き回る。初期構造でタンパク質結合サイトの外側に配置したリガンドは、タンパク質表面の様々な部位を探索後、結合サイトに侵入した。

これらの結果から、実際の創薬計算において、リガンドの結合構造に関する情報がない場合でも、まず gREST 法を用いて結合サイトとポーズを予測し、次いで gREST/REUS 法で多様な結合状態と結合経路の詳細を検討できることを確認した。

最後に、実際の創薬計算を念頭にした一例として、はじめに gREST を用いて結合ポーズを予測し、次いで得られたポーズで FEP 計算により結合親和性を評価した。検証のために、代表的なテスト系として FKBP と 10 個のリガンドの結合を対象とした。FKBP のリガンド結合については、幾つかのリガンドに対して結晶構造が得られているが、敢えてそれらを用いずに全てのリガンド結合ポーズを予測した。ドッキング計算で求めた構造で FEP 計算により結合自由エネルギー見積もる従来法と比べ、gREST+FEP 法の結果は実験データをはるかに精度良く再現する (図 4)。以上の研究成果は、現在論文として投稿中である。以上、GENESIS を用いた創薬計算を実現するための技術基盤の構築を完了し、一連の創薬応用計算フローを確立した。

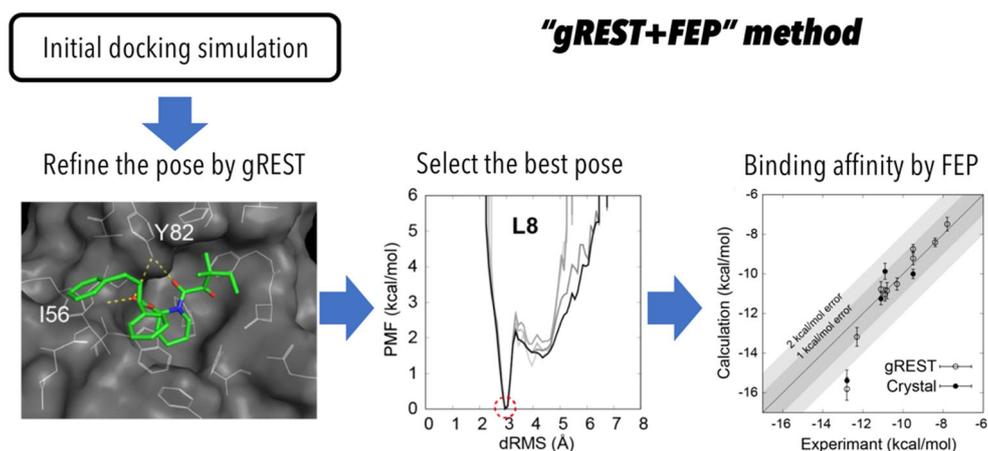


図4 gREST+FEP法の計算スキーム。はじめに、ドッキング計算により標的タンパク質とリガンドの結合構造を求め、次にgREST法で可能な結合構造を十分に探索する。その後、自由エネルギープロファイルに基づいて最も安定なポーズを選定し、FEP計算により結合自由エネルギーを求める。

③長時間ダイナミクス法（再委託先：国立大学法人東京工業大学生命理工学院 北尾彰朗）

令和元年度は、PaCS-MD法を用いた解離シミュレーション、結合・解離シミュレーションとマルコフ状態モデルを組み合わせた結合自由エネルギー・キネティクス計算を、タンパク質-ペプチドや複数の低分子複合体の長時間ダイナミクス研究に応用した。

まず複合体立体構造の知識を使わずに複合体を構成する2つの分子が解離した状態から出発し、結合・解離シミュレーションを用いて複合体モデルを生成したうえで、自由エネルギー最小の立体構造として複合体構造を予測し、長時間ダイナミクスを解明する研究を完成させた。この研究は前年度から継続してきたものである。この研究で開発した計算法は、以下の3つのステージからなる（図1）。最初のステージ1では、解離状態から出発し、PaCS-MDによってより分子間距離が近い構造を選択することで結合状態の立体構造を生成する（結合モード）。分子間距離が短くならなかったら、今度は分子間距離が長くなるようにPaCS-MDを行い、解離する過程を生成する（解離モード）。十分に分子間距離が長くなったら再び結合モードにスイッチする。ステージ1はこのように複数回、結合と解離を繰り返し、複合体モデル（デコイ）を多数生成するとともに結合・解離のパスウェイを多数生成する。後で述べる計算では結合・解離をそれぞれ11回行った。

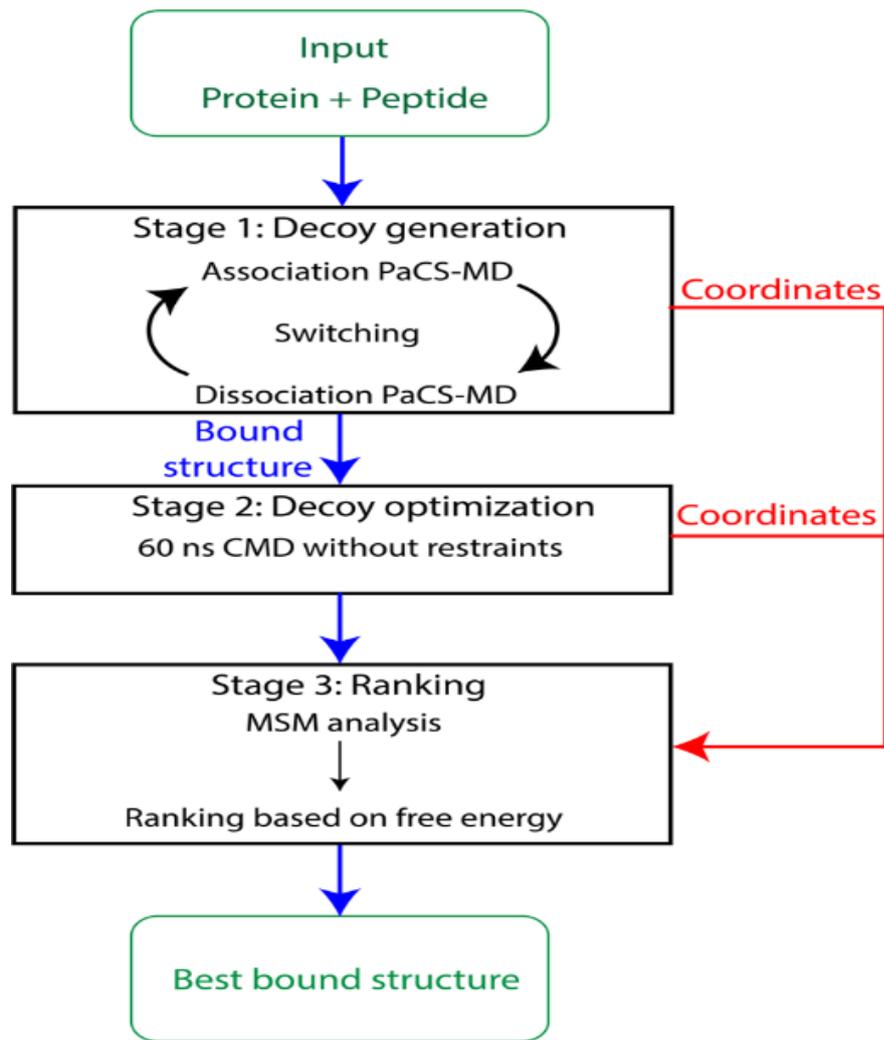


図1 複合体立体構造を予測する3つのステージ。デコイを生成するステージ1、デコイを最適化するステージ2、マルコフ状態モデル（MSM）を実行するステージ3からなる。

ステージ2では、ステージ1で生成したトラジェクトリから多数の複合体モデルを抽出し、それぞれの構造から出発した独立の最適化シミュレーションを行う。これは通常の分子動力学シミュレーションであり、この過程で複合体構造の更なる構造変化や脱水和などが起こる。今回は190の複合体モデルを選び、それぞれ60ナノ秒の分子動力学計算を行った。ステージ3では、ステージ1と2で得られたトラジェクトリをすべて統合し、マルコフ状態モデル（MSM）による解析を行った。これによりすべての状態を離散化し、定常状態での各状態の確率を決定することで、後述するエネルギー地形の解析や構造変化のパスウェイの流量（flux）解析を行う。

この計算法をタンパク質MDM2とガン抑制因子p53の天然領域ペプチドTAD-p53の複合体予測に応用した。MDM2の立体構造は単量体でも複合体でもほとんど変化しないが、TAD-p53は複合体形成時にはヘリックスを形成するものの、解離状態ではヘリックス構造が解消され特定の立体構造をとらない天然変性状態にある。シミュレーション中も解離状態ではヘリックス構造を

とらないことが多い。したがって、複合体構造予測が難しいケースであると言える。最初に、MSMによって自由エネルギー最低状態の立体構造を調べた。その結果、得られた立体構造はX線結晶解析で得られた複合体立体構造と一番近い構造であり、構造の違いも小さい(1.9 Å)ことがわかった(図2)。この結果は、我々が提案した計算法がフレキシブルドッキングの方法として利用できることを示すものである。

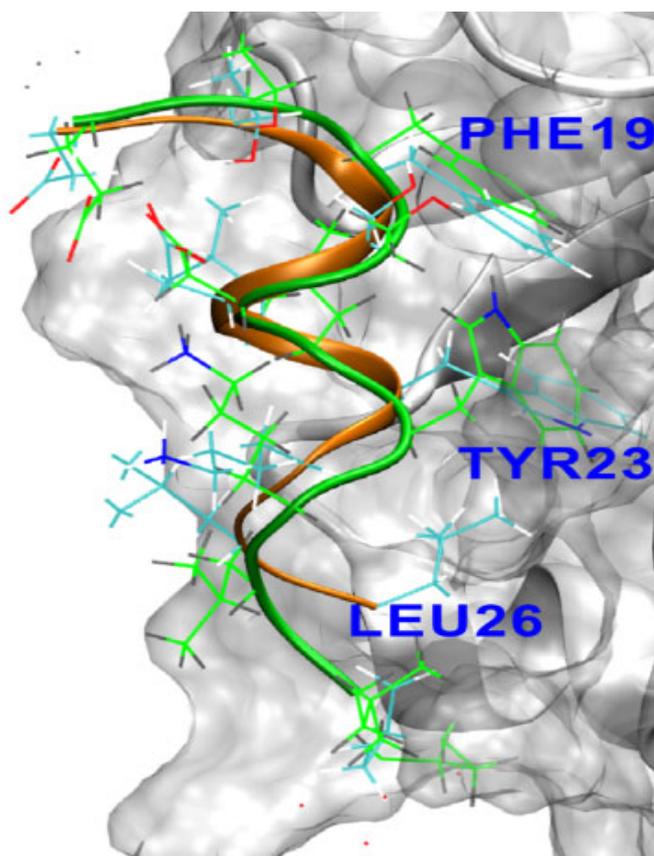


図2 X線結晶解析で得られた TAD-p53 の立体構造 (オレンジ) と自由エネルギー最低構造として予測された立体構造 (緑)。背景に表面モデルで示した MDM2 の立体構造は計算中大きく変化しない。

次に TAD-p53 と MDM2 の結合・解離経路の自由エネルギー地形の計算を行った。まず、非時間依存成分解析をおこない、その2つの軸 (TICA1 と TICA2) の上で自由エネルギー地形を示した(図3)。既に述べたように、天然の複合体立体構造 (CN) の自由エネルギーが最も低いが、それ以外にも自由エネルギー最低状態から ~ 1.5 kcal/mol 程度のエネルギー差で5つの中間状態 (CI₁~CI₅) が存在すること明らかになった。これらの中間状態で TAD-p53 は真の複合体界面の近くに結合しているものの、その配向や立体構造は天然構造とは異なるものであった。すなわち、天然構造に至るには、更なる配向や構造の変化が必要であることがわかる。

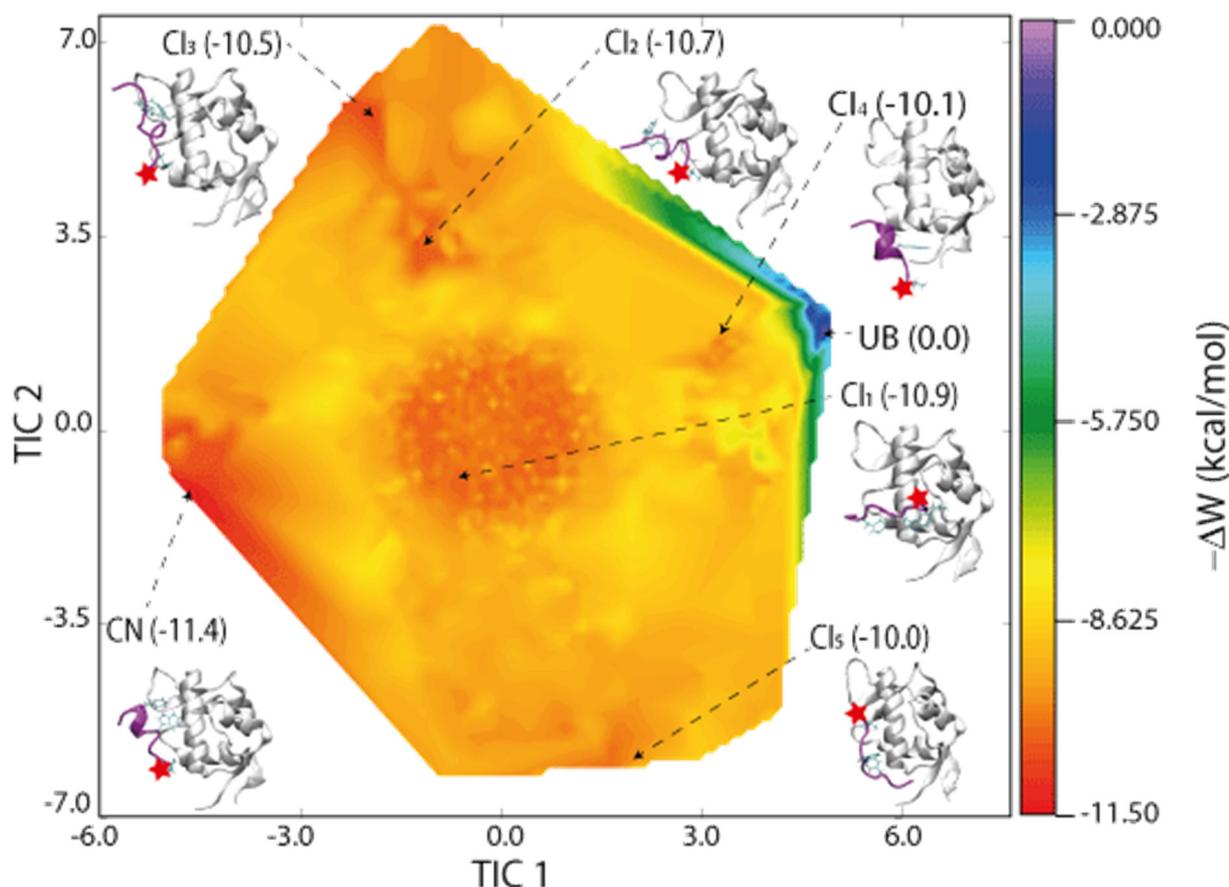


図3 2つの非時間依存成分解析の2つの軸 (TICA1 と TICA2) 上で示した TAD-p53 と MSM2 結合に関わる自由エネルギー地形と立体構造。(天然の複合体立体構造 (CN) と非結合状態 (UB) とその間にある5つの中間状態 (CI₁~CI₅) の位置を示してある。

次の図4に、非結合状態 (UB) から各状態へ向かうパスウェイとそれぞれの流量 (flux) を示した。この流量から、多数のパスウェイの中から UB から CI₂ を経て CN にいたるパスウェイが重要 (90%以上) であることが明らかになった。解離に向かうパスウェイもこれとは逆方向の CN => CI₂ => UB がメジャー (90%以上) であることも同様の解析から見出された。多数のエネルギー的にほぼ等価な中間状態があるにもかかわらず、1つの経路がメジャーであることは興味深い結果である。他の中間状態はエネルギー的には変わらないものの大きな構造変化を必要とすることから、キネティックな理由でメジャーな UB から CI₂ に至るパスウェイが選択されていると思われる。このことは非結合状態から中間状態に至る第一段階が Conformational selection 的な性質をもつことを示している。

一方、CI₂ から CN に至る結合の第二段階は第一段階と大きく異なる。第二段階では、ヘリックスの形成、側鎖のパッキングの形成、複合体界面からの脱水和がほぼ同時に起こっている。つまり複合体形成の第二段階は Induced fit 的な性格をもつことが明確になった。これらのことか

TAD-p53 の MSM2 への結合は、最初にキネティックに特定の間中状態が主に形成され、そこから Induce fit 的に最安定構造に至ることが示された。

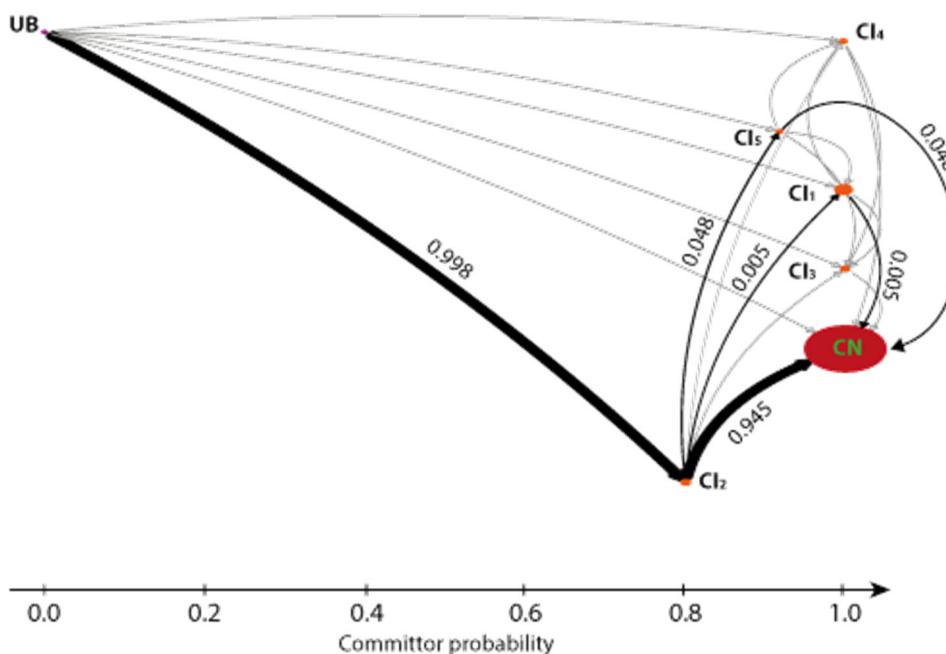


図 4 非結合状態 (UB) から天然の複合体立体構造 (CN) とその間にある 4 つの中間状態 (CI₁ ~ CI₅) の間の流量と Committed probability を示したもの。流量の比率を矢印の上の数値で示してある。

最後に、この計算で得られた標準結合自由エネルギー (ΔG°)、解離速度定数 (k_{off})、結合速度定数 (k_{on}) を我々が以前に解離型の PaCS-MD/MSM で計算した数値(Duy Phuoc Tran & Akio Kitao, *J. Phys. Chem. B*, 123, 2469-2478, 2019)、および実験 (ITC/stopped-flow と SPR) で得られた数値と下記の表 1 で比較する。前回の解離型の計算と同様に、今回の計算法でも標準結合自由エネルギー、解離速度定数、結合速度定数の実験値をほぼ再現できることが示された。 k_{off} の値からもわかるように、この複合体の解離に必要な実時間は秒のオーダーである。我々が開発した計算法は、分子動力学で合計 10 マイクロ秒程度の計算量で、秒オーダーの長時間ダイナミクスの解明を可能にすることができた。

この研究では、複合体立体構造予測と結合自由エネルギー・キネティクス計算が可能であることを示すとともに、結合解離経路の詳細な解析から複合体形成の詳しいメカニズムを明らかにすることもできた。フレキシブルな分子がタンパク質に結合した際の立体構造を計算するためには通常合算するとサブミリ秒からミリ秒程度のシミュレーションが必要であるが、我々はわずか 11.675 μs の計算量で達成していることから、画期的な計算効率を達成したと言える。この研究成果は、原著論文として発表済みである (Duy Phuoc Tran & Akio Kitao, *J. Chem. Theory Comput.*, 16, 2835-2845, 2020)。

表 1 この計算で得られた標準結合自由エネルギー (ΔG°)、解離速度定数 (k_{off})、結合速度定数 (k_{on}) を解離型 PaCS-MD/MSM および実験 (ITC/stopped-flow と SPR) との比較。

Method	ΔG° (kcal/mol)	k_{off} (s^{-1})	k_{on} ($10^7 M^{-1} s^{-1}$)
本研究	-8.9 ± 0.2	1.1 ± 0.3	0.25 ± 0.01
解離型 PaCS-MD/MSM	-8.4 ± 0.7	3.1 ± 1.3	1.4 ± 0.2
ITC/stopped-flow	-9.1 ± 0.1	2.06 ± 0.09	0.92 ± 0.004
	-9.9 ± 0.1	1.43 ± 0.24	2.2 ± 0.02
SPR	-8.8 ± 0.2	0.8 ± 0.3	0.21 ± 0.02

これ以外にも、解離シミュレーションとマルコフ状態モデルを組み合わせた PaCS-MD/MSM を応用した研究を行った。まず、以前から行ってきた細菌のシグナル蛋白質 CheY と細菌べん毛モーターの回転子タンパク質 FliM の結合安定性への圧力効果の研究を完了し、原著論文を発表することができた (H Hata, Y Nishihara, M Nishiyama, Y Sowa, I Kawagishi, A Kitao, *Sci. Rep.*, 10, 1-13, 2020)。さらにいくつかの低分子とタンパク質の複合体の標準結合自由エネルギーを計算し、実験値と一致する結果を得つつある。既に計算が終わったケースでは、タンパク質トリプシンと benzamidine の複合体で -6.1 kcal/mol (実験値: -6.4 kcal/mol)、FKBP-FK506 の複合体で -13.6 kcal/mol (実験値: -12.9 kcal/mol) という実験値をよく再現する結果を得た。このことは我々が開発した計算法が様々な複合体の結合親和性を評価するために利用できることを示したものである。

さらに酵素-阻害剤系でも同様の計算を行い、PaCS-MD/MSM をより大きなタンパク質-タンパク質複合体の評価にも応用できることを示した。具体的には解離に実時間で数時間を要するバルネースとバースター複合体の解離をわずか数十ナノ秒の PaCS-MD でシミュレーションし、MSM で標準結合自由エネルギーの実験値 (-16.8 kcal/mol) をほぼ再現することができた。

このように PaCS-MD/MSM を用いたタンパク質複合体の親和性の解析は、タンパク質-低分子複合体だけでなく、タンパク質とフレキシブルなペプチド、タンパク質とタンパク質の複合体にも応用できることを実証できた。

④ハイブリッドQM/MM自由エネルギー法 (再委託先: 国立大学法人京都大学大学院理学研究科化学専攻 林重彦)

本課題では、タンパク質の薬剤分子結合に対して、量子化学的手法を用いたハイブリッド非経験的QM/MM自由エネルギー法に基づく高精度な結合エネルギー計算法の開発及び適用を行う。令和元年度は、タンパク質の薬剤結合に対する非経験的QM/MM自由エネルギー法に基づく高精度な結合

エネルギー計算法の開発を行うと共に、量子化学的に扱う QM 領域の構造揺らぎを考慮するために必要な MM/MM 法の開発を行った。

まず、薬剤耐性の分子機構の理解に向けて、薬剤耐性が深刻である HIV プロテアーゼに対して、薬剤分子結合と酵素活性に関する解析手法の開発を行った。HIV プロテアーゼの機能阻害をする薬剤分子は、プロテアーゼ活性の反応遷移状態を模した遷移状態アナログ分子である。このような遷移状態アナログ分子結合による機能阻害に対する薬剤耐性の発現には、アミノ酸変異における薬剤分子結合能の低下に加えて、酵素活性の保持も必要な条件となる。従って、それらのアミノ酸変異による薬剤分子結合能と酵素活性の両者を同時に解析することにより、薬剤耐性の分子機構の理解や耐性に強い薬剤設計への指針が得られると期待される。

本課題では、まず、HIV プロテアーゼの Indinavir 薬剤分子結合状態に関する研究を行った。図 1 に、HIV プロテアーゼの酵素反応スキームと代表的な薬剤分子である Indinavir の分子構造を示す。ペプチド鎖を加水分解する HIV のプロテアーゼ反応においては、 sp^2 平面構造を有するペプチドのカルボニル炭素原子に水分子が攻撃し、活性中間状態として sp^3 構造を有するジェミナルジオールが生成すると考えられている。Indinavir では、対応する結合部位が、活性中間状態の切断部位を模した sp^3 構造を有する第二級アルコール構造となっている。HIV プロテアーゼはホモダイマーであり、そのインターフェースにある酵素反応部位に触媒活性に関わる二つのカルボン酸側鎖 (Asp25) が存在す

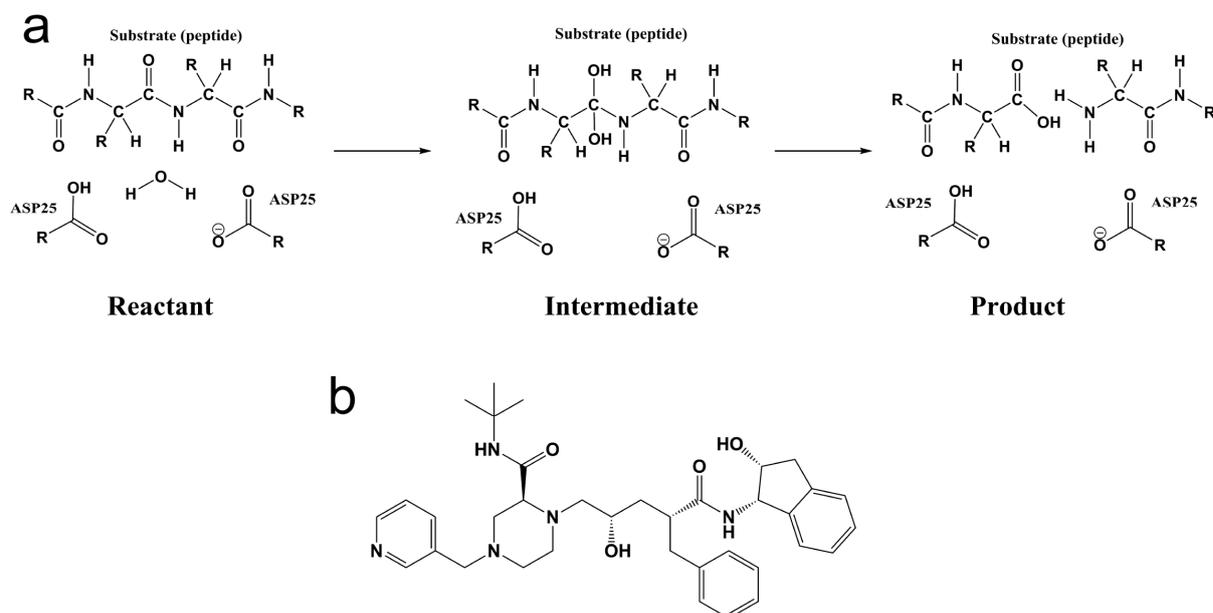


図 1 HIV プロテアーゼの酵素反応スキーム (a) と Indinavir 薬剤分子 (b)

る (図 2a)。このカルボン酸側鎖のプロトン化状態は、薬剤結合能に大きく影響を与えると考えられるが、そのプロトン化状態に関する実験的情報は得られていない。

そこで、まず、非経験的 QM/MM 自由エネルギー法を用いて、Indinavir 結合状態における触媒活性カルボン酸側鎖のプロトン化状態の決定を行った。これまでの研究で、天然基質分子に対しては、

二つの触媒活性カルボン酸側鎖の内の一つがプロトン化していることが明らかになっており、アナログ薬剤分子に関しても同様に一つのカルボン酸側鎖がプロトン化していると考えられる。

昨年度までの研究で、野生型と薬剤耐性変異の入った HIV プロテアーゼのそれぞれに対して、二つのプロトン化状態の QM/MM RWFE-SCF 法を用いた自由エネルギー構造最適化計算を行い、それらのプロトン化状態間の自由エネルギー差を自由エネルギー摂動法により計算した。

本課題で開発を行っている非経験的 QM/MM 自由エネルギー法における自由エネルギー構造最適化手法である QM/MM RWFE-SCF 法では、分子力学 (MM) による古典的力場を用いる分子動力学 (MD) シミュレーションによるタンパク質環境の構造サンプリングと、その構造サンプルにより得

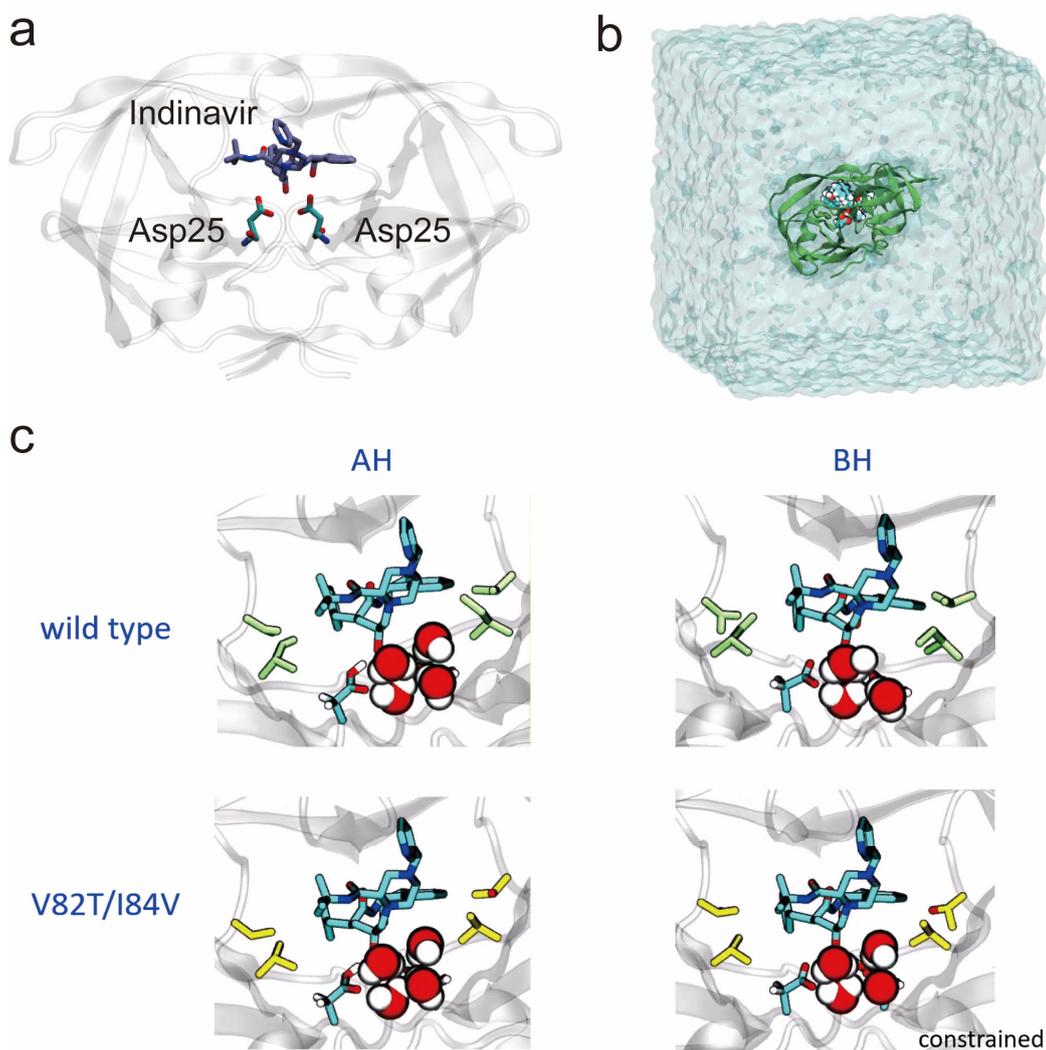


図 2 HIV プロテアーゼの Indinavir 結合状態。a : Indinavir 薬剤分子と触媒活性カルボン酸 (Asp25)。便宜上、左及び右の Asp25 それぞれ A 及び B とする。b : シミュレーション系。QM 領域は vdW 表示で描かれている。c : QM/MM 自由エネルギー最適化構造。V82T/I84V の BH 型の自由エネルギー最適化では自発的なプロトン移動を止めるために、Asp25 とプロトンの間の距離を拘束している。

られる自由エネルギー曲面上での活性部位の非経験的量子化学 (QM) 法を用いた構造最適化計算の繰り返しにより、自由エネルギー構造最適化が行われる。シミュレーション系は、水環境中に置かれた Indinavir 薬剤分子が結合した HIV プロテアーゼの周期境界条件系であり、約 76,000 原子で構成される (図 2b)。MD シミュレーションは Amber プログラムパッケージを用いて行った。シミュレーション系の中で、Indinavir 薬剤分子と触媒活性カルボン酸側鎖の計 105 原子を QM 領域として取り扱った。電子状態計算は密度汎関数法 (B3LYP-D3 汎関数) を用いた。基底関数は 6-31G** (Asp のカルボン酸の酸素原子及びそれに結合している炭素原子は 6-31+G**) であり、基底関数数は 1,059 である。QM/MM 構造最適化計算は、独自の QM/MM インターフェースを組み込んだ GAMESS プログラムパッケージを用いた。

二つの触媒活性カルボン酸側鎖のそれぞれどちらかがプロトン化した二つのプロトン化状態 (便宜上 AH と BH と呼ぶ。それぞれ図 2a の左及び右の Asp25 がプロトン化された状態である) のそれぞれに対して自由エネルギー構造最適化計算を行った。昨年度の計算で野生型タンパク質に対しては自由エネルギー最適化構造が得られている (図 2c)。また、V82T/I84V 変異体に対しても、AH 型の自由エネルギー最適化構造が得られている (図 2c)。BH 型に関しては、自由エネルギー構造最適化計算において、AH 型の自発的なプロトン移動を観測しており、AH 型のプロトン化状態を有することが決定されていた。本年度は、AH 型と BH 型の間の自由エネルギー差を求めることにより、プロトン化状態の安定性とその分子機構を明らかにした。

まず、V82T/I84V 変異体のプロトン化状態間の自由エネルギー差の計算に必要となる BH 型状態を代表する最適化構造を得るために、Asp25 とプロトンの間の距離に拘束を課し自発的なプロトン移動を止めた自由エネルギー構造最適化計算を行い、最適化構造を得た。構造最適化に要した MD シミュレーション時間は 950 ns であった。次にプロトン化状態間の自由エネルギー差を計算するために、Bennett acceptance ratio (BAR) 自由エネルギー摂動法を用いた MD 計算を行った。二つの自由エネルギー最適化構造の QM 構造と電荷を 20 分割点で線形補間し、片道 1 μ s の両方向計 2 μ s の自由エネルギー摂動計算を行った。

表 1 に自由エネルギー差 ($\Delta F_{\text{QM/MM}}$)、及びその要素である QM エネルギーの寄与 (ΔE_{QM}) と QM-MM 相互作用・MM 構造変化の寄与 ($\Delta F_{\text{QM-MM,MM}}$) を示す。天然型に関しては、プロトン化状態間の自由エネルギー差の平均は 2.8 kcal/mol となり AH 型が熱力学的に安定であると決定された。また V82T/I84V 変異体に対しても、天然型と同じく AH 型が熱力学的に安定であり、自由エネルギー差は平均として 4.6 kcal/mol と計算された。自由エネルギー計算値の行き帰りの経路の差はそれぞれ 0.6 kcal/mol であり高い統計的収束性が得られている。プロトン化状態間の自由エネルギー差は変異体の方が天然型に比べて大きく、自由エネルギー構造最適化計算における自発的なプロトン移動の挙動と一致している。図 2c に表されているように、薬剤分子の結合状態では主に B の Asp25 が水分子の水和を受けている。従って、BH 型のプロトン化状態の方が、Asp25 がプロトン化

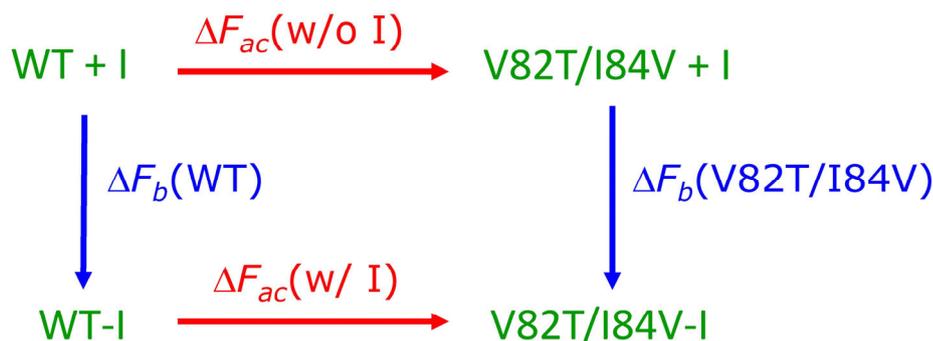
	ΔE_{QM}	$\Delta F_{QM-MM,MM}$	$\Delta F_{QM/MM}$
native			
forward	-18.1	+20.7	+2.6
backward		+21.1	+3.0
mutant			
forward	-13.2	+18.1	+4.9
backward		+17.5	+4.3

表 1 プロトン化状態間の自由エネルギー差。自由エネルギー差 ($\Delta F_{QM/MM}$)、及びその要素である QM エネルギーの寄与 (ΔE_{QM}) と QM-MM 相互作用・MM 構造変化の寄与 ($\Delta F_{QM-MM,MM}$)。正の値が BH 型の自由エネルギーが大きいことを表している。

されることにより静電ポテンシャルが弱くなるため、大きな QM-MM 相互作用の寄与 ($\Delta F_{QM-MM,MM}$) の増大を示している。一方、QM エネルギーの寄与 (ΔE_{QM}) はそれを打ち消すように大きな減少となっている。変異体では、天然型に比べてそれぞれの増大・減少幅が小さくなっているが、 ΔE_{QM} の減少幅の減少 (4.9 kcal/mol) に比べて $\Delta F_{QM-MM,MM}$ の増大幅の減少 (-3.1 kcal/mol) が小さく、変異体のプロトン化状態間の自由エネルギー差の増大を与えている。これは図 2c に示されているように、天然型では Ile84 の側鎖により BH 型のときの B のプロトン化された Asp25 の水和の水分子の数が減り疎水性相互作用が現れるが、V82T/I84V 変異体では Ile84 がより側鎖の小さいバリンに、また Val82 がより極性の高いスレオニンに置換されているため、B の Asp25 のプロトン化による水和分子数の減少が起らず、疎水性相互作用の欠如により $\Delta F_{QM-MM,MM}$ の増大幅が大きいままであることによる。一方、天然型の方が上記の BH 型の水和水分子との静電相互作用が弱いため、QM-MM 相互作用以外の ΔE_{QM} の安定化の寄与は大きくなる。このように、非経験的 QM 計算と長時間の MD シミュレーションに基づく高精度な解析により、プロトン化状態間の自由エネルギー差を決定している分子機構を明らかにすることに成功した。

次に、決定されたプロトン化状態を用いて、天然型と V82T/I84V 変異体における Indinavir の結合自由エネルギー差の計算を行った。生化学実験では、V82T/I84V 変異導入により Indinavir の解離定数が 70 倍増加 (結合自由エネルギーが 2.5 kcal/mol 減少) し、薬剤耐性の要因の一つになっている。天然型と V82T/I84V 変異体の結合自由エネルギーの差を求めるために、QM/MM 自由エネルギー構造最適化計算で得られた最適化構造に対して、alchemy 法を用いた自由エネルギー摂動法による自由エネルギー計算を行った。天然型と変異体の結合自由エネルギー差の計算は、それぞれの結合自由エネルギーを計算し比較するのが直截的であるが、薬剤分子の結合・解離過程は非常に大きな構造変化を伴い、更に結合部位の水和状況の変化による Asp25 のプロトン化状態の遷移も起きうるため、直截的なスキームでは統計的収束性に困難が生じる。そこで、図 3 の熱力学的サイクルに示されるように、天然型と変異体の薬剤分子結合エネルギーの差は、薬剤分子結合状態と非結合状態のそれぞれにおける変異導入のエネルギーの差に置き換えられるため、後者を alchemy 法を用いて計算した。

自由エネルギー計算では、QM 最適化構造の変化に起因する BAR 計算と変異導入に起因する alchemy 計算を分けて行い、前者は 10 分割で 100 ns、後者は AMBER の alchemy ルーチンを用い



$$\Delta F_{ac}(w/o I) + \Delta F_b(V82T/I84V) = \Delta F_b(WT) + \Delta F_{ac}(w/ I)$$

$$\Delta\Delta F_b = \Delta F_b(V82T/I84V) - \Delta F_b(WT) = \Delta F_{ac}(w/ I) - \Delta F_{ac}(w/o I)$$

図 3 天然型と V82T/I84V 変異体の薬剤分子結合自由エネルギーの差 $\Delta\Delta F_b$ の計算に関する熱力学的サイクル。結合自由エネルギー ΔF_b の差は alchemy 計算による変異導入の自由エネルギー ΔF_{ac} の差に置き換えられる。

て 20 分割で 200 ns の計算を行った。非薬剤分子結合状態に対しては、薬剤分子の解離により結合部位が大きく水和するため、二つの Asp25 が両者とも脱プロトン化している状態がもっともらしい状態であるが、それに加えて片方の Asp25 がプロトン化している状態に対しても計算を行った。その結果、変異体の結合自由エネルギーは天然型のそれに比べて、非薬剤分子結合状態に対して二つの Asp25 が両方とも脱プロトン化している場合は 3.8 kcal/mol、片方の Asp25 がプロトン化している場合は 5.0 kcal/mol 増加すると計算され、実験値と良い一致を得ることに成功した。上記のように、薬剤分子結合状態ではプロトン化状態が HA 型であり、変異部位を含む A の Asp25 周辺は疎水性相互作用が支配的になっている。従って、V82T/I84V 変異による疎水性の減少により結合自由エネルギーが大きく減少すると考えられる。現在論文を執筆中である。

更に、天然基質分子 (KARVLGAM 配列のペプチド鎖) に対するプロテアーゼ反応の解析を行った。シミュレーション系や計算条件は上記の薬剤分子の計算と同様である。QM 領域として切断部位のアミノ酸 (Leu と Ala) と二つの触媒カルボン酸側鎖、及び加水分解反応に関わる水分子を含む 3 個の水分子の合計 65 原子をとり、基底関数数は 587 である。昨年度の計算で、反応の活性化状態のエネルギー計算には M06-2X 汎関数を用いることが示されたため、本年度の計算では M06-2X 汎関数を用いて QM/MM 自由エネルギー構造最適化を行った。その結果、始状態と活性中間状態 I に対して自由エネルギー最適化構造を得た。更に、その後に現れる活性中間状態 II 及び III に対しても構造最適化計算を行い、計算途中ではあるがほぼ収束に近い近似的な構造を得た。図 4a にそれらの構造を示す。活性中間状態 I は加水分解反応を引き起こす水分子がカルボニル炭素原子に求核攻撃したジェミナルジオール状態、II はプロトン移動を経てアミドの窒素原子が sp^3 混成しプロトンが付加したアミン基となった状態、更に III はペプチド結合が解離した状態である。

更に図 4b にそれらの活性中間状態と Indinavir の化学構造の比較を示す。活性中間状態で特徴的な水酸基及びアミン基の位置を比較すると、Indinavir に近い化学的構造を有する状態は活性中間状態 II であることが分かる。そこで、図 4c に示すように、活性中間状態の水酸基・アミン基と

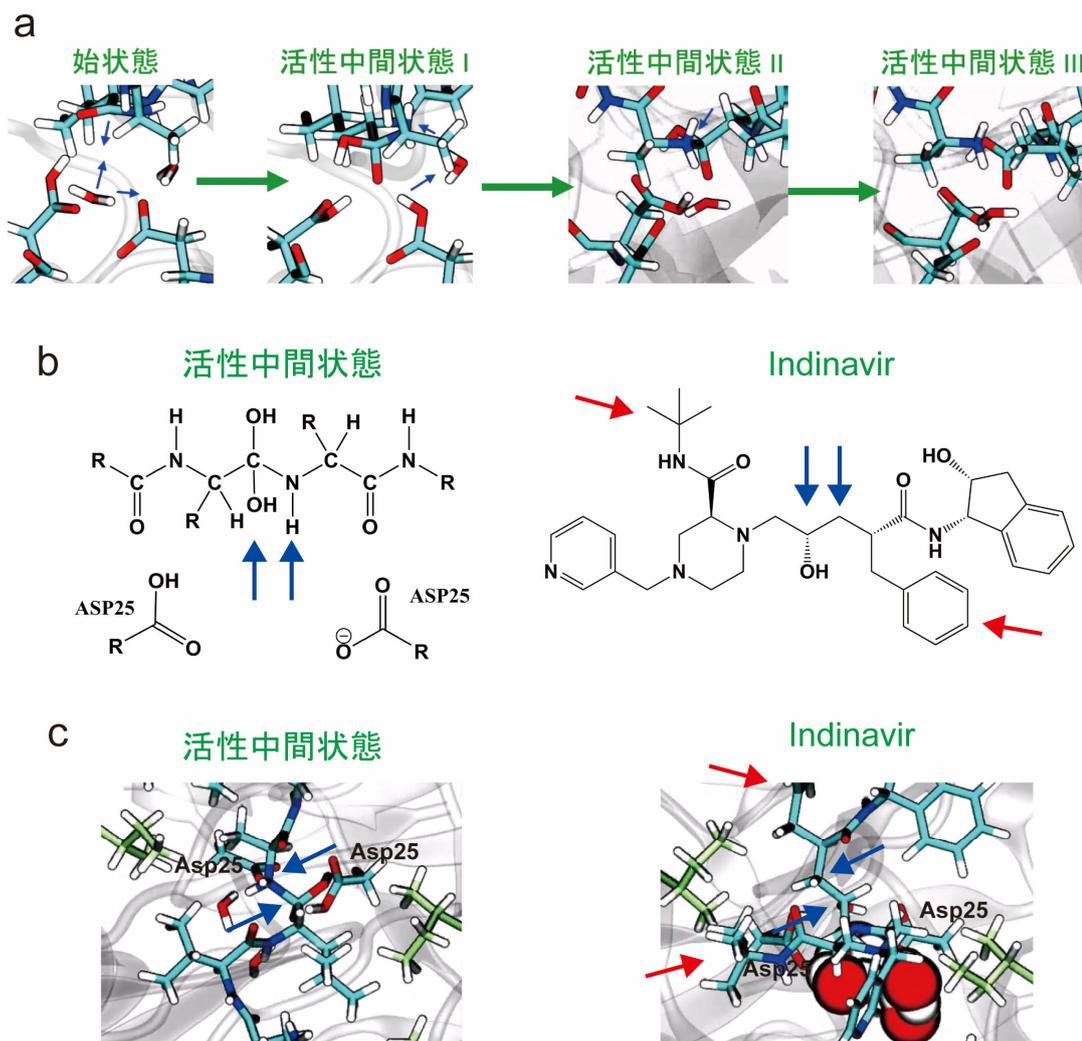


図 4 天然基質分子の反応経路と Indinavir との比較。a 反応始状態及び活性中間状態の自由エネルギー最適化構造。活性中間状態 II 及び III は計算途中の構造。b 天然基質分子の活性中間状態と Indinavir の化学構造の比較。青い矢印が天然基質分子の触媒活性部位の水酸基とアミン基、及び Indinavir における対応する部位。Indinavir の赤い矢印はタンパク質との疎水性相互作用部位。c 天然基質分子の活性中間状態と Indinavir の分子構造の比較。矢印は b で示された部位に対応している。

Indinavir の対応する部位を重ねた構造を比較すると、触媒活性を有する Asp25 のカルボン酸の位置が大きくずれており、Indinavir では天然基質分子の触媒活性を与える相互作用が完全には保持されていないことが分かった。一方、Indinavir では上記のように、触媒活性部位とは異なる薬剤分子の修飾基とタンパク質薬剤耐性変異部位との疎水性相互作用が結合に重要となっており、薬剤耐性変異導入における薬剤分子の結合エネルギーの大幅な低下及び触媒活性の保持を良く説明することが分かった。

更に、光受容体タンパク質であるアニオンチャネルロドプシンと LOV に対して、光活性化状態の QM/MM 自由エネルギー構造最適化計算を行い、性能評価を行った。LOV に関しては、光活性化による発色団分子とタンパク質側鎖の付加体生成により、離れた位置にあるシグナル伝達ヘリッ

クスの顕著な構造変化を明らかにすると共に、それと相関する発色団分子の構造変化を明らかにした。この構造変化は分子力場では記述が困難であり、QM/MM 自由エネルギー法の優位性を明らかにした。現在、論文を執筆中である。

また、リガンド分子の構造揺らぎの効果を計算するための MM/MM RWFE 法の開発を進めた。本年度は、QM/MM RWFE-SCF 法の計算コードに基づき、MM-MM 相互作用のルーチン及び MM 構造最適化のルーチンの開発を行った。

⑤粗視化モデリング（再委託先：国立大学法人京都大学大学院理学研究科生物科学専攻 高田彰二、国立大学法人名古屋大学 篠田渉）

令和元年度は、これまでに開発した大規模なタンパク質・核酸・リン脂質複合系に対する高精度な粗視化モデルを適用して、クロマチン動構造、転写制御因子とクロマチン構造の共役、生体膜とタンパク質の動態についての計算を実施した。具体的には、酵母の全遺伝子の転写開始サイトの前後に存在する -1、及び +1 ヌクレオソームの動的安定性と遺伝子発現特性の関係、山中因子 Oct4, Sox2 のヌクレオソーム結合動態の解析、膜タンパク質による生体膜の動構造変化等についての、細胞生物学的実証研究を行った。

まず、平成 29、30 年度と継続開発してきた生体膜を構成するリン脂質の高精度な粗視化分子モデル、リン脂質とタンパク質との相互作用モデルについて、代表的な不飽和リン脂質 POPC、飽和リン脂質 DPPC のパラメータチューニングを完成させた（図 1）。標的のリン脂質に対して、膜の厚み、膜面のリン脂質密度、脂質の秩序パラメータの 3 物性を同時に満たすパラメータセットを得ることに成功した。特に、前年度に見つかった問題点、長いテストシミュレーション中で稀に生体 2 重膜に穴が形成される問題、に関して、リン脂質の頭部と尾部の排除体積パラメータの比率を調整することによって解決できることを見出した。完成した粗視化リン脂質モデルを用いたテスト計算では、温度 30°C において、a) ランダム配置から自発的にリン脂質 2 重膜を形成する（図 1 中央下）、b) POPC 膜が液相様の性質を示す（側方拡散と秩序パラメータ）、c) DPPC 膜がゲル相様の性質を示す、d) POPC、DPPC 膜ともに温度変化に際して相転移を示す、e) 安定なベシクルを形成できる、f) POPC・DPPC 混合系は相分離を示す等、実験データと符合する物性を有することを確認した。さらに、リン脂質膜への膜タンパク質の挿入の実証研究を実施し、幾つかの膜貫通ヘリックスについて実験と誤差範囲で整合性のある傾き角を得ることを確認した（Ugarte & Takada, 論文投稿中）。

さらに、核酸・タンパク質系の相互作用については、平成 29 年度から開発を続けてきた PWMcos と非特異的な方向依存性ポテンシャルを、後述の転写因子・ヌクレオソーム複合系に適用する中で、細部のチューニングを行った。

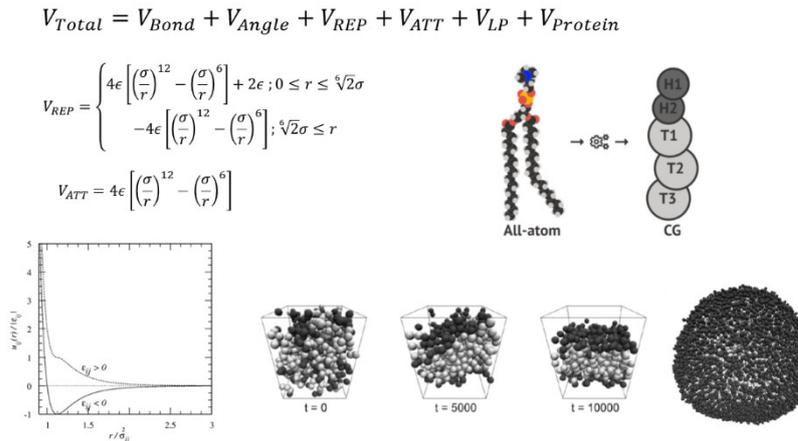


図1 リン脂質・タンパク質の粗視化モデル開発とテスト計算。左) リン脂質・タンパク質間の相互作用モデル。右上) 代表的な不飽和リン脂質 POPC の原子モデルから粗視化モデルの構築。中央下) ランダムな配置から自発的なリン脂質 2 重膜の形成に成功。右下) 安定に得られたベシクル小胞。

細胞生物学的実証研究として、これまでに開発してきた粗視化モデルを適用して、クロマチンの基本構成単位であるヌクレオソームの形成・分解についてシミュレーションを実施し、ヌクレオソーム形成・分解の動態を明らかにした。真核生物の染色体・クロマチンは、ヌクレオソームが数珠状に連なったポリマーであり、遺伝子発現等に際して、特定の位置のヌクレオソームの分解・再形成が重要な過程となる。

本研究では、まず既存の実験データを踏まえて、ヌクレオソームを構成するヒストン同士の相互作用の粗視化パラメータを決定した。次に、塩濃度依存的に、ヌクレオソームが分解することを確認した。その結果を踏まえて、実験的に詳しく調べられている 601 ヌクレオソームについて、中間的な塩濃度にあたる NaCl 400 mM 条件下で、数百本のシミュレーションを実行し、そのデータをもとにマルコフ状態モデルを構築することに成功した。得られた自由エネルギーランドスケープは、11 個の状態とそれらの間の遷移速度定数で構成される (図 2)。11 個の状態は、ヌクレオソーム、L ヘキサソーム、R ヘキサソーム、テトラソームを含み、さらにヒストンの配置、DNA との相互作用の違いによってサブグループに分かれる。601 配列の非対称性を反映して、非対称なランドスケープとなった。また、最近の実験で示唆されていた部分的開状態に相当すると思われる中間状態も発見した。これらの情報は、クロマチン動態の基礎として重要なヌクレオソーム形成・分解に新しい枠組みを与えるものとなった。

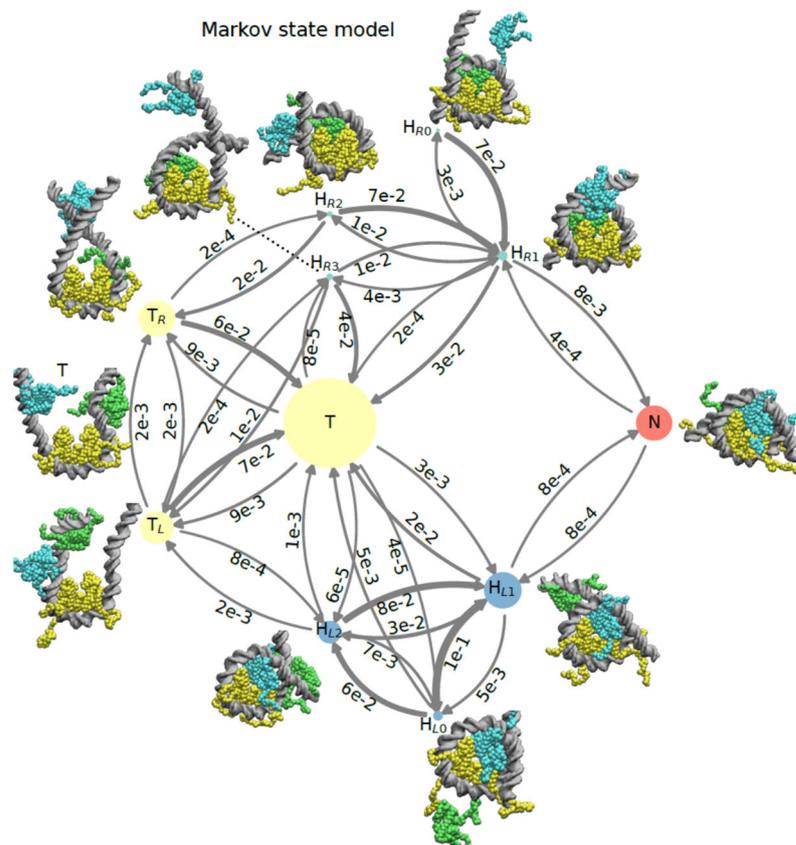


図2 塩濃度 400 mMにおける 601ヌクレオソームの形成・分解過程の自由エネルギーランドスケープ。Nはヌクレオソーム、Hはヘキサソーム、Tはテトラソームをそれぞれ表す。下付きの添え字はサブタイプを表す。(Brandani, Tan, & Takada, 論文準備中)

601ヌクレオソームは実験の便宜のために特に位置特異性を高めたものであるが、実際のゲノム配列上のヌクレオソームの動態とは異なる可能性がある。また、各遺伝子座のヌクレオソームは、その遺伝子の発現活性と関係があると期待される。とくに、転写開始地点 (TSS) の直前 (-1 とよぶ)、直後 (+1 とよぶ) のヌクレオソームに着目する。酵母ゲノムの全遺伝子座について、-1 および+1ヌクレオソームを選び、その位置が実験的に知られている場合について、粗視化 MD シミュレーションを行った。2130 個の-1ヌクレオソームについての網羅的 MD シミュレーションから、安定性の非対称性、遺伝子発現と相関する安定性を見出した (図3)。-1ヌクレオソームの下流側末端の安定性とその遺伝子の活性に有意な負の相関がみられた。すなわち、下流側末端が不安定なものほど遺伝子活性が高い傾向がある。これは、既存の実験データとも矛盾しない結果であり新しい知見を与えるものである。一方、+1ヌクレオソームについては、全体に安定性が高く、顕著な特徴は見いだせなかった。

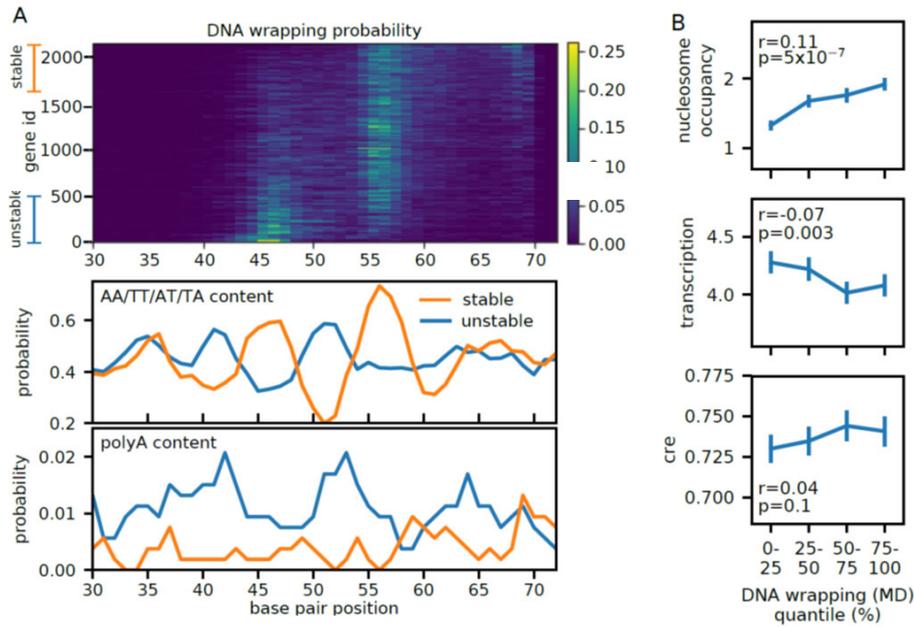


図3 酵母遺伝子座の-1ヌクレオソームの安定性と遺伝子発現。A) 2130個の-1ヌクレオソームの下流側末端の安定性(上段)。安定なもの(オレンジ)と不安定なもの(青色)でA/Tの出現パターンが逆位相になる(中段)。また、不安定なものは安定なものよりpolyAの含有量が優位に多かった。B) ヌクレオソーム安定性と転写活性の相関 (Brandani & Takada, 論文準備中)

2例目の細胞生物学的実証研究として、昨年度に引き続いて、クロマチンとタンパク質の競合について、転写因子のヌクレオソーム上への結合過程についての粗視化 MD シミュレーションを実施し、結合モードを分析した。染色体は、無数のヌクレオソームが数珠状に連なってできている。ヌクレオソーム内のDNAには多くのタンパク質(特に転写因子)がアクセスできないと考えられている。従って不活性なクロマチン状態に相当する。一方、あるクラスの転写因子などのタンパク質は、この不活性なクロマチン状態にもアクセスし、それを活性化する働きを持つことが期待される。ヌクレオソーム混雑化の染色体にアクセスし、より開いたDNA状態を作る効果を持つ転写因子は、パイオニア転写因子とよばれ、その動態に注目が集まっている。例えば、山中因子として有名な Oct4 及び Sox2 はパイオニア転写因子の例であり、パイオニア転写因子としての性質を持つことが細胞初期化に必要な可能性がある。

前年度の研究から、位置特異性の高い601ヌクレオソームの場合、Oct4は2つの結合ドメイン POU_{HD}, POU_Sをリンカーでつなぐという特徴によって、ヌクレオソーム上の並列する2つのDNA2重らせんを橋渡しするモード、結合ドメインの一つがヒストン側方に結合するモードなど多様な結合を呈した。対照的に、Sox2はより選択的な結合を示した。つまり、Sox2結合モチーフがちょうどヌクレオソームの外側を向いた位相を持っているときに限り、Sox2はそのモチーフに結合することが分かった。

この知見をもとに本年度、ヒトゲノム上のヌクレオソームへの Oct4 と Sox2 の結合を解析した。ヒト *Lin28b* は、胚性細胞において重要な働きをする遺伝子である。その遺伝子制御領域には、

Oct4, Sox2 両方の結合モチーフをもつヌクレオソームがあり、実験的にも詳しい解析の対象となっている(以下ことを Lin28b ヌクレオソームと呼ぶ)。

転写因子が結合しない状態の Lin28b ヌクレオソーム最安定配置において、Sox2 コンセンサス配列は内側に隠れており結合できない状態であった。MD シミュレーションの結果、Sox2 は擬コンセンサス配列に結合した (図 4 左)。興味深いことに、Sox2 の擬コンセンサス配列への結合によってヌクレオソーム DNA が屈曲し、それと共役してヌクレオソーム DNA が 2-3 塩基対の回転共役スライディングを起こし、その結果、Sox2 コンセンサスモチーフが外側に露出した (図 4 右)。そこで、Sox2 をもう 1 分子、系に追加して MD シミュレーションを行ったところ新しい分子はコンセンサス配列に結合することができた。

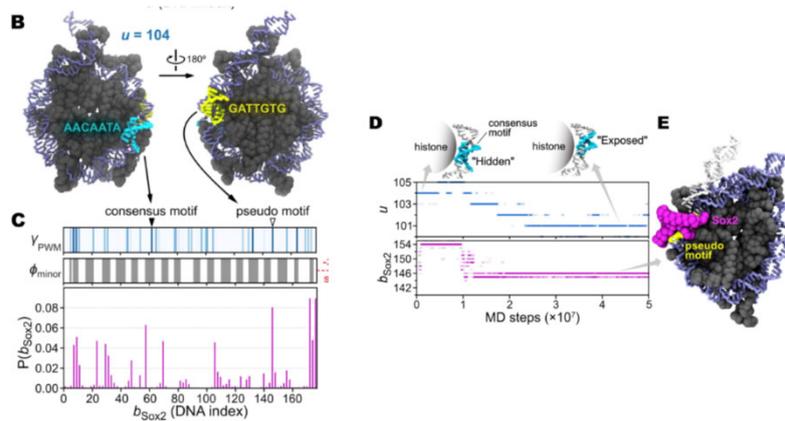


図 4 Lin28b ヌクレオソームへの Sox2 結合と状態変化。左上) Sox2 コンセンサス配列 (シアン) は内部に埋もれ、擬コンセンサス配列 (黄) が外側を向く。左下) Sox2 は、コンセンサス配列には結合できず、擬コンセンサス配列に結合した。右) Sox2 の擬コンセンサス配列への結合 (b_{Sox2}) と共役してヌクレオソーム DNA がスライディングした (シアン)。

次に、もう一つの山中因子 Oct4 の Lin28b ヌクレオソームへの結合を調べた。転写因子が結合しない状態の Lin28b ヌクレオソーム最安定配置においては、Oct4 はそのコンセンサス配列に結合しなかった。一方、上記のように Sox2 結合によって DNA がスライディングした後に Oct4 を加えて MD シミュレーションを行ったところ、Oct4 の結合パターンは変化し、一定の確率でコンセンサス配列に結合することが観察された (図 5)。

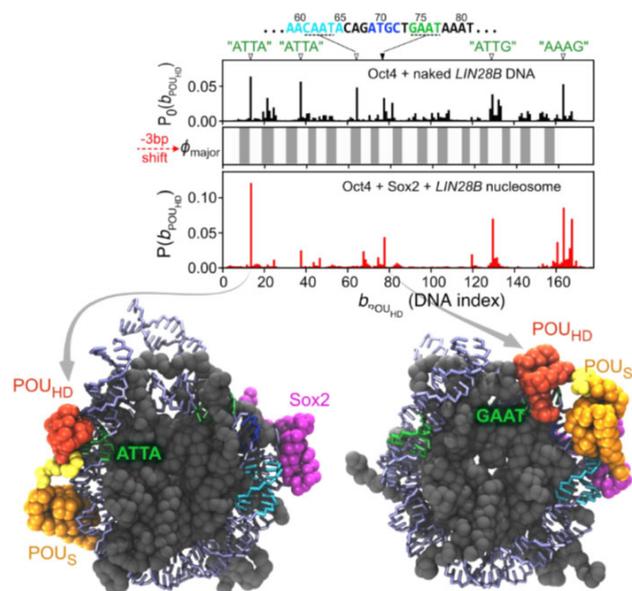


図5 Lin28bヌクレオソームへのOct4結合。転写因子が結合しない状態のLin28bヌクレオソーム最安定配置へのOct4結合パターン（上段、黒）とSox2擬コンセンサス配列結合後のヌクレオソームへのOct4結合パターン（上段、赤）。後者の構造モデル（下段）。Sox2は紫色、Oct4はその2つのドメインPOU_{HD}、POU_Sをオレンジと橙で表した。

⑥自由エネルギー計算法（再委託先：公立大学法人横浜市立大学 池口満徳）

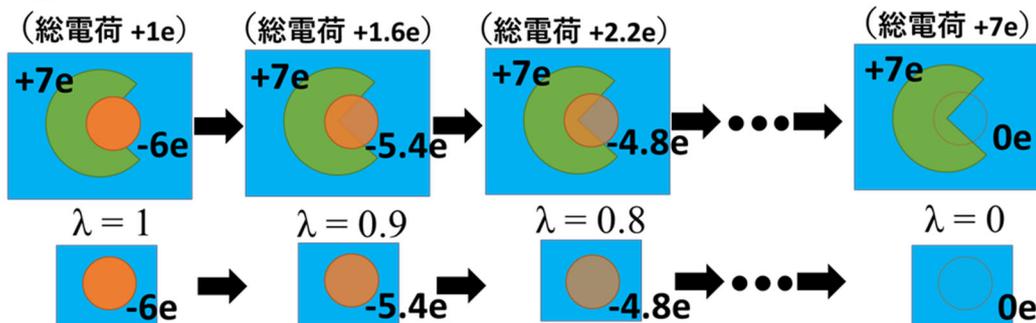
令和元年度は、医薬品候補化合物（リガンド）とタンパク質間の高効率・高精度な結合自由エネルギー計算法の拡張と性能評価を実施した。まず、平成30年度に確立した、タンパク質-リガンド間の静電相互作用に関わる有限サイズ効果の影響がない、ワープドライブ法（浴本亨、山根努、池口満徳, *J. Chem. Theory Comput.*, **14**, 6544-6559 (2018)) をフレキシブルな対象へ拡張し、汎用的な結合自由エネルギー計算法を開発した。次に、拡張したワープドライブ法と、既存の典型的な結合自由エネルギー計算法であるMP-CAFEE法を比較し、性能評価を実施した。性能評価は、電荷の大きな対象における有限サイズ効果の検証と、様々な電荷を持つ複数リガンドを用いた計算値と実験値の比較検証、の2つの検証を実施した。その結果、セルサイズを変化させると、MP-CAFEE法では数kcal/molの結合自由エネルギーの違いが生じていたが、拡張したワープドライブ法では約0.4kcal/mol程度で抑えられる事がわかった。また、様々な電荷の複数リガンドにおいては、MP-CAFEE法の結合自由エネルギー計算値では外れ値が存在し、実験値との相関がなかったが、拡張したワープドライブ法の計算値では実験値との相関がある事がわかった。

有限サイズ効果とは、周期境界条件下の分子動力学シミュレーション（MD）を用いて求めた静電溶媒和自由エネルギーが、MD中のセルサイズの大きさに依存して、計算値が変化してしまうアーティファクトである。典型的な結合自由エネルギー摂動計算では、複合体溶液系とリガンド単体溶液系それぞれで、リガンドの電荷をカップリングパラメータで線形に消していく過程が含まれているため、有限サイズ効果が生じてしまう（図1A）。各中間状態における溶質-溶媒間静電相互作用を求め、熱力学積分法によって静電溶媒和自由エネルギーを求めることができる。セルサイズの大きさを系

統的に変化させ、これまでに、小分子系 (Hummer ら)、タンパク質系 (池口ら)、リガンドタンパク質複合体系 (Rucklin ら、池口ら (平成 30 年度の成果)) において、有限サイズ効果の影響が検証されている。セルサイズの大きさを、MD の典型的な大きさから、一辺が約 100 Å の、通常は設定しない大きさのセルまで系統的に変化させたところ、中性の溶質で約 5 kcal/mol、電荷の小さな溶質で約 10 kcal/mol、電荷の大きな系で 100 kcal/mol を超える有限サイズ効果が生じることがわかった。特に、結合自由エネルギー計算の場合は、複合体溶液系とリガンド単体溶液系のそれぞれで有限サイズ効果が生じてしまい、補正がなければ、非常に大きなセルを用意する必要があることがわかった。そのように非常に大きなセルは、実際には計算できないため、有限サイズ効果の補正法を適用する必要がある。補正法の性能評価を行ったところ、電荷の大きな系であっても、有限サイズ効果を数 kcal/mol まで補正できることがわかったが、補正法の性能は、MD の典型的な大きさのセルサイズで性能が悪く、かつ、対象とする系ごとに補正値を計算する必要があるため、原理的に、有限サイズ効果の生じない計算手法が必要であることがわかった。

そこで、有限サイズ効果の補正を必要としない、ワーブドライブ法を開発し、固定構造に限定した性能評価を実施した。典型的な摂動計算 (図 1A) の検証によると、各中間状態で系の総電荷が変化してしまうことが、有限サイズ効果の原因となっていた。ワーブドライブ法では、リガンドタンパク質複合体と溶媒中のリガンドを共存させる溶液系を用い、複合体のリガンドの電荷を消した分だけ、溶媒中のリガンドの電荷が増えるようにした (図 1B)。すると、各中間状態で系の総電荷が変化しないため、セルサイズを変化させても有限サイズ効果が生じないことがわかった。また、MD で用いる典型的なセルサイズであっても、有限サイズ効果が生じておらず、計算値は、典型的な摂動法の計算値に補正を行った極限值と一致することがわかった。

(A) 典型的アルケミカル摂動法



(B) ワーブドライブ法

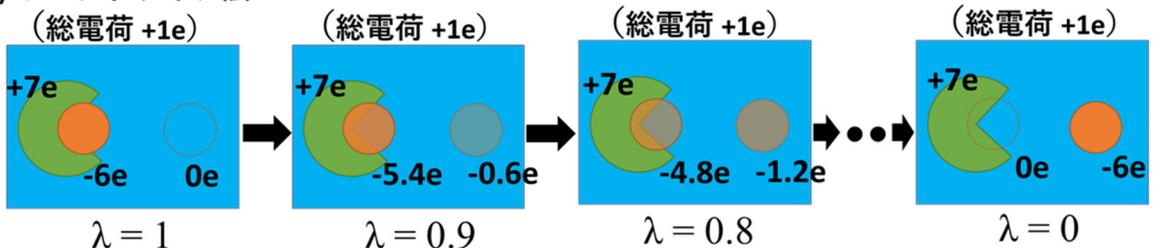


図 1 (A)既存の典型的アルケミカル摂動法と(B)ワーブドライブ法の摂動計算方法

これまで固定構造に限定し、静電相互作用の検証のみを行っていたワーブドライブ法を、フレキシブルな対象へ拡張した。ワーブドライブ法で結合自由エネルギーを求める熱力学サイクルを定義した (図 2A)。リガンドタンパク質複合体のリガンドを *site*、溶媒中の非結合リガンドを *bulk* と書く。図 2A 最上段が、結合自由エネルギー (ΔG_{bind}) であり、*bulk* のリガンドが、*site* の位置へ結合する過程である。ここで、拡張ワーブドライブ法では、3つの拘束を導入する。まず、タンパク質の重心と回転運動を拘束するため、タンパク質中の 3 原子の位置を調和ポテンシャルで拘束する (図 2A、①の×印)。次に、複合体中のリガンドとタンパク質の相対配置を拘束するクロスリンク拘束を導入する (図 2A、②の折線)。クロスリンク拘束は、Boresch らの方法を採用し、タンパク質内の 3 原子とリガンド内の 3 原子を用いて、タンパク質に対するリガンドの相対配置と向きを拘束する。最後に、*bulk* 中のリガンドに対しても、配置と向きを固定するため、空間中に固定した 3つの仮想球 (図 2A、グレーの球) と *bulk* のリガンドの 3 原子間にクロスリンク拘束を導入した。仮想球は、他の原子と相互作用しない原子で、位置は *site* と *bulk* リガンドの幾何中心間を結ぶ直線状に配置した。熱力学サイクルに従い (図 2A 半時計周り)、拡張ワーブドライブ法の結合自由エネルギー計算は、*bulk* リガンドと仮想球間にクロスリンク拘束を課す過程 ($\Delta G_{\text{cross_on}}^{\text{bulk}}$)、*site* リガンドとタンパク質間にクロスリンク拘束を課す過程 ($\Delta G_{\text{cross_on}}^{\text{site}}$)、ワーブドライブ過程 (ΔG_{warp})、*site* リガンドとタンパク質間のクロスリンク拘束を外す過程 ($\Delta G_{\text{cross_off}}^{\text{site}}$)、及び、*bulk* リガンドと仮想球間のクロスリンク拘束を外す過程 ($\Delta G_{\text{cross_off}}^{\text{bulk}}$) で実行できることがわかった。

ワーブドライブ過程 (ΔG_{warp}) は、3つの摂動計算によって求める (図 2B)。まず、ワーブ先の *bulk* リガンドの van der Waals (vdW)相互作用を導入し、*bulk* リガンドを発生させる (ΔG_{warp1} 、図 2B 上段、青色球)。次に、電荷を、*site* リガンドから *bulk* リガンドへワーブさせる (ΔG_{warp2} 、図 2B 中段、オレンジ色の移動)。 ΔG_{warp2} は、昨年度の有限サイズ効果を検証した過程である。最後に、*site* リガンドの vdW 相互作用を消し、*site* リガンドを消滅させる (ΔG_{warp3} 、図 2B 下段、青色球)。ここで、Boresch らによると、クロスリンク拘束の寄与は、クロスリンク拘束のみで、2つの分子の相対配置と向きが決定される場合に解析的に求められる。従って、 $\Delta G_{\text{cross_on}}^{\text{bulk}}$ と $\Delta G_{\text{cross_off}}^{\text{site}}$ に解析解が適用できる。また、 $\Delta G_{\text{cross_off}}^{\text{bulk}}$ では、共存系中でタンパク質からの影響が無視できると仮定すると、解析解が適用できるため、*bulk* リガンドのクロスリンク拘束では、 $\Delta G_{\text{cross_off}}^{\text{bulk}}$ と $\Delta G_{\text{cross_on}}^{\text{bulk}}$ がキャンセルアウトする。 $\Delta G_{\text{cross_on}}^{\text{site}}$ に関しては、クロスリンク拘束を外しても、リガンドはタンパク質の結合サイトに存在するため、リガンドタンパク質間の相互作用が無視できず、解析解を適用できない。そこで、 $\Delta G_{\text{cross_on}}^{\text{site}}$ は摂動計算により求める。また、共存系において、*bulk* リガンドの位置を決める寄与を体積補正項で与えると、熱力学サイクルより、結合自由エネルギーは (図 2A 時計周り)、

$$\Delta G_{\text{bind}} = -\Delta G_{\text{cross_off}}^{\text{site}} - \Delta G_{\text{warp1}} - \Delta G_{\text{warp2}} - \Delta G_{\text{warp3}} - \Delta G_{\text{cross_on}}^{\text{site}} - k_{\text{B}}T \ln(V_{\text{box}}/V^0),$$

(ここで、 V_{box} と V^0 は、MD セルの体積と標準状態の体積を示す) と表され、Boresch らの解析解を $\Delta G_{\text{cross_off}}^{\text{bulk}}$ に適用して体積補正項と整理すると、最終的に

$$\Delta G_{\text{bind}} = -\Delta G_{\text{warp1}} - \Delta G_{\text{warp2}} - \Delta G_{\text{warp3}} - \Delta G_{\text{cross_on}}^{\text{site}} + k_{\text{B}}T \ln(V^0 A),$$

(ここで、 A は解析解の力の定数等に依存する項) で計算できることがわかった。

(A) 拡張ワーブドライブ法の熱力学サイクル (B) ワーブドライブ部分の3つの摂動計算

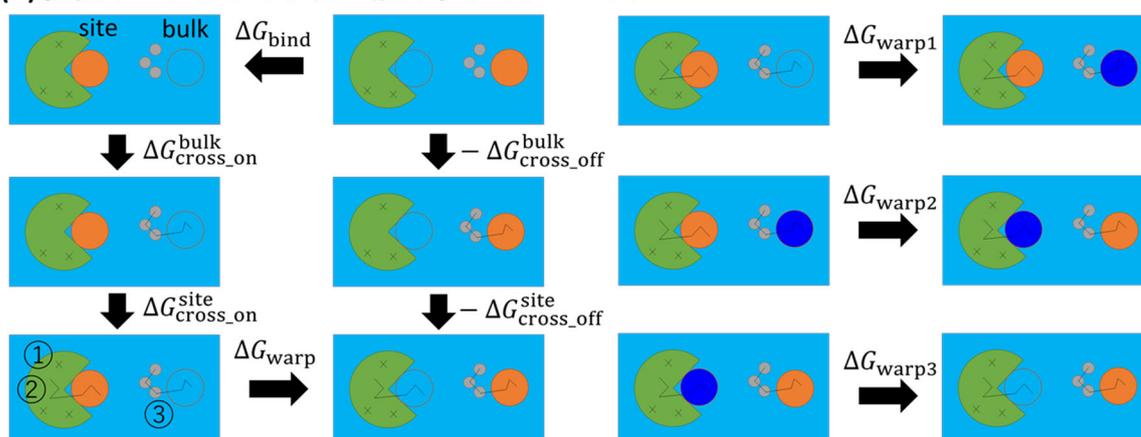


図 2 (A)拡張ワーブドライブ法の熱力学サイクルと(B)ワーブドライブ過程。オレンジは vdW 相互作用と電荷の双方が存在するリガンドを示す。青は vdW 相互作用のみのリガンドを示す。水色の円は vdW 相互作用と電荷の双方がないリガンドを示す。緑は標的タンパク質を示し、小さい円は拘束のための相互作用点を示す。

拡張ワーブドライブ法の性能を評価するために、電荷の大きなリガンドや、様々な電荷を持つリガンド群を用いて、既存手法で典型的な MP-CAFEE 法との比較解析を実施した。MP-CAFEE 法では、リガンドタンパク質複合体系 (complex 系) からリガンドを消す過程と、リガンド溶液系 (ligand 系) でリガンドを消す過程の 2 つの摂動計算を用いる (図 1 A)。先行研究同様に、標準的なプロトコルを採用し、GROMACS で摂動計算を実施した。比較を行うために、拡張ワーブドライブ法の摂動計算も GROMACS で行った。非中性電荷のリガンドでの検証を行う前に、電荷が中性のリガンドで、MD セルサイズと構造を一致させた場合に、MP-CAFEE 法と拡張ワーブドライブ法の静電相互作用寄与分の計算結果に違いがあるかを検証した。昨年度の検証によると、中性のリガンドでは有限サイズ効果が小さい。また、セルサイズが一致した状態で complex 系と ligand 系の摂動計算を行えば有限サイズ効果が効果的にキャンセルできることがわかっている。そのため、静電相互作用の摂動計算結果は MP-CAFEE 法と拡張ワーブドライブ法で一致すると想定される。共存系を用いて計算したところ、complex 系の値から ligand 系の値を引いた MP-CAFEE 法の静電相互作用の寄与 (-3.1 kcal/mol) が拡張ワーブドライブ法での計算値 (-3.1 kcal/mol) と一致することがわかった。中性リガンドの静電相互作用については、拡張ワーブドライブ法と MP-CAFEE 法で同程度の性能であることが確かめられた。

表1：結合自由エネルギー

系	8Å -layer	24Å -layer	10Å -layer	25Å -layer
手法	MP-CAFEE	MP-CAFEE	Warp	Warp
体積補正なし	-1.3	-3.5	-	-
ΔG_{bind}	-4.4	-7.2	-10.9	-10.5

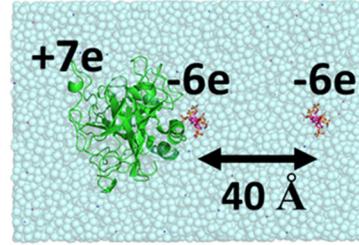


図3 ワープドライブ法の共存系

表2：MP-CAFEE法の内訳

項目	8Å-layer	24Å-layer
Complex Elec	-826.2	-830.1
Complex vdW	-20.2	-18.0
Complex dG	-846.5	-848.1
Ligand Elec	-833.1	-833.3
Ligand vdW	-12.1	-11.3
Ligand dG	-845.2	-846.7
体積補正なし ΔG_{bind}	-1.3	-3.5

表3：拡張ワープドライブ法の内訳

項目	10Å-layer	25Å-layer
$-\Delta G_{\text{warp1}}$	29.0	28.8
$-\Delta G_{\text{warp2}}$	-12.2	-9.7
$-\Delta G_{\text{warp3}}$	-34.1	-34.1
$-\Delta G_{\text{warp}}$	-17.4	-15.0
$-\Delta G_{\text{site_cross_on}}$	-4.1	-5.8
$k_B T \ln(V^0A)$	10.6	10.4
ΔG_{bind}	-10.9	-10.5

(単位 kcal/mol)

電荷の大きな対象で性能評価を実施した。ウシの酵素タンパク質トロンビンとリガンド sucrose octasulfate をテスト系に採用した。タンパク質が+7e、リガンドが-6eであり、リガンドの電荷が大きい系である。複合体の結晶構造が解かれているため、1 ns の平衡化を行った後の構造を用いて、MP-CAFEE 法の摂動計算と拡張ワープドライブ法の摂動計算を実施した。摂動計算中では、構造は固定せず、標準的なプロトコル通り、フレキシブルなままで計算を行った。有限サイズ効果の影響を検証するため、MP-CAFEE 法を、複合体を 8Å の水の層で囲む標準的なセルサイズ (8Å-layer) と、大きめの 24Å の層で囲む大きなセルサイズ (24Å-layer) で実施した。拡張ワープドライブ法は、site と bulk リガンド間距離を 40Å に設定し、共存した複合体と bulk リガンドを囲む水の層を 10Å (10Å-layer、図3) と 25Å (25Å-layer) に設定した2つの系で実施した。その結果、MP-CAFEE 法では、セルサイズに依存して、結合自由エネルギー (ΔG_{bind}) が約 3 kcal/mol 変化してしまうことがわかった (表1)。対称的に、拡張ワープドライブ法では、約 0.4 kcal/mol に抑えられることがわかった (表1)。MP-CAFEE 法の有限サイズ効果は、複合体系における静電相互作用の摂動計算 (表2、Complex Elec) で生じていた。一方で、拡張ワープドライブ法では、静電相互作用の摂動計算 ($-\Delta G_{\text{warp2}}$) で約 2.5 kcal/mol の違いがあるものの、site リガンドのクロスリンク拘束 ($-\Delta G_{\text{site_cross_on}}$) で吸収されていることがわかった (表3)。フレキシブルな対象の場合、有限サイズ効果の影響が、摂動対象のリガンドの構造揺らぎと関係していることがわかった。

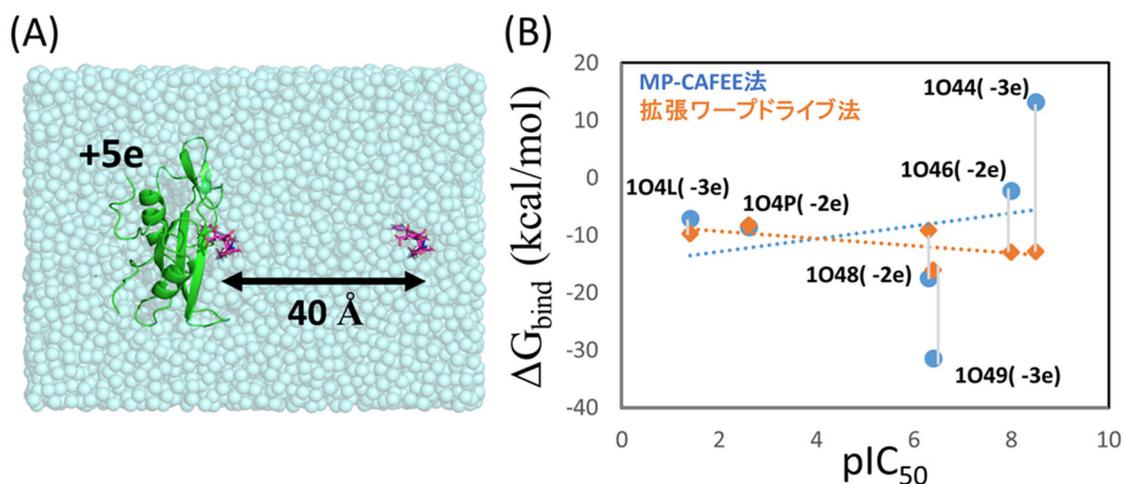


図4 (A)Src-Kinase系の共存系と(B)MP-CAFEE法と拡張ワープドライブ法の性能比較

様々な電荷のリガンド群の結合自由エネルギー計算でも性能評価を実施した。テスト系として、Src-KinaseのSH2ドメインに結合する、結合能がIC₅₀で 4×10^7 から3nMの、6つのリガンドを採用した。それぞれ複合体で結晶構造が解かれており、総電荷はタンパク質が+5e、リガンドが-2eと-3eである。MP-CAFEE法は8Å-layer系で、拡張ワープドライブ法は10Å-layer系(図4A)を用いた。その結果、MP-CAFEE法の計算値では、外れ値が存在し、実験値との相関がないことがわかった(図4B青色プロット、 $R^2 = 0.048$)。特に、電荷の大きい-3eのリガンドでは、正の値を示すものもあり、性能が悪いことがわかった。対称的に、拡張ワープドライブ法の計算値は実験値と相関があることがわかった(図4Bオレンジ色プロット、 $R^2 = 0.38$)。MP-CAFEE法で性能が悪かった、電荷-3eのリガンドにおいても、負の値を示し、他リガンドとの相関も良いことがわかった。

拡張ワープドライブ法とMP-CAFEE法の計算値を、IC₅₀から換算($k_B T \ln(IC_{50})$)した実験値と比較した。拡張ワープドライブ法の計算値は、実験値よりも低く、結合自由エネルギーを過大評価する傾向があることがわかった。実験値からの差は、pIC₅₀の小さい、結合能が弱い2つのリガンド(PDBID: 104L, 104P)では4.6 - 7.9 kcal/molと差が大きいが、結合能が強い4つのリガンドにおいては、1つ(PDBID: 1049, 7.4 kcal/mol)を除き、0.5 - 2.1 kcal/molの差であることがわかった。一方で、MP-CAFEE法の計算値は、結合能が弱い2つのリガンドについては、約5 kcal/molの違いで拡張ワープドライブ法と性能に大差がないことがわかった。しかし、結合能が強い4つのリガンドにおいては、実験値との差が8.7 - 24.9 kcal/molであり、拡張ワープドライブ法の方が実験値に近いことがわかった。これらの結果より、非中性のリガンドでは、拡張したワープドライブ法の方が、既存のMP-CAFEE法よりも性能が良いことがわかった。

(2) 次世代創薬計算技術の開発 (サブ課題 B)

①動的分子機能制御 (再委託先: 公立大学法人横浜市立大学 池口満徳)

本課題は、ポスト「京」で高効率に動作する全原子分子動力学シミュレーション(MD)を活用し、タンパク質の機能発現に重要な構造変化や動的性質の理解から、その動的機能制御を目指すもので

ある。令和元年度は、リガンド結合に際して立体構造変化が機能に重要な役割を果たす、創薬標的タンパク質ビタミン D 受容体 (VDR) について、動的アンサンブル法による動的機能解析を継続して実施した。VDR は代表的な創薬標的タンパク質であり、結合するリガンドの種類 (アゴニスト/アンタゴニスト) によって誘起される構造の違いによって機能が制御されている。平成 30 年度までに、拡張アンサンブル手法である gREST 法を用いて、異なる種類のリガンドが結合した同一の VDR 構造から、リガンドの活性に応じた構造変化をシミュレーションすることができた。本年度は、動的構造について自由エネルギー地形を利用した解析を実施し、リガンドと VDR 間の相互作用様式が、立体構造の動的構造変化の相違をもたらすメカニズムについて明らかにすることができた。さらに、近年の創薬で注目されている中分子医薬品である、シクロスポリン A (CsA) の膜透過に関わる動的構造変化の解析を継続して実施した。平成 30 年度に開発した CsA 用の新規分子力場パラメータの検証を継続して行い、新規力場の下でシミュレーションした構造群が、NMR 実験結果と矛盾のない構造分布を示すことがわかった。また、構造変化と膜透過の関係を解析するため、新規力場を用いた多数のシミュレーション群の解析を実施した。マルコフ状態遷移モデル (MSM) によって多数のシミュレーションを統合し、膜透過の確率過程モデルを構築できた。膜透過過程において、CsA の安定な領域と障壁となる領域があることがわかった。

核内受容体ビタミン D 受容体 (VDR) は、リガンド依存的に特定の遺伝子発現を制御する転写因子であり、代表的な創薬標的タンパク質である。他の核内受容体の結晶構造解析や、種々の実験により、VDR の機能発現メカニズムの初期過程は、アゴニストである活性型ビタミン D₃ の結合に伴い誘起されるヘリックス 12 (H12) 周囲の構造変化によって、コアクチベーターをリクルートできる AF-2 表面と呼ばれる特定のコンフォメーションをもった活性型構造へ構造変化する過程であると考えられている。一方で、アンタゴニストが結合した場合、H12 の配置が活性型と異なり、AF-2 表面が形成できなくなり、コアクチベーターのリクルートができないために、機能発現が阻害されると考えられている。しかし、これまでに解かれている VDR-リガンド複合体結晶構造では、リガンドの種類 (アゴニスト/アンタゴニスト) に関わらず、活性型の構造しか解かれておらず、アンタゴニスト活性を説明できていなかった (図 1A)。平成 30 年度までに、MD/MSM-SAXS 法によるアンタゴニスト複合体の溶液構造解析、及び、gREST 法によるアンタゴニスト複合体の自発的な構造変化のシミュレーションを実施した結果、アンタゴニスト活性を説明できる、溶液構造モデルを得ることに成功した。本年度は、gREST 法によって得られた溶液構造アンサンブルを用い、リガンドの種類の違いが、H12 付近の動的構造変化に及ぼす影響を解析した。

通常 MD と gREST を、活性型の結晶構造を初期構造として、1 μ s 実施した。その結果、アンタゴニスト複合体では、通常 MD (図 1B) ではサンプリングできない、活性型とは異なる構造を gREST により得られた (図 1C)。通常 MD の範囲では、代表的な溶液構造は活性型に近かった (図 1B)。対照的に、gREST の代表的な溶液構造 (図 1C、第 1 象限) では、H12 の位置が活性型とは異なっていた。H12 の位置が異なる構造的な要因を解析したところ、ヘリックス 11 末端から H12 へ至る領域 (H11-12 領域) のヘリックス構造がほどけ、不規則化することにより、H12 の N 末端部位が引っ張られ、H12 の位置がずれることがわかった。また、H11-12 領域が不規則化した要因は、ヘリックス 10/11 のキック形成と連動していることもわかった。アンタゴニスト複合体とは対比的に、アゴニスト複合体の溶液構造は、通常 MD (図 1D) と同様の、活性型構造が gREST でも安定であった (図

1E)。リガンドの種類に応じた、動的構造変化の違いを明らかにするために、リガンドとコンタクトを持つ領域と H12 の位置について解析を行った。その結果、アゴニストとアンタゴニストで、H12 付近の領域との相互作用様式が異なることがわかった。アゴニストは H12 や H11 末端とコンタクトを形成しており、H12 付近のコンフォメーションを安定化させる疎水コアを形成しているが、アンタゴニストでは疎水コアを形成できず、H12 の位置を安定化できていなかった。

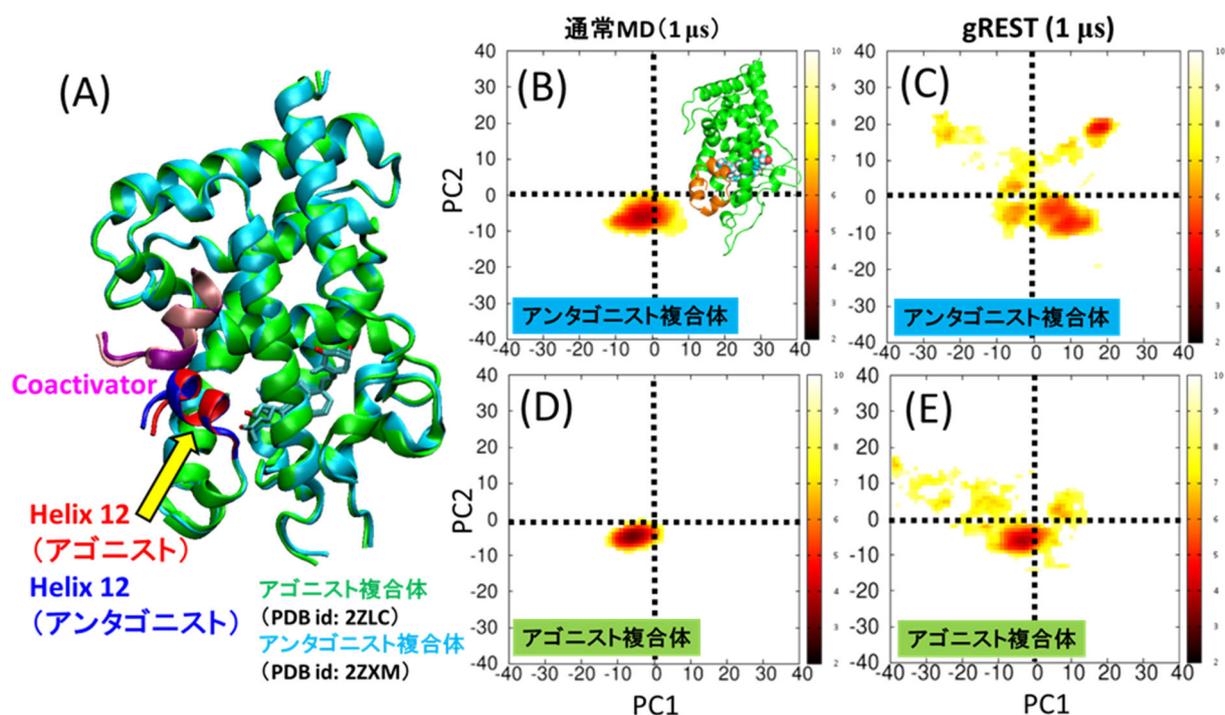


図1 (A) VDR の結晶構造と (B, D) 通常 MD と (C, E) gREST でサンプリングした構造の分布。(B-E)の色は、自由エネルギーを示している。

これらの結果より、gREST による溶液構造サンプリングによって、リガンドの種類に応じた VDR の動的構造変化の違いが、結合したリガンドの相互作用様式の違いによって自発的に誘起されることがわかった。また、gREST から得たアンタゴニスト複合体の代表的な構造の特徴や、構造変化に重要な H11-H12 領域やキックの関与は、これまでの MD/MSM-SAXS 法から得た溶液構造の特徴や構造的知見と一致しており、構造変化モデルとして妥当であることがわかった。そのため、活性型の結晶構造のみしかない状況でも、gREST を実施すれば、リガンドの活性と VDR の構造間の構造活性相関関係が得られると期待される。

シクロスポリン A (CsA) も、動的構造変化が機能に重要な役割を果たしている分子である。CsA は、*Tolypocladium inflatum* 由来の抗生物質で、免疫抑制剤として用いられる経口投与可能な中分子医薬品である。11 残基から成る環状ペプチド分子 (図 2A-C) であり、高い膜透過性とタンパク質-タンパク質相互作用のような浅く広い結合面を阻害できるといった特徴を持っている。11 残基中 8 残基は非天然アミノ酸であり (図 2A)、その内 7 残基は N メチルアミノ酸 (図 2D) である。分子量が大きいにも関わらず、高い膜透過性を持つ要因は、溶媒環境に応じて CsA が動的構造変化を起こす

ためであると示唆されている。これまでに、タンパク質（シクロフィリン）-CsA 複合体の結晶構造（PDBID: 1CWA）と、クロロホルム溶媒中の CsA 結晶構造（CSD: DEKSAN）が得られている。結晶構造において、タンパク質に結合した CsA は、環が広がった、親水性の原子を溶媒中に露出させたオープン構造をとっており（図 2B）、クロロホルム中の CsA は、環がタイトで、親水性の原子が分子内で水素結合を形成したクローズド構造をとっていた（図 2C）。そのオープン/クローズド構造が親水/疎水環境での構造である想定し、溶媒から膜中、膜中から溶媒環境へオープン/クローズド構造を変化させる膜透過モデルが提案されている（図 1E）が、実際に、脂質二重膜-水系を用いた全原子 MD シミュレーションは行われていないため、妥当性や膜透過メカニズムは明らかにされていなかった。そのため、平成 30 年度に開発した新規分子力場の検証を継続して実施し、その新規力場を用いた膜水系の全原子 MD 群から、膜透過過程を再現し、解析を行った。

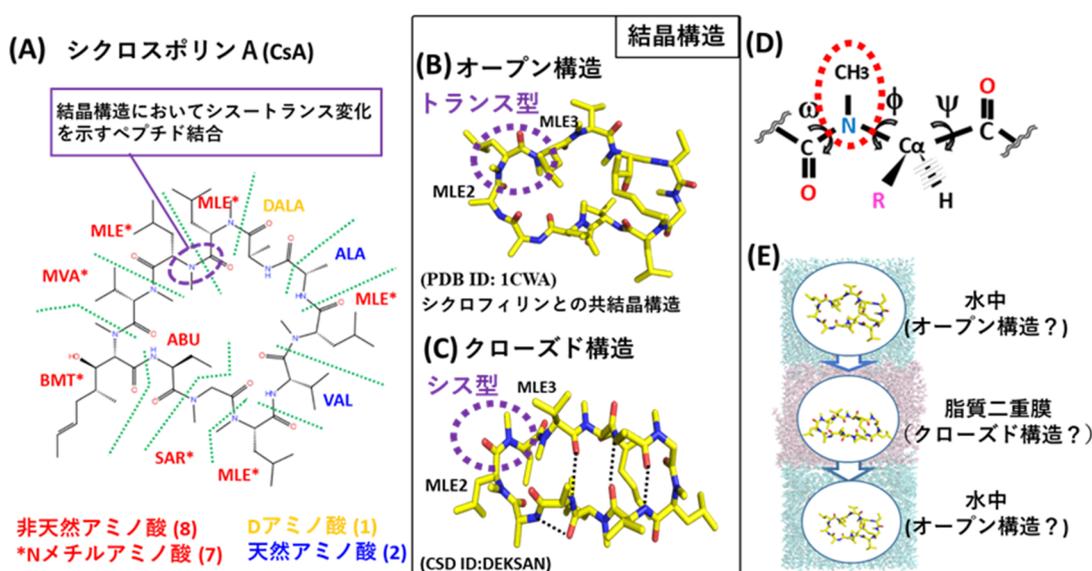


図 2 (A)シクロスポリン A の構造、(B,C) 結晶構造、(D) N メチルアミノ酸、(E)膜透過モデル

CsA 用の新規分子力場でシミュレーションされる構造群の検証を実施した。CsA は非天然の N メチルアミノ酸を持つため、通常のタンパク質用の力場が利用できない。そのため、平成 30 年度までに、CsA の CHARMM 力場の決定を行った。高精度の量子化学計算により、N メチルアミノ酸主鎖の厳密な二面角 $\phi\psi$ 空間エネルギー地形を求め、CHARMM 力場の二面角のエネルギー地形が合うように CMAP 項を決定した。そのエネルギー地形の安定領域は、結晶構造における構造分布と一致していることがわかった。本年度は、gREST により、クロロホルム中の CsA 構造をサンプリングし、力場の違いによる、構造分布の妥当性を検証した。CMAP 項を含む新規力場と、既存力場のそれぞれで、サンプリングを行った結果、新規力場から得た構造はクローズド構造の周辺（図 3A、■）に分布することがわかった（図 3A）。対照的に、既存力場では、クローズド構造だけでなく、オープン構造に近い構造（図 3B、●）まで幅広く出現することがわかった（図 3B）。溶液 NMR の実験によると、クロロホルム中の CsA の構造は、1つの状態に分布が収束する事が示されており、新規力場の構造分布の方が、実験結果と矛盾がないことが確認できた。

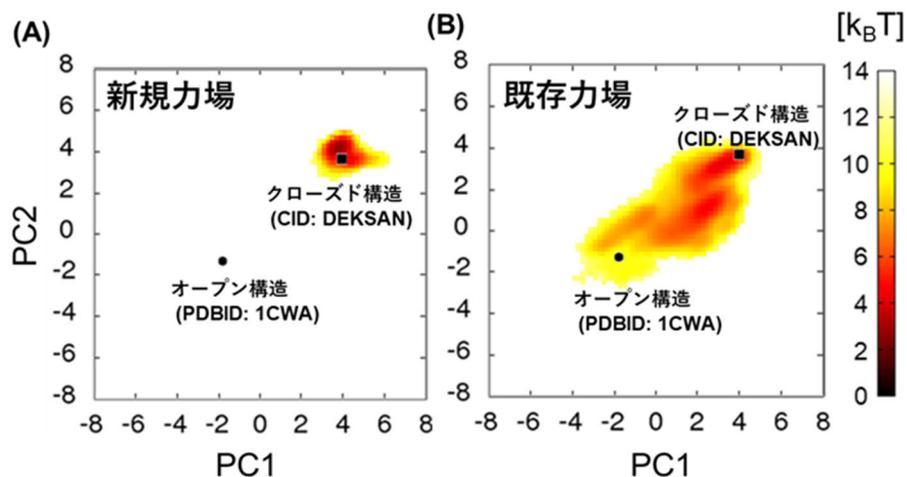


図3 (A)新規力場と (B) 既存力場を用いたクロロホルム中の構造分布

新規力場を用いて、膜水系の全原子 MD を実施した。結晶構造で得られているオープン構造とクローズド構造を初期構造とし、膜外や膜中央、及び、膜表面の様々な位置へ CsA を配置した溶液系を作成した。それぞれの溶液系で、1 - 3 μs の MD を実施し、合計で約 40 μs のトラジェクトリを得た。1 μs の MD の範囲内では、膜外から膜内への移動や、膜内から膜外への移動は起こらず、自発的な膜透過イベントは見られなかったが、クローズド構造とオープン構造の間の中間構造や、膜中で安定な位置がある事がわかった。膜外の水中では、初期構造をクローズド構造にした MD 中において、クローズド構造からオープン構造へ近い構造へ構造変化しており、中間状態の溶液構造を得ることができた。結晶構造におけるオープン/クローズド構造の比較から、2つの N メチルロイシン (図 2A-C) 間の主鎖二面角のシス/トランス状態がクローズド/オープン構造間の構造変化に寄与すると示唆されていたが、1 μs の MD の範囲では、二面角の変化は見られなかった。しかし、MD 中で、CsA の環状構造がクローズド/オープン構造とは異なる中間状態が得られたため、必ずしも、シス/トランスの状態が、クローズド/オープン構造を決定する要因ではないことがわかった。膜中では、膜・水界面の領域に CsA が安定に留まることがわかった (図 4A)。膜中央に CsA を配置した MD では、シミュレーションが進むにつれて、膜中央から脂質分子のアシル鎖付近の領域へ CsA が移動し、膜・水界面で安定に留まっていた。膜表面に CsA を半分埋めて配置した場合においても、CsA は膜中へ移動するが、膜中央ではなく、膜・水界面付近で安定に留まっていた。膜・水界面に留まる傾向は、初期配置だけでなく、初期構造がオープン/クローズド構造かにも依存しない、膜中における CsA の動的性質であることもわかった (図 4A)。

膜水系の多数のトラジェクトリをマルコフ状態遷移モデル (MSM) によって統合し、膜透過過程モデルの構築を行った。MD 中の構造を解析した結果、シス/トランス状態の変化を扱うには、通常の MD の範囲ではできないことがわかったため、トラジェクトリをトランス状態だけ、及び、シス状態だけのトラジェクトリに分類し、それぞれに対して、MSM を構築した。用いた膜水系では、2つの膜外領域は対称であり、細胞膜内外のような区別はないため、トラジェクトリを膜平面に対して反転操作をした反転トラジェクトリを作成し、通常のトラジェクトリと混ぜて MSM を構築した。約 52 μs 分のトランス状態のトラジェクトリから MSM を構築し、マイクロ状態の存在確率を解析した (図

4B)。存在確率の大きいマイクロ状態は膜表面に存在している事がわかった (図 4B, ○の大きさに対応)。また、膜表面間の遷移や、膜表面から膜外への遷移が遅い運動である事もわかった (図 4B, 赤色から青色への変化)。MSM から作成したマイクロ状態の膜透過モデルにより、膜透過の過程において、膜・水境界の領域と、膜の中央領域を横断することが障壁となっていることがわかった。この結果は、約 34 μ s 分のシス状態のトラジェクトリから作成した MSM でも、本質的に一致していることもわかった。

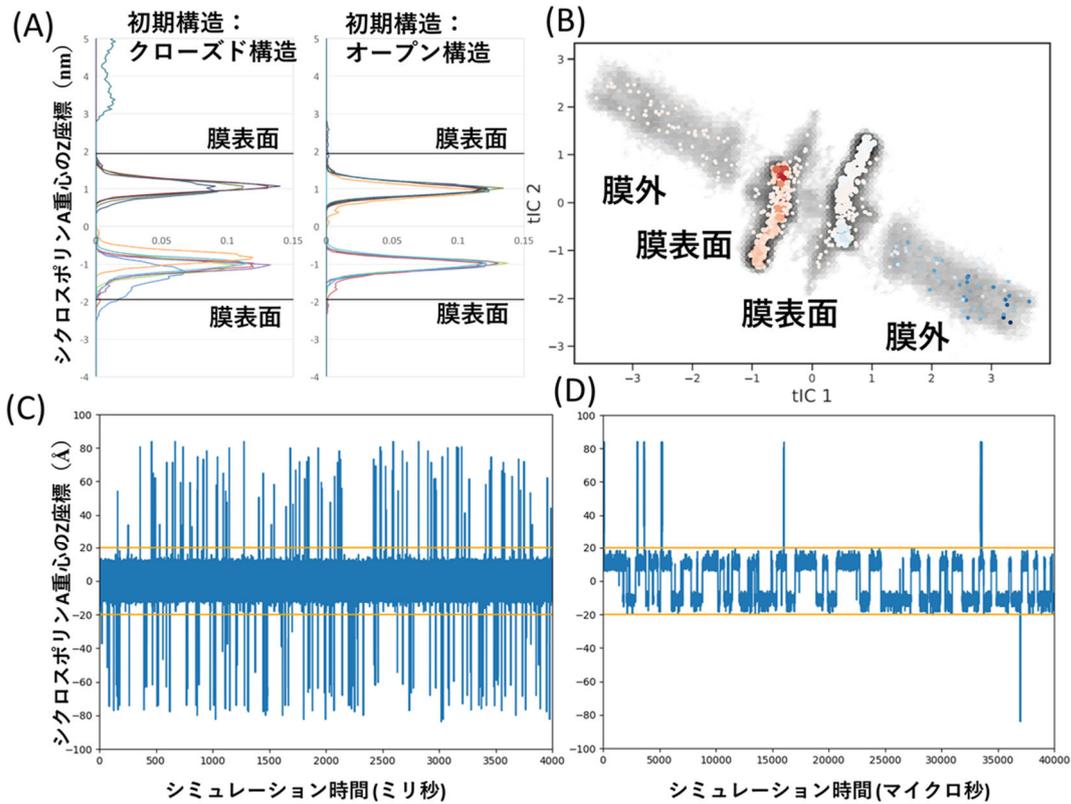


図 4 (A)CsA 重心 Z 座標の分布、(B)MSM ミクロ状態、(C, D)膜透過の長時間シミュレーション

MSM から得た状態間遷移確率を用いて、膜透過過程の長時間シミュレーションを実施した。マイクロ状態を 2つの膜外と 2つの膜表面に対応する 4つのマクロ状態に分類し、マクロ状態間の状態間遷移確率を使ったモンテカルロシミュレーションにより、4秒分の膜透過過程をシミュレーションした (図 4C)。CsA の膜中 (図 4C、オレンジ色実線内) から膜外への移動や、膜外から膜中へ移動する様子が再現できた。40 ミリ秒分のトラジェクトリを図 4D に示す。CsA は、膜中で膜・水界面領域に主に存在しているが、膜中央領域の横断や、膜・水界面領域から膜外へ移動する動的イベントが確率的に起こる事がわかる。CsA の膜透過過程について、MSM から得た長時間ダイナミクスの情報から、安定領域や障壁となる領域が存在するという動的性質を明らかにすることができた。

②タンパク質間相互作用制御 (再委託先：国立大学法人東京大学 山下雄史)

本課題は分子動力学 (Molecular Dynamics; MD) シミュレーションを用いて、タンパク質間の相互作用が目的タンパク質の機能にどのように関係しているかを解析することによりその制御を目指すもの

である。本研究課題では昨年度に引き続き多剤排出トランスポーターを対象とした。多剤排出トランスポーター複合体は近年社会問題となっている院内感染や末期がん、再発がんにみられる、“薬が効かない”という薬剤耐性化に大きな役割を担っている膜タンパク質で、複数のコンポーネントで構成される巨大超分子である（図1）。

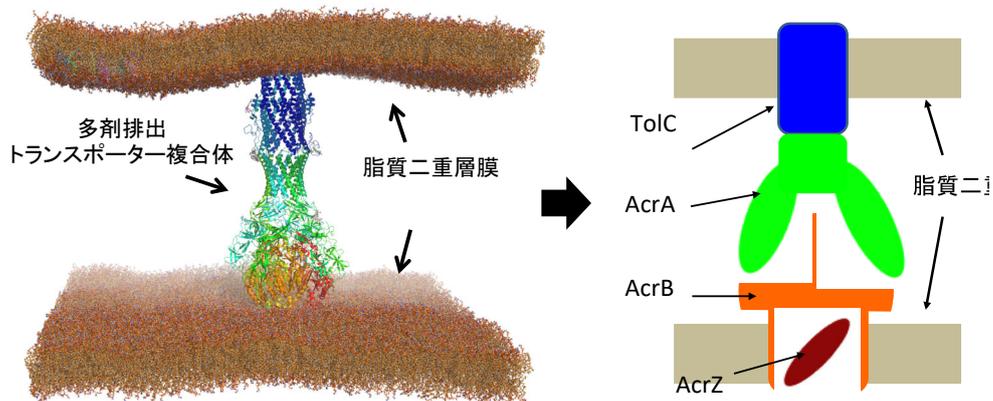
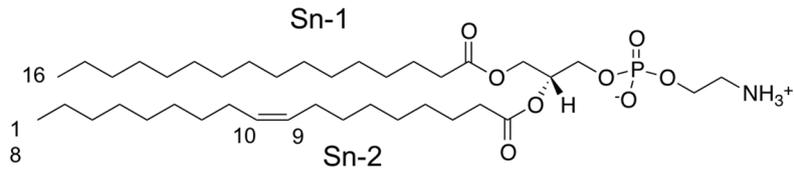


図1 多剤排出トランスポーター複合体-脂質膜系。右図は模式図

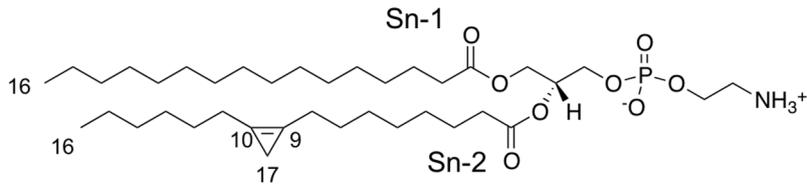
令和元年度は、実在の大腸菌膜を構成する主要脂質分子の力場作成とその大腸菌膜をミミックした多成分膜膜の構築し、AcrB-AcrZ-POPE 膜、AcrB-大腸菌膜、AcrA-AcrB-AcrZ-TolC 複合体-PMPE 膜の MD シミュレーションを行い、前年度までの結果と比較した。

令和元年度は最初に、よりリアルな膜の構築を目的にこれまで開発を進めてきた新しいFUJI 脂質力場の手法を用いて大腸菌膜の主要成分である PMPE や、PMPE を含む現実の大腸菌膜の構成比率上位 8 種類の脂質分子（PMPE、PVPE、PMPG、QMPE、PYPE、VYPE、PYPG、PVPG; 図2）の力場作成を完了し、PMPE 膜や上位 8 種類の多成分膜を構築した。大腸菌膜で約 80%を占める主要な脂質分子 PMPE は、sn-2 鎖の 9-10 位に二重結合を含む Cyclic な moiety を持つという大きな特徴がある。はじめに、moiety の影響が Area per Lipid や膜厚、電子密度にどの程度現れるのかを調べた。CHARMM-GUI を用いて PMPE1152 分子の二重膜を構築し、温度 298K、1 気圧で1 マイクロ秒の MD シミュレーションを初速度を変えて 3 本実行した。POPE 膜についても温度 298K、1 気圧で1 マイクロ秒の MD シミュレーションを 3 本行った。

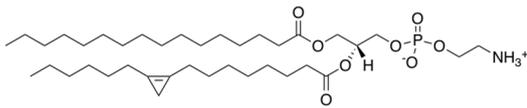
POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine)



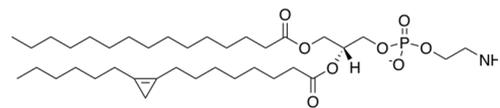
PMPE (1-palmitoyl-2-cis-9,10-methylene-hexadecanoic-acid-sn-glycero-3-phosphoethanolamine)



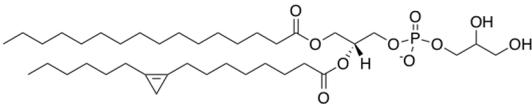
PMPE



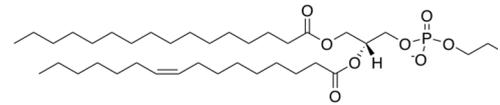
QMPE



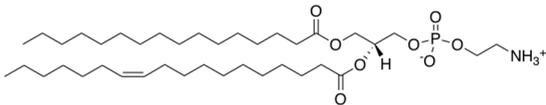
PMPG



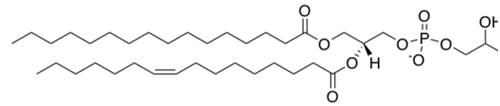
PYPE



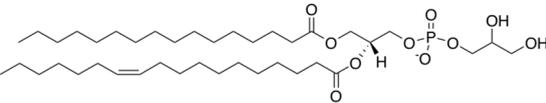
PVPE



PYPG



PVPG



YYPE

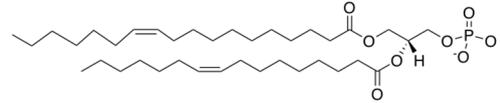


図 2 A:大腸菌膜で約 80%を占める主要な脂質分子 PMPE とそのモデル脂質分子の POPE の構造
B:今年度作成した 8 種類の大腸菌膜の構成脂質分子。

図 3 はそれぞれの膜の電子密度である。図 3 をみると、POPE 膜の方がブロードになっているのがわかる。膜厚は、Total の電子密度の 2 つのピーク間距離から算出することができるが、POPE 膜は 3.92 ± 0.01 、PMPE 膜は 3.55 ± 0.01 となり、PMPE 膜の方が膜厚は小さいことがわかった。Area per lipid は POPE 膜が 57.0 ± 0.01 PMPE 膜が 58.8 ± 0.01 と PMPE 膜の方が大きくなり膜厚の結果と矛盾しないことがわかった。

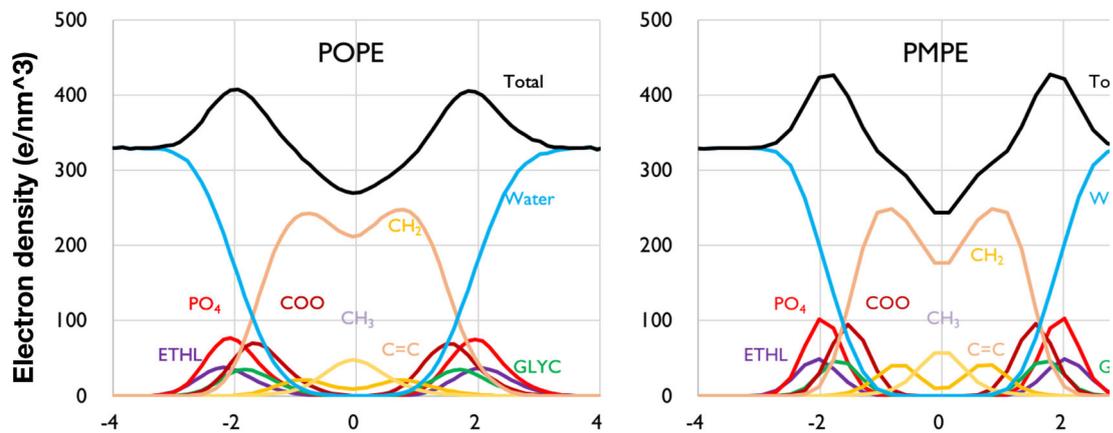


図3 POPE (左) と PMPE (右) 膜の膜面に対して垂直方向の電子密度。

図4はPOPE膜とPMPE膜のオーダーパラメータ S_{CD} を示す。 S_{CD} は、脂質分子の2つのアシル鎖 (Sn-1鎖、Sn-2鎖) のゆらぎ具合を反映する変量で、数値が小さくなるほど、アシル鎖はディスオーダーしていることを示す。両膜とも、Sn-1鎖はリン酸側 (C2) から末端までなだらかに減少しているが、

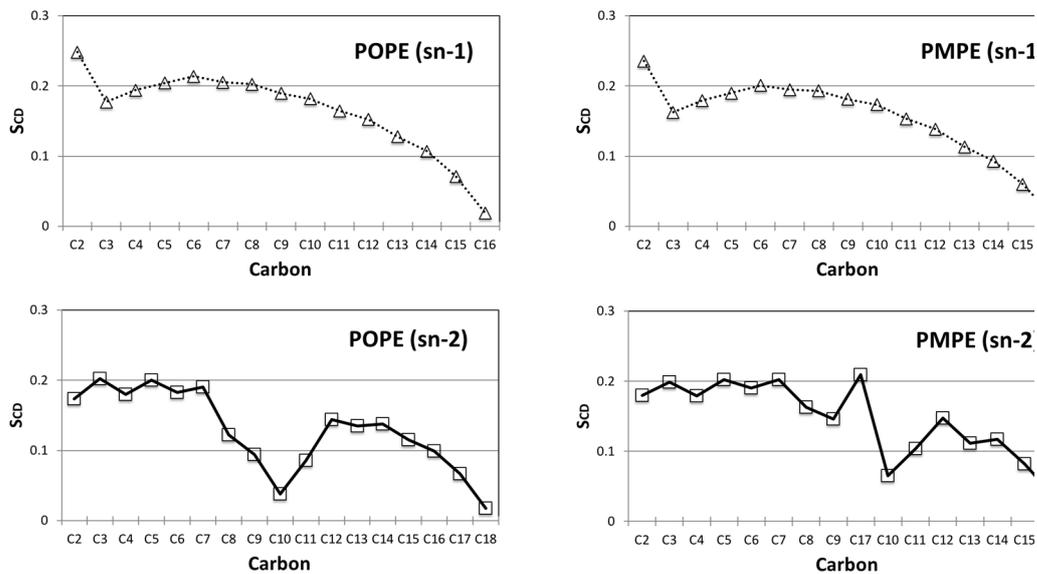


図4 オーダーパラメータ (S_{CD})

Sn-2鎖は2つの脂質分子で大きく異なっている。POPE (図2) の sn-2鎖の9-10位に見られるような二重結合は脂質分子の局所的なディスオーダーを引き起こすことは知られており10位を中心にV字型になっている。PMPE分子も sn-2鎖の9-10位は小さくなっているが、9位と10位の間に位置する17位 (C17) の値が大きくなっており、この3つのC原子による cyclic moiety が立体障害として影響していることが示唆された。

次に AcrZ の役割について解析を行った。AcrZ は、ポンプの役目をする AcrB の 3 つのモノマーの transmembrane ドメインにそれぞれ結合するタンパク質で、ある種の基質の排出を調整すると考えられているが詳細はわかっていない。図 5 は AcrZ の有無で AcrB のモノマー間の相互作用は変化するかを調べた結果である。AcrZ を含まない系では -2162 ± 116 kJ/mol、AcrZ を含む系では -1583 ± 57 kJ/mol となり AcrB のモノマー間の相互作用は AcrZ が結合していると弱まることがわかった。

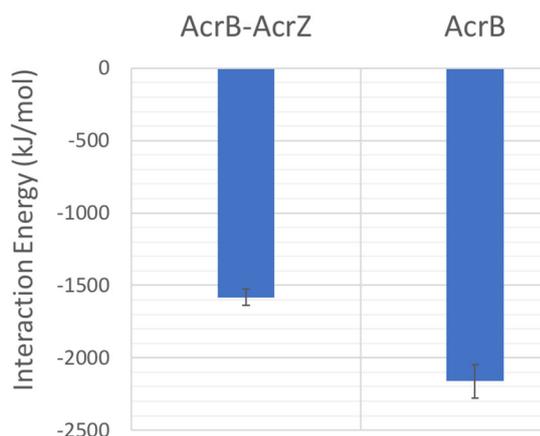


図 5 AcrB のモノマー間の相互作用。

図 6 は膜と AcrB や AcrB-AcrZ との相互作用 (900ns から $1 \mu s$ までの平均値) を示したものである。AcrB-POPE 膜系では -13031 ± 364 kJ/mol であるのに対し、AcrB-AcrZ-POPE 膜系では -12301 ± 307 kJ/mol で、AcrB と POPE 膜との相互作用よりも AcrB-AcrZ と POPE 膜との相互作用の方が小さいことがわかった。図 7 は AcrZ を構成する各アミノ酸と POPE 膜との相互作用を示したものである。これを見ると、両末端のメチオニンやヒスチジン、43番目のアスパラギン酸がやや強く POPE 分子と相互作用していることが見てとれるが、半数近くのアミノ酸が -10 kJ/mol よりも小さいことがわかる。この AcrZ と POPE 膜との比較的弱い相互作用が、AcrB-AcrZ 全体と POPE 膜との相互作用を弱めていると考えられる。

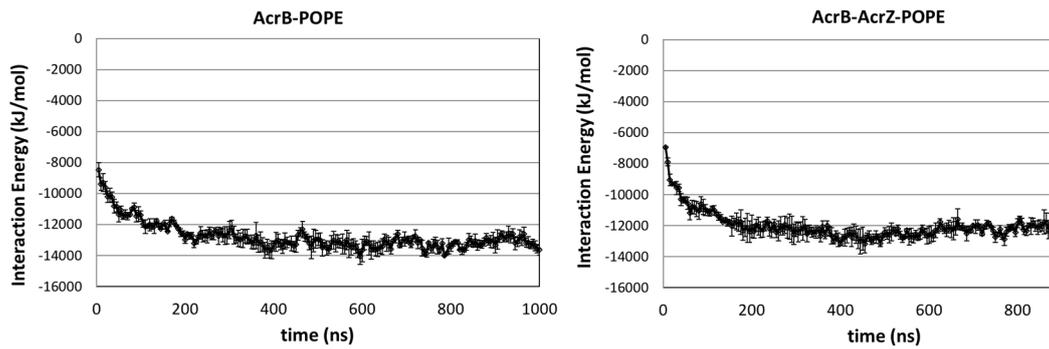


図 6 膜タンパク質—膜の相互作用。左図が AcrB-POPE 膜系で、右図が AcrB-AcrZ-POPE 膜系である。それぞれ 3 本の $1\mu\text{s}$ のトラジェクトリの平均を示した。Error bar は標準誤差である

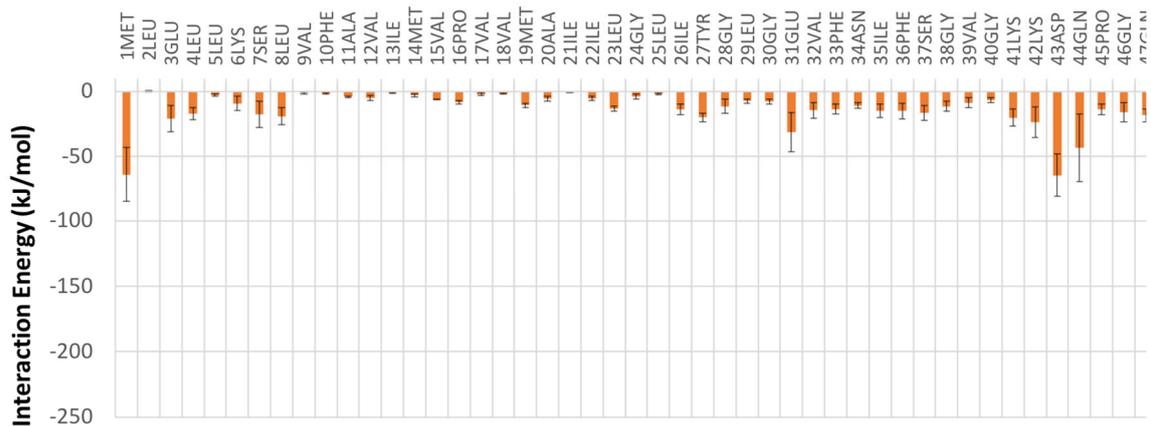


図 7 AcrZ の各アミノ酸と POPE 膜との相互作用。それぞれ 3 本の $1\mu\text{s}$ のトラジェクトリの平均を示した。Error bar は標準誤差である。

次に基質排出に関わる AcrB のポーター部分のダイナミクスに AcrZ が影響を与えるかどうかをみるために、主成分解析を行った。図 8 は AcrB-POPE 膜系、AcrBZ-POPE 膜系の第一主成分を黒い矢印で示した図である。解析は 500ns - $1\mu\text{s}$ のトラジェクトリ 3 本で行った。各トラジェクトリは Transmembrane 領域の backbone 原子にフィッティングした後、AcrB のポータードメインの水素原子以

外の原子に対して共分散行列を計算した。図8を見ると、AcrBの方は α ヘリックス（シリンダー形状）に対して垂直に近い角度の矢印が多いのに対し、AcrB-AcrZの方は α ヘリックスに対して上向きに傾いているのがわかる。これはAcrB-AcrZ-POPE膜系では、AcrBのポータードメインから離れたtransmembrane領域に結合しているAcrZがポータードメインのダイナミクスに影響を与えているこ

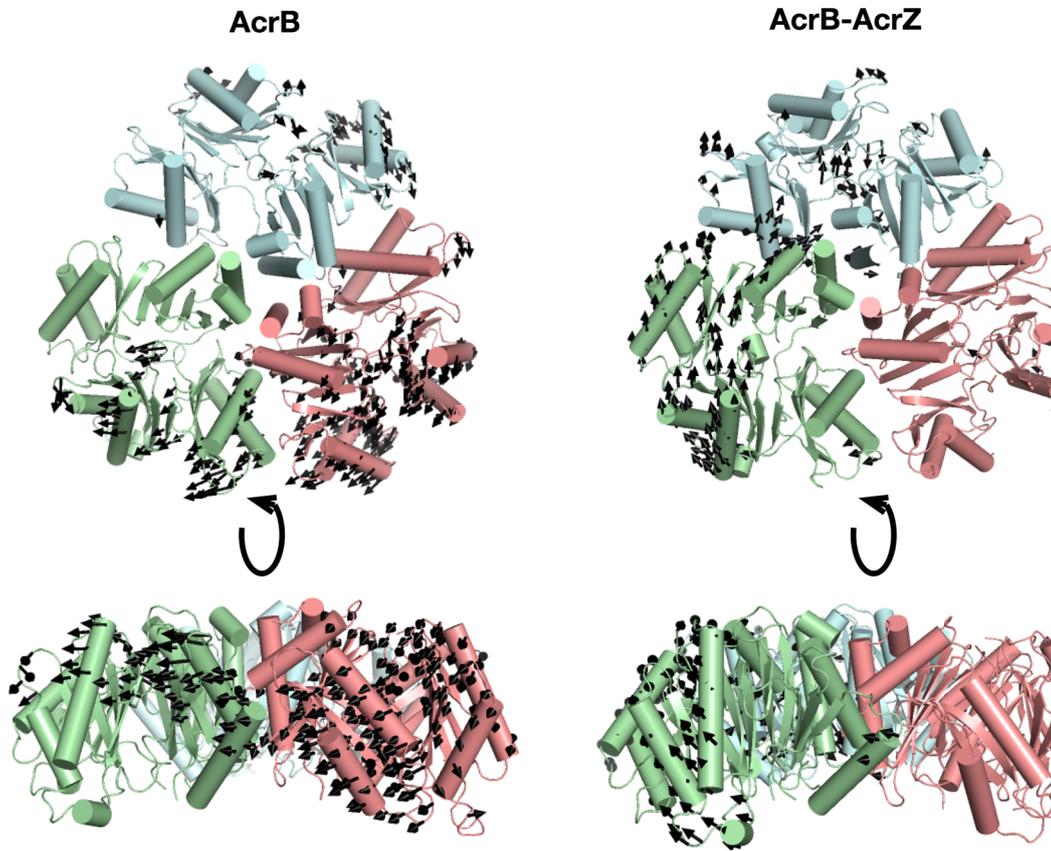


図8 AcrBポータードメインの主成分解析結果。第1主成分は黒矢印で示される。モノマーの色は赤が取り込み型、青が結合型、緑が排出型を示す。

とを示唆している。図9はAcrA-AcrB-AcrZ-TolC-POPE膜系の主成分解析の結果である。AcrB-AcrZの場合と比べると、AcrA-AcrB-AcrZ-TolCは上向きの主成分がポータードメインの中央部にもあることがわかる。複合体になると、ポータードメインの動きがより全体的に上向きに変化することを示唆する結果である。

AcrABZ-TolC-POPE

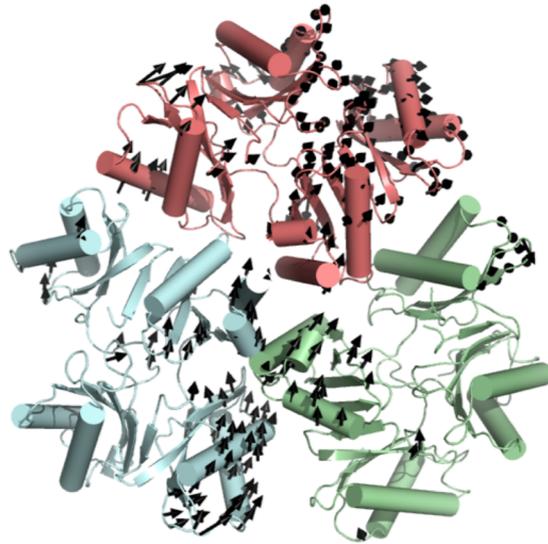


図9 AcrA-AcrB-AcrZ-TolC-POPE 膜の主成分解析の結果。

図10は脂質膜の違いがAcrBのモノマー間の相互作用に影響を与えるか調べた結果である。これを見ると、脂質膜がPMPEやモデル大腸菌膜になるとモノマー間の相互作用が弱まることになる。

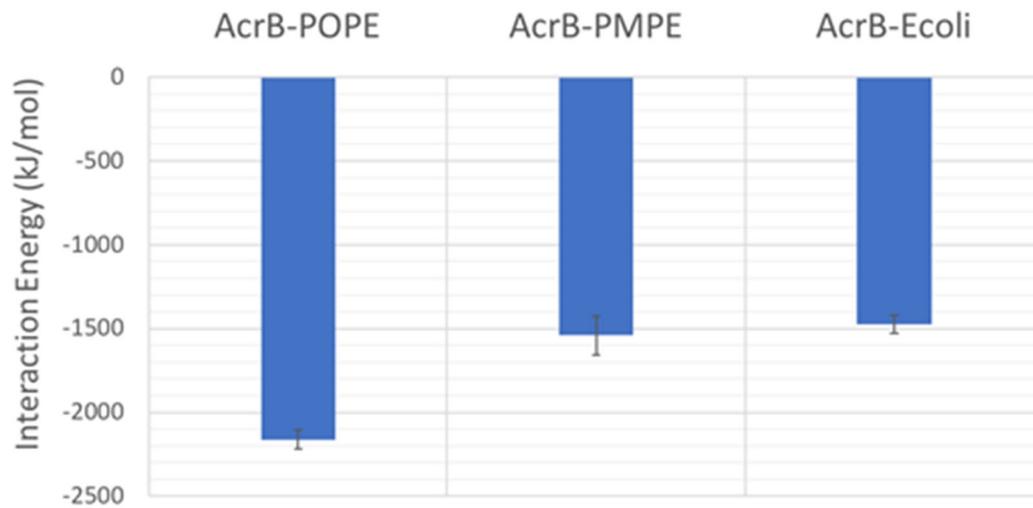


図10 AcrBのモノマー間の相互作用。

図 11 は AcrA-AcrB-AcrZ-TolC-POPE 膜、AcrA-AcrB-AcrZ-TolC-PMPE 膜における AcrB のモノマー間相互作用の結果である。この結果から POPE 膜よりも PMPE 膜に埋め込まれた方が AcrB のモノマー間相互作用が弱まることがわかった。AcrA-AcrB-AcrZ-TolC 複合体の主成分解析と合わせて考えると、AcrZ や脂質分子が AcrB モノマー間の相互作用に影響を与え、AcrB のポータードメインの動きに変化を与えることがわかった。

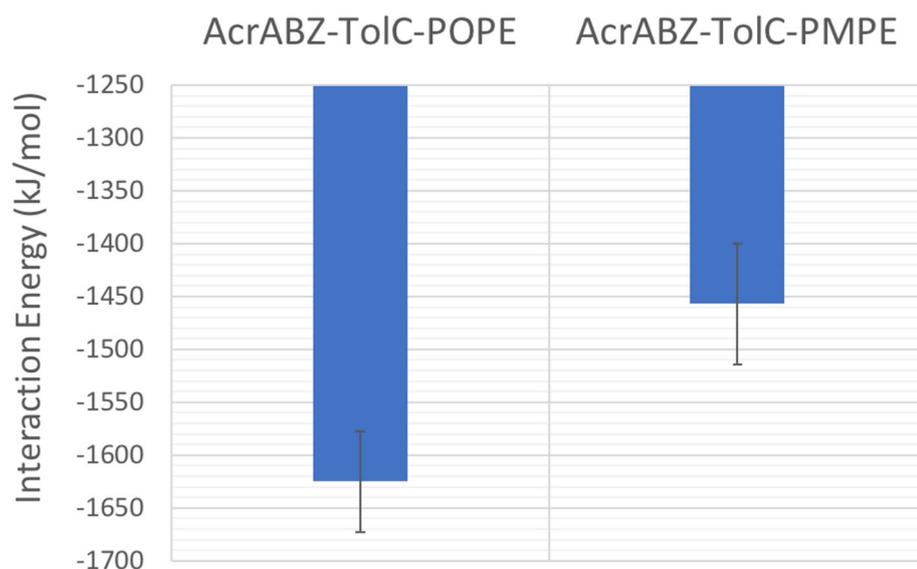


図 11 AcrA-AcrB-AcrZ-TolC-POPE 膜 (左) と AcrA-AcrB-AcrZ-TolC-PMPE 膜における AcrB モノマー間の相互作用

タンパク質間相互作用制御はバイオ医薬品とも密接に関係する。図 12 は進行ガンを治療するために開発中のプレターゲット治療プラットフォームの全体像を示している。右上にガン患者に薬を投与する順番を示しているように、まず抗体 (Antibody) を付けた Cupid (ヒト型化ストレプトアビジン四量体) を患者に投与する。図 8 左上で VL (軽鎖) (図中で黄色の楕円) と VH (重鎖) (図中で緑の楕円) と書かれているのが抗体の先端部分だけからなる scFv で、ガン細胞の表面に特異的に発現している標的タンパク質 (target protein) (図中で茶色の楕円) に選択的に結合する。Cupid は scFv を先端に持ったタンパク質分子が四量体を形成したもので、人の血管に注入されると体内を循環して標的タンパク質を表面に持ったガン細胞にのみ結合する。scFv を持った Cupid はタンパク質なので人体内を巡ってガン細胞に結合して血液中から無くなるまで待っても副作用の心配はない。イミノビオチン (iminobiotin) 構造を両端に持つ分子の真ん中に放射性物質などを付着させるためのドータを付けた分子が Psyche である。分子量が小さい Psyche は短時間で体内を巡りガン細胞に付着している Cupid と結合して、残りは腎臓で尿中に排出されるので一時間ほどで血液中から居なくなる。この為に Psyche に付けた放射性物質による副作用は、抗体に放射性物質を直接付けた場合より遥かに少なく、ガン組織だけにとどまる Psyche の放射線が効率よくガン細胞を消滅させる。

進行ガン治療プラットフォームとタンパク質間相互作用の最適化

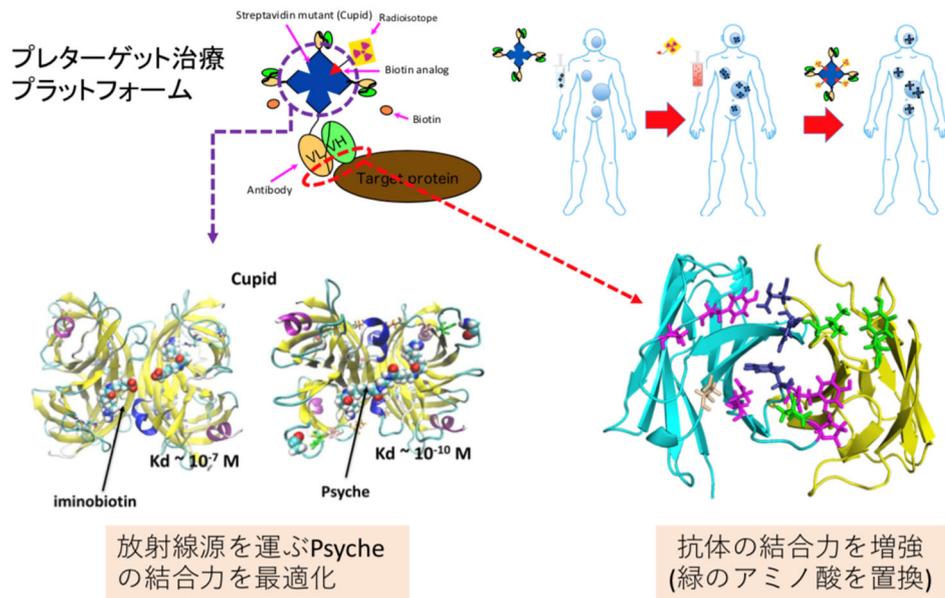


図 12 進行ガンのプレターゲット治療法に用いる scFv、Cupid、Psyche

このプレターゲット技術の開発に、抗体 (scFv) と標的タンパク質の結合や Cupid の四量体形成などタンパク質間相互作用が本質的に関わる問題が内包される。本課題においても、計算科学の観点から様々な開発支援および技術開発をおこなってきた。最近、成果の一部が論文として発表された (Sugiyama et al., Proc. Jpn. Acad. Ser. B, 2019)。昨年度には、抗体の親和性向上において、界面水分子や 3 つ組塩橋という鍵になる特徴を見出した (Yamashita, Structure, 2019)。

これを受けて、本年度は塩橋に着目して、塩橋の安定性が制御される要因について調査した。鶏卵白リゾチーム (hen egg-white lysozyme, HEL) と抗体 HyHEL-10 の間には 3 つ組塩橋が存在するが、この 3 つ組塩橋はアラニン置換によって親和性への寄与が確認されている (図 13a)。

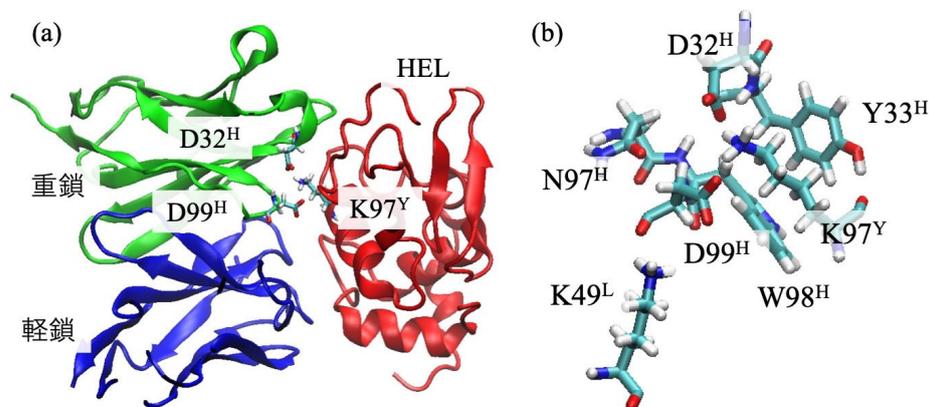


図 13 (a) HEL-HyHEL-10 複合体と 3 つ組塩橋 (b) 3 つ組塩橋とそれを囲む残基

まず、我々はこの3つ組塩橋の安定性を調べるために、MD シミュレーションをおこなった。力場には我々のグループで開発している国産高精度力場 FUJI を用いた。結晶構造では D32^H-K97^V ペアも D99^H-K97^V ペアも安定に塩橋を形成していたが、MD シミュレーションでは D99^H-K97^V ペアの塩橋形成割合は 10%を下回った。おそらく、実験では結晶中の特殊な環境によって D99^H-K97^V 塩橋が安定化されていたと考えられる。もう一方の D32^H-K97^V ペアの塩橋形成割合は 90%と高かった。

C α 間距離では D99^H-K97^V ペアよりも D32^H-K97^V ペアの方が遠いので、単純な 2 残基の相互作用ではこれら 2つの塩橋の安定性の違いは理解できない。実際に、塩橋 2 残基からなるモデル(Res2)を調べた。本簡約モデルでは主鎖重原子を固定しており、側鎖の振る舞いのみを MD シミュレーションによって計算する。C α 間距離や側鎖の伸びる方向が塩橋の安定性に大きく影響していたが、この 2 残基だけでは実際の複合体中での D99^H-K97^V 塩橋と D32^H-K97^V 塩橋の安定性を再現できないことが分かった。

そこで、塩橋を取り囲む残基の影響を考えるため、さらに 3つ組塩橋からなる 3 残基モデル(Res3)・K49^Lを加えた 4 残基モデル(Res4)・N97^Hを加えた 5 残基モデル(Res5)・芳香族環を持つ Y33^Hと W98^Hを加えた 7 残基モデル(Res7M)を構築した。また、Y33^Hと W98^Hの側鎖が実際の複合体中では特定の構造に強く束縛されていることを模倣し、これらの側鎖重原子の位置も制限した Res7S モデルも構築した。これらのモデルによって、塩橋の形成割合を算出した結果が図 14 である。Res3-5 のように親水性残基だけでは複合体中での塩橋安定性を再現することはできない。実際の複合体における塩橋形成割合は、Res7S を導入することで初めて再現性を高めることが可能であった。このことは、タンパク質界面で固定され壁のように塩橋残基の運動に影響を与える Y33^Hと W98^Hの存在が非常に重要であることを意味している。

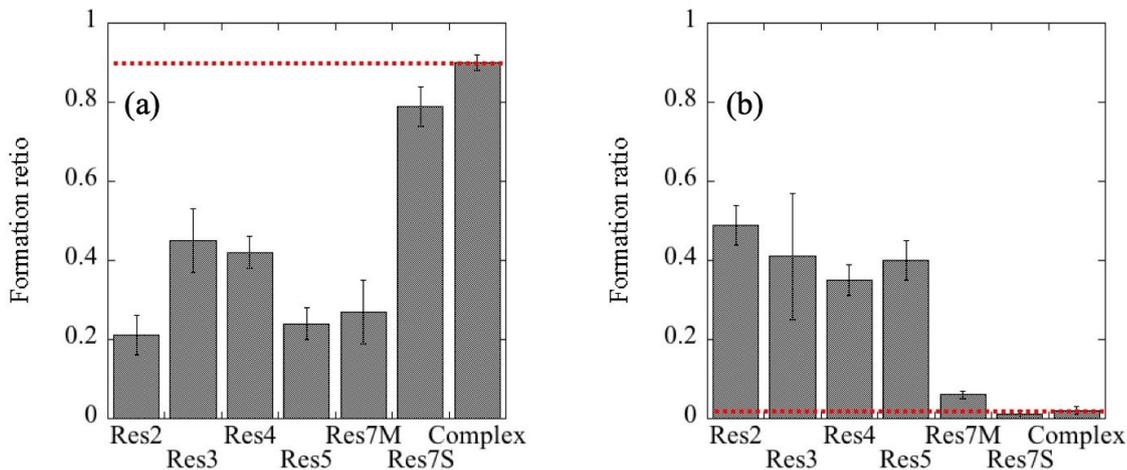


図 14 塩橋形成割合の比較：(a) D32^H-K97^V ペア (b) D99^H-K97^V ペア

本研究では、再現性のあるモデルを構築していくことで、系の本質を読み解く構成論的アプローチにより、抗原抗体界面の塩橋安定性の要因について理解を深めた。特に、親水性残基との相互作用だけでなく、疎水性残基からの効果も重要であることが分かった。特に、芳香族環がタンパク質界面において強く固定されるという効果も本質的であり、塩橋の振る舞いに大きく影響を与えていることが分かった。本研究の結果は、塩橋の安定性を予測し安定性を高める設計を可能にするには、塩橋を

取り囲む環境の効果を考慮する必要であることを示唆している。したがって、合理的な抗体医薬品開発には大規模 MD シミュレーションによるタンパク質界面の緻密な解析が極めて重要だと考えられる。

③ウイルス標的創薬計算技術（再委託先：国立大学法人名古屋大学 篠田渉）

ウイルス環境の分子シミュレーションによる再現・構築とウイルスカプシドと抗ウイルス剤の相互作用解析に基づく創薬計算技術の開発を実施している。令和元年度は B 型肝炎ウイルス（HBV）カプシドを対象とし、薬剤（逆転写酵素阻害剤）のカプシド吸収における自由エネルギー解析技術の開発・改良を進めた。さらに複数の阻害剤や核酸アナログのカプシド吸収効率の評価を予定していたが、自由エネルギー計算の収束に予定以上の計算時間が必要と判明したため、現行で最も有効な逆転写酵素阻害剤、エンテカビルに特化して、カプシドへの吸収自由エネルギーの評価に取り組んだ。昨年度来、行っている荷電溶質の溶媒和自由エネルギー解析を、高速多重極子展開法を用いて行う方法論の開発を継続して行った。本手法を汎分子動力学ソフトウェア MODYLAS に移植し、高並列演算環境において計算パフォーマンスの最適化を行った。方法論に関する論文は現在、科学雑誌に投稿中であり、より簡単な荷電溶質の水溶液中での計算は非常に高い収束性を示すことが分かった。本手法を用いて、(三リン酸化された)エンテカビルを真空から HBV カプシド中心へと移動する際の自由エネルギー変化(ΔG_{capsid})の計算に取り組み、京コンピュータ等を用いて収束計算に成功した。同系は約 600 万粒子系で、このような大規模系での自由エネルギー計算は他にほぼ例がない。 $\Delta G_{\text{capsid}} = -3270$

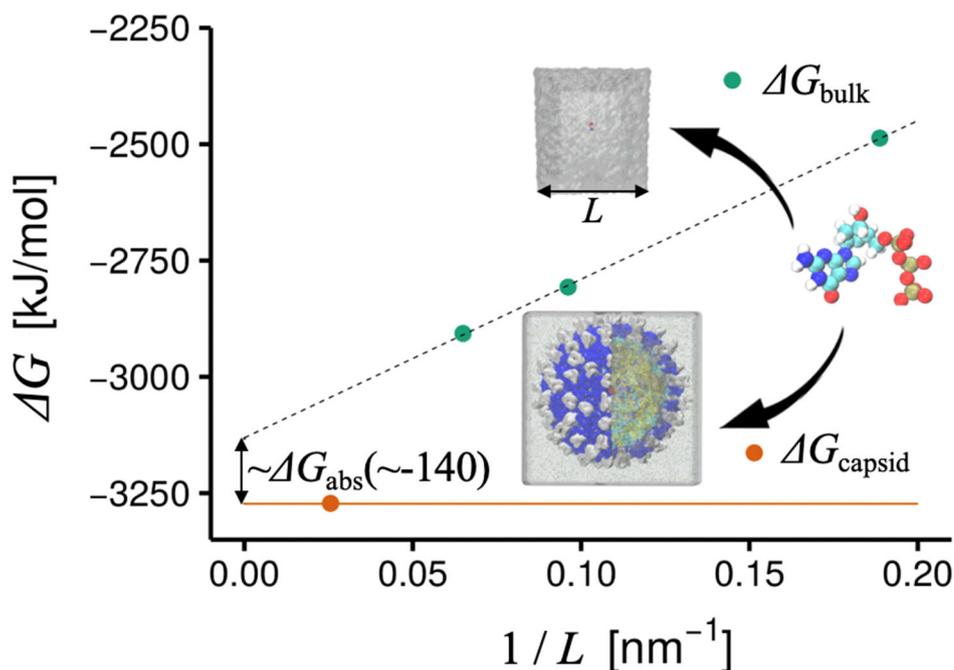


図1 カプシド中心とバルク水溶液中へのエンテカビル三リン酸の溶媒和自由エネルギー計算。後者はシステムサイズの逆数($1/L$)の関数として計算し、無限希釈極限($1/L=0$)の自由エネルギーを算出。

±10 kJ/mol と見積もられた(図 1)。エンテカビルの外部からカプシド内への吸収自由エネルギーを見積もるには、同じカプシド系の外部溶液中へのエンテカビルの溶媒和自由エネルギーを計算すべきであるが、今年度は京の運用が停止し計算負荷の許容量を超えたため、近似的にエンテカビルを真空から肝細胞条件のバルク電解質溶液へ移動する際の溶媒和自由エネルギー計算(ΔG_{bulk})を行った。図 1 には計算された自由エネルギーをシミュレーションボックスサイズ(L)の逆数の関数としてプロットしたものである。荷電溶質の溶媒和自由エネルギー計算では、有限サイズ効果が顕著であり、理論的にも($1/L$)に比例するサイズ効果が予想される。その補正を行い無限希釈系での値から吸収自由エネルギー ΔG_{abs} は約-140kJ/mol と大きく負の値となった。これは、先に述べた近似的な扱いであるが、エンテカビル三リン酸がカプシド内部に入って大きな自由エネルギー利得を示すことを示唆している。さらに、カプシド内外での溶媒和自由エネルギー評価は、粗視化モデルを用いても行った(後述)。

図 2 には HBV カプシド全体系の全原子分子動力学(MD)計算(200ns)のトラジェクトリから、イオンの分布をカプシド中心から動径方向に算出した結果である。カプシドタンパク質の殻が $R \sim 13-15$ nm に存在し、この領域ではカプシド細孔内にしかイオンは分布できないため、密度が低くなっている。また、カプシド内の pgRNA が存在する $R \sim 8-11$ nm の領域では、RNA のリン酸を中和する形でカチオンが多く存在している。特に Mg^{2+} は RNA に強くバインドしており、RNA の安定化に寄与すると実験的に示唆されていたため初期配置から RNA 近傍に配置したが、この領域から計算時間中に離れていく

様な運動は見られなかった。これらの特徴に比べて、図では少し見にくいだが、カプシド内部の水溶液 ($R < 5$ nm)とカプシド外部の水溶液($R > 20$ nm)を比べると、イオン濃度がカプシド内部でより高くなっていることがわかる。カプシド内外でのイオンの交換は計算時間の 200ns の間に十分起こっていることから、平衡濃度としてイオンがカプシド内部でより多く存在することが示されている。同様なことが、荷電した溶質にも言えるため、+4 価に帯電しているエンテカビル三リン酸がカプシド内部でより安定化されるような場がカプシドによって形成されているものと考えられ、これはエンテカビルの負の吸収自由エネルギーをよく説明する。図 2 には破線で SPICA 粗視化力場によるイオン分布もプロットしている。我々が開発している SPICA 粗視化力場は、全原子 MD の結果をよく再現することがわかる。

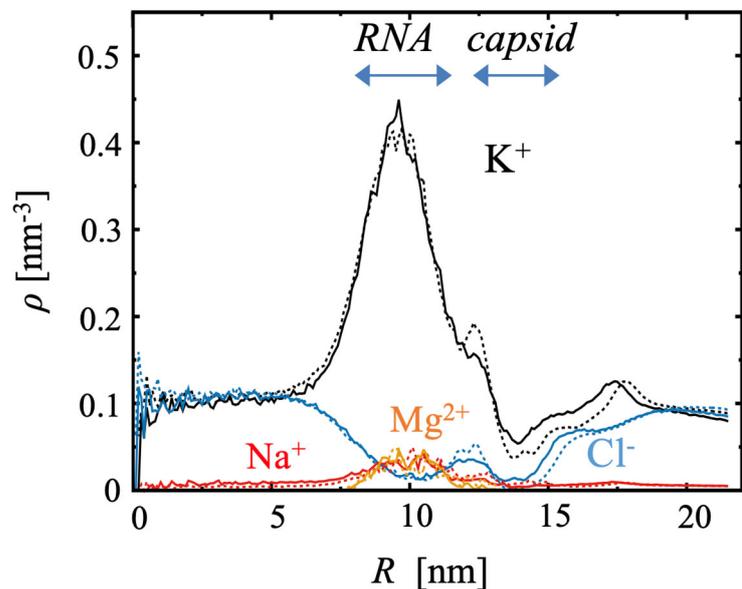


図 2 MD 計算から得られたイオン分布。カプシド中心から動径距離の関数として計算。pgRNA やカプシド殻の位置も参照のため示す。全原子 MD を実線、粗視化 MD を破線で示す。

上記では、カプシド中心への吸収という限定された空間への自由エネルギー計算を行っていたため、カプシド内部全体での分布(自由エネルギー)を概算することを全原子 MD で試みた。そのため、エンテカビル三リン酸を 100 分子カプシド内部にランダムに配置し、その配置から緩和計算を 30 ns 行った(図 3)。最初の 20ns を平衡化として、最後の 10ns のイオン分布から自由エネルギー差を算出した(図 4)。むろん時間に限りがあるため、完全に自由エネルギーとして収束したものではないが、計算された自由エネルギー差から、カプシド内部ではエンテカビル三リン酸が RNA 内及び表面($R \sim 8-10$ nm)に集まる傾向が見られた。カプシド外へのエンテカビルの拡散は計算時間内には見られなかった。SPICA 粗視化力場を用いた同様な計算も行っているが、500 ns (粗視化時間: 物理的な時間スケールは 4~10 倍程度であり、2~5 μ s に相当) の計算から、ほぼ同様な自由エネルギープロファイルが得られており(図には載せていない)、また、この時間でもカプシド外へのエンテカビルの脱出は見られなかった。これらのことから、内部に吸収された三リン酸化された逆転写酵素阻害剤(エンテカビル)は、RNA 近傍で安定化されることが示唆される。

すでに SPICA 粗視化力場の結果を少し述べているが、SPICA 粗視化力場は元々脂質膜系での計算のため作られてきたが、タンパク質、DNA/RNA を扱えるような拡張を行ってきた。基本的に脂質膜系のモデリングは界面自由エネルギーを合わせ、分子分配係数、溶媒和自由エネルギーなどを実験、全原子モデルと整合させることで、非常に高い定量性を保つ粗視化モデルとなっている。タンパク質や DNA/RNA の各残基、塩基について、脂質膜を横切る自由エネルギープロファイルを全原子モデルと合わせることで、全原子モデル(特に CHARMM36 力場に合わせた)で期待される結果を再現するモデルである。昨年報告の通り、HBV カプシドの外部バルク水溶液中にエンテカビル三リンを 55 分子加えた初期配置から、SPICA

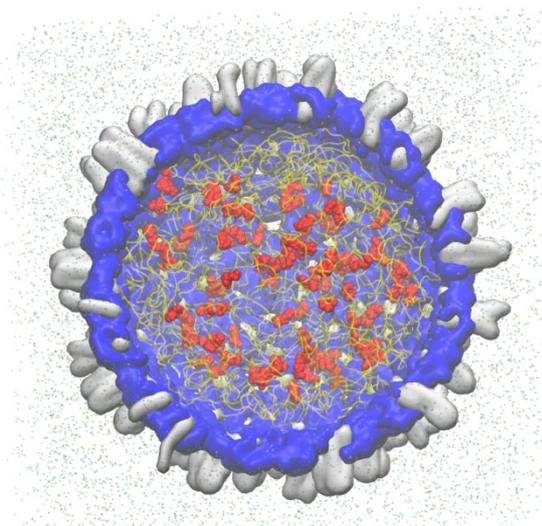


図 3 エンテカビル三リン酸(赤)をカプシド(青、灰色)内部で緩和させた全原子 MD 計算のスナップショット。RNA は黄線。イオンを点で示した。

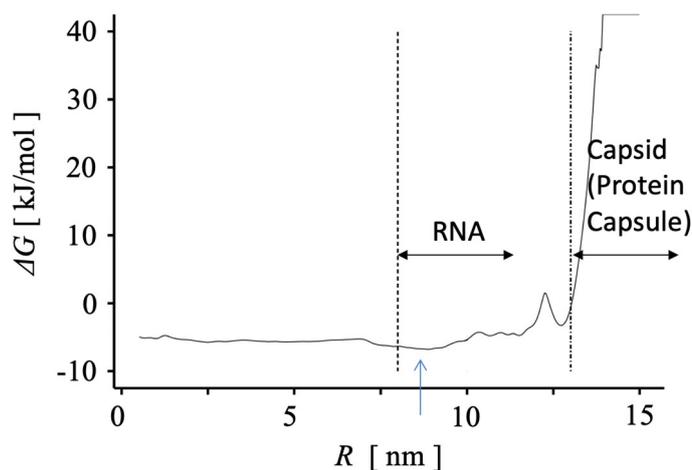


図 4 エンテカビルの分布から計算したカプシド内部での自由エネルギープロファイル。カプシド中心からの距離の関数として計算。青色の矢印部分に浅いポテンシャルウェルが見られる。

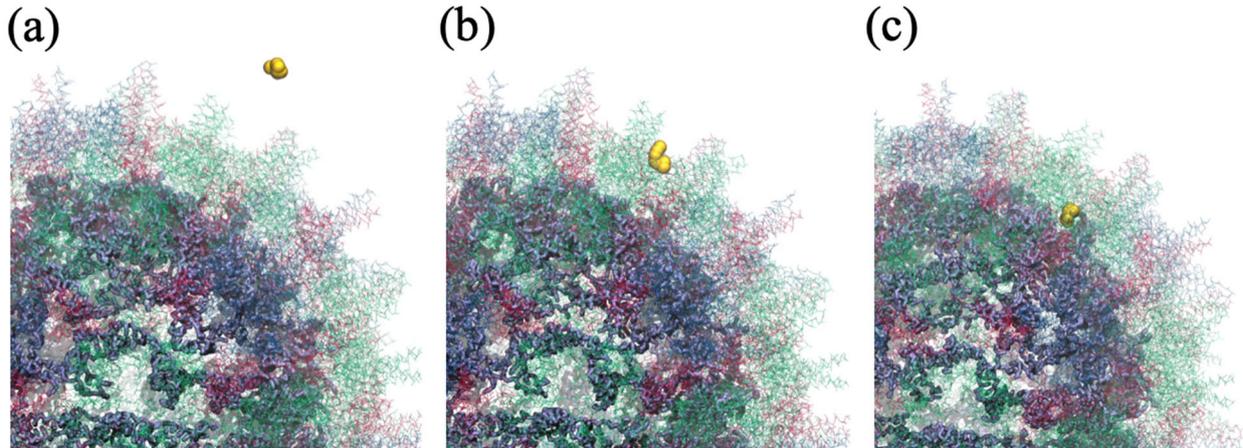


図 5 エンテカビル(黄色)のカプシド内(透明な線)への侵入。グレーの鎖は RNA。(a) エンテカビルが溶液中、(outer-capsid site)(b) 3 回回転対称軸周りの入口の結合サイトに到達(entrance site)、(c)カプシド内部に到達し RNA 表面に吸着した状態。(inner-capsid site)

力場を用いて粗視化 MD 計算を行った。計算開始後 0.5 μs 程度ではほぼすべての阻害剤がカプシド表面に吸着し、500 ns 過ぎに阻害剤 1 分子が 3 回回転対称軸上の細孔を抜けて、カプシド内側に入っていく様子を確認した。3 μs (実効 12-30 μs に相当) まで計算時間を延ばしたが、カプシド表面に吸着した阻害剤は比較的長い滞在時間を示し、自発的に入っていく阻害剤はこの他に見られていない。3 回回転対称軸周りの細孔が親水的で孔径が大きく、その他の細孔は阻害剤透過に有利ではなく、その透過を考える必要がほとんどないことがわかった。そのため、初期配置で阻害剤を 3 回回転対称軸周辺に配置した粗視化 MD を追加で行った。その結果、1 μs の間に数分子が細孔を抜けてカプシド内部の RNA 表面に到達する様子が観測されている。

図 5 には 3 回回転対称軸周りの細孔に侵入する阻害剤の様子を示している。外にいる阻害剤は特別な配向を持たず、ランダムに回転しながらカプシド表面に吸着する。(図 5a) 昨年度までの全原子 MD から得られた電場解析などにより予測した経路と同様に、負に帯電した阻害剤はスパイクを避けながら、カプシド表面に到達することが、SPICA 力場の粗視化 MD でも観測されている。3 回回転対称軸周りの細孔入り口付近には荷電残基が形成する吸着ポケットが存在し、阻害剤はそこに一旦吸着を示した。(図 5b) さらに数百 ns 後、細孔内部への拡散を示し RNA 表面に到達し、安定な吸着を示した。(図 5c) 3 μs まで、そのカプシド内部のポケットか

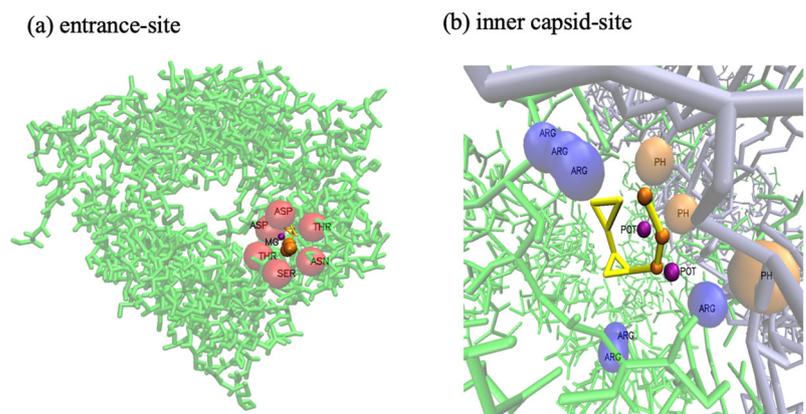


図 6 2つの安定サイトに到達したエンテカビル三リン酸。(a) 細孔入り口付近のサイト (b) カプシド内部の結合サイト。エンテカビル三リン酸を黄色の線とオレンジの球で表示。緑色でタンパク質、グレーで RNA を表示。サイト近傍の荷電残基を球 ((a)の赤と(b)の青、オレンジ) で表示した。

ら離れることがなく、RNA 表面で留まった。これら阻害剤が安定に滞在するポケットは荷電残基の存在で特徴づけられる。図 6 にこれらのポケットを定義する周辺残基を球で示す。これら残基と阻害剤の距離によりポケットを定義し、それぞれのポケットにいる阻害剤の溶媒和自由エネルギーを、位置拘束を用いて計算し、カプシド外部の水溶液中での溶媒和自由エネルギーと比較することで、吸着、吸収の自由エネルギーを見積もることとした。全体系は粗視化力場を用いても、約 93 万粒子系という大規模計算であるが、 μs の計算が可能のため、非常に高精度に自由エネルギーを見積もることができた。

SPICA 力場用のソフト LJ ポテンシャルを用いた自由エネルギー摂動法を LAMMPS コードに移植し、計算を実行した。 λ を 100 点で刻み、全体 $2\mu\text{s}$ の計算を各サンプルについて行うことで 1 kcal/mol 以内の統計誤差での計算が実現した。図 7 に溶媒和自由エネルギー計算の結果を λ の関数で示す。粗視化力場では溶媒和自由エネルギーの絶対値は全原子モデルよりも小さくなるが、異なる状態間の自由エネルギー差は正しく評価できると期待できる。結果は $\Delta G_{\text{outer-capsid}} = -65.9 \pm 0.1 \text{ kcal/mol}$ 、 $\Delta G_{\text{entrance-site}} = -74.9 \pm 0.3 \text{ kcal/mol}$ 、 $\Delta G_{\text{inner-capsid-site}} = -85.9 \pm 0.6 \text{ kcal/mol}$ と見積もられ、バルク溶液から細孔入り口の結合サイトに来ることで約 -9 kcal/mol の自由エネルギー利得があり、入り口からカプシド内側の RNA 表面に到達することでさらに約 -11 kcal/mol の利得があることがわかった。バインディングサイトへの拘束による自由エネルギー補正を考えるとさらに -1.9 kcal/mol 程度カプシド内部で安定となり、全体でカプシド外にいる場合に比べて、阻害剤はカプシド内の RNA 表面に到達することで約 -21.9 kcal/mol のエネルギー利得を示した。これは非常に大きな値であり、一旦吸収された阻害剤はほとんどカプシド外に排出されることがないことがわかる。逆転写酵素が機能する RNA 表面に阻害剤が安定に存在することは、三リン酸化された阻害剤の効率の高さをよく説明するものである。

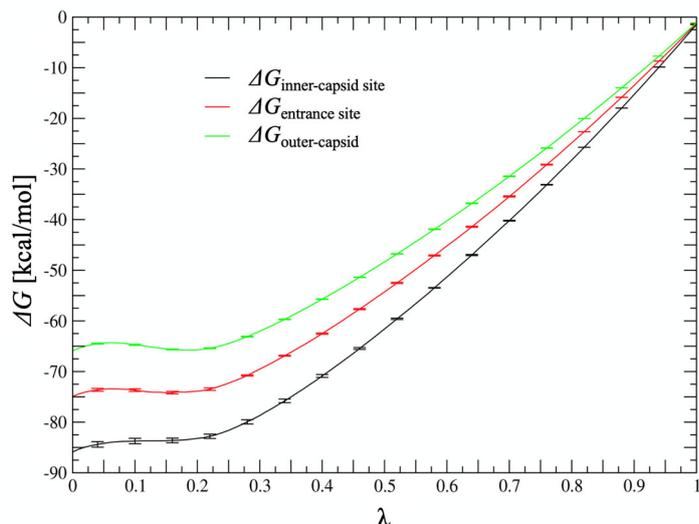


図 7 自由エネルギー摂動法の計算結果。3 つの異なる状態での溶媒和自由エネルギーをそれぞれ計算している。

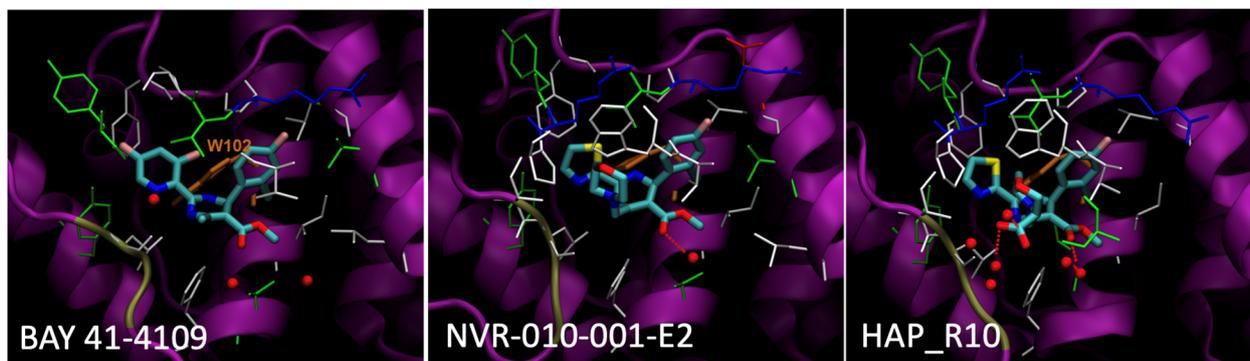


図 8 カプシドタンパク 4 量体(紫 Cartoon 表示)の界面結合ポケットに結合した HAPs のスナップショット。HAP を太い Stick で、周辺のアミノ酸残基を細い Stick で表示した。

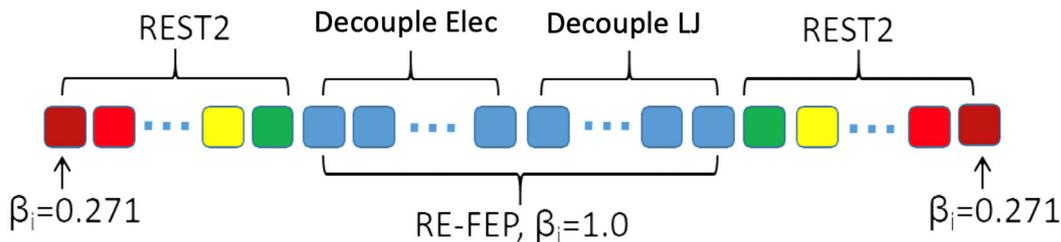


図 9 reduced FEP/REST2 法。レプリカ交換は隣り合うレプリカ間でのみ行う。 β_i はリスケーリング強度、色は異なる強度に対応。

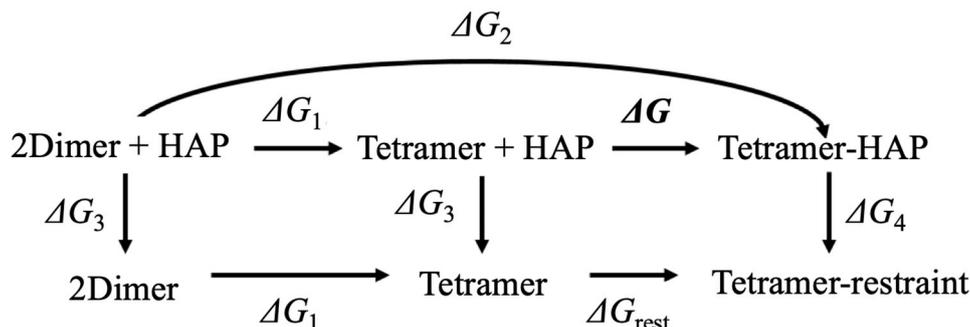


図 10 熱力学サイクル。HAP による 4 量体の自由エネルギー利得 ΔG を計算するため、 ΔG_3 、 ΔG_4 、 ΔG_{rest} を計算する。 ΔG_{rest} は HAP をポケットに拘束する restraint に必要な自由エネルギー寄与。 ΔG_1 は実験値が提案されており、 ΔG の算出により ΔG_2 の計算も可能となる。

また、カプシド形成阻害剤によるカプシドタンパク複合体の構造安定性への影響について、HAPs (heteroaryl pyrimidines) のカプシドタンパク 4 量体の結合ポケットへの自由エネルギー解析を、全原子モデルを用いて実施した。HAPs の中から 3 つの阻害剤 ; BAY41-4109, NVR-010001E2, HAP_R10 を選び、これらの結合自由エネルギー計算を行った。結合構造をドッキングシミュレーションと MD による緩和計算よりモデリングした(図 8)。通常の BAR 法では自由エネルギー計算の収束性に問題が見られたため、aMD (accelerated MD) によって改善を試みたが不十分であった。そのためさらに FEP/REST2 (free energy perturbation/replica exchange solute tempering) により、阻害剤の溶媒和自由エネルギー計算を行った。これはレプリカ交換法の拡張法であり、図 9 に示すようにレプリカを並べて交換するスキームであり、FEP の端点の系のみ REST2 で拡張する手法である。FEP/REST2 によって自由エネルギーの収束が改善され、3 つの阻害剤の結合自由エネルギーの差を議論するのに十分な精度の結果を得た。図 10 に示すような、熱力学サイクルを考えることで、HAPs によるカプシド 2 量体間の結合自由エネルギーの変化 ΔG を議論することができる。 ΔG の大きさは BAY41-4109<NVR-100-001-E2<HAP_R10 の順で大きくなり、 pEC_{50} (50%効果濃度)とよく相関した結果を与えた(図 11)。また、HAPs が結合することにより、タンパク質の結合配向が変化することも見いだされた。この平面配向を強める局所的なタンパク質配向の変化は、HAPs がエンブティーカプシドだけでなく、集合構造としてナノシートなどの他の構造も誘起する可能性を示す結果である。

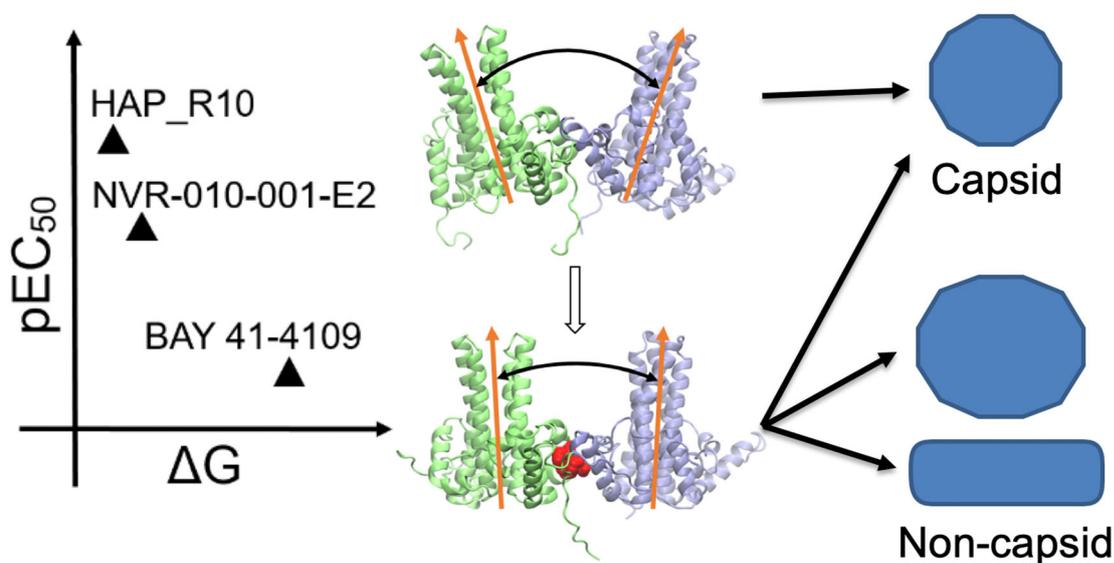


図 11 自由エネルギー ΔG と pEC_{50} の相関。HAP の結合による 4 量体構造の変化、またその集合体構造への影響の模式図。

④核酸-タンパク質相互作用制御 (再委託先: 国立研究開発法人量子科学技術研究開発機構 河野秀俊)

令和元年度は、引き続き核内の DNA 収納状態の動構造を調べた。全原子分子動力学シミュレーションでは数ミリ秒の現象を捉えることは未だ難しい。そこで、振じれや引っ張りの外力を加え、その応答をみることで動構造を調べた。

真核生物の DNA は、ヌクレオソームという構造体の形で、細胞の中の核に収納されている。ヌクレオソームはヒストンタンパク質 8 量体の周りに約 150 塩基対の DNA が 2 回巻き付いた構造体であり、核の中ではこのような構造体は何千何万と存在する。DNA に刻まれた遺伝子が働くためには、個々のヌクレオソームに巻き付いた DNA が他の分子に読まなければならない。そのためには、巻き付いた DNA は解離しなければならないが、その詳細な過程についてはよく分かっていない。そこで、引っ張りと振じれに対するヌクレオソームの応答を調べた。

まず、引っ張りに対するヌクレオソームの応答を調べた。ヌクレオソームを形成するヒストンには、様々なバリエーションが存在し、異なる役割を果たしていることが分かってきている。これまでに、我々は通常のヒストン H3 から構成されたヌクレオソームに対して、DNA の引っ張りに対する自由エネルギー地形を報告した (Kono et al. PLoS Comp. Biol. 2018)。今回は、ヒス

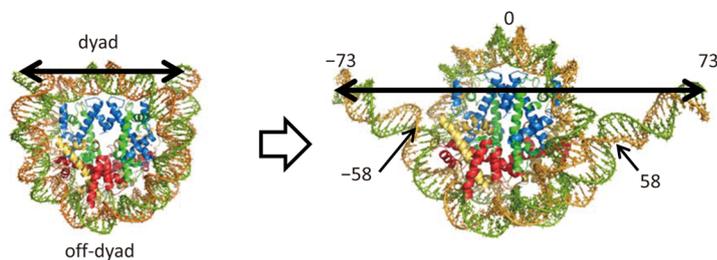


図 1 ヌクレオソーム構造と反応座標。矢印で示したように DNA の末端間の距離を反応座標に取り、DNA の解離に対する自由エネルギー変化を求めた。計算は溶媒分子、イオンを含む 40 万原子の系で実行した (表示では、溶媒分子、イオンを省略)。

トン H3 のバリエント、CENP-A から構成されるヌクレオソームの引っ張りに対するヌクレオソームの応答を調べた。図 1 に示したように、ヌクレオソームに巻き付いた DNA の両端間の距離を反応座標に取り、DNA の解離に必要な自由エネルギーとその過程でのヒストンと DNA の相互作用

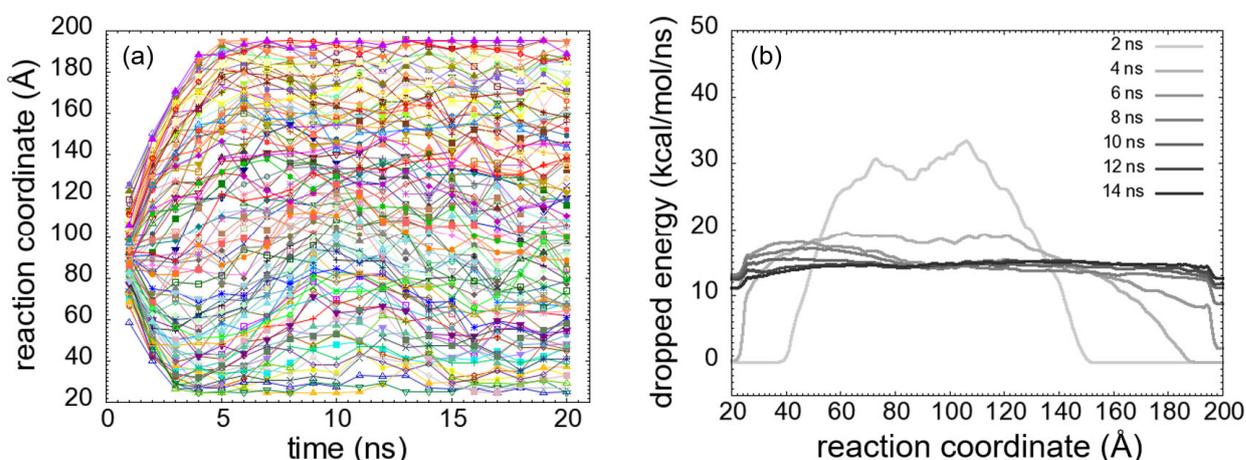


図 2 (a) 100 個のレプリカの反応座標上での位置の時間遷移。異なる色は、異なるレプリカを表す。5 ナノ秒で反応座標全体をレプリカが埋め尽くしているのが分かる。(b) 反応座標上の各位置に付加したエネルギーの変化。1 ナノ秒間に付加したエネルギーの値を表示。14 ナノ秒後では、反応座標全体にわたって一様にエネルギーが付加されているのが分かる。

用の変化を調べた。計算には、ポスト「京」重点課題 1 で高度化した分子動力学計算プログラム SCUBA を用いた。SCUBA に実装されている ABMD 法 (メタダイナミクス法の一つ) により、ヒストンに巻き付いた DNA をゆっくりと解離させていった。ABMD 法では、ある状態の構造がサンプリングされるとその状態に対応する反応座標にペナルティとなるエネルギーを付与していくことで、まだサンプリングされていない領域をサンプリングすることを促す。計算は、反応座標に対して付与したエネルギーと実際の自由エネルギーの合計値が一樣になるまで行う。このような方法で様々な状態を効率よくサンプリングするために、同時に 100 個のレプリカ (各レプリカには 48 ノードを割り当て) を用いて計算し、その結果を解析した。図 2a に示すように、時間とともに 100 個のレプリカが反応座標全体に拡がり、様々な構造がサンプリングされているのが分かる。図 2b には、1 ナノ秒間に各反応座標に付与したエネルギーの時間遷移を示す。この図から、14 ナノ秒後には付与するエネルギーが反応座標に対して一定になっているのが分かる。これは、反応座標上をランダムウォークするための必要条件が満たされたことを示す。次に、各反応座標まわりで十分な構造集団を得るために、反応座標を 2Å 刻みの区間に分け、各区間の位置に拘束をかけたアンブレラサンプリングを行った。十分な構造集団が取得できていれば、それから計算される自由エネルギー値

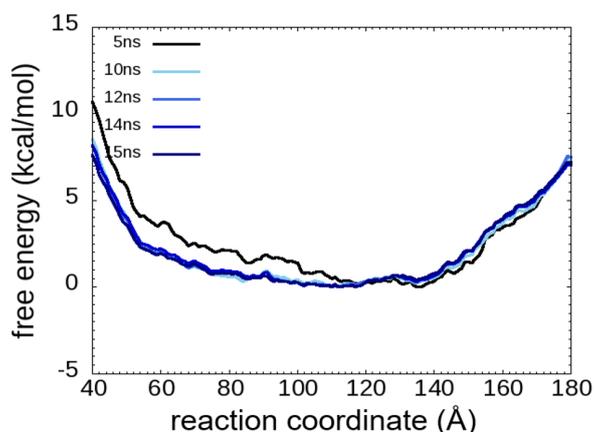


図 3 DNA 解離における自由エネルギー変化。サンプリング時間に対して、自由エネルギーが収束していく様子が分かる。

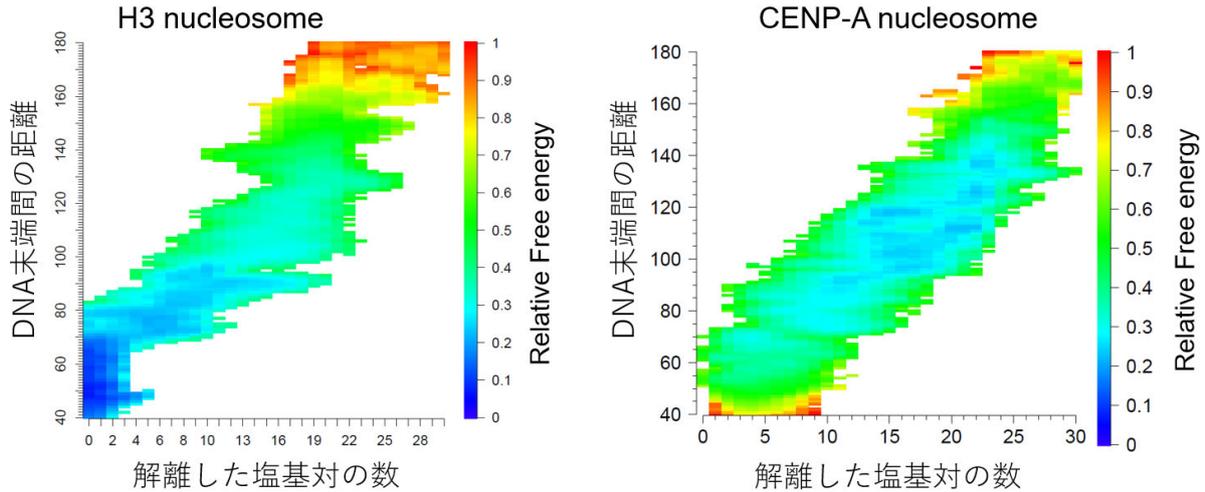


図4 自由エネルギー地形。左が H3 ヌクレオソーム、右が CENP-A ヌクレオソーム。エネルギーの低い領域は青色で示す。CENP-A ヌクレオソームは比較的能量差が少ない領域が広がっているのが分かる。また、10 から 25 塩基対の解離した状態で安定であることが分かる。

は変化しない。図3に示すように、15ナノ秒のアンブレラサンプリングで自由エネルギーの収束が見られた。その結果をもとに、以前に報告した H3 ヌクレオソームとの自由エネルギー地形の比較を行った。構造変化を理解しやすいように、自由エネルギーを、DNA 末端間の距離と解離した塩基対の数の2次元で示した(図4)。図4に示すように、H3 ヌクレオソームは、巻き付いた状態(図4の左下の青色)が最も安定であるのに対し、CENP-A ヌクレオソームは15から20塩基対がヒ

ストンから解離した状態が最も安定であることが分かった。また、自由エネルギーの差がx軸、y軸の変化に対して小さいことから、DNAの解離状態の変化が起こりやすいことが分かった。CENP-A ヌクレオソームは、染色体分配の時に形成される動原体に特異的に存在することが知られている。CENP-A ヌクレオソームは、他の動原体形成に必要なタンパク質への結合部位を提供しなければならない。本計算結果は、小さなエネルギー変化でDNAを解離した様々な構造を取ることができ、CENP-A ヌクレオソームの特徴をよく反映していると考えられる。さらに、解離状態を詳細に調べると、H3 ヌクレオソーム同様にDNAの解離に非対称性が見られた(図5)。2つの末端のうち片側のDNAが解離すると、10塩基対までDNAの解離が進行し、その後もう片方が10塩基対解離する、といった具合にDNAの解離が進むことが示された。しかし、この傾向はH3 ヌクレオソームに比べて弱く、状態間の自由エネルギー差が小さいことを反映して、様々な構造を取ることが示されている(図5)。

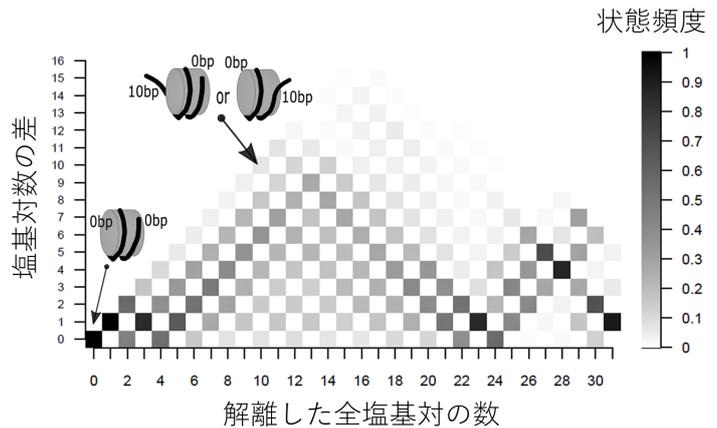


図5 ヌクレオソームの状態分布。縦軸は、両末端の解離している塩基数の差。四角の濃淡は、解離した総塩基数に対して、両末端の解離状態の頻度を表す。

次に、捩じれに対するヌクレオソームの応答を報告する。細胞核内の DNA は、絶えず引っ張りや捩じれといった力学的なストレスを受けている。例えば、RNA ポリメラーゼが DNA の遺伝子情報を mRNA に転写する際、ポリメラーゼの進行する先にはプラスのトルクが、進行した後ろ側はマイナスのトルクがかかると言われている。このような捩じれがヌクレオソーム構造の安定性にどのような影響を与えるかを明らかにすることが本計算の目的であった。シミュレーション計算は、CENP-A ヌクレオソームの計算同様に、SCUBA を用いて実施した。延べ 11 マイクロ秒の計算を行った。

まず、DNA が巻き付いた結晶構造から DNA が解離した状態の構造を、ABMD 法を用いて用意した。その各状態の構造に対して、図 6 に示すように DNA の末端にプラス、マイナスのトルクをかけてアンブレラサンプリングを行うことで、トルク存在下での DNA 解離の自由エネルギー地形を求めた。

大きなトルクをかけ過ぎると一気に捩じれてしまい、DNA が不自然に壊れてしまう。一方、トルクが小さすぎると DNA は捩じれない。試行錯誤の上、15 pN・nm 前後のトルク存在下で計算を行うことが適切であることが分かった。図 7 にプラス、マイナスのトルクをかけた場合の DNA の捩じれを示す。この図をみると、DNA の解離が進むのにしたが、トルクが無い時には DNA の捩じれに変化があまり見られないのに対し、プラス、マイナスのトルク存在下では DNA の捩じれがそれぞれ増加、減少することがわかる。次に、得られた構造アンサンブルを用いて自由エネルギー解析を行った。自由エネルギーは、プラスのトルク、マイナスのトルク、トルク無しの 3 条件で、反応座標が 150Å まではほぼ同様のカーブを描くこ

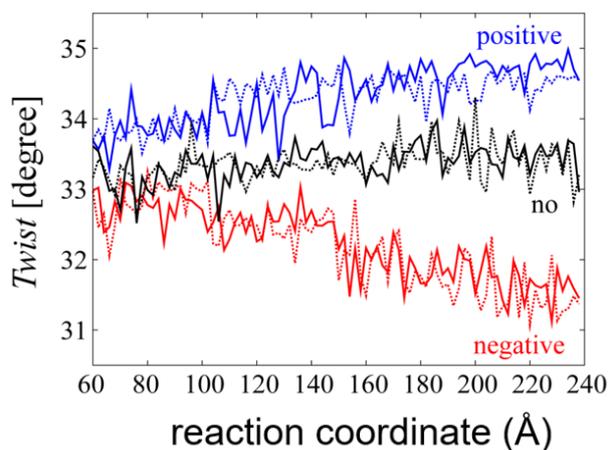


図 7 反応座標と DNA の捩じれ (Twist) の関係。DNA の解離が進むほど、捩じれが大きくなっていくのが分かる。

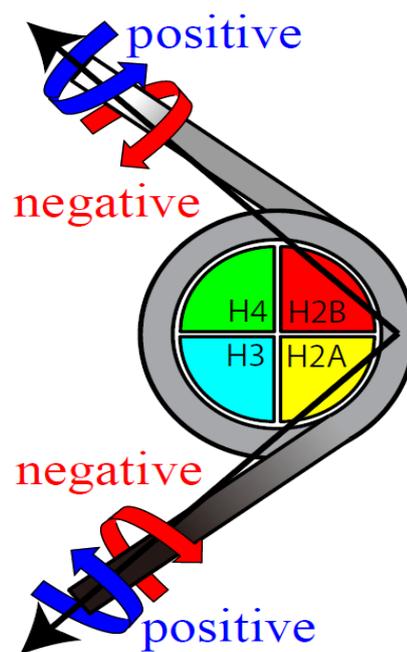


図 6 ヌクレオソーム両端の捩じり。青がプラスに捩じった場合、赤がマイナスに捩じった場合。

とが分かった。その後、反応座標が大きくなるにつれて、プラスのトルクの場合は自由エネルギーが増加しないのに対し、マイナスのトルクの場合はエネルギーの上昇がみられた。この振る舞いの違いを調べるために、引っ張りの計算と同様に、ヒストンから解離した塩基対の数で構造状態を見直した。結果、プラスのトルクの場合、反応座標が小さい領域でもヒストンから DNA が解離していることが分かった。図 8 に示すように、プラスのトルク存在下の反応座標 120 Å の DNA 解離状態は、マイナスのトルク存在下の反応座標 165 Å での解離状態と相当することが分かつ

た。つまり、プラスのトルク存在下の方が小さな自由エネルギーで解離が進行していくことが示された。一方、マイナスのトルク存在下では DNA 自体の構造が変形してヒストンと相互作用しやすくなり、解離が進行しにくいことが分かった。さらに、予備的な解析ではあるが、DNA とヒストンとの相互作用部位や相互作用エネルギーはトルクの有無やトルクの向きに関係しないことが示された。結果的に、解離の自由エネルギーを生み出しているのは、DNA 自体の変形のしやすさに起因することが示唆された。つまり、プラスに振じることで DNA 自体が硬くなり、曲がってヒストンに巻き付くには大きなエネルギーが必要となる。一方、マイナスに振じる場合、DNA 自体は曲がりやすくなるため、ヌクレオソームの構造をとるために必要な DNA の曲げに対するエネルギーが不要となる。結果、ヌクレオソーム状態が安定化されると結論付けられた。図 9 に、DNA の解離状態の詳細を表す。図 5 と比べて、対角線上にはっきりと状態があるのが分かる。横軸から離れるほど、片側の DNA が解離している状態を表しており、プラスに振じることで、片側の DNA が一気に解離した状態が多いのが分かる。また、15 塩基対片側が解離すると、もう片方が 15 塩基対解離する状態が多いことがはっきりと見えることができ、プラスのトルクが解離過程の非対称性を促進していることを表していると考えられる。このように、本計算でトルク存在下での DNA 解離の進行状態の詳細を明らかにすることができた。

以上、当初に掲げた、引っ張りと振じれに対する応答を調べる計画を達成することができた。引っ張り応答では、ヒストンバリエーションの違い

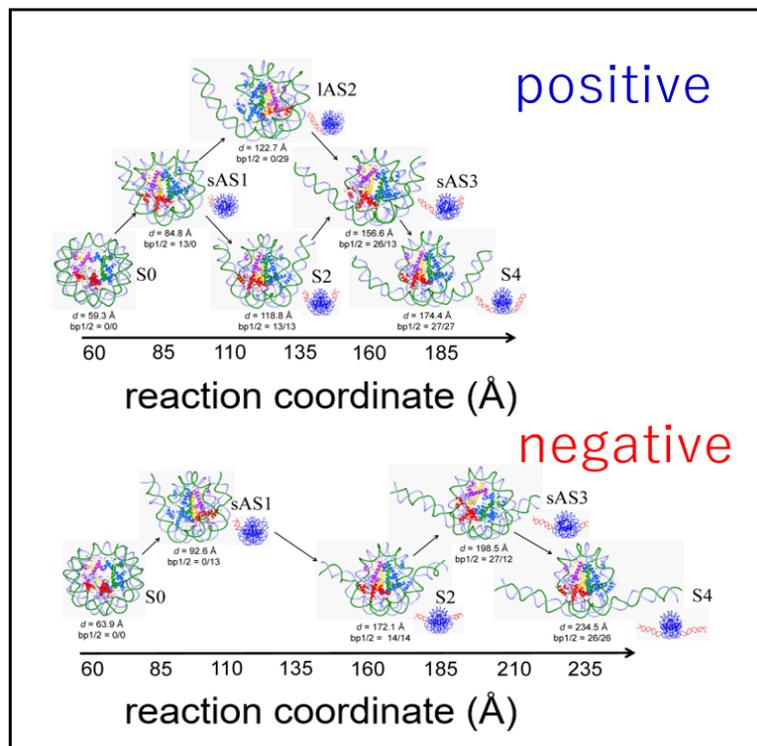


図 8 トルク存在下での DNA の解離状態。プラスのトルクの場合、DNA 末端間の距離 (= 反応座標) が短い状態で DNA の解離が進む。一方、マイナスのトルクの場合、DNA が解離するには末端間の距離が大きくなる。

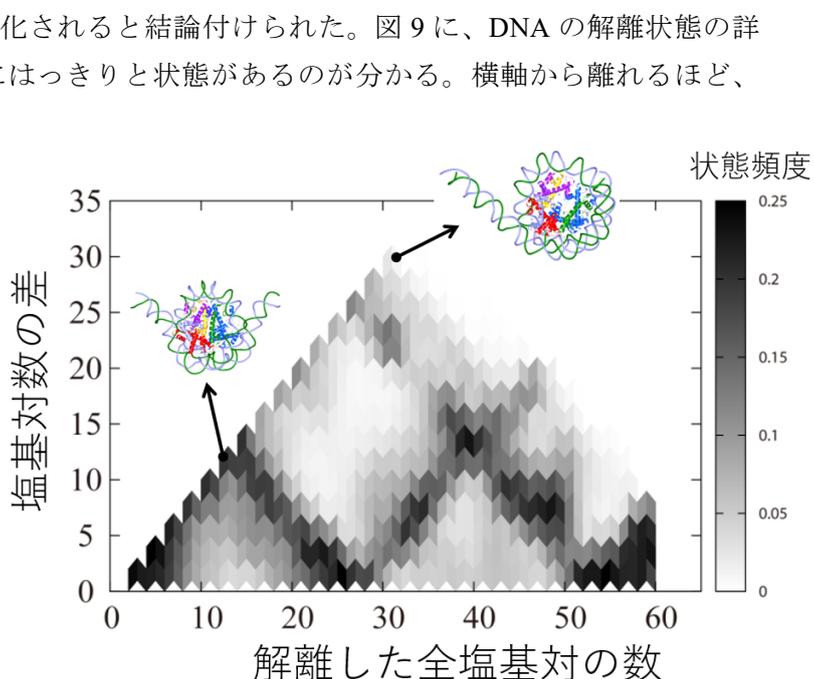


図 9 プラスのトルク存在下での DNA の解離状態。縦軸が、2つの DNA 末端領域での解離した塩基対の差を表す。色の濃淡は、図 5 と同じ。グラフに対応したヌクレオソームの解離状態を示す。

を明らかにすることができた。また、振じれ応答に対しては、振じれの方向によって、DNA の解離にしやすさや解離状態が大きく変わることが分かった。特に、プラスのトルク存在下、つまり、ポリメラーゼの進行に伴って生じるトルクによって、DNA は解離しやすくなっている様子がシミュレーション結果から示唆された。これらの結果は、DNA に作用する力の向きによっても、遺伝子の発現が制御される可能性を示唆している。このように、これまでは不可能であった大規模シミュレーションを京コンピュータで実行することができるようになったおかげで、より細胞内に近い状態での生体高分子のダイナミクスを明らかにすることができた。

⑤細胞内環境シミュレーション

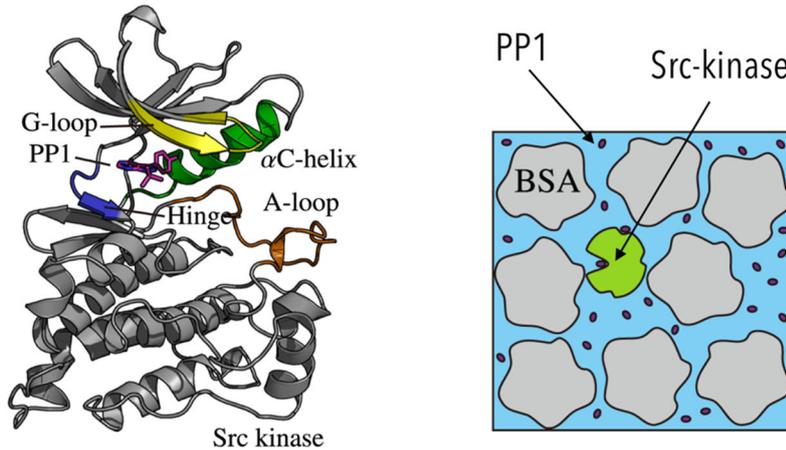
細胞内にあるタンパク質は、試験管内の希薄な溶液条件とは異なり、非常に混み合った環境（混雑環境）の中で互いを認識し機能している。試験管内では標的タンパク質に強く結合する化合物が、生体内では期待した薬効を示さない例は珍しくなく、薬剤候補化合物の設計でも混雑環境の影響の解明は急務である。ここ数年、混雑環境での生体分子の構造、動態、機能発現機構を調べる研究は実験的にも理論的にも急速に進歩しているものの、原子、分子レベルの理解は十分に得られていない。令和元年度は、1) 細胞内分子環境を取り込んだ構造モデリングを行うツールの開発と、2) 分子細胞環境がタンパク質・基質結合に及ぼす影響を明らかにするシミュレーションを実施した。

すでに確立した感のある溶液中あるいは脂質二重膜中のタンパク質の分子動力学と異なり、混み合った環境における複数のタンパク質の初期モデルを構築し、分子動力学計算を安定に長時間行うことは容易ではない。これまでの研究において、我々は Michigan 州立大学の Michael Feig 教授と連携して、マルチスケールモデリング法を開発してきた。この手法では、まず分子サイズを考慮した複数の球のシミュレーションを実施する。次にタンパク質の主鎖原子のみを含む粗視化分子モデルを用いた短い分子動力学を行い、原子衝突を回避する。さらに、水やイオン、代謝物などを加えた全原子分子モデルを構築して、安定で長時間の計算を実行する。従来の手法ではそれぞれのタンパク質の位置は最初の球の計算で決まっていた、粗視化分子モデルから全原子モデルに移行する過程では大きく位置を変化させることができなかつた。この問題を回避するために、我々は全てのステップを分子動力学ソフトウェア GENESIS で実行できるように新しい Python スクリプトを開発した。このスクリプト、Cellular-Scale Simulation Builder (CSSB)では、球を用いた計算、粗視化分子モデル、全原子モデルのいずれも GENESIS の温度一定あるいは圧力一定条件下で計算することができるため、より安定で信頼性の高い構造モデルが構築できることがわかった。この手法については現在論文としてまとめるべく準備中である。また、このツールを使うことによって安定な細胞質分子モデルを構築することができ、「富岳」を用いた大規模シミュレーションの実現につながった。

2 番目の課題では、分子混雑環境がタンパク質・基質結合に及ぼす影響を明らかにするために、ウシ血清アルブミン (BSA : Bovine Serum Albumin) タンパク質から成る混雑環境中でのキナーゼと阻害剤の結合機構を解析した (図 1)。比較のために、二次元レプリカ交換法 (gREST/REUS 法) を用いた計算による結果が得られている、Src キナーゼと PP1 阻害剤の結合を対象とした。希釈溶液系に加え、異なる数の BSA タンパク質からなる 3 つの混雑環境系を構築し、複数のマイクロ秒規模の分子動力学シミュレーションを実施した。希釈溶液系と混雑環境系の両シミュレーションで、PP1 阻害剤と Src キナーゼとの結合が見られた。得られた結合構造は、X 線結晶構造と良く一致しており、以前に報告されて

いる計算結果とも良く一致する。統計量が少なく、混雑環境が結合を促進するのか、あるいは妨げるかといった議論は出来ないが、希釈溶液系と混雑環境とでは阻害剤の結合経路が違っており、細胞環境での結合メカニズムを考える上で貴重な知見が得られた。

PP1 binding to Src-kinase in a crowded environment



PP1 interacts with hinge region during the binding

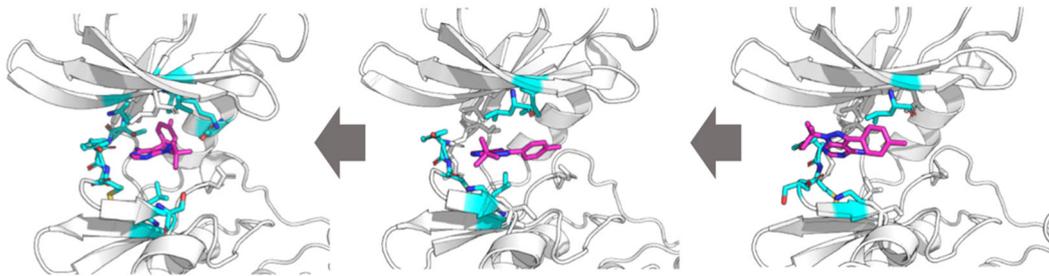


図1 ウシ血清アルブミン (BSA : Bovine Serum Albumin) タンパク質から成る混雑環境中でのキナーゼと PP1 阻害剤の結合の長時間分子動力学計算。下に典型的な PP1 結合のトラジェクトリを示す。

キナーゼ付近の阻害剤濃度は、結合に関する重要な指標である。希薄溶液では、PP1 分子は Src キナーゼの ATP 結合サイトだけではなく、タンパク質表面に広く分布している。系に含まれる BSA タンパク質の数が増えるにつれて、Src キナーゼ表面上の PP1 分布は大幅に減少し、逆にキナーゼ表面上の BSA 分布が増加している (図 2)。つまり、希薄溶液中に見られた PP1 分布の多くが BSA の分布に取って代わられている。この結果は、混雑環境下でキナーゼが複数の BSA タンパク質に取り囲まれることで、PP1 結合サイトが立体的に遮断されていることを示唆する。

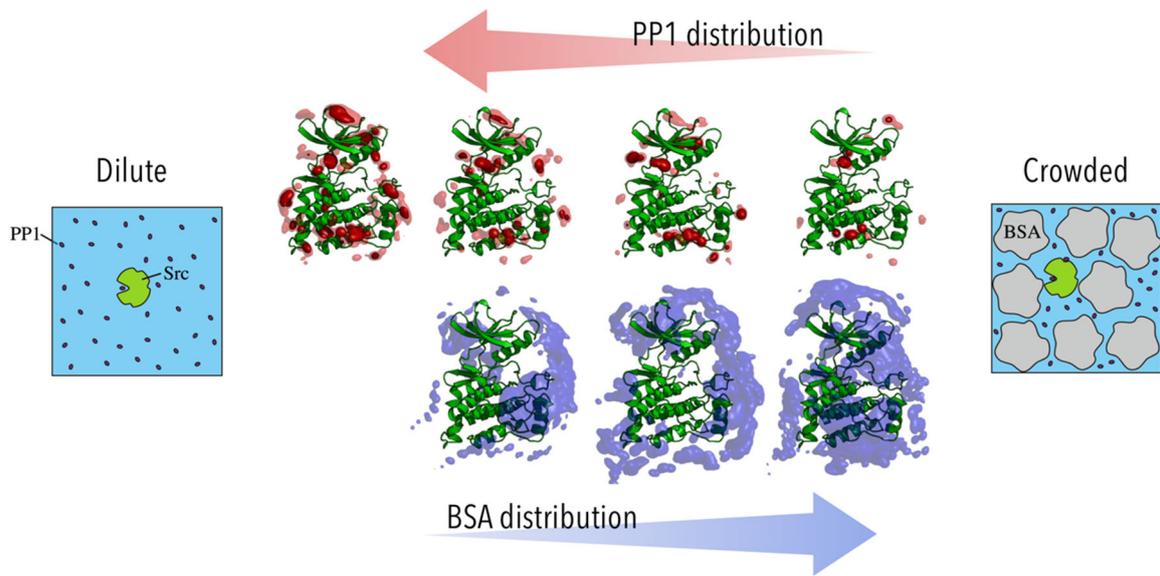


図2 希釈環境（左）と混雑環境（右）でのキナーゼタンパク質表面上の阻害剤分布変化。右にゆくほど混雑度合いは増す。阻害剤分子 PP1 の分布を赤、BSA タンパク質の分布を青で示した。混雑環境では、希釈環境で見られた阻害剤分子 PP1 の分布（赤）の多くが BSA の分布（青）に置き換わっている。

一方、BSA 表面の PP1 分布を調べた結果、PP1 分子は BSA タンパク質とも良く結合することがわかった。PP1 分子が BSA タンパク質表面に捉われることで、Src キナーゼ付近の有効濃度が低下していることを示唆する。以上の解析結果から、混雑タンパク質（off-target タンパク質）は標的タンパク質の阻害剤結合サイトを立体的に遮蔽するとともに、阻害剤を吸着することで標的タンパク質付近の阻害剤濃度を実効的に下げていることが明らかになった。さらに、PP1 阻害剤の拡散を調べた結果、混雑環境では、阻害剤と Src キナーゼ・BSA との相互作用により阻害剤の拡散が遅くなることも示した。現在、これらの研究成果を論文として執筆中である。

PP1 distribution on BSA surface

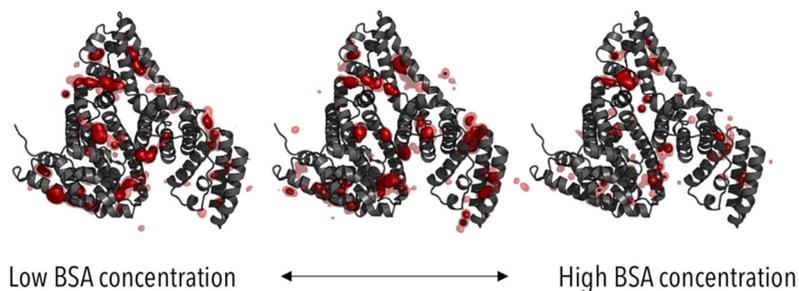


図3 混雑環境での BSA タンパク質表面上の阻害剤分布。右にゆくほど混雑度合いは増す。阻害剤分子 PP1 は BSA タンパク質表面にも広く分布していることがわかる。

これらの計算例から明らかになってきたのは、通常の計算や *in vitro* 実験が行われる希薄溶液中と実際の細胞環境の相違である。生体分子が固有の立体構造により固有の機能を発現するという古典

的な考えだけでは説明のできない環境要因も今後はきちんと考察していく必要があることがわかった。生命科学の大きな流れの中で、細胞内、細胞膜表面、クロマチン形成などにおける大きな構造体（ドロレット）の存在が頻りに議論されるようになってきたこととも関係し、より大きなスケールで、かつ、分子レベルの解像度で生命現象を解析していくことがより大きな意義を持つようになってきた。スーパーコンピュータ「京」やポスト「京」を用いた大規模なスケールの分子動力学計算はまさにこれらの問題と連動し、生命現象の新しい理解に貢献していくことができる。

(3) 創薬ビッグデータ統合システムの開発 (サブ課題 C)

①高精度薬剤デザイン (再委託先：国立大学法人京都大学大学院医学研究科 奥野恭史)

令和元年度は、創薬ビッグデータ統合システムのプロトタイプにサブ課題 A-C で開発される各種創薬計算手法の実装を行った。本システムでは、ドッキングシミュレーションと(拡張型)分子動力学シミュレーションの組み合わせによって、大規模な化合物ライブラリから標的タンパク質に対して高い結合親和性を有する医薬品候補化合物を予測する(図 1)。大規模化合物ライブラリの構築には、東大・津田チームの成果物である化合物生成手法(ChemTS 法)を用いる。

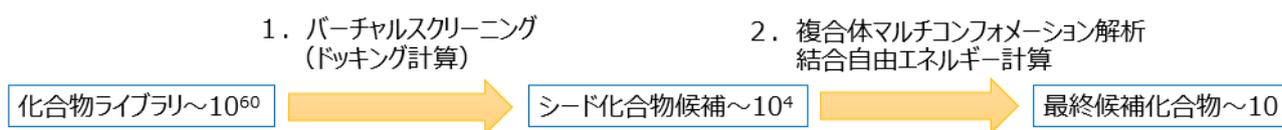


図 1 創薬におけるインシリコスクリーニングの概念図

大規模な化合物ライブラリを対象に、2種類のステップ (1. ドッキング計算 (計算コスト低)、2. 分子動力学シミュレーションによる複合体マルチコンフォメーション解析・結合自由エネルギー計算 (計算コスト高)) によって最終候補化合物への絞り込みを行う。

1 化合物あたりの計算フローは下記に示す 4 つのパートに分かれる (図 2)。計算フローの入力となるタンパク質構造は、産業技術総合研究所・広川チームの成果物であるタンパク質機能部位データベースから抽出する。

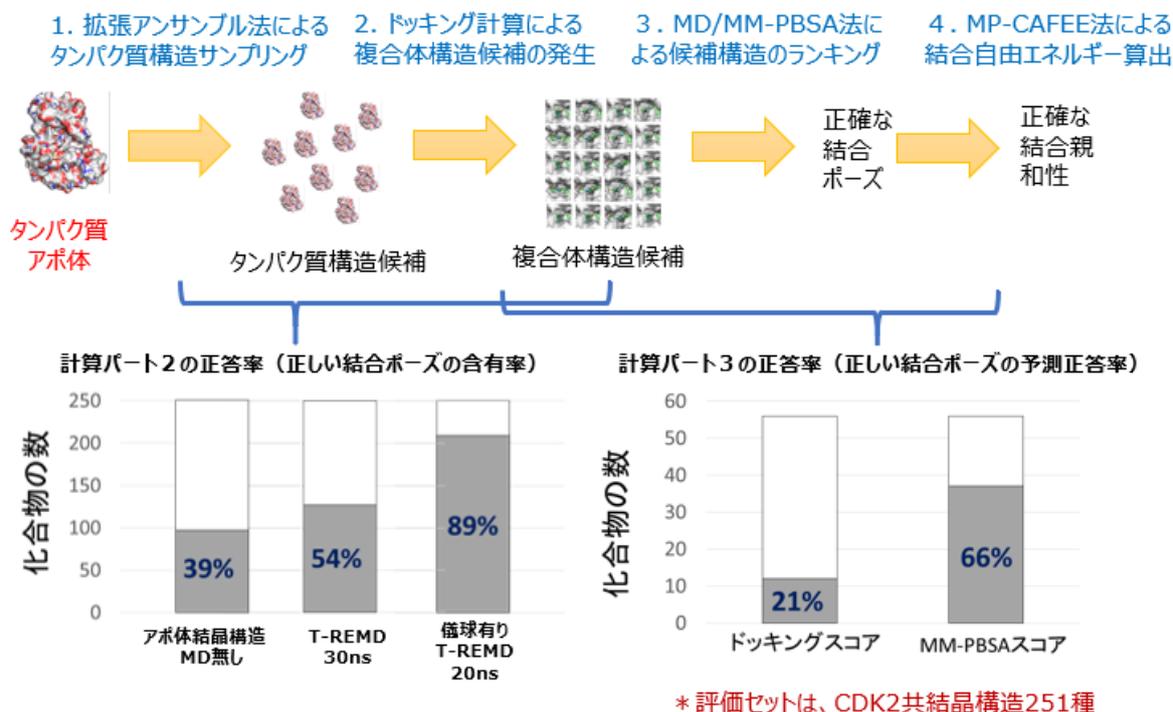


図2 1 化合物あたりの計算フロー

1. 拡張アンサンブル法によるタンパク質構造サンプリング：

サンプリング効率を高めることを目的とした拡張アンサンブル型分子動力学シミュレーションによって、タンパク質の動的性質を捉え、溶液中で取り得るタンパク質構造を複数取得する。医薬品が結合していないタンパク質アポ構造を初期値とした場合には、シミュレーション中に薬剤結合ポケットが閉じてしまう傾向があり、化合物ドッキングの際に結合ポーズのバリエーションを担保することが困難であった。そこで、本研究課題では、バーチャルリガンドを結合させてポケットを広げた状態で構造サンプリングを行う「Virtual Ligand Method (儀球有り T-REMD)」を開発し、論文発表に至っている (M. Araki, et al., *J. Comput. Chem.* 39(32):2679-2689, 2018)。

2. ドッキング計算による複合体構造候補の発生：

タンパク質-化合物ドッキングシミュレーションによって、医薬品化合物の結合ポーズ候補を複数発生させる。ドッキングプログラムとしては、無償のタンパク質-化合物ドッキングプログラム：rDockに基づいてスーパーコンピュータ「京」による大規模並列計算やタンパク質側鎖を可変的に取り扱うフレキシブルドッキングを可能にするための改良を施したバージョンを搭載している。また、1.「儀球有り T-REMD」との組み合わせによって、正しい化合物結合ポーズを高確率で発生させることに成功している (左下の棒グラフ)。

3. MD/MM-PBSA 法による候補構造のランキング：

2.で取得したタンパク質-化合物ドッキング構造モデルのそれぞれに対して化合物の結合安定性を評価し、熱力学的に最安定な化合物結合ポーズを予測する。本研究課題では、結合安定性評価の方法として、複合体の分子動力学トラジェクトリに基づいた MM-PBSA (Molecular Mechanics

Poisson-Boltzmann Surface Area) スコアリングを検証し、従来法 (ドッキングスコア) よりも大幅に予測正答率を上昇させることに成功している (右下の棒グラフ)。

4. MP-CAFEE 法による結合自由エネルギーの算出：

3.で取得した複合体構造モデルに対して化合物の結合親和性を高精度に予測する。結合親和性予測には、アルケミカル自由エネルギー摂動法によって化合物の結合自由エネルギーを算出する MP-CAFEE 法 (Massively Parallel Computation of Absolute binding Free Energy with well-Equilibrated states) を使用する。我々はこれまでに、MP-CAFEE 法をスーパーコンピュータ「京」に実装し、結合自由エネルギー計算を高速かつ効率的に実施できる環境を整えている。

令和元年度は、AI 創薬手法として、タンパク質-化合物結合ポーズの正誤 (図 2 のパート 2, 3) を判定する Convolutional Neural Networks (CNN) を実装し、従来法であるドッキング計算のみでは結合ポーズの予測に失敗した幾つかのリガンドにおいて、本法によって正しく予測する事に成功した。また、結合自由エネルギー (ΔG) 計算手法 (図 1 のパート 4) については、スーパーコンピュータ「京」での計算実績を有するアルケミカル自由エネルギー摂動法の一つである MP-CAFEE 法に基づいて、立体構造柔軟性が高いタンパク質に対しても化合物の結合自由エネルギーを高精度に推定する multi-conformations 法を開発し、論文発表に至っている (M. Araki, et al., *J. Chem. Inf. Model.* 56:2445-2456, 2016)。さらに、本年度は、図 2 のパート 1-3 に置き換わる手法として、薬剤の結合部位・結合構造を 1 種類のシミュレーションによって同時に推定する ColDock 法 (東京工業大学・北尾チームの成果物) を実装した。

また、各種創薬計算手法を、協力機関と連携し具体的な創薬ターゲットに適用した場合の性能評価を引き続き行ったところ、複数の創薬ターゲットにおいて本システムが高い予測性能を示すことを確認した。具体的には、がん、腎疾患、乾癬といった疾患を対象とした治療薬開発において、図 1、2 に示す計算フロー (の一部) を適用することで標的タンパク質に強く結合する薬剤候補 (群) を予測しており、その中の幾つかにおいては、協力機関によるセルベースドアッセイや生化学的試験によって一定の薬効が確認されている。また、味覚受容体の活性物質予測も行っており、様々な味覚受容体において新規アゴニストの導出に成功している。さらに、ゲノム医療・創薬への応用を実現する計算基盤の構築として、平成 30 年度に構築したゲノム医療分子シミュレーション基盤を実践的に活用する事で、計算フローのブラッシュアップを行った。具体的には、EGFR 遺伝子上の変異による薬剤耐性メカニズムや PI3KCA 遺伝子上に生じた多重変異によるがん化メカニズムを明らかにするための分子シミュレーションを行い、論文発表に至った。下記に詳細な研究内容を示す。

がんゲノム医療の普及により膨大な種類の遺伝子変異が同定されている。その中には、治療薬 (分子標的薬) の効果が明らかでないものも多く存在している。この種の変異に対しては、有効な治療薬を選ぶ方法がなく、がんゲノム医療を確立する上で大きな課題となっていた。本研究では、国内最大の肺がん遺伝子スクリーニングネットワーク LC-SCRUM.Japan と協力し、日本最大のがん臨床ゲノムデータベースに登録された 2,164 人の肺がん患者における EGFR 遺伝子の稀な変異の分布を明らかにした (図3上)。それぞれの遺伝子変異に対して、どのような治療薬の効果が高いの

かを細胞実験で調べ、薬剤に対する効果に変異によって大きく異なることを明らかにした。さらに、各々の変異体に対する薬剤感受性をスーパーコンピュータ「京」により予測した。具体的には、EGFR野生型-オシメルチニブ (EGFR阻害薬)共結晶構造 (PDBID=4ZAU) に基づいてEGFR変異体9種類の立体構造をモデリングし (図3左下)、MP-CAFEE法によってオシメルチニブとの結合自由エネルギー (ΔG) を評価した。算出した ΔG を実験的に測定された細胞増殖50%阻害濃度 (IC_{50}) に対してプロットしたところ、両者の間に良好な相関関係が見られ、薬剤との結合親和性が低下するEGFR変異体を発現させた細胞では、薬剤が効きにくくなる傾向が見出された (図3右下)。以上の結果より、個々の遺伝子変異に対して適した薬剤をMP-CAFEE法によって高精度に予測できる可能性が示唆された事に加えて、一連のシミュレーションから得られたEGFR変異体-薬剤結合に関わる立体構造情報は、今後、耐性を克服する新規薬剤の分子設計に役立つと期待される (Ikemura S., et al., *PNAS*, 116(20):10025-10030, 2019 : 慶應義塾大学医学部 安田浩之専任講師、副島研造教授との共同研究)。

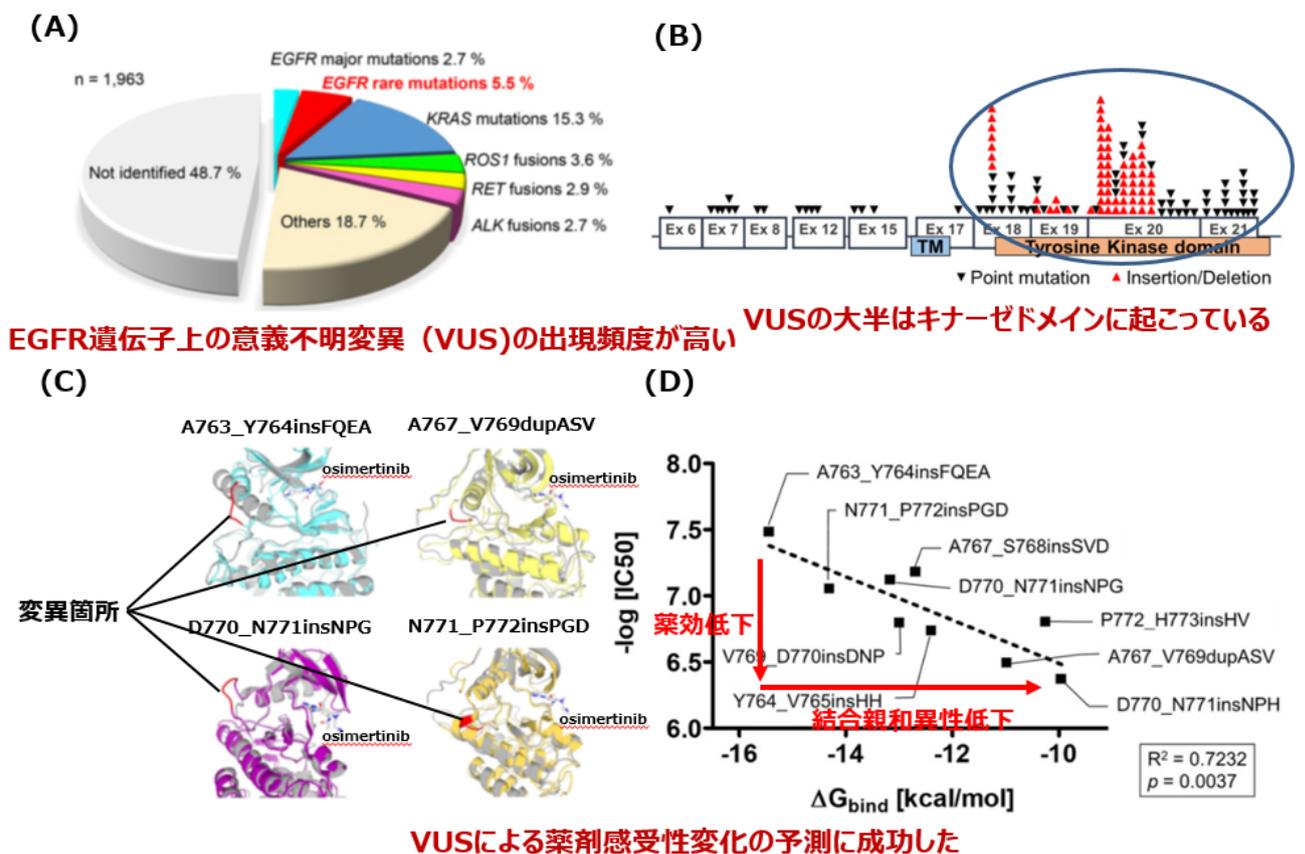


図3 (A) NSCLC 患者における遺伝子変異の出現頻度、(B) EGFR 遺伝子における意義不明変異の分布、(C) EGFR 変異体の立体構造モデリング。野生型-オシメルチニブ共結晶構造に基づいて、薬剤耐性変異体をコンピュータ上でモデリングした。薬剤をスティックモデルで表示し、変異部位を赤色でハイライトした。(D) EGFR キナーゼタンパク質とオシメルチニブの結合親和性 (結合自由エネルギー) の予測値と実験的に測定した薬効値 (50%阻害濃度の対数) のプロット。各点は EGFR 変異体 9 種類それぞれに対応している。

これまでがん遺伝子には従来単独で変異が生じることが多いと考えられてきたが、6万例（150がん種以上）を超える大規模ながんゲノムデータを対象とした遺伝子解析により、PIK3CA など一部のがん遺伝子では複数の変異が生じやすい傾向が明らかになり、単独変異よりもより強い下流シグナルの活性化や当該遺伝子への依存度が認められた（図 4A）。そこで、本研究では代表的ながん遺伝子である PIK3CA に着目し、野生型、R88Q、H1047R、R88Q/H1047R の分子動力学シミュレーション（500ns×3）を行った。PIK3CA では、ABD、Kinase、iSH₂ といった複数のドメイン同士が 4 次構造を形成する事により、不活性化状態を安定化しているが、2 重変異体 (R88Q/H1047R) では、ABD ドメインと Kinase ドメインをつなぐ水素結合の不安定化が認められた（図 4B）。また、各変異体の ABD、Kinase、iSH₂ ドメインの配向を野生型と比較すると、2 重変異体では PI3KCA を構成するドメイン同士のパッキングが弱まってキナーゼドメインが露出する傾向が見られた（図 4C）。以上の結果より、PIK3CA 遺伝子に複数変異が生じた場合、野生型や単独変異と比べて活性化状態に移行しやすい事が示唆され、分子レベルでの発がん機構を推定する事に成功した（Saito Y., et al., *Nature*, In press, 2020 : 国立がん研究センター 斎藤優樹任意研修生、片岡圭亮分野長との共同研究）。

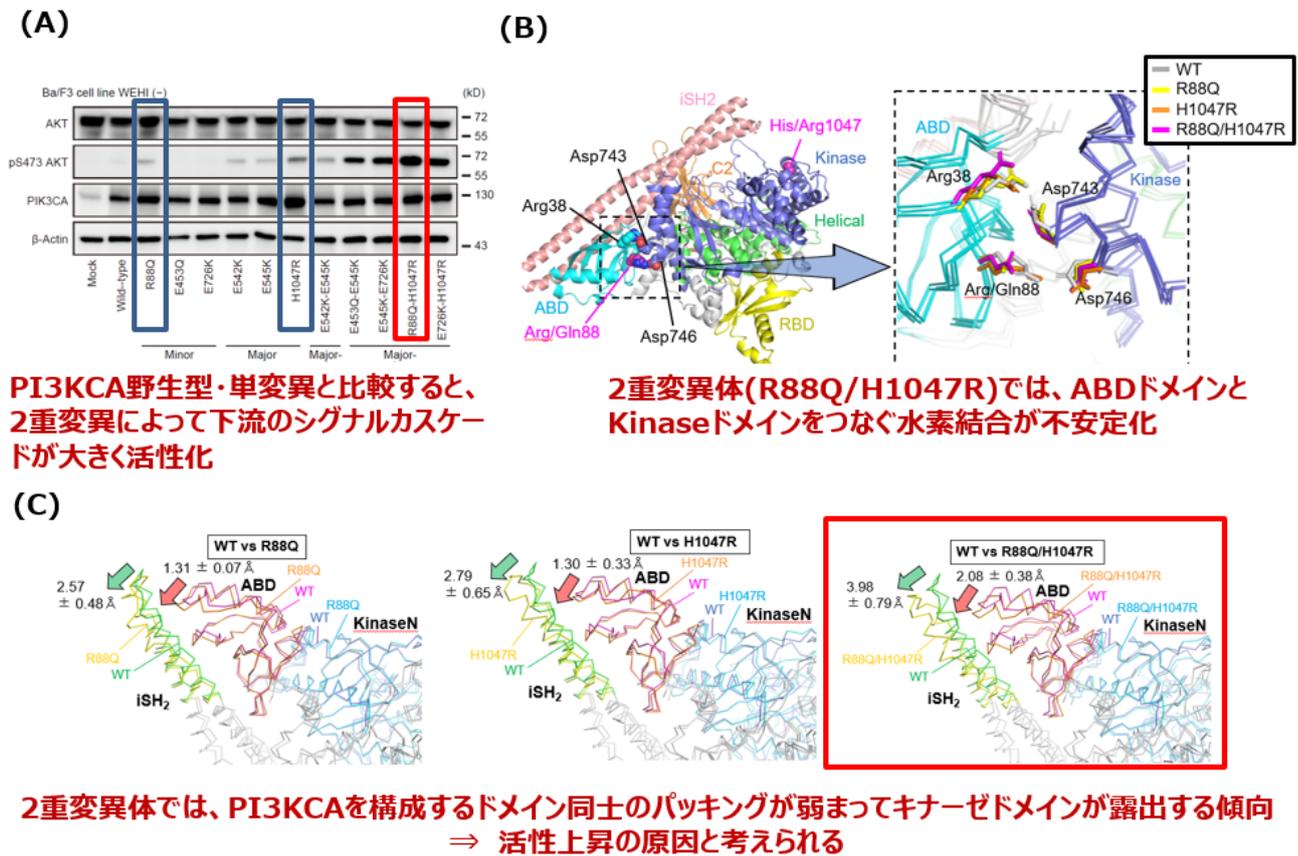


図 4 (A) 免疫沈降実験、(B-C) PIK3C α 野生型・変異体の MD シミュレーション(500ns×3)結果 (B) ABD, Kinase ドメイン間で形成される水素結合の比較、(C) ABD, Kinase, iSH₂ ドメインの配向の比較

②機能制御部位データベース（再委託先：国立研究開発法人産業技術総合研究所 広川貴次）

令和元年度は、(1) 機能部位予測データベースのシステム改良とサブ課題 C①と連携した創薬ビックデータ統合システムへの組み込み、(2) アンサンブル構造を活用したバーチャルスクリーニングや

ドッキングによる機能部位予測のシステム実装、並びに共同研究を通じて創薬標的タンパク質への適用性能評価と手法の高精度化に取り組んだ。以下に詳細の成果を述べる。

(1) 機能部位予測データベースのシステム改良とサブ課題 C①と連携した創薬ビックデータ統合システムへの組み込み

これまで構築した DrugBank に基づく機能制御部位データベースは、タンパク質構造データをヒトタンパク質データへ変換し、ホモロジーモデリング法により、野生型配列及び欠損部位の修復と最適化し、リガンド結合部位 (SiteFinder 及び PLBindex 法を適用)、ドラッグビリティスコア (PLBindex 値)、rDock によるドッキング計算環境が実装されていた。このシステムを創薬ビックデータ統合システムへの組み込み、ユーザーが簡便に利用可能なデータベースへ改良すべく、コマンドラインユーザインタフェース (Drug Docking pipeline system: DDpipe) を開発した。DDpipe は、創薬標的タンパク質データベース内の標的分子モデル構造に対する候補化合物のドッキングシミュレーションを半自動で行う計算パイプラインで、モデル構造情報の表示や、この抽出情報に基づくドッキングシミュレーションのインプットファイル作成、そしてドッキングシミュレーションの実施といった一連の処理を行う Linux 環境向けのコマンドラインプログラムである。使用したプログラム言語は、erl v5.16.3, perl-DBI, perl-DBD-MySQL. x86_64, Bash、データベースは、MySQL Ver 14.14 Distrib 5.7.27、ドッキングプログラムには、rDock v2013.1 を用いた。

DDpipe の主な機能分類は、以下の通りである。

1. DrugBank モデリングデータベースからの情報取得と表示
2. DrugBank モデリングデータベースからのドッキングシミュレーションの実施用のデータ取得
3. 取得した情報を基にしたタンパク質-リガンド間ドッキングシミュレーション実行

図 1 は、DDpipe 実行コマンドリストである。これらのコマンド実行環境は、sasanqua サーバーへの実装を完了した。

Commands	
list	:display the list of protein names and ids in the database
search	:show information of docking target proteins
parameter	:generate input files of docking simulation
get-model	:copy a model structure to CWD(current working directory)
analysis	:process the .score file to display the max and min of each ligand
-h/-help	:show the help message
Options and target (at 'DDpipe list')	
-db	: display the list of protein names and ids in database
Options and target (at 'DDpipe search')	
-u <UniProt ID>	: show information associated with the UniProt ID
-n <UniProt Name>	: show information associated with the UniProt title
-d <DrugBank ID>	: show information associated with the DrugBank ID
Options and target (at 'DDpipe parameter')	
-idu <UniProt ID>	-include/exclude : generate parm. files extracted by/except for the UniProt ID
-idp <PDB ID>	-include/exclude : generate parm. files extracted by/except for the PDB ID
-idd <DrugBank ID>	-include/exculde : generate parm. files extracted by/except for the DrugBank ID
-target <parameter file>	: display information of target proteins selected by parameter file
-generation <parameter file>	: generate input files for docking simulation
Options and target (at 'DDpipe getmodel')	
<PDB ID>	: obtain a model structure file (.pdb) from model data library
Options and target (at 'DDpipe analysis')	
<.score>	:calcutate the max and min S ^{total} values by reading of .score file

図 1 DDpipe のコマンド一覧

(2) アンサンブル構造を活用したバーチャルスクリーニングやドッキングによる機能部位予測のシステム実装

DDpipe を用いた複数の構造鋳型や化合物ライブラリを用いたバーチャルスクリーニングやドッキングシステムの実装例を以下に示す。この例は、本課題で構築した DrugBank に基づく標的データベース中から標的タンパク質 Cyclin-dependent Kinase (以下、CDK2、Uniprot ID: P24941) を選択し、CDK2 に結合可能な医薬品を FDA の 1616 品目のデータベース (ZINC15 よりダウンロード) からスクリーニングするものである (図 2 右)。まず、標的タンパク質検索では、parameter.inp に Uniprot ID をはじめバーチャルスクリーニングで重要な構造の解像度やドラッグビリティのスコア (PLB) の上限下限値、そして標的そのものが存在しない場合でも配列相同性 (Identity) を利用することより類縁タンパク質を代替に実施することが可能である。図 2 左は、CDK2 を検索する際のパラメータファイルの記入例と検索結果を示している。検索結果では、2 つの CDK2 の X 線構造 (3QTU、3RNI) がヒットしている。本システムでは、このように 2 つ以上の構造でもアンサンブル構造として入力構造に利用が可能となっている。

DDpipe parameter -target parameter.inp

タンパク質構造予測データベースから目的のパラメータを有する構造候補を検索。ポケットは自動的に検出されている

```
(base) [mki@sasanqua 190928_CDK2_step2]$ DDpipe paramete
=====
The target protein number is 2
=====
UniProt ID: P24941
=====
PDB ID: 3QTU (chain A)
Site      : 1
PLB      : 4.74831
Size     : 438
Pocket Residues:
ILE10:GLY11:GLU12:GLY13:THR14:TYR15:VAL18:ALA31:LYS33:ILE35:ALA48:GLU51:ILE52:LEU55:LYS56:L
:GLU81:PHE82:LEU83:HIS84:GLN85:ASP86:LYS89:HIS125:ASP127:LYS129:GLN131:ASN132:LEU134:LEU143
151:PHE152:GLY153:VAL154:VAL164:THR165:LEU166:ARG169
=====
PDB ID: 3RNI (chain A)
Site      : 1
PLB      : 4.65434
Size     : 925
Pocket Residues:
ILE10:GLY11:GLU12:GLY13:THR14:TYR15:GLY16:VAL18:ALA31:LYS33:ILE35:ARG36:LEU37:ASP38:THR41:V
:ILE63:VAL64:LYS65:LEU66:ASN74:LEU76:LEU78:PHE80:GLU81:PHE82:LEU83:HIS84:GLN85:ASP86:LYS88:
EU128:LYS129:PRO130:GLN131:ASN132:LEU134:LEU143:ALA144:ASP145:PHE146:GLY147:LEU148:ALA149:A
6:ARG157:THR158:TYR159:GLU162:VAL163:VAL164:THR165:LEU166:TRP167:TYR168:ARG169:ALA170:ILE17
P185:TRP187:SER188:GLU195:ARG199:ARG200:ALA201:PRO204:GLY205:ASP206:SER207:GLU208:ILE209:AS
```

パラメータファイルの内容

IDTYPE	uniprot_id
IDFLAG	include
ID	P24941
Resolution(A)	2.5
Identity(%)	98.0263157894737
Method	X-RAY_DIFFRACTION
PLB_UPPER	4.74831
PLB_LOWER	4.600

ベンチマークプロトコル

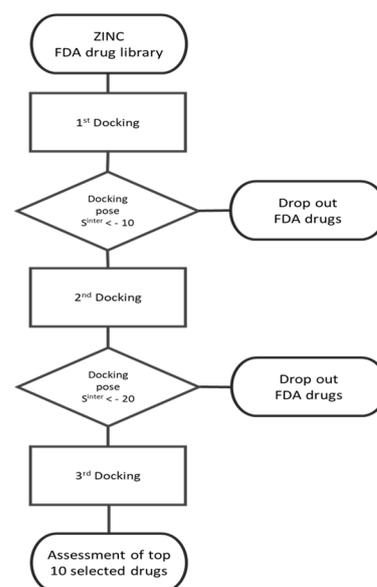
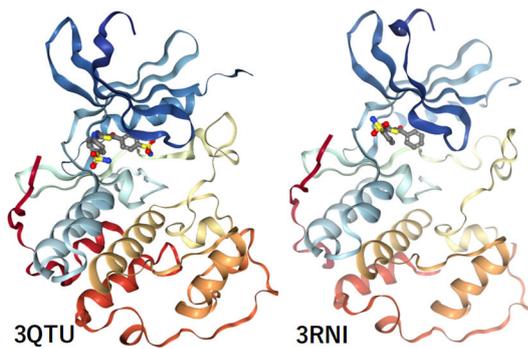


図 2 DDpipe によるバーチャルスクリーニング実装例。標的タンパク質探索コマンド例 (左) とバーチャルスクリーニングワークフロー

標的タンパク質を決定後、スクリーニングライブラリとなる化合物データベースと共に rDock を使用したタンパク質-リガンドドッキングシミュレーションが実行される。ドッキングシミュレーションは計 3 回 (1st, 2nd, 3rd docking) 実施し、1 回目 (各リガンドのドッキング試行数 5 回) と 2 回目 (各リガンドのドッキング試行数 10 回) において rDock のドッキング評価関数の Sinter を使用し、評価関数を満たすリガンドライブラリの絞り込みを行う。1st docking では、3QTU と 3RNI のどちらかのモデル構造で評価関数を満たすリガンド (集合和) が出力され、2nd docking の入力となる。今回のベンチマークでは、3QTU もしくは 3RNI のどちらかにドッキングポーズを示した化合物数 (和集合) は、1312 個であった。2nd docking では、1312 個が入力リガンドとなり、その結果、両方にドッキングポーズを示した化合物数 (共通部分) は、158 個であった。最後の 3rd docking では、158 個を入力として、ドッキング試行回数を 50 回に設定し、ドッキングを実施し、評価関数 (Stotal) の値によりドッキング結果を評価し、Top10 のリガンドに関して、医薬品として出力した (図 3 右)。副作用やドラッグリポジショニングの場合は、異なる標的タンパク質の阻害剤が注目されるが、10 品目の中には、Cabozantinib や Imatinib などの既知 Kinase 阻害剤もスクリーニング結果に含まれており、選択性とポジティブコントロールも同定しており、システム実装として精度も確認することができた。



S^{total} Top 10 FDA ligands (ordered in increasing the S^{total} value)

<ZINC ID>	<Ligand name>	<Function>
ZINC000003830391	Cefadroxil	Inhibition of bacterial cell wall synthesis
ZINC000000000469	Orciprenaline	A beta-adrenergic agonist
ZINC000003830500	Cephalexin	Inhibition of bacterial cell wall synthesis
ZINC000005733652	Diosmetin	Not Available
ZINC000030691754	Betrixaban	competitive and reversible inhibition of the factor Xa
ZINC000003873160	Losartan	Prevention of the angiotensin II binding to the AT1 receptor
ZINC000001481815	Exjade	Not Available
ZINC000070466416	Cabozantinib	Non-specific tyrosine kinase inhibitor
ZINC000019632618	imatinib	Kinase inhibitor
ZINC000018324776	Vardenafil	Inhibition of the cGMP specific phosphodiesterase type 5t

図 3 ベンチマークで用いた 2 つの CDK2 の X 線構造 (左) と FDA ライブラリへのバーチャルスクリーニング結果 (上位 10 品目)

(3) 共同研究を通じて創薬標的タンパク質への適用性能評価と手法の高精度化

これまで「アンサンブル構造に基づくバーチャルスクリーニングやドッキングの精度を高める構造選定法」の適用例として、多くのがんに関わる MAP キナーゼ経路 (RAS/RAF/MEK/ERK) の最下流にある ERK2 を標的タンパク質として、バーチャルスクリーニングで得られた候補化合物にアロステリック部位に作用する新しい“抗がん剤の種”の創製を目指してきた。実施内容は、①MD 計算とサンプリング法により構造変化のパスウェイを獲得、②サンプリング構造からのアロステリック部位同定と化合物データベースによるバーチャルスクリーニング (in silico スクリーニング) の実施である (図 4)。

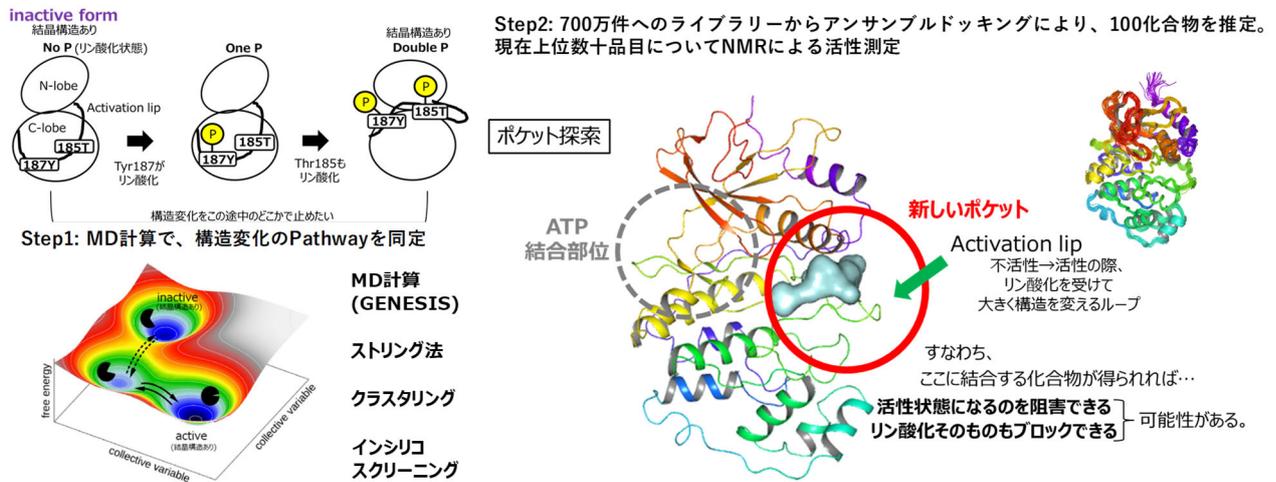


図 4 ERK2 を標的としたアンサンブル構造を活用したバーチャルスクリーニングの流れ

本年度は、適用性能を評価するために上記バーチャルスクリーニングの選定化合物 50 品目を NMR および MAPK Cascade Panels による評価を行った。図 5 は、MEK1-ERK2-基質のカスケードに対する化合物の阻害%を化合物ごとにプロットしたものである。その結果、100 uM (白抜き棒グラフ) にて 50%以上阻害がかかる化合物が 8 個、見出された (赤い矢印にて指示)。また、NMR 実験にて ERK-pY に選択的に結合する可能性が示唆されている group2-#33 (オレンジの矢印にて指示) についても、100 uM 添加時に約 40%の阻害が認められた。

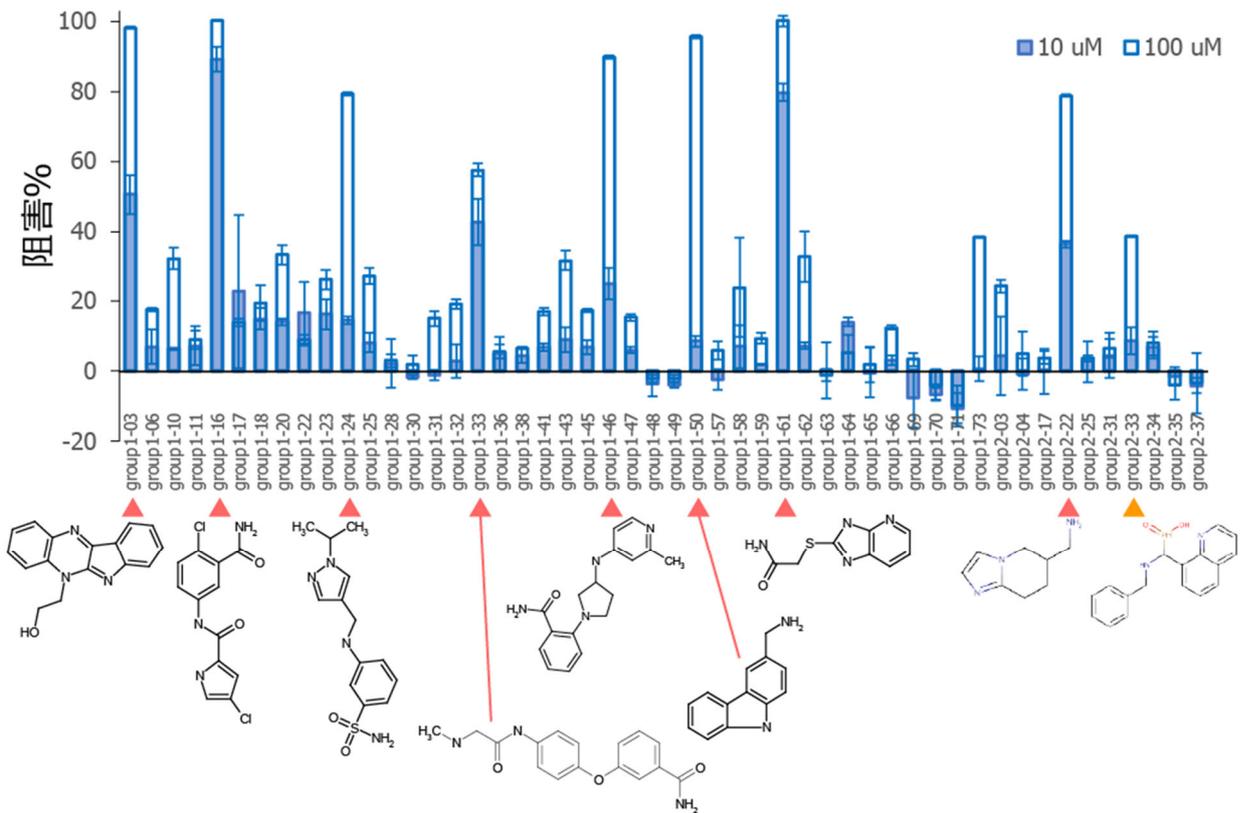


図 5 MEK1-ERK2-基質のカスケードに対する 50 品目の化合物の阻害評価結果

今後は、上記で注目した 9 化合物について、①化合物が ERK2 に結合するかを化合物観測の STD-NMR 実験、②化合物が ERK2 に部位特異的に結合するか、およびその結合部位を ERK2 観測の NMR 実験にて確認、③化合物が MEK1 による ERK2 のリン酸化を阻害するか（リン酸化された ERK2 が基質を阻害するステップではなく）を Phos-tag SDS-PAGE による ERK2 のバンドシフトを指標としたリン酸化アッセイにて更なる評価する予定である。本課題の目的に対し、ERK2 を通じて、ストリング法を中心としたアンサンブル構造及びドラッグブルポケット探索を融合した、新しいアロステリック中間体選定方法のプロトコル構築とバーチャルスクリーニングで選定した化合物群が、実際に実験による活性が確認されたことより、手法の適用性能が高く評価された。

また手法の更なる高度化を目指して、非天然構造状態を持つ標的タンパク質について適用を試みた。標的は、HIV ウイルス由来の Tat タンパク質で、単体では非天然構造状態であるが、転写の際、ヒト由来の P-TEFb、および TAR RNA と複合体形成する際に特定の折り畳み構造を形成する。本適用では、非天然構造状態から複合体構造の間の折り畳み中間体をレプリカ交換分子動力学シミュレーション (REMD) でサンプリングし、中間体アンサンブル構造に対してドラッグブルポケット探索の探索を試みた。これは前述の ERK2 のアロステリック部位同定より更なる高度な機能部位の同定となる (図 6)。

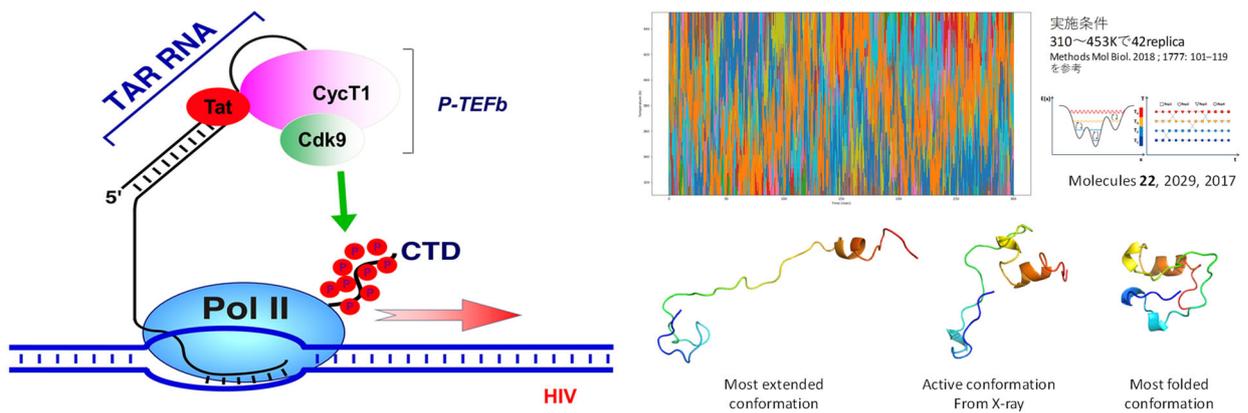


図 6 HIV 転写における Tat 複合体モデル (左) および REMD による Tat タンパク質の中間体サンプリング (右)

図 7 は、アンサンブル構造を RMSD による主成分分析の第 1-2 成分でランドスケープにマッピングした結果である。アンサンブル構造に対するドラッグブルポケット探索の結果、広範囲にわたってドラッグブルポケット探索を有する中間体構造が存在していることが予想された (図中では代表の 8 構造を表示)。特にアンサンブル構造の出現密度が高く複合体の構造に近いドラッグブル中間体構造の存在も確認でき、この状態を標的とすることで、Tat タンパク質の折り畳みを制御するケミカルシャペロン機能の阻害薬の探索が期待される。現在、市販ライブラリによるスクリーニングを実施しており、名古屋市立大学にてアッセイ評価に進む予定である。

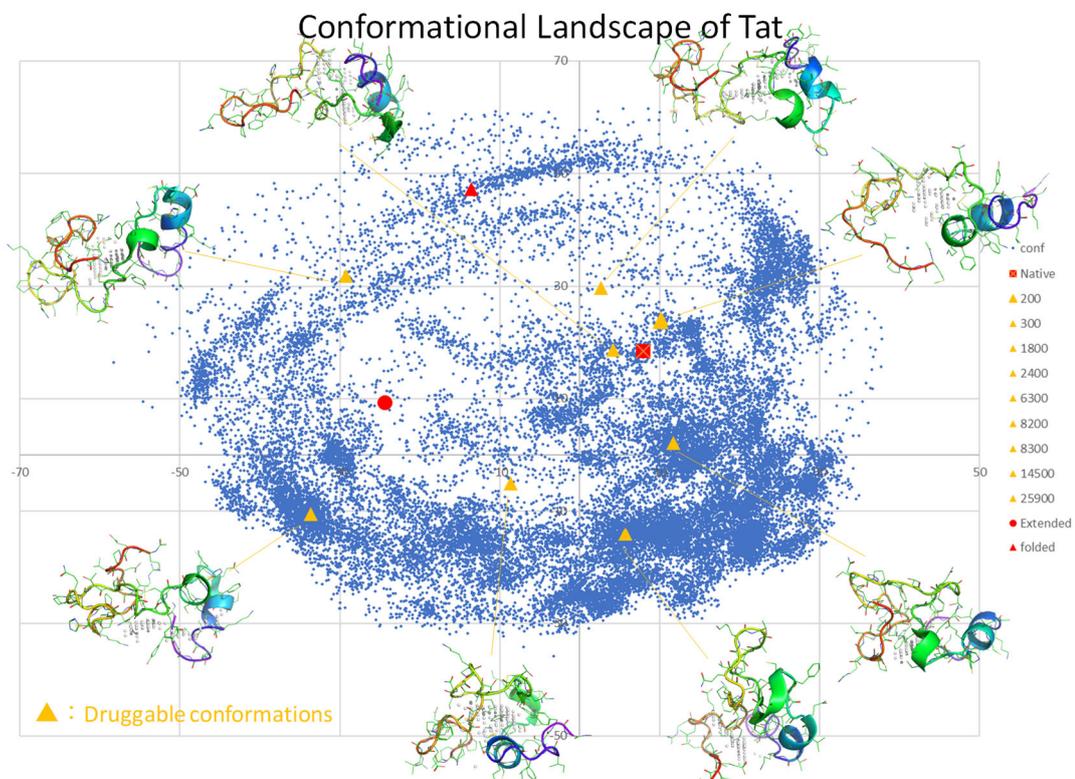


図 7 Tat タンパク質の中間体構造ランドスケープとドラッグブル構造の同定

③創薬関連ビッグデータ（再委託先：国立大学法人東京大学大学院新領域創成科学研究科 津田宏治）

本研究では、膨大なタンパク質構造と化合物の組み合わせから薬剤候補化合物を高速に探索する機械学習を基礎としたシステムを構築することを目的としている。令和元年度は、強化学習と分子動力学計算を用いた構造変化パスウェイ探索アルゴリズムを開発した。現在ポスト「京」プロジェクトで使用されている PaCS-MD 法は、短い時間の MD を直列に組み合わせることで、初期構造から目標構造への軌跡を短時間で得ることができるが、反面、局所最適構造にトラップされやすい傾向がある。本研究では、木探索を用いることで、一旦トラップされても、後戻りして抜け出すことを可能にする手法で TS-MD 法を開発した。また、実際のタンパク質構造に適用し、性能を評価した。

以下に、TS-MD 法の詳細を示す。TS-MD 法では、短時間の MD 計算を階層的に展開することで、初期構造からターゲット構造への構造変化パスウェイを探索する（図 1a）。TS-MD では、Selection, Expansion, Backpropagation の三段階を繰り返すことで、徐々に探索木を大きくしていく。一方、Pacs-MD では、独立な短時間 MD を複数走らせて、最もターゲットに近づいた数個を採用し、次回の短時間 MD の初期構造として用いる。そのような方式のため、TS-MD のように、これまでの探索過程を記憶しておらず、後戻りして、違う経路を選択するということができない。

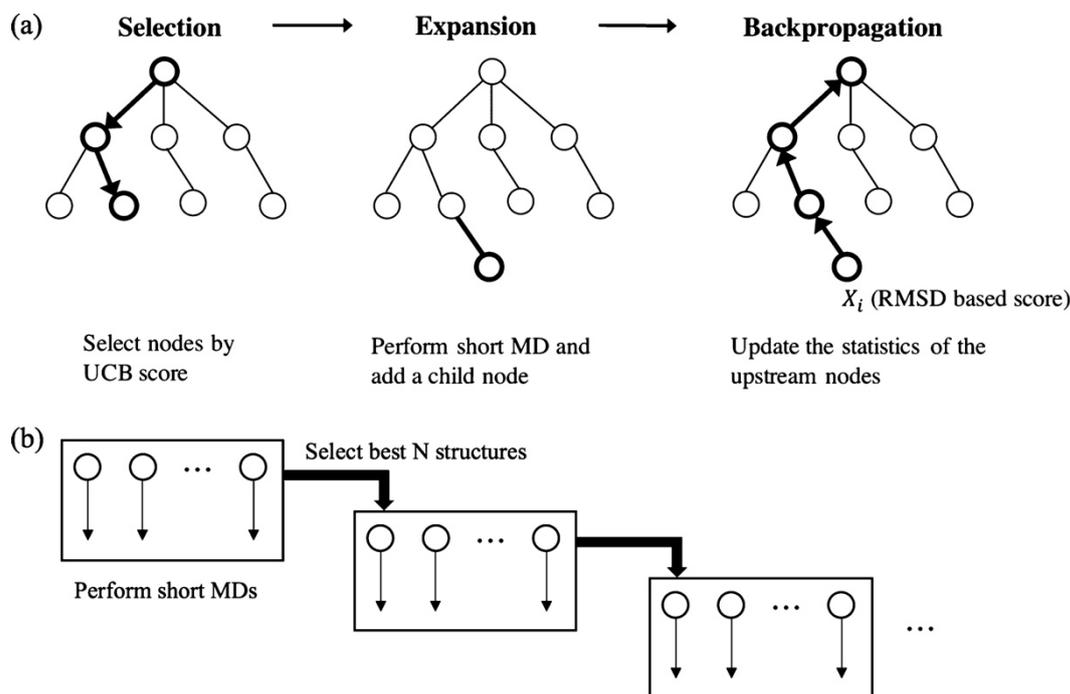


図 1 TS-MD(a)と、Pacs-MD(b)の動作の違い

TS-MD と、Pacs-MD の性能を比較するため、10 残基からなるタンパク質 Chignolin と 20 残基からなるタンパク質 Trp-cage のフォールディングに関して経路探索を行なった。これらの構造を図 2 に示す。

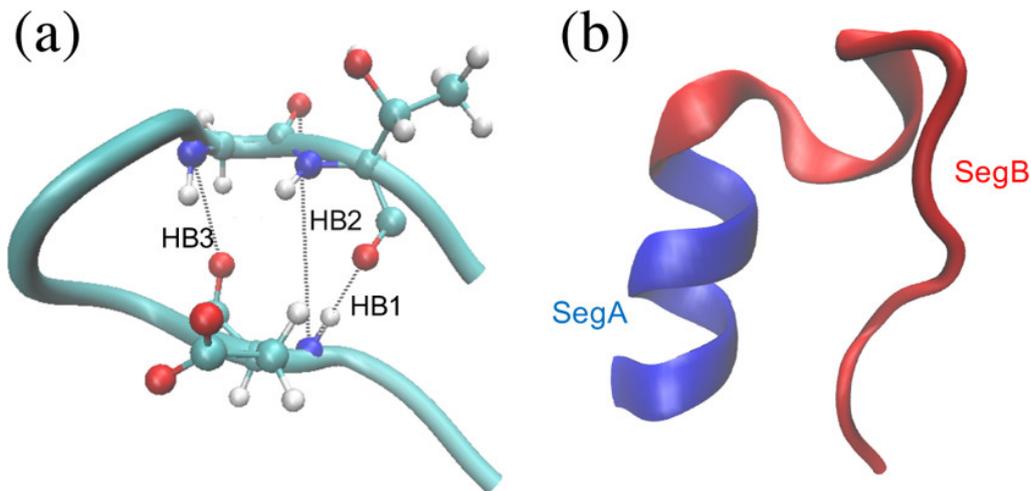


図2 (a)Chignolin の分子構造 (b)Trp-cage の分子構

Chignolin のフォールディングに関して、両手法の比較を 2000MD シミュレーションまで行なった。10 回の試行の中で、Pacs-MD は、9 回ターゲットから RMSD が 1Å 以内の構造に到着したのに対し、TS-MD は 10 回すべてで 1Å 以内の構造に到着した (図 3)。基準を厳しくして 0.5Å 以内とすると、Pacs-MD の到着回数は 4 回であったのに対し、TS-MD の到着回数は 8 回と、優れた性能を見せた。Trp-Cage に関しても、同様に、TS-MD の方が優れていた。

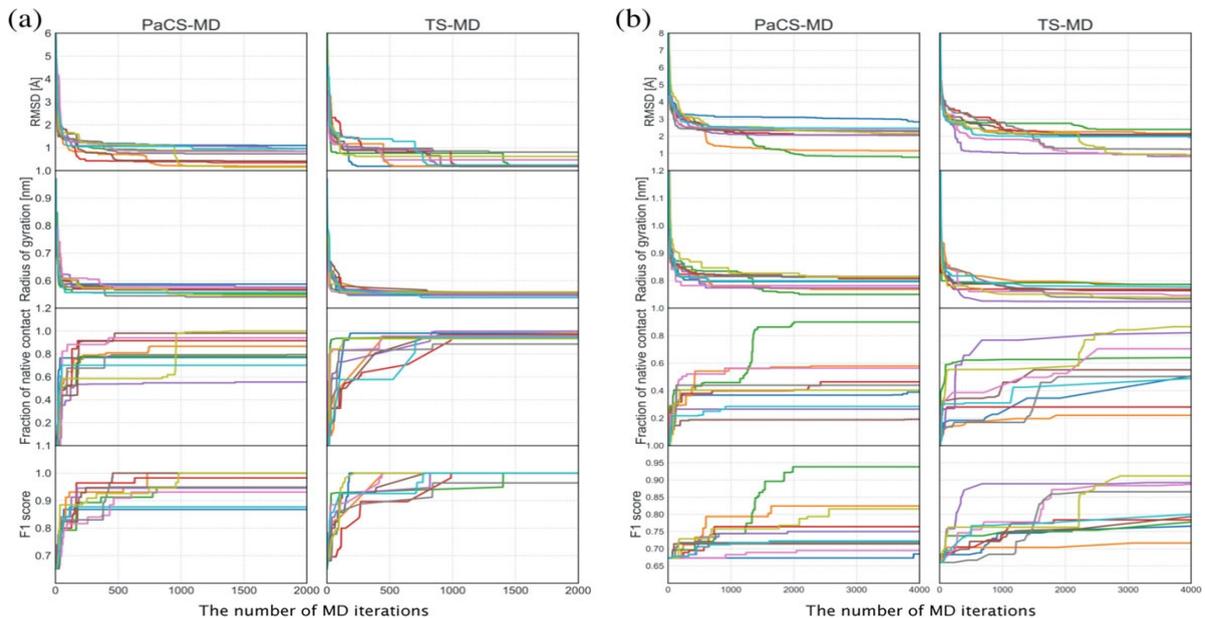


図3 (a)Chignolin に対する Pacs-MD, TSMD の評価 (b)Trp-Cage に関する評価

本研究で示されたように、木構造による探索を分子動力学計算と組み合わせることには、大きなメリットがある。このような研究の方向性は、今回示したような小さなタンパク質だけでなく、より大きなタンパク質にも適用可能であり、今後富岳における研究においても、広く用いられることが期待される。

④標的分子ネットワーク（再委託先：国立大学法人東京工業大学生命理工学院 北尾彰朗）

令和元年度は、前年度から継続して行ってきた分子動力学計算とエネルギー表示の理論を組み合わせたタンパク質-タンパク質複合体モデル構造評価法 **evERdock** の効率化を完了させた。前年度までに開発してきた **evERdock** 法では、1つの複合体モデル（デコイ）について MD 計算による構造の精密化および複数回の計算より精度が向上できる。しかしながら、この取り扱いは計算コストの上昇を伴うため、津田グループと共同で機械学習を用いて計算コストの削減を行い **evERdock** 法によるモデル評価の効率化を試みた。具体的にはバンディット問題の一つとし知られるベストアームアイデンティフィケーション (BAI) に対するアルゴリズムを導入した (**evERdock-BAI** 法)。図 1 に機械学習を用いずに各デコイ同じ回数評価する場合 (**uniform sampling**) と **evERdock-BAI** 法を用いる場合のフォローチャートを示した。限られた計算資源で複合体モデルの評価を行う状況を想定し、**evERdock** による評価数を N 、評価したいデコイ数を M とすると、機械学習を用いない場合は各デコイを等しく N/M 回評価し、平均の結合自由エネルギーより評価する。一方、**evERdock-BAI** 法では **evERdock** による評価ごとにスコアを更新して、最もスコアの高いデコイの評価を行う。図 1(B)には BAI アルゴリズムとして UCBE を採用した例を示した。図に示したとおり UCBE のスコアは2つの項からなるが、第1項はこれまでの結合自由エネルギーが低いとスコアが高くなり、第2項これまでの探索数が多いとスコアが低くなるように設定されている。すなわち第1項により、結合自由エネルギーが高く正解である確率の低いデコイに対する計算コストを削減し、より尤もらしい構造を更に調べるために計算コストを利用することができる（知識利用）。一方、第2項により評価数が十分に多いデコイのスコアは小さくなるため、様々なデコイを調べる探索に計算コストを利用することが出来る。このようにして知識利用と探索をバランスよく実行する手法である。

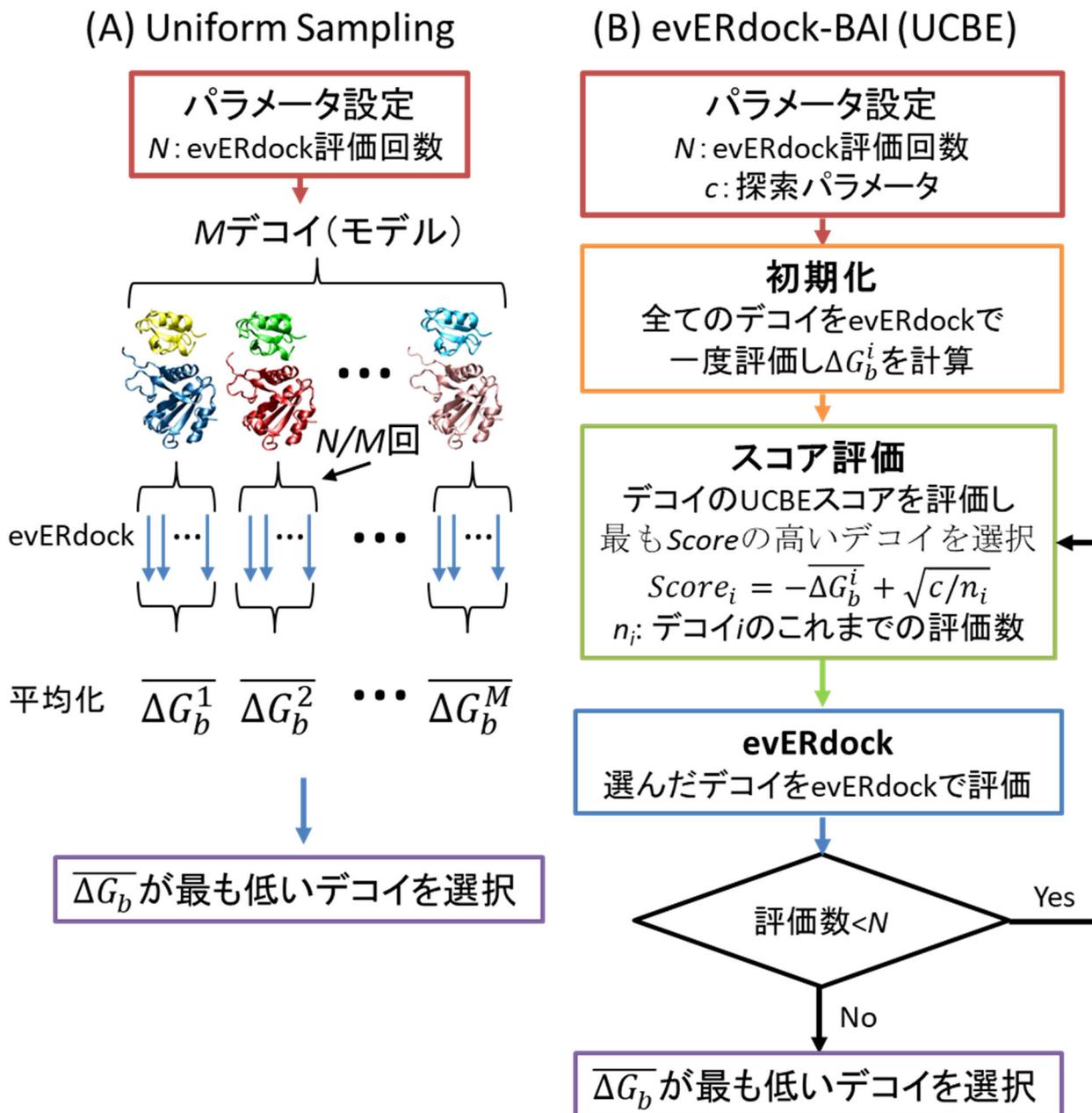
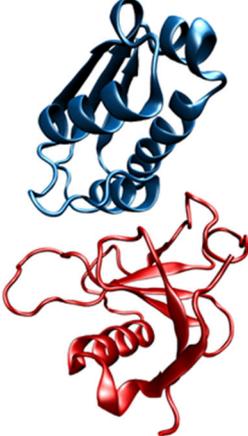
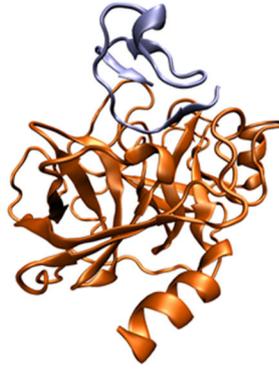
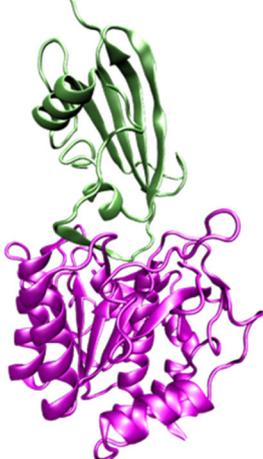


図1 evERdockによる評価数 N 、デコイ数 M とした時の(A)Uniform Sampling と(B)evERdock-BAI アルゴリズムを用いた場合のフローチャート、BAI アルゴリズムに UCBE を用いた例を示した。

evERdock-BAI 法を図2に示した3種類の複合体に適用した。それぞれ対象となるデコイ数やデコイ中の正解構造数などが異なる状況を想定している。図2の表に示したように95%の正解率を得るために必要なevERdockによる評価数を比較するとUniform Samplingと比較していずれの複合体でもモデル構造当たりにおこなうMDの実行回数を軽減することに成功し、特に3SICに関しては336回から83回と4.05倍の高速化を達成した。この研究で得られた成果を平成元年度にJ. Chem. Phys.誌に発表した(K Terayama, A Shinobu, K Tsuda, K Takemura, A Kitao, J. Chem. Phys. 151 (21), 215104, 2019)。

複合体 (PDB ID)	1AY7	1PPE	3SIC
構造	Barstar  RNase SA	Trypsin inhibitor  Trypsin	Subtilisin inhibitor  Subtilisin BPN'
複合体	1AY7	1PPE	3SIC
デコイ数	11	11	21
正解構造数	2	4	1
$N_{\text{evERdock-runs}}$ (Uniform Sampling)	33	33	336
$N_{\text{evERdock-runs}}$ (BAI/UCBE)	15 (2.20)	27 (1.22)	83 (4.05)

$N_{\text{evERdock-runs}}$: 95%の正解率を得るために必要なevERdock計算の数

図2 evERdock-BAI法を適用した複合体構造、デコイ数、正解構造数および95%の正解率を得るために必要なevERdockの評価数(最終行の括弧内はBAIアルゴリズムを用いることによる計算コストの削減割合)

これまでに高精度でタンパク質-タンパク質複合体モデル評価を行うevERdock法を開発し、上述したようにBAIアルゴリズムを用いて効率化を行うevERdock-BAI法を開発してきた。既に報告してきたようにevERdock法は短時間のMD計算で評価を行うため、MD計算を用いる高精度の手法としては多くのモデル評価が可能な方法であり、数百モデルに適用した例も報告している。BAIアルゴリズムを用いた効率化も考慮し、十分な計算資源があれば千程度の複合体モデルを高精度で評価することも十分可能である。しかしながら、ZDOCK等の剛体ドッキングを用いて初期構造モデル群を生成する場合、数万以上の候補構造が生成されるため、evERdock-BAI法で取り扱う構造数との間にはまだギャップがある。そこで、図3に示したように、剛体ドッキングで生成した複合体モデルをevERdock-BAI法で評価する流れを考え、evERdockでモデル評価の対象とする複合体モデル構造を粗視化レベルで絞り込んでおくためのクラスタリング法の高精度化を

行うことにより、高い効率と精度を伴うタンパク質-タンパク質複合体構造予測手法の流れが確立出来ると考えた。

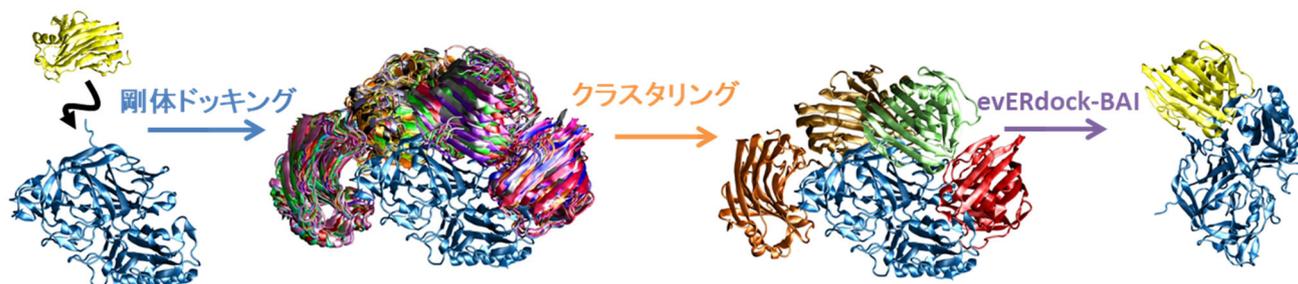


図3 剛体ドッキング、クラスタリング、evERdock-BAI 法用いる流れ

クラスタリングの手法を検討する上で着目したのは、一般的なドッキングを行った際のドッキングスコア上位構造間の類似性である。図4に例を示したように、一般的な剛体ドッキングである ZDOCK を用いて複合体構造予測を行うと、スコア上位構造は比較的近い構造を取っている。仮にドッキングのスコア関数が正確な場合は、上位構造に含まれる正解構造数を増やす目的には適している。しかしながら、これまで示してきたようにドッキングの経験的なスコア関数には問題がある場合が多いため、上位構造群だけに注目していると正解に近い構造を見逃す可能性が高くなる。

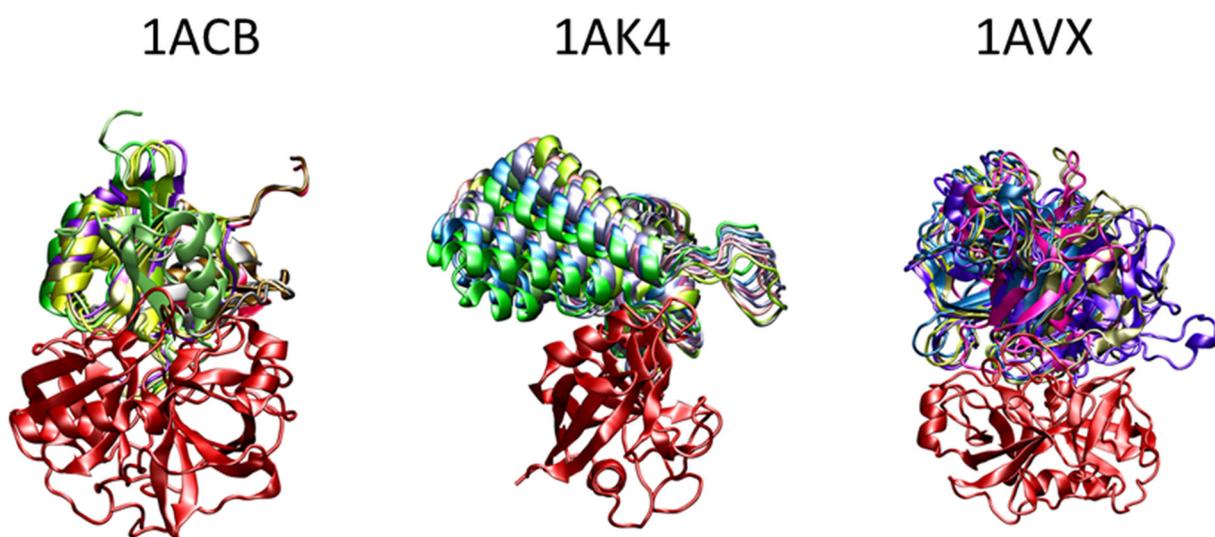


図4 剛体ドッキング ZDOCK3 で生成したドッキングスコア上位 10 構造

そこで図5に示したように、ドッキングのスコア関数を基準としつつ、類似した構造は一つのクラスターとして取り扱うことにより、正解構造を含む候補構造数の削減を試みた。タンパク質-タンパク質複合体構造予測に用いられるベンチマークから 185 種類のタンパク質-タンパク質複合体を対象として解析した。はじめに ZDOCK3 を用いて剛体ドッキングを行い 54,000 構造生

成し、ドッキングスコアが最も高い構造をクラスター中心として抽出する。その後、複合体を構成する大きいタンパク質を重ね合わせた後の小さなタンパク質の RMSD で計算されるリガンド RMSD (L-RMSD)、界面残基の L-RMSD (iL-RMSD)、複合体界面における残基レベルでのコンタクト類似率 (FCC) を用いてクラスター中心と類似した構造を探索する。クラスター中心と類似構造を 1 つのクラスターとして取り除く。この操作をデコイが無くなるまで繰り返す。クラスターの評価を行う際は、クラスター中心のみを代表構造として評価するため、候補構造がクラスター数に減少する。類似性を判定するための RMSD のカットオフ半径 (R_c) やカットオフ類似率 (f_c) などの最適なクラスタリングパラメータを決定し、evERdock の対象とすべきモデル構造数の削減を試みた。

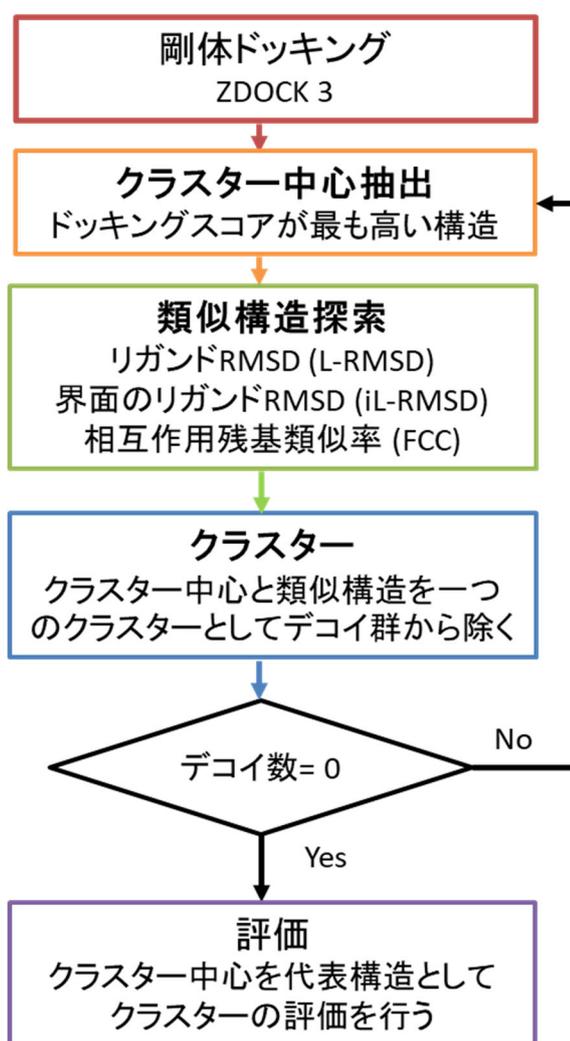


図 5 ドッキングスコアを基準としたクラスタリング解析の流れ

図 6 に示したように、90 %の確率で正解に近い構造を含む候補構造群を生成するために必要な構造数($N_{90\%}$)は、パラメータに大きく依存し、L-RMSD, iL-RMSD, FCC それぞれについてカットオフ 9 Å, 8 Å, 0.3 を最適なパラメータとして選択した。これらの最適なパラメータを元にクラスタリングを行うと $N_{90\%}$ はクラスタリング実行前の 4,808 から最少 320 まで減少し、90 %の evERdock

の対象とすべきモデル構造数を最大 93 % 軽減することができた。また上位千構造が正解に近い構造を含んでいるかどうかを検討すると、大きな構造変化を伴わない 121 種類の複合体のうち最大 117 種類の複合体で正解構造を含んでおり、報告されているいずれのドッキング手法よりも高い 97 % の成功率を達成した。

この研究で得られた成果は平成元年度に *Biophys. Physicobiol.* 誌に発表した (K Takemura, A Kitao *Biophys. Physicobiol.* 16, 295-303, 2019)。

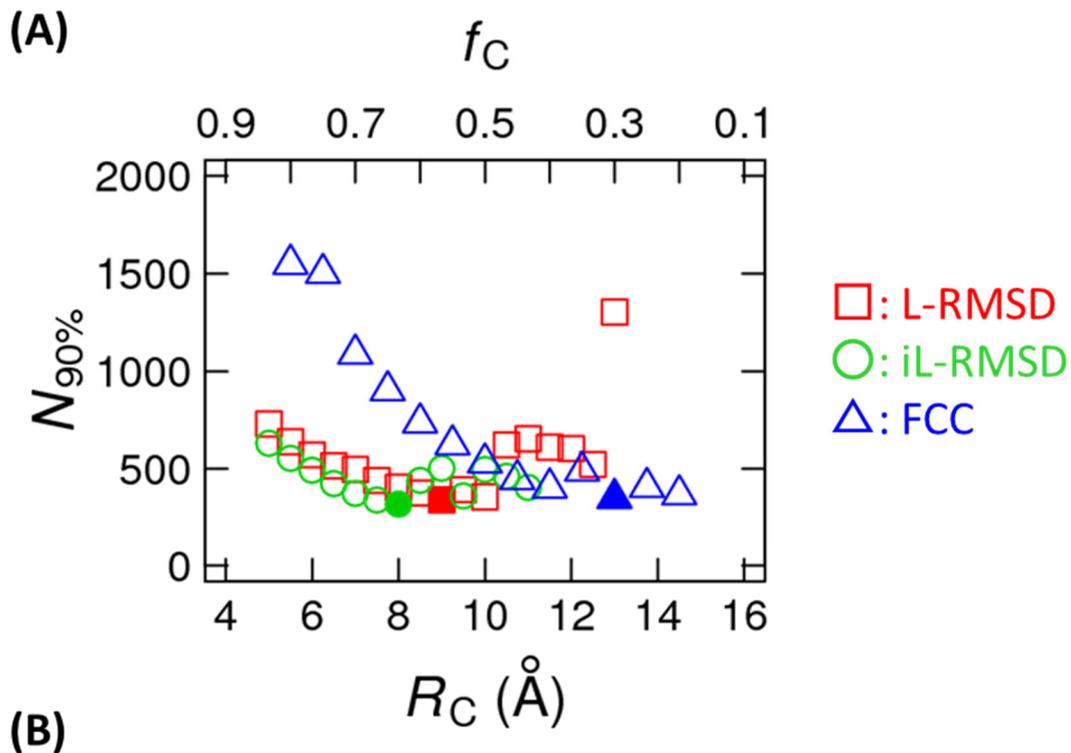


図 6 (A)90%の正解率を得るために必要な構造数($N_{90\%}$)のパラメータ依存性 (B) 最適なパラメータを用いた際の $N_{90\%}$ および上位千構造の成功数 (121 種類の複合体での成功率)

⑤生体系マルチスケールモデリング (再委託先：国立大学法人東京大学大学院情報学環 寺田透)

本研究では、心臓シミュレータ UT-Heart を用いた、遺伝子変異や薬剤による不整脈誘発リスクの予測システムの構築を目指して、分子シミュレーションに基づく、電位依存性イオンチャネルの動作モデル構築法の開発と、薬剤と心筋イオンチャネルの相互作用予測法の開発を行った。

まず、電位依存性イオンチャネルの動作モデル構築法の開発の成果を述べる。昨年度は、心筋イオンチャネル hERG の電位センサードメインについて、分極状態 ($V_m = -0.1$ V) と脱分極状態 ($V_m = 0.0$ V) の 2 つの条件で、立体構造変化の推定経路に沿った 101 個のモデルをそれぞれ構築した。令和元年度はまず、これらのモデルを初期構造として、それぞれ 200 ns の MD シミュレーションを行った。図 1 にこの MD シミュレーションによって生成された構造アンサンブルの分布を示す。ここから十分に広い立体構造空間を探索できていることがわかる。

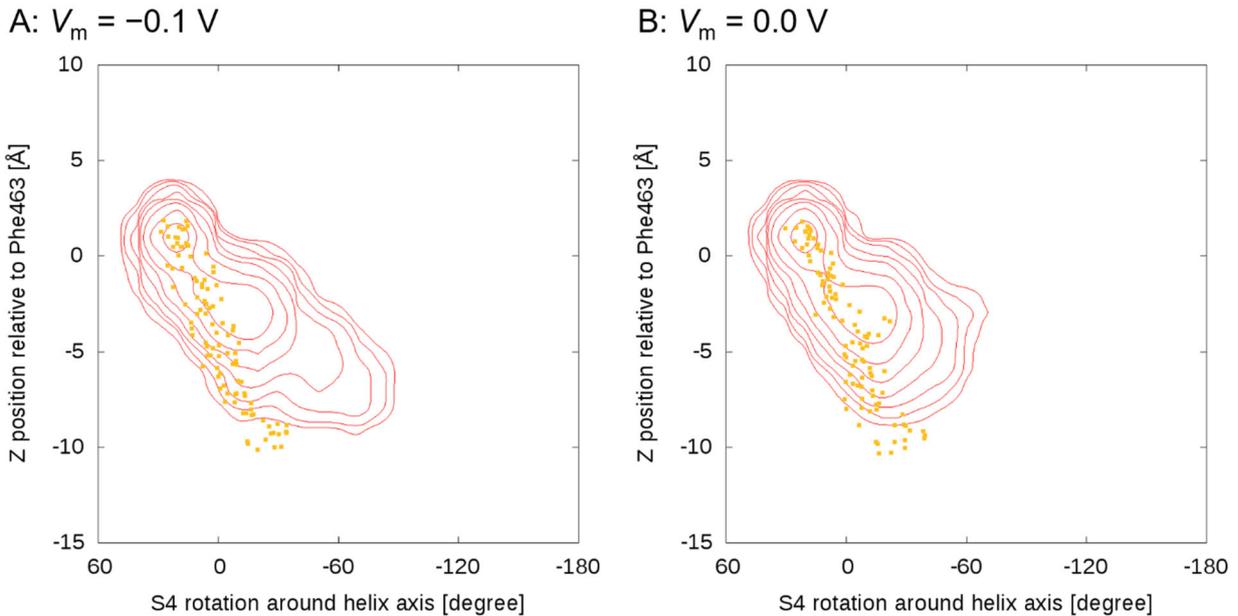


図 1 立体構造変化経路に沿った 101 個のモデルを初期構造とした 200 ns の MD シミュレーションによって生成された構造アンサンブルの分布を、S4 ヘリックスの回転を横軸、Phe463 の Ca の Z 座標に対する Arg531 の Ca の Z 座標 (Z_{R3}) を縦軸としたときの確率密度分布を等高線で示した。橙色の点は初期構造を示す。(A) 分極状態 ($V_m = -0.1$ V)。(B) 脱分極状態 ($V_m = 0.0$ V)。

次いで、この構造アンサンブルから、Phe463 の Ca の Z 座標に対する Arg531 の Ca の Z 座標 (以降 Z_{R3} と呼ぶ) に応じて 160 個の構造を抽出した。ここでは、 Z_{R3} の -11 Å から 5 Å の範囲を 1 Å ずつの区間に分け、それぞれの区間に入る構造を 1 つずつランダムに選択するセットを 10 セット作成した。それぞれのセットを初期構造とする、 Z_{R3} を反応座標とした 16 レプリカからなるレプリカ交換アンブレラサンプリングを、分極状態について 80 ns、脱分極状態について 60 ns、それぞれ独立に 10 回実施した。アンブレラポテンシャルには各区間の中心を平衡位置とする調和ポテンシャルを用いた。図 2A にレプリカ交換アンブレラサンプリングの結果得られた自由エネルギープロファイルを示す。ここでは 10 回のランを平均した結果を示している。図 2B の青色の点は分極状態 ($V_m = -0.1$ V) の自由エネルギープロファイルから、脱分極状態 ($V_m = 0.0$ V) の自由エネルギープロファイルを差し引いた値を示している。このデータは、以下の関数で近似することができる。

$$f(z) = \begin{cases} 0.375(z + 4) & z < -4 \\ 0 & -4 \leq z < 2 \\ 0.375(z - 2) & z \geq 2 \end{cases}$$

この関数の値を図 2B に橙色の実線で示した。

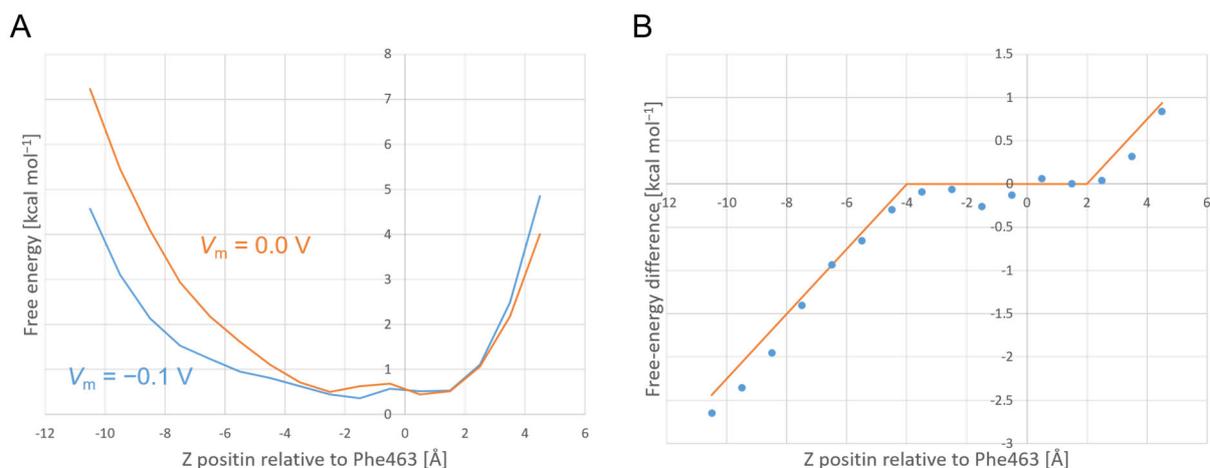


図 2 (A) レプリカ交換アンブレラサンプリングの結果得られた、 Z_{R3} を反応座標とする自由エネルギープロファイル。10 回のランの平均を示した。分極状態 ($V_m = -0.1$ V) の自由エネルギープロファイルを青色の実線で、脱分極状態 ($V_m = 0.0$ V) の自由エネルギープロファイルを橙色の実線で示した。(B) 分極状態の自由エネルギープロファイルと脱分極状態の自由エネルギープロファイルの差を青色の点で示した。これを近似した関数を橙色の実線で示した。

次に、分極状態におけるレプリカ交換アンブレラサンプリングの 40 ns 時点のスナップショットを抽出し、これを初期構造として、分極状態において 1 μ s の MD シミュレーションを行った。初期構造は 160 個 (16 レプリカ \times 10 ラン) あるので、合計 160 μ s に相当する。このトラジェクトリを解析・統合し、まず、分極状態におけるマルコフ状態遷移モデルを構築した。ここではまず、ラグタイムを 5 ns から 200 ns まで変え、遷移行列から計算される緩和時間スケールのラグタイム依存性を検討し (図 3)、マルコフ性が成り立つ、すなわち緩和時間スケールが一定になるラグタイムを決定した。ここではラグタイムを 200 ns とした。

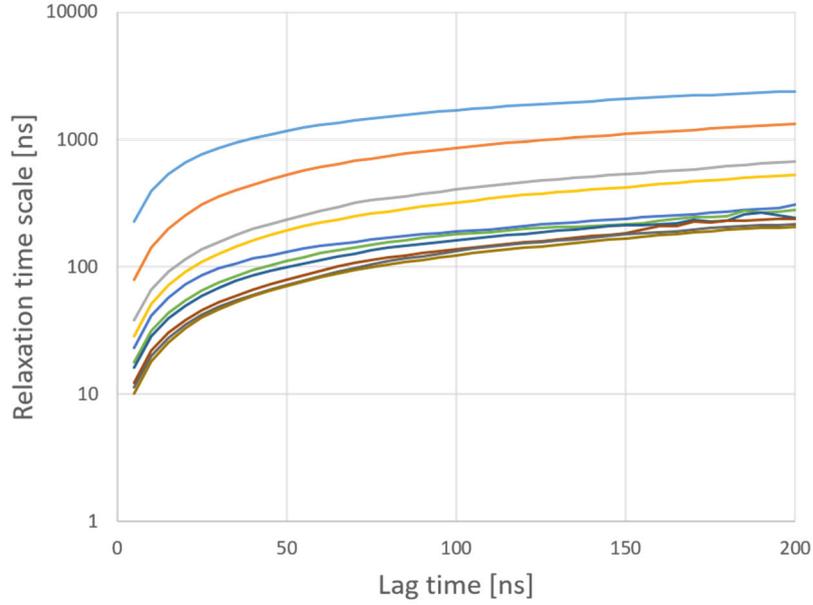


図3 マルコフ状態モデルの遷移行列から計算された、緩和時間スケールのラグタイム依存性。

次に、この分極状態におけるマルコフ状態遷移モデルに、レプリカ交換アンブレラサンプリングによって計算された、自由エネルギー差の Z_{R3} 依存性 ($f(z)$) を取り込み、電位依存的なマルコフ状態遷移モデルを構築した。ここでは、以下の通り、分極状態における遷移行列 $T_{ij}(V_0)$ から、電位依存的な遷移行列 $T_{ij}(V_m)$ を計算した。

$$T_{ij}(V_m) = \alpha_{ij}(V_m) T_{ij}(V_0) \quad (i \neq j)$$

$$\alpha_{ij}(V_m) = \min \left[1, \exp \frac{(f(Z_j) - f(Z_i)) (V_m - V_0)}{RT V_0} \right]$$

$$T_{ii}(V_m) = 1 - \sum_j T_{ij}(V_m)$$

ここで、 V_m は膜電位、 $V_0 = -0.1$ V、 Z_i は状態 i における Z_{R3} の座標の平均値、 R はガス定数、 T は温度 (300 K) である。この電位依存的な遷移行列を用いて計算された、分極状態と脱分極状態の自由エネルギー地形を図4に示す。ここから Z_{R3} の小さい領域 ($Z_{R3} < -5$) で、分極状態の方が、自由エネルギーの低い領域が広いことがわかる。一方で、 Z_{R3} の大きい領域 ($Z_{R3} > -5$) で自由エネルギー地形があまり変わらないように見える。これは図2Bに示す通り Z_{R3} が -4 \AA と 2 \AA の間で $f(z) = 0$ となっているためである。これは、この領域で見かけの誘電率が高くなっており、膜電位に伴う電場を打ち消していることを示唆している。この原因やモデルの妥当性については、今後さらに検討を行う。

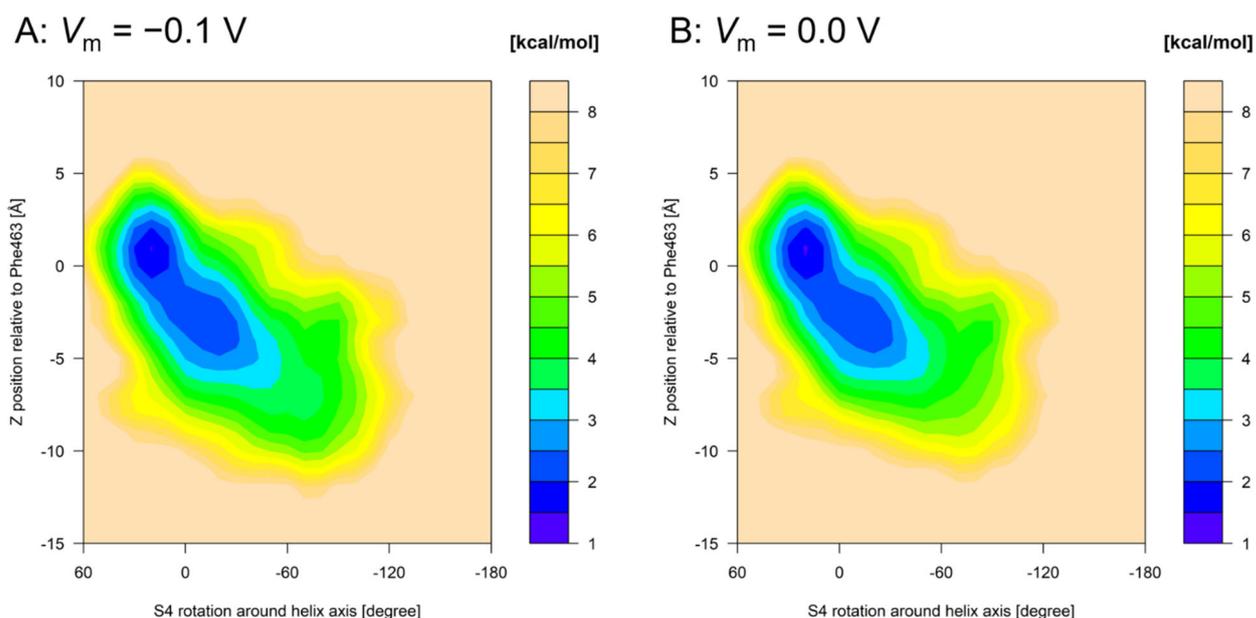


図4 電位依存的な遷移行列から計算された自由エネルギー地形。横軸は、S4ヘリックスの回転、縦軸はPhe463のCaのZ座標に対するArg531のCaのZ座標 (Z_{R3}) である。

最後に、電位依存的な遷移行列から計算された mean first passage time (MFPT) の膜電位依存性を図5に示す。ここでは $Z_{R3} > 0 \text{ \AA}$ をUp状態、 $Z_{R3} < -5 \text{ \AA}$ をDown状態と定義した。ここから、膜電位が正で大きくなるにつれて、Down状態への遷移が遅くなるのがわかる。膜電位が 0.1 V における、Down状態への遷移のMFPTは 118 \mu s で、単独のMDシミュレーションでは追跡できない、非常に遅い運動が再現されているといえる。一方、Up状態への遷移は、膜電位が正で大きくなるにつれて速くなっている。また、全体的にUp状態への遷移の方が、Down状態への遷移よりも速いことから、Up状態の方が安定であることがわかる。レプリカ交換アンブレラサンプリング(図2)と自由エネルギー地形(図4)からも同じ結果が得られている。分極状態では、本来Down状態の方が安定であるはずであるが、シミュレーションの結果では、分極状態ではDown状態がやや安定化するものの、依然としてUp状態の方が安定である。本研究において平成29年度までに実施した、Shakeファミリーのカリウムチャネルの電位センサードメインを用いたシミュレーションでは、分極状態ではDown状態が安定、脱分極状態ではUp状態が安定という、実験事実と一致した結果が得られていた。Shakeファミリーのカリウムチャネルは、電位センサードメインに隣接するポアドメインが、別のポリペプチド鎖由来というドメインスワップ構造をとっているが、hERGでは、同じポリペプチド鎖由来のドメインが隣接している。このタンパク質アーキテクチャの違いが、電位センサードメインを取り出してシミュレーションを行った時の、電位依存的な振舞いに影響を与えている可能性が考えられる。今後「富岳」を用いることで、ポアドメインを含むチャネル全体のシミュレーションが可能となり、より精度の高いマルコフ状態遷移モデルを構築することが可能となると期待される。

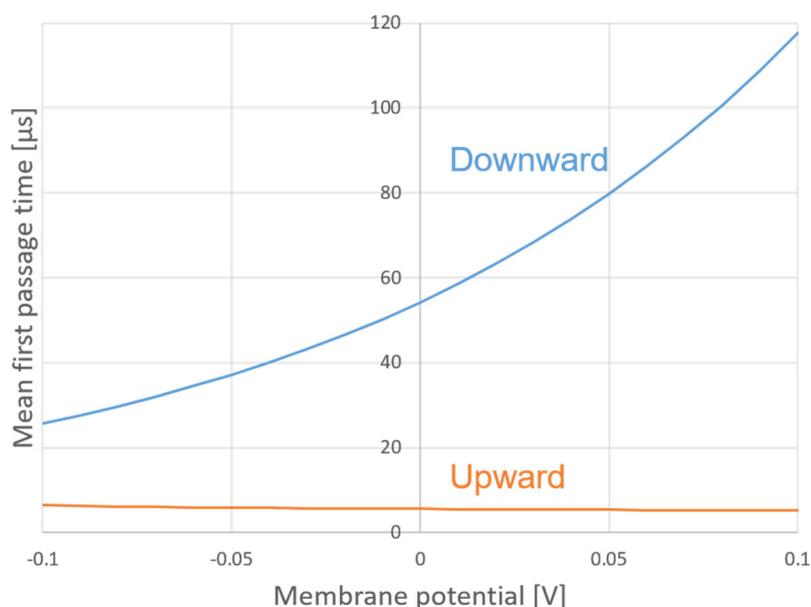


図5 Mean first passage time (MFPT) の膜電位依存性。Up 状態から Down 状態への遷移 (Downward) の MFPT を青色の実線で、Down 状態から Up 状態への遷移 (Upward) の MFPT を橙色の実線で示した。

次に、薬剤と心筋イオンチャネルの相互作用予測法の開発の成果を述べる。令和元年度はまず、前年度までに開発した、心筋カリウムチャネル hERG に対する薬剤の親和性予測法について、精度の検証と、他の予測法との精度の比較を行った。hERG に対する薬剤の親和性予測法では、hERG の電子顕微鏡構造 (PDB ID: 5VA2) のポア領域に対して、薬剤をドッキングし、MP-CAFEE 法を用いて結合自由エネルギーを計算する。昨年度までに多様な構造を持つ 12 種類の薬剤について、結合自由エネルギーを求めた。この計算値を結合自由エネルギーの実験値に対してプロットすると、非常によく相関しており、決定係数 R^2 は 0.909 となった (図 1A)。令和元年度は、これまでに計算を行った 12 種類の薬剤 (=トレーニングセット) とは異なる 4 種類の薬剤 (=テストセット) について、同様に結合自由エネルギーを計算し、実験値と比較することで、精度の検証を行った。図 6A に令和元年度に計算した 4 種類の薬剤のデータを橙色の点で示す。ここで予測値 ($\Delta G_{\text{bind}}(\text{pred})$) は、回帰直線 $y = ax + b$ の係数 $a = 2.48$ 、 $b = 7.05$ を用いて、関係式 $\Delta G_{\text{bind}}(\text{pred}) = (\Delta G_{\text{bind}}(\text{calc}) - b)/a$ から求めた。この値と実験値 ($\Delta G_{\text{bind}}(\text{exp})$) との差が誤差となる。4 種類の薬剤の根二乗平均誤差 (RMSE) は 0.73 kcal mol⁻¹ となった。これは解離定数に換算すると約 3 倍となる。実験では解離定数が 10 倍程度ばらつくとされているが、本予測手法では、95% のケースで誤差が実験における誤差の範囲内に収まるといえ、十分な精度を持つと言える。

さらに、既存に手法との精度の比較を行った。ここでは、Schrodinger Suite の QikProp を用いて、hERG への親和性の予測値である QPlogHERG スコアを計算した。上記と同じ 12 種類の薬剤 (トレーニングセット) のうちスコアの得られた 11 種類の薬剤について、QPlogHERG スコアを求め、結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対してプロットした (図 6B)。ここから回帰直線を求めたところ、決定係数 R^2 は 0.765 となった。これは我々の手法におけるトレーニングセットの値 ($R^2 =$

0.909) より低い値となった。次いで、4種類の薬剤（テストセット）について QPlogHERG スコアを求め、回帰直線を用いて、結合自由エネルギーの予測値 ($\Delta G_{\text{bind}}(\text{pred})$) を求めたところ、根二乗平均誤差 (RMSE) は $1.40 \text{ kcal mol}^{-1}$ となった (図 6B 参照)。こちらも、我々の手法におけるテストセットの値 (RMSE = $0.73 \text{ kcal mol}^{-1}$) より大きい値となった。以上の結果から、我々の手法の方が既存の手法 (QPlogHERG) より精度が高いことが確かめられた。本研究の成果は論文にまとめ、Scientific Reports 誌にて公表した[Negami *et al. Sci. Rep.* **9**, 16586 (2019)]。

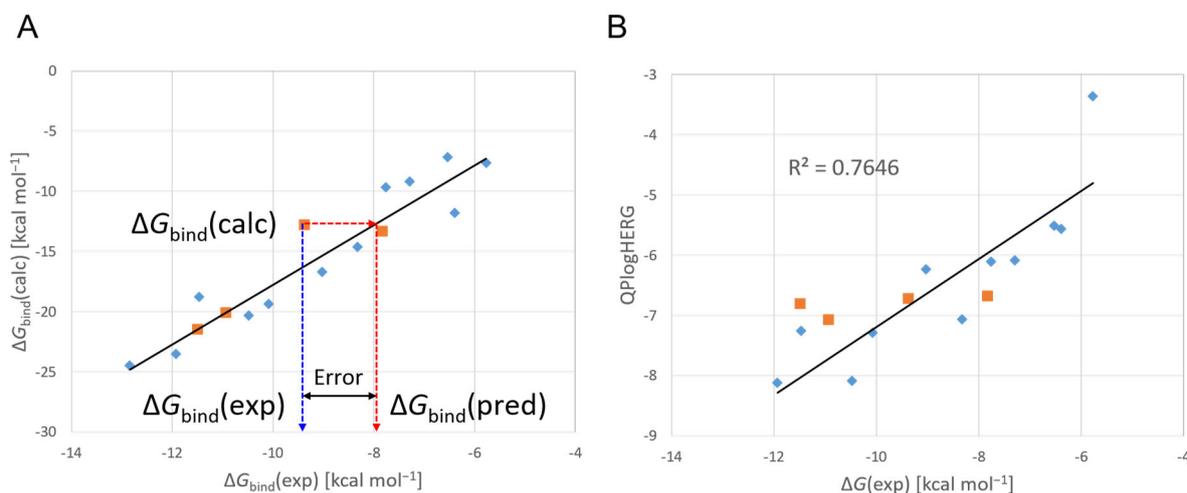


図 6 hERG に対する薬剤の結合自由エネルギー計算。(A) 結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対する、計算値 ($\Delta G_{\text{bind}}(\text{calc})$) のプロット。昨年度までに計算を行い、回帰直線を求めるのに用いた 12 種類の薬剤 (トレーニングセット) のデータ点を青色で示す。令和元年度に計算を行い、精度の検証に用いた 4 種類の薬剤 (テストセット) のデータ点を橙色で示す。結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対する、計算値 ($\Delta G_{\text{bind}}(\text{calc})$) のプロット。(B) 結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対する、QPlogHERG スコアのプロット。回帰直線を求めるのに用いた 11 種類の薬剤 (トレーニングセット; 計算値の得られなかった 1 種類の薬剤を除く) のデータ点を青色で示す。精度の検証に用いた 4 種類の薬剤 (テストセット) のデータ点を橙色で示す。

次に心筋ナトリウムチャンネル Nav1.5 に対する薬剤の親和性予測法開発の成果について述べる。ヒトの Nav1.5 の立体構造は実験的に決定されていないため、昨年度は、類似のアミノ酸配列を持つ、デンキウナギのナトリウムチャンネル (eeNav1.4) の電子顕微鏡構造 (PDB ID: 5XSY) と、ヒト骨格筋のナトリウムチャンネル hNav1.4 の電子顕微鏡構造 (PDB ID: 6AGF) をそれぞれ鋳型としてモデリングを行った。予備的なドッキングシミュレーションの結果、結合自由エネルギーの実験値とドッキングスコアの相関が、hNav1.4 の電子顕微鏡構造を鋳型としたモデルの方が優れていることが明らかとなった。したがって、令和元年度は、この hNav1.4 の電子顕微鏡構造を鋳型としたモデルの改良に取り組んだ。ここでは、昨年度構築したモデルをモデル 1 とし、これを、MD シミュレーションを用いて精密化して得られたモデルをモデル 2 と呼ぶ。このモデルは、モデル 1 を初期構造として 500 ns の MD シミュレーションを行い、得られた構造アンサンブルを立体構造類似性に基づいてクラスターに分類した結果得られた、最大クラスターの代表構造である。図 7 に 2 つのモデルの重ね合わせ

を示す。それぞれのモデルに対し、同じ 10 種類の薬剤をドッキングし、各薬剤においてスコアの最も良いポーズについて、MP-CAFEE 法を用いて結合自由エネルギー計算を行った (図 8)。ここで、結合自由エネルギー実験値は、IC50 値から、関係式 $\Delta G_{\text{bind}}(\text{exp}) = RT \ln(\text{IC50})$ を用いて推定している。モデル 1 を用いた時、実験値 ($\Delta G_{\text{bind}}(\text{exp})$) と計算値 ($\Delta G_{\text{bind}}(\text{calc})$) の間の相関の決定係数 R^2 は 0.356 であった。一方、モデル 2 を用いた時、決定係数 R^2 は 0.729 となった。したがって、ホモロジーモデルを MD シミュレーションにより精密化することで、実験値と計算値の間の相関が改善したと言える。今後は、モデル 2 をベースに改良を加え、さらなる精度の向上を目指す。

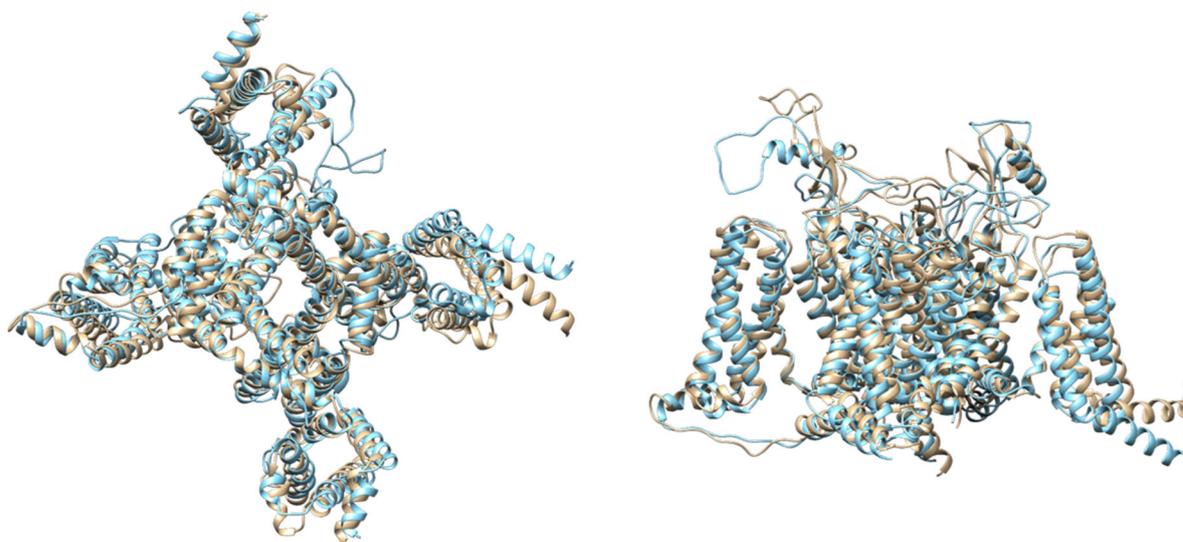
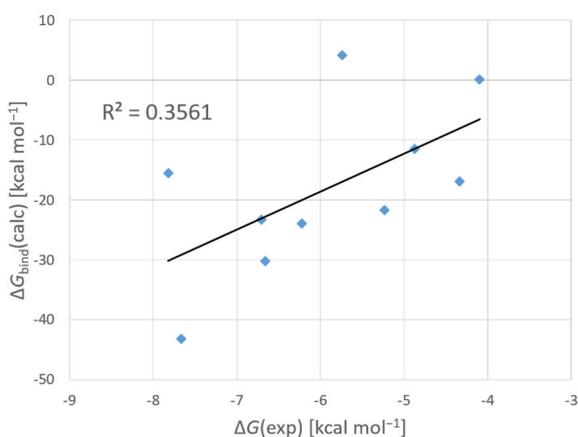


図 7 Nav1.5 のモデル、モデル 1 (茶色) とモデル 2 (水色) の重ね合わせ。主鎖構造をリボンモデルで示した。(左) 細胞外から見た図。(右) 横から見た図。

A: モデル1



B: モデル2

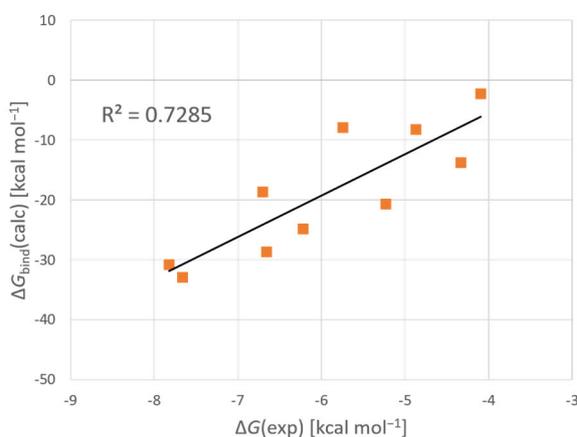


図 8 (A)Nav1.5 のモデル 1 に対する 10 種類の薬剤の結合自由エネルギー計算。(B)Nav1.5 のモデル 2 に対する 10 種類の薬剤の結合自由エネルギー計算。結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対して、計算値 ($\Delta G_{\text{bind}}(\text{calc})$) をプロットした。回帰直線を決定係数 R^2 とともに示す。

最後に心筋カルシウムチャネル Cav1.2 に対する薬剤の親和性予測法開発の成果について述べる。Cav1.2 も立体構造が実験的に決定されていないため、令和元年度は、まず、Cav1.2 の立体構造モデリングを行った。ここでは、ウサギの骨格筋由来の rCav1.1 の電子顕微鏡構造 (PDB ID: 6JPA) を鋳型としてモデリングを実施した。rCav1.1 のアミノ酸配列と心筋カルシウムチャネル Cav1.2 のアミノ酸配列に対するカバー率は 70%、配列一致度も 70%であった。構築したモデルを図 9 に示す。



図 9 Cav1.2 のモデル。主鎖構造をリボンモデルで示した。(左) 細胞外から見た図。(右) 横から見た図。

このモデルに対して、10 種類の薬剤のドッキングシミュレーションを行い、各薬剤においてスコアの最も良いポーズについて、MP-CAFEE 法を用いて結合自由エネルギー計算を行った。ここでは、ポーズ探索の中心や、ポア内の Ca^{2+} の有無など、複数のドッキング条件の検討を行い、実験値 ($\Delta G_{\text{bind}}(\text{exp})$) と計算値 ($\Delta G_{\text{bind}}(\text{calc})$) の間の相関が最も良くなる条件を探索した。その結果得られた条件における計算結果を図 10 に示す。実験値と計算値の間の相関の決定係数 R^2 は 0.695 となり、比較的高い相関が得られた。今後は、このモデルに改良を加え、さらなる精度の向上を目指す。

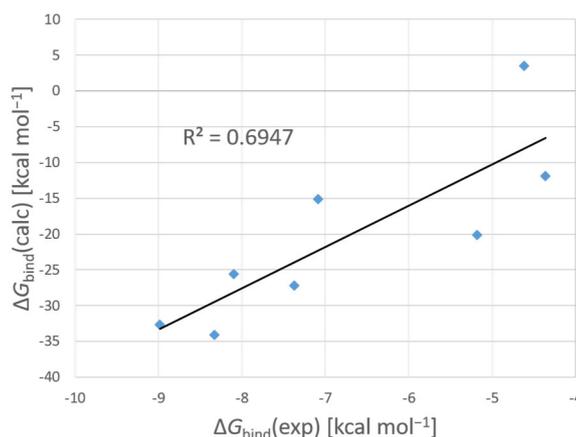


図 10 Cav1.2 のモデルに対する 10 種類の薬剤の結合自由エネルギー計算。結合自由エネルギーの実験値 ($\Delta G_{\text{bind}}(\text{exp})$) に対して、計算値 ($\Delta G_{\text{bind}}(\text{calc})$) をプロットした。回帰直線を決定係数 R^2 とともに示す。

(4) プロジェクトの総合的推進

プロジェクト全体の連携を密としつつ円滑に運営していくため、執行部会議を定期的実施し、プロジェクト運営全般の検討・決裁・指揮を行った。重点課題1 ワークショップ（全体会議）を1月に開催し、プロジェクト全体の進捗状況を確認、最終成果の共有を行った。プロジェクトで得られた成果については、シンポジウムやウェブサイト、新聞メディアなどの記事掲載、研究紹介コンテンツなどにより積極的に公表した。

特に、令和元年度では、ポスト「京」プロジェクトで同じく健康長寿社会の実現をターゲットとする重点課題2との連携による一般公開型のシンポジウムを8月に開催し、プロジェクトで得られた成果についての討論や意見交換、一般向け成果報告パンフレットの配布により、さらに知見を広めるとともに本プロジェクトのプレゼンスを高めることができた。

また、日本生物物理学会などバイオ計算に関心の高い研究者を数多く集める学会との共催により本課題に関連した発表枠を設け、効果的な討論、意見交換を行ったほか、若年層が研究者として将来的な参画を図ることを目的に、高校生対象のイベントや出張講義などのアウトリーチ活動を通して今後の展開に役立てた。加えて、産業連携コンソーシアムとの連携ワークショップや報告会で討論や意見交換を行い、プロジェクトの成果創出を担う産業界、医療機関との連携、協力を図った。

課題推進にあたっては、第三者の視点で進捗状況の把握・評価・改善提言・指導等を行う諮問委員会を8月に開催し、これまでの成果を報告するとともに、次期の発展的プロジェクトに向けた具体的な助言を得た。

4-3. 活動（研究会等）

(1) 重点課題1 ワークショップ

- ・第7回 神戸理化学研究所生命機能科学研究センター 2月10日

(2) ポスト「京」重点課題1×重点課題2 連携シンポジウム

「個別化医療・創薬基盤に衝撃 —スパコン「富岳」が起こす地殻変動—」

8月8日 秋葉原 UDX ギャラリーネクスト NEXT-1

- ・一般公開型 ポスターセッション、口頭発表

(3) 人材育成

- ・HPCI 人材育成タスクフォース

4月10日 文科省へ説明

10月2日 ポータルサイト運用、今後のタスクフォース方針

3月24日 今後の体制、引継ぎ

- ・神戸大学計算科学教育センター「計算生命科学の基礎VI」10月2日～1月29日、全15回
- ・2019年度情報計算科学セミナー 5月30日～3月26日、全11回日、3月26日は中止)

(4) アウトリーチ活動

- ・ASEAN HPC ワークショップ (計算科学振興財団) 4月15日

- ・スーパーサイエンスハイスクール山梨県立韮崎高校 関西科学研修特別講義 7月24日

- ・第6回「京」を中核とする HPCI システム利用研究課題 成果報告会 11月1日

- ・石川県高等学校文化連盟理科部秋期行事「高校生のための実験・実習セミナー」11月8日

- ・第12回サイエンスフェア in 兵庫 1月26日

- ・HPC フォーラム スーパーコンピュータ「京」の時代の成果から、「富岳」への期待

3月4日開催中止のため、見える化セッション WEB 展示

→https://www.hpci-office.jp/pages/mieruka_2020

(5) 学会・シンポジウム発表

- ・第57回日本生物物理学会年会共催シンポジウム 9月25日

- ・大学 ICT 推進協議会 2019 年度年次大会 12月12日

(6) プレスリリース

- ・「LC-SCRUM-Japan で構築した日本最大臨床ゲノムデータを活用しスーパーコンピュータで治療薬の効き目を予測 ～がんゲノム医療における新たなツールの開発～」奥野恭史 (2019/05/07)

- ・「酵素－阻害剤結合の初期会合体を予測 –初期結合過程を標的とした新たな創薬分子設計の可能性を拓く–」杉田有治 (2019/09/06)

(7) 記事掲載

- ・「G20 見据え AI 創薬で産官学シンポジウム、日本の強みをどう発揮するか」奥野恭史 (日経 xTECH、2019/5 : WEB)

- ・「総論 第4次産業革命における創薬イノベーション」奥野恭史 (月刊細胞、2019/5 : 雑誌)

- ・「AI 創薬へ、各社のデータ共有体制が必要」奥野恭史 (日刊薬業、2019/7/22 : 新聞)

- ・「スパコンの道広げた「京」」奥野恭史 (読売新聞、2019/9/15 : 新聞)

- ・「スパコン「京」が開いた分子シミュレーション研究の扉」奥野恭史 (週刊医学界新聞 3340 号、2019/9/30 : 新聞)

- ・「人工知能のすべて」「創薬 AI 薬として使えそうな物質を AI が提案！」奥野恭史 (Newton、2019/9 号 : 雑誌)

- ・「学習する AI 広がる夢」奥野恭史 (読売新聞、2019/10/4 : 新聞)

- ・「100社で挑む創薬 AI、初の商用化事例を披露するも課題は山積み」奥野恭史（日経 XTECH、2019/10/15：WEB）
- ・「AI創薬 コスト半減」奥野恭史（日本経済新聞、2019/10/7：新聞）
- ・「AI創薬の異業種連携進む」奥野恭史（薬事日報、2019/10/16：新聞）
- ・「AI創薬の異業種連携進む—十分なデータ確保に課題 ライフインテリジェンスコンソーシアム」奥野恭史（薬事日報、2019/10：WEB）
- ・「サイエンス 追う BOX-学習する AI 広がる夢」奥野恭史（読売新聞（近畿版）、2019/10：新聞）
- ・「進化する AI、発症予測から原因究明へ」奥野恭史（Medical Tribune、2019/11：WEB）
- ・「AI創薬で LINC が成果 データとモデルの共有に課題」奥野恭史（化学工業日報、2019/12：新聞）
- ・研究最前線「GENESIS で創薬が変わる」杉田有治（理研ニュース 2020 年 2 月、2020/2/5：広報誌）

(8) イベント

- ・理化学研究所一般公開
和光 4 月 20 日、横浜 9 月 21 日、神戸（R-CCS）11 月 9 日、大阪 11 月 23 日

4-4. 実施体制

業務項目	担当機関	担当責任者
(1) ポスト「京」での MD 高度化とアルゴリズム深化 (サブ課題 A)	国立研究開発法人理化学研究所	生命機能科学研究センター チームリーダー 杉田 有治
①ポスト「京」での MD 高度化	国立研究開発法人理化学研究所	生命機能科学研究センター チームリーダー 杉田 有治
②拡張アンサンブル法	国立研究開発法人理化学研究所	生命機能科学研究センター チームリーダー 杉田 有治
③長時間ダイナミクス法	国立大学法人東京工業大学	生命理工学院 教授 北尾 彰朗
④ハイブリッド QM/MM 自由エネルギー法	国立大学法人京都大学	大学院理学研究科 教授 林 重彦
⑤粗視化モデリング	国立大学法人京都大学 国立大学法人名古屋大学	大学院理学研究科 教授 高田 彰二 大学院工学研究科 准教授 篠田 渉

⑥自由エネルギー計算法	公立大学法人横浜市立大学	生命医科学研究科 教授 池口 満徳
(2) 次世代創薬計算技術の開発 (サブ課題 B)	公立大学法人横浜市立大学	生命医科学研究科 教授 池口 満徳
①動的分子機能制御	公立大学法人横浜市立大学	生命医科学研究科 教授 池口 満徳
②タンパク質間相互作用制御	国立大学法人東京大学	先端科学技術研究センター 特任教授 藤谷 秀章 及び 特任准教授 山下 雄史
③ウイルス標的創薬計算技術	国立大学法人名古屋大学	大学院工学研究科 准教授 篠田 渉
④核酸-タンパク質相互作用制 御	国立研究開発法人量子科学 技術研究開発機構	量子生命科学領域 グループリーダー 河野 秀俊
⑤細胞内環境シミュレーション	国立研究開発法人理化学研 究所	生命機能科学研究センター チームリーダー 杉田 有治
(3) 創薬ビッグデータ統合システ ムの開発 (サブ課題 C)	国立大学法人京都大学	大学院医学研究科 教授 奥野 恭史
①高精度薬剤デザイン	国立大学法人京都大学	大学院医学研究科 教授 奥野 恭史
②機能制御部位データベース	国立研究開発法人産業技術 総合研究所	創薬分子プロファイリング研 究センター 研究チーム長 広川 貴次
③創薬関連ビッグデータ	国立大学法人東京大学	大学院新領域創成科学研究科 教授 津田 宏治
④標的分子ネットワーク	国立大学法人東京工業大学	生命理工学院 教授 北尾 彰朗
⑤生体系マルチスケールモデリ ング	国立大学法人東京大学	大学院情報学環 准教授 寺田 透
(4) プロジェクトの総合的推進	国立研究開発法人理化学研 究所	生命機能科学研究センター 客員主管研究員 奥野 恭史

様式第2-1

別添1

学会等発表実績

重点課題1:「生体分子システムの機能制御による革新的創薬基盤の構築」

代表者氏名:奥野恭史

サブ課題A:ポスト京でのMD高度化とアルゴリズム深化

サブ課題代表者:杉田 有治

1. 学会等における口頭・ポスター発表(国際会議)

No.	発表した成果(発表題目)	発表者氏名	発表した場所(学会名等)	発表した時期	国内外の別
1	Nucleosome dynamics studied by coarse-grained molecular simulations; sliding, disassembly, and transcription factor binding	Shoji Takada(Kyoto University)	ADVANCED INNOVATION LECTURE FOR STRUCTURAL BIOLOGY	2019年5月	国際
2	SIMULATING PROTEIN COMPLEX FORMATION AND	Akio Kitao(Tokyo Institute of Technology)	Telluride Science Research Center Workshop on Protein Dynamics	2019年7月	国際
3	Simulating protein complex formation and dissociation	Akio Kitao(Tokyo Institute of Technology)	The 10th Toyota RIKEN International Workshop on Science of Life Phenomena Woven by Water and Biomolecules	2019年9月	国際
4	Association and dissociation mechanisms of proteins and peptides	Akio Kitao(School of Life Science and Technology, Tokyo Institute of Technology)	WRHI International Workshop on Advanced Laser Spectroscopy for Soft Molecular Systems	2019年11月	国際

学会等における口頭・ポスター発表(国内学会)

No.	発表した成果(発表題目)	発表者氏名	発表した場所(学会名等)	発表した時期	国内外の別
1	低計算量の拡張アンサンブル法の開発:レプリカ交換と Gaussian accelerated Molecular Dynamicsの組合せ	尾嶋拓(理化学研究所生命機能科学研究センター), 李秀栄(理化学研究所生命機能科学研究センター), 杉田有治(理化学研究所生命機能科学研究センター, 理化学研究所計算科学研究センター, 理化学研究所杉田理論分子科学研究室)	第19回日本蛋白質科学会年会	2019年6月	国内
2	2次元レプリカ交換分子動力学(gREST/REUS)シミュレーションによるSrc タンパク質阻害剤の結合自由エネルギー解析	李秀栄(理研BDR),尾嶋拓(理研BDR),杉田有治(理研BDR)	第19回日本蛋白質科学会年会	2019年6月	国内
3	クラスタリングを用いた蛋白質-蛋白質ドッキング複合体候補構造の効率的絞込み	竹村 和浩, 北尾 彰朗(東工大生命理工)	第19回日本蛋白質科学会年会 第71回日本細胞生物学会大会 合同年次大会	2019年6月	国内
4	圧力による蛋白質の水和を利用した細胞機能の制御	畑 宏明, 西原 泰孝, 西山 雅祥, 川岸 郁朗, 北尾 彰朗(東京工業大学 生命理工学院)	第19回日本蛋白質科学会年会	2019年6月	国内
5	計算から予測するタンパク質複合体	北尾 彰朗(東京工業大学生命理工学院)	2019日本生化学会関東支部例会	2019年6月	国内

6	光合成系IIマンガングラスターのプロトン化状態の理論的同一	成せい(Kyoto University), 林重彦(Kyoto University)	新学術領域「革新的光物質変換」第2回合同班会議	2019年7月	国内
7	GENESISを用いた創薬計算基盤技術	李秀栄(理研BDR),尾嶋拓(理研BDR),笠原健人(理研BDR)	ポスト「京」重点課題1×重点課題2シンポジウム	2019年8月	国内
8	タンパク質複合体の構造予測と結合・解離過程の解明	畑宏明, 竹村和浩, Duy Phuoc TRAN, 北尾彰朗(東京工業大学 生命理工学院)	ポスト「京」重点課題1×重点課題2連携シンポジウム	2019年8月	国内
9	Determination of Protonation States of HIV-1 Protease Binding Indinavir and Its Drug-Resistant Mutant by Free Energy Calculations	Masahiko Taguchi(Kyoto University), Ryo Oyama(Kyoto University), Masahiro Kaneko(Kyoto University), Shigehiko Hayashi(Kyoto University)	ポスト「京」重点課題1×重点課題2連携シンポジウム	2019年8月	国内
10	Free-energy analysis of protein-ligand binding pose using generalized ensemble methods	Hiraku Osima(RIKEN Center for Biosystems Dynamics Research), Suyong Re(RIKEN Center for Biosystems Dynamics Research), Yuji Sugita(RIKEN Center for Biosystems Dynamics Research, RIKEN Center for Computational Science, RIKEN Cluster for Pioneering Research)	第57回日本生物物理学会年会	2019年9月	国内
11	拡張アンサンブル法を用いたタンパク質-リガンド結合ポーズおよび親和性の高精度予測	尾嶋拓(理化学研究所生命機能科学研究センター), 李秀栄(理化学研究所生命機能科学研究センター), 杉田有治(理化学研究所生命機能科学研究センター, 理化学研究所計算科学研究センター, 理化学研究所杉田理論分子科学研究室)	日本物理学会 2019年秋季大会	2019年9月	国内
12	Free energy analysis of ATP competitive inhibitor-protein kinase bindings	Suyong Re (RIKEN BDR), Hiraku Oshima (RIKEN BDR), Yuji Sugita (RIKEN BDR)	第57回日本生物物理学会年会	2019年9月	国内
13	An efficient screening, an accurate evaluation, and a simple prediction of protein complex structures	Kazuhiro Takemura, Akio Kitao(Tokyo Institute of Technology)	第57回日本生物物理学会年会	2019年9月	国内
14	Pressure effects on protein-protein interactions studied by molecular dynamics simulations	Hiroaki Hata, Yasutaka Nishihara, Masayoshi Nishiyama, Ikuro Kawagishi, Akio Kitao(Tokyo Institute of Technology)	57th Annual Meeting of Biophysical Society of Japan	2019年9月	国内
15	Determination of protonated states for native and mutant structures of HIV-1 protease with indinavir by free energy calculations	Masahiko Taguchi(Kyoto University), Ryo Oyama(Kyoto University), Masahiro Kaneko(Kyoto University), Shigehiko Hayashi(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
16	GtACR1の分子機構に関する理論的研究	鹿倉啓史(京都大学), 成せい(京都大学), 林重彦(京都大学)	ISSPワークショップ「レチナルタンパク質の光機能発現の物理と化学」	2019年9月	国内
17	Molecular Dynamics of Nucleosome Assembly	Giovanni Brandani(Kyoto University), Shoji Takada(Kyoto University), Cheng Tan(RIKEN)	第57回日本生物物理学会年会	2019年9月	国内

18	Molecular dynamics simulation of the mechanical properties of lipid membranes in the presence of proteins	Diego Ugarte(Kyoto University), Shoji Takada(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
19	Theoretical study on an enzymatic reaction of the hammerhead ribozyme	Ayaka Matsuyama(Kyoto University), Masahiko Taguchi(Kyoto University), Shigehiko Hayashi(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
20	Theoretical study on molecular mechanics of natural anion channelrhodopsin GtACR1	Takafumi Shikakura(Kyoto University), Cheng Cheng(Kyoto University), Shigehiko Hayashi(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
21	Theoretical study on molecular mechanism of a light-driven ion transport of Halorhodopsin from Natronomonas pharaonis	Ryo OYAMA(Kyoto University), Taisuke HASEGAWA(Kyoto University), Shigehiko HAYASHI(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
22	V字型アントラセンパネルを持つ両親媒性分子から形成される分子カプセルのクラスタリング解析	田口真彦(京都大学), 山本裕生(京都大学), 田中大地(京都大学), 林重彦(京都大学)	第13回分子科学討論会	2019年9月	国内
23	チャネルロドプシンに関する理論的研究	成せい(京都大学)	ISSPワークショップ「レチナルタンパク質の光機能発現の物理と化学」	2019年9月	国内
24	ハロロドプシンCl ⁻ イオンポンプの光活性化の分子機構に関する理論的研究	Ryo OYAMA,(Kyoto University), Taisuke HASEGAWA,(Kyoto University), Shigehiko HAYASHI(Kyoto University)	ISSPワークショップ「レチナルタンパク質の光機能発現の物理と化学」	2019年9月	国内
25	液体ピレンの分子動力学シミュレーションと理論的解析	山本裕生(京都大学), Lu Fengniu(国際ナノアーキテクトニクス研究拠点(MANA)), 中西尚志(国際ナノアーキテクトニクス研究拠点(MANA)), 林重彦(京都大学)	第13回分子科学討論会	2019年9月	国内
26	金属タンパク質における酸化還元電位の第一原理計算法の開発	Cheng Cheng(Kyoto University), Shigehiko Hayashi(Kyoto University)	第57回日本生物物理学会年会	2019年9月	国内
27	計算で予測するタンパク質複合体・立体構造と結合親和性	北尾彰朗(東京工業大学生命理工学院)	第7回LiHubフォーラム メゾスケール分子と創薬	2019年11月	国内

2. 学会誌・雑誌等における論文掲載

No.	掲載した論文(発表題目)	発表者氏名	発表した場所(学会誌・雑誌名等)	発表した時期	国際共著 (○を記入)
1	Nucleosome Crowding in Chromatin Slows the Diffusion but Can Promote Target Search of Proteins	Ryo Kanada(RIKEN), Tsuyoshi Terakawa(Kyoto University), Hiroo Kenzaki(RIKEN), Shoji Takad(Kyoto University)	Biophysical Journal, Vol.116, pp.2285-2295	2019年6月	

2	Overcoming the Bottleneck of the Enzymatic Cycle by Steric Frustration	Wenfei Li(Nanjing University), Jun Wang(Nanjing University), Jian Zhang(Nanjing University), Shoji Takada(Kyoto University), Wei Wang(Nanjing University)	PHYSICAL REVIEW LETTERS 122, 238102	2019年6月	
3	De Novo Prediction of Binders and Nonbinders for T4 Lysozyme by gREST Simulations	Ai Niitsu (RIKEN BDR), Suyong Re (RIKEN BDR), Hiraku Oshima (RIKEN BDR), Motoshi Kamiya (RIKEN RCCS), Yuji Sugita (RIKEN CPR)	J. Chem. Inf. Model., 59, 3879–3888 (2019)	2019年8月	
4	Encounter complexes and hidden poses of kinase-inhibitor binding on the free-energy landscape	Suyong Re (RIKEN BDR), Hiraku Oshima (RIKEN BDR), Kento Kasahara (RIKEN BDR), Motoshi Kamiya (RIKEN RCCS), and Yuji Sugita (RIKEN CPR)	Proc. Natl. Acad. Sci. USA, 116, 18404–18409 (2019)	2019年8月	
5	Essential ion binding residues for Na ⁺ flow in stator complex of the Vibrio flagellar motor	Yasuhiro Onoue(Nagoya University), Masayo Iwaki(Nagoya University), Ai Shinobu(Tokyo Institute of Technology), Yasutaka Nishihara(University of Tokyo), Hiroto Iwatsuki(Nagoya University), Hiroyuki Terashima(Nagoya University), Akio Kitao(Tokyo Institute of Technology), Hideki Kandori(Nagoya Institute of Technology), Michio Homma(Nagoya University)	Scientific Reports, Vol.9 (2019)	2019年8月	
6	Replica-Exchange Umbrella Sampling Combined with Gaussian Accelerated Molecular Dynamics for Free-Energy Calculation of Biomolecules	Hiraku Oshima(RIKEN Center for Biosystems Dynamics Research), Suyong Re(RIKEN Center for Biosystems Dynamics Research), Yuji Sugita(RIKEN Center for Biosystems Dynamics Research, RIKEN Center for Computational Science, RIKEN Cluster for Pioneering Research)	J. Chem. Theory Comput.2019, 15, 10, 5199–5208	2019年9月	

7	Phagocytosis is mediated by two-dimensional assemblies of the F-BAR protein GAS7	Kyoko Hanawa-Suetsugu, Yuzuru Itoh, Maisarah Ab Fatah, Tamako Nishimura, Kazuhiro Takemura, Kohei Takeshita, Satoru Kubota, Naoyuki Miyazaki, Wan Nurul Izzati Wan Mohamad Noor, Takehiko Inaba, Nhung Thi Hong Nguyen, Sayaka Hamada-Nakahara, Kayoko Oono-Yakura, Masashi Tachikawa, Kenji Iwasaki, Daisuke Kohda, Masaki Yamamoto, Akio Kitao, Atsushi Shimada, Shiro Suetsugu(1Nara Institute of Science and Technology, University of Tokyo, School of Life Science and Technology, Tokyo Institute of Technology, RIKEN SPring-8 Center, Institute for Protein Research, Osaka University, Theoretical Biology Laboratory, RIKEN, Division of Structural Biology, Medical Institute of Bioregulation, Kyushu University.)	Nature Communications, 10, 4763 (2019)	2019年10月	
8	More efficient screening of protein-protein complex model structures for reducing the number of candidates	Kazuhiro Takemura, Akio Kitao(School of Life Science and Technology, Tokyo Institute of Technology)	Biophysics and Physicobiology 16, 295-303 (2019)	2019年11月	
9	Structural Mechanisms Underlying Activity Changes in an AMPA-type Glutamate Receptor Induced by Substitutions in Its Ligand-Binding Domain	Masayoshi Sakakura(Graduate School of Medical Life Science, Yokohama City University), Yumi Ohkubo(Graduate School of Medical Life Science, Yokohama City University), Hiraku Oshima(RIKEN Center for Biosystems Dynamics Research), Suyong Re(RIKEN Center for Biosystems Dynamics Research), Masahiro Ito(Graduate School of Medical Life Science, Yokohama City University), Yuji Sugita(RIKEN Center for Biosystems Dynamics Research, RIKEN Center for Computational Science, RIKEN Cluster for Pioneering Research), Hideo Takahashi(Graduate School of Medical Life Science, Yokohama City University)	Structure, 27, 1698-1709(2019)	2019年11月	

10	evERdock BAI: Machine-learning-guided selection of protein-protein complex structure	Kei Terayama, Ai Shinobu, Koji Tsuda, Kazuhiro Takemura, Akio Kitao(RIKEN Center for Advanced Intelligence Project, Medical Sciences Innovation Hub Program, RIKEN Cluster for Science, Technology and Innovation Hub, Graduate School of Medicine, Kyoto University, School of Life Sciences and Technology, Tokyo Institute of Technology, Graduate School of Frontier Sciences, The University of Tokyo, Research and Services Division of Materials Data and Integrated System, National Institute for Materials Science)	J. Chem. Phys. 151, 215104 (2019)	2019年12月	○
11	High pressure inhibits signaling protein binding to the flagellar motor and bacterial chemotaxis through enhanced hydration	Hiroaki Hata, Yasutaka Nishihara, Masayoshi Nishiyama, Yoshiyuki Sowa, Ikuro Kawagishi, Akio Kitao(School of Life Science and Technology, Tokyo Institute of Technology, Institute of Molecular and Cellular Biosciences, The University of Tokyo,Department of Physics, Kindai University, Department of Frontier Bioscience, Hosei University)	Scientific Reports, 10, 2351 (2020)	2020年2月	

「生体分子システムの機能制御による革新的創薬基盤の構築」
実施計画

平成31年2月7日改訂版

国立研究開発法人 理化学研究所

奥野 恭史

【改訂履歴】

版	項目	ページ	主な改訂内容（概要）
1.0 (2017. 12. 22)	—	—	中間評価指摘事項を受け、版数変更して対応。
	目次	—	「中間評価における指摘事項への対応状況」目次項目追加。
	1. (3)	3～6	指摘事項に対応した内容を追記。
	1. (5)	8～9	表に指摘事項に対応した内容を追記。
	1. (6)	10	図に協力機関を追加。
	1. (6)	11	指摘事項に対応した内容を追記。
	2－1. (5)	17	図に協力機関を追加。
	2－2. (5)	22	図に協力機関を追加。
	2－3. (2)	24	指摘事項に対応した内容を追記。
	2－3. (5)	27	図に協力機関を追加。
	2－3. (5)	28	表に協力機関と連携内容を追記
	4.	31～32	中間評価における指摘事項への対応状況を追記。
別紙1	2	表に実施期間を追加。	
2.0 (2018. 3. 26)	1. (3)	3～4	指摘事項に対応した内容と図を追記。
	1. (7)	13	H30年度、H31年度の「京」と「京」以外の計算資源量を変更。資源量の変更に関する内容を削除。
	2－3. (2)	26	指摘事項に対応した内容を追記。
	別紙2	3～10	H30年度、H31年度の数値を変更。
3.0 (2019. 2. 7)	—	—	研究体制の変更。
	1. (3)	5	指摘事項に対応した内容を追記。
	1. (3)	7	ポスト「京」運用後、5年後・10年後のアウトカムイメージを追加。
	2－3. (2) ①	26	指摘事項に対応した内容を削除。
	2－3. (2) ③	26	指摘事項に対応した内容を追記。

目次

1. 実施概要	1
(1) 目的・意義	1
(2) 研究開発内容	1
(3) 目標・期待される成果.....	3
(4) 周辺領域への波及効果、課題全体における計算科学やシミュレーションの位置づけ	8
(5) 年次計画	9
(6) 実施体制	11
(7) 必要計算資源.....	13
2. 研究開発内容詳細.....	14
2-1. サブ課題A. ポスト「京」でのMD高度化とアルゴリズム深化.....	14
(1) 目的・意義	14
(2) 実施内容	14
(3) 目標・期待される成果.....	16
(4) 「京」でできていること、ポスト「京」でなければならないこと.....	17
(5) 実施体制	17
2-2. サブ課題B. 次世代創薬計算技術の開発	19
(1) 目的・意義	19
(2) 実施内容	19
(3) 目標・期待される成果.....	20
(4) 「京」でできていること、ポスト「京」でなければならないこと.....	21
(5) 実施体制	22
2-3. サブ課題C. 創薬ビッグデータ統合システムの開発.....	25
(1) 目的・意義	25
(2) 実施内容	25
(3) 目標・期待される成果.....	27
(4) 「京」でできていること、ポスト「京」でなければならないこと.....	27
(5) 実施体制	28
3. 採択時の留意事項への対応状況	30
4. 中間評価における指摘事項への対応	32

1. 実施概要

(1) 目的・意義

この十数年、製薬業界では、新薬の承認数が横ばい状態（20品目程度／年）であるのに対し、研究開発費が増え続けているという深刻な問題に直面している。特に、近年では、医薬品開発が容易なターゲット疾患のほとんどは開発し尽され、残された疾患の多くは従来型の創薬アプローチでは治療薬の創出が困難であると考えられている。このことから、「開発費を抑えながら、新薬創出を加速すること」は創薬・医療分野にとっての最重要課題となっている。そこで本研究では、ポスト「京」を基軸とした創薬計算基盤を開発し、製薬会社やアカデミア創薬研究で利活用することで、これらの課題を克服することを目的としている。具体的には、ポスト「京」の演算能力を最大限に活かす分子シミュレーション技術を開発し、生体分子システムの時間的空間的機能解析に資する新たな構造生命科学を開拓することで、次世代の創薬アプローチを切り拓く計算技術の開発を目指す。さらに、これらの要素計算技術を創薬計算フローに沿って連結した創薬ビッグデータ統合システムを開発することで、高精度かつ超高速の革新的な創薬計算基盤を確立する。

本プロジェクトで開発するポスト「京」を基軸とする創薬計算基盤が製薬現場で利用されることで、これまでの創薬の実験プロセスが計算機シミュレーションに置き換わるなどの開発プロセスの効率化につながり、医薬品開発コストの劇的削減、ひいては医療費の削減につながるものと期待できる。さらには、標的タンパク質の動的機能制御創薬、タンパク質-核酸の超分子複合体を標的とするエピジェネティック創薬、超大規模生体分子システムのシミュレーションに基づくシステム創薬などの新たな創薬の革新的アプローチを開拓することで、これまで開発が困難であった新薬の創出や医薬品開発のスピードアップにつながり、結果として新薬を求める患者全体への貢献が期待できる。

(2) 研究開発内容

本研究は、ポスト「京」とのコデザインによって分子動力学（以下、MD）計算の超高速化と創薬計算要素技術の深化を図るサブ課題A、ポスト「京」仕様のMD計算法を新たな構造生命科学の開拓と次世代創薬計算技術に応用するサブ課題B、およびサブ課題Aとサブ課題Bで開発された計算技術を実践的な創薬計算フローとして統合化し製薬会社に提供するサブ課題Cの3つのサブ課題で研究遂行する。なお、研究実施の年次計画としては、調査研究・準備研究フェーズ（H26、27）において研究体制の整備と研究実施の準備を行う。本格実施フェーズの当初2年間（H28、29）においてコデザインによるポスト「京」向けのMDカーネルの高速化と各種計算手法の技術開発（主にサブ課題Aとサブ課題B）と標的分子系のモデリング（主にサブ課題Bとサブ課題C）を行う。続く2年間（H30、31）において、サブ課題Aがフルアプリとしてポスト「京」に最適化され最新の計算手法が導入されたMDプログラムを完成させるとともに、次いでサブ課題Bとサブ課題Cが、利用可能になった計算手法から順次標的分子系への計算適用を行い、性能評価と精密化を図る。これにより、ポスト「京」の運用開始直後から創薬プロセスの革新につながるMDシミュレーションが実施できる。

各サブ課題の主な研究開発概要は次のとおりである。

サブ課題A：ポスト「京」でのMD高度化とアルゴリズム深化

理化学研究所計算科学研究機構を中心に開発しているMD計算ソフトGENESISをターゲットアプリとしてコデザインによるMD計算の超高速化をはかるとともに、高精度かつ超高速の創薬計算基盤を

(3) 目標・期待される成果

本重点課題では、創薬分野および構造生命科学の進展にとって重要な 3 つのテーマ「長時間分子シミュレーション」「超大規模生体分子システムシミュレーション」「創薬ビッグデータ統合システム」を 3 つのサブ課題共通の全体目標として設定した。

- 長時間分子シミュレーション

ポスト「京」を用いることで、タンパク質が機能発現している時間スケールであるミリ秒以上の分子シミュレーションを実現し、動的機能に着目した新たな構造生命科学を革新するとともに、創薬標的分子の動的機能制御を指向したより効果的な新薬の開発を可能にする。本テーマは、サブ課題 A のポスト「京」での MD 高度化、長時間ダイナミクス法、自由エネルギー計算法、サブ課題 B の動的分子機能制御、サブ課題 C の機能制御部位データベース、高精度薬剤デザインが連携して研究開発する。

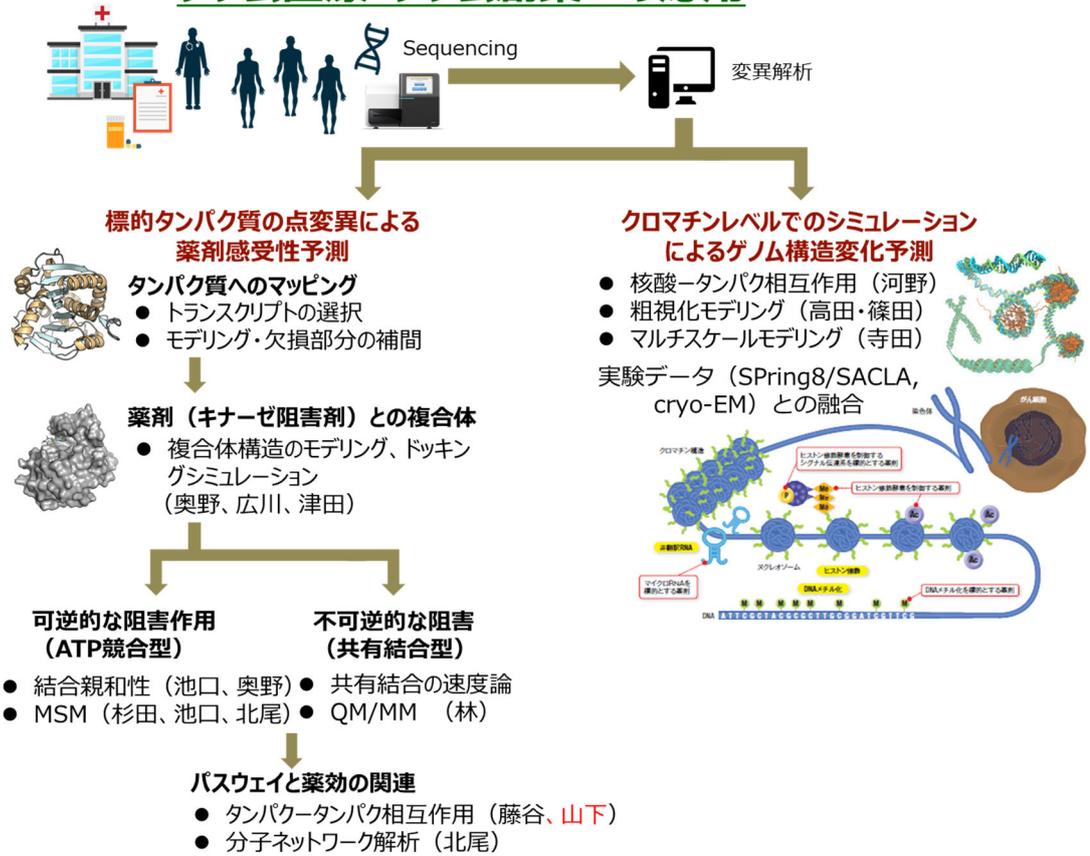
- 超大規模生体分子システムシミュレーション

マルチスケール（全原子および粗視化）MD の Capability 計算を最適化することで、「京」ではサブマイクロ秒であった計算時間を、ポスト「京」によって数十マイクロ秒に延長した超大規模生体分子システム（多数のタンパク質および核酸等からなる細胞内環境やウイルスなど）のシミュレーションを実現する。これにより、単独のタンパク質のみの環境とは異なる多種多様な生体分子が複雑に混みあった細胞内環境のダイナミクスを解明するとともに、エピジェネティック創薬やシステム創薬などの次世代の創薬手法の開拓を目指す。本テーマは、サブ課題 A のポスト「京」での MD 高度化、粗視化モデリング、サブ課題 B のタンパク質間相互作用制御、核酸-タンパク質相互作用制御、ウイルス標的、細胞内環境シミュレーション、サブ課題 C の高精度薬剤デザイン、創薬関連ビッグデータ、標的分子ネットワーク、マルチスケールモデリングが連携して研究開発する。

- 創薬ビッグデータ統合システム

ポスト「京」とのコデザインで高度化された MD 計算ソフトを基盤として、各種創薬計算手法を創薬計算フローに沿って連結した創薬ビッグデータ統合システムを開発する。これにより、大規模な候補化合物と複数の創薬標的・副作用関連タンパク質の莫大な組み合わせから、より効果的で副作用の少ない医薬品候補物質を高精度かつ超高速にデザインできる創薬基盤を構築し、製薬会社等に提供する。本テーマは、サブ課題 A とサブ課題 B が開発する全ての要素計算技術をサブ課題 C が中心となって創薬計算フローに沿ってシステム化を図る。全体の共通テーマとしてゲノム医療・ゲノム創薬を応用ターゲットにする。ゲノム医療・ゲノム創薬は、現在世界的に急速に臨床実装、創薬実装が進むテーマであり、本テーマにサブ課題 A、B、C の全員でチャレンジすることでポスト「京」の創薬応用の課題が網羅されると考える。具体的には、ゲノム医療・ゲノム創薬では、「標的タンパク質の点変異による薬剤感受性予測」と「クロマチンレベルでのシミュレーションによるゲノム構造変化予測」にフォーカスを絞り、サブ課題 A、B、C で開発する要素技術を下図のとおり、適用する。

ゲノム医療・ゲノム創薬への応用



定量的、定性的目標について、上記3つの主要テーマに加え、「MD 計算高度化 (コデザイン)」と「分野振興 (人材育成等)」の5項目で、以下に記す。

アウトプット成果 (平成 29 年度終了時) :

- ① MD 計算高度化 (コデザイン) - コデザインにより、ターゲットアプリ GENESIS における MD カーネルをポスト「京」向けのアルゴリズムに変更し、単独の MD として数倍の高速化を図る (具体的な目標値はポスト「京」の性能設計における守秘義務から明記できない。)
- ② 長時間分子シミュレーション-「京」等を用いて、標的タンパク質について10マイクロ秒以上の動的構造解析を実現する。
- ③ 超大規模生体分子システムシミュレーション-サブミクロンサイズに迫るウイルスカプシド系の薬剤相互作用解析を想定し、直径 30nm 程度の B 型肝炎ウイルスカプシド系の分子モデリング (全原子及び粗視化モデル) を行う。また、細胞環境を考慮した混み合い系でのマイクロ秒以上のシミュレーション、100 万原子を超えるタンパク質間相互作用解析、および核酸タンパク質間相互作用における2次元以上の多次元自由エネルギー解析を実施する。
- ④ 創薬ビッグデータ統合システム-創薬計算フローに実装する要素計算技術の調査、各種創薬計算手法の評価を行い創薬ビッグデータ統合システムの設計を完了する。低分子化合物とタンパク質間の結合自由エネルギー計算法について、「京」等を用いて従来法より数倍以上少ない演算量での計算を可能とする新手法を開発する。
- ⑤ 分野振興 (人材育成等) - 20 社以上が参加する製薬コンソーシアム (バイオグリッドセンター

関西主催)により、「京」での利用が進むとともに、ポスト「京」に対する意見集約がなされ、重点課題1での研究計画に反映される。また、大学、研究機関、製薬コンソーシアム等と協力し人材育成プログラムの企画を行い、実施検討する。

アウトプット成果（本格実施フェーズ終了時）：

- ① MD 計算高度化（コデザイン）ーコデザインにより、GENESIS をフルアプリとして「京」で行われた創薬応用 MD と比較して、ポスト「京」において Capacity 計算の総量として数十倍から百倍程度の総演算を行うことを可能にする。さらに超大規模生体分子システムのマルチスケール（全原子および粗視化）MD 計算をポスト「京」のアーキテクチャーに最適化する。ポスト「京」の世界最高レベルの性能をフル活用できる高速な MD シミュレーションをターゲットアプリ GENESIS で実現する。
- ② 長時間分子シミュレーションー「京」等を用いて、薬剤と標的タンパク質との薬剤反応について 10 マイクロ秒以上の動的構造解析を実現する。レプリカ交換 MD などの国際的優位性のある拡張アンサンブルやマルコフ状態モデルなどのアルゴリズムを活かし、効率的動的構造解析を実現する。生命現象の非常に幅広い時間スケールに対応して、詳細な電子状態から臓器レベルに至る幅広い時間スケールに対応したマルチスケールシミュレーション法を開発する。さらに、構造生物学実験と連携し、実験情報を統合する方法を確立する。
- ③ 超大規模生体分子システムシミュレーションー「京」等において、直径 30nm 程度の B 型肝炎ウイルスカプシドへの低分子化合物の取り込み機構（結合、透過）について、自由エネルギー評価を可能とする計算手法を開発する。タンパク質間相互作用制御では、「京」等において 100 万原子数を超える系で短時間(10 マイクロ秒以下)シミュレーションによる低分子の活性予測を可能にする。さらに、核酸-タンパク質間相互作用制御では、翻訳後修飾を受けた DNA 結合タンパク質の認識機構について 2 次元以上の多次元自由エネルギープロファイル解析を可能とする計算手法を開発する。また、細胞内環境を考慮した巨大な生体分子システムでのタンパク質・基質結合の定量的評価法を確立する。全原子と粗視化モデルを接続する独自性の高いマルチスケールシミュレーション手法を確立し、それらを連成することで世界最大レベルの超大規模生体分子シミュレーションシステムの長時間ダイナミクスを計算できる方法を確立する。
- ④ 創薬ビッグデータ統合システムー上記①から③の各種創薬計算手法を具体的な創薬テーマ（造血器悪性腫瘍・固形腫瘍の抗癌剤、腎疾患治療薬、アルツハイマー治療薬、薬剤耐性シミュレーションや製薬会社の個別テーマ等）に適用し、性能評価を実施する。特に、ゲノム医療、ゲノム創薬を共通のターゲットテーマとして上記①から③の各種計算手法を応用、評価する。ここで具体的な標的タンパク質としては、各種がんのドライバー変異や薬剤耐性変異が報告され、がんのゲノム医療に重要な標的タンパク質である ALK, EGFR, RET を中心ターゲットとする。特にこれらの標的タンパク質は、国内のがんゲノム研究者とのコラボレーションの対象にもなっており、実験検証による計算の精度評価も可能である。本計算では、これらの標的タンパク質に様々なアミノ酸変異が生じた際の薬剤との結合自由エネルギーを従来比 2 倍の計算効率で算出することを目指す。さらに上記①から③を統合化することで、ポスト「京」を用いた一連の創薬計算フローが実行できる基盤「創薬ビッグデータ統合システム」の初版を完成させる。さらに、人工知能 (AI)

技術を分子シミュレーションと融合させ、大幅な効率化を図る。

- ⑤ 分野振興（人材育成等）－製薬コンソーシアム、大学、研究機関に対して、創薬ビッグデータ統合システムの初版を公開する。また、実施してきた人材育成プログラムにより、製薬コンソーシアム、大学、研究機関に創薬ビッグデータ統合システムを利用できる人材が育成され、システムの利用が開始される。また、人工知能の創薬応用のコンソーシアム（製薬企業、IT 企業、アカデミア 90 団体以上参加）である LINC と連携し、人工知能技術のポスト「京」応用を推進する。

アウトプット成果（ポスト「京」運用開始 5 年後）：

- ① MD 計算高度化（コデザイン）－分子動力学プログラム GENESIS をポスト「京」を用いて評価し、さらなる最適化と並列化を行うことで「京」と比較して数十倍から百倍程度の Capacity 計算を実現する。また、Capability 計算として世界最大級の MD 計算を長時間実行するソフトウェア基盤を評価するとともに、さらなる最適化と高速化を図る。MD 計算ソフトウェアの高度化は、以降のサイエンス上の成果を得る上での基盤技術となる。
- ② 長時間分子シミュレーション－ポスト「京」において、標的タンパク質のミリ秒以上の動的構造解析を実現し、低分子化合物が結合・解離する過程や、それに対応した標的タンパク質の立体構造変化過程など、動的構造の変化を解析する。創薬における薬剤の結合・解離過程のような速度論的シミュレーションはポスト「京」によって初めてなし得る。さらに、生命現象の幅広いスケールに対応したマルチスケールシミュレーションを実施し、酵素反応など、電子状態が関わる詳細な化学反応から、臓器運動のような巨視的運動（重点課題 2 の UT-Heart との連結）まで、非常に幅広い時間スケールのシミュレーションを実現する。さらに、ポスト「京」ならではの長時間分子シミュレーションと日本が誇る高度な構造生物学実験施設と連携することで、静止した実験構造から動的な機能発現過程の構造生物学を展開する新たな分野を開拓する。
- ③ 超大規模生体分子システムシミュレーション－ポスト「京」を用いて、サブミクロンサイズのウイルスカプシド系について大規模分子系シミュレーションを実行し、薬剤との相互作用、送達・作用機構を解析する。タンパク質間相互作用制御では、ミリ秒オーダーで起きる構造変化を持ったタンパク質複合体に作用する低分子化合物の活性予測を可能にする。さらに、核酸-タンパク質間相互作用制御では、多次元自由エネルギープロファイル解析を行い、翻訳後修飾の DNA 結合に対する影響を明らかにする。また、チャレンジングな超大規模生体分子システム・ネットワーク系（多数のタンパク質および核酸等からなる細胞内環境や核内環境など）における薬剤等低分子化合物の結合計算等を実現し、創薬応用に必要な知見を得る。超大規模生体分子システムシミュレーションでは、全原子モデルと粗視化モデルを高度に連成したシミュレーションを実現し、ポスト「京」で初めて実現する長時間ダイナミクスを行い、シミュレーションから生物学実験へ新たな視点と課題を与える新しい分野を開拓する。
- ④ 創薬ビッグデータ統合システム－ポスト「京」を用いて、「京」時代と比較して数十倍から百倍多くのスクリーニングを実現し、創薬ビッグデータ統合システムを構造生命科学研究所や創薬研究に応用する。特にゲノム医療・ゲノム創薬を応用ターゲットとして核内のマルチスケールシミュレーションを実現し、世界的に注目される Precision Medicine の予測基盤とする。さらに、製薬会社やアカデミア創薬研究の具体的な創薬テーマに創薬ビッグデータ統合システムを適用すること

で創薬応用を促進する。創薬テーマとしては 100 テーマがポスト「京」で計算されることを目標とする。

- ⑤ 分野振興（人材育成等）－製薬コンソーシアム参加企業を中心とした製薬会社での現場利用を促進し、製薬会社で定期的にシステム利用がなされる状態を目指す。また、人材育成プログラムにより、現場利用に十分な人材が育成される。

アウトカム効果（ポスト「京」運用開始 5 年後）：

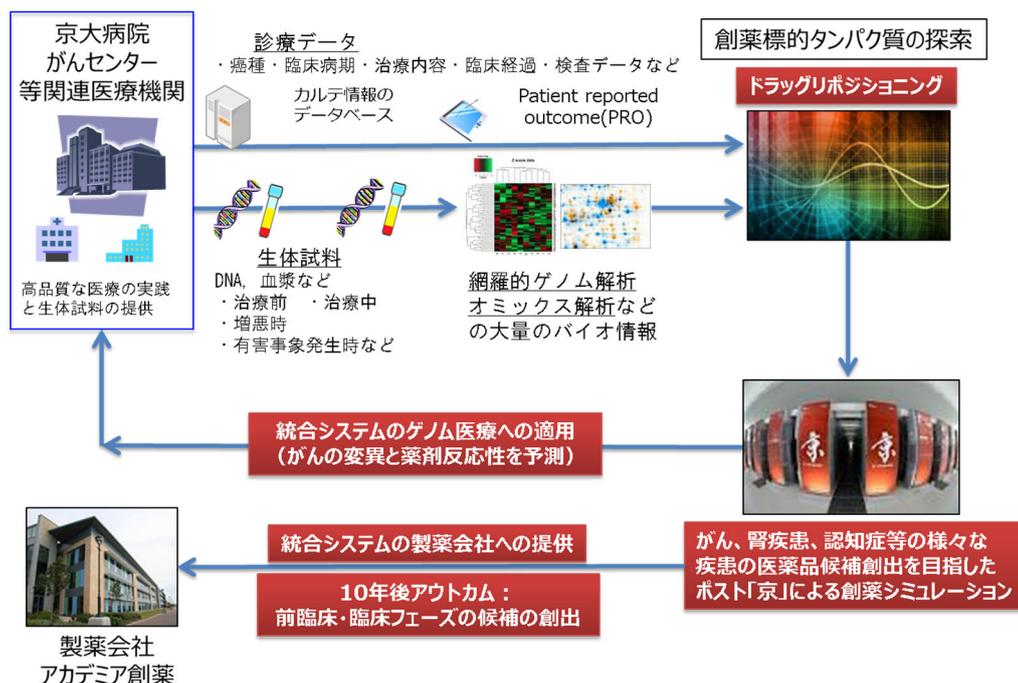
創薬ビッグデータ統合システムを通じて、アカデミア創薬および製薬各社の創薬テーマのリード化合物の創出に成功していること、およびポスト「京」での MD 計算結果や先端計測実験との連携で創出された研究成果が重要な学術論文として世界に発信されていることを目標とする。具体的な、計算ターゲットとしては、転写因子・ヒストン・DNA 等からなる核内生体分子系、生体膜と膜タンパク質からなる生体膜分子系、酵素タンパク等の機能性マルチドメイン分子系を対象とし、創薬テーマとしては、造血管悪性腫瘍・固形腫瘍の抗癌剤、腎疾患治療薬、アルツハイマー治療薬、薬剤耐性シミュレーション、ドラッグリポジショニングや製薬会社の個別テーマ等を対象にする。特に、世界的に急速な進展が進むゲノム医療・ゲノム創薬への臨床実装・社会実装を目指す。

また、タンパク質・薬剤の結合状態の予測精度の改善だけでなく、タンパク質の機能制御という観点から新しい概念を、計算科学を中心に導入し、実験によって検証する。そのために、タンパク質の動的性質と細胞環境（溶媒との相互作用、分子混雑、生体膜との相互作用など）のより信頼性の高いモデリングを実現する。

アウトカム成果（ポスト「京」運用開始 10 年後）：

創薬ビッグデータ統合システムを通じて、前臨床・臨床フェーズに進む候補化合物の創出に成功していることを目指す。計算科学と構造生物学とが融合した新たな構造生命科学に創成に資する。

ポスト「京」運用後、5年後・10年後のアウトカムイメージ



(4) 周辺領域への波及効果、課題全体における計算科学やシミュレーションの位置づけ

医薬品開発の成功確率は約 20,000 分の 1 と低く、製薬業界では、開発コストの高騰と承認新薬数の低迷が重要課題の一つとなっている。例えば、Nature Reviews Drug Discovery 9, 203-214 (March 2010) での報告では、一品目当たりの平均医薬品開発費は約 1,800 億円と見積もられているが、ポスト「京」による分子シミュレーションの直接的効果を示す開発プロセスは初期のシード化合物探索とリード化合物最適化の段階であり、670 億円の開発費と 4.5 年の開発期間を要すると報告されている。したがって、ポスト「京」による創薬ビッグデータ統合システムを開発することで、実験を代替する予測、膨大な実験結果の解析、新規知見の発見が促進され、開発の効率化、低コスト化、新薬創出につながるものと期待されている。さらに、分子シミュレーション技術は、製薬産業以外の領域（化学系、農林水産系、食品系等）にも応用可能であり、バイオ産業全般の活性化に資するものと考えられる。

また、現在、製薬業界では、医薬品開発が容易なターゲット疾患のほとんどは開発し尽され、残された疾患はこれまでの創薬アプローチでは新薬の創出が困難なものばかりであると危惧されている。これに対して本研究は、ポスト「京」の演算能力を最大限活かすことで、動的機能制御創薬、エピジェネティック創薬、超大規模生体分子系シミュレーションに基づくシステム創薬などの新たな革新的創薬アプローチを開拓する可能性を有する。また、現在の創薬プロセスの大半は実験を中心に展開されているが、ポスト「京」により高精度かつ超高速な創薬計算を達成することで、創薬の実験プロセスが計算機シミュレーションに置き換わるなどのプロセス革新を誘起し、我が国の製薬産業発展の求心力になるものと考えられる。

保健医療における波及効果としては、ポスト「京」による革新的創薬基盤が普及することで、医薬品開発のスピードアップや新薬の創出が実現されることにより、結果として新薬を求める患者全体への貢献が期待できる。また、従来法では発見できなかった医薬品候補の発見により、難治性疾患の創薬に繋がることも期待できる。さらには、医薬品開発の効率化につながることから、開発コストの劇的削減とそれによる医療費削減が期待され、超少子高齢化を迎える我が国の社会保障問題解消の一助につながるものと考えられる。

更に、ポスト「京」で初めて実現される超大規模生体分子系シミュレーションや長時間スケール分子シミュレーションによって、生体分子システムの時間的空間的機能解析が可能になり、これに実験科学が融合することで、新たな構造生命科学の開拓と生命科学全体の発展に資することが期待される。

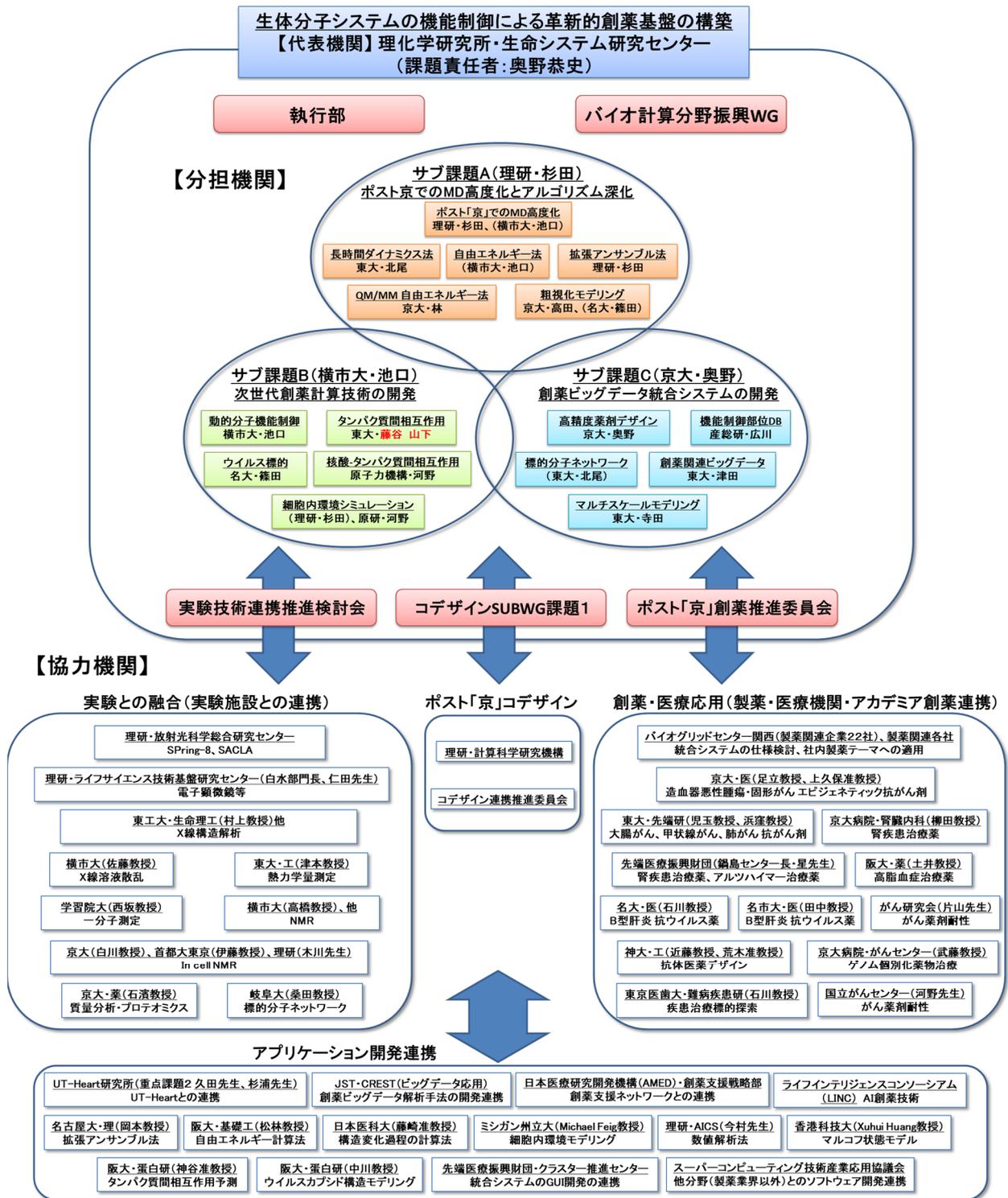
(5) 年次計画

課題全体	中間目標 (平成29年度)	ポスト「京」仕様のMDと各種創薬計算手法を連結した統合システムのプロトタイプを完成する。
	最終目標 (平成31年度)	ポスト「京」仕様のMDと各種創薬計算手法を統合化した創薬ビッグデータ統合システムを完成する。

サブ課題名 (分担機関・責任者)	調査研究・準備研究フェーズ		本格実施フェーズ			
	平成26年度	平成27年度	平成28年度	平成29年度	平成30年度	平成31年度
サブ課題A ポスト「京」でのMD高度化とアルゴリズム深化 (理化学研究所・杉田有治)	(目標) ポスト「京」でのMD高度化の方針を検討 (実施内容) 計算機関係者との連携を開始	(目標) MD高度化を開始し、創薬応用手法を検討 (実施内容) 既存手法のテストとコードのデザインを実施	(目標) 新規な創薬応用手法の開発を推進 (実施内容) 各実施者がアルゴリズム開発とテストを実施	(目標) 大規模生体システムの効率的計算法を開発 (実施内容) 新規手法のテストを行い、既存手法と比較	(目標) コデザインと新規手法を結合 (実施内容) 創薬応用計算を効率化する構造探索・自由エネルギー計算と大規模システムのマルチスケールシミュレーションの開発・改良	(目標) コデザインと新規手法を結合 (実施内容) 長時間ダイナミクスを創薬応用計算において実現し、超大規模な生体分子系シミュレーション法を確立することで国際的優位性をもつ計算手法を確立
サブ課題B 次世代創薬計算技術の開発 (横浜市立大学・池口満徳)	(目標) 課題実施内容の検討 (実施内容) 動的アンサンブル法 (DA法)、タンパク質間相互作用 (PPI)、ウイルス標的について検討	(目標) 課題実施内容の現状技術検討 (実施内容) DA法、PPI、ウイルス標的、核酸・タンパク質相互作用 (NPI) について、現状技術を検討	(目標) 次世代創薬計算技術開発の本格実施開始 (実施内容) DA法、PPI、NPI、ウイルス標的、細胞標的について、本格開発を開始	(目標) 次世代創薬計算技術プロトタイプ完成 (実施内容) DA法、PPI、NPI、ウイルス標的、細胞標的について、計算手法の目処を確立	(目標) 次世代創薬計算技術プロトタイプ改良 (実施内容) DA法、PPI、NPI、ウイルス標的、細胞標的について、国際的優位性のある計算手法を導入・改良	(目標) 次世代創薬計算手法完成 (実施内容) DA法、PPI、NPI、ウイルス標的、細胞標的について、国際的優位性のある計算手法を確立

<p>サブ課題C 創薬ビッグデータ 統合システムの開 発（京都大学・奥 野恭史）</p>	<p>（目標） 統合システム開 発のグランドデ ザイン （実施内容） サブ課題AとBと 議論をしながら、 開発計画を検討</p>	<p>（目標） 統合システムの 仕様決定 （実施内容） 製薬会社との議 論を行い、要素技 術・システム仕様 を調査検討</p>	<p>（目標） 創薬計算フロー の要素アプリの 開発 （実施内容） 創薬計算フロー の各要素アプリ を順次開発</p>	<p>（目標） 統合システムの プロトタイプ完 成 （実施内容） 創薬計算フロー をテスト実行し、 プロトタイプを開 発</p>	<p>（目標） サブ課題AとBの 方法論の組み込 み開始 （実施内容） 統合システムに サブ課題AとBで 開発された方法 論を実装、および ゲノム医療・ゲノ ム創薬への応用</p>	<p>（目標） 統合システムの 完成 （実施内容） 創薬計算フロー をテストし、統合 システムを改良、 およびゲノム医 療・ゲノム創薬へ の応用</p>
--	--	---	---	--	--	--

(6) 実施体制



上図に示すとおり、代表機関・理化学研究所生命システム研究センターのもと、執行部、ポスト「京」創業推進委員会、実験技術連携推進検討会、バイオ計算分野振興WG、コデザインサブWGを組織化し、分担機関による研究実施と協力機関との強力な連携を図りながら、プロジェクトを遂行する。協力機関は、4つのグループに大別し、創薬・医療応用を主旨とした製薬会社・医療機関・アカデミア創業

研究機関のグループ、先端機器計測など実験との融合を主旨とした実験施設・研究機関のグループ、アプリケーション開発や他創薬系プロジェクトとのアプリケーション連携を主旨とした研究機関等のグループ、ポスト「京」とのコデザイン連携を図るグループから成る。

各組織の役割は次の通りである。

執行部：課題責任者奥野（理化学研究所/京都大学）、サブ課題責任者杉田（理化学研究所）、池口（横浜市立大学）と事務（理化学研究所 QBiC 推進室）で構成し、毎月定期的に開催する運営会議のほか、適宜、TV 会議やメール討論を通じて、プロジェクト運営全般の検討・決裁・指揮をする。

重点課題 1 ワークショップ（全体会議）：サブ課題実施者や若手研究者を含むプロジェクト参加者全員を対象とした研究ワークショップを年 2 回開催し、進捗確認とサブ課題間・分担機関間での相互連携を推進する。年に 2 回開催するワークショップにおいて、初回は製薬会社へ公開型で開催し、2 回目は諮問委員会委員の参加型で行う。ワークショップの参加対象企業には、KBDD コンソーシアム（「京」の創薬応用コンソ）、LINC（AI 創薬コンソ）の企業へと参加対象を広げ、製薬現場の広範囲にわたるニーズを受けた開発を展開する。

ポスト「京」創薬推進委員会：製薬会社・医療機関・アカデミア創薬研究機関等の協力機関と連携し、創薬・医療応用を促進するための委員会。製薬会社や医療機関からの有識者を中心に組織化し、創薬ターゲットの選定や創薬ビッグデータ統合システムの産業化利用などの具体的な創薬・医療応用に関する検討を行う。また、サブ課題 C の課題を A, B にも反映するために連携促進担当 2 名（京大 1 名、理研 1 名）を設置する。

実験技術連携推進検討会：先端実験施設や実験研究者との連携を促進するとともに、先端機器計測などの様々な実験技術とポスト「京」による分子シミュレーション技術とを融合を検討し、新たな構造生命科学の創成を目指す。また、サブ課題 C の課題を A, B にも反映するために連携促進担当 3 名（理研）を設置する。

コデザインサブ WG：MD 計算ソフト GENESIS をターゲットアプリとして、ポスト「京」のコデザインを推進する。また、各サブ課題で開発する要素計算技術の数理的枠組みを検討するために、アプリケーション開発連携の協力機関メンバーの協力を得て、数理モデル検討ミーティングを適宜開催する。

バイオ計算分野振興 WG：当該研究の成果普及とバイオ計算人材の人材育成を図る。具体的には、当該研究で開発する各種創薬計算技術やそれらを実装した創薬ビッグデータ統合システムの利用説明会・講習会やバイオ計算分野関連の各種勉強会・講演会の開催などが挙げられる。また、これらの実施に当たっては重点課題 2 等と連携し、相乗効果を生み出していく。

また、当該 WG では、本プロジェクトで雇用する若手研究者の研究・発表活動を奨励し、若手人材育成とキャリアパス形成を促進する。具体的な人材育成計画は、以下の通りである。

1. プロジェクトで雇用する若手研究者全員より希望キャリアパスの個別ヒアリングを実施する。
2. 担当 PI 教員とともに、各若手研究者の希望・適性・年齢・研究内容等を考慮して、キャリアパス確保にむけた育成計画を検討する。（1.と 2. は H27 年度と H28 年度中に実施）
3. 各 PI 教員が所属する大学や他大学で開催されるセミナーへの聴講を若手研究者に紹介し、基礎生命科学・創薬・医療などの異分野の視野と知識の習得を推進する。
4. 各 PI 教員や協力機関を通じたアカデミア・産業界へのネットワークを活用し、求人情報を積極的に収集するとともに、各若手研究者の研究内容・業績等を広く紹介する広報を行う。

本プロジェクトで雇用する研究者予定数は、それぞれのチームについて、奥野 3名、杉田 2名、池口 2名、北尾 2名、林 1名、藤谷 (2018.12月まで) および山下 (2019.1月以降) 1名 (ポスドク以外教員等 2名)、篠田 2名、広川 1名、津田 1名、寺田 1名 (ポスドク以外教員 1名)、河野 1名であり、若手研究者 (ポスドク) の雇用予定数は計 17名となる。これらの若手研究者を対象に上記の人材育成プランを実施し、適時、キャリアアップをはかる。

なお、現在、我が国ではバイオ計算人材が枯渇しており、バイオ系のアカデミア、企業ともに若手人材の需要が高い状況であることから、こうした人材育成は非常に重要視されている。就職先としては、アカデミアでは薬学・医学などの応用系も含むバイオ系の大学研究科や研究機関、企業としては製薬、化学、食品関連企業が有力な就職先となる。

(7) 必要計算資源

必要計算資源は、ターゲットアプリケーション GENESIS、および、各課題で用いられる MP-CAFEE (GROMACS)、MODYLAS、SCUBA (これらのアプリは HPCI 戦略プログラムで活用されており、「京」で実行実績あり) の見積もりから積算した。サブ課題 A では、GENESIS および、その QM/MM 応用である QM/GENESIS の利用で 570 万ノード時間/年 (「京」) 107 万ノード時間/年 (「京」以外) である。この「京」の計算資源は、現在の「京」での GENESIS のベンチマーク値を用いると 10 万原子系 10 μ s, 100 万原子系 1 μ s, 1000 万原子系 0.1 μ s 弱程度に相当する。ストレッチ目標では、超大規模な系を対象として見積もった。コデザインはサブ課題 A と連携しており、その必要資源量も含まれている。サブ課題 B では、GENESIS に加えて、現在 HPCI 戦略プログラムで活用されており、ポスト「京」でも各課題で利用予定である MP-CAFEE, MODYLAS, SCUBA での見積もりを行い、824 万ノード時間/年 (「京」) である。GENESIS は動的分子機能制御、MP-CAFEE はタンパク質間相互作用制御、MODYLAS はウイルス標的総逆計算技術、SCUBA は核酸-タンパク質相互作用制御の各研究開発で用いる。開発サブ課題 C では、GENESIS、GROMACS の利用で、394.4 万ノード時間/年 (「京」) とした。この「京」の計算資源は、おおよそ、結合自由エネルギー計算 (120 ノード時間/ケース) を評価化合物 33 \times 創薬関連タンパク質 100 \times 標的構造数 10 に対して行う計算資源に相当する。

「京」の計算資源量

(単位：ノード時間/年)

	H27 年度	H28 年度	H29 年度	H30 年度	H31 年度
	1,200,000	17,884,000	23,249,200	23,249,200	23,249,200

「京」以外の計算資源量

(単位：ノード時間/年)

	H27 年度	H28 年度	H29 年度	H30 年度	H31 年度
	6,770,000	6,770,000	6,770,000	6,770,000	6,770,000

2. 研究開発内容詳細

2-1. サブ課題A. ポスト「京」での MD 高度化とアルゴリズム深化

(1) 目的・意義

創薬において、構造の柔軟性や溶媒との相互作用を考慮するために、MD 計算の活用が不可欠になってきた。また、生命科学において、様々な実験や計測結果を検証し解釈するためにも MD 計算は重要な役割を果たしている。HPCI 戦略プログラム分野 1（以下、戦略分野 1）では、従来考慮されていなかった細胞環境（細胞質や核内混み合い環境）を露わに考慮した MD 計算、タンパク質の柔軟性を考慮した薬剤結合自由エネルギー計算などを実施してきた。しかし、タンパク質と薬剤候補化合物の莫大な組み合わせ、薬剤によるタンパク質の機能制御を理解するための長時間（ミリ秒）の分子運動、リボゾームやウイルスなどの生体超分子複合体や細胞環境を考慮した巨大な系などを MD により計算するためには、現在の計算機上でのパフォーマンスは十分なものではない。本サブ課題では、ポスト「京」開発主体とのコデザインを通してポスト「京」の演算能力を最大限に生かす事が可能な MD ソフトウェアの開発と高度化を行う。さらに新規性の高いアルゴリズムを開発することで、医薬品開発を加速化し画期的新薬を創出する基盤を構築する。MD 計算による応用計算は、現在行われているタンパク質と薬剤候補化合物に関する莫大な組み合わせ計算と、将来の創薬応用につながると期待される大規模生体分子システムをターゲットとした計算に分けられる。HPC の用語では、前者は小・中規模の計算を大量に行う Capacity 計算に、後者は複雑で大規模な計算の限界に挑戦する Capability 計算に対応する。本課題では、Capacity 計算と Capability 計算のいずれにおいてもポスト「京」の演算性能を最大限に発揮する MD 計算エンジンを開発し、新たな創薬応用計算の可能性を探索する。

(2) 実施内容

MD 計算の高度化研究を行うために、ポスト「京」開発主体と連携したコデザインを実施する。対象とするソフトウェアは、理化学研究所計算科学研究機構を中心に開発している GENESIS である。MD 計算では演算コストの高いのは短距離力の非共有結合相互作用と、クーロン力などの長距離相互作用である。両者とも CPU の特性によって最適化の指針が異なることがわかっているため、ポスト「京」で採用される CPU の特性を生かした最適化を実施する必要がある。短距離相互作用計算では特に、SIMD 化率を向上させる必要があり、理化学研究所計算科学研究機構との緊密な連携が必要である。長距離相互作用については、通常用いられる Particle mesh ewald 法ではノード通信を必要とする FFT 計算が含まれるため、高速な通信ライブラリの利用等の検討が必要である。これらを通して創薬応用 MD を効率的にポスト「京」を用いて実施する基盤を構築する。また、ポスト「京」の演算性能を最大限に発揮する全原子 MD による大規模な系での高速化も必要である。大規模な並列化を実現するため高速な通信ライブラリーを用いて高い並列化率を維持する必要がある。また、超大規模生体分子システムのシミュレーションを実施するためには、粗視化とマルチスケール化が必須であることから京都大学・高田を新たにメンバーに追加し、名古屋大学・篠田や理化学研究所・杉田と連携する

さらに、高度化研究と並行して、生命科学の研究に用いられて来た最先端のアルゴリズム手法を創薬の現場で使いやすい形で活用できるように発展させる。具体的には、拡張アンサンブル法の改良によるタンパク質薬剤複合体構造予測法、熱力学理論を用いた自由エネルギー評価法、多数の独立な MD 計算を組み合わせた長時間ダイナミクス法、粗視化モデル、QM/MM 自由エネルギー法などの開発と生体分子

系でのテスト評価を行う。また、創薬や生命科学に必要な手法は上に挙げた手法に限定されないため、調査研究フェーズにアカデミアと産業界からの要望や意見も募り、ポスト「京」で行う MD 計算に必要なアルゴリズムを多く開発する。

- ① ポスト「京」での MD 高度化（理化学研究所 杉田 有治、横浜市立大学 池口満徳（サブ課題 B 実施者））分子動力学アプリケーション「GENESIS」をポスト「京」において高度化する。GENESIS は戦略分野 1 では大規模な生体分子システム（例えば、分子混雑環境を考慮した 1 億原子系）の MD 計算において高い並列性能を示した。ポスト「京」では創薬応用 MD に適した Capacity computing を効率良く実施するための高度化を行い、Capacity computing を実施した場合に「京」での総演算量と比較し数十倍から百倍多くの創薬応用 MD 計算を実現する。また、Capability computing としても世界最大級の MD 計算を実現するための並列化の効率アップを図る。
- ② 拡張アンサンブル法（理化学研究所 杉田 有治）レプリカ交換分子動力学法（REMD）などの拡張アンサンブル法は従来、タンパク質折れ畳みや凝集機構の解明に用いられてきた。本課題では、創薬応用 MD に利用しやすい形に改良を加え、タンパク質と基質の結合自由エネルギーを高精度に求める。より具体的には、基質が結合することでタンパク質の構造や揺らぎが変化する可能性がありうるが、そのようなタンパク質自身の変化と基質結合の両方の効果を含む拡張アンサンブル計算法を開発する。この手法はコストを下げるためというよりは、ポスト「京」などの利用した大規模長時間の計算を実施した場合に、従来はできなかった計算を実現するものである。
- ③ 長時間ダイナミクス法（東京大学 北尾 彰朗）多数の独立な MD 計算を統合化することによって長時間ダイナミクスを実現する方法論の開発を行う。Markov State Model では将来の状態は現在の状態にのみ依存すると仮定する。この仮定に基づき、異なる初期構造からスタートする多数の独立な MD 計算から異なる状態間の遷移確率を求めることができる。全ての状態変化についての遷移確率を求めることができれば、自由エネルギー表面の解析ができたことと同等の情報が得られたことになり、長時間ダイナミクスを解析できる。この手法を標的タンパク質に応用し、機能的長時間運動や薬剤結合が生じた場合のタンパク質構造変化を予測可能にする。この手法も拡張アンサンブル法と同様で、コストを下げるというよりも大規模計算によって従来不可能であった計算を実現するためのものである。
- ④ ハイブリッド QM/MM 自由エネルギー法（京都大学 林 重彦）ハイブリッド法による薬剤結合自由エネルギー計算法を確立するとともに、創薬応用として量子的相互作用が重要となる薬剤結合に対する計算の精密化をはかる。この手法では、特に薬剤結合部位におけるタンパク質と薬剤の相互作用エネルギーを、第一原理量子化学計算を用いて計算することで、古典分子力場に含まれるモデルパラメータに依存しない正確な予測が可能になる。量子化学計算は粒子数が増加した場合に急激に演算コストが上昇するが、多くの薬剤候補化合物との相互作用に適用可能にするための計算コストの減少を図る必要がある。
- ⑤ 自由エネルギー計算法（横浜市立大学 池口満徳（サブ課題 B 実施者））低分子化合物と標的タンパク質との結合自由エネルギーの各種計算手法について、精度と計算速度の両面から性能評価する。この手法は従来の結合自由エネルギーと異なり、結合状態と非結合状態の間に中間状態を挟むことをせず、結合自由エネルギー予測を行う手法の開発を行う。創薬応用計算においては、多段階のスクリーニングを行うことで標的タンパク質に対して高い結合能を示す薬剤候補化合物を探索する

必要があるが、ここで開発する自由エネルギー計算法を用いることで、従来法と比較して計算コストが少なくなるため、より多くの薬剤候補化合物の探索が可能になる。これにより、サブ課題 C が開発する統合システムに実装する方法を選定するとともに、それらのポスト「京」での実装を行う。

- ⑥ 粗視化モデリング (京都大学 高田 彰二、名古屋大学 篠田 渉 (サブ課題 B 実施者)) 大規模生体分子複合系の構造予測のための粗視化技術および、粗視化・全原子のハイブリッド系の分子シミュレーション技術を開発することで、超大規模生体分子システムのシミュレーションを実現する。粗視化モデリングを行うことで、全原子 MD では到達できない長時間の生命現象に迫ることができる。

(3) 目標・期待される成果

「京」で行われた創薬応用 MD と比較して、Capacity 計算の総量としてポスト「京」では数十倍から百倍程度の総演算を行うことが可能な基盤を、コデザインを通して開発する。それによって様々な計算アルゴリズムを用いた小・中規模のタンパク質・薬剤候補化合物の結合自由エネルギー計算や動的制御の計算が、ポスト「京」を用いて大量に、かつ、高速に計算可能になる。さらに「京」では不可能であった大規模な生体分子システムでの更なる長時間計算を実現するために、ポスト「京」での Capability 計算を最適化する。これにより、「京」ではサブマイクロ秒であった計算時間を超大規模系に対してもマイクロ秒から数十マイクロ秒に延長し、生体超分子複合体の分子ダイナミクスが関わる重要な生命科学の課題を解決する。さらには、将来の創薬応用のシーズを探り、実験家と連携して計算で得られた知見を検証する。

アウトプット成果 (平成 29 年度終了時) : ポスト「京」での MD 高度化では、ポスト「京」では数十倍から百倍以上の Capacity 計算を可能にすることを想定し、「京」等において単独の MD 計算として数倍の高速化を図る (具体的な数字はポスト「京」の性能設計に関わるため記述できない)。拡張アンサンブル法と長時間ダイナミクス法の開発においては、タンパク質構造変化と薬剤結合の両方を考慮した新しい手法の開発を行う。ハイブリッド QM/MM 自由エネルギー法や自由エネルギー計算法では、従来用いられている計算手法からコストダウンを図り、数倍以上少ない演算量での計算を可能とする。大規模システムの計算効率アップのためには、時間積分法の改善や長距離相互作用計算におけるパラメタの検討などを行い、同等の計算精度を維持しながら数倍の高速化を実現する。また、粗視化モデリングにおいても並列化技術を検討して、より多くの CPU を用いた並列化を実現し、大規模でかつ長時間ダイナミクスを実現し、生命現象の解明に繋げる。

アウトプット成果 (本格実施フェーズ終了時) : 平成 29 年度終了時の研究開発を発展させ、ポスト「京」に最適化された MD 計算エンジンを有する GENESIS 新バージョンに対して、拡張アンサンブル法、長時間ダイナミクス法、QM/MM 自由エネルギー法、自由エネルギー計算法などを導入する。これにより、アカデミアのみならず製薬企業等の研究者が、創薬応用 MD として様々なアルゴリズムを用いてポスト「京」上で研究開発を実施できるようになる。さらに、このソフトウェアを、サブ課題 C で開発する創薬ビッグデータ統合システムに組み込み、課題実施者のみならず産業界にも公開する。また、超大規模生体分子システムのシミュレーションを実現するためのマルチスケール (全原子および粗視化) MD 計算を、ポスト「京」のアーキテクチャーに最適化する。

アウトプット成果 (ポスト「京」運用開始 5 年後): 拡張アンサンブル法や長時間ダイナミクスによって、従来のシミュレーションでは不可能であった薬剤結合ポーズの予測などを実現し、薬剤候補化合物の結合自由エネルギーの予測を向上させる。一方で、QM/MM 自由エネルギー法や自由エネルギー計算のコストを減少させることによって、ポスト「京」を用いて、「京」時代と比較して数十倍から百倍多くのスクリーニングを実現し、より信頼性の高い In silico 創薬を実現する。また、手法のみならず創薬応用に関する学術的な論文、例えば、重要なタンパク質と薬剤候補化合物の自由エネルギーなどに関する新しい知見を報告する。

アウトカム効果 (ポスト「京」運用開始 5 年後): 開発された手法の信頼性がアウトプットとして学術論文などで示されているため、アカデミアのみならず製薬企業なども GENESIS を用いた創薬応用 MD 計算を実施する。これにより、製薬企業と重点課題 1 実施機関との連携研究が進展し始める。

アウトカム成果 (ポスト「京」運用開始 10 年後): 5 年間に行われた大規模計算の結果を踏まえて、手法の練り直しと更なる高度化を図り、次世代創薬計算のスタンダードとしての位置付けを確立する。すなわち、国内のアカデミアや製薬企業のみならず、海外の研究者や企業も創薬応用計算に、重点課題 1 で開発した手法を用いるようになる。

(4) 「京」でできていること、ポスト「京」でなければならないこと

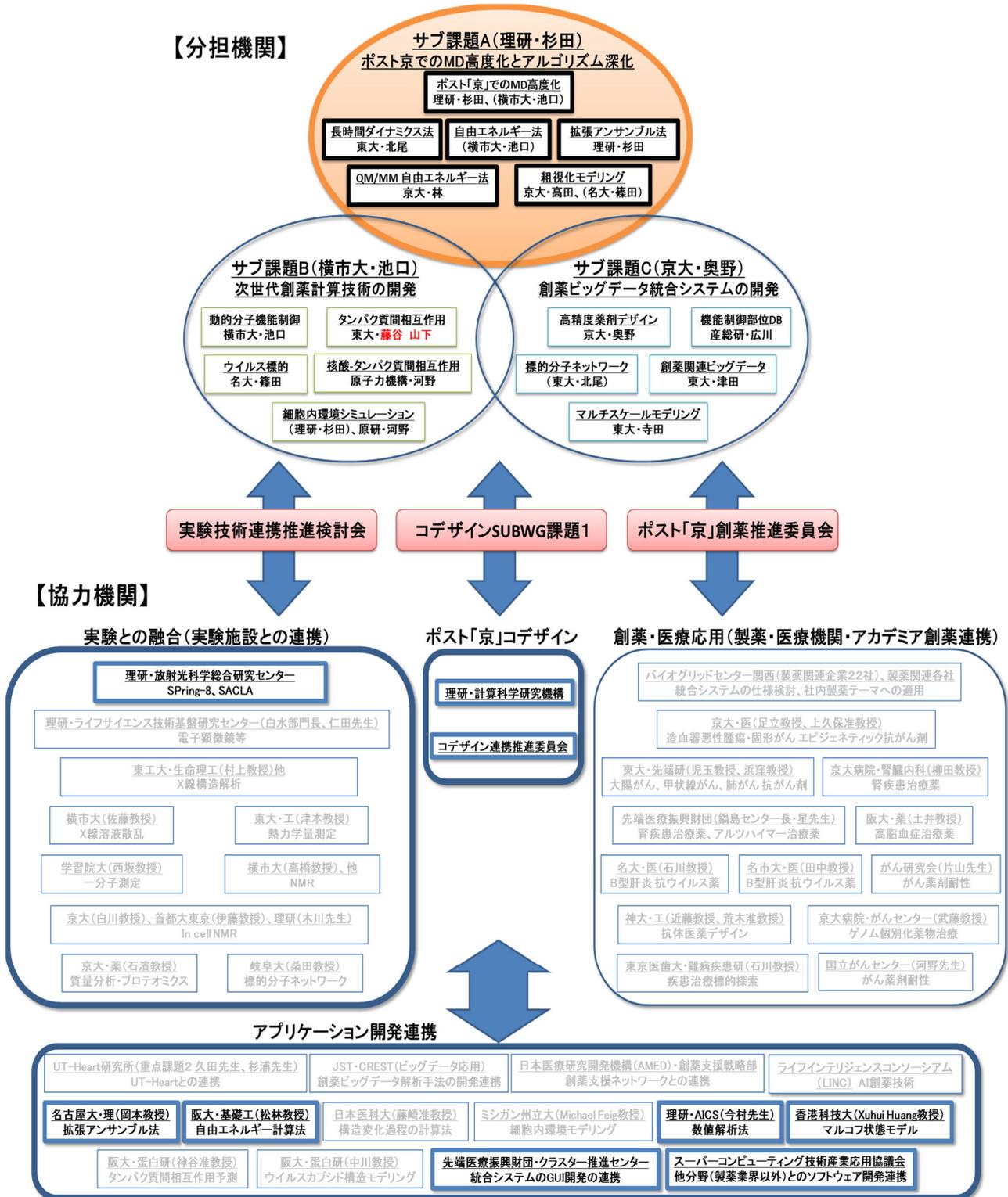
「京」を大規模に利用することで藤谷（東京大学）らは MP-CAFEE 法という MD 計算を用いた自由エネルギー計算でタンパク質と薬剤の相互作用が高精度に評価できることを示した（戦略分野 1）。また、杉田（理化学研究所：戦略分野 1）、岡崎（名古屋大学：戦略分野 2）らは、従来は不可能であった数千万から数億原子を含む巨大な生体系（細胞内分子混雑系、ウイルスなど）の全原子 MD 計算を実現した。しかし、いずれの計算においても「京」上での MD 計算のパフォーマンス限界により計算可能な時間が短く（数百ナノ秒程度）、タンパク質と薬剤の正確な結合構造が実験的に不明な場合などの結合自由エネルギーや、巨大な生体系における長時間の分子運動の予測などが困難であった。

ポスト「京」で初めてできると期待されるのは、「コデザイン」による MD 計算自身の高速化であり、タンパク質等が働く時間スケール（マイクロ秒以上）でのシミュレーションや創薬応用計算が可能になる。また、本重点課題の目的である創薬応用計算を行うために必要な要素を検討した結果として、長時間ダイナミクスによるタンパク質と薬剤の複合体構造の正確な予測、効率的な自由エネルギー計算に基づく低コスト・高速な薬剤候補化合物のスクリーニング、分子力場に含まれるパラメタ依存性に起因する問題を解決するために QM/MM 自由エネルギー計算による正確な相互作用の検討などが実現される。ポスト「京」の利用に向けては、「コデザイン」によって高速化された MD ソフトウェアである GENESIS をベースに新規アルゴリズムを導入することで、高速化と高機能化の両方を実現することができる。課題 C で開発する統合システムに、高速・高機能分子動力学法を導入することによって、信頼性の高い創薬応用 MD を実現するプラットフォームを開発する。

(5) 実施体制

サブ課題責任者・杉田（理化学研究所）の指揮のもと、4 名の当該サブ課題実施者との密接な連携を図りながら研究を遂行する。各研究テーマの進捗管理は、3 ヶ月ごとサブ課題研究進捗会議で確認する。他のサブ課題との連携は、執行部・運営会議を通して促進していく。特に、各種計算アルゴリズム

ムや創薬計算手法の選定や評価はサブ課題間の共同で実施する。協力機関は、下図に示すとおり、アプリケーション開発の連携を主旨とした研究機関等のグループ、ポスト「京」とのコデザイン連携を図るグループ、先端機器計測など実験との融合を主旨とした実験施設・研究機関のグループの3つのグループから成る。詳細の連携内容はリストの通りである。



*当該サブ課題の協力機関は太線で囲ったもののみである。

	協力機関などの連携先	連携内容
実験との融合 (実験施設との連携)	理化学研究所 放射光科学総合研究センター	SPring-8やSACLAなどの放射光施設を用いた生体高分子および生体超分子複合体の立体構造解析を行う。
アプリケーション開発連携	大阪大学 基礎工学研究科	エネルギー表示による自由エネルギー計算法を用いた手法の開発と応用計算を、松林伸幸教授らと連携する。
	理化学研究所 計算科学研究機構	分子動力学計算に必要な数値解析手法の高速化に関する知見を得るために、大規模並列数値計算技術研究チームの今村俊幸チームリーダーと連携する。
	名古屋大学 大学院理学研究科	拡張アンサンブル法による自由エネルギー計算法の開発と応用計算に関する議論を行い、岡本祐幸教授と連携する。
	Hong Kong University of Science and Technology	Markov State Modelを用いた長時間ダイナミクスの手法と応用計算について、Xufei Huang教授と連携する。
	先端医療振興財団 クラスター事業推進課	創薬ビッグデータ統合システムのCUI, GUI開発において岩田浩明主任研究員と連携する。
	スーパーコンピューティング技術産業応用協議会	製薬業界以外の産業界とのソフトウェア開発連携に関する情報交換を行う。

2-2. サブ課題B. 次世代創薬計算技術の開発

(1) 目的・意義

現在、広く用いられている創薬アプローチは分子標的創薬と呼ばれ、単一の標的タンパク質の阻害に主眼をおいたアプローチであり、これにより多くの医薬品が開発されてきた。しかしながら、製薬業界では、この十数年、新薬の承認数が低迷しており、医薬品開発が容易なターゲット疾患のほとんどは開発し尽され、残された疾患はこれまでの創薬アプローチでは医薬品の創出が困難なものばかりであると考えられている。そこで本サブ課題では、ポスト「京」の演算能力を最大限に活かす分子シミュレーション技術を開発し、生体分子システムの動的な機能発現や、巨大な生体分子システムの機能解析に資する新たな計算生命科学を開拓することで、次世代の創薬アプローチを切り拓く計算技術の開発を目指す。また、それらの高度な計算技術を生命科学における様々な実験計測手法と連携して新しい知見をもたらす解析技術へと発展させる。サブ課題Aで開発するポスト「京」に最適化したMD計算ソフトを高度に活用し、サブ課題Cで開発する統合システムへ組み込むべく、次世代創薬におけるHPC計算技術を開発する。

具体的には、ポスト「京」を活用した創薬研究に必要な標的タンパク質の動的機能制御、タンパク質間相互作用制御、ウイルス環境標的、細胞環境標的、核酸-タンパク質相互作用制御のための計算技術の開発を目指す。将来の創薬研究を見越したチャレンジングなテーマである動的機能制御や、巨大な生体システム・ネットワーク系（多数のタンパク質および核酸等からなるウイルス環境、細胞内環境や核内環境など）を対象とした計算を実施する。これらのシミュレーションは、ポスト「京」において初めて達成できる計算機実験であり、新たな構造生命科学の開拓につながることを期待される。また創薬応用においても、現在の主流の創薬手法である分子標的創薬の限界（単独の創薬標的タンパク質の枯渇）を打破する次世代の創薬手法であるエピジェネティック創薬やシステム創薬につながり得るものと期待できる。このように、ポスト「京」によるシミュレーションは、産業面および学術面の両面において国際的に優位性のある革新的でチャレンジングなテーマである。

(2) 実施内容

ポスト「京」を活用した次世代創薬研究の基盤となる技術開発を目標とする。現在の計算創薬では困難な、標的タンパク質における動的な構造変化の制御、タンパク質間相互作用の制御、さらには巨大分子系である、ウイルス全体、細胞環境、核酸-タンパク質複合体を標的とする計算創薬技術を開発する。また、ポスト「京」によって実現する高度な計算技術と生命科学における様々な実験計測技術との融合

手法の開発を行い、新たな構造生命科学の開拓を目指す。

① 動的分子機能制御（横浜市立大学 池口 満徳）

アロステリック制御やアゴニスト・アンタゴニスト作用など、受容体タンパク質に結合し構造変化を誘起する薬剤を合理的に設計することは、製薬業界でのニーズが非常に高いが、これまでの創薬技術では非常に困難な課題であった。そのため、ポスト「京」を用いた超長時間アンサンブル計算により、標的タンパク質の動的性質を理解し、標的タンパク質の構造変化を制御する薬剤の創出のための基盤技術開発を目指す。

② タンパク質間相互作用制御（東京大学 藤谷 秀章（2018.12月まで）山下 雄史（2019.1月以降））

タンパク質間相互作用を制御する薬剤や抗体医薬などのバイオ医薬品は製薬業界でも非常に期待されている。しかしながら、タンパク質の立体構造は低分子と比較して格段に複雑であり、その立体構造の柔軟性も考慮するためには、大規模な計算が不可欠であるため、ポスト「京」を高度に活用したタンパク質間相互作用解析の計算技術を開発する。

③ ウイルス標的創薬計算技術（名古屋大学 篠田 渉）

ウイルスによる感染症を対象とした創薬計算技術の開発を行う。巨大分子系であるウイルスカプシドをモデリングし、ポスト「京」によるシミュレーションを実施することでウイルスカプシド存在下の環境特性の知見を獲得し、そこでの薬剤動態や結合などを評価する技術開発を行う。

④ 細胞環境標的創薬計算技術（理化学研究所 杉田 有治）

リボゾームなどの生体超分子が複数存在し、複雑に混み合った細胞環境下のシミュレーションを行い、そのような環境での薬剤動態や結合などを評価する技術開発を行う。

⑤ 核酸-タンパク質相互作用制御（日本原子力研究開発機構 河野 秀俊）

核内タンパク質やRNA結合タンパク質など、核酸結合タンパク質も重要な創薬ターゲットと考えられている。しかしながら、核酸もタンパク質と同様に非常に複雑な生体高分子であるため、ポスト「京」によるシミュレーションを実施することで、核酸-タンパク質相互作用の詳細解析の計算技術を開発する。

（3）目標・期待される成果

ポスト「京」向けに高度に最適化されたMD計算ソフトを創薬へ繋げる計算技術を実現する。

アウトプット成果（平成29年度終了時）：動的機能分子制御では、ポスト「京」での計算実施時に標的タンパク質の機能制御に関わるミリ秒以上の動的構造解析を可能にすることを想定し、「京」等において10マイクロ秒以上の動的構造解析を可能とする計算システムを開発する。タンパク質間相互作用制御では、ポスト「京」での100万原子数を超えるタンパク質複合体計算の精度を向上させるための力場を開発する。ウイルス標的創薬計算技術では、サブミクロンサイズに迫るウイルスカプシド系の薬剤相互作用解析を想定し、「京」等において、直径30nm程度のB型肝炎ウイルスカプシド系の分子モデリング（全原子及び粗視化モデル）及び構造解析を可能とする計算システムを開発する。細胞環境標的創薬計算技術では、細胞環境を考慮した大規模MD計算を、アプリケーションの高速化によって、「京」等においてマイクロ秒以上のシミュレーションを可能とする計算システムを開発する。核酸-タンパク質相互作用制御では、ポスト「京」での計算実施時に、転写因子など塩基配列特

異的に結合する DNA 結合タンパク質の認識機構について多次元自由エネルギープロファイル解析を可能とすることを想定し、「京」等において 2 次元以上の多次元自由エネルギープロファイル解析を可能とする計算システムを開発する。

アウトプット成果 (本格実施フェーズ終了時) : 動的機能分子制御では、平成 29 年度終了時の成果を進展させ、「京」等において、低分子化合物結合などの複数の条件において、それぞれ 10 マイクロ秒以上の動的構造の変化を解析可能とする計算システムを開発する。タンパク質間相互作用制御では、「京」等において 100 万原子数を超える系で短時間(10 マイクロ秒以下)シミュレーションによる低分子の活性予測を可能にする。ウイルス標的創薬計算技術では、「京」等において、直径 30nm 程度の B 型肝炎ウイルスカプシドへの低分子化合物の取り込み機構(結合、透過)について、自由エネルギー評価を可能とする計算システムを開発する。細胞環境標的創薬技術では、細胞内でのタンパク質の安定性や基質認識を理解するために必要な自由エネルギー計算法を開発して、高速化した分子動力学エンジンをを用いて、定量的な予測を行う方法論を確立する。核酸-タンパク質間相互作用制御では、平成 29 年度終了時の成果を進展させ、翻訳後修飾を受けた DNA 結合タンパク質の認識機構について 2 次元以上の多次元自由エネルギープロファイル解析を可能とする計算システムを開発する。

アウトプット成果 (ポスト「京」運用開始 5 年後) : 動的機能分子制御では、ポスト「京」において、標的タンパク質のミリ秒以上の動的構造解析を実現し、低分子化合物結合等による動的構造の変化を解析する。タンパク質間相互作用制御では、ミリ秒オーダーで起きる構造変化を持ったタンパク質複合体に作用する低分子化合物の活性予測を可能にする。ウイルス標的創薬計算技術では、サブミクロンサイズのウイルスカプシド系について、分子モデリング(全原子及び粗視化)・構造解析を実行し、薬剤との相互作用、送達・作用機構を解析する。細胞環境標的創薬技術では、細胞内環境を考慮した計算において、一つのタンパク質における基質の競合や、一つの基質に対する複数のタンパク質の競合などについての自由エネルギー計算を実現し、創薬応用に必要な知見を得る。核酸-タンパク質間相互作用制御では、ポスト「京」において、多次元自由エネルギープロファイル解析を行い、翻訳後修飾の DNA 結合に対する影響を明らかにする。

アウトカム成果 (ポスト「京」運用開始 5 年後) : 製薬企業等、実業界などでの応用事例を積み上げるとともに、生命科学として、一分子測定、放射光、XFEL など先端計測実験と連携し、生命科学分野に新しい知見と信頼性の高い構造情報をもたらす。

アウトカム成果 (ポスト「京」運用開始 10 年後) : 実業界からのフィードバックによる創薬計算技術の高度化を実施しつつ、産学連携で、大規模 HPC を用いた画期的新薬開発に貢献する。

(4) 「京」でできていること、ポスト「京」でなければならないこと

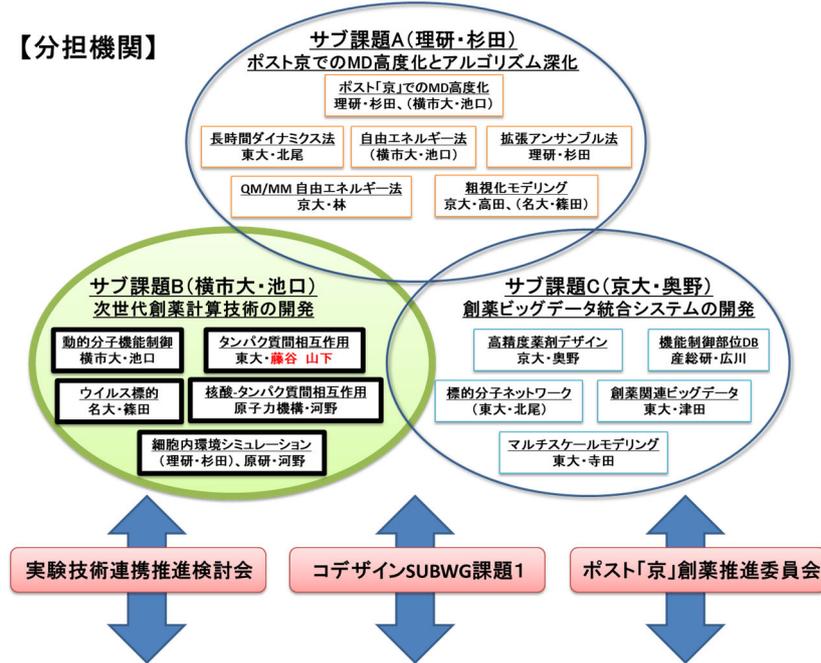
「京」を活用することで、河野(日本原子力研究開発機構)、高田(京都大学)、池口(横浜市大)は MD 計算により DNA 結合タンパク質における動的性質の解析、藤谷(東京大学)は MP-CAFEE 法により標的タンパク質と薬剤候補化合物の結合自由エネルギー計算が高精度に実現できることを示した。さらに、巨大系では、杉田(理化学研究所)、河野(日本原子力研究開発機構)、岡崎(名古屋大学)は、複雑に混み合った細胞環境やヌクレオソームなどの巨大核酸タンパク質複合体、ウイルス全体の計算が可能であることを示した。しかし、現在の「京」を用いたこれらの計算では、それぞれ数百ナノ秒の計算が限界であり、標的タンパク質が大幅な構造変化を示す場合やタンパク質間相互作用、細胞やウイル

スなど巨大系における十分なアンサンブルの実現は困難であった。ポスト「京」では、数十倍～百倍のアプリ性能向上を背景として、これらの難しい対象においても、十分なアンサンブルの収束を実現することを目指す。これら種々の巨大分子系のシミュレーションは、ポスト「京」において初めて達成できる計算機実験であり、新たな構造生命科学の開拓および次世代の創薬手法であるエピジェネティック創薬やシステム創薬につながり得るものと期待できる。

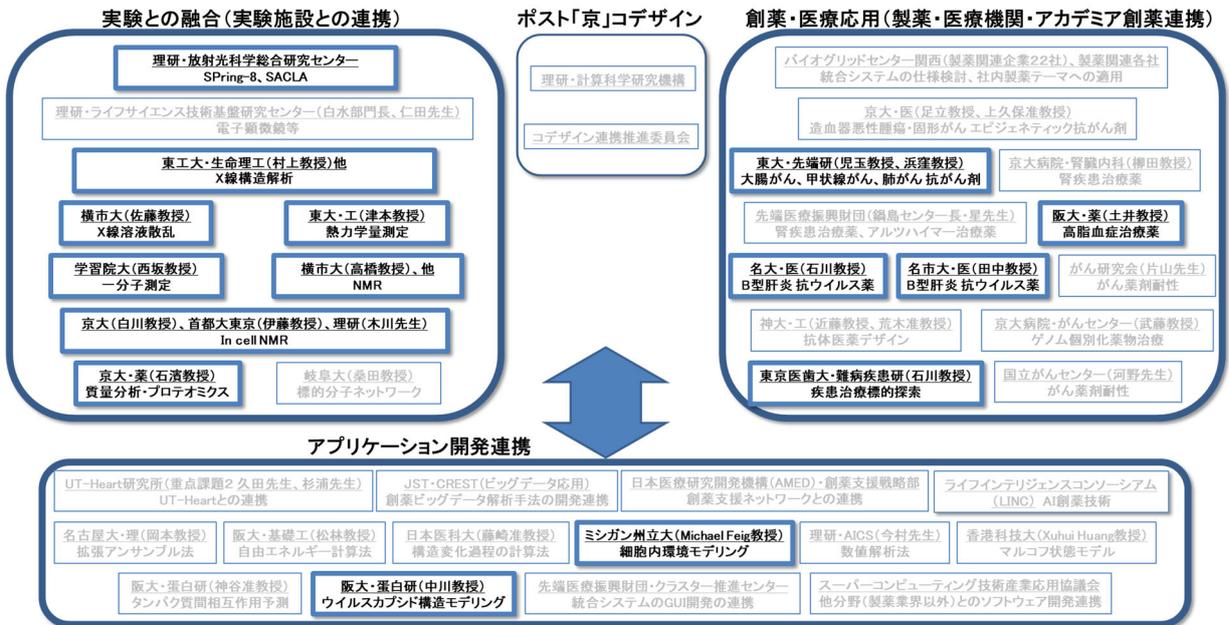
(5) 実施体制

サブ課題責任者・池口（横浜市立大学）の指揮のもと、4名の当該サブ課題実施者との密接な連携を図りながら研究を遂行する。各研究テーマの進捗管理は、3ヶ月ごとサブ課題研究進捗会議で確認する。他のサブ課題との連携は、執行部・運営会議を通して促進していく。特に、各種計算アルゴリズムや創薬計算手法の選定や評価はサブ課題間の共同で実施する。協力機関は、下図に示すとおり、先端機器計測など実験との融合を主旨とした実験施設・研究機関のグループ、創薬・医薬応用を主旨とした製薬会社・医療機関・アカデミア創薬研究機関のグループ、アプリケーション開発の連携を主旨とした研究機関等のグループの3つのグループから成る。詳細の連携内容は次頁リストの通りである。

【分担機関】



【協力機関】



*当該サブ課題の協力機関は太線で囲ったもののみである。

	協力機関などの連携先	連携内容
実験との融合 (実験施設との連携)	理化学研究所 放射光科学総合研究センター	SPring-8やSACLAなどの放射光施設を用いた生体高分子および生体超分子複合体の立体構造解析を行う。
	東京工業大学 生命理工学研究科	多剤排出トランスポーターなどのX線結晶構造解析について、村上聡教授と連携する。
	横浜市立大学	SAXS測定について佐藤衛教授と、NMR測定について高橋栄夫教授と連携する。
	京都大学 工学研究科	細胞環境中の動態のシミュレーション結果を、白川昌宏教授のin cell NMR実験と比較検討する。
	首都大学東京 理工学研究科	細胞環境中の動態のシミュレーション結果を、伊藤隆教授のin cell NMR実験と比較検討する。
	東京大学 工学系研究科	タンパク質間相互作用などの分子結合の熱力学測定について、津本浩平教授と連携する。
	学習院大学 理学部	生体分子の1分子測定について、西坂崇之教授と協力する
	京都大学 薬学研究科	石濱泰教授(製剤機能解析学分野)と連携し、質量分析によるプロテオミクスにおける実験検証の協力を得る。
創薬・医療応用 (製薬・医療機関・アカデミア 創薬との連携)	名古屋市立大学 大学院医学研究科	B型肝炎ウイルスの抗ウイルス剤の開発を田中靖人教授らと共同で行う。
	名古屋大学 大学院医学系研究科	B型肝炎ウイルスの抗ウイルス剤の開発を石川哲也教授らと共同で行う。
	東京大学 先端科学技術研究センター	大腸がん、甲状腺がん、肺がんの抗がん剤開発に関して、実験協力者としては児玉教授、浜窪教授と連携する。
	大阪大学 薬学研究科	高脂血症治療薬の開発に関して、実験協力者として土井健史教授と連携する。
	東京医科歯科大学	ゲノム病理に関する情報提供や疾患標的の探索とシミュレーション結果の検証に関する協力
アプリケーション開発連携	Michigan State University	細胞環境のモデリングについて、Michael Feig教授と協力する。
	大阪大学 蛋白質研究所	ウイルスカプシド構造の分子モデリングについて中川敦史教授らに助言を得る。

2-3. サブ課題C. 創薬ビッグデータ統合システムの開発

(1) 目的・意義

医薬品開発プロセスにおいて、計算技術の利活用が最も期待されている工程は、「化合物ライブラリーから疾患原因タンパク質に結合し制御する化合物を探索するスクリーニング段階」と、「スクリーニングで特定した候補化合物の活性向上と副作用回避を目指して化学構造を変換するリード最適化段階」である。製薬現場では市販のソフトなどを用いて候補化合物の探索や分子デザインがなされているが、予測精度の低さや評価できる化合物数・標的タンパク質数の限界などの課題が未だ残されている。本サブ課題では、これらの課題を克服するために、サブ課題AとBで開発される要素計算技術を創薬計算フロー（スクリーニングからリード最適化）にそって連結し、ポスト「京」の演算能力を活かし医薬品の候補化合物の探索や分子デザインを実現する創薬ビッグデータ統合システムの開発を行う。

具体的には、ポスト「京」のスケールメリットを最大に活かし、多数の創薬関連タンパク質（副作用関連タンパク質を含む）やそれらの分子ネットワークと化合物ライブラリーとの膨大な組合せから最適解を高速に計算し、ベストな薬剤候補化合物を自動でデザインできる創薬計算基盤を構築し、製薬会社に提供することを目指す。また、サブ課題AとBで開発される各種創薬計算手法を協力機関が有する具体的な創薬ターゲットに適用し、性能評価も行う。

(2) 実施内容

サブ課題Aとサブ課題Bで開発されるMD計算法および各種創薬計算手法を創薬計算フローに沿って連結した統合システムを開発することで、大規模な候補化合物と複数の創薬標的・副作用関連タンパク質の莫大な組み合わせからベストな薬剤候補化合物を高精度かつ超高速に予測できる創薬計算基盤を構築する。またサブ課題Aとサブ課題Bで開発される各種創薬計算手法を、協力機関と連携し具体的な創薬ターゲットに適用し、性能評価を行う。現時点で決定している創薬ターゲット、造血器悪性腫瘍・固形腫瘍におけるエピジェネティック抗癌剤の開発（京都大学医学部 足立教授・上久保准教授との連携）、腎疾患治療薬の開発（京都大学病院腎臓内科 柳田教授、先端医療振興財団 鍋島先生との連携）、アルツハイマー治療薬の開発（先端医療振興財団 鍋島先生・星先生との連携）、抗癌剤の薬剤耐性予測（京都大学病院がんセンター 武藤教授、がん研究会 片山先生、国立がんセンター 河野先生、京都大学薬学部 石濱教授との連携）、ドラッグリポジショニング、GPCR機能制御化合物探索等である。また、適宜、製薬会社と共同研究を結び、各種創薬計算手法を製薬会社の実践的創薬テーマに応用する。さらに、奥野チームと津田チームが中心となり各種創薬シミュレーション技術とAI技術の融合を図るとともに、AI創薬コンソーシアム（製薬企業、IT企業、アカデミア90団体以上参加）であるLINCと連携し、ビッグデータ統合システムにAI創薬技術を導入することで世界に類を見ない創薬計算基盤を実現する。具体的には、最適化アルゴリズムによるMD回数の劇的削減、モンテカルロ木探索による効率的化合物生成、AIスコア導入によるドッキングスコアの高精度化、AIによる力場補正、長時間MDトラジェクトリーのAI分類などのAI技術開発を行い、分子シミュレーションアプローチの限界を補完することを目指す。

① 高精度薬剤デザイン（京都大学 奥野 恭史）

新規化合物（バーチャル化合物）を計算機上でデザインする化合物生成手法やHit to Leadを行うための化学構造変換手法の開発を行う。新規化学構造を計算機上で自動発生させる計算法はデ

ノボデザイン法と呼ばれているが、既存の方法論はいずれも薬理活性を有する化合物を高精度にデザインする性能に至っていない。そこで、サブ課題 A と B で開発される要素計算技術を適用することで、高精度に薬剤候補化合物をデザインできる新規なデノボデザイン法を開発する。また、サブ課題 A と B で開発された各種 MD 計算法および創薬計算手法を統合化し、化合物のスクリーニングからリード最適化において製薬現場利用ができる創薬計算フローを構築する。さらに、サブ課題 A と B で開発される各種創薬計算手法を協力機関が有する具体的な創薬ターゲットに適用し、性能評価も行う。

② 機能制御部位データベース (産業技術総合研究所 広川 貴次)

創薬標的タンパク・副作用関連タンパクに関するホモロジーモデリング法、マルチコンフォメーション生成法、機能制御部位推定法を開発する。さらに、これらの計算法を統合化し、タンパク質標的の部位データベースを構築する。これまでの創薬計算では X 線結晶構造の活性部位ポケットに結合する化合物をデザインすることが主流であったが、当該研究では、MD 計算で生成される動的なタンパク質構造を対象にして機能発現を制御する部位を推定する計算法を開発し、それによって推定された創薬関連タンパク質の機能制御部位をデータベース化する。

③ 創薬関連ビッグデータ (東京大学 津田 宏治)

サブ課題 A と B ではポスト「京」の演算能力を活かしタンパク質の長時間 MD を可能にするが、このことは同時に膨大なタンパク質の動的構造を生成することを意味する。そのため、多数の創薬関連タンパク質について、長時間 MD によって生成される複数の標的部位 (タンパク質機能制御部位) と大規模な化合物数との組み合わせが膨大になることから、当該研究では、膨大なタンパク質構造と化合物の組合せから薬剤候補化合物を高速に探索する組合せ最適化アルゴリズムの開発を行う。具体的には、ベイズ最適化や多腕バンディット等の最適化アルゴリズムを適用し、結合自由エネルギーを簡易に算出する MM-PBSA を評価関数として、最小の MD 計算時間で最小自由エネルギー結合構造を推定する方法論を開発し、MD 計算時間の従来比 8 分の 1 の削減を目指す。

④ 標的分子ネットワーク (東京大学 北尾 彰朗 (サブ課題 A 実施者))

生体系において単一で機能しているタンパク質はほとんど無く、複数のタンパク質が互いに相互作用することで生体内での機能を発現している。従来型の創薬手法は分子標的創薬と呼ばれ単一のタンパク質を標的とする創薬が主流であり、複数のタンパク質からなるタンパク質複合体を標的として創薬を合理的に行うことは非常に困難であると考えられている。当該研究では、タンパク質複合体を対象にした創薬を目指した要素計算技術として、分子ネットワークやタンパク質複合体の立体構造予測と作用機構解明を行う。さらに、複合体立体構造と作用機構から最適な創薬ターゲットを選択する技術を開発する。

⑤ 生体系マルチスケールモデリング (東京大学 寺田 透)

心臓シミュレータ UT-Heart では、膜電位依存的なイオンチャネルの動作モデルとして、閉状態と開状態間を遷移する Markov state model を採用している。現状では、状態間の遷移速度は、実験に基づいて与えられている。これを分子シミュレーションに基づく動作モデルに置き換えれば、個人ゲノムにおけるイオンチャネル配列の変異 (一塩基多型等) が致命的な不整脈を引き起こすメカニズムを明らかにすることが可能になると期待される。当該研究では、心筋イオンチャネルのモデルとして、立体構造が決定している別の膜電位依存性イオンチャネルを用いて、方法論の

開発と精度の評価を行う。ここでは長時間 MD 等を用いて、膜電位依存的なイオンチャネルの構造変化を予測し、さらにマルチコピー・マルチスケール MD 法に基づいて、その間の構造変化パスウェイとこれに沿った構造変化速度を予測する。さらに、ホモロジーモデリング法を用いて心筋イオンチャネルの粗視化モデルを構築し、これを用いて不整脈を誘発することが知られている薬剤がイオンチャネルに結合して阻害する過程を再現する粗視化 MD 法を開発する。

(3) 目標・期待される成果

本サブ課題でポスト「京」を基幹とする世界唯一の創薬ビッグデータ統合システムを開発するとともに、当該システムを製薬会社が現場利用できる運用体制を構築し、医薬品開発の加速化を目指す。

アウトプット成果（平成 29 年度終了時）：

創薬計算フローに実装する要素計算技術の調査、評価等を行い創薬ビッグデータ統合システムの設計を完了する。また、バーチャル化合物の化合物生成手法、標的タンパク質の機能制御部位推定法、タンパク質複合体の立体構造予測法、タンパク質-化合物相互作用予測のための最適化アルゴリズムを開発する。さらに、心筋イオンチャネルの膜電位依存的な構造変化のシミュレーションを実現する。

アウトプット成果（本格実施フェーズ終了時）：

3 つのサブ課題で開発されるすべての要素計算技術を実装した創薬ビッグデータ統合システムの初版を完成させる。初版の統合システムでは、標的タンパク質の入力に対し、ポスト「京」を中核とした一連の創薬計算フローが実行され、最適な薬剤候補化合物が自動デザインできる仕様を目指す。

アウトプット成果（ポスト「京」運用開始 5 年後）：具体的な創薬テーマに創薬ビッグデータ統合システムを適用することで当該システムの改良と精密化をはかるとともに、製薬会社での現場利用を促進し、製薬会社で定常的にシステム利用がなされる状態を目指す。

アウトカム成果（ポスト「京」運用開始 5 年後）：創薬ビッグデータ統合システムを通じて、アカデミア創薬および製薬各社の創薬テーマのリード化合物の創出に成功していることを目標とする。

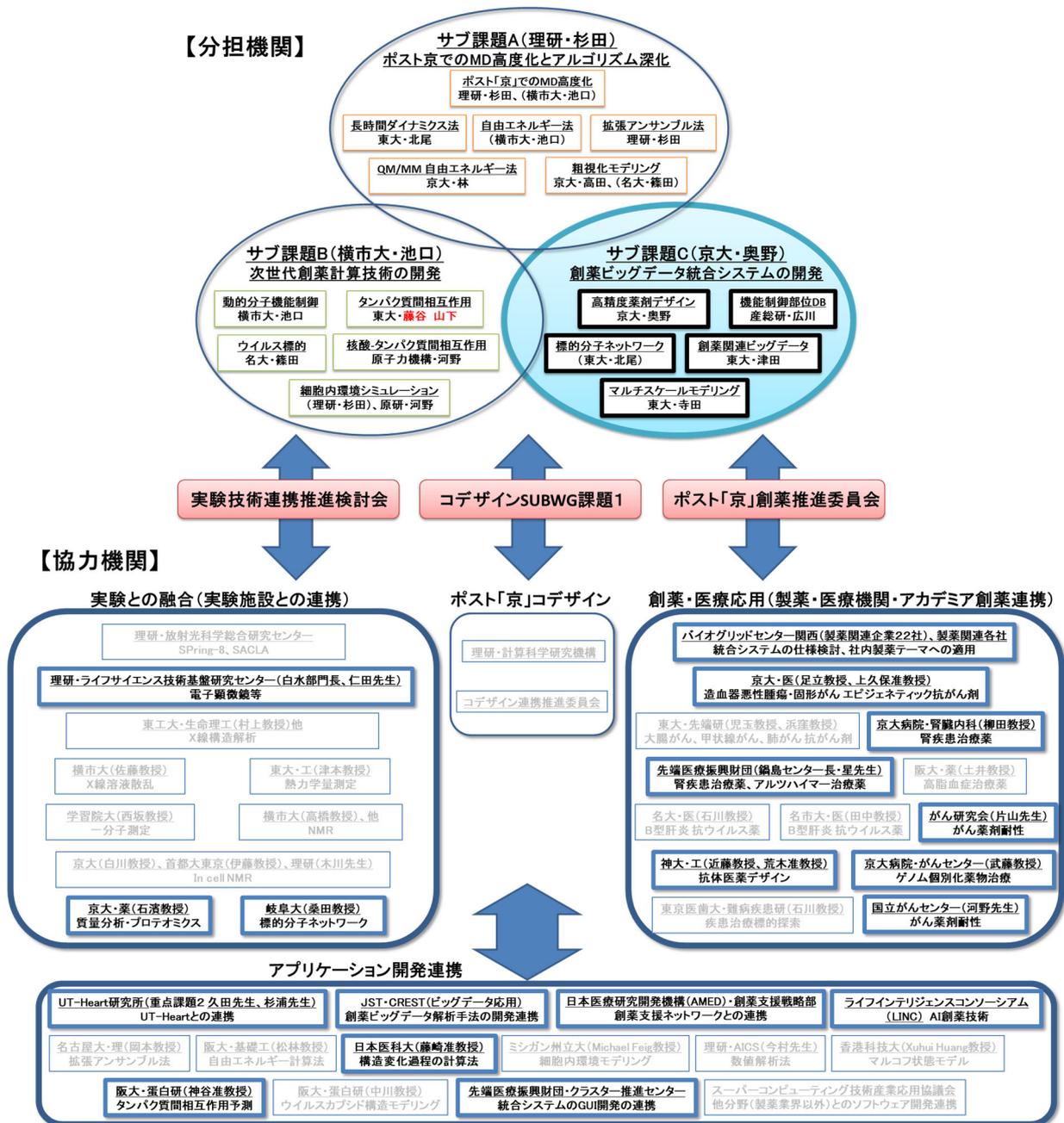
アウトカム成果（ポスト「京」運用開始 10 年後）：創薬ビッグデータ統合システムを通じて、前臨床・臨床フェーズに進む候補化合物に創出に成功していることを目指す。

(4) 「京」でできていること、ポスト「京」でなければできないこと

「京」では、タンパク質と化合物との高精度な結合自由エネルギー計算ができる MP-CAFEE 法の実時間実行が可能となった。しかし、タンパク質の機能を制御する薬物をデザインするためには、タンパク質などの分子運動を長時間観測する必要があるが、「京」ではマイクロ秒程度の MD 計算しか出来なかった。これに対してポスト「京」ではミリ秒程度の MD 計算が可能になることから、タンパク質の動的制御を指向した薬物デザインが期待できる。また、「京」の全ノードを利用できたとしても、MP-CAFEE 法では 1 週間で 150 化合物程度の計算しかできないが、ポスト「京」では 1 週間で 15,000 種の化合物が計算できることになる。したがって、ポスト「京」によって、化合物ライブラリーと多数の創薬関連タンパク質との大規模な組合せの結合予測が可能となる。さらに、化合物とタンパク質との大規模な組合せの結合計算が可能となることから、「京」では不可能であった実効性のある一連の創薬計算フローの構築がポスト「京」によって初めて実現できる。

(5) 実施体制

サブ課題責任者・奥野（京都大学）の指揮のもと、4名の当該サブ課題実施者との密接な連携を図りながら研究を遂行する。各研究テーマの進捗管理は、3ヶ月ごとサブ課題研究進捗会議で確認する。他のサブ課題との連携は、執行部・運営会議を通して促進していく。特に、各種計算アルゴリズムや創薬計算手法の選定や評価はサブ課題間の共同で実施する。協力機関は、下図に示すとおり、創薬・医薬応用を主旨とした製薬会社・医療機関・アカデミア創薬研究機関のグループ、アプリケーション開発の連携を主旨とした研究機関等のグループ、先端機器計測など実験との融合を主旨とした実験施設・研究機関のグループの3つのグループから成る。詳細の連携内容は次頁リストの通りである。



* 当該サブ課題の協力機関は太線で囲ったもののみである。

	協力機関などの連携先	連携内容
創薬・医療応用 (製薬・医療機関・アカデミア 創薬との連携)	NPOバイオグリッドセンター関西	製薬会社22社、IT会社2社からなる製薬コンソーシアム(当該機関が代表)と連携し、産業界への技術移転をはかる。
	京都大学 医学研究科人間健康科学専攻	造血器悪性腫瘍・固形がんの治療薬、特に、エビジェネティック抗がん剤の開発に関して、実験協力者として足立教授、上久保准教授と連携する。さらに、臨床の視点から創薬ターゲットの選定に関する助言を頂く。
	京都大学医学部附属病院 腎臓内科	腎疾患における標的タンパク質の探索やその治療薬の開発に関して、実験協力者として柳田素子教授(腎臓内科)と連携する。さらに、臨床の視点から創薬ターゲットの選定に関する助言を頂く。
	京都大学医学部附属病院 がんセンター	武藤学教授(がん薬物治療科)より、抗がん剤の薬理作用、創薬、ゲノムを含む医療情報に関して知見提供と、計算結果の実験検証の協力を得る。さらに、臨床の視点から創薬ターゲットの選定に関する助言を頂く。
	先端医療振興財団 先端医療センター研究所	腎疾患やアルツハイマーを治療薬開発に関して、実験協力者として鍋島陽一センター長と星美奈子先生と連携する。
	がん研究会 がん化学療法センター	抗がん剤の薬理作用やゲノム変異による薬剤耐性について、片山量平研究員(基礎研究部)より実験検証の協力を得る。
	国立がん研究センター	抗がん剤の薬理作用やゲノム変異による薬剤耐性について、河野隆志研究分野長(ゲノム生物学研究分野)より実験検証の協力を得る。
実験との融合 (実験施設との連携)	神戸大学 自然科学系先端融合研究環/工学研究科	近藤昭彦教授、荒木通啓准教授より抗体医薬品の分子設計における実験検証と産業界への技術移転の協力を得る。
	理化学研究所 ライフサイエンス技術基盤研究センター	白水美香子部門長、仁田亮上級研究員の協力を得て、電子顕微鏡を用いた生体高分子および生体超分子複合体の立体構造解析を行い、実験との融合を図る
	京都大学 薬学研究科	石濱泰教授(製剤機能解析学分野)より、質量分析によるタンパク質ネットワーク解析における実験検証の協力を得る。
アプリケーション開発連携	岐阜大学 大学院連合創薬医療情報研究科	桑田一夫教授より標的分子ネットワークに関する実験検証の協力および助言を得る。
	JST・CREST「ビッグデータ応用」船津プロジェクト	創薬ビッグデータ解析手法の開発について、船津公人教授(東大・工)、泰地真弘人副センター長(理研QBIC)との連携および共同開発を行う。
	日本医療研究開発機構(AMED) 創薬支援戦略部	創薬支援ネットワークとの連携について中西理部長より助言を得る。
	(株)UT-Heart研究所	久田俊明締役、杉浦清了明締役から分子シミュレーションの心臓シミュレータUT-Heartとの連成法および薬剤を用いた心筋イオンチャネル阻害に関する助言および情報を得る
	先端医療振興財団 クラスタ推進センター	創薬ビッグデータ統合システムのCUI、GUI開発において岩田浩明主任研究員と連携する。
	大阪大学 蛋白質研究所蛋白質解析先端研究センター	タンパク質間相互作用予測や拡張アンサンブル法について、神谷成敏准教授(蛋白質情報科学研究室)の協力を得る。
	日本医科大学	立体構造変化の過程とその速度定数の計算法について、藤崎弘士准教授から助言を得る。
ライフインテリジェンスコンソーシアム	AI創薬コンソーシアム(製薬企業、IT企業、アカデミア90団体以上参加)である当コンソーシアムと連携し、ビッグデータ統合システムにAI創薬技術を導入する。	

3. 採択時の留意事項への対応状況

以下、指摘事項は斜体で記載し、回答を>以下に記載している。

- (1) 課題全体として達成すべき成果を明確にするとともに、その成果実現に向けた定量的・定性的な目標（年間目標及び最終目標）を明確にすること。その際、現在実現されている計算手法からの発展だけではなく、国際的な優位性のある革新的でチャレンジングな目標についても明確にすること。

>課題全体として達成すべき成果：1－（1）、1－（5）に反映済み

>国際的な優位性のある革新的でチャレンジングな目標：1－（2）に反映済み

- (2) 成果の展開先（製薬企業等）の要求を考慮した目標値を明確にすること。その際、インシリコの研究者だけではなく、実験系の研究者の要求も十分に取り込めるようにすること。

>成果の展開先（製薬企業等）の要求を考慮した目標値を明確にすること：各サブ課題の実施内容に反映済み。各サブ課題の研究項目は、H26年度にNPO バイオグリッドセンター関西の製薬コンソーシアム（国内製薬会社（22社））のスパコン創薬のニーズ調査を受けて選定したものである。さらに、当該プロジェクト開始時点より、ポスト「京」重点課題1の開発内容に関する製薬各社への個別ヒアリングを開始し、現在13社の製薬会社の現場ニーズヒアリングを済ませ、これらの要求も十分に反映した研究計画にしている。目標値としては、本研究で開発する全てのアプリを統合化する創薬ビッグデータ統合システムを完成することを設定した。

>実験系の研究者の要求も十分に取り込めるようにすること：1－（6）に反映済み。実験技術連携推進検討会を設置し、実験研究者との密な連携を推進する。

- (3) 研究対象とするターゲットタンパク質を選定する仕組みを明確にするとともに、ターゲットタンパク質の妥当性について助言を得る方法（例えば医療機関を実施体制に加えるなど）を明確にすること。

>1－（6）に反映済み。製薬会社や医療機関からの有識者を中心に構成するポスト「京」創薬推進委員会を設置し、創薬ターゲットの選定や創薬ビッグデータ統合システムの産業化利用などの具体的な創薬・医療応用に関する検討を行う。（構成員には、現在、塩野義製薬・坂田氏、バイオグリッドセンター関西・理事志水氏の内諾済み。京都大学病院関係者にも参画依頼中。）

- (4) 積極的に成功事例を出していくため、特に重点的に成果を創出するテーマをしばって設定すること。

>反映済み。予算（人件費）配分に傾斜をつけることで、重点的に成果を創出するテーマを促進するようにした。具体的なテーマは、MD 高度化、動的機能制御、創薬ビッグデータ統合システム、UT-Heart による心毒性予測である。また、創薬ターゲットについては、現在臨床現場において治療薬が不十分なターゲットを優先することになっている。

- (5) サブ課題Aからサブ課題B、サブ課題Cという成果創出の流れだけではなく、逆向きのフィードバ

ックや、製薬企業・実証実験からのフィードバックを取り込んで、成果の最大化を図れるようにすること。

> 1 - (6)、各サブ課題の実施体制欄に反映済み。毎月の運営会議や年2回の重点課題1ワークショップ（全体会議）サブ課題間・分担機関間での相互連携を推進するとともに、各種計算アルゴリズムや創薬計算手法の選定や評価をサブ課題間の共同で実施する。

(6) 相乗効果が期待される他プロジェクト等と効果的に連携して取り組むとともに、役割分担を明確にすること。

> 1 - (6)、各サブ課題の実施体制欄に反映済み。

(7) 人件費の占める割合が大きいが、本プロジェクトで雇用する研究者（ポスドク等）の人員計画を明確にするとともに、将来のステップアップまで見据えた登用、人材育成の取組の計画を明確にすること。

> 1 - (6)に反映済み。バイオ計算分野振興WGを設置し、人材育成や若手研究者のプロモーションを強力に推進する。

(8) 当該全体提案は、全体提案の方針に沿った形で、手法高度化の相乗効果を見込むものとして、当該部分提案の一部である分子シミュレーションに関する取組部分（分担機関：東京大学大学院農学生命科学研究科・寺田透）を取り込むように調整すること。

> 反映済み。サブ課題Cに東京大学・寺田先生（生体系マルチスケールモデリング）をサブ課題実施者に追加した。

4. 中間評価における指摘事項への対応

以下、指摘事項は斜体で記載し、回答を>以下に記載している。

(1) 課題全体として達成すべき成果を明確にするとともに、その成果実現に向けた定量的・定性的な目標（年間目標及び最終目標）を明確にすること。その際、現在実現されている計算手法からの発展だけではなく、国際的な優位性のある革新的でチャレンジングな目標についても明確にすること。
>計画書4ページから5ページの「アウトプット成果（本格実施フェーズ終了時）」部分および8ページから9ページの「年次計画」部分に加筆した。

(2) 情報科学技術分野における研究開発の論文数、学会発表数は、事業の成果を議論する上で1つの指標となりうるため、分野の特性を考慮の上、論文数、学会発表数の達成目標値を設定すること。

>

論文発表目標数は、1年間あたり25件とする。

学会発表目標数は、1年間あたり70件とする。

(3) 予備計算などを通じて、サイエンス的な目標を明確にすること。その目標に対して、ポスト「京」でいつまでに何をどこまで明らかにすることを目指すのかを明確にすること。その時点でポスト「京」で初めてできる画期的な利活用について具体的に説明すること。

>計画書5ページから6ページの「アウトプット成果（ポスト「京」運用開始5年後）」部分に加筆した。

(4) 成果の論文発表数が少ない。論文を含め成果の発信が望まれる。

>報告書に記載した論文発表数（16件）を改めて集計したところ、本事業開始から現在に至るまでに37件の論文発表をされており、着実に成果発表が出来ているものとする。

(5) 採択時の指摘事項（サブ課題Aからサブ課題B、サブ課題Cという成果創出の流れだけではなく、逆向きのフィードバックや、製薬企業・実証実験からのフィードバックを取り込んで、成果の最大化を図れるようにすること。）への対応が必ずしも十分ではない。サブ課題A、Bの深化とサブ課題Cの創薬出口戦略の間のより有機的な連携とそれらの間のより高い整合性が得られるよう、サブ課題間の連携について再検討すること。

>毎年2回開催するワークショップにおいて、初回は製薬会社へ公開型で開催し、2回目は諮問委員会委員の参加型で行っている。本ワークショップでは参加した製薬企業・実験研究者からポスト「京」の創薬現場利用における有意義な示唆があり、それらをサブ課題全体の次期開発に反映している。今年度は、ワークショップの参加対象企業を、KBDDコンソーシアム（「京」の創薬応用コンソ）を対象にしたが、次年度以降ではLINC（AI創薬コンソ）の企業へと参加対象を広げ、製薬現場の広範囲にわたるニーズを受けた開発を展開する予定である。

さらに、サブ課題全体の共通テーマとしてゲノム医療・ゲノム創薬を応用ターゲットにすることに決

定した。ゲノム医療・ゲノム創薬は、現在世界的に急速に臨床実装、創薬実装が進むテーマであり、本テーマにサブ課題 A、B、C の全員でチャレンジすることでポスト「京」の創薬応用の課題が網羅されると考えられる。

連携を促進するために、サブ課題 C の奥野チームに連携促進担当 5 名をアサインし、サブ課題 C からサブ課題 A と B との連携を促進する体制を図る。(計画書 11 ページに反映)

(6) 連携を促進する具体的な運営体制・方法を再検討すること。

> 上記 (5) の回答中に記載。

(7) AI の取り込みなども含めて、将来の利活用を検討し、ダイナミックに製薬業界が進んでいる中でそれに対応した基盤技術となるよう留意すること。それを示すため、他の方法と比較して、優位性を示すこと。

> 奥野チームと津田チームが中心となり各種創薬シミュレーション技術と AI 技術の融合を図るとともに、AI 創薬コンソーシアム(製薬企業、IT 企業、アカデミア 90 団体以上参加)である LINC と連携し、ビッグデータ統合システムに AI 創薬技術を導入することで世界に類を見ない創薬計算基盤を実現する。具体的には、最適化アルゴリズムによる MD 回数の劇的削減、モンテカルロ木探索による効率的化合物生成、AI スコア導入によるドッキングスコアの高精度化、AI による力場補正、長時間 MD トラジェクトリーの AI 分類などの AI 技術開発を行い、分子シミュレーションアプローチの限界を補完することを目指す。(計画書 24 ページの反映)

(8) 採択時の指摘事項(積極的に成功事例を出していくため、特に重点的に成果を創出するテーマをしぼって設定すること)への対応が必ずしも十分ではない。課題として優れた成果が期待できる点を明確にし、そこに向けて努力を集中することを再度検討すること。

> 全体の共通テーマとしてゲノム医療・ゲノム創薬を応用ターゲットにすることに決定した。ゲノム医療・ゲノム創薬は、現在世界的に急速に臨床実装、創薬実装が進むテーマであり、本テーマにサブ課題 A、B、C の全員でチャレンジすることでポスト「京」の創薬応用の課題が網羅されると考える。

(計画書 3 ページに反映)

(別紙1) 実施機関一覧

	実施機関	備考	
	理化学研究所(奥野 恭史)	代表機関(課題責任者)	
サブ課題A	理化学研究所(杉田 有治)	分担機関(サブ課題責任者)	
	東京工業大学 生命理工学研究科	分担機関(北尾 彰朗)	
	京都大学 大学院理学研究科 化学専攻	分担機関(林 重彦)	
	京都大学 大学院理学研究科 生物科学専攻	分担機関(高田 彰二)	
	理化学研究所 放射光科学研究センター	協力機関	
	大阪大学 大学院基礎工学研究科	協力機関	
	理化学研究所 計算科学研究センター	協力機関	
	名古屋大学 大学院理学研究科	協力機関	
	Hong Kong University of Science and Technology	協力機関	
	先端医療振興財団 クラスター事業推進課	協力機関	
	スーパーコンピューティング技術産業応用協議会	協力機関	
	サブ課題B	横浜市立大学(池口 満徳)	分担機関(サブ課題責任者)
東京大学 先端科学技術研究センター		分担機関(藤谷秀章・山下雄史)	
名古屋大学 大学院工学研究科		分担機関(篠田 渉)	
国立研究開発法人量子科学技術研究開発機構		分担機関(河野 秀俊)	
理化学研究所 放射光科学研究センター		協力機関	
東京工業大学 生命理工学研究科		協力機関	
横浜市立大学		協力機関	
京都大学 大学院工学研究科		協力機関	
首都大学東京 大学院理工学研究科		協力機関	
東京大学 工学系研究科		協力機関	
学習院大学 理学部		協力機関	
京都大学 薬学研究科		協力機関	
名古屋市立大学 大学院医学研究科		協力機関	
名古屋大学 大学院医学系研究科		協力機関	
東京大学 先端科学技術研究センター		協力機関	
大阪大学 薬学研究科		協力機関	
東京医科歯科大学		協力機関	
Michigan State University		協力機関	
大阪大学 蛋白質研究所		協力機関	
サブ課題C		京都大学(奥野 恭史)	分担機関(サブ課題責任者)
		産業技術総合研究所 創薬分子プロファイリング研究センター	分担機関(広川 貴次)
	東京大学 大学院新領域創成研究科	分担機関(津田 宏治)	
	東京大学 情報学環・学際情報学府	分担機関(寺田 透)	
	NPOバイオグリッドセンター関西	協力機関	
	京都大学 医学研究科人間健康科学専攻	協力機関	
	京都大学 医学部附属病院 腎臓内科	協力機関	
	京都大学 医学部附属病院 がんセンター	協力機関	
	先端医療振興財団 先端医療センター研究所	協力機関	
	がん研究会 がん化学療法センター	協力機関	
	国立がん研究センター	協力機関	
	神戸大学 自然科学系先端融合研究環/工学研究科	協力機関	
	京都大学 大学院薬学研究科	協力機関	
	岐阜大学 大学院連合創薬医療情報研究科	協力機関	
	JST-CREST「ビッグデータ応用」船津プロジェクト	協力機関	
	日本医療研究開発機構(AMED) 創薬支援戦略部	協力機関	
	(株)UT-Heart研究所	協力機関	
	先端医療振興財団 クラスター事業推進課	協力機関	
	大阪大学 蛋白質研究所蛋白質解析先端研究センター	協力機関	
	日本医科大学	協力機関	
	ライフインテリジェンスコンソーシアム	協力機関	