

令和5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

(次世代超高速電子計算機システム利用の成果促進)

「「富岳」で目指すシミュレーション・AI 駆動型次世代医療・創薬」

成果報告書

令和6年5月30日

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

奥野 恭史

目次

補助事業の名称.....	- 1 -
1. 補助事業の目的.....	- 1 -
2. 令和5年度（報告年度）の実施内容	- 1 -
2-1. 当該年度（令和5年度）の事業実施計画.....	- 1 -
2-2. 実施内容（成果）	- 3 -
2-3. 活動（研究会の活動等）	- 34 -
2-4. 実施体制.....	- 35 -

補助事業の名称

「富岳」成果創出加速プログラム

(次世代超高速電子計算機システム利用の成果促進)

「富岳」で目指すシミュレーション・AI 駆動型次世代医療・創薬

体系的番号： JPMXP1020230120

1. 補助事業の目的

がん・心疾患・難病をターゲット疾患とした病態解明・創薬・診断支援に「富岳」を実践的応用することで、現代の医療・創薬が抱える根本課題克服に挑戦するシミュレーション・AI 駆動型次世代医療・創薬の開拓を目指す。

2. 令和5年度（報告年度）の実施内容 2

ー 1. 当該年度（令和5年度）の事業実施計画

社会実装グループ：

①疾患多階層データベースの構築と応用（京大 鎌田・中澤）：ゲノム変異、分子構造、ダイナミクス、マルチオミクス、パスウェイ、細胞レベルでの疾患多階層データベースの構築と、疾患原因・創薬標的分子の探索と病態メカニズム解明に取り組む。R5 年度は、多階層データを統合するためのデータ設計とユースケースに基づくクエリの実装、小規模 MD の実施とスコア化のための検討を行う。

②インシリコ心疾患データベースの公開と応用（UT-Heart 研 杉浦、京大 糀谷）：心臓シミュレーションによる心疾患データベースの構築と、実臨床データに基づく AI との融合による拡張性心不全の発症を早期に検出するバイオマーカーの探索に取り組む。R5 年度は、心不全(hfpEF)に関わるパラメータを変動させた心臓シミュレーションと心房モデル性能試験、実臨床データに基づく心不全(hfpEF)検出の AI の開発、及び API 開発を行う。また、データベースの管理・運用法についても検討する。

③創薬 DX プラットフォームを用いた実践的創薬（京大 荒木・岩田・大塚）：難病のドラッグリポジショニング、がんの個別化医療・創薬の実践、アカデミア創薬テーマへの創薬応用に取り組む。R5 年度は、下記の3項目を実施する。(i) 難病のドラッグリポジショニング：PathwayGCN を用いて難病 800 種の各々の創薬標的遺伝子を抽出し、当該遺伝子産物（タンパク質）の立体構造モデリングを行う。(ii) がんの個別化医療・創薬：10 名のがん患者に対するマルチオミクスデータをベイジアンネットワーク解析法・PathwayGCN 法を用いて解析し、創薬標的遺伝子を同定する。次いで、標的タンパク質の立体構造をモデリングし、PathwayGCN 法・ドッキング計算・ColDock 法を用いた既存薬ライブラリからのインシリコスクリーニングを開始する。(iii) 共同研究による創薬応用：各創薬標的タンパク質について、T- REMD によるタンパク質の立体構造サンプリングを実施する。

④インシリコ心毒性評価システムによる抗がん剤服用患者の安全性確保およびシステムの高精度化（東大 寺田、理研 井阪）：心筋細胞イオンチャネルへの抗がん剤のドッキングシミュレーションと UT-Heart の融合による不整脈発生リスク評価に取り組む。R5 年度は、iPS 細胞モデル及びがん患者固有の薬剤受容性に関する調査、及び 4 種類の抗がん剤の選択及び hERG, NaV1.5, CaV1.2 チャネルへのドッキングシ

ミュレーションを行う。

次世代アプリケーション開発グループ：

⑤遺伝子と臓器を結ぶ究極のマルチスケール心臓シミュレーション (UT-Heart 研 鷺尾、京大 岡本)：収縮タンパクの MD と心臓の有限要素拍動解析を連成させたシミュレータの高速化 (富岳 NEXT を見据えた次世代 UT-Heart 開発) と肥大型心筋症の解明・治療薬への応用に取り組む。R5 年度は、肥大型心筋症の分子実験に関する研究報告調査と分子モデル検討、それに基づく心臓シミュレーション、及び MD 計算高速化とそれに基づく連成シミュレーション実施・分析を行う。

⑥MD×AI に基づく次世代創薬計算技術開発 (横浜市大 寺山、京大 松本)：MD (分子動力学計算) に基づく次世代バーチャルスクリーニング技術と、次世代化合物生成 AI - MD 融合手法の開発に取り組む。R5 年度は、MAVS および MADD のプロトタイプを実装し「富岳」上でテストを実施する。

⑦ニューモダリティ創薬のための創薬計算手法開発 (理研 千葉、阪大 信夫、東工大 Tran、筑波大 吉野)：抗体・タンパク質医薬品・ペプチド医薬品・PROTAC などの新規モダリティ創薬に資する次世代計算手法の開発に取り組む。R5 年度は、薬剤-標的分子の複合体構造を再現する計算条件の検討及び構造アンサンブルを拡張サンプリングで取得する方法の検討を実施する。

⑧大規模生体系ネットワーク処理技術開発 (京大 小島、弘前大 玉田)：分子・細胞・臓器・個人・疾患の多階層マルチスケールの超巨大ネットワークを構築・予測する次世代 AI 技術の開発に取り組む。R5 年度は、PathwayGCN の富岳対応した分散計算のコアライブラリの選定と実装、および INGOR の新規並列ネットワーク推定アルゴリズムの選定と実装を行う。

プロジェクトの総合的推進 (成果発表・広報普及・アウトリーチ)

- ・プロジェクト専用サイトを新設し、メディアなどの記事掲載、研究紹介コンテンツにより、専門家から一般まで幅広い層に積極的にプレゼンスを高める。
- ・年 2 回のワークショップ開催や各種関連学会で研究発表を行い、学术界や産業界へのアウトリーチを行う。
- ・主要アプリケーションや AI アプリを統合した創薬 DX プラットフォームのユーザーガイドを整備し、「富岳」などの計算機環境で計算できるよう学生、博士研究員、企業を対象とした利用講習会を年数回開催する。
- ・創薬・医療の DX を進めている産学連携コンソーシアム (ライフインテリジェンスコンソーシアム：LINC) との連携を継続し、産業界へのアウトリーチを広く行う。
- ・参加メンバー所属大学での講義や各種学会・講演会などで、本事業の研究内容を紹介し、大学生、大学院生、研究機関や企業の研究者の人材育成を図る。特に、若手・女性研究者の促進・育成については、若手・女性研究者促進 WG を設置し、生命科学の生命科学・情報科学人材の拡大を図る。

2-2. 実施内容（成果）

社会実装グループ：

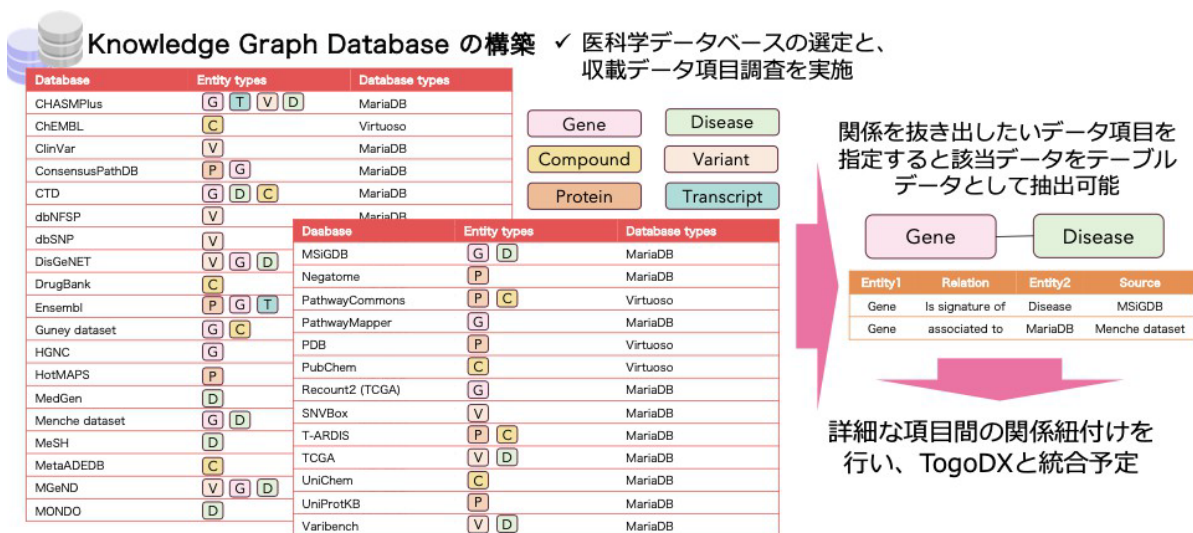
①疾患多階層データベースの構築と応用（京大 鎌田・中澤）：

ゲノム変異、分子構造、ダイナミクス、マルチオミクス、パスウェイ、細胞レベルでの疾患多階層データベースの構築と、疾患原因・創薬標的分子の探索と病態メカニズム解明に取り組む。R5 年度は、多階層データを統合するために、協力研究機関とデータモデル策定およびクエリ設計のためのユースケース検討を行なった。さらに、疾患関連変異体構造のモデリングと小規模 MD の実施、スコア化のための検討を行なった。

多階層データベース構築に向け、まず医科学に関連する 32 の主要な公共データベース・データセットに対し、統合するためのデータ項目整理を実施した。各データに収載されているデータ項目を「Gene」

「Compound」 「Protein」 「Disease」 「Variant」 「Transcript」 の主要 6 項目に整理し、ラベルを付与することで、例えば「Gene」と「Compound」の関連を抜き出すというクエリを与えると、収載しているデータベース・データセット横断的にデータを抽出するシステムの構築を行なった。医科学におけるデータセットに加え、基本的な生命科学データやエピゲノムや分子間相互作用等の階層のデータと接続するために、

協力研究機関であるライフサイエンス統合データベースセンター（DBCLS）とともに、生命科学データにおけるデータモデルに関する調査と検討を進めた。これまでの様々なデータモデルが提案されている。そこで、生命科学データにおけるデータモデルとして、BioLink [Unni, D.R. et al., *Clin Transl Sci.*, 2022]に着目し、いくつかのユースケースを想定し、すでに調査済みのデータベースに収載されるデータ項目が表現可能かについて調査を行なった。多数のデータ項目およびその関連は表現が可能であることが確認できたが、いくつかの項目では表現ができないもしくは適切ではない可能性があることがわかった。今後、協力研究機関および BioLink 開発者との協働でデータモデルの改良を進めていく予定である。



図①-1

疾患関連変異体構造のモデリングと小規模 MD の実施のため、計算対象とする難病に関連する変異体の絞り込みを行なった。米国で開発され世界的にも広く用いられている疾患バリエーションデータベース

ClinVar に記載されている約 100 万変異を対象に、難病に関連する変異の抽出を行なった。難病の定義は日本固有のものである。そのため、国内外のデータベースに記載されているデータから難病に関連するものかを判断するために、DBCLS で開発されている NANDO (Nanbyo Disease Ontology) を用いた。NANDO は日本の難病を整理し構造化するとともに、ClinVar など主要なデータベースでも用いられている統合疾患オントロジーMONDO との紐付けを行なっている。変異体構造のモデリングを行うために、ゲノムレベルの変異をタンパク質、タンパク質立体構造にマッピングするためのパイプラインを構築し、立体構造情報への紐付けを行なった。ClinVar に記載される疾患関連性のアノテーション (Pathogenicity/Benign) と疾患ラベル、NANDO 情報をもとに絞り込みを行い、難病に関連する 22,072 変異の特定を行なった。構造情報にマップされた情報に基づき、「富岳」に実装された OpenFold を用いて変異体および変異有さない野生型のタンパク質立体構造推定を実施した。構造推定が完了した約 1000 種のタンパク質に対して 10ns の短時間の全原子分子動力学計算 (ShortMD) を実施した。病原性を有する変異の特徴を検出することを目指し、取得したトラジェクトリデータの解析を進めている。

難病に紐づいたバリエーション

	内容	Variants	Missense variants	NANDOに紐づいた variants	NANDOに紐づいた variantsの疾患数	NANDOに紐づいた variantsのタンパク数
1	Pathogenic と Benign どちらもある疾患(n=1,535)のバリエーションリスト	99,609	33,671	22,072	837	1,087
2	1. 以外でPathogenicラベルのあるバリエーションリスト	13,845	4,295	4,462	346	268
3	1. 以外で Benignラベルのあるバリエーションリスト	330	298	84	27	21
4	1. 以外で Likely pathogenic ラベルのあるバリエーションリスト	37,213	18,725	10,397	752	933

図①-2

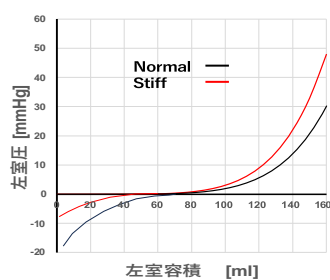
疾患原因・創薬標的分子の探索と病態メカニズム解明に向け、疾患患者から得られる遺伝子発現データとベイジアンネットワークアプリを用いた疾患ネットワークの推定を実施した。2023 年度は、米国で実施された大規模がんゲノム解析コンソーシアムである The Cancer Genome Atlas (TCGA) データを対象に、正常組織と比較可能ながん種 (10 種) を選定し、疾患ネットワークの推定を行なった。対象とした癌腫は、乳がん(1092 人)、肺がん(498 人)、甲状腺がん(504 人)、前立腺がん(495 人)、胃がん(414 人)、肝細胞がん(369 人)、大腸がん(288 人)、膵臓がん(178 人)である。各がん種を対象に、患者のサブ集団定義に関する 10 の条件を設定し、合計 100 のネットワーク推定を実施した。推定されたネットワークの解析を進めることで、疾患標的の探索および病態メカニズムの推定に取り組む。

②インシリコ心疾患データベースの公開と応用 (UT-Heart 研 杉浦、京大 糀谷) :

平成5年度前期においては収縮機能の保たれた心不全 (Heart failure with preserved ejection fraction: HFpEF) 早期発見のバイオマーカー検出のために以下の12要素の変動を網羅的に組み合わせた合計23040ケースを含む約3.9万ケースのシミュレーションを施行した。

イオン電流関連 ;

1. Na ピーク電流 (INa-peak) : 100 or 70%,
2. Ca 調節系 : L-type Ca 電流 (ICaL) , 筋小胞体 Ca ポンプ (SERCA)、筋小胞体 Ca 放出チャネル (CaRC), Na-Ca 交換体 (NCX), Na 電流 late component (INaL)
 - Ca0: CaL 100%, SERCA 100%, CaRC 100%
 - Ca1: CaL 80%, SERCA 70%, CaRC 150%
 - Ca2: CaL 80%, SERCA 70%, CaRC 150%, NCX 200%
 - Ca3: CaL 80%, SERCA 70%, CaRC 150%, INaL 200%
 - Ca4: CaL 80%, SERCA 70%, CaRC 150%, NCX 200%, InaL 200%
3. 遅延整流性 K 電流の遅い成分 (IKs) : 100 or 60%
4. 遅延整流性 K 電流の速い成分 (IKr) : 100 or 70%
5. 一過性外向き電流 (Ito) : 100 or 60%
6. 内向き整流性 K 電流 (Ik1) : 100 or 70%
7. 興奮伝播速度 (G) : 細胞長軸/短軸方向 100/100 or 80/130%
8. 左室形状 (LV) : 正常 or 拡張型 or 肥大型
9. 興奮伝播様式 (EX) : 正常 or 左脚ブロック
10. 細胞分布 (CELL) : 心内膜側から endocardial cell/M-cell/epicardial cell の割合
10/75/15 or 20/45/35 or 0/100/0
11. 心室の弛緩時物性 (Passive) : normal or stiff (下図②-1)

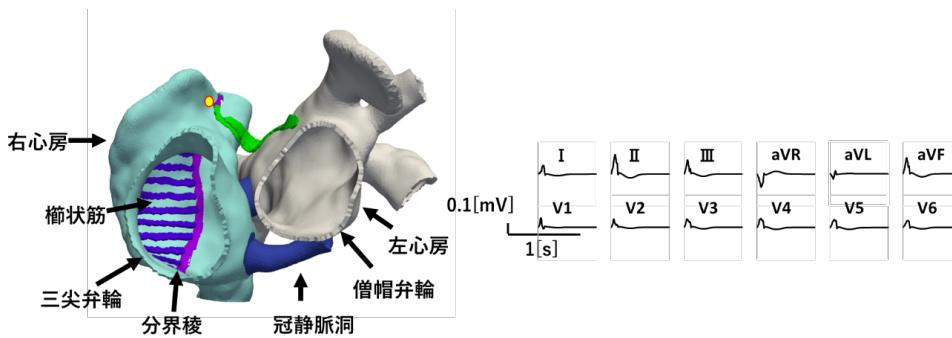


図②-1

12. 心電図の PQ 間隔 : 170ms or 200ms

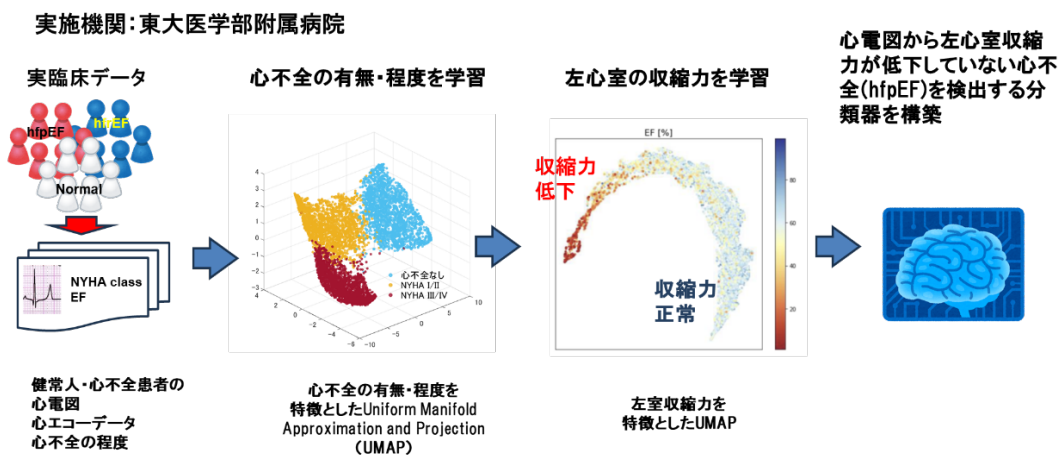
心電図から HFpEF を判別する AI は後述のように今年度から製作を開始したが、ひとまず以上のデータを既に作製されている HFrEF (収縮機能の低下した心不全) 判別用の AI に入力した結果、正常 2652、軽症～中等症 (New York Heart Association (NYHA) 機能分類 I および II 度) 3732、重症 (NYHA III および IV 度) 16656 に分類された。

なお既に述べた通りここで使用した AI はHF_rEF を対象にしたものであり最終的な結果ではない。また要素変動を追加したシミュレーションを 2023 年度下期に実施しており、その結果を加えた解析を行う予定である。また次年度には HF_pEF において注目されている心房の変化を検出するために心房のシミュレーションを行う予定であるが、そのために今年度は組織構造を忠実に再現したシミュレーションモデル（下図左）を開発し、正常な興奮伝播に伴う心電図（P 波）（下図右）の再現を確認する性能試験を実施した。



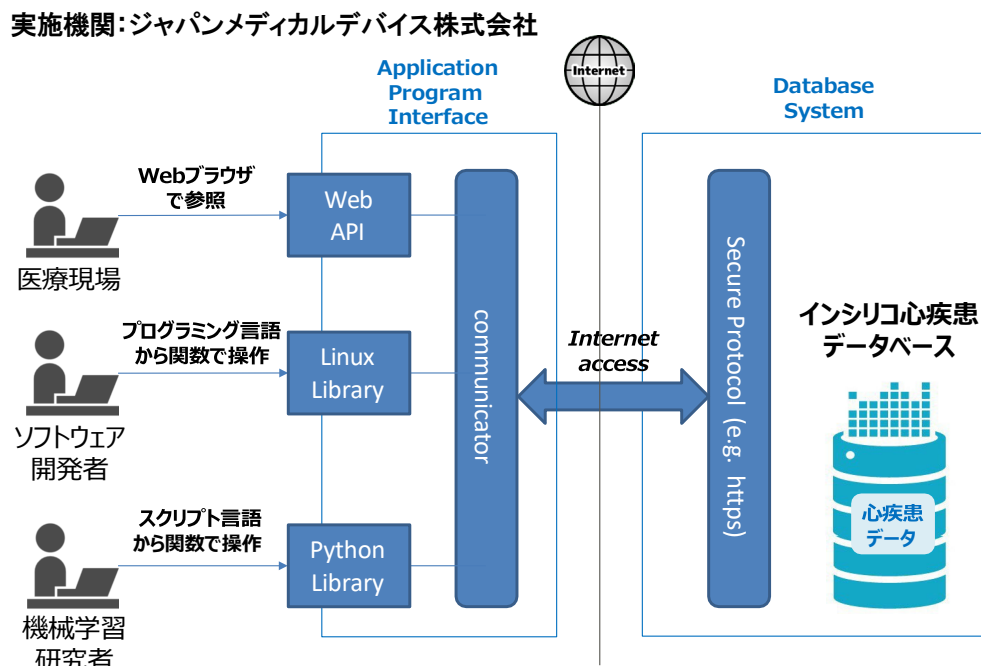
図②-2 心房モデルとシミュレーションによる心電図

HF_pEF を検出する AI（分類器）については、東京大学附属病院において本年度より下図に示す手順で開発を開始した。次年度も引き続き AI 開発を継続する予定である。



図②-3 HF_pEF 検出のための AI 心電図システムの概要

さらにジャパンメディカルデバイスにおいて、以上のように構築されつつあるデータベースの幅広い活用を促進するための Application Program Interface (API) の開発を開始した (下図)。



図②-4 データベース活用のための API

本年度は API の基本設計・システム機能要件として公開画面、データ項目、アクセス手法の検討を、非機能要件としてはシステム方式、ストレージ容量、応答時間、信頼性、拡張性、セキュリティの定義等のデータベースの管理・運用法につき検討を行なった。次年度も引き続き API 開発を継続する予定である。またストレージ容量約 2PB を想定したデータサイズの削減については、心臓シミュレーション結果 1 ケースの容量を 33.6GB から 26.9GB に削減することで、3 万ケースを超える全体の容量を約 1PB から約 0.8PB まで削減した。

③創薬 DX プラットフォームを用いた実践的創薬（京大 荒木・岩田・大塚）：

本課題では、難病のドラッグリポジショニング、がんの個別化医療・創薬の実践、アカデミア創薬テーマへの創薬応用に取り組む。それぞれのテーマについて、R5 年度に達成した成果を述べる。

(i) 難病のドラッグリポジショニング：難病データベース（NANDO）及び疾患に関連する変異が登録されている ClinVar データベースに基づいて、創薬ターゲットとする難病及び創薬標的遺伝子を選定した（図③-1）。NANDO に登録されている難病の中で、病原性の遺伝子変異（Pathogenic）と非病原性の遺伝子変異（benign）の両方が報告されている疾患は 837 種類であり、情報量が多いこれらの疾患を創薬ターゲットとした。また、変異が生じている遺伝子（タンパク質）1087 種を創薬標的遺伝子（タンパク質）とした。

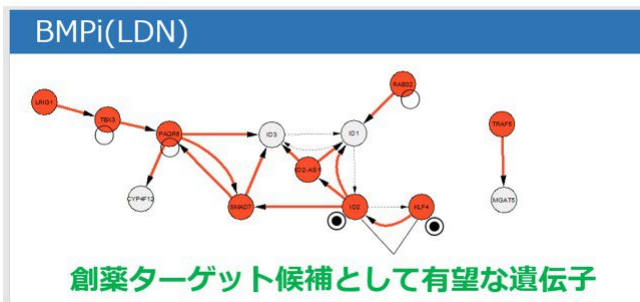
	内容	Variants	Missense variants	NANDOに紐づいた variants	NANDOに紐づいた variantsの疾患数	NANDOに紐づいた variantsのタンパク数
1	Pathogenic と Benign どちらもある疾患(n=1,535)のバリエーションリスト	99,609 (99,586)	33,671	22,072	837	1,087
2	1. 以外で Pathogenic ラベルのあるバリエーションリスト	13,845 (13,840)	4,295	4,462	346	268
3	1. 以外で Benign ラベルのあるバリエーションリスト	330	298	84	27	21
4	1. 以外で Likely pathogenic ラベルのあるバリエーションリスト	37,213 (37,205)	18,725	10,397	752	933

※ 配列の一致&置換ができたものは括弧の中の数

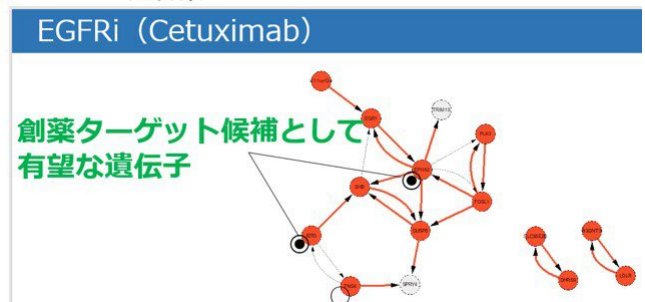
図③-1：NANDO 及び ClinVar データベースに基づいて抽出した変異リスト

(ii) がんの個別化医療・創薬：10 名の大腸がん患者由来のオルガノイド（CTOS）に 5 つの分子標的薬（BMP, MEK, WNT, PI3K, EGFR 阻害剤）を投与し、その前後のタイミングで取得した遺伝子発現データ（トランスクリプトーム）からベイジアンネットワークを推定し、遺伝子間の発現ネットワーク解析を行なった。個々のサンプルに対して、薬剤の投与前後で遺伝子間の関係に変動がみられたサブネットワークを抽出した。次に、薬剤の効果が良好であったグループに共通したサブネットワークに着目し、その中で抵抗性サンプルでは見られにくい遺伝子間関係の原因側遺伝子 7 種類を、創薬標的候補として選抜した。（図③-2）また、多剤に抵抗性となった 2 名の患者由来サンプルに特徴的なサブネットワークにも注目し、その中に含まれる遺伝子から創薬標的候補を抽出した。更に、「富岳」上でドッキング計算・ColDock 法の動作確認を行い、インシリコスクリーニングの準備を整えた。

<BMP阻害薬>



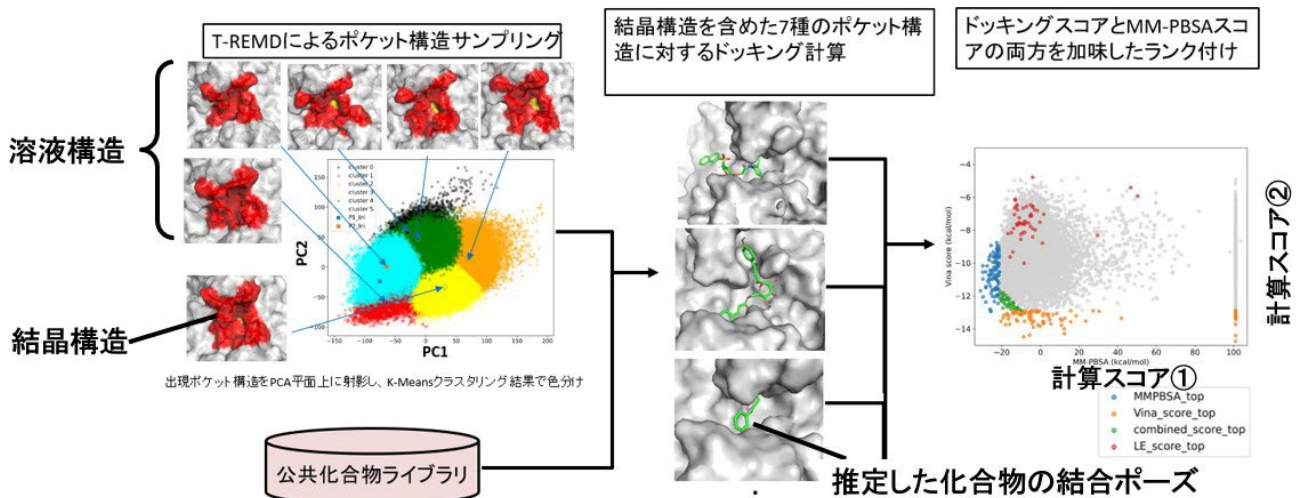
<EGFR阻害薬>



図③-2：ベイジアンネットワーク解析によって抽出した遺伝子サブネットワーク。代表的な 2 種類の薬剤（BMP 阻害薬、EGFR 阻害薬）の各々について、薬剤抵抗性サンプルで変動が見られないサブネットワークを表示し、創薬ターゲットとして有望な遺伝子を矢印で示した。

(iii) 共同研究による創薬応用：

本テーマでは、これまでに実験機関と連携して行なってきた基礎生物学的研究の創薬応用を目指す。R5 年度は、共同研究テーマ 11 種類の各々について、T-REMD による創薬標的タンパク質の立体構造サンプリング、及びタンパク質-化合物ドッキング計算による化合物ライブラリからのインシリコスクリーニングを開始した。代表的な研究事例を図③-3 に示す。



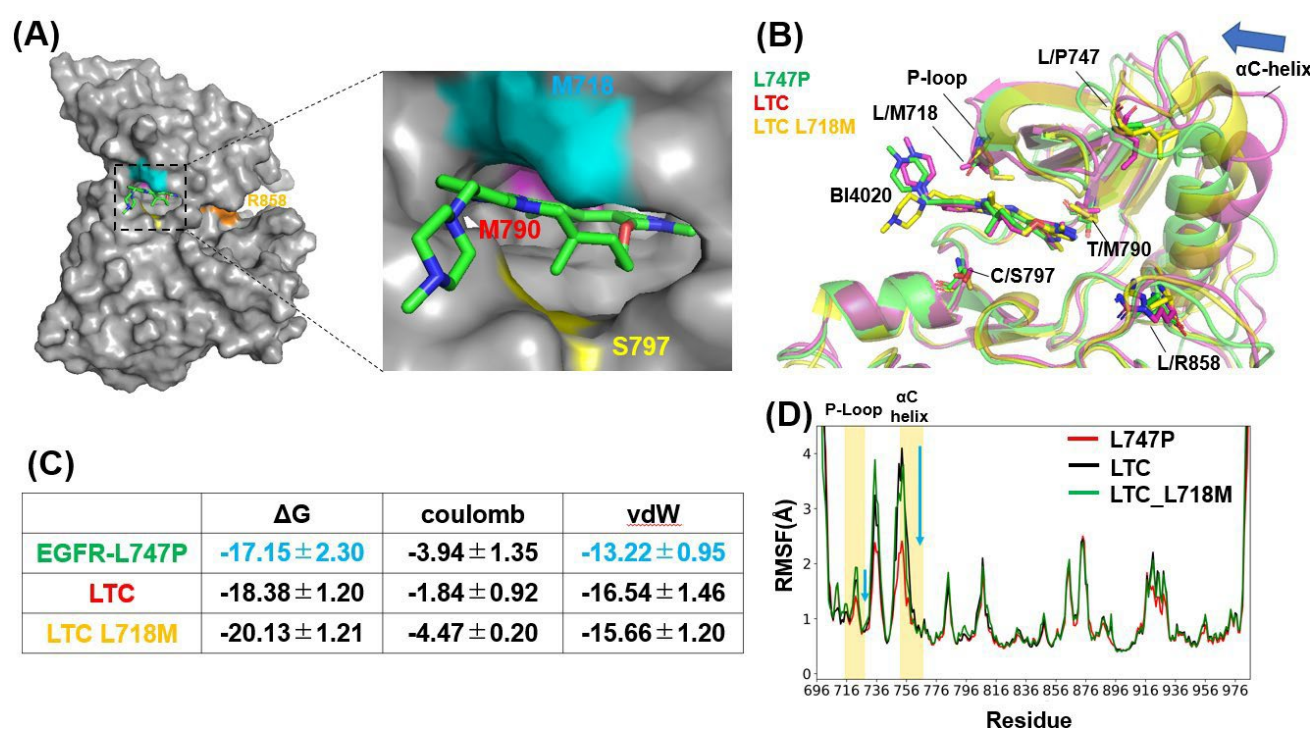
図③-3：細胞死抑制因子を標的とした機能阻害剤開発

温度レプリカ交換分子動力学シミュレーション（T-REMD）でサンプリングしたポケット構造を用いたドッキング計算により、ドッキングスコアと MM-PBSA スコアの両方を加味した指標に基づいて約 300 種の候補化合物を選抜した。

更に、上述のテーマと関連した研究成果として、EGFR キナーゼドメイン上の変異による薬剤耐性の分子メカニズムを推定した。下記に詳細な研究内容を記載する。

EGFR阻害薬 (BI4020) における薬剤耐性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、 L858R/T790M/C797S (LTC変異)、及び L858R/T790M/C797S/L718M (LTC-L718M変異)を対象に、第4世代EGFR阻害剤 (BI4020) の感受性の違いにおける分子メカニズムを解析した。EGFR-L747P変異体、LTC変異体、及びLTC-L718M変異体について、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、LTC変異体及びLTC-L718M変異体についてはATPポケットに対するBI4020の結合が安定化した一方で (図③-4A-C)、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、ATPポケットに歪みが生じてBI4020の結合親和性が低下すると推定された (図③-4B-D)。これらの結果より、BI4020はLTC変異体及びLTC-L718M変異体に対しては一定の薬効が認められるものの、L747P変異体に対しては感受性が低いという実験結果をサポートした。



図③-4 EGFR-L747P, L858R/T790M/C797S (LTC), L858R/T790M/C797S/L718M (LTC-L718M)変異体のシミュレーション結果。

(A) MDシミュレーションによって推定したEGFR-L747P変異体に対するBI4020の結合モード。タンパク質、薬剤はそれぞれサーフェスマデル、スティックモデルで表示した。薬剤結合ポケットの拡大図を右に示した。変異アミノ酸 (M718, M790, S797) はそれぞれシアン、赤、黄色でハイライトした。

(B) 1 μ 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-BI4020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。L747P, LTC, LTC-L718M変異

体をそれぞれ緑、赤、オレンジで表示し、LTC, LTC-L718M 変異体と比較した際の L747R 変異体における α C-helix の配向の変化を矢印で示した。

(C) MP-CAFEE によって算出した BI4020 の結合自由エネルギー (ΔG)。 ΔG における、静電相互作用 (coulomb)、ファンデルワールス相互作用 (vdW) の寄与を示した。

(D) 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, α C-helix 領域を黄色でハイライトした。

④インシリコ心毒性評価システムによる抗がん剤服用患者の安全性確保およびシステムの高精度化（東大 寺田、理研 井阪）：

本項目では、ドッキングシミュレーションによるチャネル-薬剤複合体構造予測および MD シミュレーションに基づく結合自由エネルギー計算を組み合わせた方法を用いることで心筋細胞イオンチャネルに対する阻害活性の予測を目指す。これまでの研究において、Comprehensive *in vitro* Proarrhythmia Assay (CiPA)においてリファレンスとして定められている薬剤について、心筋細胞イオンチャネル hERG、Nav1.5、に対する結合自由エネルギーの計算結果が阻害活性の実験データに基づく pIC50 と相関することが確認されている (Negami *et al.*, *Biophys. Physicobiol.* 20, e200016 (2023))。このことから、新たな薬剤について結合自由エネルギーを計算し、これまでに得られている回帰分析の結果に当てはめることで阻害活性を推定することができると思われる。また別プロジェクトにて、Ca_v1.2 チャネルの電子顕微鏡構造 (PDBID: 8FD7) の MD シミュレーションによる平衡化構造に対して CiPA リファレンス薬剤 24 種類のドッキングシミュレーションを行い、ドッキングスコアと pIC50 を比較したところ、従来使用している Glide XP では決定係数が 0.005 であるのに対して、深層学習に基づくスコア関数を用いる Gnina 1.0 では、決定係数が 0.627 になるという結果を得た (図④-1 左)。さらに、Gnina のドッキング構造を用いて MP-CAFEE 法により結合自由エネルギーを計算し pIC50 と比較したところ、決定係数は 0.414 となった (図④-1 右)。このため、本プロジェクトでも Ca_v1.2 チャネルに対するドッキングシミュレーションでは Gnina を用いることにした。

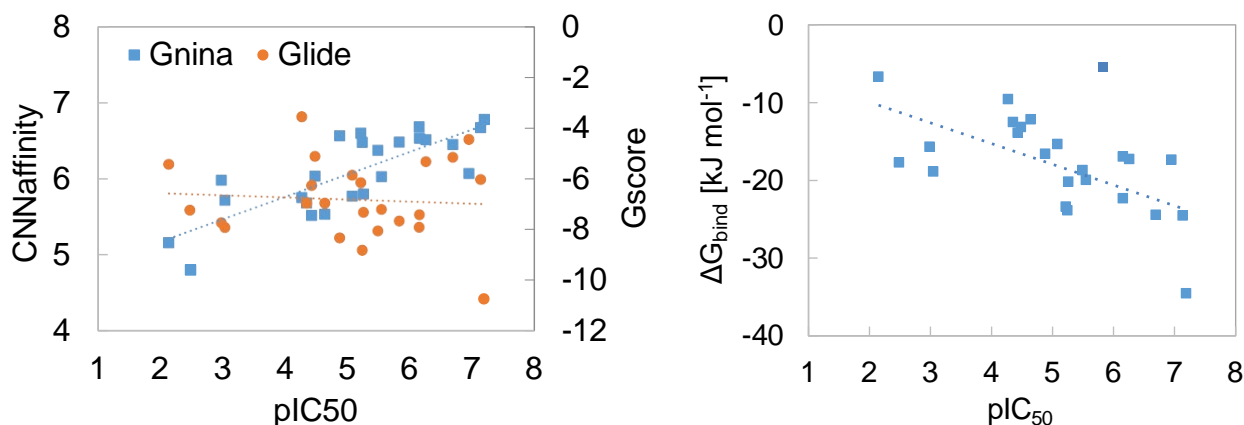
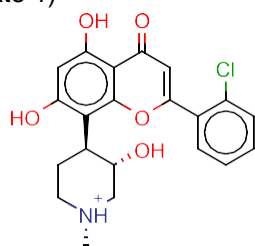


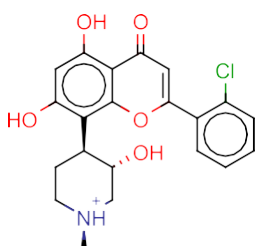
図 ④-1. 左 : Ca_v1.2 チャネルにおける、阻害活性の実験データ pIC50 に対する Gnina によるドッキングスコア (CNAffinity) および Glide のドッキングスコア (Gscore) のプロット。右 : 結合自由エネルギー計算値 ΔG_{bind} と実験データ pIC50 の比較。

令和 5 年度は、4 種類の抗がん剤について hERG、Nav1.5、Ca_v1.2 各イオンチャネルへのドッキングシミュレーション、平衡化計算を行なった。先行研究 (Xian *et al.*, *Toxicol. Appl. Pharmacol.* 437, 115886 (2022)) において、不整脈を誘起することが報告されている 7 種類の抗がん剤のうち、alvocidib、dasatinib、sorafenib、sunitinib の 4 種類を計算対象として選定した。alvocidib と dasatinib については、それぞれ 2 つのプロトン化状態を考慮した (図④-2)。hERG、Nav1.5 チャネルの構造には電子顕微鏡構造 (PDBID: 5VA2, 6LQA) を、Ca_v1.2 チャネルには、電子顕微鏡構造 (PDBID: 8FD7) を MD シミュレーションにより平衡化した構造を用いた。ドッキングプログラムは、hERG につい

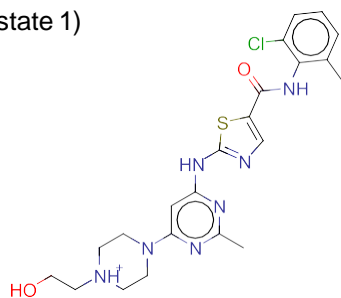
Alvocidib :
(state 1)



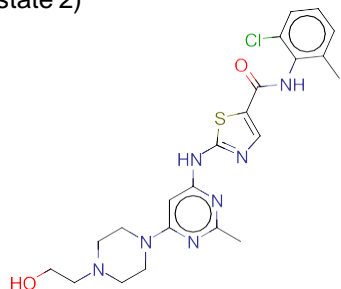
(state 2)



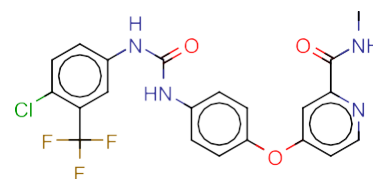
Dasatinib :
(state 1)



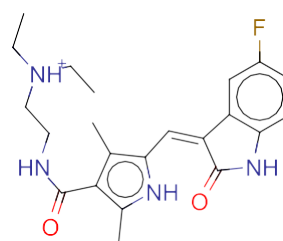
(state 2)



Sorafenib :



Sunitinib :

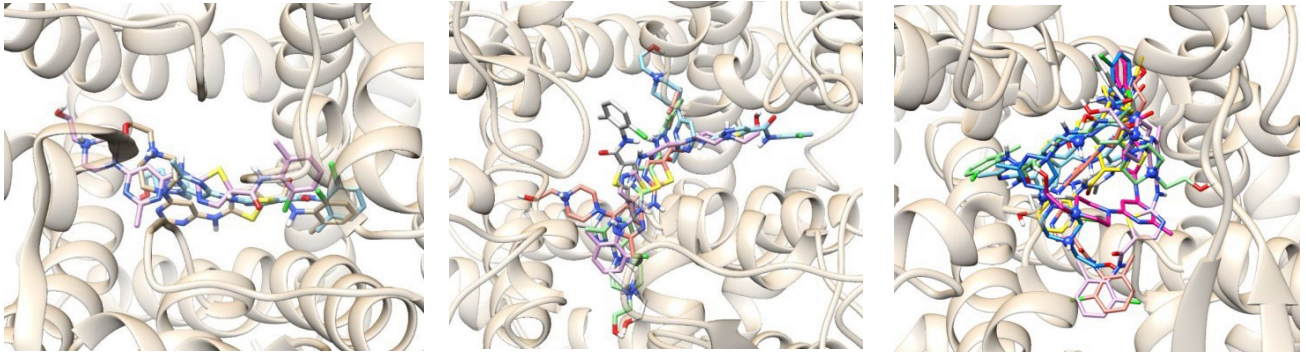


図④-2. 計算対象とした抗がん剤の構造。

ては Glide Induced-fit docking (IFD)、Nav1.5 については Glide XP を、Cav1.2 については Gnina 1.0 を用いた。ドッキングシミュレーションの結果、各系について複数のポーズが得られた。それぞれの系のポーズ群に対し非水素原子の RMSD を指標とした average linkage クラスタリングを実施した。カットオフ距離を hERG、Nav1.5 について 4 Å、Cav1.2 について 5 Å とした結果、各プロトン化状態について 1-8 種類のクラスタが得られた (表④-1)。一例として、図④-3 に dasatinib の 1つのプロトン化状態について各チャネルにおけるクラスタ代表構造を示した。続いて、各クラスタ代表構造に対し、平衡化計算を実施した。それぞれの複合体構造を水和した POPC 脂質二重層中に配置した後、50 ns の MD シミュレーションを、初速を変えて 5 回実施した。得られたスナップショット全体の中からチャネル-薬剤分子間の静電相互作用と van der Waals 相互作用の和が最も小さい (負で絶対値が大きい) ものを選択し、平衡化構造とした。

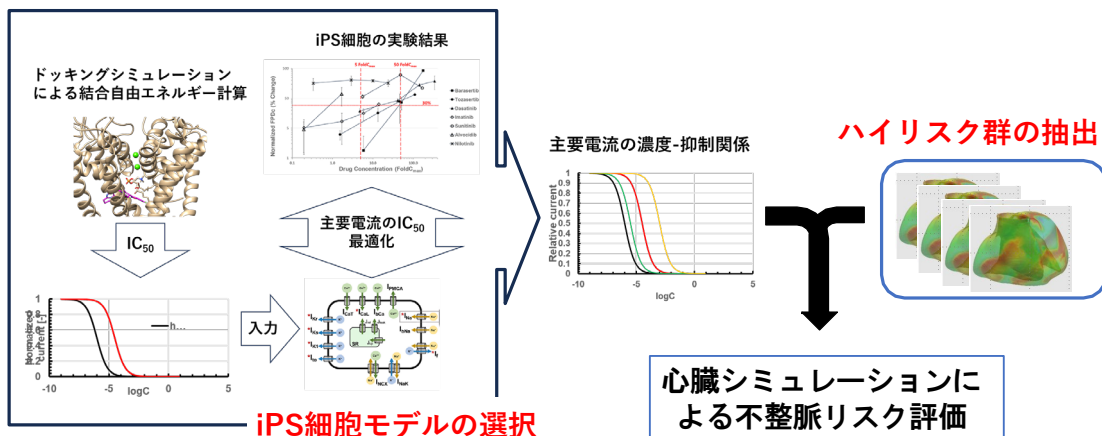
表④-1 各系におけるドッキングポーズのクラスタ数

Channel	Alvocidib		Dasatinib		Sorafenib	Sunitinib
	state 1	state 2	state 1	state 2		
hERG	6	2	3	3	5	7
Nav1.5	1	1	2	5	2	3
Cav1.2	5	7	7	8	4	2



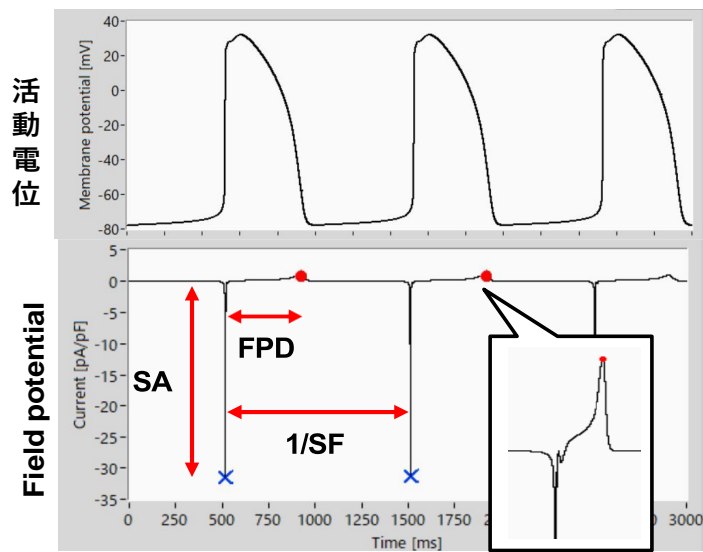
図④-3. Dasatinib (state 1)のドッキングポーズにおけるクラスタ代表構造。左、中央、右がそれぞれ hERG、Nav1.5、Cav1.2 チャンネルに対する結果を表す。茶色のリボンがチャンネル構造、異なる色のスティックモデルがそれぞれ異なるクラスタ代表構造を示す。

上述の分子シミュレーションから得られた薬剤のチャンネルへの結合自由エネルギー計算値から導かれる IC50 値を、iPS 細胞を用いた実験結果と比較対照することで確定し、確定した IC50 値を用いてハイリスク患者心臓モデルにおけるシミュレーションによる不整脈リスク評価を行うことが最終目的である。令和 5 年度はこの計画で iPS 細胞モデルの選定とその実装、およびハイリスク群の調査を進めた (図④-4)。



図④-4 インシリコ心毒性評価システム

これまでに報告されている主な iPS 心筋細胞モデルとしては(1) Paci et al. *Ann Biomed Eng* **41**, 2334-2348. (2013), (2) Koivumaki et al. *Front Physiol* **9**, 80. (2018), (3) Kernik et al. *J Physiol* **597**. 17 (2019)があるが(1)(2)についてはともに参照している電流の実験データが1編の論文のみであるのに対し(3)は INa, ICaL, IKr, Ito, IKs, If, IK1 電流のそれぞれにつき 2-3 の source を参照しそれらの平均モデルを作製している。さらに実験において予想される細胞間のばらつきについても考察しているなどの点を考慮して(3)を選択した。iPS 細胞を用いた薬理実験においては細胞外誘導から測定される Field potential duration (FPD)と Spike amplitude (SA)および Spike frequency (SF)が評価されるため、iPS 細胞モデルを実装し、計算結果からこれらの指標を自動で検出する application を作製した(図④-5,赤●、青×)。さらにこの機能を使用してパラメータ最適化のプログラムを開発中である。



図④-5 iPS細胞モデルと諸指標

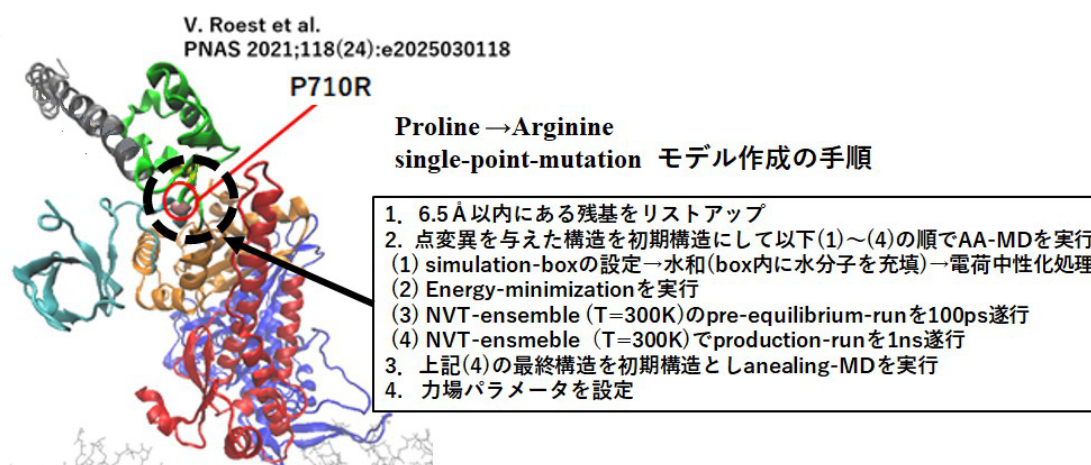
抗ガン剤の不整脈リスク評価を実施する予定である。

ハイリスク群の抽出については文献調査を行なった。不整脈は基質に誘因が加わることによって発生する。ガン患者固有の問題としては 1) 抗ガン剤の副作用による嘔吐、下痢および輸液がもたらす血清電解質異常 2) 抗菌薬、制吐剤、鎮痛剤などによる催不整脈作用の相乗効果および薬剤代謝酵素の抑制による血中濃度の上昇などの薬剤相互作用 3) 抗ガン剤による心筋障害（炎症、線維化）を確認した。次年度において心電図データベースから選んだ心電図QT 間隔の正常上限例にこれらの条件を組み合わせた 20程度の心臓をハイリスク例として生成し、

次世代アプリケーション開発グループ：

⑤ 遺伝子と臓器を結ぶ究極のマルチスケール心臓シミュレーション (UT-Heart 研 鷲尾、京大 岡本)：

本年度は、先ず肥大型心筋症の分子実験に関する研究報告調査と分子モデル検討を行なった。肥大型心筋症 (HCM) は遺伝性の心臓病であり、1,000 を超える変異と関連していると言われるが、その多くは β -心筋ミオシン (MYH7) にあるとされている。これまでの HCM 変異を持つミオシンの分子研究により、ATPase および負荷に応じたアクチンからの解離速度に対するさまざまな影響が検討されてきたが、このような多様な分子効果がどのように組み合わせられて細胞レベルの収縮力に影響を及ぼし、さらには HCM の表現型をもたらすかについては依然として研究途上にある。Morck ら (Morck et al. Hypertrophic cardiomyopathy mutations in the pliant and light chain-binding regions of the lever arm of human β -cardiac myosin have divergent effects on myosin function. eLife 2022;11:e76805) は、 β -心筋ミオシンのレバーアーム近傍の柔軟領域および軽鎖結合領域にある変異の効果を調査し、変異が柔軟領域にあるか軽鎖結合領域にあるかに応じて、ミオシンの機能に対して多様なアロステリック効果を及ぼすことを示した。特にレバーアーム近傍の柔軟領域の変異はミオシン分子の運動機能や出力にさまざまな影響を及ぼすことを明らかにした。また Roest ら (Roest et al. Hypertrophic cardiomyopathy β -cardiac myosin mutation (P710R) leads to hypercontractility by Disrupting super relaxed state. PNAS 2021 Vol. 118 No. 24 e2025030118) はモータードメイン内の P710R (710 番目残基の Proline が Arginine に変異) に焦点を当て、単一ミオシン分子の光トラップ測定により、この変異が運動速度と ATPase を劇的に低下させることを示した。またこの変異によりミオシンモーターのステップサイズとアクチン解離速度の負荷感受性が低下することを明らかにした。変異によるミオシンの負荷感受性の低下は前述の Morck らの論文でも指摘されている。以上の文献調査を踏まえ、今年度の研究対象として HCM に関わる変異として特に力学的観点から興味深い P710R を取り上げシミュレーションを実施することとした。粗視化分子モデルへの変異導入方法の手順を下図に示す。



図⑤-1 今回対象とする変異 P710R と粗視化分子モデル作成手順

シミュレーションを実施するに当たって行なった高速化・効率化は以下の通りである。

先ず、分子動力学計算の時間積分に関しては、これまで下図に示されるように慣性力が無視できる overdamped Langevin 方程式に対する通常の陽解法に較べ、粘性力項およびポテンシャル力項を工夫した準陰解法「Semi-implicit Hessian Correction (SimHec)」(Washio et al. Semi-Implicit Time Integration with Hessian Eigenvalue Corrections for a Larger Time Step in Molecular Dynamics Simulations. Journal of Chemical Theory and Computation. 2021;17(9):5792-804)の開発により 50 倍の時間刻み幅 (10 倍の高速化) が達成されてきたが、今年度はランダム力項にも補正を加えた手法「SimHec-RC」に発展させることにより更に 50 倍の時間刻み幅 (50 倍の高速化) が実現された (Washio et al. On random force correction for large time steps in semi-implicitly discretized overdamped Langevin equations. AIMS Mathematics 査読中)。

Overdamped Langevin Equation

$$\mathbf{G} \frac{d\mathbf{X}(t)}{dt} = -\frac{dU}{d\mathbf{X}}(\mathbf{X}) + \boldsymbol{\varepsilon}(t) \quad \boldsymbol{\varepsilon}(t)dt \otimes \boldsymbol{\varepsilon}(t)dt = 2k_B T \mathbf{G} dt$$

陽解法 (Explicit method, 伊藤積分に従う)

$$\mathbf{G} \frac{\mathbf{X}_h(t+h) - \mathbf{X}_h(t)}{h} = -\frac{dU}{d\mathbf{X}}(\mathbf{X}_h(t)) + \boldsymbol{\varepsilon}_h(t) \quad \langle \boldsymbol{\varepsilon}_h(t)h \otimes \boldsymbol{\varepsilon}_h(t)h \rangle = 2k_B T \mathbf{G} h$$

前富岳プロジェクトの成果



時間刻み幅を50倍 20fs→1ps(計算速度10倍)

準陰解法 (Semi-implicit Hessian Correction (SimHec), JCTC 2021 Washio et al.)

$$(\mathbf{G} + h\tilde{\mathbf{H}}) \frac{\tilde{\mathbf{X}}_h(t+h) - \tilde{\mathbf{X}}_h(t)}{h} = -\frac{dU}{d\mathbf{X}}(\tilde{\mathbf{X}}_h(t)) + \boldsymbol{\varepsilon}_h(t)$$

R5年度の成果



時間刻み幅をさらに50倍 1ps→50ps(計算速度さらに50倍)

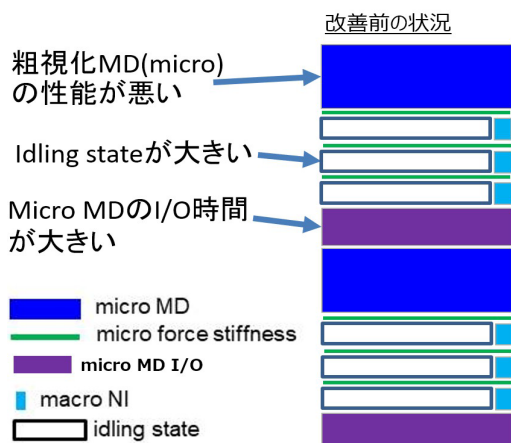
準陰解法 (SimHec-RC RC: random force correction)

$$(\mathbf{G} + h\tilde{\mathbf{H}}) \frac{\tilde{\mathbf{X}}_h(t+h) - \tilde{\mathbf{X}}_h(t)}{h} = -\frac{dU}{d\mathbf{X}}(\tilde{\mathbf{X}}_h(t)) + \boldsymbol{\varepsilon}_h(t) + \tilde{\boldsymbol{\varepsilon}}_h(t) \quad \begin{aligned} \langle \boldsymbol{\varepsilon}_h(t)h \otimes \tilde{\boldsymbol{\varepsilon}}_h(t)h \rangle &= 0 \\ \langle \tilde{\boldsymbol{\varepsilon}}_h(t)h \otimes \tilde{\boldsymbol{\varepsilon}}_h(t)h \rangle &= 2k_B T h^2 \tilde{\mathbf{H}} \end{aligned}$$

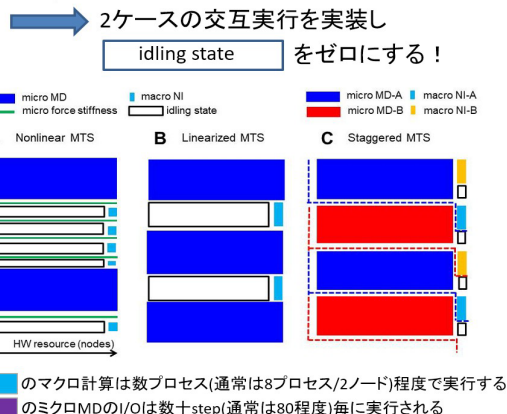
図⑤-2 分子動力学計算における高速化 (1)

次に、高速化・効率化の一環として MPI 並列区間の拡大を図った。これまでマトリクス生成と線形ソルバの計算は、1 プロセスのみが実行するシリアル処理となっていたため、本処理中は他の計算ノードが待ち状態となっていた。そこで、これらの処理の MPI 並列化に加え、待ち時間を利用して別の心臓モデルを並行して計算する「Staggered Multiple Time Step 法」を考案し、計算ノードのロードバランスを改善した。下図上段は、検討した手法ごとの計算資源のアイドル状態 (黒枠部分) を示したもので、Staggered Multiple Time Step 法では、ミクロの計算 (MD シミュレーション) とマクロの計算 (有限要素シミュレーション) を同時に処理することで、アイドル時間を大幅に削減することが可能となった。また、同図下段に示すように並列化計算のループ構造見直しを行なった。即ち、これまでの並列化計算においては、CG 粒子によるスレッド分割を行なっていたが、ループ長の観点からこれを MD 計算単位のスレッド分割にする等プログラムの大幅な書き換えによる高速化を実施した。以上の合計で、従来に比べ 5.82 倍の高速化が達成された。さらに、I/O 専用プロセスによる非同期 I/O により、I/O 時間を完全に隠蔽することが出来、60 万プロセスの実行時間が 1.49 倍に高速化された。

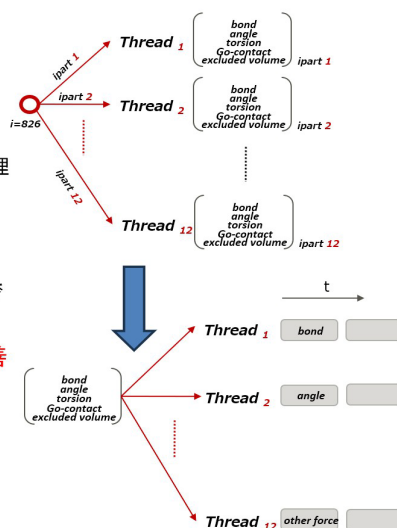
• 性能上の問題



Staggered Multiple Time Step Schemeの開発



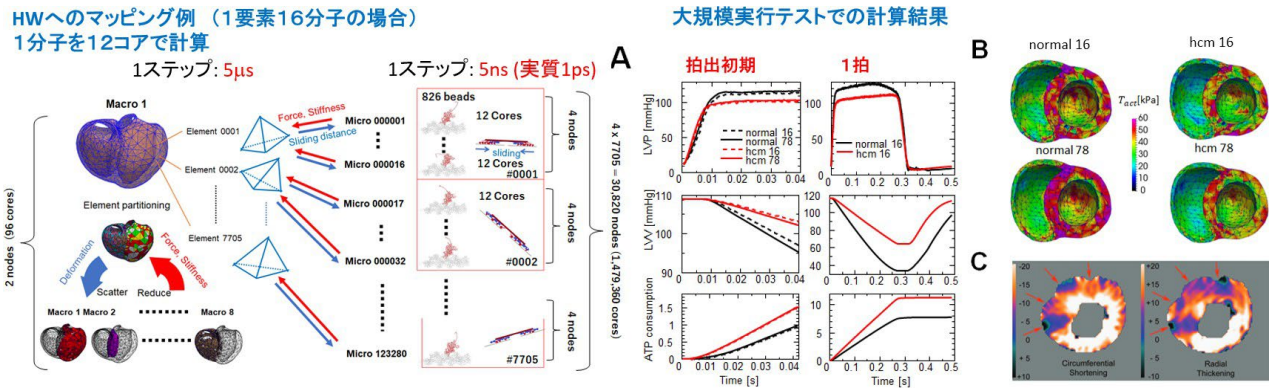
- 粗視化MDでは1分子内のCG粒子を振り分けて計算
 - アミノ酸に含まれるα炭素の位置へCG粒子を配置(=826粒子)
 - 計算対象となるCG粒子をスレッドへ分散し各計算を平行して処理(=約69粒子/スレッド)
- SIMD/ソフトウェアパイプライン(SWPL)が半数は適用できていない→**ループ長不足と命令スケジューリング不足**
- 並列化レベルをスレッド単位から関数単位へ引き上げることで**ループ長を確保**
- ループ構造を見直すことで**命令スケジューリングを大幅に改善**
 - 投機的プログラムによるブランチ命令の排除
 - 一時データ保存・ループ分割によるループ構造の単純化



図⑤-3 分子動力学計算における高速化 (2)

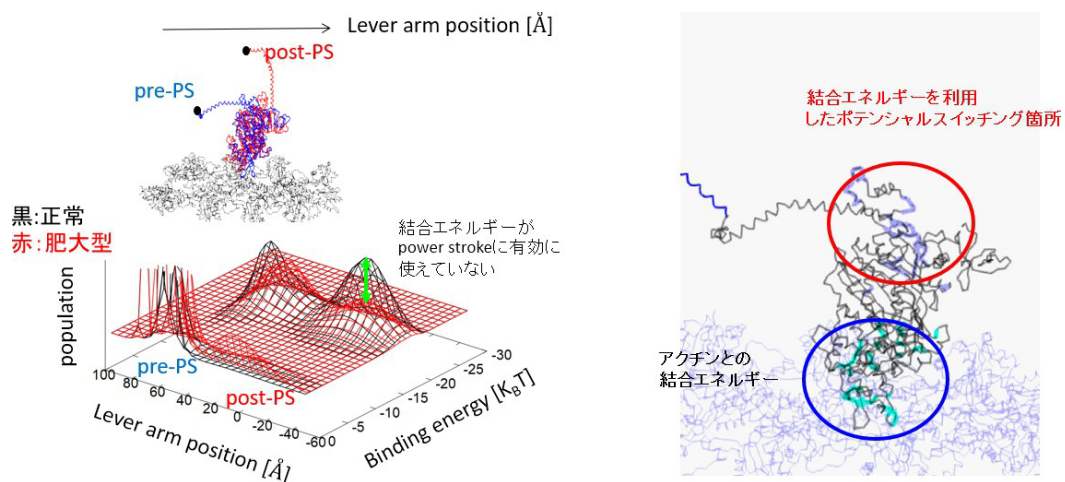
以上の高速化・効率化を経て、両心室モデルによる大規模シミュレーションを実施した。心臓は正常と HCM の 2 ケースを Staggered Multiple Time Step 法により同時実行し、実際にマクロ計算時における大多数のマイクロ担当ノードのアイドルングを回避できることを確認した。ここではマクロ 1 ステップ (5ms) につきマイクロ 1000 ステップを計算した。なお、本大規模実行においては分子動力学計算の時間積分には SimHec 法を用いたため、分子モデルのポテンシャル調整により状態遷移速度を高め 1ps 時間刻みを 5ns と見做してマクロ有限要素心臓モデルとカップリングした。マクロ有限要素心臓モデルの要素数は 7,705 要素である。1 有限要素あたりの分子数 (マイクロモデル) は 16~78 の範囲で変化させ分子数に対するマクロ解の収束状況を調べた。マクロ 0.5s の計算に 3 万ノードを使用して 24h を要した (下図左)。下図 A はシミュレーション結果の左心室圧、左心室容積、エネルギー消費について正常と HCM を比較したものである。なお正常と HCM 心臓の違いはミオシンの分子モデルの違いのみであり、有限要素モデルを始め他の条件は総て同一である。1 つのアミノ酸のみの置換がマクロレベルにおいて相応の違いをもたらす得ることを実際に示したことは意義があると考えられる。同図B 上段は 1 有限要素に 16 分子を、下段は 78 分子を割り当てた場合の収縮力の分布を示すが、いずれの分子数でも HCM にはムラ (空間

的なばらつき)が見られる。これに対し、同図 C は実際の HCM 心臓周方向と半径方向の歪み分布を計測したものであり (Aletras et al. Heterogeneity of intramural function in hypertrophic cardiomyopathy: mechanistic insights from MRI late gadolinium enhancement and high-resolution displacement encoding with stimulated echoes strain maps. Circulation Cardiovascular imaging, Vol. 4, No. 4, pp. 425-34, 2011)、シミュレーション結果との類似性が指摘される。



図⑤-4 正常および HCM 心臓シミュレーション結果の比較

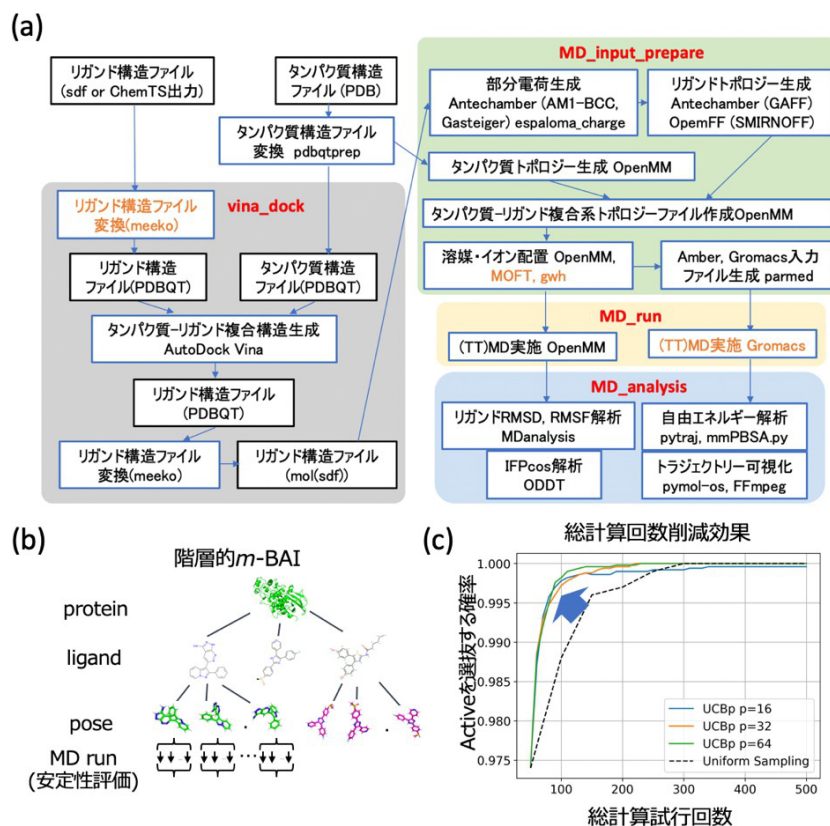
本ミオシン分子モデルは、下図右に示されるように、アクチンとの結合エネルギーを利用して分子内の一部領域のポテンシャルをスイッチしている。同図左は、パワーストローク前後の結合エネルギーを正常と HCM で比較したものであり、その頻度分布から、正常に対し HCM では結合エネルギーがパワーストロークに有効に使えていないことが分かる。前出の図 A に示されるようなマクロレベルにおける拍出性能の差異をマイクロレベルから順を追って分析・説明できるところにマルチスケールシミュレーションの真の意義がある。



図⑤-5 正常および HCM 心臓の分子レベルでの効率の比較

⑥MD×AI に基づく次世代創薬計算技術開発（横浜市大 寺山、京大 松本）：

本課題では、MD を含む先進的分子シミュレーションを用いた活性・物性値推定を活用し、現状のデータ駆動型創薬の欠点を克服する新たな創薬 AI 手法の開発を行う。既存のバーチャルスクリーニングでは主に分子ドッキングが用いられるが、鋳型(タンパク質)は基本的に固定して取り扱うため、タンパク質の柔軟なダイナミクスに起因するリガンドとタンパク質間の複雑な相互作用を十分に取り込むことは難しい。ドッキングでは偽陽性、つまり実験的には結合親和性が弱いにも関わらずドッキング計算上は結合親和性が高いと判定されてしまう例が多く、バーチャルスクリーニングにおける大きな課題となっている。ドッキングした構造から短時間でも MD 計算に基づくリガンドの安定性評価などにより、ドッキングでは見積もれない結合親和性の強さを評価できれば、簡易でより高精度なバーチャルスクリーニングにつながる可能性がある。また、リガンド-タンパク質間の結合親和性の推定はスクリーニングだけでなく、hit 化合物が発見された後の hit-to-lead あるいは lead 最適化の際にも重要になる。特に近年発展している分子生成 AI による設計を考えた時に、活性の実験データが非常に少ない、あるいは無い場合は、分子シミュレーションによるコストパフォーマンスに優れた結合親和性推定が必要になる。しかし、MD 計算はドッキングシミュレーションや機械学習による予測に比べると計算コストが大きいいため、総計算コストを削減する工夫や「富岳」の大規模な計算リソースを効率よく利用する方法論が求められる。そこで、本課題では分子シミュレーション技法と、生成 AI や強化学習など様々な AI 手法を連携させることにより、効果的な創薬 AI 手法の開発を目指す。具体的には、MD（分子動力学計算）に基づく高精度なバーチャルスクリーニング技術(MAVS: MDxAI-based Virtual Screening)と、薬候補分子を計算機上で設



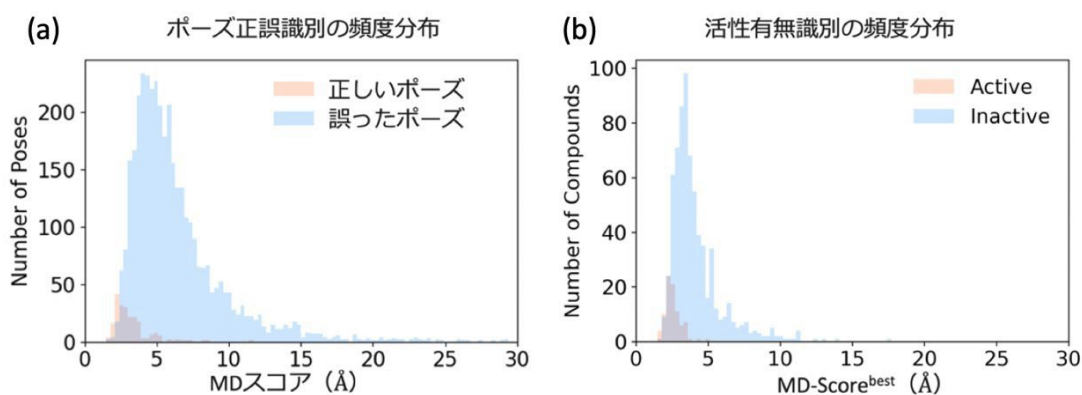
図⑥-1. MAVS のワークフローと強化学習による計算最適化。(a)短時間 MD 計算による結合親和性評価のワークフロー。(b)結合ポーズを考慮したリガンド選抜の概要図。(c)階層的 m -BAI による計算効率化の検証例。

計するために化合物生成 AI と分子シミュレーションを融合した手法(MADD: MDxAI-based Drug Design)の開発に取り組む。以下、MAVS と MADD の R5 年度の成果について記述する。

MAVS に関して、ワークフローの実装、ワークフローの「富岳」での動作確認、6 種類のタンパク質を対象にタンパク質-リガンド複合体の短時間 MD の計算評価を実施した。まず、MD 計算に基づくバーチャルスクリーニングを「富岳」上で実施するためには、多数のリガンドに対して自動的に計算を実施するワークフローを構築する必要がある。そこで、リガンドの構造ファイルとターゲットタンパク質の構造ファイルのみを入力として、ドッキング計算・MD 計算・結合親和性の評価を行うワークフローを実装した(図⑥-1 (a))。また、このワークフローが「富岳」上で動作することを確認した。実際に MD 計算を用いてバーチャルスクリーニングをする場合、非常に大きな計算コストが掛かってしまうため、強化学習の一種で複数の候補から最も有望な m 個を効率的に探索する手法である m Best-Arm Identification (m -BAI)を発展させた階層的 m -BAI を新たに提案した(図⑥-1 (b))。実際に階層的 m -BAI による計算削減効果を簡易的な例で検証しており(図⑥-1 (c))、何も工夫しない場合(Uniform Sampling)に比べ少ない計算回数でより高精度に活性のあるリガンドの選出が可能であることを確認している。

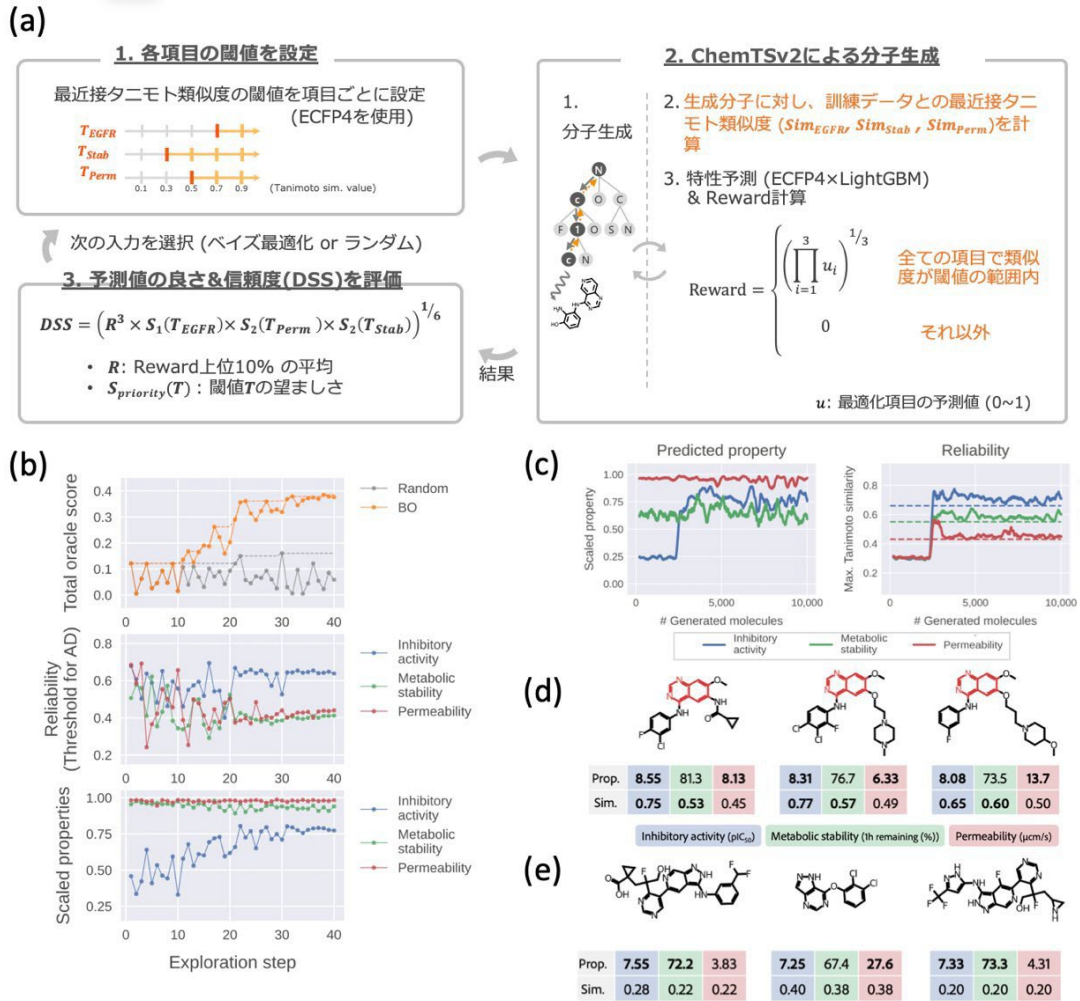
さらに、「富岳」上で、6つのターゲットタンパク質(EGFR, PTPN1, ERK2, PPARG, BACE1, P38)に対して MAVS の短時間 MD 計算を行い、リガンド-タンパク質複合体の安定性評価の方法論について検討を行った。まず、正確にリガンド-タンパク質複合体の結合親和性を推定するためには、結合ポーズを高精度に求める必要がある。そこで各ターゲットタンパク質に対して、19 個以上、総計で 425 個のリガンドに対して結合ポーズ検証用データセットを作成し、評価を実施した。ドッキング計算で得られた 10,172 個の複合体構造に対して 5 ns の MD 計算(各リガンドに対し 5 初速度)を実施し、リガンドの RMSD を安定性スコアとして評価した。その結果、図⑥-2 (a)に示すように正しいポーズは安定している一方、誤ったポーズは比較的不安定で、ドッキング計算で得られた正誤判別精度 (60.2%) より MD 計算による精度 (71.0%) が高いことを確認できた。また、活性値の有無についても Active/Inactive 化合物を集めたデータセットを用意し、同様の計算手法で評価した結果、図⑥-2 (b)に示すように Active 化合物は安定しており、Active 化合物を判別する ROC-AUC スコアは 0.89 と非常に高いことを確認した。今後より詳細に、計算時間や評価手法・データセットの検討を実施する予定である。

MADD に関して、データ駆動型の分子生成 AI の課題を明確にするために予測信頼度を考慮した分子生成 AI の開発を行い、また分子生成 AI に GPU を用いた分子ドッキングや拡散モデルを用いたドッキング手



図⑥-2. MAVS の短時間 MD 計算による結合ポーズおよび活性の評価。(a)正しいポーズおよび誤ったポーズの MD 計算による安定性評価スコアの分布。(b)Active および Inactive 化合物の MD 計算による安定性評価スコアの分布。

法の組み込みを実施した。データ駆動型分子生成 AI の課題として、適用範囲を考慮せずに分子生成を行なった場合、生成 AI が学習データとかけ離れた分子構造（適用範囲外の構造）を生成しても、予測モデルはある程度の値を返す。そのため、活性や ADMET などの特性が良好な値を持っているように見えても、実際には適用範囲外で、実際にはそれらの値を持たない分子をモデルが生成する可能性がある。さらに、創薬で必要な複数の項目を考慮する多目的最適化の場合、項目ごとに学習データで定義される適用範囲が異なるため、すべての項目で適用範囲内に収めながら特性値を達成することは非常に難しい作業となる。そこで、これまで開発している分子生成 AI ChemTSv2 に、適用範囲を考慮しながら分子生成を行う手法 DyRAMO を実装した。DyRAMO は予測が信頼できる範囲を信頼度という形でパラメータ化し、最適化項目ごとに適切な信頼度を自動的に調整しながら多目的最適化を行う手法である (図⑥-3 (a))。DyRAMO の目標は、全項目の信頼度基準を満足しながら、予測値が最適化された分子を生成することである。目標を達成するために、以下のプロセスを繰り返す： 1. 各項目の信頼度 (閾値) を決定し、2. ChemTSv2 で分子を生成し、3. 目標達成度を評価する。信頼度を無闇に探索するのは非効率であると想定されるため、ベイズ最適化を用いた効率的な探索アルゴリズムを実装した。EGFR の活性値と代謝安定性、膜透過性を対象として、ベイズ最適化で最適な信頼度閾値を探索した例を図⑥-3 (b)に示す。ベイズ最適化によって目標 (Total oracle score) がランダムに比べて効率よく上昇していることがわかる。また最適化された信頼度閾値での分子生成の様子 (図⑥-3 (c)) と生成された分子例 (図⑥-3 (d)) から各特性の最適化と信頼度 (訓練データとの類似度の高さ) が両立していることがわかる。信頼度を考慮しないと図⑥-3 (e) のように特性値だけが最適化されることと比較すると、DyRAMO の効果が大きいことが確認できる。今後 DyRAMO で予測信頼度が一定の低い候補に関して MD 計算による活性予測値を付与するなどして、信頼度と速度を両立した生成 AI 手法の開発を予定している。また、分子生成 AI と計算手法の連携として、既存のドッキング手法よりも数十倍高速な GPU を用いたドッキング手法を ChemTSv2 の評価関数に組み込み、また、精度が高いと期待される拡散モデルを用いた分子ドッキング手法 (DiffDock) も評価関数に実装した。



図③-3. 予測信頼度を考慮した化合物生成手法の生成。(a)創薬に求められる特性の最適化と予測信頼度を両立するための枠組み。最適な信頼度(訓練データとの類似度の閾値)をベイズ最適化により調整。(b)ベイズ最適化によるパラメータの探索の様子。(c)最も良い信頼度での分子生成の様子。(d)(c)で生成された分子の例。(e)信頼度を考慮せずに特性の最適化を行なった場合の例。

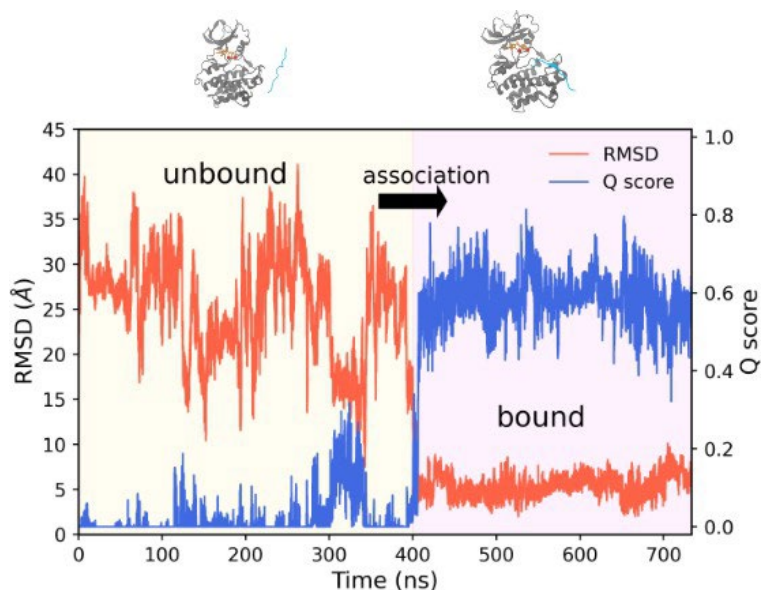
⑦ニューモダリティ創薬のための創薬計算手法開発（理研 千葉、阪大 信夫、東工大 Tran、筑波大 吉野）：

抗体・タンパク質医薬品・ペプチド医薬品・PROTAC などの新規モダリティ創薬に資する次世代計算手法の開発に取り組む。R5 年度は、薬剤-標的分子の複合体構造を再現する計算条件の検討及び構造アンサンブルを拡張サンプリングで取得する方法の検討を実施する。

(i) 中分子ペプチドへの拡張サンプリングの最適化

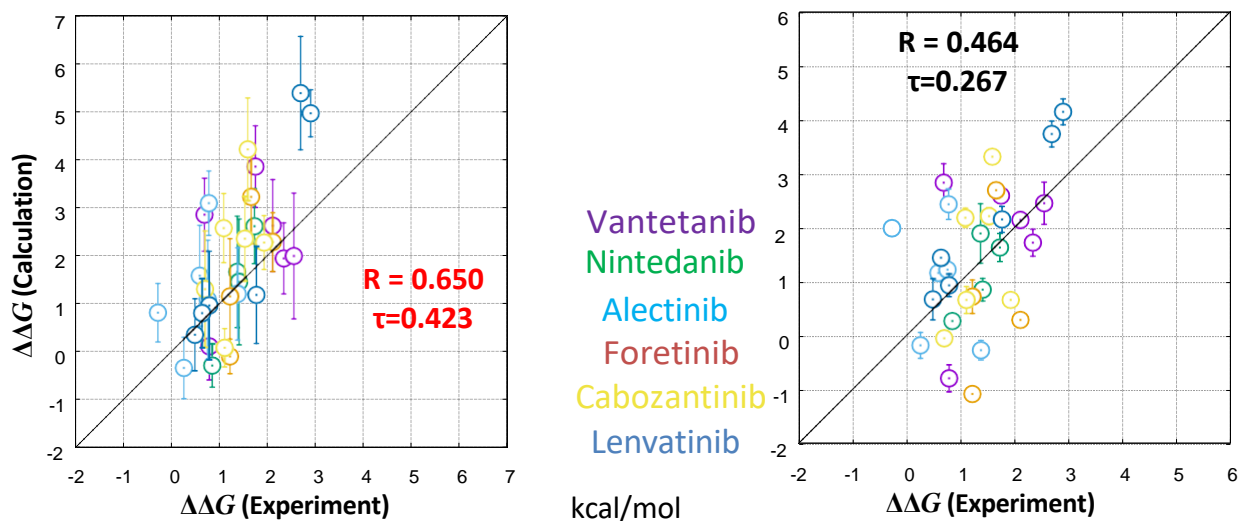
タンパク質-タンパク質間の広い面による相互作用のように、低分子では制御が困難である標的に対しては、より大きな相互作用面を有する抗体などの高分子による創薬アプローチが有効と考えられている。しかしながら、抗体の有する性質、すなわち細胞内分子を標的とできない点や経口薬として投与できないなどの課題を克服するために、低分子と高分子の中間（両方）の性質を有する中分子ペプチドの開発が活発である。

中分子ペプチドの開発のためには、構造に基づく分子デザインが、低分子医薬品開発の場合と同様に、重要である。拡張サンプリング手法である gREST/REUS 法は、いままでのおもに低分子の標的タンパク質分子への結合のシミュレーションに適用されてきた[Shinobu et al., Front Mol Biosci 2022, Re et al., Proc Natl Acad Sci USA 2019]。しかしながら、gREST/REUS を中分子ペプチドと標的タンパク質の複合体の結合過程解析に適用するためには、系の大きさによる適切な Collective Variable (CV) の選び方、ペプチドの柔軟性によるサンプリングの不足といった困難がある。今年度は、この困難を、設定パラメータの試行錯誤による最適化により克服し、中分子ペプチドが解離状態から中間状態を経て標的タンパク質へ結合する過程のシミュレーションに成功した（図⑦-1）。今回は、特定のペプチド-タンパク質ペアについて試行したが、今後、この拡張サンプリングを中分子ペプチドとタンパク質の結合経路サンプリング法として一般的に適用できるようにし、ペプチド創薬の加速を目指している。



図⑦-1. 標的タンパク質（グレー）に対して中分子ペプチド（シアン）が結合する過程を拡張サンプリングによって実現できた。結合は 400 ns 前後で観測された。結合過程のシミュレーションの解析によって、結合力の強い性質や医薬品としてより好ましい性質を有するペプチドデザイン（アミノ酸変異導入）を実現できる。

医薬品候補ペプチドをデザインするためには、ペプチドと標的タンパク質が結合している状態での、ペプチドへのアミノ酸変異導入に伴う結合力の変化 ($\Delta\Delta G$) を見積もることができると便利である。2022 年度まで文部科学省より支援を受けた「富岳」成果創出加速プログラム「プレジジョンメディシンを加速する創薬ビッグデータ統合システムの推進」における、小野らの報告では、タンパク質-低分子化合物の変異導入に伴う結合力変化を算出する MutationFEP 法を開発し[Ono et al., Sci Rep 2020.]、「富岳」に実装して利用してきた。本プロジェクトでは、対象とする分子が、低分子から中分子ペプチドに変わること、探索する構造空間が広がることが予想されるため、2023 年度は拡張サンプリングを MutationFEP 法に組み合わせた Mutation FEP-gREST 法を開発した。この新手法は、実験結果を従来法より再現できることを示した (図⑦-2)。この開発は当初の計画を超えた成果である。この手法がタンパク質-ペプチド間の $\Delta\Delta G$ 計算の精度向上に寄与できれば、ペプチド医薬品の最適化などに応用できる。



図⑦-2. RET キナーゼについて、約 30 種の変異と薬剤の組合せの親和性変化の実験データを利用し、MutationFEP-gREST と従来法が、変異導入に伴う低分子-タンパク質間の結合力の変化 ($\Delta\Delta G$) 予測の性能を比較した。拡張サンプリングを取り入れた本プロジェクトの MutationFEP-gREST 法は性能 (相関係数) が向上することを確認できた。

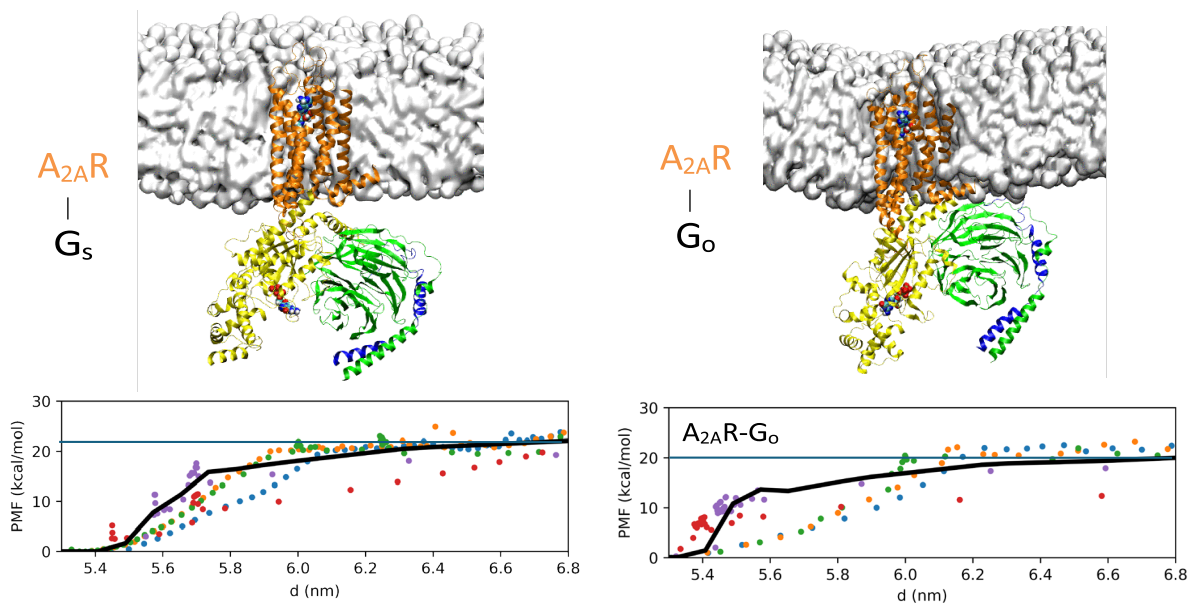
さらに、医薬品候補ペプチドをデザインするためには、ペプチド結合によって引き起こされる標的タンパク質の構造変化とそれに伴う機能変化 (「ペプチド結合-タンパク質構造-機能変化のダイナミクス」) をとらえられることも重要である。例えば、基質ペプチドが Src キナーゼに結合すると、Src キナーゼは構造変化とともに活性型から非活性型へ機能変化する。2023 年度は、Src キナーゼに拡張サンプリング手法である gREST を適用し、このダイナミクスの再現に成功した。

(ii) タンパク質-タンパク質相互作用の拡張サンプリングによる解析

高分子同士であるタンパク質-タンパク質間相互作用では、その結合過程や解離過程を中間状態も含めて原子レベルの構造情報を分子シミュレーションで取得することは、探索すべき構造空間の広さに起因する計算コストの増大によって難易度が高い。しかしながら、タンパク質-タンパク質間相互作用の結合解離過程の解析も、分子構造に基づく創薬にとって重要であることは変わらない。

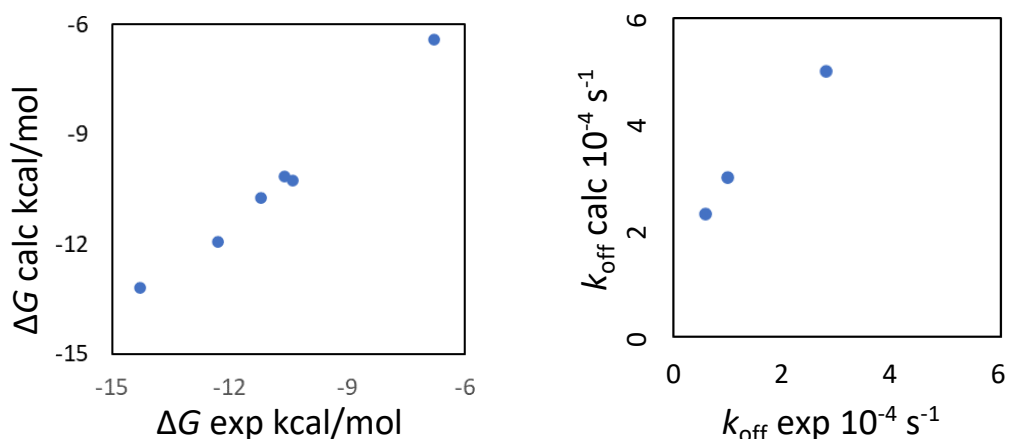
前項目 (i) で述べた低分子-タンパク質間の結合解離過程を解析できる手法は、gREST/REUS 以外にも提案されている。例えば、2022 年度まで文部科学省より支援を受けた「富岳」成果創出加速プログラム「プレジジョンメディスンを加速する創薬ビッグデータ統合システムの推進」における、北尾らの報告では、拡張サンプリング手法である PaCS-MD によって、タンパク質-低分子化合物の複合体の解離シミュレーションが実施可能であることが示されている [Hata, Tran et al, BPPB, 2021]。しかし、タンパク質-タンパク質複合体の解離のシミュレーションは計算コストの増大のため困難であった。そこで、2023 年度は、タンパク質-タンパク質複合体の解離過程のシミュレーションのために、PaCS/MD のパラメータ最適化を行なった。題材として選択した異なる二つの「タンパク質-タンパク質相互作用の結合力」

の相違の解析では、実験による観測データの再現に成功した (図⑦-3)。



図⑦-3. 膜 (グレーの表面表示) 中に存在する $A_{2A}R$ タンパク質 (オレンジ) に対して、2 種類の G タンパク質 (G_s (左) と G_o (右)) が解離する過程を拡張サンプリング (dissociation PaCS-MD) によって実現できた。d-PaCS-MD によるトラジェクトリの解析によって、 $A_{2A}R$ と G タンパク質の距離と平均力ポテンシャル (PMF) を算出できる。 $A_{2A}R$ -G タンパク質間距離が十分離れた位置と結合位置の PMF の差 (ΔG) (図下の緑線) の、 G_s と G_o での差 ($\Delta\Delta G =$ 約 2 kcal/mol) は、実験による観測データを再現できた。

拡張サンプリングである PaCS-MD を低分子-タンパク質間の結合解離過程への適用例を増やすことは、中分子ペプチド-タンパク質間の結合解離過程への応用時に役に立つ。そこで、dPaCS-MD を、6 種類の低分子と標的タンパク質について適用し、タンパク質-リガンド間 ΔG と k_{off} を再現することに成功した (図⑦-4)。



図⑦-4. A_{2A} 受容体と 6 種類の低分子化合物結合力 (ΔG) (左) のおよび解離速度定数 (k_{off}) (報告されている実験データは 3 種類) の実験値と計算値の比較 (右)。dPaCS-MD は実験データの定量的な再現に成功していることを示している。

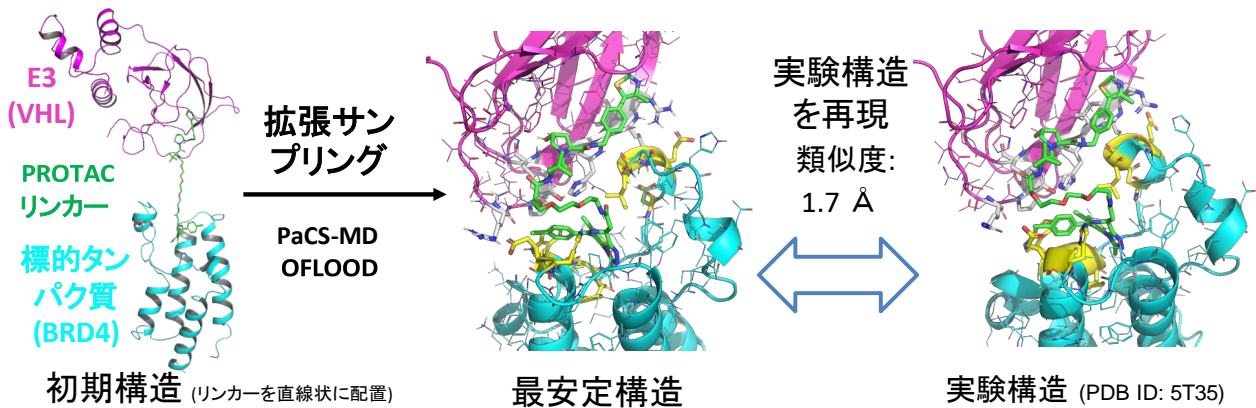
(iii) PROTAC デザインのための拡張サンプリング手法の最適化

E3 ユビキチンリガーゼと標的タンパク質を認識する機能性低分子である PROTAC (PROteolysis-Targeting Chimeras) は、標的タンパク質のユビキチン化を強制的に誘導することによって、標的タンパク質を細胞内から排除する新規創薬モダリティである。PROTAC の活性の発現には、PROTAC の、E3 結合部位と標的タンパク質結合部位を繋ぐ適切なリンカーの設計が重要であるため、複合体モデルの予測手法が報告されている。しかしながら、従来法はタンパク質を剛体として扱うため、リンカーを含めたタンパク質間相互作用や結合過程などが考慮されておらず、適切なリンカー選択の基準となる正確な複合体モデルを得られない。また、E3 と標的タンパク質は天然状態で結合するペアではなく PROTAC によって弱く安定化される結合界面をもつため、従来法でモデリングすることは困難である。そこで、本研究では拡張サンプリングである PaCS-MD を利用し、PROTAC 設計のための複合体モデル予測手法を開発している。

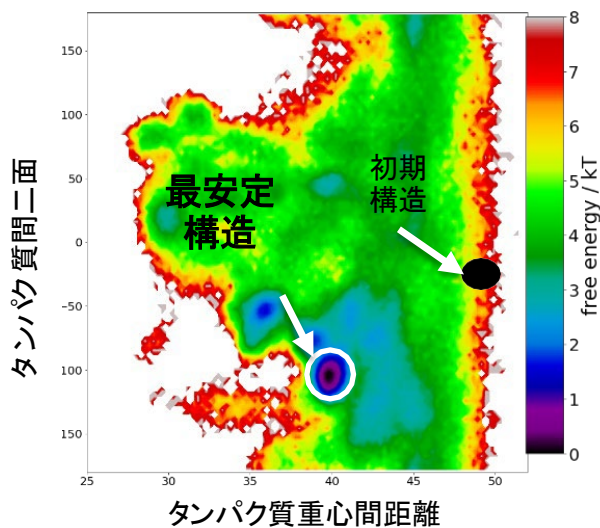
2023 年度は、PaCS-MD の設定を、標的・PROTAC・E3 複合体の構造予測に最適化した。具体的には、両タンパク質が接近する軌跡を採用する PaCS-MD と、出現確率が低い構造を初期構造として短時間 MD をリスタートさせる OFLOOD を適用することで、正確性の高い立体構造の再現に成功した (図⑦-5)。さらに、E3 と標的タンパク質の結合過程の自由エネルギー面の解析にも成功したことで、より詳細な分子デザインのための情報を取得できることを示した [Hirao et al., CBI 学会 2023 年大会、Excellent Poster Award、2023 年 10 月 24 日。Kudo et al., CBI 学会 2023 年大会、Excellent Poster Award、2023 年 10 月 24 日]。

以上のように開発した手法は、すでに具体的な創薬課題 (AMED BINDS 事業の研究支援課題「PROTAC 化合物の最適化とその検証」) において実践し、分子認識機構の解明及び高活性 PROTAC の設計支援を行な

っている。



図⑦-5. E3 ユビキチンリガーゼ (マゼンタ)・リンカーを直線状に配置した PROTAC (緑)・標的タンパク質 BRD4 (シアン) を初期構造 (左) にして拡張サンプリング法 PaCS-MD を実施した。得られた最安定構造 (中央) は実験構造 (右) を正確に再現できた。



図⑦-6. PROTAC のリンカーを直線状に配置して初期構造を構築し、PaCS-MD を実施した。得られた構造アンサンブルの解析から、E3-標的タンパク質間距離とタンパク質間二面角に対して、自由エネルギー面を描画し、最安定構造を決定できた。

(iv) AI による抗体の変異体の構造予測モデル開発

抗体医薬品は新しい創薬モダリティの中でも先行して創薬手法の開発が進んでいる。すでに、特別な修飾のない抗体（通常抗体）は医薬品売上 Top10 のうちのおよそ半分を占めるまでに成長している。こうした通常抗体に対して、現在は、抗体-薬物複合体（Antibody-drug conjugate, ADC）や、二重特異性をもつ抗体（Bispecific antibody）などの複雑な構造を有する抗体医薬品や、低分子化抗体などの開発が増加しており、効率的な抗体医薬品開発手法が求められている。

抗体医薬品開発の上流工程では、初期に得られた抗体の抗原に対する結合親和性を保ったまま、抗体の安定性などの多数の医薬品としての特性値を向上させる必要がある。このために、複数のアミノ酸を同時に変異させる協同的変異体[Chiba et al., Sci Rep, 2020]による、多数の特性値の同時最適化が有用と考えられるが、評価対象の変異体の組合せが膨大となる点が課題となる。そこで、「富岳」の多ノード性を活用するため、2023 年度は千葉らが開発したアミノ酸改変を導入した抗体の予測構造の精度を判定する機械学習プログラムを「富岳」に実装した。

また、この機械学習プログラムは、構造に基づくタンパク質間相互作用を記述子化するプログラム

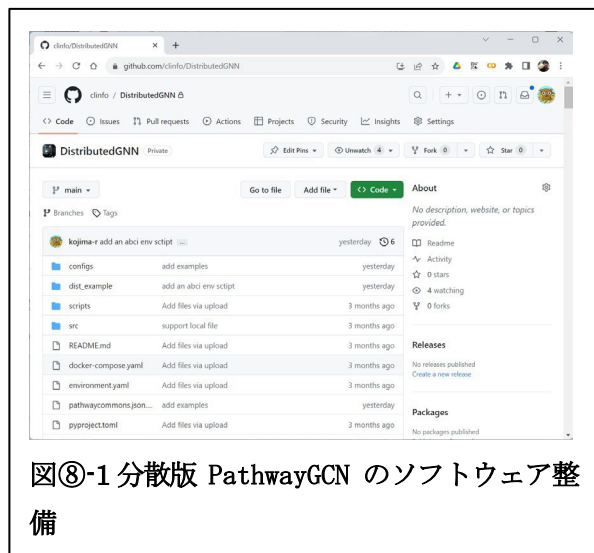
（intDesc）を利用している。この記述子を、通常分子動力学シミュレーションや gREST/REUS などの拡張サンプリング手法から得られる膨大な構造アンサンブルの相互作用解析にも利用するため、2023 年度は intDesc の「富岳」での大量並列実行の効率向上を実現した。さらに、この記述子化プログラムを、本プロジェクトメンバーが実行可能な計算環境を「富岳」上に整備した。

⑧大規模生体系ネットワーク処理技術開発（京大 小島、弘前大 玉田）：

本グループは、PathwayGCN および INGOR を用いた大規模生体系ネットワーク処理技術開発を行っており、以下にそれぞれの本年度での成果を記載する。

PathwayGCN の開発とソフトウェア整備

本年度は、PathwayGCN の基本的なクラスタマシン上での実行環境整備を行なった（右図）。より具体的には、データ共有、Pathwaycommons などのグラフを使った例題の提供、GNN（Graph Neural Network）によるリンク予測の事前学習、および「富岳」や ABCI（産総研スパコン）上での実行に関する検討を行なった。具体的には、大規模グラフをクラスタマシン上でのデータ共有を用いて分割したファイルを共有し、それらを用いた学習が容易に行える環境を整備した。また、Pathwaycommons などのグラフデータを活用した例題を用意し、実行環境の構築とデータの取り扱いについての実践的な手順例を提供した。



図⑧-1 分散版 PathwayGCN のソフトウェア整備

学習方法としては、GNN によるリンク予測の事前学習を採用し、この機能も併せてソフトウェアに組み込んだ。これにより、教師データが少量しかない場合においても、自己教師あり学習により、リンクの有無を予測するモデルを効率的に学習することが可能となった。「富岳」ノードにおいても、本システムの実行環境を整備した。「富岳」のハードウェア環境に合わせて最適化を行い、効率的な実行が可能となるよう検討を進めている。また、現在は分散に関する検討を行っており、「富岳」ノード上での分散実行も可能性の一つとして検討中である。

PathwayGCN の ABCI 上での複数ノード分散実行

GPU を用いた効率的な実行のために ABCI（産総研：AI 橋渡しクラウド）上でも、複数ノードを利用した分散実行環境を整備した。これにより、大規模なグラフデータに対する処理や GNN のトレーニングなど、計算資源を効果的に利用した高速な実行が可能となった。

PathwayGCN の他グループへの提供および機能強化

他グループとの連携として、PathwayGCN のコードの共有と利用可能にするための機能強化も行っており、特に、「①疾患多階層データベースの構築と応用」グループと共同で本ソフトウェアを活用した分子ネットワークを考慮した新しい PathwayGCN の活用方法である Net-DMPred (<https://github.com/clininfo/Net-DMPred>) を提案した。Net-DMPred では、個々の変異の特徴とグラフ部分で学習したグラフ特徴を組み合わせて、変異がドライバー変異であるかどうかの予測を可能にし、従来よりも高い性能で予測が可能であることを示した。

INGOR の開発とソフトウェア整備

INGOR はベイジアンネットワークを用いた大規模生体ネットワーク推定ソフトウェアである。本年度は、多サンプル実行時の処理時間の検討、多サンプル対応に向けたアルゴリズムの検討、高速化、そして高機能化を行なった。以下にその概要を記す。

INGOR のスレッド並列化と高速化

課題だったスレッド並列に対応した。スレッド並列化により、1スレッド 45 分の実行時間が 12 スレッドで 13 分まで短縮された。また数値計算コードの再検討とチューニングを行い、7.4 倍（実行時間 1/8）の高速化に成功した。

多サンプル実行時の処理時間の検討と対応アルゴリズムの検討

多サンプル実行の前提条件として 26000 サンプルでの実行時間の検討を行なった。「富岳」実測値を用いた検討では、384 ノード使用時に 75 日、12288 ノード使用時に 2 日と 8 時間、という見積もりであった。比較対象として、FX700 を利用したシングルノードでの測定からの外挿では 384 ノードで 9 時間、12288 ノードで 16 分であった。これは通信や同期コストを含まない値の理論性能と言えるので、今後のチューニングでどこまで近づけるかが一つの大きな目標となり得る。

INGOR 高機能化

前項の通り、多サンプル実行では、「富岳」の最大連続実行時間を超えてしまうため、現状では実行が不可能なことがわかった。そこで暫定的な対応として、計算の中断・再開機能を実装した。またネットワーク計算途中の全サブネットワーク出力機能を実装した。最終的なネットワーク構造だけでなく、途中のサブネットワーク計算結果を出力することで、これを GCN などによる機械学習のデータとして活用可能となり、これまで利用できなかった情報を用いたネットワーク構造の学習が可能になると考えられる。

プロジェクトの総合的推進（成果発表・広報普及・アウトリーチ）

令和5年度は、令和5年4月17日及び令和5年10月5日に上期の研究計画・進捗を発表する「成果創出ワークショップ」を課題内で開催した。

分野を超えた情報交換・交流としては、令和5年12月1日のシンポジウム「富岳百景」において、課題代表の奥野恭史が「富岳」を用いた研究成果として、『「富岳」で目指すシミュレーション・AI駆動型次世代医療・創薬』について紹介した。また、令和6年3月7日に開催された「第3回「富岳」成果創出加速プログラム 研究交流会」では、当課題から次世代研究者によるポスター発表に2名が参加した。更に、TV報道として、「計算で切り開く新時代!スーパーコンピューター富岳」というタイトルで当課題の研究内容がNHKサイエンスZERO（令和5年9月24日）に取り上げられた。また、富岳NEXTに向けて、生命科学・創薬・医療分野における計算科学及びAI for Scienceのロードマップ作成に協力した。

これらの活動を通して、連携する機関や研究者以外の一般市民、他の分野へも当課題の研究成果のプレゼンスを高めることができ、さらに若手研究者の人材育成を目指した活動につなげることができた。

また、研究成果の推進では、分子動力学シミュレーションソフトウェア「GENESIS」の活用を大学や研究機関、産業界などに普及するため、一般社団法人 ライフインテリジェンスコンソーシアムと連携し、令和5年5月16日、8月23日、11月30日にGENESIS講習会を開催した。更に、LINCの春季年会である「LINC SHOWCASE 2024 Spring」（令和6年3月21-22日）において創薬DXプラットフォームのデモをおこなうことで、産業界へ広くアウトリーチした。

2-3. 活動（研究会の活動等）

令和5年度は、以下の（1）ワークショップ（課題内進捗報告会）を主催した。

（1）ワークショップ（課題内進捗報告会）

概要	課題内 8 チームの研究進捗報告、富岳利用、ツール利用などの情報交換
日時	令和5年10月5日 13:00-16:00
参加者	本プロジェクト参加者 約 50 名
開催方法	対面（グランフロント大阪）

2-4. 実施体制

業務項目	担当機関	担当責任者
社会実装グループ：		
①疾患多階層データベースの構築と応用	京都府京都市左京区吉田本町 国立大学法人京都大学大学院医学研究科	鎌田 真由美 准教授 中澤 麻衣 特定助教
②インシリコ心疾患データベースの公開と応用	千葉県柏市若柴178-4-4 国立大学法人東京大学柏の葉キャンパス駅前サテライト3階 株式会社 UT-Heart 研究所	杉浦清了 取締役社長
	京都府京都市左京区吉田本町 国立大学法人京都大学大学院医学研究科	糀谷 泰彦 特定助教
③創薬 DX プラットフォームを用いた実践的創薬	京都府京都市左京区吉田本町 国立大学法人京都大学大学院医学研究科	荒木 望嗣 特定准教授 岩田 浩明 特定准教授 大塚 教雄 特定准教授
④インシリコ心毒性評価システムによる抗がん剤服用患者の安全性確保およびシステムの高精度化	東京都文京区本郷 7-3-1 国立大学法人東京大学大学院 農学生命科学研究科	寺田 透 准教授
	埼玉県和光市広沢 2 番 1 号国立研究開発法人 理化学研究所	井阪 悠太 特別研究員
次世代アプリケーション開発グループ：		
⑤遺伝子と臓器を結ぶ究極のマルチスケール心臓シミュレーション	千葉県柏市若柴178-4-4 国立大学法人東京大学柏の葉キャンパス駅前サテライト3階 株式会社 UT-Heart 研究所	鷺尾 巧 取締役副社長
	京都府京都市左京区吉田本町 国立大学法人京都大学大学院医学研究科	岡本 有司 特定助教
⑥MD×AIに基づく次世代創薬計算技術開発	横浜市鶴見区末広町 1-7-29 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
	京都府京都市左京区吉田本町 国立大学法人京都大学大学院医学研究科	松本 篤幸 特定准教授
⑦ニューモダリティ創薬のための創薬計算手法開発	埼玉県和光市広沢 2 番 1 号国立研究開発法人 理化学研究所	千葉 峻太郎 研究員

<p>⑧大規模生体系ネットワーク 処理技術開発</p> <p>プロジェクトの総合的推進 (成果発表・広報普及・アウトリーチ) :</p>	<p>大阪府吹田市山田丘 2 番 8 号テ クノアライアンス棟 AB 棟 5 階 A507 国立大学法人大阪大学 ヒューマン・メタバース疾患研 究拠点</p> <p>東京都目黒区大岡山 2-1-2- 1 M6-13 国立大学法人東 京工業大学 生命理工学院</p> <p>茨城県つくば市天王台 1-1-1 国立大学法人筑波大学 医学医 療系 生命医科学域</p> <p>京都府京都市左京区吉田本町 国立大学法人京都大学大学院医 学研究科</p> <p>青森県弘前市在府町 5 国立大 学法人弘前大学 医学研究科健 康未来イノベーションセンター</p> <p>京都府京都市左京区吉田本町 国立大学法人京都大学大学院 医学研究科</p>	<p>信夫 愛 特任准教授</p> <p>Tran Phuoc Duy 助教</p> <p>吉野 龍ノ介 助教</p> <p>小島 諒介 講師</p> <p>玉田 嘉紀 教授</p> <p>奥野 恭史 教授</p>
--	--	---

別添1 学会等発表実績

1. 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
1	Network-based prediction approach for cancer-specific driver missense mutations using a graph neural network	Narumi Hatano, Mayumi Kamada, Ryosuk Kojima, Yasushi Okuno	BMC Bioinformatics, Bio Med Central, Vol. 24 No . 1 pp. 1-15	10 October 2023
2	Inhibition of the bone morphogenetic protein pathway suppresses tumor growth through downregulation of epidermal growth factor receptor in MEK/ERK-dependent colorectal cancer	Shota Shimizu, Jumpei Kondo, Kunishige Onuma, Roberto Coppo, Kasumi Ota, Mayumi Kamada, Yohei Harada, Yoshihisa Tanaka, Mai Adachi Nakazawa, Yoshinori Tamada, Yasushi Okuno, Kenji Kawada, Kazutaka Obama, Robert J. Coffey, Yoshiyuki Fujiwara, Masahiro Inoue	Cancer science	25 June 2023
3	Low-energy defibrillation using a base-apex epicardial electrode	Jun-ichi Okada, Takumi Washio, Seiryō Sugiura, Toshiaki Hisada	Pacing Clin Electrophysiol 46:1509-1518	2023 September
4	Transition mechanisms from atrial flutter to atrial fibrillation during anti-tachycardia pacing therapy	Jun-ichi Okada, Takumi Washio, Seiryō Sugiura, Toshiaki Hisada	Pacing Clin Electrophysiol 46:1325-1332	2023 October
5	An explainable artificial intelligence-enabled electrocardiogram analysis model for the classification of reduced left ventricular function.	Katsushika S, Kodera S, Sawano S, Shinohara H, Setoguchi N, Tanabe K, Higashikuni Y, Takeda N, <u>Fujiu K</u> , Daimon M, Akazawa H, Morita H, Komuro I.	Eur Heart J Digit Health 17;4(3):254-264.	2023, April
6	Deep Learning Models for Predicting Left	Sato M, Kodera S, Setoguchi N, Tanabe K,	Circ J. 88(1):146-156	2023 Dec

	Heart Abnormalities From Single-Lead Electrocardiogram for the Development of Wearable Devices	Kushida S, Kanda J, Saji M, Nanasato M, Maki H, Fujita H, Kato N, Watanabe H, Suzuki M, Takahashi M, Sawada N, Yamasaki M, Sawano S, Katsushika S, Shinohara H, Takeda N, <u>Fujiu K</u> , Daimon M, Akazawa H, Morita H, Komuro I.		
7	A mutually communicable external system resource in remote monitoring for cardiovascular implantable electronic devices	Hasumi E, <u>Fujiu K</u> , Nakamura K, Yumino D, Nishii N, Imai Y, Shoda M, Komuro I.	Pacing Clin Electrophysiol. 2024 Jan;47(1):127-130	2024, January
8	A macrocyclic kinase inhibitor overcomes triple resistant mutations in EGFR-positive lung cancer	Suzuki M, Uchibori K, Oh-Hara T, Nomura Y, Suzuki R, Takemoto A, Araki M, Matsumoto S, Sagae Y, Kukimoto-Niino M, Kawase Y, Shirouzu M, Okuno Y, Nishio M, Fujita N, Katayama R	NPJ Precis. Oncol., 8(46), 1-13	2024年2月
9	「タンパク質の分子動力学シミュレーションに基づく創薬計算技術の開発とゲノム医療分野での応用研究」	荒木望嗣	SAR News No. 45, Cutting Edge, p10-18	2023年10月1日
10	Establishment of an MR1 Presentation Reporter Screening System and Identification of Phenylpropanoid Derivatives as MR1 Ligands	Matsuoka T, Hattori A, Oishi S, Araki M, Biao Ma, Fujii T, Arichi N, Okuno Y, Kakeya H, Yamasaki S, Ohno H, Inuki S	Journal of medicinal chemistry, 2023, 66(17), 12520-12535	2023年8月28日
11	Quantitative analysis of protein dynamics using a deep learning technique combined with experimental cryo-EM density data and MD simulations	Shigeyuki Matsumoto, Shoichi Ishida, Kei Terayama, Yasuhshi Okuno	Biophysics and Physicobiology 20(2), e200022.	2023年6月
12	Application of In Silico Technologies for Drug Target Discovery and Pharmacokinetic Analysis	Iwata H	Chemical and Pharmaceutical Bulletin, Vol. 71 No. 6, 2023.	2023年6月1日

13	Improving Compound-Protein Interaction Prediction by Self-Training with Augmenting Negative Samples	Koyama T, Matsumoto S, Iwata H, Kojima R, Okuno Y	Journal of chemical information and modeling, 63.15: 4552-4559, 2023	2023年7月17日
14	VGAE-MCTS: a New Molecular Generative Model combining Variational Graph Auto-Encoder and Monte Carlo Tree Search	Iwata H, Nakai T, Koyama T, Matsumoto S, Kojima R, Okuno Y	Journal of chemical information and modeling, 63.23: 7392-7400, 2023	2023年11月22日
15	Design of antimicrobial peptides containing non-proteinogenic amino acids using multi-objective Bayesian optimisation	Y. Murakami, S. Ishida, Y. Demizu, K. Terayama	Digital Discovery, 2023, 2, 1347-1353	2023/8
16	ChemTSv2: Functional molecular design using de novo molecule generator	S. Ishida, T. Aasawat, M. Sumita, M. Katouda, T. Yoshizawa, K. Yoshizoe, K. Tsuda, K. Terayama	WIREs Computational Molecular Science. 2023. e1680	2023/7

2. 国際会議・シンポジウムにおける口頭・ポスター発表

No.	発表した成果（発表題目、口頭・ポスター発表の別）	発表者氏名（所属機関）	発表した場所（学会名等）	発表した時期
1	Network-Based Prediction Method for Cancer Driver Missense Mutations.	Narumi Hatano, Mayumi Kamada, Ryosuke Kojima, Yasushi Okuno	Conference: Intelligent Systems for Molecular Biology/European Conference on Computational Biology (ISMB/ECCB) 2023	25 July 2023
2	Cancer prognosis prediction integrating gene expression data and a background biomolecule network	Kazuma Inoue, Ryosuke Kojima, Mayumi Kamada, Yasushi Okuno	Conference: Intelligent Systems for Molecular Biology/European Conference on Computational Biology (ISMB/ECCB) 2023	25 July 2023
3	Latent State Estimation of Cancer Patients Treated with Nivolumab Using Deep State Space Model	Aya Nakamura, Ryosuke Kojima, Yuji Okamoto, Yohei Mineharu, Yohei Harada, <u>Mayumi Kamada</u> , Yasushi Okuno	Conference: Intelligent Systems for Molecular Biology/European Conference on Computational	25 July 2023

			Biology (ISMB/ECCB) 2023	
4	A Novel Digital Health Technology “ped UT-Heart” to Support Decision-Making of Surgical Procedures for Congenital Heart Disease (口頭)	Isao Shiraishi, Kenichi Kurosaki, Shigeru Iwai, (National Cerebral and Cardiovascular Center) Takumi Washio, Seiryu Sugiura, Toshiaki Hisada (UT-Heart Inc.)	The 88 th Annual Meeting of the Japanese Circulation Society	2024年3月9日
5	Ped-UT-Heart project supports precision medicine for congenital heart disease by providing detailed anatomical and functional information (口頭)	Seiryu Sugiura Jun-ichi Okada Takumi Washio Toshiaki Hisada (UT-Heart Inc.) Isao Shiraishi, (National Cerebral and Cardiovascular Center)	The 101 th Annual Meeting of The Physiological Society of Japan	2024年3月28日
6	Remote Monitoring of Single-lead Electrocardiography Enables Detection of Heart Failure Status	Katsuhito Fujiu University of Tokyo	American Heart Association Scientific Sessions 2023	2023年11月1日
7	A MULTISCALE MOLECULAR SIMULATION APPROACH FOR ASSESSING THE EFFECTS OF MUTATION ON PROTEIN ACTIVITY AND DRUG SENSITIVITY OF ONCOGENIC KINASES (口頭発表)	<u>Araki M</u> and Okuno Y	Biophysical Society 68th annual meeting	2024年2月13日
8	Structural origins of site-specific inhibitor binding to Dihydrofolate Reductase unveiled by molecular dynamics simulations combined with Markov state modeling (口頭発表)	<u>Araki M</u> and Okuno Y	Spring 2024 ACS National Meeting	2024年3月17日
9	Iterative Data Augmentation of Near Boundary Negative Samples Improves the Model Generalizability in Compound-Protein Interaction Prediction.	Takuto Koyama, Shigeyuki Matsumoto, Hiroaki Iwata, Ryosuke Kojima, Yasushi Okuno.	Intelligent Systems for Molecular Biology/European Conference on Computational Biology (ISMB/ECCB) 2023, Centre de Congrès de Lyon, France	2023/07/24

10	Integrated platform development on Fugaku for drug target discovery	Iwata H	第 97 回日本薬理学会 年会	2023 年 12 月 14 日
11	医薬品開発プロセスにおける分子設計	岩田浩明	日本化学会第 104 春季 年会	2024 年 3 月 20 日
12	事前学習済み畳み込みニューラルネットワークを活用した医薬品添加剤のクラスタリング (ポスター)	岩田浩明, 林祥弘, 小山拓豊, 長谷川亜樹, 寺山慧, 奥野 恭史	日本薬剤学会第 38 年 会, Aichi, Japan	2023/5
13	Comprehensive Clinical Risk Assessment of Drug Candidate Compounds using ADMET-AI Models.	Shoma Ito, Hiroaki Iwata, Shigeyuki Matsumoto, Ryosuke Kojima, Yoshinobu Igarashi, Masataka Kuroda, Hitoshi Kawashima, Reiko Watanabe, Hiroshi Yamada, Kenji Mizuguchi, Yasushi Okuno.	日本薬物動態学会第 38 回年会/第 23 回シト クロム P450 国際会議 国際合同大会	2023 年 9 月 25 日 (月) ~ 29 日 (金)
14	A New Molecular Generation Model Combining Deep Learning and Reinforcement Learning,	Hiroaki Iwata, Taichi Nakai, Takuto Koyama, Shigeyuki Matsumoto, Ryosuke Kojima, Yasushi Okuno	情報計算化学生物学会 2023 年大会(CBI2023), Tokyo, Japan	2023/10
15	Exploring Chemical Structural Insight of ADME Properties via Interpretable Deep Learning (ポスター)	Shoma Ito, Takuto Koyama, Hiroaki Iwata, Shigeyuki Matsumoto, Ryosuke Kojima, Yasushi Okuno.	情報計算化学生物学会 2023 年大会(CBI2023), Tokyo, Japan,	2023/10
16	Insight into Federated Learning for Compound-Protein Interaction Prediction (口頭発表)	Takuto Koyama, Hiroaki Iwata, Shigeyuki Matsumoto, Ryosuke Kojima, Takao Otsuka, Aki Hasegawa and Yasushi Okuno.	情報計算化学生物学会 2023 年大会(CBI2023), Tokyo, Japan,	2023/10/ 25
17	Practical Evaluation of Protein-Compound Interactions with Computational Approaches (口頭発表)	Matsumoto S	第 97 回日本薬理学会 年会	2023 年 12 月 14 日
18	招待講演 マルチスケール心臓シミュレータ UT-Heart はどのように生物物理と医療を結ぶ懸け橋となりうるか	鷺尾巧 (株式会社 UT-Heart 研究所)	第 37 期 CAMM フォー ラム」11 月例会本例会	Nov. 2023
19	招待講演 心臓シミュレータ UT-Heart はどのように医学・医療に貢献できるか	鷺尾巧 (株式会社 UT-Heart 研究所)	第 22 回 日本心臓血 管発生研究会	Dec. 2023

20	口頭発表 分子モーターの確率的協調性を組み込んだマルチスケール心臓シミュレータのマクロスコピックな医療指標での精度について	鷺尾巧 (株式会社 UT-Heart 研究所)	第 101 回日本生理学会大会	Mar. 2024
21	分子生成 AI による創薬に向けて: 多目的最適化の課題と展望 (口頭、招待講演)	寺山慧 (横浜市大)	よこはま NMR 研究会第 73 回ワークショップ「AI 創薬」	2024 年 3 月 22 日
22	生成モデルと強化学習による 分子設計: 創薬から材料まで (口頭、招待講演)	寺山慧 (横浜市大)	神奈川県・横浜市・川崎市主催オンラインセミナー AI 創薬 before /after	2024 年 1 月 16 日
23	複数の特性予測モデルの信頼性を考慮した分子の多目的最適化 (口頭、SAR Presentation Award)	吉澤竜哉 (横浜市大)、石田祥一、佐藤朋広、大田雅照、本間光貴、寺山慧	第 51 回構造活性関連シンポジウム	2023 年 11 月 21 日
24	How the bacterial Flagellar Motor stator works: an insight from molecular dynamics simulation (Oral talk)	D. P. Tran, A. Kitao (Tokyo Tech, School of Life Science and Technology)	The 34th IUPAP Conference on Computational Physics (Kobe, Japan)	2023/8/4-8
25	Sampling the conformational transition of the monomer of Nsp15 of the SARS-nCOV-2 gives a hint to inhibit its hexamer (Invited talk)	D. P. Tran, A. Kitao (Tokyo Tech, School of Life Science and Technology)	The 1st International Symposium of Nano Life Science: Nanobiotechnology, Biosenser, Computation (Quy Nhon, Vietnam)	2023/9/17-19
26	Unveiling the dynamics of Adenosine A2A receptor coupling to G proteins: a molecular dynamics simulation (Poster)	D. P. Tran, S. Hagimoto, A. Kitao (Tokyo Tech, School of Life Science and Technology)	The 61th Annual Meeting of the Biophysical Society of Japan (Nagoya, Japan)	2023/11/14-16
27	Extensive sampling of protein-inhibitor binding landscapes using molecular dynamics simulations (Oral talk)	A. Shinobu ^{1,2} , S. Re ³ , H. Oshima ^{2,4} , Y. Sugita ² ¹ Osaka University, WPI Premium Research Institute for Human Metaverse Medicine (WPI-PRIME), ² RIKEN BDR, ³ National Institutes of Biomedical Innovation, Health, and Nutrition, ⁴ University of Hyogo	ICMS, Taipei Taiwan	2023/10/7
28	Atomic-Level Characterization of Protein Kinase Inhibitors binding through massive	A. Shinobu ^{1,2} , S. Re ³ , H. Oshima ^{2,4} , Y. Sugita ² ¹ Osaka University, WPI Premium Research Institute	The 61th Annual Meeting of the Biophysical Society of Japan, Nagoya, Japan	2023/11/15

	Molecular Dynamics Simulations (Poster)	for Human Metaverse Medicine (WPI-PRIME), ² RIKEN BDR, ³ National Institutes of Biomedical Innovation, Health, and Nutrition, ⁴ University of Hyogo		
29	Development of a discriminative model to determine the structural accuracy of an antibody-antigen mutant models based on geometrical interaction descriptors (Poster)	S. Chiba, Y. Okuno, M. Ikeguchi, M. Ohta (RIKEN, R-CCS)	The 6th R-CCS International Symposium	2024/1/29
30	A novel method for gene knockdown simulation using Markov chain Monte Carlo methods (MCMC) (ポスター発表)	中澤 麻衣, 玉田 嘉紀, 奥野 恭史	第12回 生命医薬情報学連合大会 (IIBMP 2023) @ 柏の葉カンファレンスセンター (千葉県柏市)	Sep. 7-9, 2023
31	Multi-omics network analysis for Elucidating Drug Sensitivity Mechanisms of Individual Cancer Patients) (ポスター発表)	太田 可純, 原田 陽平, 中澤 麻衣, 玉田 嘉紀, 奥野 恭史	第12回 生命医薬情報学連合大会 (IIBMP 2023) @ 柏の葉カンファレンスセンター (千葉県柏市)	Sep. 7-9, 2023
32	INGOR / SiGN-BN - Gene Network Estimation Software (ポスター発表)	玉田 嘉紀, 中澤 麻衣, 太田 可純, 原田 陽平, 藤本 健二, 鎌田 真由美, 奥野 恭史	第12回 生命医薬情報学連合大会 (IIBMP 2023) @ 柏の葉カンファレンスセンター (千葉県柏市)	Sep. 7-9, 2023