

令和2年度高性能汎用計算機高度利用事業
「富岳」成果創出加速プログラム
「プレシジョンメディスンを加速する創薬ビッグデータ統合
システムの推進」
成果報告書

令和3年5月28日
国立研究開発法人 理化学研究所

奥野 恭史

目次

1. 補助事業の目的.....	1
2. 令和2年度（報告年度）の実施内容.....	1
2-1. 当該年度（令和2年度）の事業実施計画.....	1
2-2. 実施内容（成果）.....	4
(1) 分子病態解析.....	4
(2) 薬剤反応性推定.....	10
(3) 薬剤分子設計.....	22
(4) AI・データサイエンス・社会実装基盤.....	29
(5) プロジェクトの総合的推進.....	37
2-3. 活動（研究会の活動等）.....	37
2-4. 実施体制.....	38
別添1 学会等発表実績.....	40

補助事業の名称

「富岳」成果創出加速プログラム

「プレシジョンメディスンを加速する創薬ビッグデータ統合システムの推進」

1. 補助事業の目的

ポスト「京」重点課題1で「富岳」のために開発してきた GENESIS、創薬ビッグデータ統合システムを用いることで、患者由来の遺伝子多型・変異が、タンパク質の構造やダイナミクスに与える影響を分子シミュレーションにより明らかにする。これによって得られる分子レベルでの病態解明・薬剤反応性・薬剤設計に関する知見を臨床現場、創薬現場に提供することで、プレシジョンメディスンの加速を目指す。

2. 令和2年度（報告年度）の実施内容

2-1. 当該年度（令和2年度）の事業実施計画

（1）分子病態解析

①動的構造機能解析（横浜市大・池口）

本研究課題では、キナーゼやその他の希少疾患関連タンパク質を対象として、変異導入による立体構造変化、構造ダイナミクスの変化を研究する。令和2年度は、MSMに向けた構造モデリング、変異体の構造モデリングを行うと共に、対象タンパク質1種類について分子動力学計算のテストを行い、構造安定性計算などの方法の改良を図る。

②タンパク質活性予測（京大理・林）

ハイブリッド QM/MM 自由エネルギー法である QM/MM RWFE-SCF 法を適用し、キナーゼのリン酸化反応解析、及び共有結合型阻害剤の反応性に関する計算を行う。令和2年度は、キナーゼ野生型1種類のモデル分子とモデル反応に対して、効率的な反応活性化状態最適化法の開発を行う。

（2）薬剤反応性推定

①結合自由エネルギー計算（京大医・奥野）

Virtual ligand 法、MP-CAFEE 法や MutationFEP 法を適用し、がんゲノムで観測された各種変異に対して、キナーゼタンパク質と非共有結合型阻害剤や ATP との結合自由エネルギーを推定することで、各種変異型キナーゼと薬剤との反応性を見積もり、各種変異に対して最適な薬剤選択を可能にする。令和2年度は、計算対象のモデリングを行い、キナーゼ-阻害剤数十ペアについて、各種計算手法を試行する。

②結合経路・ポーズ推定（理研 BDR・杉田）

本課題では、ポスト「京」重点課題1で開発した gREST/REUS 法を適用し、がんに関連するキナ

ーゼと阻害剤の結合に対して、X線結晶構造などで観測しうる最安定な結合ポーズだけでなく、準安定的・過渡的なポーズを含めた複数の結合ポーズと、それらに至る結合経路を自由エネルギー地形に基づいて特定する。さらに、阻害剤との結合だけでなく、基質タンパク質（ペプチド）の結合ポーズを特定して、特異性の高い阻害剤の創出を目指す。令和2年度は、計算対象のモデリングを行い、キナーゼ-阻害剤1ペアを対象として計算条件を最適化し、必要であれば計算手法を改良する。

③結合速度論解析（東工大・北尾）

PaCS-MD/MSM法を用いて、キナーゼと阻害剤などの結合や解離の過程を詳しく調べ、結合のメカニズムを明らかにする。令和2年度は、これまで開発してきたPaCS-MD/MSM法を対象複合体1ペアに適用し、必要があれば計算法の改良を行う。

（3）薬剤分子設計

①低分子デザイン（産総研・広川）

ポスト「京」重点課題1で構築した、MD計算と標的タンパク質モデリングおよびドッキング計算法を融合したSBDD技術とChemTSによるデノボデザインを可能にする「創薬ビッグデータ統合システム」を用いて、薬剤耐性キナーゼを回避する低分子デザインおよび希少疾患関連タンパク質を標的とした低分子探索とデザインを行う。令和2年度は、薬剤耐性変異タンパク質のモデリング法の改良、希少疾患関連タンパク質の選定を行う。

②抗体医薬デザイン（東大先端研・山下）

これまで「京」コンピュータを活用して構築した大規模MDシミュレーション技術を駆使して、タンパク質抗原が変異することによる抗体医薬品薬効への影響の予測を可能にしていく。さらに、抗原-抗体相互作用が低下する抗原変異に対して、抗体改良によって再び抗原との相互作用が回復できないかをMDシミュレーションに基づき推定する。令和2年度は計算対象のモデリングを行い、変異体抗原-野生型抗体4種について、各種計算手法を試行する。

（4）AI・データサイエンス・社会実装基盤

①疾患ゲノム構造機能データベース（京大医・鎌田）

本研究は、京都大学大学院医学研究科で開発している日本人の疾患関連ゲノム情報（遺伝子多型・変異）を集積するデータベースMGeND（Medical Genomics Japan Variant Database）において、明確な疾患との関連が同定されていないゲノム情報を標的タンパク質の探索に活用することを目的としている。令和2年度は、MGeNDにおいて臨床的意義が不明な遺伝子多型・変異を中心に、PDBに立体構造情報が登録されており非同義置換の対象となるタンパク質を選出する。

②AI・シミュレーション融合（横浜市大・寺山）

（3）①低分子デザインと連携し、これまで開発してきたChemTS等のAIを用いた新規分子生成手法を用いてターゲットタンパク質に対する分子設計を実施する。令和2年度はターゲットタンパク質に適した分子設計手法のパラメータ検討や高速に探索するための並列化の検討を行う。具体的

には、ChemTS を用いて多数の化合物を発生させるとともに、キナーゼ・阻害剤 1 ペアを対象として COMBO を用いた MD シミュレーションの効率化を図る。

③創薬ビッグデータ統合システム（理研 MIH・本間）

神戸医療産業都市推進機構の分子シミュレーション GUI である K4 や京都大学と理研の AI 構築インターフェイス Knime-GCN に基づいて、計算時間のかかるシミュレーションや AI のハイパーパラメータ最適化について「富岳」と連携できるプラットフォームを構築する。令和 2 年度は、環状ペプチドを対象とした FMO 力場 AI 開発のためのプログラム開発を開始する。

（5）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の会議を適宜、開催する。プロジェクト内の研究の進捗状況および成果の発表のためのワークショップ等を開催する。プロジェクトで得られた成果については学会発表等により、積極的に公表する。また、他のプロジェクトとの連携などにより、効率的・効果的な研究の推進を行う。

2-2. 実施内容（成果）

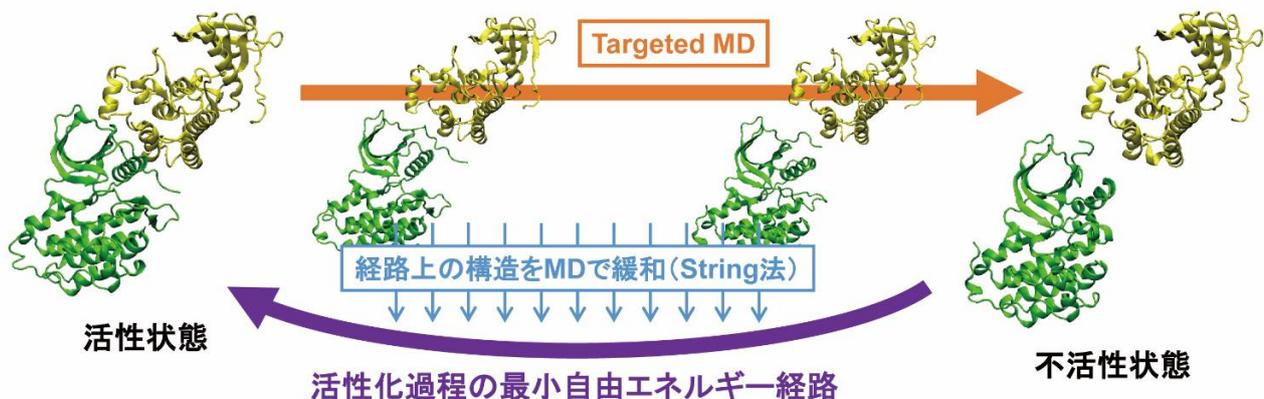
（1）分子病態解析

①動的構造機能解析

横浜市立大学 生命医科学研究科 池口満徳

本課題は、スーパーコンピュータ「富岳」で大幅な効率化が期待される全原子分子動力学シミュレーション（MD）を活用し、変異導入によるタンパク質の立体構造変化、構造ダイナミクス変化の解明を目指すものである。令和2年度は、上皮成長因子受容体（EGFR）キナーゼドメインについて、MD計算のシステムセットアップ、及び、「富岳」にて構造変化過程のサンプリング計算のテストを実施することで、次年度以降の大規模計算の準備を整えた。EGFRは、細胞外ドメインに上皮成長因子（EGF）が結合するとダイマーを形成して活性状態となり、核内へとシグナルを伝達させる。EGFRが活性化する際には、細胞内のキナーゼドメインはアシンメトリックダイマーを形成し、ダイマーの片側（レシーバー側）のキナーゼが活性化することが知られているが、ダイマー化に伴う構造変化や構造ダイナミクス変化の詳細は解明されておらず、発がんやがんの増殖に関与するEGFRキナーゼの遺伝子変異の影響について、立体構造の変化や構造ダイナミクスへの影響が明らかになっていなかった。

EGFRキナーゼの活性化に関連する構造変化や構造ダイナミクス変化は、キナーゼドメイン内の活性化過程だけでなく、ダイマー化も関与するため、通常MDのタイムスケールではシミュレーションできない。そのため、活性化過程の構造変化経路の探索と、マルコフ状態遷移モデル（MSM）による状態間遷移の統合を組み合わせた、長時間ダイナミクスのシミュレーションを実施する。MSMでは、複数の短いMD計算で得られた多数のトラジェクトリを統合し、クラスタリングによって定義した状態間の遷移確率を基に、長時間の構造ダイナミクスを得ることができる。しかし、EGFRキナーゼがダイマーを形成して活性化する過程については、全体での構造変化が大きく、活性状態に至る構造変化の経路が複数通り想定される。そのため、MSMを実施する前に、構造変化の経路探索をストリング法により実施した。ストリング法を用いると、端点となる2つの構造間の構造変化の経路を、自由エネルギー最小となるように最適化できる。具体的には、構造変化の始状態と終状態を繋ぐ経路上に構造が少しずつ異なるコピー系を配置し、それらをMD計算でまとめて動かすことで最小自由エネルギー経路を効率よく探索する。図(1)-1では、主にホモロジーモデリングを用いて作成したEGFRキナーゼの活性状態の構造（ダイマー）と不活性状態の構造（ダイマー形成直前）、及びストリング法の概要についてまとめている。



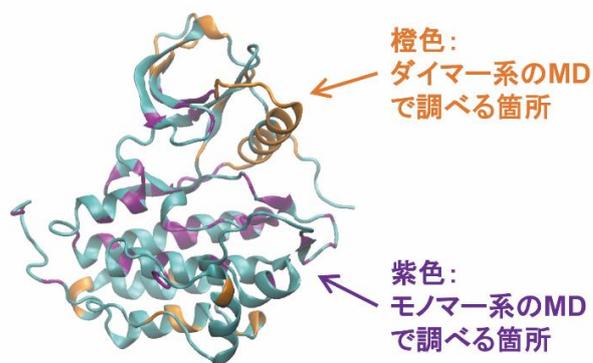
図(1)-1 EGFRの2量体における活性状態と不活性状態の構造変化経路。初期経路をTargeted MDにより生成し、経路上の構造をSTRING法によりMDで緩和し、活性化過程の最小自由エネルギー経路を得

る。

手順としては、先ず、活性状態を始状態、不活性状態を終状態に指定し、拘束力を掛けて始状態と終状態を繋ぐ Targeted MD を実施することで、ストリング法の初期経路を取得する。その後、この初期経路上の構造を緩和させ、最小自由エネルギー経路を探索する。令和2年度には、EGFR キナーゼダイマーの活性状態と不活性状態の構造モデリング、及び、Targeted MD による初期経路作成とストリング法を実施した。構造モデリングでは、活性状態の構造は、ダイマーの結晶構造を基にホモロジーモデリングで作成した。一方で、不活性状態の構造については、ダイマー形成直前の結晶構造が存在しないため、レシーバー側のキナーゼが活性状態であるダイマー結晶構造と不活性状態のモノマー結晶構造を組み合わせることで作成した。ストリング法を実施した結果、緩和させた構造変化経路は初期経路に強く依存し、キナーゼドメイン同士のダイマー形成過程と、レシーバーの不活性状態から活性状態への構造変化過程のカップルの仕方に複数のパターンがあることが分かった。そのため、初期経路作成の段階で、ダイマー化とレシーバーの構造変化が段階的に起こるパターンや、構造変化のうち N-lobe の構造変化がカップルするモデルなど、複数の初期経路の計算が必要なことが分かった。

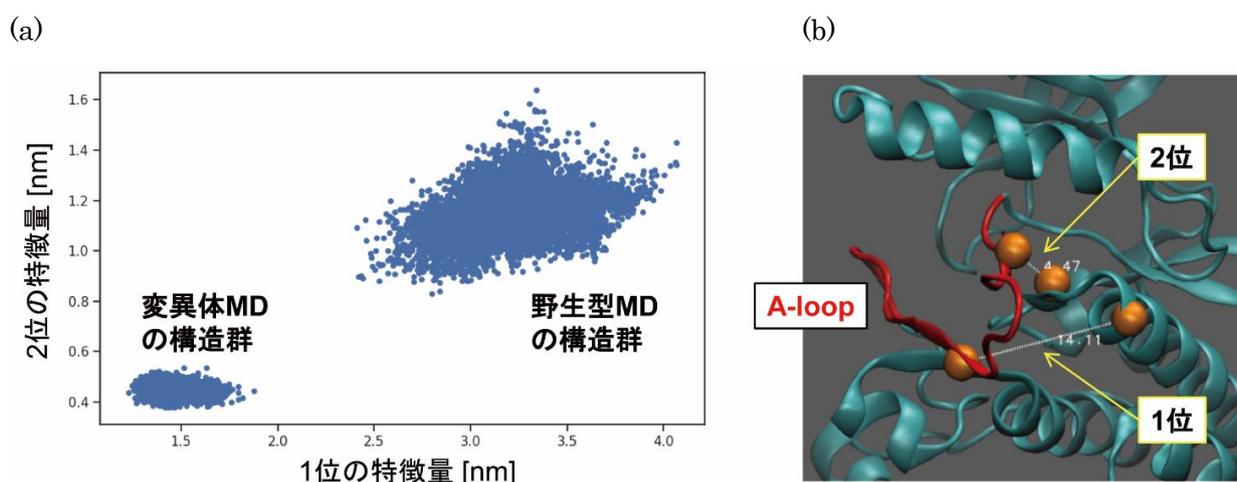
MSM で解析する多数のトラジェクトリを得るための大規模 MD 計算に向けて、「富岳」にて MD 計算速度のベンチマークを実施した。使用した MD 計算ソフトは GENESIS (バージョン 2.0 beta) である。その結果、EGFR キナーゼダイマーの系 (水分子込みで原子数約 100,000) では、40 ノード使用で 136 ns/day の計算速度となった。また、EGFR キナーゼモノマーの系 (原子数約 54,000) では、24 ノード使用で 147 ns/day となった。比較として、「京」で同計算ソフト (バージョンは異なる) を使用した場合には、EGFR キナーゼモノマーと同等サイズの系で 17 ns/day の計算速度であったため、「富岳」では GENESIS による MD 計算速度が大きく向上していることを確認できた。

EGFR キナーゼに変異が生じた場合の立体構造変化、及び、構造ダイナミクスの変化については、野生型と変異体の MD 計算を実施し、それらのトラジェクトリを比較して解析する。臨床の現場では現在までに 100 種類を超える EGFR キナーゼの変異体が確認されており、それらの中には、立体構造からだけでは理解が困難な、意義不明変異も含まれている。EGFR キナーゼの様々な残基に変異が生じることも確認されているが、それらが活性化の過程に及ぼす影響を考えた場合、ダイマー化の過程に影響する変異と、モノマーの構造変化に影響する変異に分類される。ダイマー化の過程に影響する変異はダイマー系の MD 計算で調べる必要があるが、モノマーの構造変化に影響する変異は原子数の少ないモノマー系の MD 計算で調べることができる。そこで、本課題で提供を受けた EGFR キナーゼドメインの変異体リストにおいて、各変異体がダイマー化に影響するか、あるいはモノマーの構造変化に影響するかを変異箇所から予測し、ダイマー系の MD で調べるものとモノマー系の MD で調べるものに分類した (図(1)-2)。更に、この分類の結果と前述した MD 計算のベンチマーク結果を合わせることで、EGFR キナーゼの変異を MD 計算で網羅的に調べるために必要な計算資源を見積もった。



図(1)-2 EGFR の変異箇所。ダイマー系の MD で調べる箇所 (橙) とモノマー系の MD で調べる箇所 (紫) に分類した。

また、令和2年度は、野生型と変異体で1 μ s 程度の比較的短時間の MD トラジェクトリから、変異導入による構造変化、構造ダイナミクス変化を見出す解析手法の確立を行った。2つのトラジェクトリ間の違いを決定木ベースの機械学習アルゴリズムで見出す GINI importance を導入し、キナーゼドメインの変異である L858R 変異による構造変化への影響が推定できるか試した。L858R 変異は、キナーゼドメインの A-loop の二次構造を変化させ、異常活性を引き起こすことが知られている。EGFR キナーゼの野生型と L858R 変異体のモノマーに対して、それぞれの結晶構造を初期状態とした長さ 800 ns の MD 計算を実施し、GINI importance によって解析した。この機械学習では、特定残基の C α 原子間距離や二面角などの特徴量を説明変数とした決定木をランダムに生成し、野生型と変異体両方の MD トラジェクトリを合わせた構造群を2つのグループに分けるというプロセスを複数回実行させた（ランダムフォレスト）。この時、構造群をより正確にグループ分けできた特徴量、即ち、野生型と変異体のトラジェクトリの違いが特に反映されている特徴量を、GINI importance で順位付けて選出した。図(1)-3(a)、(b)では、C α 原子間距離という特徴量に対して、GINI importance で1位、2位として選出された特徴量についてまとめている。



図(1)-3 GINI-importance による特徴量の抽出。野生型と変異体の MD から、GINI importance によって、構造群の分類を特徴づける説明変数を選出した。(a) 1位と2位の特徴量の分布。(b) 1位と2位の特徴量の構造上の位置。

GINI importance で1位、2位となった特徴量で構造群をプロットすると、野生型と変異体で構造群が綺麗に分かれている (図(1)-3(a)、(b))。また、これらの特徴量はどちらも A-loop の動きに関連する部位の C α 原子間距離であった ((図(1)-3(b))。上記の結果は、決定木の数 10,000 のランダムフォレストから得られた結果の一つであるが、同条件で何度か解析を実行しても、選出される C α 原子間距離の箇所に若干の変化は見られたが A-loop の動きに関連する箇所が選出されるという点は変わらなかった。また、決定木の数を100くらいまで減らしても似たような箇所の C α 原子間距離が選出されることも確認できた。以上の結果から、野生型と L858R 変異体の MD トラジェクトリでは、A-loop の振舞いに大きな違いが見られるという結論が得られた。この結論は、L858R 変異が A-loop の二次構造を変化させるという先行研究の結果とも一致するため、GINI importance による解析の有用性を示している。

②タンパク質活性予測

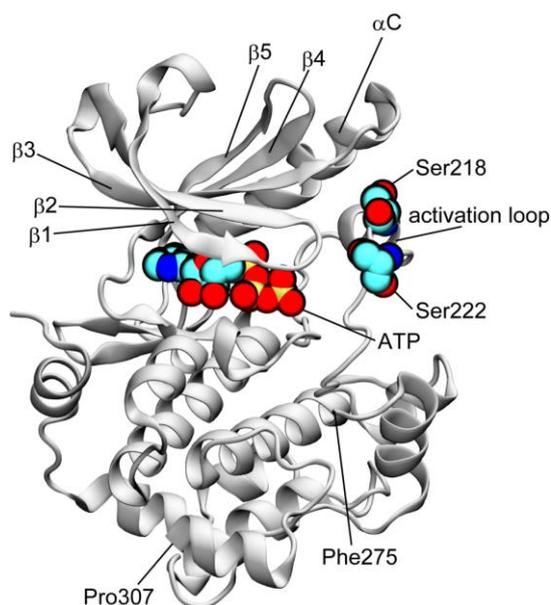
京都大学大学院 理学研究科 林重彦

本課題では、ハイブリッド QM/MM 自由エネルギー法である QM/MM RWFE-SCF 法を適用し、キナーゼのリン酸化反応解析、及び共有結合型阻害剤の反応性に関する計算を行う。令和2年度は、キナーゼ野生型1種類のモデル分子とモデル反応に対して、効率的な反応活性化状態最適化法の開発を行った。具体的には MEK1 キナーゼのリン酸化による ATPase 活性化状態の遷移状態構造を決定する手法を開発し、リン酸化による活性化の分子機構を明らかにした。

まず、キナーゼ野生型モデル分子として MAPK シグナル伝達経路の MAPK キナーゼである MEK1 (図(1)-4) を対象として計算を行った。MEK1 は、自身の activation loop に存在する Ser218 と Ser222 がリン酸化を受けることにより、シグナル伝達経路下流の MAP キナーゼのリン酸化を行う。同様に、MEK1 のリン酸化により、ATPase も活性化されることが知られている。ATP 結合部位と activation loop は離れた位置にあるため、リン酸化による ATPase 活性化の分子機構はこれまで不明であった。そこで、MEK1 キナーゼのリン酸化反応活性化の分子機構を明らかにするために、MEK1 キナーゼの ATPase 酵素反応を理論的に解析した。

QM/MM RWFE-SCF 自由エネルギー計算は、タンパク質環境の古典的力場を用いた長時間分子動力学 (MD) シミュレーションと酵素反応部位の非経験的 QM/MM 計算のサイクルを繰り返すことにより、自由エネルギーを最適化する。MD シミュレーションに関しては、AMBER プログラムパッケージを用いて行った。周期条件の水分子ボックスの中に MEK1 タンパク質を配置した約 68,000 原子系のシミュレーション系を用いた。QM/MM 計算は、GAMESS に本研究室で開発した QM/MM インターフェイスを組み込んだプログラムを用いた。QM 領域には ATP 分子の三リン酸部位 とそれに配位する Mg^{2+} イオン、及びその周辺のアミノ酸側鎖と水分子を含めた。QM 領域の量子化学計算は B3LYP/6-31++G** を用いて行った。QM/MM RWFE-SCF 法の自由エネルギー構造最適化では、1 サイクルあたり 5 ns の平衡化、及び 10 ns のサンプリング計算を行い、20,000 構造サンプルを用いて統計平均を行った。

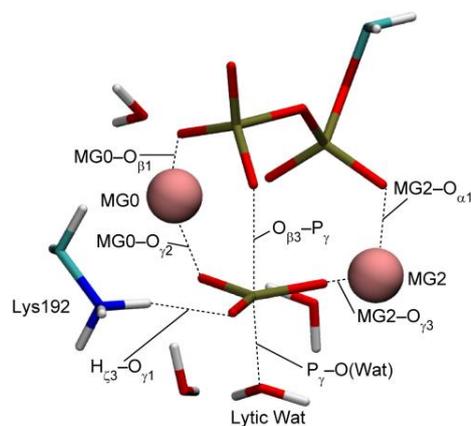
まず、activation loop のセリン側鎖がリン酸化されていない状態 (WT_mg2) に対して、反応始状態の QM/MM 自由エネルギー構造最適化を行った。構造最適化には 1.05 μ s の MD トrajекトリ計算を要した。次に、その最適化構造から出発して、activation loop の 2 つのセリン側鎖を電荷 -1 のリン酸基で修飾したもの (11mg2)、及び電荷 -2 のリン酸基で修飾したもの (22mg2) に関して自由エネルギー構造最適化を行った。構造最適化にはそれぞれ 90 ns の MD 計算を要した。更にそれらの WT_mg2、



図(1)-4 1 MEK1 キナーゼの構造。Activation loop に存在する Ser218 と Ser222 がリン酸化を受けると ATPase が活性化される。

11mg2、及び 22mg2 の 3 状態に対して、それぞれ ATP 加水分解反応の遷移状態を決定した。それらの遷移状態決定に要した MD シミュレーション時間は、それぞれ 135 ns、60 ns、及び 210 ns である。決定した反応始状態と遷移状態の間の自由エネルギー差、すなわち反応活性化自由エネルギーを計算するために、Bennett acceptance ratio (BAR) 法を用いた自由エネルギー計算を行った。それぞれのリン酸化状態の始状態から遷移状態への変化を QM 領域の構造を 42 分割して表し、各点で 5 ns 平衡化、及び 20 ns サンプリング、合計 1050 ns の MD シミュレーションを行った。同様の計算を遷移状態から始状態への方向についても行い 2 方向の計算で得られた自由エネルギー差の平均を最終値とした。従って、合計の BAR 自由エネルギー計算は 6.30 μ s である。更に、タンパク質の大規模構造変化を正確に観測するために、それぞれの自由エネルギー最適化構造に対して 1 μ s の平衡化シミュレーション、また、BAR 計算の最終構造から 2 μ s の平衡化シミュレーションを行い、それぞれ 1 μ s のサンプルからタンパク質の平均構造を決定した。合計の平均構造決定の MD シミュレーションの長さは 18 μ s である。

図(1)-5 に構造最適化計算により得られた WT_mg2 の遷移状態構造を示す。遷移状態では、解離する γ -メタリン酸基と残る二リン酸が近傍の Lys192 及び 2 個の Mg^{2+} イオンと相互作用をしている。表(1)-1 に計算で得られた活性化自由エネルギーをまとめた。Activation loop のセリン側鎖にリン酸化を受けていないタンパク質 (WT_mg2) の場合は、活性化自由エネルギーは 22.3 kcal/mol と計算された。リン酸化された MEK1 の ATPase 酵素反応の活性化自由エネルギーの実験値は 17.6 kcal/mol であり、WT_mg2 のそれは 5.7 kcal/mol ほど高く、実験観測結果と同様に、セリン側鎖がリン酸化されていなければ ATPase 活性が抑制される結果が得られた。また、電荷 -1 のリン酸基の修飾の場合には、活性化自由エネルギーは 14.4



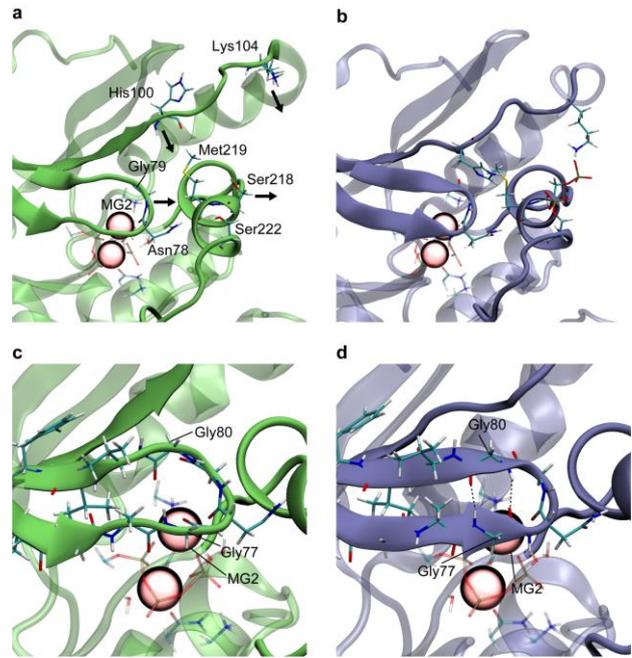
図(1)-5 ATP 加水分解反応の遷移状態 (WT_2mg)

kcal/mol と実験値に近づき、リン酸化を受けていないタンパク質の場合よりも、活性化自由エネルギーが大きく低下する、すなわち ATPase が活性化されることが明らかになった。また、電荷 -2 のリン酸基の修飾の場合には、更に活性化自由エネルギーが低下し 9.8 kcal/mol となり、実験値を大きく過小評価することになった。従って、リン酸化されるリン酸基は 1 つプロトン化された電荷 -1 のリン酸基であることが示唆される。構造解析の結果、activation loop のセリン側鎖のリン酸化により、加水分解の求核攻撃をする水分子が反応始状態より γ -リン酸基に近づき、更に反応遷移状態で Lys192 及び Mg^{2+} が γ -メタリン酸基により近づくことが明らかになった。このような、リン酸化がもたらす活性部位のコンパクトな構造の形成が大きな活性化自由エネルギーの低下をもたらしていることを見出した。

表(1)-1 各リン酸化状態における ATPase 酵素反応の反応活性化エネルギー

RWFE + BAR	ΔE_{QM}	ΔE_{ZERO}	ΔF_{VIB}	$\Delta F_{QM-MM,MM}$	Reaction Barrier
WT_mg2	93.80	-1.49	1.47	-71.48	22.30
11mg2	75.76	-1.06	0.03	-60.34	14.39
22mg2	70.76	-0.94	0.88	-60.92	9.78

上記の様に、ATP 加水分解反応部位と activation loop のセリン側鎖のリン酸化部位は離れており、その相関の分子機構は不明であった。そこで、長時間 MD シミュレーションにより得られた平均構造を解析し、反応性の顕著な変化の分子論的な起源の特定を行った。図(1)-6 にセリン側鎖のリン酸化によるタンパク質構造変化を示す。リン酸化に起因するより強い水合と近傍のループに存在する Lys104 との塩橋形成により、activation loop がよりバルク水側にシフトしている。更に、その activation loop に水素結合している $\beta 1$ 及び $\beta 2$ の β ヘアピン部分も同様にシフトし、より拡張した β シート構造が形成される。この β ヘアピン部分は ATP 加水分解反応部位を形成しており、この構造変化が反応活性化を与えるコンパクトな反応活性部位構造の分子論的起源となっていることを明らかにした。本成果は、現在、論文投稿準備中である。



図(1)-6 Activation loop のセリン側鎖のリン酸化によるタンパク質の構造変化。a WT_mg2、b 11mg2、c WT_mg2 の活性部位、d 11mg2 の活性部...

(2) 薬剤反応性推定

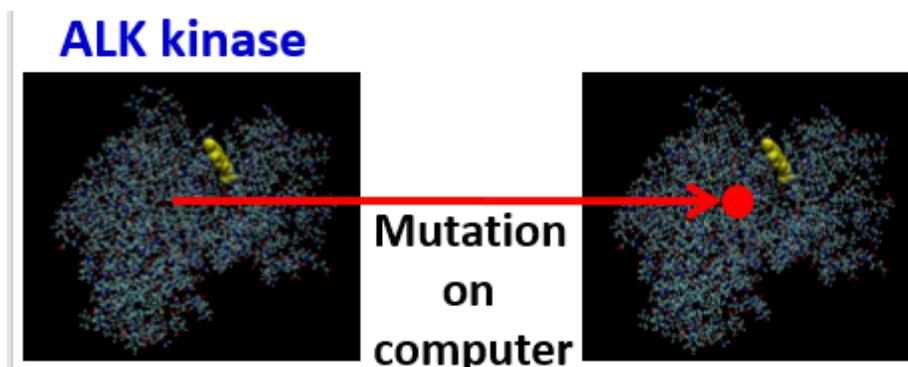
①結合自由エネルギー計算

京都大学大学院 医学研究科 奥野恭史

令和2年度は、タンパク質と化合物の結合ポーズ・結合自由エネルギーを推定する Virtual ligand 法、MP-CAFEE 法、MutationFEP 法といった各種計算手法を「富岳」上に実装すると共に、キナーゼ-阻害剤数十ペアに対して MP-CAFEE 法を実施した。具体的には、非小細胞肺癌患者のゲノム解析から同定された EGFR・ALK 遺伝子上の変異を対象に、変異型キナーゼ-薬剤複合体の分子動力学 (MD) シミュレーション並びに薬剤の結合自由エネルギー計算によって、変異に起因するタンパク質活性化・薬剤耐性化の分子メカニズムを推定した。下記に詳細な研究内容 ((i)-(iii)) を記載する。

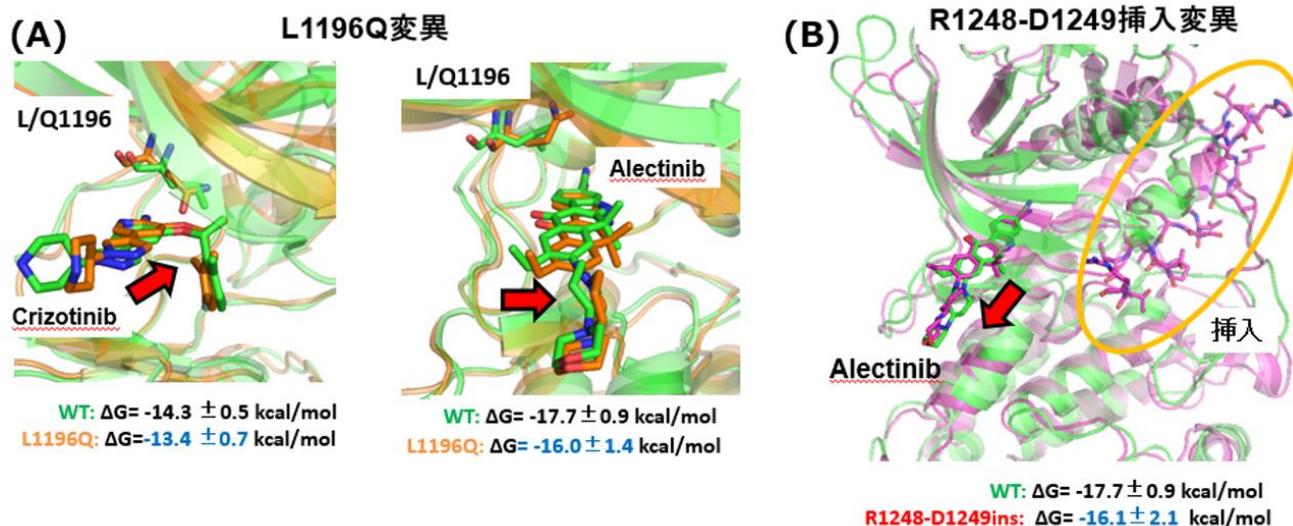
(i) 既存 ALK 阻害薬の薬剤耐性メカニズムの解明

薬物治療を継続していると、遺伝子変異によって薬剤耐性を獲得する場合がある。非小細胞肺癌 (NSCLC) の治療に用いられる ALK キナーゼ阻害薬の場合では、薬剤耐性を引き起こす ALK キナーゼ配列上のアミノ酸変異が複数報告されている。本研究では、非小細胞肺癌患者のゲノム解析から新たに同定された ALK 遺伝子上の2種類の変異[L1196Q 置換変異、R1248-D1249 間の挿入変異 (R1248-D1249ins)] を対象に、ALK キナーゼ変異体と3種類の ALK 阻害薬 (クリゾチニブ、アレクチニブ、セリチニブ) の結合親和性を算出した。各々の薬剤との共結晶構造 (PDBID=2XP2:クリゾチニブ、3A0X:アレクチニブ、4MKC:セリチニブ) に基づいて ALK 変異体2種類の立体構造をモデリングし (図(2)-1)、MP-CAFEE 法によって薬剤との結合自由エネルギー (ΔG) を評価した。その結果、L1196Q 変異体においては、セリチニブの結合には影響を与えなかったものの、薬剤近傍の疎水性残基が親水性残基に変異することによってクリゾチニブとアレクチニブの結合位置がわずかに変化し、これらの薬剤の結合親和性が低下する結果となった (図(2)-2A)。R1248-D1249ins 変異体においては、薬剤に直接接していない領域に24残基のアミノ酸配列が挿入されることにより、アレクチニブの結合位置がわずかに変化し、タンパク質-薬剤間の静電相互作用が部分的に失われることで結合親和性が低下する結果となった (図(2)-2B)。一連のシミュレーションから得られた ALK 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 (Furuta H., et al., *Journal of Thoracic oncology*, 16(3):477-482, 2020; 愛知県がんセンター 真砂博士との共同研究)。



図(2)-1 ALK 変異体の立体構造モデリング。野生型-薬剤共結晶構造 (左) に基づいて、薬剤耐性変異体 (右) をコンピュータ上でモデリングした。薬剤、変異アミノ酸をそれぞれ黄色、赤色でハ

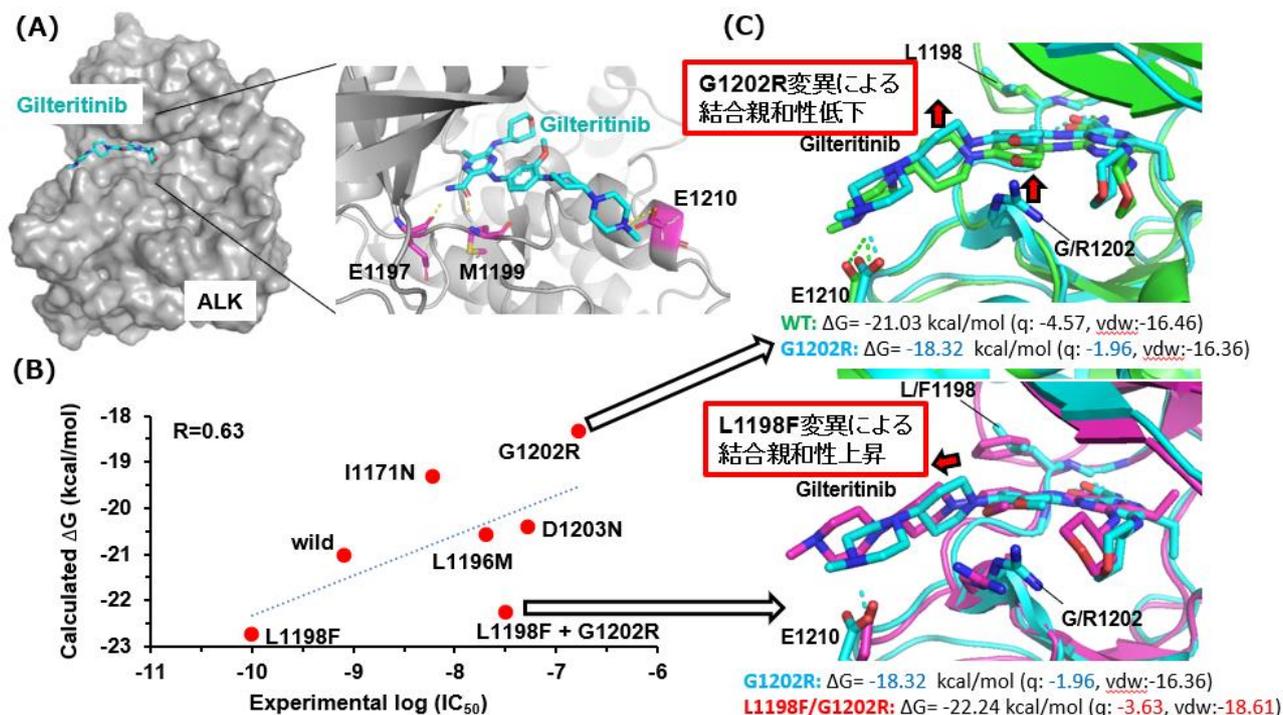
イライトした。



図(2)-2 ALK 変異体-薬剤複合体のシミュレーション構造。(A) L1196Q 変異体に対する (左) クリゾチニブ及び (右) アレクチニブの結合構造。(B) R1248-D1249ins 変異体に対するアレクチニブの結合構造。タンパク質、薬剤は、それぞれリボンモデル、スティックモデルで表示した。ALK 野生型、L1196Q 変異体、R1248-D1249ins 変異体をそれぞれ緑、オレンジ、マゼンタで表示し、野生型と比較した際の薬剤の結合位置の変化を矢印で示した。また、MP-CAFEE によって算出した結合自由エネルギー (ΔG) をシミュレーション構造の下に記載した。

(ii) 新規 ALK 阻害薬の ALK に対する結合ポーズ推定、及び薬剤耐性メカニズムの解明

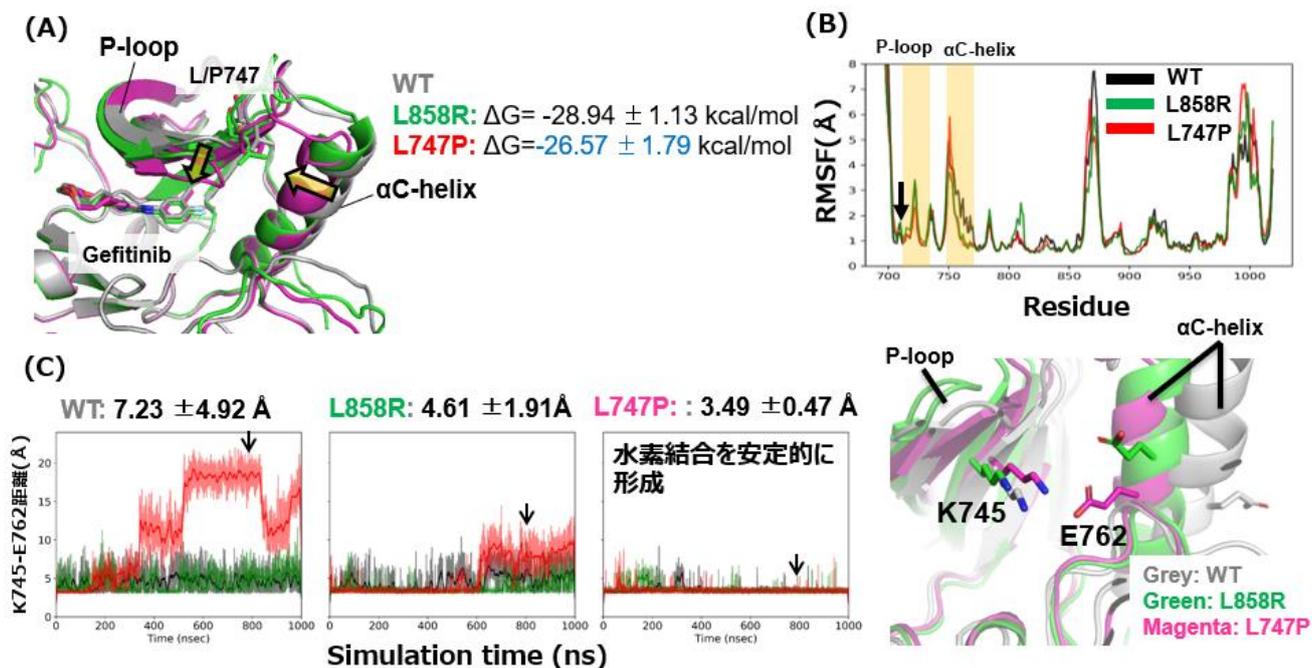
タンパク質-化合物ドッキング計算と複合体の MD シミュレーションの組み合わせによって、新たな ALK 阻害薬として同定したギルテリチニブの ALK に対する結合ポーズを推定した (図(2)-3A)。更に、代表的な薬剤耐性変異体 6 種を対象に、MP-CAFEE 法によって変異型 ALK とギルテリチニブの結合親和性 (ΔG) を推定した。算出した ΔG を実験的に測定された細胞増殖 50%阻害濃度 (IC_{50}) に対してプロットしたところ、両者の間に概ね良好な相関関係が見られたことから、推定した ALK-ギルテリチニブ結合ポーズの妥当性が実証されると共に、変異に伴う薬剤感受性の変化を精度高く推定することに成功した (図(2)-3B)。特に、薬剤近傍のグリシン残基がかさ高いアルギニンに変異した G1202R 変異体では、ギルテリチニブとの静電相互作用が部分的に失われることによって結合親和性が低下するものの (図(2)-3C 上)、薬剤近傍において L1198F 変異が更に導入されることで、より強固な静電相互作用・van der Waals 相互作用を形成するようにギルテリチニブが再配向すると推定された (図(2)-3C 下)。(Mizuta H., et al., *Nature Communications*, 12:1-16, 2021; がん研究会 片山博士との共同研究)



図(2)-3 ALK 変異体-ギルテリチニブ複合体のシミュレーション構造。(A) タンパク質-化合物ドッキング計算と複合体のMDシミュレーションによって推定したALK野生型に対するギルテリチニブの結合ポーズ。(B) MP-CAFEE法によって推定したギルテリチニブの結合自由エネルギー(ΔG)と実験的に測定した薬効値(IC_{50})の相関。(C) G1202R変異体(上)及びL1198F/G1202R変異体(下)に対するギルテリチニブの結合構造。タンパク質、薬剤は、それぞれリボンモデル、スティックモデルで表示した。ALK野生型、G1202R変異体、L1198F/G1202R変異体をそれぞれ緑、シアン、マゼンタで表示し、G1202R変異およびL1198F変異が入った際の薬剤の結合位置の変化を矢印で示した。また、MP-CAFEEによって算出した結合自由エネルギー(ΔG)をシミュレーション構造の下に記載した。

(iii) EGFR阻害薬 (Gefitinib) の薬剤耐性化及びEGFR活性化の分子メカニズム解明

非小細胞肺癌患者の半数近くにおいて、上皮成長因子受容体 (epidermal growth factor receptor: EGFR) の活性化変異ががんの原因遺伝子として報告されている。本研究では、患者のゲノム解析から新たに同定されたEGFR上のL747P変異を対象に、薬剤感受性並びにキナーゼ活性化機構を解析した。EGFR野生型、L858R変異体、およびL747P変異体のマイクロ秒タイムスケールのMDシミュレーション結果を比較したところ、L747P変異体では、変異アミノ酸の近傍に位置する α C-helixやP-loopの配向及び構造柔軟性が顕著に変化することでgefitinibの結合親和性を低下させると示唆され、当該変異体に対してgefitinibの感受性が低下するという実験結果をサポートした (図(2)-4A、図(2)-4B)。また、EGFR活性型コンフォメーションの代表的な特徴として、P-loop近傍に位置する745番目のリジン (K) と α C-helix上に位置する762番目のグルタミン酸 (E) の間で塩橋が形成されることが挙げられるが、L747P変異体ではこの塩橋が恒常的に安定化されていることが示唆されたことから (図(2)-4C)、当該変異体のキナーゼ活性上昇の分子メカニズムを推定することに成功した。

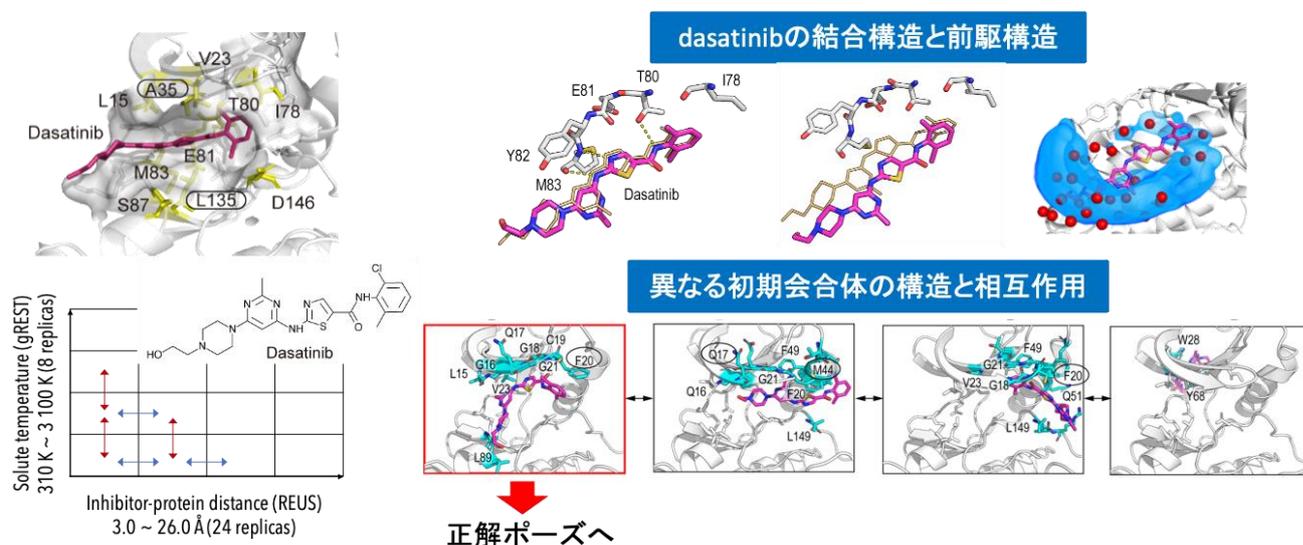


図(2)-4 EGFR-L747P 変異体のシミュレーション構造。(A) $1\mu s$ の MD シミュレーションを独立に 3 本実施して得られた EGFR-Gefitinib 複合体の平均構造。タンパク質、薬剤は、それぞれリボンモデル、スティックモデルで表示した。EGFR 野生型、L858R 変異体、L747P 変異体をそれぞれ灰色、緑、マゼンタで表示し、ALK 野生型・L858R 変異体と比較した際の L747R 変異体における P-loop 及び α C-helix の配向の変化を矢印で示した。また、MP-CAFEE によって算出した結合自由エネルギー (ΔG) をシミュレーション構造の右側に記載した。(B) EGFR 主鎖における構造柔軟性(Root-Mean-Square Fluctuation; RMSF) の解析結果。L747R 変異体において構造柔軟性が顕著に低下した P-loop 領域を矢印で示した。(C) K745-E762 間で形成される塩橋 (水素結合) の安定性。独立した 3 本の $1\mu s$ シミュレーションの結果を赤、緑、黒色で示し、シミュレーション全体から算出した K745-E762 間の平均的距離をグラフの上に示した。また、グラフ中に示した矢印のシミュレーション時間に相当する立体構造をグラフの右に示した。

②結合経路・ポーズ推定

理化学研究所 生命機能科学研究センター 杉田有治

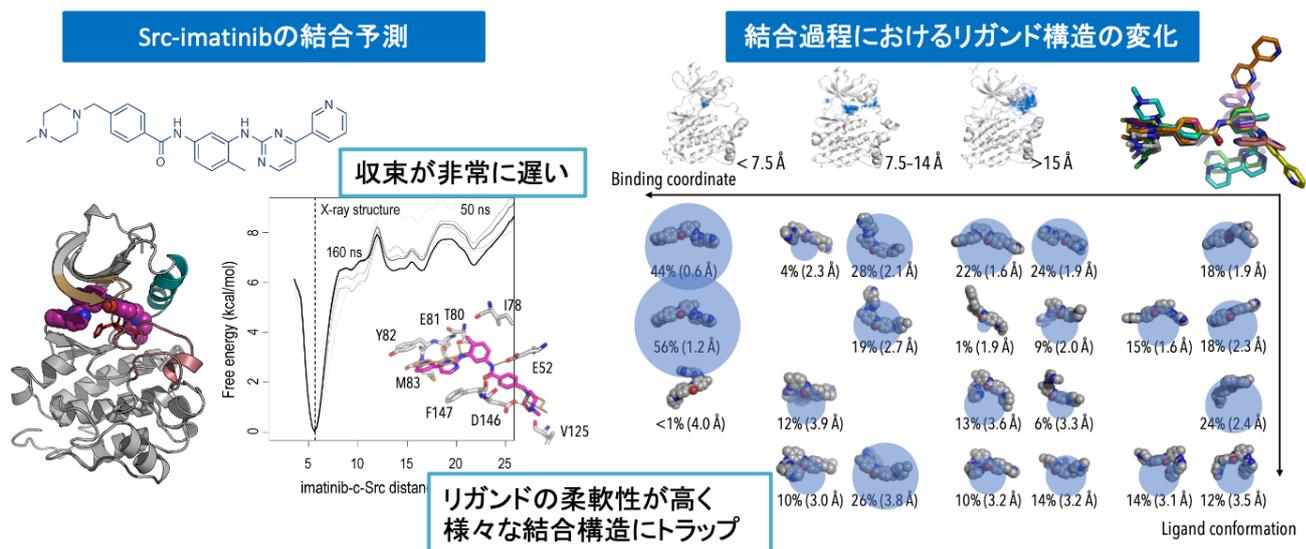
令和2年度は、ポスト「京」重点課題1で開発した generalized Replica Exchange with Solute Tempering (gREST) 法と Replica-Exchange Umbrella Sampling (REUS) 法を組み合わせた gREST/REUS 法を、Src キナーゼとその阻害剤である dasatinib の結合予測に応用し、結合ポーズおよび結合経路を特定した(図(2)-5)。この計算では、gREST 法を用いることでタンパク質結合サイト内残基と dasatinib の構造柔軟性を向上させ、同時に REUS 法で dasatinib の結合・解離を行うことでサンプリング効率を向上できる。共結晶構造での dasatinib の結合ポーズや周囲の水分子配置を再現できるだけでなく、実験では観測することができない準安定な前駆構造を観測することができた。さらに過渡的な初期会合体の構造と相互作用を明らかにした。初期会合体形成は、リガンド結合の成否を左右する重要なステップであるが、実験で捉えることは難しく、詳細についてよくわかっていない。今回の結果は、初期会合体でリガンドは Glycine-rich loop と呼ばれる保存性の高いアミノ酸残基群と相互作用しており、Glycine-rich loop 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示唆している。初期会合体形成における Glycine-rich loop の重要性は阻害剤 PP1 を用いた計算でも示されており (S. Re, H. Oshima, K. Kasahara, M. Kamiya, Y. Sugita, Proc. Natl. Acad. Sci. U. S. A. 116, 18404-18409 (2019))、Src キナーゼとリガンドの結合の普遍的なメカニズムだと考えられる。



図(2)-5 gREST 法で Src キナーゼの残基と dasatinib の運動性・柔軟性を向上させ (左上)、REUS 法と組み合わせることでサンプリングを向上させる (gREST/REUS 法、左下)。得られた結合ポーズ、前駆構造、周囲の水分子の分布 (右上)。Glycine-rich loop と dasatinib との相互作用と結合経路 (右下)。

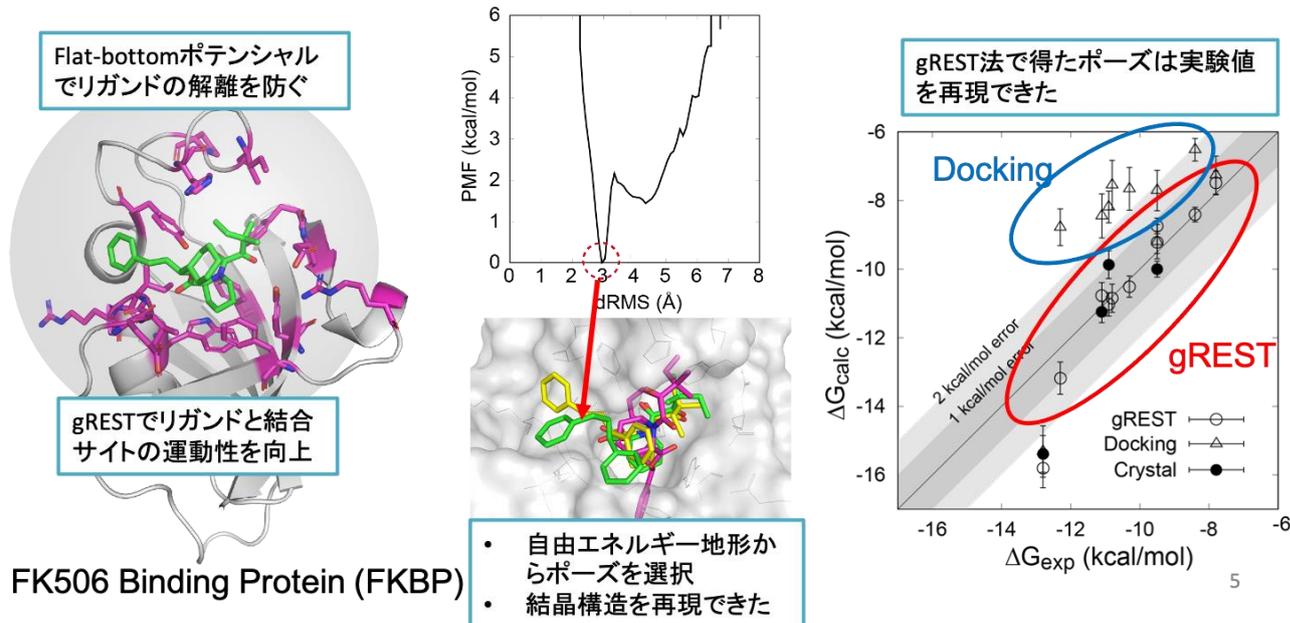
一方、dasatinib と同様の gREST/REUS 計算を阻害剤 imatinib の結合予測で行ったところ、いくつか課題点が見えてきた(図(2)-6)。imatinib では自由エネルギー地形の収束が非常に遅く、160 ns の計算では収束する傾向が見られなかった。imatinib は dasatinib よりも大きく、柔軟性が高いため、様々なコンフォメーションを取ることができる。そのため dasatinib と比べて可能な結合構造の数が急増し、系

が様々な結合構造にトラップされてしまうことで、自由エネルギー計算の収束性が悪くなったと考えられる。また、リガンドのコンフォメーションの重要性が上がったため、リガンドと結合サイトの重心間距離のような単純な反応座標では結合経路を特定することができなくなった。imatinibのような柔軟性の高いリガンドの結合メカニズムの解明には、gREST/REUS法のパラメータの調整に加え、新たな反応座標の導入などが必要である。これらの課題の解決は、imatinibよりもさらに柔軟性が高く、結合構造の組合せが多いペプチドの結合においてさらに重要になると考えられる。



図(2)-6 imatinibの構造およびSrc-imatinib重心間距離の自由エネルギー地形(左)、重心間距離ごとに取り出したimatinibのコンフォメーション(右)。

また、gREST法を用いた効率的なリガンド結合ポーズ予測法を開発した。リガンドの結合ポーズ予測にはドッキング計算がよく用いられているが、タンパク質やリガンドの構造揺らぎ等を考慮しないため正確に予測することが難しい。一方、MD計算では溶媒も含めた全自由度を考慮することで分子力場パラメータの精度範囲で厳密に予測できるが、リガンドが準安定なポーズに捕らわれてしまい、最適なポーズへの収束が非常に遅い。我々は、ポーズ探索を効率化するためにgREST法を用い、FK506結合タンパク質(FKBP)の結合サイトとリガンドの構造揺らぎ・運動性を向上させ、準安定ポーズに捕らわれないようにした(図(2)-7)。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から離れないよう工夫し、リガンドが結合サイト内で自由に動き回れるようにすることで、探索効率を向上させた。得られたトラジェクトリからdistance root-mean-square displacement (dRMS)の自由エネルギー地形を求めることで、最安定な結合ポーズが得られる。FKBPとの共結晶構造が得られている4個のリガンドについて、ドッキング計算でポーズを予測すると2個のリガンドは結晶構造を再現できなかったが、gREST法では結晶構造を全て再現することに成功した。ただし、構造柔軟性が高いリガンドではgREST法によって構造揺らぎが大きくなりすぎ、元々の化学構造が壊れてしまうことがある。その場合は、リガンドの構造に拘束をかけることで、正しくサンプリングできることがわかった。

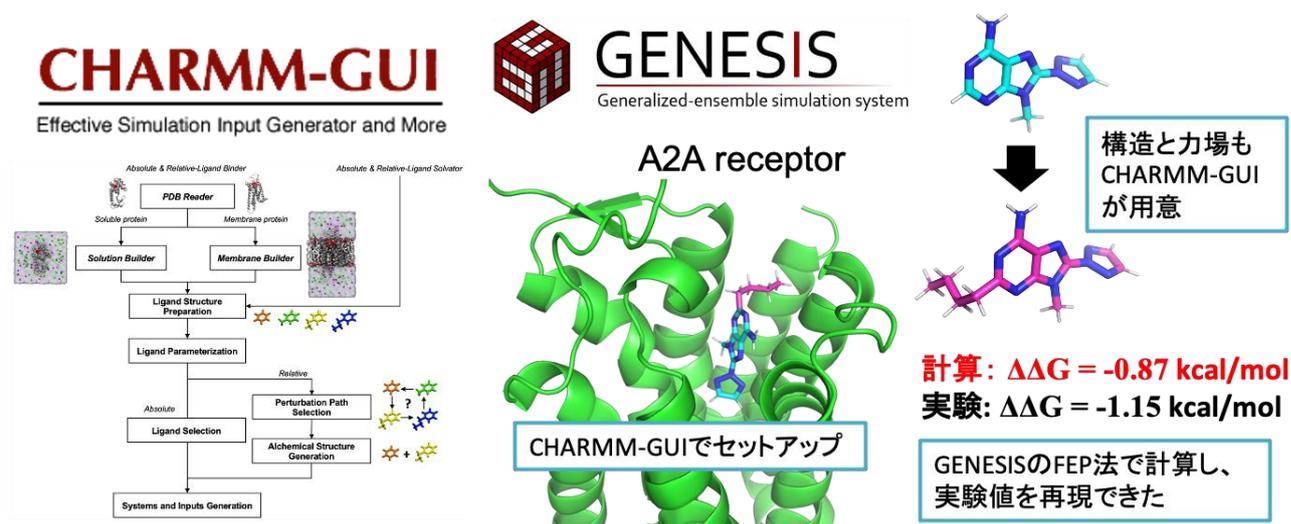


図(2)-7 gREST 法と Flat-bottom ポテンシャルを用いた結合ポーズ探索スキーム (左)、dRMS 自由エネルギー地形からのポーズ決定 (中央)、得られたポーズの結合親和性 (右)。Flat-bottom ポテンシャルはグレー球で示しており、リガンド分子 (緑) は球の中を自由に動き回る。

gREST 法で得られた結合ポーズの結合親和性をアルケミカル自由エネルギー摂動法 (FEP 法) の一種である Massively Parallel Computation of Absolute binding Free Energy with well-Equilibrated states (MP-CAFEE)法を用いて予測した (図 3)。gREST 法で得たポーズの親和性は実験結果と 1kcal/mol の範囲で良く一致した。一方、ドッキング法で得たポーズの親和性を FEP 法で求めた場合は、実験値から大きくはずれた。このことからポーズは親和性予測に大きく影響し、正しいポーズで FEP 計算を行うことで親和性を正確に予測できることがわかった。我々は、gREST 法で結合ポーズを予測し、FEP 法で結合親和性を予測する一連のプロトコルを「gREST+FEP 法」と名付けた。我々のシミュレーションでは共結晶構造は用いておらず、結合ポーズに関する事前情報は一切含まれていない。結晶構造がない系や新規化合物にも応用可能な高精度かつ汎用性の高い予測手法である。gREST+FEP 法の研究成果は国際誌に掲載された (H. Oshima, S. Re, Y. Sugita, J. Chem. Info. Modeling 60, 5382 (2020))。

GENESIS version 1.6.0 には高精度な創薬計算を実現するために FEP 法が実装されている。摂動部分の取り扱いとして、従来のデュアルトポロジーを用いる方法に加え、ハイブリッドトポロジーを用いた方法を導入することで、CHARMM 力場だけでなく AMBER 力場を用いた計算もサポートしている。FEP 法で結合親和性の高精度予測は可能になるが、計算前の準備 (複合体構造の構築、リガンドの力場パラメータの設定、MD シミュレーションソフトウェアへの入力ファイルの作成等) がシステム毎に必要となる。これらの準備は複雑で煩雑なため、時間もかかり、実用化の大きな妨げとなっていた。特に、初心者はセットアップの段階でミスを引き起こすことが懸念される。我々は Wonpil Im (Lehigh 大学)らが開発しているモデリングツール CHARMM-GUI (<http://www.charmm-gui.org>)と連携することで、Web サイト上で簡単に FEP の入力ファイルを作成し GENESIS での計算を行うことができる環境を整

えた (図(2)-8)。CHARMM-GUI の Free Energy Calculator は、リガンドの水和自由エネルギーと結合親和性の絶対値および相対値の計算に必要な力場ファイル、構造ファイル、GENESIS 用入力ファイルと解析用スクリプトを半自動的に生成する。結合親和性の絶対値は MP-CAFE 法を用いて、相対値はハイブリッドトポロジーを用いて計算できる。また、結合親和性計算においては膜あり・なしが選べるようになっている。



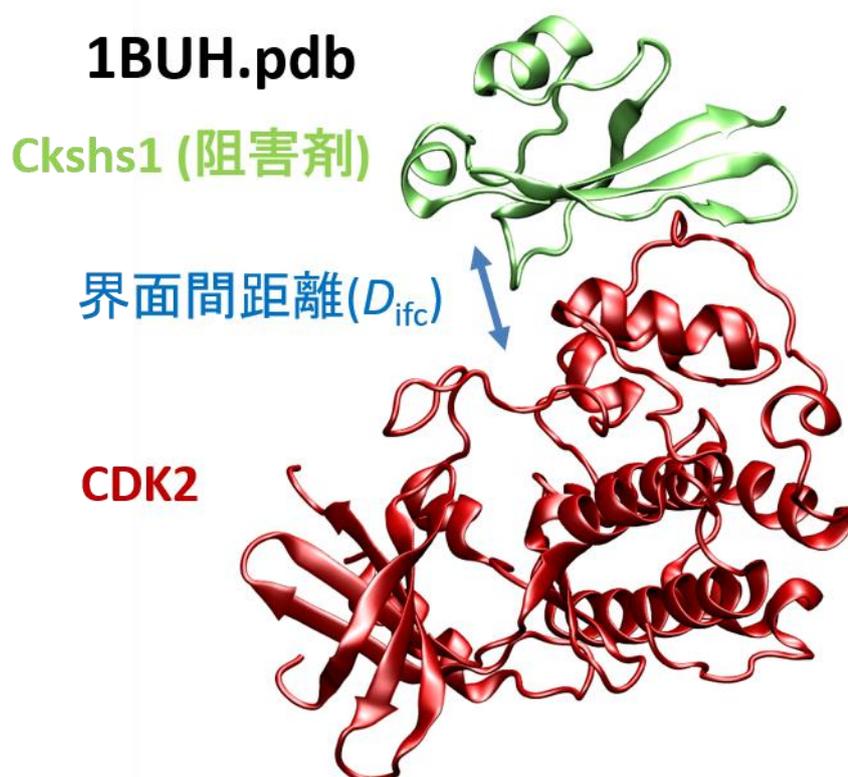
図(2)-8 CHARMM-GUI のセットアップのスキーム (左)、CHARMM-GUI で作成した A2A receptor とリガンドの分子モデル (中央)、リガンド 2 種間の相対結合親和性 $\Delta \Delta G$ の計算値と実験値。

CHARMM-GUI の動作チェックのため、様々な生体分子系の自由エネルギーを GENESIS で計算した。膜タンパク質 A2A receptor のリガンド相対結合親和性の場合、CHARMM-GUI でハイブリッドトポロジー用の入力ファイルを作成し、GENESIS の FEP 法で計算した(図(2)-8)。得られた相対結合親和性 $\Delta \Delta G$ は実験値と非常に良く一致しており、CHARMM-GUI のセットアップおよび GENESIS での FEP 計算が正しく機能していることが確認された。CHARMM-GUI と GENESIS の連携の成果は国際誌に掲載された (S. Kim, H. Oshima, H. Zhang, N. R. Kern, S. Re, J. Lee, B. Roux, Y. Sugita, W. Jiang, W. Im, J. Chem. Theory. Comp. 16, 7207 (2020))。

③結合速度論解析

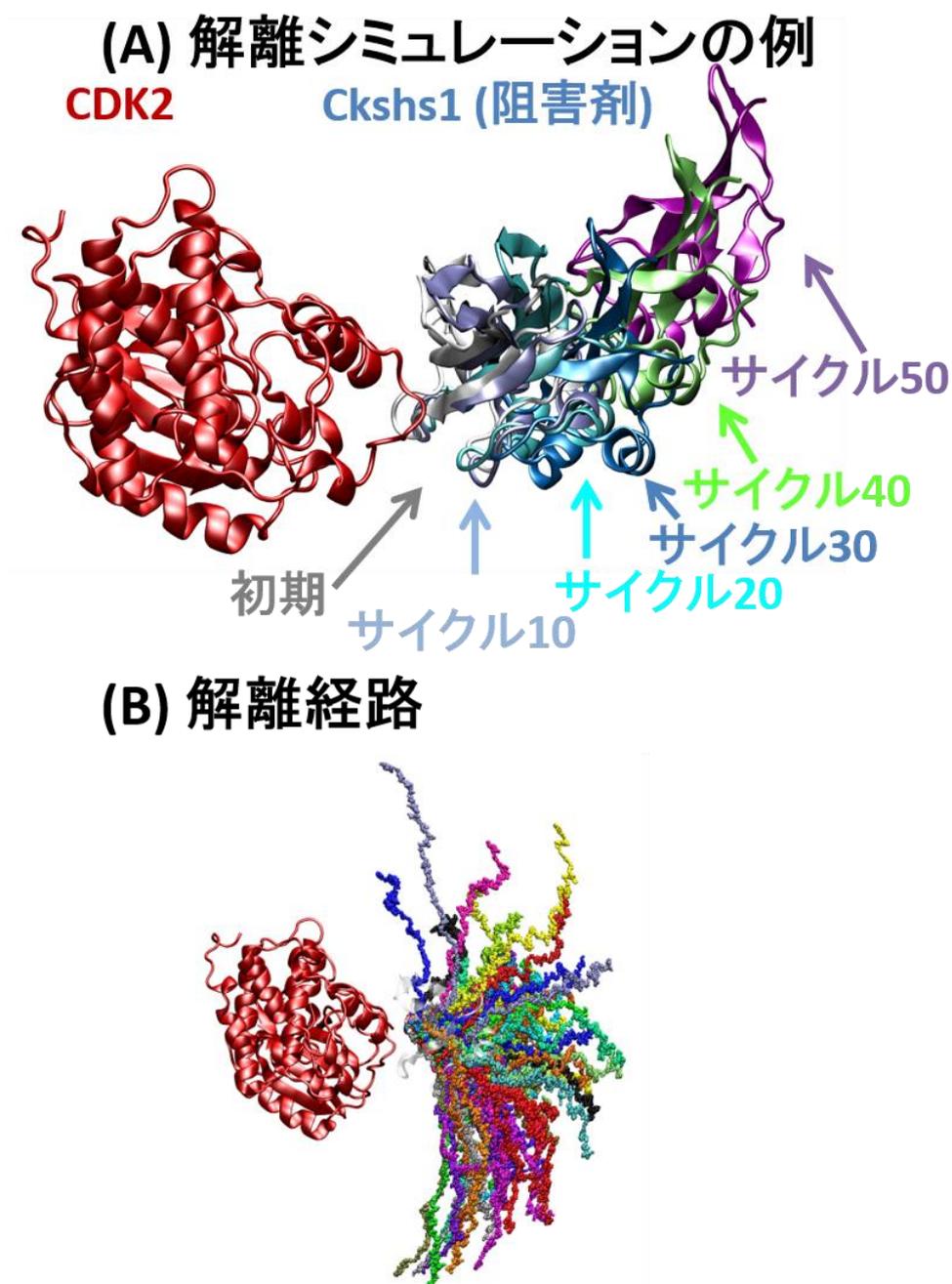
東京工業大学 生命理工学院 北尾彰朗

令和2年度は、これまで開発してきた PaCS-MD/MSM 法を対象複合体 1 ペアに適用し、解離シミュレーションにおいて解離を促進する計算法の改良を行った。PaCS-MD 法は、①比較的短時間の分子動力学計算を同時並列で実行、②得られた構造からターゲットに近い構造を選択、③それらを初期構造として分子動力学計算を再度行う、というサイクルを繰り返す手法で、外力を与えることなしに蛋白質の構造変化や蛋白質複合体の結合・解離を誘起することができる。解離シミュレーションでは、上記②の「ターゲットに近い構造」として複合体を構成する分子間の距離が大きい構造を選択することにより、効率よく複合体の解離を観測できる。これまでも蛋白質-リガンド複合体や蛋白質-ペプチド複合体の解離シミュレーションに成功しており、PaCS-MD 法で生成したトラジェクトリをマルコフ状態モデル (MSM) で解析することにより、結合自由エネルギー、結合・解離速度定数の評価にも成功している。PaCS-MD/MSM 法の蛋白質-蛋白質複合体への適用例として、図(2)-9 に示したキナーゼタンパク質である野生型 CDK2 と阻害タンパク質 Ckshs1 の解離シミュレーションに PaCS-MD 法を適用した。蛋白質-リガンド複合体や蛋白質-ペプチド複合体の解離シミュレーションでは、分子間距離として、分子全体の重心間距離を用いていたが、蛋白質-蛋白質複合体の場合は、蛋白質の形状等を考慮し、図 1 に示した複合体界面を形成する残基の重心間距離 (界面間距離: D_{ifc}) を用いた。



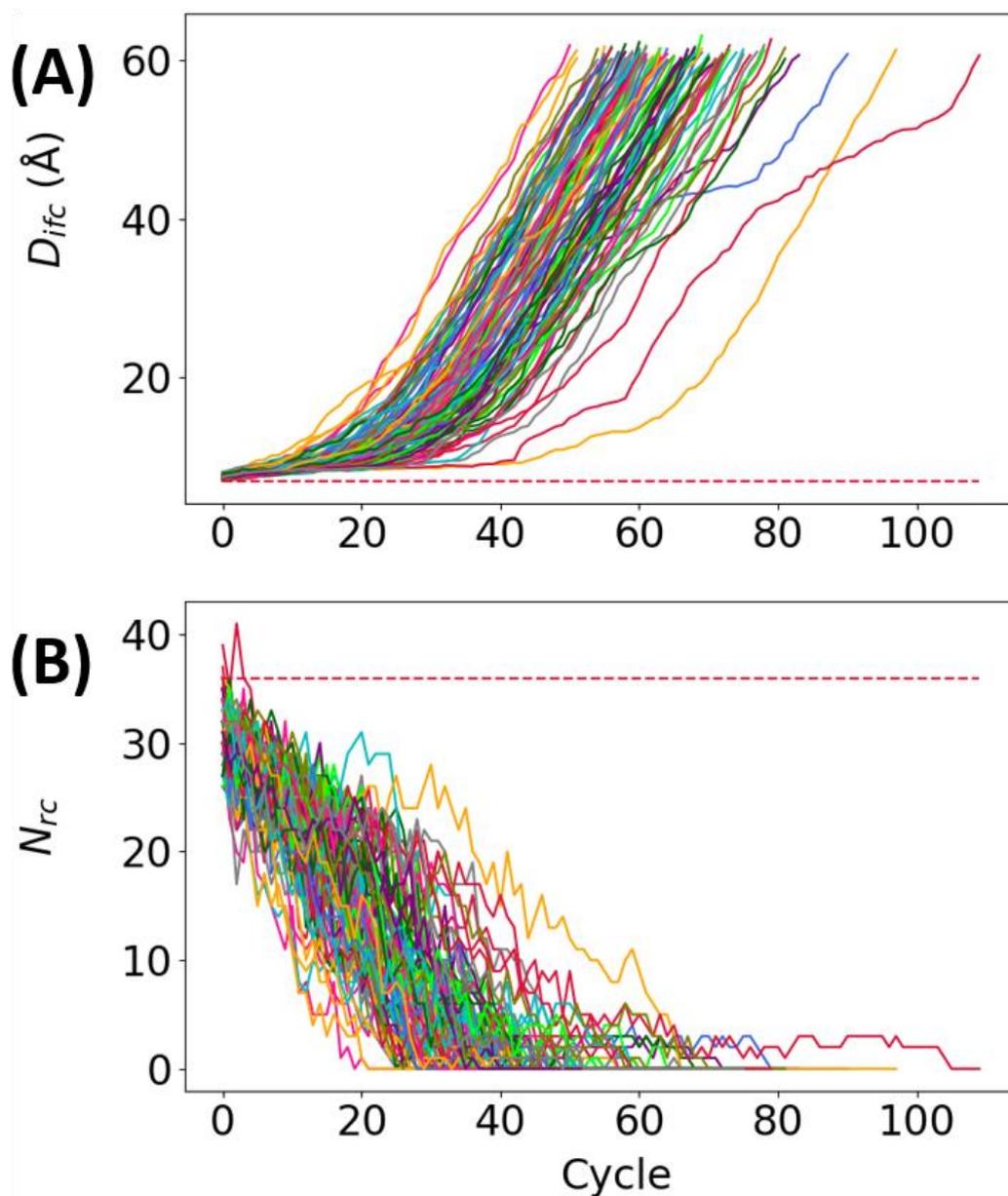
図(2)-9 野生型 CDK2 (赤) と阻害タンパク質 Ckshs1 (緑)

図(2)-10(A)に例を示したように、サイクルが進むにつれて蛋白質間の距離が大きくなる解離シミュレーションを実行した。蛋白質-蛋白質複合体への適用テストとして、1サイクルにおけるMDを100 ps、同時実行するMDの数を10、各サイクルで界面間距離が大きい構造を選択、界面間距離が60 Åに到達するまで計算する条件で解離シミュレーションを100回実行した。図(2)-10(B)に阻害タンパクの重心位置を示したように、全ての試行で蛋白質-蛋白質複合体の解離に成功し、100通りの解離パスウェイの生成に成功した。



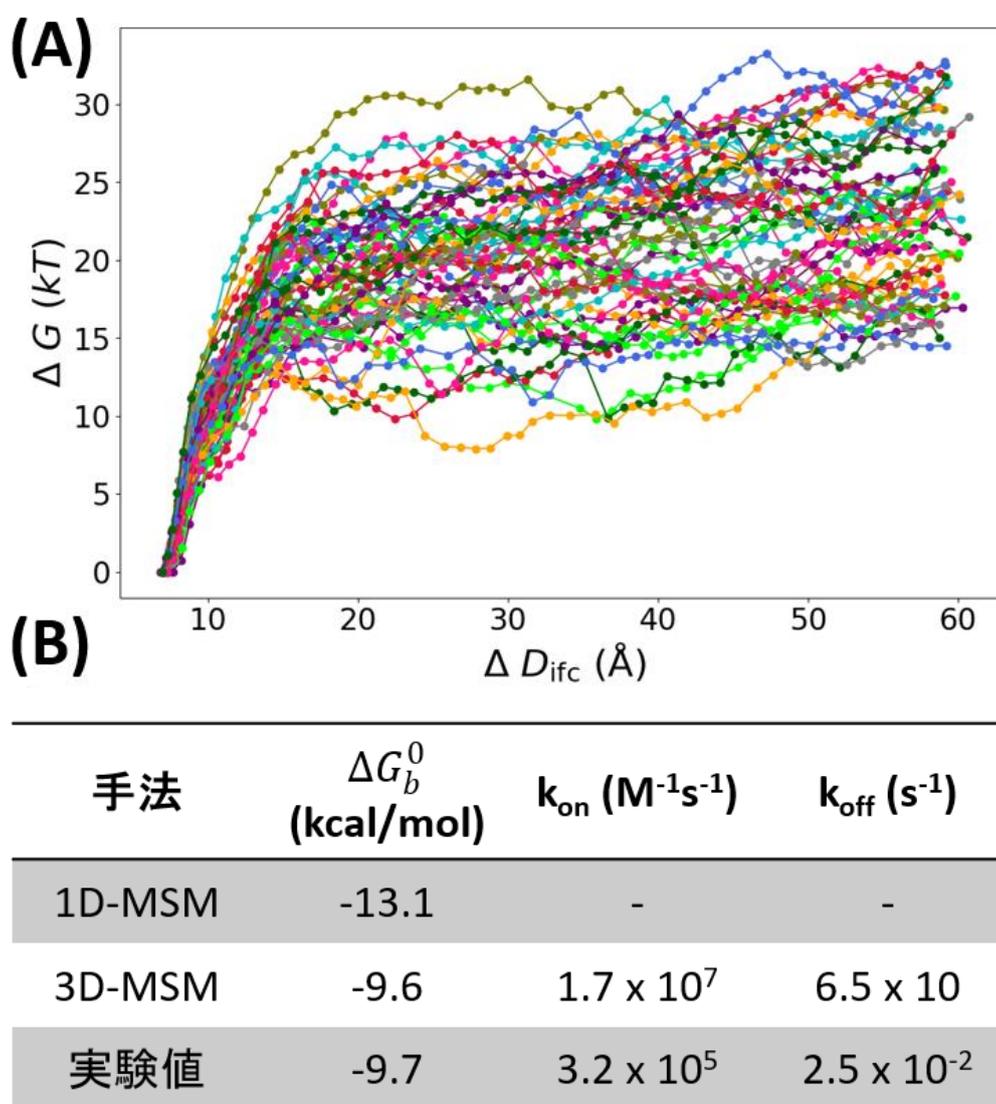
図(2)-10 (A) 解離シミュレーションの例 (B)解離経路 (阻害蛋白質は界面残基の重心を表示)

図(2)-11(A)に示したように、界面間距離が 60 Å に到達するまで必要なサイクル数は 51~109 サイクルと試行ごとに異なり、平均で 66 サイクル要した。図(2)-11(B)に示したように、蛋白質間のコンタクト残基ペア数は全ての試行で 0 に到達しており完全に解離していることが分かる。解離パスウェイによっては、ある程度界面距離が大きい場合でも蛋白質間に接触している残基ペアがあるため、解離に必要なサイクル数が多くなると考えられる。界面間距離に加えて残基ペア数を構造選択時の基準に入れるなど、より効率的な方法が考えられる。しかしながら、複合体の解離に必要なサイクル数が最大の 109 サイクルの場合でもシミュレーション時間は 109 ns (10 並列 x 100 ps x 109 サイクル) であり、複合体の解離速度定数が $2.5 \times 10^{-2} \text{ s}^{-1}$ であることを考慮すると非常に効率が良いことが分かる。以上のように PaCS-MD は蛋白質 - 蛋白質複合体の解離シミュレーションにも有効であることが分かった。



図(2)-11 (A) 界面距離と(B)コンタクト残基ペア数 破線(赤)は初期値

PaCS-MDにより生成したトラジェクトリを用いてMSM解析を行った。各試行より得られる界面間距離を用いてMSM解析(1D-MSM)を行った結果、図(2)-12(A)のように界面距離が30 Å以上では平坦になっている自由エネルギー地形が得られたことから、界面距離30 Å以上を非結合状態と定義した。また、界面距離ではなく、界面間の相対位置を用いてMSM(3D-MSM)解析を行い、標準結合自由エネルギー(ΔG_b^0)、結合速度定数(k_{on})、解離速度定数(k_{off})を評価した。図(2)-12(B)に示したように、結合自由エネルギーに関しては3D-MSMにより実験値と近い値が得られている。速度定数に関しては、大きく外れているが、蛋白質-ペプチド複合体を対象にした際の経験から、より高次元のMSM解析より改善されることが予想される。蛋白質-蛋白質複合体を対象にした場合の、MSM解析に用いる物性や解析におけるクラスター数などの適切な条件を現在検証している。以上のようにMSM解析によって結合自由エネルギー評価を行い、さらに結合速度論解析も開始した。



図(2)-12 (A)1D-MSM 解析より得られたエネルギー地形 (B)標準結合自由エネルギー及び解離定数の比較

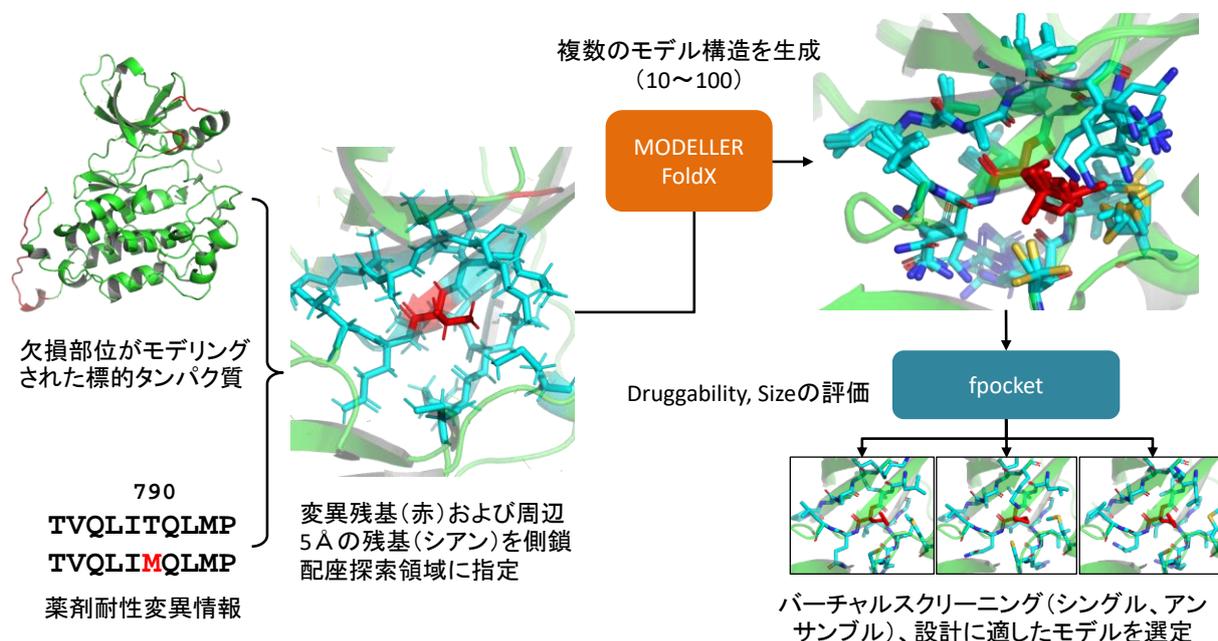
(3) 薬剤分子設計

①低分子デザイン

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 広川貴次

本課題では、ポスト「京」重点課題1で構築した、MD計算と標的タンパク質モデリングおよびドッキング計算法を融合したSBDD技術と化合物生成手法(ChemTS法)によるデノボデザインを可能にする「創薬ビッグデータ統合システム」を用いて、薬剤耐性キナーゼを回避する低分子デザインおよび希少疾患関連タンパク質を標的とした低分子探索とデザインを行う。令和2年度は、薬剤耐性変異タンパク質のモデリング法の改良、希少疾患関連タンパク質の選定を行った。

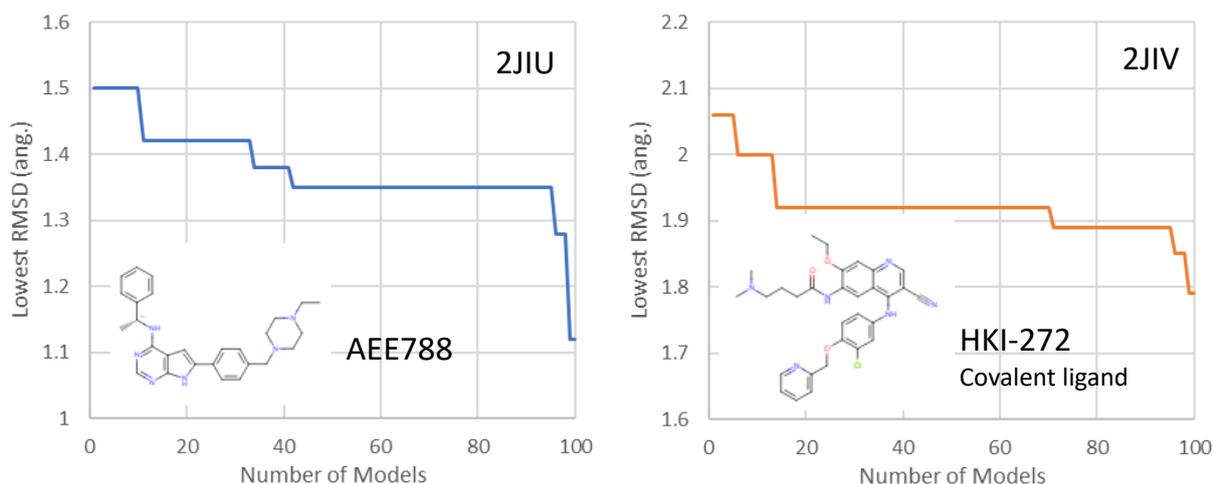
薬剤耐性変異タンパク質モデリング法の改良は、ポスト「京」重点課題1で構築した「タンパク質変異体立体構造予測システム」を基盤に行った。これまでのシステムは、標的タンパク質にdisorder領域が存在する場合、disorder領域を入力PDBファイルから自動で検出し、disorder領域のモデル構造の構築を行う「Disorder領域モデリング機能」、構築したモデル構造または任意の構造のPDBファイルを入力として、任意の変異を加えた変異体モデル構造の構築を行う「変異モデリング機能」、構築したモデル構造または任意の構造のPDBファイルを入力として、任意の配列を欠損・挿入したモデル構造の構築を行う、「挿入・欠損モデリング機能」の3機能から構成され、欠損残基の補完や変異情報を反映したモデルの初期構造構築が目的であった。本年度は、低分子探索を考慮し、変異箇所周辺の構造のアンサンブル構築機能を追加した(図(3)-1)。



図(3)-1 タンパク質の変異体立体構造構築システムの変異箇所構造アンサンブル構築機能拡張

本機能は、目的とする変異残基を基準に周辺5Åのアミノ酸残基に対して初期構造に依存しない複数の側鎖の配座を生成し、アンサンブル構造を構築するものである。側鎖の生成には、MODELLERを利用した。生成されたアンサンブル構造は、薬剤耐性変異に対応した化合物インシリコスクリーニングに用いられることになるが、その際、各アンサンブル構造が形成するポケットのDruggabilityスコアや大

きさ等で鋳型構造の優先順位をつける他、代表的なアンサンブル構造を複数鋳型構造としたアンサンブルスクリーニングの実施が想定される。しかし、それら上記のいずれにおいてもアンサンブル構造の中に、リガンドが結合している状態の側鎖構造が正確にサンプリングできていることが求められる。よって、現状のシステムの精度を評価するためにベンチマークとして、EGFR を例に、野生型構造 (PDB: 4I23) を初期構造に、T790M 変異体を想定した、本システムによる周辺側鎖構造のアンサンブル構造構築を行い、作用機序の異なる 2 つのリガンド (非共有結合リガンド (AEE788、PDB: 2JIU)、共有結合リガンド (HIKI-272、PDB: 2JIV) との Holo 構造との比較を行った (図(3)-2)。アンサンブル構造生成数の探索もかねて行った結果、100 構造サンプリングで、AEE788、HIKI-272 結合型の側鎖構造の RMSD が、それぞれ 1.2 Å、1.8 Å 以下の正解に類似した構造がサンプリングされていたことが確認できた。今後は、現バージョンをプロトタイプとしつつ、構造生成プログラムに FoldX を追加し、更に周辺残基 (5~8 Å) とサンプリング数 (100~200) を最適化することでシステムの改善と網羅的モデリングに適用する予定である。また、化合物インシリコスクリーニングに向けて、アンサンブル構造に対してポケットの Druggability スコアや大きさの計算し、活性化合物を用いてヒット率の濃縮係数との相関解析も行い、化合物スクリーニングに最適化構造選定法も確立させる。



図(3)-2 Wild type EGFR をテンプレートにした T790M モデリングの検証

希少疾患関連タンパク質の選定については、奥野グループと連携し、ベンチマークタンパク質として EGFR を設定した。EGFR 変異リストは、MGEND (日本人のみ) データベースでは登録変異数が少ないため (EGFR 変異 13 種。内訳：遺伝性腫瘍：845 種 (EGFR の変異は 5 個)、希少疾患：547 種 (EGFR の変異は 0 個)、体細胞がん：838 種 (EGFR の変異は 8 個))、登録件数が多い ClinVar (欧米) も考慮し (変異 349 種)、計 362 種がモデリングの対象として選定された。図(3)-3 は、ClinVar のデータ内容

chr	start	end	ref	alt	gene	variant_type	significance	condition	conflict_significance	Ensemble_1	ppos_1	Uniprot_1	PDB_1	Ensemble_2	ppos_2	Uniprot_2	PDB_2
1	899892	899892	C	A	KLHL17	nonsynonymous SNV	Likely_pathogenic	Tracheoesophageal fistula		ENST00000338591	p.A561E	Q6TDP4	6HRL_A,6HRL_B				
1	1147462	1147462	A	G	TNFRSF4	nonsynonymous SNV	Likely_benign	Immunodeficiency_16		ENST00000379236	p.I165T	P43489	2HEV_R,2HEY_R,2HEY_T,6OGX_G,6OKM_R,6OKN_E,6OKN_R				
1	1147484	1147484	C	A	TNFRSF4	nonsynonymous SNV	Uncertain_significance	Immunodeficiency_16		ENST00000379236	p.A158S	P43489	2HEV_R,2HEY_R,2HEY_T,6OGX_G,6OKM_R,6OKN_E,6OKN_R				

図(3)-3 ClinVar データコンテンツ (一部)

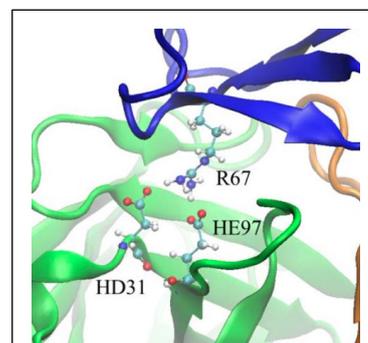
の一部を示しているが、変異体の中には、既に結晶構造が存在することがわかる（図(3)-3：PDB_1、PDB_2列）。今後は、網羅的モデリングの際、構造既知の変異体 EGR においては、解像度や Holo 体情報を考慮して初期構造を選定し、それ以外は、前述の変異体モデリングシステムにより初期構造を構築し、変異部位周辺のアンサンブル構造生成を行う予定である。

②抗体医薬デザイン（東大先端研・山下）

東京大学 先端科学技術研究センター 山下雄史

本研究では、スーパーコンピュータ「富岳」を活用した大規模分子動力学（MD）シミュレーションを活用して、抗体の改変設計とその原理に挑む。抗体は、人間の体において自己と非自己を見分け、非自己を抗原として認識し強く結合することができるタンパク質である。抗体は免疫系で重要な働きを持つだけでなく、医薬品としても注目される。例えば、がん治療に限っても様々な作用機序を持つ抗体ベースの医薬品が登場している。例えば、乳がん治療に使用されるハーセプチン（トラスツズマブ）はHER2に特異的に結合し、抗体依存性細胞傷害作用（ADCC）を引き起こす。また、2018年にノーベル医学生理学賞を受賞した本庶佑博士らが開発したオプジーボ（ニボルマブ）も抗体医薬品である。オプジーボはT細胞に発現するPD-1に結合して、がん細胞が免疫系の攻撃から逃れる作用を抑制する。すなわち、免疫系のブレーキを停止することで免疫系ががん細胞を活発に攻撃できるようにする医薬品となっている。いずれにせよ、抗体が強く抗原を認識することが抗体を医薬品として利用するために重要な条件である。したがって、抗原との親和性を向上させる抗体改変設計は、抗体医薬品開発を効率化する非常に重要な技術になりうる。

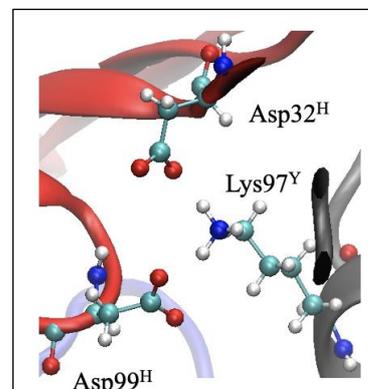
我々は前年度に肝臓がん等で特異的に発現しているROBO1を認識する抗体（B5209B）の親和性を人為的に向上させる研究成果を発表した（Yamashita et al., Structure, 2019）。B5209Bの重鎖P103や軽鎖Y30をアラニンに置換したもの（HP103A、LY30A）で有意な親和性向上を達成した。アラニンの側鎖はメチル基であり特徴的に示す相互作用はないが、HP103Aではエンタルピー的に、LY30Aではエントロピー的に利得があった。MDシミュレーションによる解析によるとHP103A変異では少し離れた場所にある3つ組の塩橋構造（R67-HD31-HD97）、図(3)-4の安定性を改善することで、LY30A変異ではタンパク質界面に固定し



図(3)-4: ROBO1-B5209B 界面に見られる塩橋構造

ている水分子の個数を減らすことで、親和性を改善していることが分かった。これらのアラニン置換の効果は新しい改変設計のアイデアを提供する基盤となりうる。従来の考え方では、構造に基づき特徴的な相互作用をアミノ酸への置換を導入することで相互作用を追加する設計を行うため、アラニンへの置換はまず除外される設計となる。したがって、この研究で見出した3つ組塩橋や界面水は改変設計の成功の鍵となる構造モチーフとみなすことができる。

抗原-抗体間に張られる塩橋の安定性は、親和性の向上を左右する重要な因子であることが分かってきたが、塩橋がどのように安定化されているのかは良く分かっていない。我々は、卵白リゾチーム(HEL)とこれに結合する抗体HyHEL-10が複合体化する際に形成する塩橋に注目し、その安定性の起源を調査した(Okajima et al., J. Phys. Chem. B, 2021)。HEL-HyHEL-10界面にはK97-HD32-HD99という3つ組塩橋（図(3)-5）が存在している。我々が取った手法は、構成論的アプローチと呼ばれるもので、複雑な現象から重要な要因を取り出すことに長けている。ロボット工学においては、人間に近いロボットを組み立ててみることで人



図(3)-5: HEL-HyHEL-10 界面に見られる塩橋構造

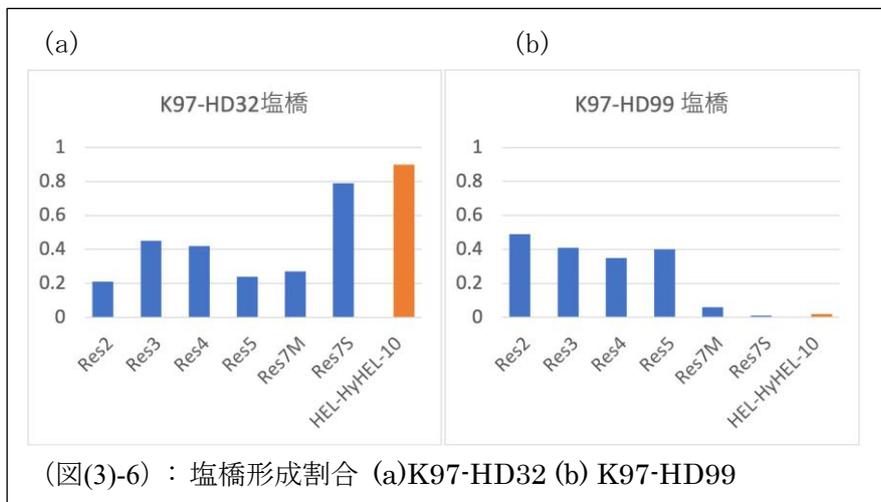
間にとって重要な因子を見いだしていくものとなる。本研究においては、塩橋を持つ様々な要素を逐次的に追加していくことで抗原-抗体の界面における塩橋の安定性が達成されるのかを調べた。

まず、塩橋を形成する残基は、主鎖の形により位置が束縛されている。したがって、まず、最初に2つの塩橋アミノ酸（リジンとアスパラギン酸）の α 炭素（ $C\alpha$ ）を束縛したMDシミュレーションを行った。R67-HD31とR67-HD99の $C\alpha$ 炭素間距離は、それぞれ約10.0と8.5Åである。抗原抗体界面の環境下であれば、R67-HD31が安定であるのに対して、R67-HD99は不安定である。塩橋の形成率はそれぞれ、0.90と0.02である。しかし、上記の2アミノ酸モデル系ではどちらも0.2程度の塩橋形成率であった。さらに、側鎖の向きが主鎖構造により束縛されることを考慮に入れ側鎖の向きを固定した2アミノ酸モデルRes2のMDシミュレーションも行った（図(3)-6）。側鎖の向きが固定される効果は認められたが、抗原抗体界面での塩橋の振る舞いは再現できていなかった。塩橋の安定性には環境の効果が大きいと推定される。

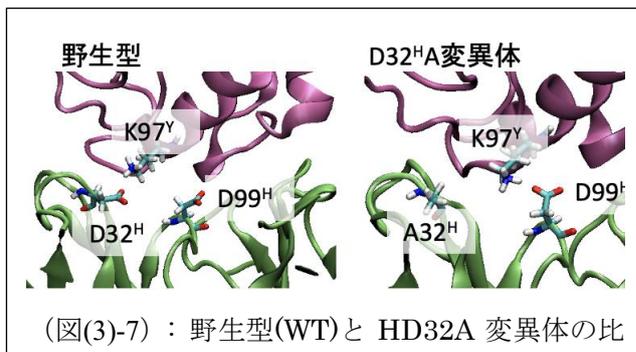
次に、3つ組塩橋の形から3つのアミノ酸の相互作用が重要になる可能性があるため3アミノ酸モデルRes3を構築した（図(3)-6）。さらに、静電相互作用的な影響を持ちうるアミノ酸を加えた4アミノ酸モデルRes4（LK49を追加）、5アミノ酸モデルRes5（HN97を追加）を構築した（図(3)-6）。し

しかし、依然として、抗原-抗体界面での塩橋の振る舞いは再現しきれなかった。そこで、塩橋付近の疎水性アミノ酸残基の影響も調べることにした。HY33とHW98を加えた7アミノ酸モデルRes7Mを導入した（図(3)-6）。これはK97-HD99塩橋の安定性を再現していたが、K97-HD32塩橋の安定性は再現できていない。さらに、不自然に揺らいでいたHY33とHW98の側鎖を固定したモデル(Res7S)を試したところ、両方の塩橋の安定性をうまく再現していた（図(3)-6）。この結果は、疎水性アミノ酸HY33とHW98が塩橋安定性に重要であることを示している。また、これらのアミノ酸側鎖が背後に位置するアミノ酸残基によって適切な位置に固定されていることも重要であることを意味する。このような適切な位置に配置する“アミノ酸の社会学”を知ることが塩橋安定性に重要であり、ひいては抗原-抗体間相互作用の強さに重要であることが分かった。このようなアミノ酸同士の相互作用の総体を知るにはMDシミュレーションが強力な道具となると考えられる。

以上のような構成論的なモデル解析を用いると、変異効果の詳細も解き明かすことができることが分かった。例えば、HEL-HyHEL-10系で強い塩橋を形成しているHD32をアラニン置換（HD32A）することを考える（図(3)-7）。実験では、HD32A変異による親和性低下が予想に反し

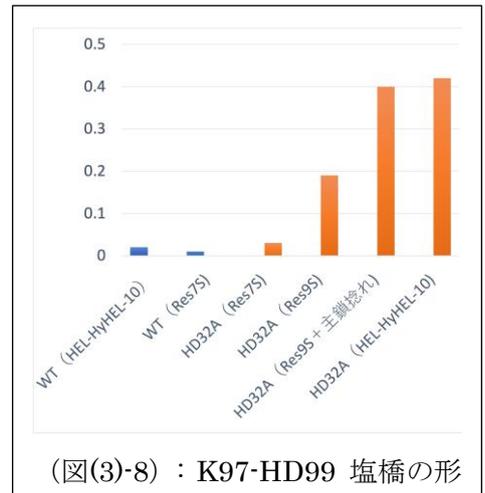


(図(3)-6)：塩橋形成割合 (a)K97-HD32 (b) K97-HD99



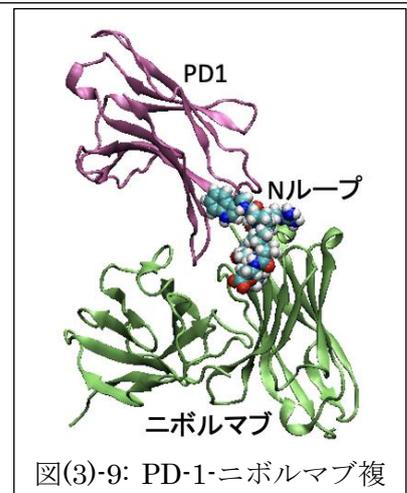
(図(3)-7)：野生型(WT)と HD32A 変異体の比

て小さい。その原因は K97-HD99 塩橋の安定化であることが MD シミュレーションによる解析で分かった。では、どうして K97-HD99 塩橋が安定化できたのかという疑問が残る。もし HD32 からの静電的反発が失われたことによるのであれば、上記の Res7S モデルでも同様の現象が起きるはずである。しかし、Res7S モデルの HD32A 変異体では K97-HD99 塩橋はほとんど安定化しなかった。この安定化を支えるメカニズムを構成論的に調べたところ、2つの要素が重要であることが分かった。第1の要因は L75 と HD101 という他のアミノ酸がさらに関与していたこと、第2の要因は変異により主鎖の捻れが生じていたことであった (図(3)-8)。我々は、このように構成論的な簡易塩橋モデルを用い、直感的に考えると誤解を招く複雑現象の正確な解析に成功した。



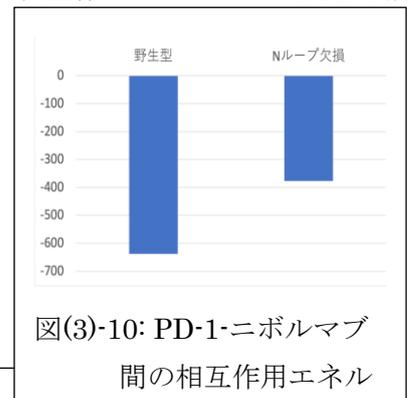
(図(3)-8) : K97-HD99 塩橋の形

さらに、MD シミュレーションを用いて、免疫チェックポイント阻害剤として使われる抗 PD-1 抗体ニボルマブを用いて、抗原-抗体相互作用を自在に変化させる設計を考える。今年度は、抗原 (PD-1) 側の変異により相互作用が減退する可能性を調査した。PD-1 の N 末端側には N-loop と呼ばれる構造があり、その部位がニボルマブとの結合で重要な役割を果たしている。実際に、N-loop を欠損させると親和性が大きく低下することが実験で示されている。まず、図(3)-9 のように N-loop を含んだ系と N-loop を欠損させた系の MD シミュレーションを実施した。相互作用エネルギー解析から、N-loop 欠損により相互作用エネルギーが大きく減



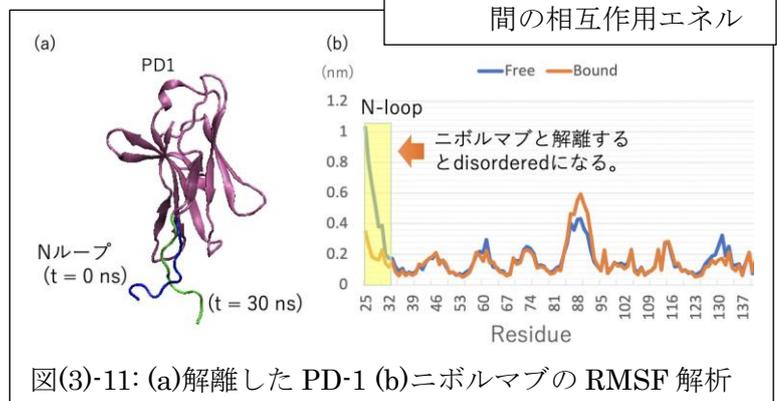
図(3)-9: PD-1-ニボルマブ複

退することが分かった (図(3)-10)。これは実験結果と矛盾しないものである。また、解離状態での PD-1 を MD シミュレーションしたところ、この N-loop は大きく揺らいでおりディスオーダー領域になっていた (図(3)-11)。したがって、ニボルマブと結合することではじめて、N-loop は図(3)-9 のような特定の構造に固定化されることになる。



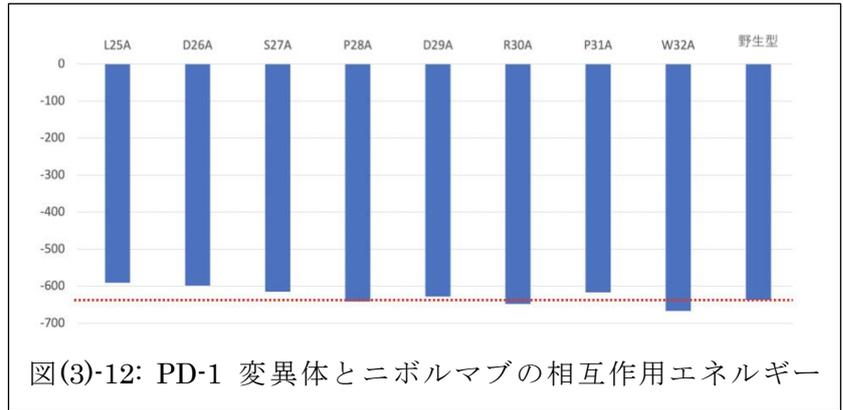
図(3)-10: PD-1-ニボルマブ間の相互作用エネルギー

この N-loop での相互作用をより深く理解するために、in silico アラニンスキャンを行った。具体的には、各 N-loop アミノ酸残基をアラニン置換した構造を計算機で生成し、さらに MD シミュレーションを実施した。そのシミュレーションを元に相互作用エネルギー解析をした結果を図(3)-12 に示す。L25A・D26A・S27A の変異は有意に相互作用を低下させていることが分かった。これは L25・D26・S27 がニボルマブとの重要な界面を形成していることを意味し、複合体構造



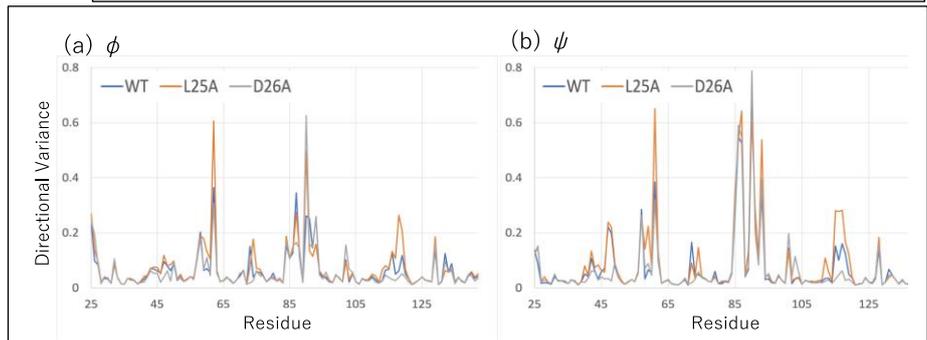
図(3)-11: (a)解離した PD-1 (b)ニボルマブの RMSF 解析

からの結論とも合致する。さらに興味深いことに、P31 と W32 はニボルマブとほとんど接していないが、アラニン置換することで相互作用エネルギーを優位に変化させている。特に、W32A 変異は相互作用エネルギーを増大する傾向を持っていた。

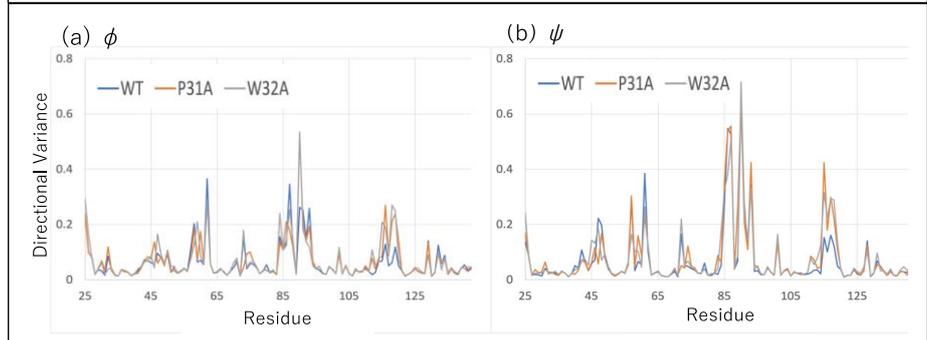


図(3)-12: PD-1 変異体とニボルマブの相互作用エネルギー

これらの変異がタンパク質構造にどのような影響を与えているのかを詳しく調べるために、主鎖の ϕ 角および ψ 角に対し角度統計解析 (Yamashita, AIP Conf. Proc. 2016, Miyanabe et al., Biochemistry, 2018) を行った。その結果、アミノ酸 1 つの置換だけでもタンパク質全体に渡って、ダイナミクス・平均構造が変化していることが分かった (図(3)-13、図(3)-14)。特に、60 番目のアミノ酸残基付近で有意な影響が見られるがこの領域もニボルマブとの界面の一部である。



図(3)-13: L25A および D26A 変異体の角度分散



図(3)-14: P31A および W32A 変異体の角度分散

こうした影響が相互作用エネルギーの変化を生み出していると考えられる。我々は、このような構造・ダイナミクスの変化を MD 計算と機械学習の組み合わせによってある程度予測できると考えている。すでに、いくつかの手法で成果が出始めている (Yamashita and Shoji, AIP Conf. Proc. 2021, Yamashita et al., AIP Conf. Proc. 2018)。

抗体は薬物との複合体 (ADC) というアイデアでさらに応用範囲を広げている。例えば、トラスツズマブをベースにしたカドサイラはすでに承認されている。セツキシマブをベースにしたアキシャルクスは光免疫療法を開拓している。我々が開発してきたシミュレーション技術はこの分野でも新たな特許の創出を数多く支援している (Sugiyama et al., Proc. Jpn. Acad. Ser. B, 2019)。また、我々のシミュレーション技術はタンパク質間相互作用に普遍的に利用できるものである。現在大きな社会問題となっている新型コロナウイルスと ACE2 の相互作用を調査することにも活用されている。

(4) AI・データサイエンス・社会実装基盤

①疾患ゲノム構造機能データベース

京都大学大学院 医学研究科 鎌田真由美

令和2年度は、臨床ゲノム情報統合データベース Medical Genomics Japan Variant Database (MGeND) に登録されている遺伝子多型・変異(バリエント)を対象に、計算候補となるバリエントリストを作成した。具体的には、ゲノムポジションで示されるバリエント情報を、トランスクリプトレベル、タンパク質立体構造レベルにマッピングするアノテーションパイプラインを構築し、Protein Data Bank (PDB) に立体構造情報が登録されておりかつ非同義置換の対象となるバリエントをそれぞれのデータベースから選出した。以下に、対象バリエントおよびアノテーションパイプラインの詳細、そしてアノテーション結果について記載する。

(1) 対象バリエント

MGeND (<https://mgend.med.kyoto-u.ac.jp/>) は、日本医療研究開発機構 (Japan Agency for Medical Research and Development, AMED) による「臨床ゲノム情報統合データベース整備事業 (以下、本事業)」のもと、「がん」「稀少・難治性疾患」「感染症」「認知症」「難聴」の疾患領域を対象に、臨床ゲノム解析を実施する国内の研究班、そして京都大学・国立国際医療研究センター・慶應義塾大学との協働で、開発した日本人疾患ゲノム情報統合データベースである。2018年3月の運用以来、20万以上のバリエント情報を公開している。

本年度の計算ターゲット策定では、検証可能な事例を有する「稀少・難治性疾患領域」、「がん領域」を対象とした。MGeNDには2020年6月時点で、稀少・難治性疾患には14,878バリエント、がん領域では157,456バリエントが登録されている。これらには登録者指定による未公開バリエントも含むため、本年度の策定では、公開されている、稀少・難治性疾患の3,266バリエント、がん領域(体細胞・生殖細胞の区別なし)145,662バリエントを対象とした。

(2) アノテーションパイプラインとバリエントリストの作成

バリエントリストの作成は、(i) 構造情報へのマッピング、(ii) 構造および結合情報の詳細アノテーションの2段階で実施した。以下に、具体的な流れと結果を記載する。

(i) 構造情報マッピング

データベースに登録されているバリエントはゲノム位置で表現される。そこで、アノテーションソフトウェア ANNOVAR (Wang K. et al., Nucleic Acids Research, 38:e164, 2010) を用い、全バリエント一様にトランスクリプトレベルの情報 (Ensemble transcript ID) を付与する。Ensemble Transcript IDをもとに、EMBL-EBI UniProt (<https://www.ebi.ac.uk/uniprot/>) から取得可能な ID 相互連携情報 (ID Mapping) を用いてタンパク質レベルの情報 (UniProt ID) にマッピングし、構造情報 (PDB ID) との紐付けを行った。ここで、1つの遺伝子に対して複数トランスクリプトが存在する場合、UniProtにおいて配列長などをもとに規定される代表的なエントリ (canonical isoform) を採用した。また、PDB構造情報において、対象タンパク質全てのアミノ酸領域を含んでいない(一部構造のみ構造決定されている)場合がある。そこで、

EMBL-EBI により提供される PDB 配列と UniProt 配列のアラインメント情報” Structure integration with function, taxonomy and sequence (SIFTS)” を用い、該当変異が PDB 構造内に含まれるかを確認した上で紐付けを行った。

構造マッピングを行なった結果を【表(4)-1】に示す。稀少・難治性疾患領域は 3,266 バリエントのうち 683 バリエントが、がん領域は 145,662 バリエントのうち 13,691 バリエントが、立体構造情報へ紐付けすることが可能であった。

【表(4)-1】 構造情報マッピング結果

	All Cancer	Rare Disease
Total variants	145, 662	3, 266
Unique variants	54, 716	2, 170
Exonic variants (without synonymous)	28, 303	1, 835
Variants with PDB ID	13, 691	683

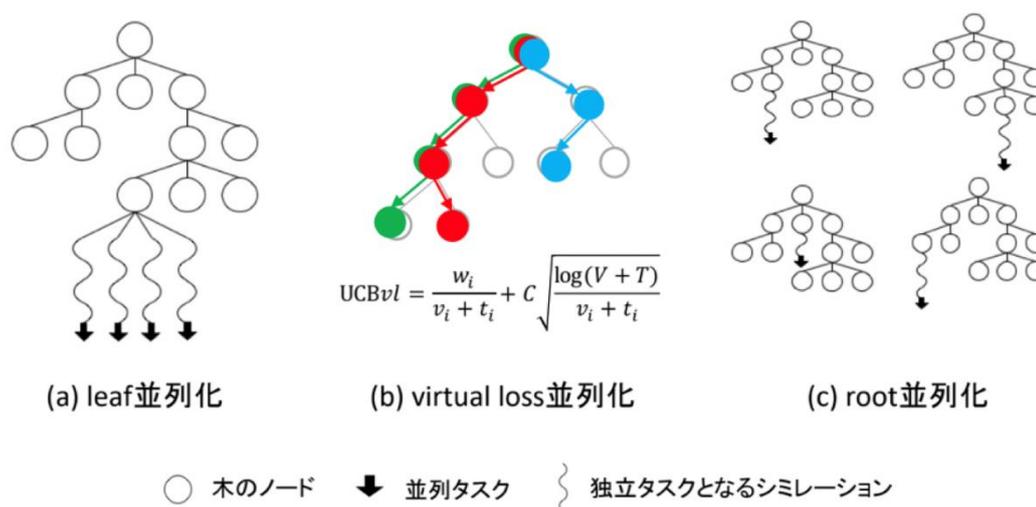
(ii) 構造詳細アノテーション

さらに、変異体構造モデリングおよび検証で必要となる、詳細な構造および結合情報を抽出する。まず、構造の詳細情報として PDB より提供される mmCIF ファイルより、解像度や修飾アミノ酸情報など 21 項目を抽出する。また、薬剤感受性計算では化合物との複合体構造を対象とすることから、抽出した PDB から化合物との結合詳細情報を抽出する。公共データベース BindingMOAD (<http://bindingmoad.org/>)、PDBBind (<http://www.pdbbind.org.cn/>)、BindingDB (<https://www.bindingdb.org/bind/index.jsp>)を対象に、IC₅₀をはじめとした化合物との結合詳細情報 16 項目を抽出し、先のマッピングで得られた PDB に紐付けを行う。【表(4)-1】にアノテーション結果を示す。(i) で抽出した、がん 13,691 バリエント、稀少難治性疾患 683 バリエントのうち、各 5,611, 185 バリエントが複合体構造に紐付き持ち且つ結合情報を持つことがわかった。次年度以降さらに詳細な検討を行い、検証用バリエントと計算用バリエントの絞り込みを行う。

【表(4)-1】 構造詳細アノテーション結果

	All Cancer	Rare Disease
Total Variants (with PDB IDs)	13, 691	683
Variants with binding data	5, 611	185
Variants with K _i	1, 965	98
Variants with IC ₅₀	2, 531	45
Variants with K _d	4, 281	96
Variants with EC ₅₀	782	14

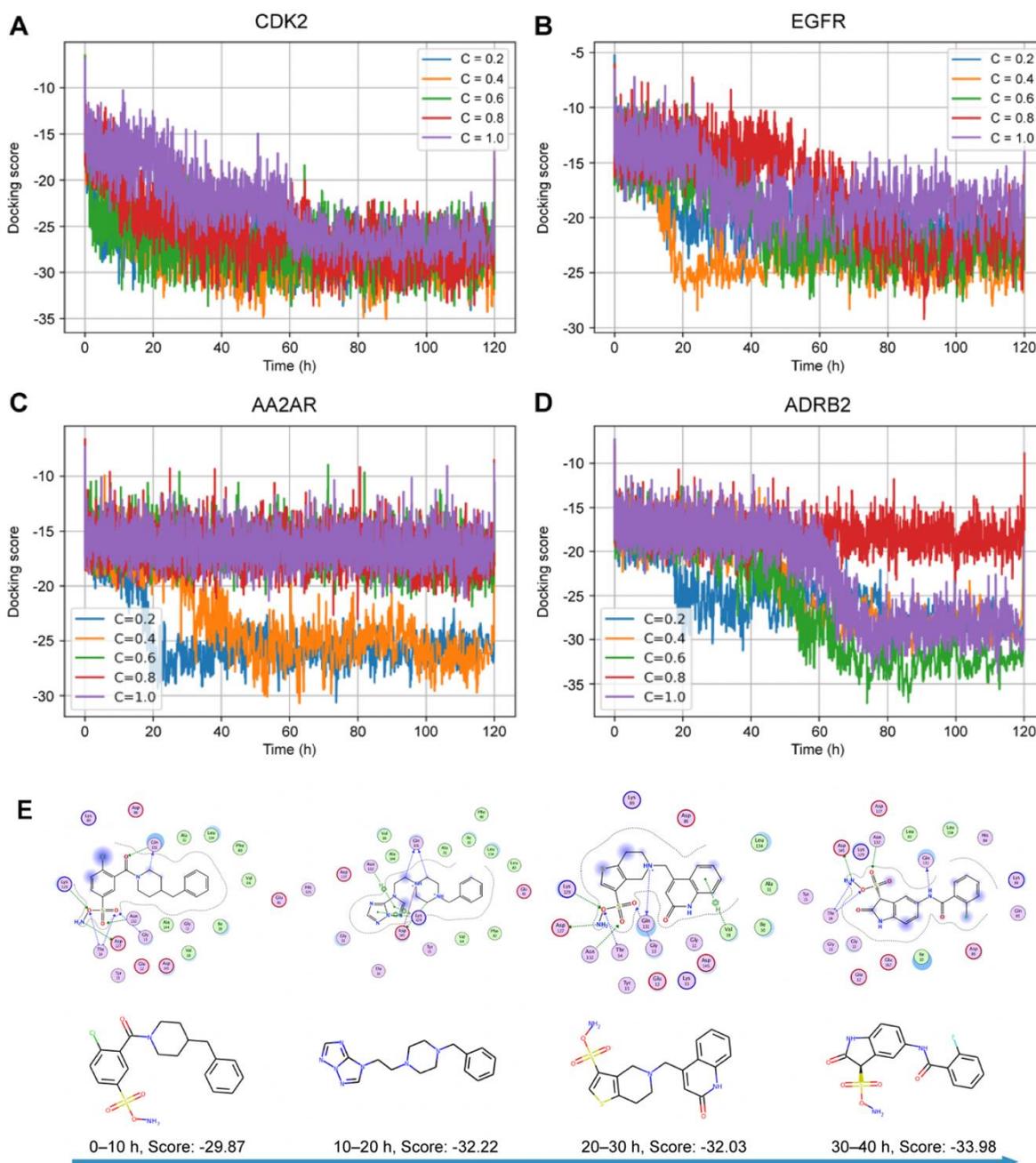
シミュレーションは計算コストが高いため、シミュレーションと連携して分子デザインを実施するためには、計算の効率化が不可欠である。本年度は ChemTS の高速化として、Leaf 並列、Virtual loss 並列、Root 並列、の三つの並列化実装に取り組んだ(図(4)-1)。Leaf 並列はモンテカルロ木探索におけるロールアウト部分の並列化である。Virtual loss 並列は、シミュレーションで評価するノードを virtual loss と呼ばれるスコアの工夫によってうまく分散させる方法である。Root 並列は探索木自体を複数用意することで並列化を図る手法である。特に Leaf 並列は単純な方法ながらドッキングシミュレーション(rdock)のように一つのリガンドに対して複数の計算が必要な際には有効になる。後述の分子発生実験では Leaf 並列を行うことで 120 時間、10CPU の並列計算によりドッキングシミュレーションと合わせつつ 2.5 万程度の分子生成に成功した。今後手法による分子生成の傾向の違いなどを検証する予定である。



図(4)-1 実装した ChemTS の並列化手法。(a) Leaf 並列、(b) Virtual loss による並列、(c)Root 並列による並列化を示している。

シミュレーションと連携した分子デザイン手法の改良と検証のために、ChemTS とドッキングシミュレーション(rdock)及び合成可能性等を考慮した手法の開発及び評価を実施した。CDK2、EGFR、AA2AR、ADRB2 の 4 種類のタンパク質をターゲット、rdock の計算値を含むスコアを評価関数として分子生成を実施した。また現実的な分子を生成するために、合成容易性(SA score)や既知化合物における局所パターンの出現頻度に基づくフィルター等も考慮したスコアを設計した。ChemTS の探索パラメータ C を 0.2 から 1 まで変化させつつ 120 時間分子生成を行なった結果が図(4)-2(A)-(D)である。時間経過に従って全体的にドッキングスコアが落ちており、これはターゲットタンパク質とより親和性の高いと期待される分子が生成されていることを示している。また図(4)-2(A)-(D)から、理論通り探索パラメータを小さくするほど短時間で収束する傾向も確認できる。一方探索パラメータを大きくすると収束は遅い傾向がある。生成した化合物を可視化したところ(図(4)-3)、C が大きいほど探索空間は広い傾向があり、状況に応じて適切に探索条件を変える必要があ

ることを確認できた。さらに、実際に生成された分子の例を図(4)-2(E)に示す。ドッキングシミュレーションと組み合わせることにより単にスコアだけでなくタンパクとの相互作用を図 2(E)のように考慮できることが本手法の強みである。これらの結果については論文にまとめ現在投稿中である。



る。

図(4)-2 4つのターゲットタンパク質に対する分子生成の結果。それぞれ120時間の計算で探索パラメータCを0.2から1.0まで変化させてドッキングシミュレーション(rdock)の結果を評価関数に組み込み分子生成を実施している。EはCDK2に対して生成した分子の例と、ドッキングシミュレーションで最もスコアが高かったコンフォメーションとタンパク質残基間の相互作用の様子を示している

図(4)-3 生成した分子を PCA で可視化した結果。灰色の点は、ChemTS において分子生成に用いる RNN の訓練データに用いた分子を示している。探索パラメータとターゲットによるものの探索の初期段階では訓練データに近い分子を生成しているが、探索が進むにつれて徐々に新規な分子を生成している様子が確認できる。

また、分子シミュレーションを効率化するために、シミュレーションとグレイボックス最適化 (Gray Box Optimization, GBO) や COMBO を組み合わせた手法開発を実施した。まず、分子の安定構造を GBO を用いて効率的に推定する手法を開発した。分子の安定したコンフォメーションを推定する方法として、多数のコンフォメーションを発生させ、それぞれを量子化学シミュレーション等で構造緩和し、最安定構造を得る方法が存在する。しかし全ての候補に対して構造緩和を実施すると多大な計算コストがかかる。そこで GBO に基づいて、不適な構造候補の計算を途中で停止し、有望な候補に計算を集中することで効率化を図る手法を構築した。ジペプチドやミコフェノール酸等創薬の対象になりうる分子に対して提案手法を適用したところ、全候補を緩和する方法に比べておよそ 100 分の 1 程度の計算ステップ数で最安定構造を発見できた。これらの結果は論文にまとめ現在投稿中である。また、粗視化分子動力学シミュレーション(CGMD)において、効率的に大きな構造変化を再現する手法についても開発を行った。通常 CGMD でタンパク質をシミュレーションするには実験的に得られた結晶構造に基づく力場を設定することで、結晶構造の周りの運動を再現できる。一方より大きな構造変化を実現するためには、異なる力場の設定が必要になる。本年度は様々な力場構築のパラメータに対してどのような構造変化が起こるかを網羅的に検証し、適切なパラメータを探索する機械学習手法(COMBO によるベイズ最適化・能動学習)について検討を進めた。来年度、シミュレーション結果とパラメータ探索の効率化をより詳細に検証し、論文としてまとめる予定である。

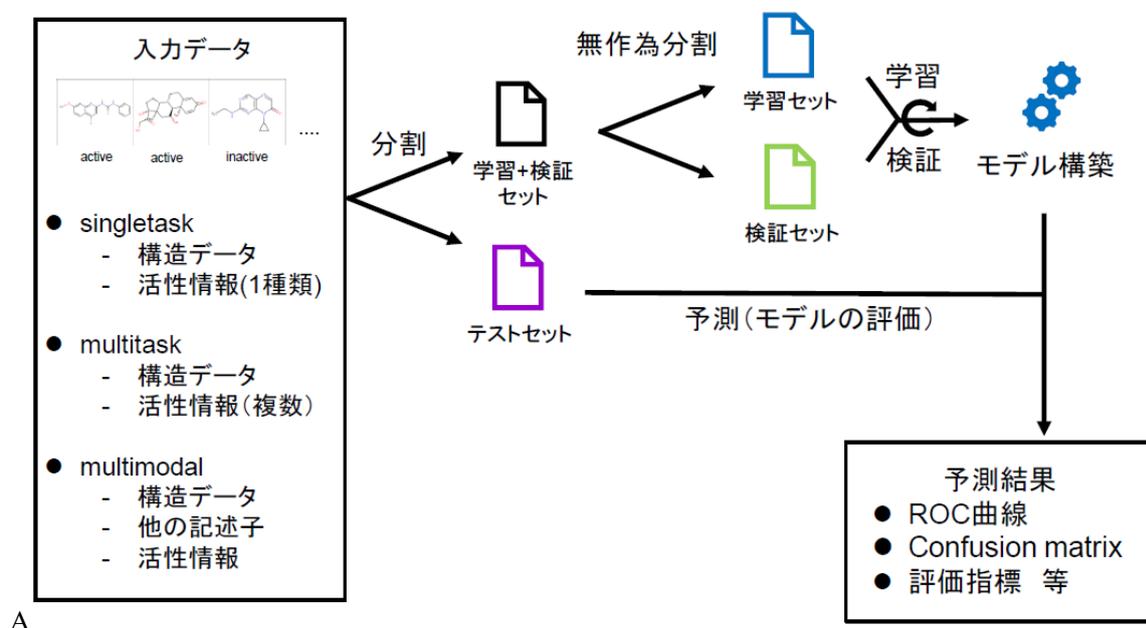
③創薬ビッグデータ統合システム

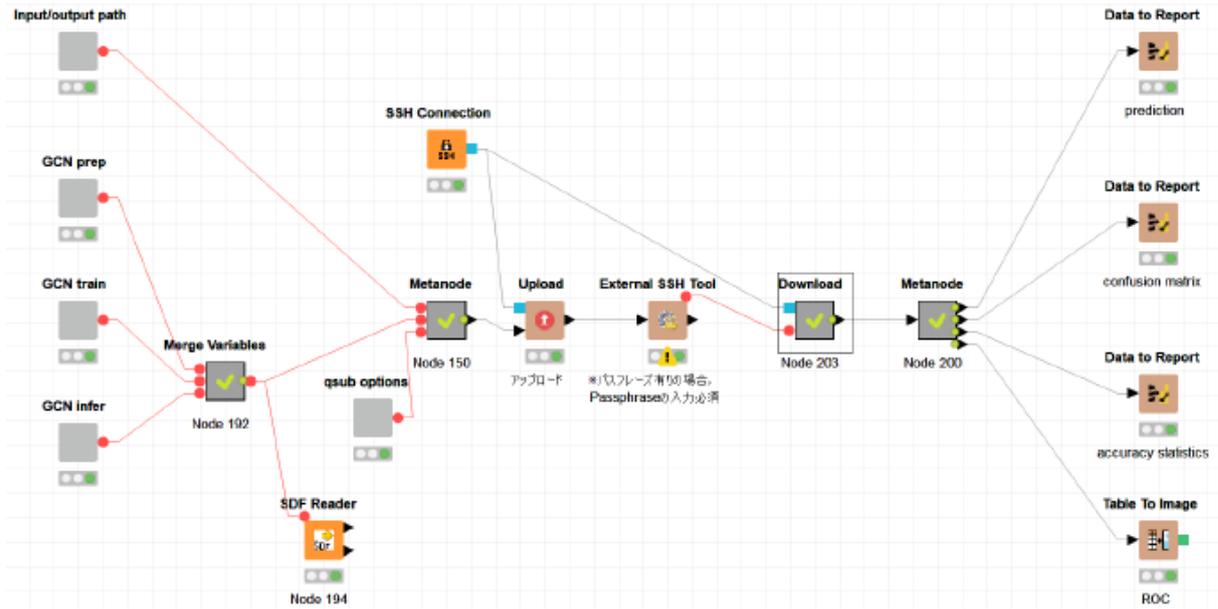
理化学研究所 医科学イノベーションハブ推進プログラム 本間光貴

今年度は、KNIME 上で京都大学と連携して開発している k-GCN に基づく AI 予測モデル構築のワークフローを作成した。また、シミュレーションに関しては、量子力学効果を考慮した AI 力場 (FMO-AI 力場) のためのデータ収集を行っており、複数のタンパク質や環状ペプチドの系におけるそれぞれ 1000 件以上の MD トラジェクトリの FMO 計算を実施した。環状ペプチドについては、指定したペプチドの計算が自動的に実施できるスクリプトを開発した。これらの成果として、世界で初めての量子力学計算による電荷 (FMO 電荷) を高精度で予測できる AI を開発して出版した。以下に詳細を述べる。

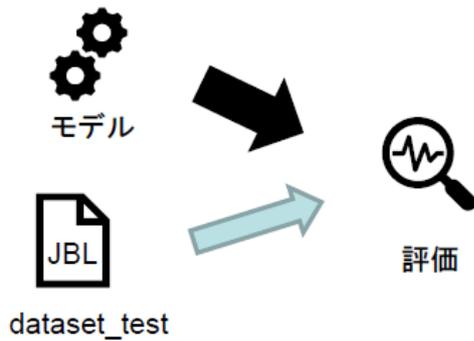
(1) KNIME 上で AI 予測モデルを構築するワークフロー

AI 予測モデルの構築と実行について、AI 開発者がコマンドラインで実施するだけでなく、創薬研究のエンドユーザーが GUI を通して実施できるように KNIME 上でインターフェイスを作成した。KNIME は、計算研究者が行うコマンドライン操作のフローを可視化し、フローの修正や、フローの簡単な GUI を効率的に構築することができるソフトウェアである。今回、図(4)-4A に示すフローについて図(4)-4B に示す KNIME ワークフローを作成した。エンドユーザーは、このフローによる GUI から、モデルの構築、予測の実行を実施でき、図(4)-4C に示す予測性能指標や図(4)-4D に示す予測結果における寄与度も確認できる。





B



予測値

prediction > 0.5 ⇒ 1
prediction ≤ 0.5 ⇒ 0

Row ID	I name	I true_label	I predicted_label	D prediction	S smiles
0	11112411	0	0	0.937	c1ccc2c(c1)C(=O)N(C)CC2
1	11111245	0	1	0.51	CCOC1=CC=C(C=C1)C2=CC=CC=C2
2	4239914	0	1	0.54	COc1ccc2c(c1)nc3cc(C)cc3
3	4254362	0	0	0.304	Nc1nc(N)c(O)c2ccnc2c1
4	4237787	1	1	0.573	O=C1C=CC(=O)N1

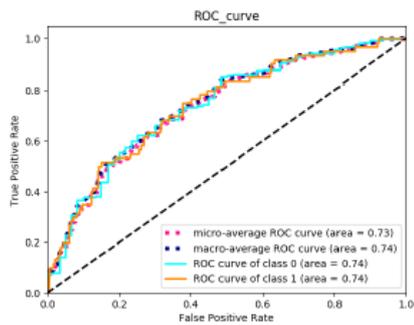
混同行列

Row ID	I Predict (Positive)	I Predict (Negative)
Actual (Positive)	41	19
Actual (Negative)	100	199

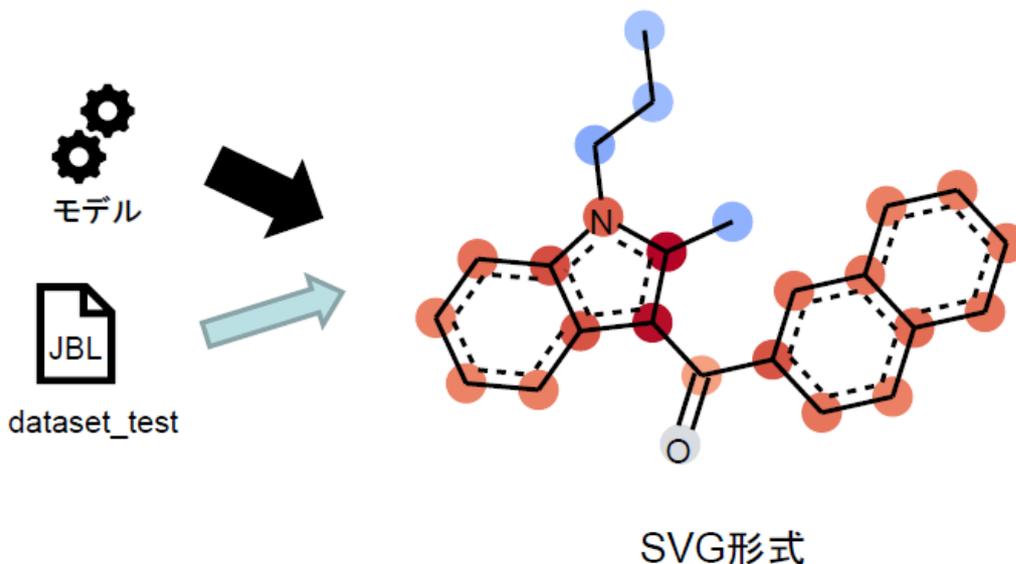
性能指標

Row ID	D Coll
True_Positive(TP)	41
True_Negative(TN)	199
False_Positive(FP)	100
False_Negative(FN)	19
Sensitivity(Recall,TPR)	0.683
Specificity(Selectivity or TNR)	0.656
Precision(PPV)	0.291
Negative_predictive_value(NPV)	0.913
Miss_rate(FNR)	0.317
Fail-out(FPR)	0.334
False_discovery_rate(FDR)	0.334
False_omission_rate(FOR)	0.087
Threat_score(TS,Critical_success_index)	0.256
Accuracy(ACC)	0.659
F1-score	0.408
Matthewd_correlation_coefficient(MCC)	0.257
Informedness(Bookmaker_informedness)BI	0.349
Markedness(MK)	1.204
Balanced_ACC	0.674
Average_prediction_score(AP)	0.252
Cohen_Kappa	0.227

ROC曲線



C



D

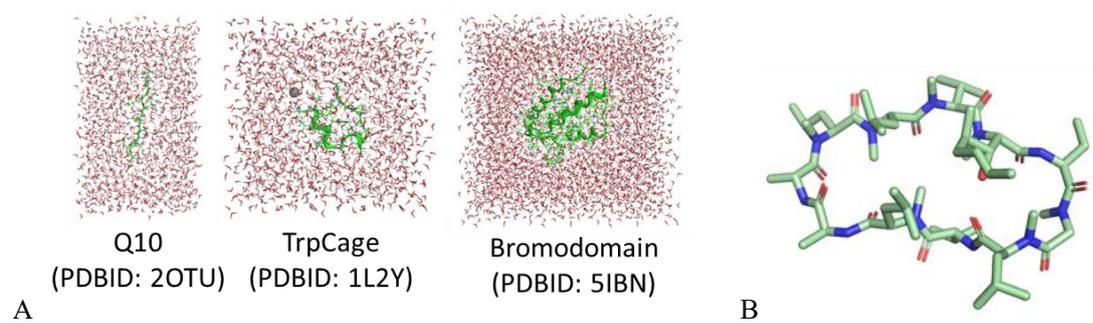
図(4)-4 KNIME 上で k-GCN による DNN 予測モデルを構築するワークフロー

(2) AI 力場構築のための量子力学計算 (FMO) データの収集

タンパク質を含む系の量子力学計算は、FMO 法などの手法を用いて並列計算を行っても数時間程度の計算時間がかかる。一方、既存の AMBER ベースの分子力場では、分散力は角度の要素を含まない Lenard-Jones 型のポテンシャルであり、 π 電子に代表される分子軌道同士の相互作用を適切に考慮することができない粗いものである。研究室では、計算速度と量子力学的な相互作用の考慮を両立できる AI 力場の構築を目指して、集中的な FMO 計算の実施によるデータの収集を行っている。今年度は、以下の図(4)-5A に示す 3 種のタンパク質について、それぞれ 100 ns の MD 計算を行い、得られたトラジェクトリからそれぞれ 1 万構造を抽出し、FMO 計算を実施した。計算結果の一部は、世界初のタンパク質の量子化学計算値データベースである FMO DB

(<https://drugdesign.riken.jp/FMODB/>) で公開している他、FMO 電荷を予測する AI の構築に利用され、成果を出版した。(J. Chem. Inf. Model. 2020, 60, 7, 3361–3368)

また、従来の分子力場では、タンパク質と低分子については、力場パラメータ決定のための実測値が多いため、ある程度の精度が担保されているが、環状ペプチド等の中分子については、適切な力場パラメータの設定が難しい。近年注目されている医薬品モダリティである中分子のシミュレーションの精度を向上させるために、図(4)-5B に例示される環状ペプチドについて、ユーザーが指定した配列から構造を立ち上げて配座探索を行い、各配座に対して FMO 計算を実施するスクリプトを作成した。



図(4)-5 FMO-AI 力場構築のためのデータ収集

(5) 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の会議を月1回開催した。プロジェクト内の研究の進捗状況および成果の発表のために、プロジェクト内の研究ワークショップ(キックオフミーティング)を2020年8月6日に開催し、産業界、臨床研究者も参加する「富岳」創薬活用ワークショップを2021年3月4日に開催した。プロジェクトで得られた成果については学会発表等により、積極的に公表を行った。また、他のプロジェクトとの連携などにより、効率的・効果的な研究の推進を行った。具体的には、「富岳」成果創出加速プログラム「全原子・粗視化分子動力学による細胞内分子動態の解明」が主催する国際ワークショップ "The 1st Fugaku Bio-supercomputing Workshop on Cellular-scale Molecular Dynamics Simulations" (2021年1月6日、7日開催)の共催などを行った。さらに共同で GENESIS を用いた分子動力学計算を推進した。

2-3. 活動(研究会の活動等)

2020年度は、以下の(1)プロジェクト・キックオフミーティング、(2)スーパーコンピュータ「富岳」創薬活用ワークショップを開催した。

(1) プロジェクト・キックオフミーティング

【概要】関係者による研究計画に基づくプロジェクトの進め方、「富岳」利用、ツール利用などの情報交換を行った。

【日時】2020年8月6日(木) 10:00-12:30

【参加者】本プロジェクト参加者 41名

【開催方法】WEB会議

【プログラム】

10:00-10:30 プロジェクト全体概要説明(理研/京大 奥野恭史 教授、本プロジェクト代表)

10:30-11:00 「富岳」での GENESIS の実装状況と

新型コロナウイルスタンパク質の分子動力学シミュレーション(理研 杉田有治 チームリーダー)

11:00-11:15 疾患変異データベースの紹介(京都大 鎌田真由美 准教授)

11:15-11:45 各先生の計算題材・シミュレーション内容について

(全体概要、疾患変異データベースの説明を受けて質疑応答)

11:45-12:00 計算機の使用ルールについて(理研/京大 奥野恭史 教授)

12:00-12:30 「富岳による新型コロナウイルスの治療薬候補同定」の紹介(理研/京大 奥野恭史 教授)

以上

(2) スーパーコンピュータ「富岳」創薬活用ワークショップ

【概要】「富岳」供用開始に向け、プロジェクトの研究内容について説明を行い、関連する研究機関、企業関係者と共に「富岳」の創薬分野での活用について意見交換を行った。

【日時】2021年3月4日

【参加登録者】172名(ライフ系企業82名、IT系企業30名、大学・研究機関54名、官公庁6名)

【開催方法】 WEB 会議

【プログラム】

15:00-15:10	1	全体概要	理研/京大 奥野恭史 教授 (本プロジェクト 研究開発課題責任者)
15:10-15:30	2	「富岳」の活用事例	
		(1) 富岳コロナ対策プロジェクト	理研/京大 奥野恭史 教授
		(2) 超並列分子動力学ソフトウェア GENESIS の紹介	理研 杉田有治 チームリーダー
15:30-16:10	3	研究課題紹介 (本プロジェクト 研究課題代表より)	
		(1) 分子病態解析	横浜市立大 池口満徳 教授
		(2) 薬剤反応性推定	京都大 荒木 望嗣 准教授
		(3) 薬剤分子設計	筑波大 広川貴次 教授
		(4) AI・データサイエンス・社会実装基盤	京都大 鎌田真由美 准教授
16:10-16:30		質疑応答	
16:30-16:45	4	臨床研究サイドから「富岳」に期待する事	国立がん研究センター 河野隆志 分野長
16:45-17:00	5	産業分野からの「富岳」に期待する事	アステラス製薬 デジタルテクノロジーグループ 上村 丈二 課長
17:00-17:10		質疑応答および意見交換	
17:10-17:15	6	閉会挨拶	理研/京大 奥野恭史 教授

以上

2-4. 実施体制

業務項目	担当機関	担当責任者
(1) 分子病態解析 ①動的構造機能解析	公立大学法人横浜市立大学 生命医科学研究科	池口 満徳 教授
②タンパク質活性予測	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 教授
(2) 薬剤反応性推定 ①結合自由エネルギー計算	国立大学法人京都大学 大学院 医学研究科 人間健康科学系専攻 ビッグデータ医科学分野	奥野 恭史 教授
②結合経路・ポーズ推定	国立研究開発法人理化学研究所	杉田 有治 チームリーダー

③結合速度論解析	国立大学法人東京工業大学 生命理工学院	北尾 彰朗 教授
(3) 薬剤分子設計 ①低分子デザイン	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	広川 貴次 研究チーム長
②抗体医薬デザイン	国立大学法人 東京大学 先端科学技術研究センター	山下 雄史 特任准教授
(4) AI・データサイエンス・社会実装基盤 ①疾患ゲノム構造機能データベース	国立大学法人京都大学大学院医学研究科 人間健康科学系専攻 ビッグデータ医科学分野	鎌田真由美 准教授
②AI・シミュレーション融合	公立大学法人横浜市立大学 生命医科学研究科	寺山 慧 准教授
③創薬ビッグデータ統合システム	国立研究開発法人理化学研究所 医科学イノベーションハブ推進プログラム医薬プロセス最適化プラットフォーム推進グループ	本間 光貴 副グループディレクター
(5) 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 医科学イノベーションハブ推進プログラム	奥野 恭史 副プログラムディレクター

別添 1 学会等発表実績

※別紙に記載

学会等発表実績

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	East Asian population-specific variants and cross-disease variants: Detailed analysis based on MGeND records and database improvement	Rei Ashine, Mayumi Kamada, Masahiko Nakatsui, Ryosuke Kojima, Yasushi Okuno	American Society of Human Genetics (ASHG) 2020 Annual Meeting	2020年10月	国際
2	The effect of conformational flexibility of protein-ligand binding free energy computation and its significance in predicting mutation-induced drug sensitivity changes in oncogenic kinases.	Araki M and Okuno Y	Fall 2020 ACS National Meeting	2020年8月	国際
3	Protein-drug binding free energy computation incorporating protein flexibility and its significance in predicting mutation-induced drug resistance in oncogenic kinases.	Araki M and Okuno Y	Spring 2020 ACS National Meeting	2020年4月	国際
4	Molecular dynamics simulations reveal the protein-protein interaction mechanism	Takefumi Yamashita	16th International Conference of Computational Methods in Sciences and Engineering (CC symposium of ICC MSE 2020)	2020年9月	国際
5	Relationship between the void formation and stress change during the deformation of epoxy resin	Naoyuki Shoji and Takefumi Yamashita	16th International Conference of Computational Methods	2020年9月	国際

			in Sciences and Engineering (CC symposium of ICCM SE 2020)		
6	Conformational flexibility and stability of SARS-CoV-2 spike protein in solution	Yuji Sugita	ACS Spring 2021	2021 年 4 月	国際
7	Intrinsic Conformational Flexibility of SARS-CoV-2 Spike Protein Simulated on Fugaku	Yuji Sugita	2021 R-CCS Symposium	2021 年 2 月	国際
8	Multi-scale MD simulations using GENESIS on Fugaku	Yuji Sugita	The 1st Fugaku Bio-supercomputing Workshop on Cellular-Scale Molecular Dynamics Simulations	2021 年 1 月	国際
9	Use of enhanced conformational sampling for the analysis of protein-ligand binding processes	Yuji Sugita	Virtual Workshop "Biomolecular Interactions in Cellular Environments" in Telluride Science Research Center	2020 年 8 月	国際
10	Replica-exchange simulations on the conformational dynamics of spike protein on the surface of SARS-CoV-2	Yuji Sugita	Molecular Basis of Proteinopathies	2021 年 2 月	国際
11	Developments and Applications of Free-Energy Analysis for Protein-Ligand Binding	Hiraku Oshima	RIKEN BDR Symposium 2021	2021 年 3 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	Regulation mechanism of agonist/antagonist activities of vitamin D receptor elucidated by generalized ensemble method	T.Ekimoto, T.Kudo, T.Yamane, and M.Ikeguchi	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
2	Comparison of molecular dynamics of cyclosporin A and cyclosporin E	A.Ito, T.Yamane, T.Ekimoto, and M.Ikeguchi	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
3	Elucidation of interaction between neurotrophin receptor TrkAd5 and its binding peptide	M.Takahashi, T.Ekimoto, T.Yamane, R.Suzuki, H.Takahashi, and M.Ikeguchi	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
4	3D-RISM-AI: Prediction of protein-ligand binding free energy based on machine learning using 3D-RISM	K.Ohsaki, T.Ekimoto, T.Yamane, and M.Ikeguchi	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
5	Structural basis for the switching of binding partners by Ser298 phosphorylation in UHRF1	S.Kori, T.Jimenji, T.Ekimoto, M.Sato, F.Kusano, T.Oda, M.Unoki, M.Ikeguchi, and K.Arita	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
6	Comprehensive analysis of the hydration of small molecule binding sites in ligand-free protein structures: 3D-RISM approach	T.Yoshidome, M.Ikeguchi, and M.Ohta	第 20 回日本蛋白質科学会年会(新型コロナウイルスのため要旨のみ)	2020 年 6 月	国内
7	中分子シクロスポリン A とシクロスポリン E の分子ダイナミクスの比較	伊藤朱里、浴本亨、山根努、池口満徳	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内
8	ニューロトロフィン受容体 TrkAd5 と結合ペプチド間の相互作用様式の解明	高橋真帆、浴本亨、鈴木里佳、山根努、高橋栄夫、池口満徳	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内

9	機械学習を用いた 3D-RISM ベースの結合自由エネルギー予測	大崎和、浴本亨、山根努、池口満徳	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内
10	シクロスポリン A の CHARMM カ場の開発	山根努、浴本亨、池口満徳	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内
11	幾何学的相互作用解析と機械学習による抗原-抗体複合体側鎖モデル構造妥当性予測	千葉峻太郎、小甲裕一、野沢優翼、柳田駿介、佐藤美和、池口満徳、大田雅照	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内
12	ディープラーニングを用いたタンパク質水和分布予測法の研究	河間光祐、吉留崇、池口満徳、大田雅照	第 48 回構造活性相関シンポジウム(オンライン)	2020 年 12 月	国内
13	クライオ EM 密度マップからの生体分子ダイナミクス情報推定-データベース・シミュレーション・機械学習の融合-	寺山慧	ISSP ワークショップ	2021 年 2 月	国内
14	化学空間や配座空間をより自由に探索するために: 様々な機械学習・最適化手法とシミュレーションの連携	寺山慧	第 48 回構造活性相関シンポジウム	2020 年 12 月	国内
15	AI の可能性と実応用: 魚・サンゴモニタリングから創薬・材料科学まで	寺山慧	第 28 回海洋工学シンポジウム	2020 年 9 月	国内
16	Atomistically Deciphering Functional Activation Processes of Proteins with Molecular Simulations	Shigehiko Hayashi	第 58 回日本生物物理学会年会	2020 年 9 月	国内
17	Free Energy Calculations of HIV-1 Protease Binding Indinavir and Its Drug-Resistant Mutant	Masahiko Taguchi, Ryo Oyama, Masahiro Kaneso, Shigehiko Hayashi	第 58 回日本生物物理学会年会	2020 年 9 月	国内
18	Theoretical study on molecular mechanics of natural anion channel rhodopsin GtACR1	Takafumi Shikakura, Cheng, Shigehiko Hayashi	第 58 回日本生物物理学会年会	2020 年 9 月	国内
19	Theoretical study on molecular mechanism of a light-driven ion transport of Halorhodopsin	Ryo Oyama, Taisuke Hasegawa, Shigehiko Hayashi	第 58 回日本生物物理学会年会	2020 年 9 月	国内

20	「産学連携コンソーシアム LINC から見た AI 創薬の現状と未来」	奥野恭史	知的財産戦略ネットワーク 講演会 医療 AI 「AI 創薬 について」～Aish リーズパ ート 2～	2021 年 3 月	国内
21	「医薬プロセス最適化プラットフォーム開発の現状と今後」	奥野恭史	理化学研究所 医科学イ ノベーション推進プロ グラム シンポジウム	2021 年 3 月	国内
22	「プレジジョンメディスンを加速する創薬ビッグデータ統合シ ステムの推進」	奥野恭史、荒木 望嗣、鎌田真 由美	富岳の創薬活用に向けた ワークショップ	2021 年 3 月	国内
23	「データ駆動型 医療・医薬品・ヘルスケア産業」	奥野恭史	JIQVIA セミナー	2021 年 2 月	国内
24	「AI が拓く創薬イノベーション」	奥野恭史	第 16 回がんトランスレー ショナルリサーチ (TR) ワ ークショップ AI が創る医 薬品開発のカットニング・ エッジ)	2021 年 1 月	国内
25	スパコンを用いたタンパク質-化合物結合親和性の高精度 予測と医療・創薬への応用	荒木望嗣, 奥野恭史	第 48 回構造活性相関シン ポジウム	2020 年 12 月	国内
26	「AI・シミュレーションが拓く創薬・医療の未来」	奥野恭史	けいはんな R&D イノベ ーションフォーラム	2020 年 12 月	国内
27	「「富岳」がもたらす創薬・医療へのインパクト」	奥野恭史	第 13 回 スーパーコンピュ ーティング技術 産業応用 シンポジウム	2020 年 12 月	国内
28	「スーパーコンピュータ・AI が拓く創薬の未来」	奥野恭史	Life Science Startup Eco system	2020 年 12 月	国内

29	「スーパーコンピュータ・AI が拓く創薬の未来」	奥野恭史	第 3 回産学官連携情報交流セミナー	2020 年 12 月	国内
30	日本人疾患バリエーションデータベース MGeND	中津井雅彦, 鎌田真由美	第 43 回分子生物学会 企画フォーラム「ゲノム研究と医療をつなぐデータベース」	2020 年 12 月	国内
31	MGeND: Database updates and detailed analysis	Mayumi Kamada, Rei Ashine, Masahiko Nakatsui, Ryosuke Kojima, Yasushi Okuno	日本人類遺伝学会第 65 回大会	2020 年 11 月	国内
32	「AI・シミュレーションが拓く創薬・医療の未来」	奥野恭史	JCHUGAI DIGITAL DAY ～ヘルスケア×デジタルの 2030 未来予想図	2020 年 11 月	国内
33	「創薬における AI の現状と未来」	奥野恭史	第 21 回日本毒性病理学会	2020 年 11 月	国内
34	日本人疾患バリエーションデータベース MGeND アップデート	中津井雅彦, 芦根怜, 鎌田真由美, 小島諒介, 奥野恭史	トーゴーの日 2020	2020 年 10 月	国内
35	「深層学習技術による cryo-EM 密度マップデータに隠された動的情報の抽出」	松本篤幸	第一回クライオ電子顕微鏡画像からの高度情報処理研究会	2020 年 10 月	国内
36	「スーパーコンピュータ・AI が拓くがん分子標的治療戦略」 「Cancer molecular targeted therapy strategy challenged by supercomputer and AI」	奥野恭史	第 24 回日本がん分子標的治療学会学術集会	2020 年 10 月	国内
37	「スーパーコンピュータ・AI が拓く創薬・医療の未来」	奥野恭史	人事労災対策委員会・総務委員会,	2020 年 9 月	国内

38	「創薬 AI とこれからの医薬品開発、新たな産業の役割」	奥野恭史	IQVIA セミナー(製薬企業向け)	2020年7月	国内
39	「シミュレーション・AI 駆動型創薬」	奥野恭史	俯瞰ワークショップ ナノテクノロジー・材料分野 区分別分科会	2020年7月	国内
40	「創薬 AI とこれからの医薬品開発、新たな産業の役割」	奥野恭史	IQVIA セミナー(メディア向け)	2020年7月	国内
41	「AI・シミュレーションによる薬効・毒性予測」	奥野恭史	第 47 回日本毒性学会学術年会シンポジウム,	2020年6月	国内
42	Dynamics of Close-Open State of Candida antarctica Lipase B Obtained by Parallel Cascade Molecular Dynamics Simulation and The Markov State Model	Wijaya Tegar Nurwahyu, Akio Kitao	第 10 回日本生物物理学会関東支部会 オンライン開催	2021年3月	国内
43	Unguided Binding MD of Protein-Protein Complexes by PPI-ColDock	竹村和浩,北尾彰朗	第 20 回日本蛋白質科学会年会 オンライン学会	2020年7月	国内
44	Simulating the association/dissociation process of flexible protein complexes	Duy Phuoc Tran,Akio Kitao	第 20 回日本蛋白質科学会年会 オンライン学会	2020年7月	国内
45	Protein-protein complex dissociation simulated by Parallel Cascade Selection Molecular Dynamics	宮澤佳希, Tran Phuoc Duy, 竹村和浩, 北尾彰朗	第 20 回日本蛋白質科学会年会 オンライン学会	2020年7月	国内
46	Unguided Binding MD of Protein-Protein Complexes by PPI-ColDock	竹村和浩, 北尾彰朗	第 58 回日本生物物理学会年会 オンライン学会	2020年9月	国内
47	PaCS-MD/MSM を用いたタンパク質複合体の速度定数評価	宮澤佳希, Tran Phuoc Duy, 竹村和浩, 北尾彰朗	第 58 回日本生物物理学会年会 オンライン学会	2020年9月	国内

48	Investigating the dissociation process of DBD-p53/DNA complex by PaCS-MD and MSM	Sobeh Mohamed Marzouk, Akio Kitao	第 58 回日本生物物理学 会年会 オンライン学会	2020 年 9 月	国内
49	分子動力学計算による p53-C 末端部位の DNA 結合機構の解明	平悠太, Tran Phuoc Duy, 北尾 彰朗	第 58 回日本生物物理学 会年会 オンライン学会	2020 年 9 月	国内
50	分子動力学計算による p53-C 末端部位の DNA 結合機構の解明	平悠太, Tran Phuoc Duy, 北尾 彰朗	第 34 回分子シミュレーシ ョン討論会 オンライン学会	2020 年 12 月	国内
51	PaCS-MD/MSM を用いたタンパク質-タンパク質複合体の結合親和性評価	宮澤佳希, Tran Phuoc Duy, 竹 村和浩, 北尾彰朗	第 34 回分子シミュレーシ ョン討論会 オンライン学会	2020 年 12 月	国内
52	PaCS-MD によるアクチオシンの結合解離予測	尾川拓巳, Tran Phuoc Duy, 藤 井高志, 難波啓一, 北尾彰朗	第 34 回分子シミュレーシ ョン討論会 オンライン学会	2020 年 12 月	国内
53	抗原-抗体界面における塩橋安定性の理論的解析	山下雄史	第 1 回生体分子シミュレー ション・モデリング研究会	2021 年 3 月	国内
54	MD シミュレーションによるタンパク質-リガンド相互作用の研究 (Molecular dynamics study on protein-ligand interaction)	山下雄史	日本生物物理学会年会 (第 9 回 Biophysics and P hysicobiology 論文賞受賞 講演会/The 9th Award Seminar for outstanding Biophysics and Physicobi ology paper)	2020 年 9 月	国内
55	タンパク質-リガンド結合の自由エネルギー計算法の開発と応用	尾嶋拓	CBI 学会 2020 年大会	2020 年 10 月	国内
56	New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems	Jaewoon Jung, 小林千草, Cheng Tan, 笠原健人, Michael Feig, 杉田有治	第 34 回 分子シミュレーシ ョン討論会	2020 年 12 月	国内
57	新型コロナウイルス表面のタンパク質動的構造予測	杉田有治	HPCI フォーラム	2021 年 3 月	国内

58	細胞環境はどのようにタンパク質の構造・ダイナミクス・機能に影響を与えるか？	杉田有治	CBI 学会 2020 年大会	2020 年 10 月	国内
59	All-atom molecular dynamics simulations of spike protein on the surface of SARS-CoV-2 in solution	Yuji Sugita	RIKEN-NRC HPC Workshop	2020 年 10 月	国内
60	富岳を用いた新型コロナウイルス表面のスパイクタンパク質動的構造予測	杉田有治	近畿化学協会コンピュータ化学部会	2021 年 3 月	国内
61	新型コロナウイルス 表面のタンパク質動的構造予測	杉田有治	第 9 回 JCAHPC セミナー	2020 年 10 月	国内
62	計算機シミュレーションで細胞の中を観る	杉田有治	京都大学 MACS コロキアム	2020 年 11 月	国内
63	スーパーコンピュータが明らかにする蛋白質の動的構造と機能	杉田有治	理研・東北大学連携シンポジウム「計測科学が拓く生命科学の新展開」	2020 年 12 月	国内
64	Developments and applications of generalized-ensemble methods for free-energy analysis of protein-ligand binding	尾嶋拓	第 20 回日本蛋白質科学会年会	2020 年 7 月	国内
65	In silico drug discovery using the structural data and molecular simulation	広川貴次	第 43 回日本分子生物学会年会	2020 年 12 月	国内
66	創薬支援研究における計算化学の活用法	広川貴次	第 8 回食・触コンソーシアムワークショップ	2020 年 12 月	国内
67	薬剤分子設計および低分子デザイン	広川貴次	スーパーコンピュータ「富岳」創薬活用ワークショップ	2021 年 3 月	国内
68	新型コロナウイルスによるパンデミックに対する FMO DB の取り組み	渡邊千鶴、神坂紀久子、大山達也、高谷大輔、加藤幸一郎、川嶋裕介、半田佑磨、山本亜美、渡邊一樹、福澤薫、本間光貴	第 48 回構造活性相関シンポジウム	2021 年 12 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 <small>(○を記入)</small>
1	Imidazolium-based multiblock amphiphile as transmembrane anion transporter	M.Mori, K.Sato, T.Ekimoto, S.Okumura, M.Ikeguchi, K.V.Tabata, H.Noji, and K.Kinbara	Chemistry-An Asian Journal, Volume 16, Issue 2, Pages 147-157	2021年1月	
2	Structural and dynamical insights into the PH domain of p62 in human TFIIH	M.Okuda, T.Ekimoto, J.Kurita, M.Ikeguchi, and Y.Nishimura	Nucleic Acids Research, Volume 49, Issue 5, Pages 2916-2930	2021年3月	
3	X線小角散乱と分子動力学シミュレーションを組み合わせた生体分子の溶液構造解析	浴本亨、小甲裕一、笠口友隆、池口満徳	分子シミュレーション学会誌「アンサンブル」、22巻、3号、239-251頁	2020年7月	
4	Black-Box Optimization for Automated Discovery	Kei Terayama, Masato Sumita, Ryo Tamura, and Koji Tsuda	Accounts of Chemical Research	2021年2月	
5	Ab Initio Evaluation of the Redox Potential of Cytochrome c	Cheng, Hayashi Shigehiko	Journal of Chemical Theory and Computation	2021年1月	
6	Extraction of protein dynamics information from cryo-EM maps using deep learning	Matsumoto S, Ishida S, Araki M, Kato T, Terayama K, Okuno Y	Nature Machine Intelligence, 3, 153-160	2021年2月	
7	Gilteritinib overcomes lorlatinib resistance in ALK-rearranged cancer	Mizuta H, Okada K, Araki M, Adachi J, Takemoto A, Kutkowska J, Maruyama K, Yanagitani N, Oh-hara T, Watanabe K, Tamai K, Friboulet L, Katayama K, Ma B, Sasakura Y, Sagae Y,	Nature Communications 12, 1261	2021年2月	

		Kukimoto-Niino M, Shirouzu M, Takagi S, Simizu S, Nishio M, Okuno Y, Fujita N, and Katayama R			
8	Novel resistance mechanisms including L1196Q, P1094 H, and R1248_D1249insertion in three patients with non-small-cell lung cancer following ALK tyrosine kinase inhibitor treatment	Furuta H, Araki M, Masago K, Sagae Y, Fujita S, Seto K, Shimizu J, Horio Y, Sasaki E, Hosoda W, Katayama R, Okuno Y, Hida T	Journal of Thoracic Oncology (3):477-482	2021年3月	
9	「富岳」が拓く計算科学の未来	荒木望嗣、鎌田真由美、奥野恭史	milsil 13(6),2020	2020年11月	
10	“Trp53 mutation in Krt5-expressing basal cells facilitates the development of basal squamous-like invasive bladder cancer in the chemical carcinogenesis of mouse bladder”	Norihiko Masuda, Kaoru Murakami, Yuki Kita, Akihiro Hamada, Mayumi Kamada, Yuki Teramoto, Toru Sakatani, Keiyu Matsumoto, Takeshi Sano, Ryoichi Saito, Yasushi Okuno, Osamu Ogawa, Takashi Kobayashi	American Journal of Pathology, pii: S0002-9440(20)30202-9. doi:10.1016/j.ajpath.2020.04.005.	2020年8月	
11	Edge expansion parallel cascade selection molecular dynamics simulation for investigating large-amplitude collective motions of proteins	Kenichiro Takaba, Duy Phuoc Tran, Akio Kitao	J. Chem. Phys. v152, p225101(1-13)	2020年6月	
12	Regulation of caveolae through cholesterol-depletion-dependent tabulation mediated by PACSIN2	Aini Gusmira, Kazuhiro Takemura, Shin Yong Lee, Takehiro Inaba, Kyoko Hanawa-Suetsugu, Kayoko Oono-Yakura, Kazum Yasuhara, Akio Kitao. Shiro Suetsugu,	J. Cell. Sci., V133(19), P246785(1-11)	2020年9月	

13	Delineating the conformational landscape of the adenosine A2A receptor during G protein coupling	Shuya Kate Huang, Aditya Pandey, Duy Phuoc Tran, Nicolas L. Villanueva, Akio Kitao, Roger K. Sunahara, Adnan Sljoka, Scott R. Prosser	Cell., V184(7), P1884-1894 (e14)	2021年3月	○
14	PaCS-MD/MSM: Parallel Cascade Selection Molecular Dynamic Simulation in Combination with Markov State Model as an Efficient non-Bias Sampling Method	Duy Phuoc Tran, Hiroaki Hata, Takumi Ogawa, Yuta Taira, Akio Kitao	分子シミュレーション学会誌アンサンブル, 22, 151-156	2020年4月	
15	バイオ分野での量子コンピュータへの期待	北尾彰朗	現代化学、v6(No.591)、p50-51	2020年6月	
16	Singular Spectrum Transformation for Detecting Molecular Motion Mode Change of Protein Systems	Takefumi Yamashita and Naoyuki Shoji	AIP Conf. Proc. 2343, 020010 (2021); https://doi.org/10.1063/5.0047736	2021年3月	
17	Relationship Between Void Formation and Stress Change During Deformation of Epoxy Resin	Naoyuki Shoji and Takefumi Yamashita	AIP Conf. Proc. 2343, 020009 (2021); https://doi.org/10.1063/5.0047734	2021年3月	
18	Modification and Validation of the DREIDING Force Field for Molecular Liquid Simulations (DREIDING-UT)	Kohei Sasaki and Takefumi Yamashita	J. Chem. Inf. Model., 61, 3, 1172-1179 (2021). https://doi.org/10.1021/acs.jcim.0c01169	2021年2月	
19	Environmental effects on salt bridge stability in the protein-protein interface: The case of hen egg-white lysozyme and its antibody, HyHEL-10	Ryo Okajima, Shuichi Hiraoka, and Takefumi Yamashita	J. Phys. Chem. B 125, 1542-1549 (2021). https://doi.org/10.1021/acs.jpcc.0c09248	2021年1月	

20	Prediction of Protein-Ligand Binding Pose and Affinity Using the gREST+FEP Method	Hiraku Oshima, Suyong Re, Yuji Sugita	J. Chem. Inf. Model. Vol. 60, 5382-5394 (2020)	2020年7月	
21	CHARMM-GUI Free Energy Calculator for Absolute and Relative Ligand Solvation and Binding Free Energy Simulations	Seonghoon Kim, Hiraku Oshima, Han Zhang, Nathan R. Kern, Suyong Re, Jumin Lee, Benoît Roux, Yuji Sugita, Wei Jiang, Wonpil Im	J. Chem. Theory Comput. 2020, 16, 7207-7218	2020年10月	○
22	Unraveling the Coupling between Conformational Changes and Ligand Binding in Ribose Binding Protein Using Multiscale Molecular Dynamics and Free-Energy Calculations	Weitong Ren, Hisham M Dokainish, Ai Shinobu, Hiraku Oshima, Yuji Sugita	J. Phys. Chem. B, 125,2898-2909 (2021)	2021年3月	
23	分子動力学ソフトウェア GENESIS を用いたタンパク質-リガンド結合の自由エネルギー計算	尾嶋 拓, 李 秀榮, 新津 藍, 杉田 有治	シミュレーション, 40, 22-28 (2021)	2021年3月	
24	Group-based evaluation of temperature and pressure for molecular dynamics simulation with a large time step	Jaewoon Jung, Yuji Sugita	The Journal of Chemical Physics, 153, 234115 (2020)	2020年12月	
25	New parallel computing algorithm of molecular dynamics for extremely huge scale biological systems	Jaewoon Jung, Chigusa Kobayashi, Cheng Tan, Akiyoshi Kuroda, Kazuo Minami, Yutaka Ishikaw, Kento Kasahara, Yuji Sugita, Shigeru Ishiduki, Tatsuo Nishiki, Hikaru Inoue, Michael Feig	Journal of Computational Chemistry, 42, 231-241(2021)	2021年2月	

26	FMODB: The World's First Database of Quantum Mechanical Calculations for Biomacromolecules Based on the Fragment Molecular Orbital Method.	Takaya D, Watanabe C, Nagase S, Kamisaka K, Okiyama Y, Moriwaki H, Yuki H, Sato T, Kurita N, Yagi Y, Takagi T, Kawashita N, Takaba K, Ozawa T, Takimoto-Kamimura M, Tanaka S, Fukuzawa K, Honma T	J Chem Inf Model. 61, 777-794, 2021,	2021 年 1 月	
27	High-Precision Atomic Charge Prediction for Protein Systems Using Fragment Molecular Orbital Calculation and Machine Learning.	Kato K, Masuda T, Watanabe C, Miyagawa N, Mizouchi H, Nagase S, Kamisaka K, Oshima K, Ono S, Ueda H, Tokuhisa A, Kanada R, Ohta M, Ikeguchi M, Okuno Y, Fukuzawa K, Honma T.	J Chem Inf Model. 60, 3361-368, 2020.	2020 年 6 月	