

果てなくつづく医学の険路  
「富岳」が拓く  
創薬の未来

「富岳」成果加速プログラム

領域2 国民の生命・財産を守る取組の強化

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

## もくじ

- 02 「富岳」成果創出加速プログラムについて
- 03 ごあいさつ  
研究目的と課題概要
- 04 「富岳」で初めて可能となる超大規模計算・データ解析
- 05 研究体制と研究内容
- 07 研究成果
- 12 創薬ビッグデータ統合システム
- 14 プレスリリース
- 16 「富岳」による新型コロナウイルスの治療薬候補同定
- 18 COVID-19の治療薬候補を既存薬の中から計算で探し出す
- 20 プロジェクトの統合的推進

## 「富岳」成果創出加速プログラムについて

2021年に本格運用を開始した「富岳」は、これまでに開発したアプリケーションを活用した研究開発を行い、成果を創出するフェーズに入りました。各分野における世界最先端の成果の創出が期待されています。また「富岳」は、Society 5.0や持続可能な開発目標（SDGs）等の実現のための大規模計算基盤であり、国が実施する他の研究開発プロジェクトや産業界、行政組織等との連携体制を構築しながら、最先端の科学的成果の創出や成果の社会実装を強力に推進する必要があります。

そのため、2020年4月からスタートした「富岳」成果創出加速プログラムでは、「富岳」の計算資源を活用して、4つの領域で価値創造につながる成果が期待されています。



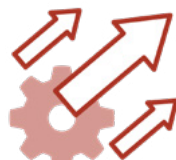
領域1

人類の普遍的課題への  
挑戦と未来開拓



領域2

国民の生命・財産を守る  
取組の強化



領域3

産業競争力の強化



領域4

研究基盤

# ごあいさつ

近年、個人個人により最適な治療を提供するプレジジョンメディスン (Precision Medicine) が最先端医療の一つとして注目されています。わが国でも、患者のゲノムを解析し、がんの治療薬の選択や希少難病の診断につなげるゲノム医療として、研究開発および臨床実装が進められています。しかしながら、患者のゲノム解析を行い、患者固有のゲノム配列 (遺伝子多型・変異) が明らかになっても、臨床現場では、10%未満の解釈可能な遺伝子多型・変異のみで治療選択を行っているのが現状です。すなわち、ゲノム医療での真の実現には、患者のゲノム配列を解析するだけでは不十分であり、ゲノム配列の臨床的意義を解明することがプレジジョンメディスンにおける世界共通の根本課題となっています。

本研究では、ポスト「京」重点課題Iのプロジェクトで「富岳」のために開発してきたGENESISや様々な創薬計算アプリケーションを用いることで、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より明らかにし、これによって得られた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、プレジジョンメディスンの加速を目指します。

プレジジョンメディスンの加速を推進することは、我が国の製薬産業発展の求心力となり、結果として新薬を求める患者への貢献が期待できます。このような日本の産業と科学の持続的発展を支える「富岳」のプロジェクトに、今後も皆様からの益々のご理解とご支援、ご協力をたまわりますようよろしくお願い申し上げます。

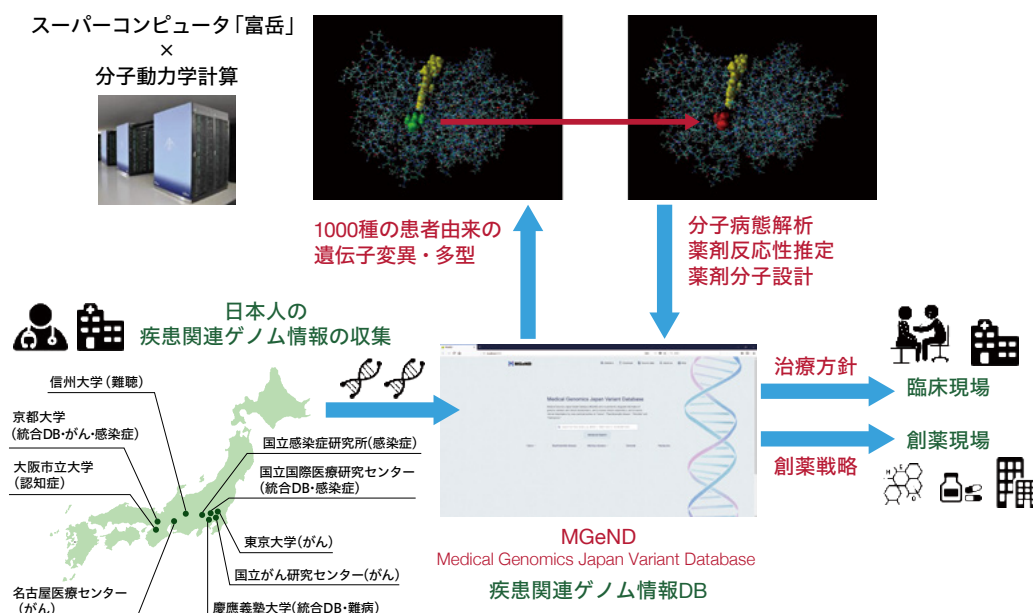


課題代表  
理化学研究所 計算科学研究センター  
京都大学大学院医学研究科

奥野 恭史

## 研究目的と課題概要

がんやウイルス感染症などの多くの疾患の病態に、遺伝子多型・変異が関係しています。しかし、ゲノム解析ではその解明は不十分です。本研究では、「京」の125倍の実行性能の向上に成功したGENESISを中核として「富岳」×「GENESIS」のスケールメリットを最大限に生かし、1,000種の疾患関連遺伝子多型・変異がタンパク質のダイナミクスに与える影響を明らかにします。これによって分子病態解析、薬剤反応性推定、薬剤分子設計に関する知見を臨床現場、創薬現場に提供することを目的としています。



# 「富岳」で初めて可能となる 超大規模計算・データ解析

「京」の125倍の実効性能を発揮できる GENESISと創薬ビッグデータ統合システムにより、1,000種の疾患関連遺伝子多型・変異に対するタンパク質の分子レベルでの病態解明、薬剤反応推定、薬剤分子設計を実現します。

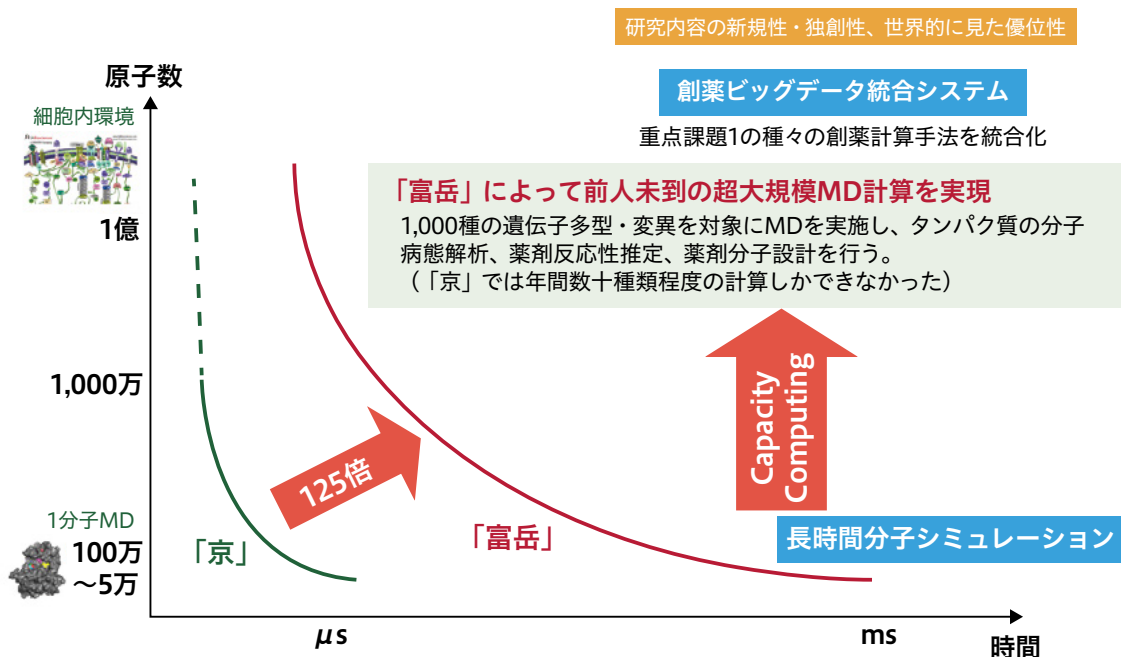
## 「富岳」の必要性

「富岳」では、「京」では困難であったミリ秒オーダーの分子シミュレーションが可能です。タンパク質の構造ダイナミクスや薬剤反応の分子動力学計算を実施するという前人未達の超大規模計算を行えます。「京」で同様の計算をすると、年間数十種類しか計算できなかったことから、「富岳」が必要なのです。

## 期待される成果

「富岳」を用いた分子シミュレーションにより、日本人の疾患関連ゲノム情報（遺伝子多型・変異）の1,000種を対象にしたタンパク質の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得ることができます。そのシミュレーション結果をMGeNDデータベースで登録、公開し、さらに創薬ビッグデータシステムを用いることで、製薬会社による「富岳」を用いた創薬計算の現場利用を実現します。

## 「富岳」で初めて可能となる超大規模計算・データ解析



「富岳」によってミリ秒オーダーの分子シミュレーションが初めて可能になり、タンパク質の長時間ダイナミクスや薬剤結合反応の精密な予測ができる。

# 研究体制と研究内容

「富岳」のために開発を進め、「京」の125倍の実行性能の向上に成功した「GENESIS」を中核として、  
「富岳」、「GENESIS」のスケールメリットを最大限に生かし、  
1,000種の疾患関連遺伝子多型・変異に対するタンパク質構造レベルでの  
①分子病態解析、②薬剤反応性推定、③薬剤分子設計を実施します。  
それらの知見をもとにAI・データサイエンスの技術を「富岳」による分子シミュレーションと組み合わせ、  
④AI・データサイエンス・社会実装基盤を構築することにより世界を圧倒する創薬計算を実現し、  
プレジジョンメディシンの加速を目指します。

## ④AI・データサイエンス・社会実装基盤

### AI-シミュレーション融合(理研AIP・寺山)

AI技術による分子設計の  
高精度化・MD計算に対する  
最適化アルゴリズム適用による効率化



### 疾患ゲノム構造データベース(京大医・鎌田)

疾患関連変異とタンパク質立体構造の集約  
および解析結果の公開



### 創薬ビッグデータ統合システム(理研MIH・本間)

各創薬基盤と「富岳」の連携プラットフォーム構築  
および産業界との連携  
および「富岳」の利用促進



## ①分子病態解析

変異によるタンパク質の構造、安定性、  
活性の変化を推定し、病態との関連を解析する



## ②薬剤反応性推定

変異によるタンパク質と薬剤の反応性の変化を推定し、  
患者に最適な薬剤治療の指針を提案する



## ③薬剤分子設計

希少疾患、薬剤耐性腫瘍などを対象にした  
精密な薬剤分子設計を実現する

### 動的構造機能解析(横浜市立・池口)

MSMを用いた変異の導入による  
構造ダイナミクス変化解析

希少疾患タンパク質安定性評価 200 構造  
キナーゼ等・長時間ダイナミクス 2 構造



### タンパク質活性予測(京大理・林)

QM/MMによるキナーゼのリン酸化反応解析・  
共有結合型阻害剤の結合反応性の計算

キナーゼリン酸化反応 10 構造  
共有結合型阻害剤反応 12 構造



### 結合自由エネルギー計算(京大医・奥野)

MP-CAFEE法やFEP法による各種変異に対する  
非共有結合型阻害剤やATPとの  
結合自由エネルギー推定

キナーゼ-ATP-阻害剤 800 構造



### 結合経路・ポーズ推定(理研BDR・杉田)

gREST/RUESIによるキナーゼと  
阻害剤の結合経路推定・ポーズ推定

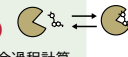
キナーゼ 阻害剤 10 構造  
キナーゼ 基質ペプチド 5 構造



### 結合速度論解析(東工大・北尾)

PACS-MD・MSMIによるキナーゼと阻害剤の結合過程計算

キナーゼ-阻害剤・  
基質ペプチド 6 構造



### 低分子デザイン(筑波大・広川)

創薬ビッグデータ統合システムを  
用いたドラッグデザイン

薬剤耐性キナーゼ 500 構造  
希少疾患関連タンパク質 500 構造



### 抗体医薬デザイン(東大先端研・山下)

PD1抗体など有名な抗体デザイン

抗原抗体 30 構造



# 1. 分子病態解析

変異によるタンパク質の構造、安定性、活性の変化を推定し、  
病態との関連を解析します。

## 1-1 動的構造機能解析 横浜市立大学大学院 生命医科学研究科 池口満徳

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します。



## 1-2 タンパク質活性予測 京都大学大学院 理学研究科 林 重彦

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、および共有結合型阻害剤の反応性に関する計算を行います。



## 2. 薬剤反応性推定

変異によるタンパク質と薬剤の反応性の変化を推定し、患者に最適な薬剤治療の指針を提案します。

### 2-1 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

MP-CAFE 法やFEP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の結合自由エネルギーの変化を推定することで、キナーゼ変異体に対する薬剤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します。



### 2-2 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

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します。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します。



### 2-3 結合速度論解析 東京工業大学 生命理工学院 北尾彰朗

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します。



## 3. 薬剤分子設計

希少疾患、薬剤耐性腫瘍などを対象にした精密な薬剤分子設計を実現します。

### 3-1 低分子デザイン 筑波大学 医学医療系 広川貴次

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います。



### 3-2 抗体医薬デザイン 東京大学 先端科学技術研究センター 山下雄史

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の薬効への影響の予測を可能にしていきます。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します。



## 4. AI・データサイエンス・社会実装基盤

### 4-1 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

京都大学大学院医学研究科で開発した日本人疾患関連ゲノムバリエーションデータベース MGeND(Medical Genomics Japan Variant Database)や代表的な疾患バリエーションデータベースを対象に、明確に疾患との関連が同定されていないバリエーションに対するタンパク質立体構造情報との統合や標的タンパク質の探索、解析結果の公開を行います。



### 4-2 AI・シミュレーション融合 横浜市立大学大学院 生命医科学研究科 寺山 慧

他研究課題と連携しつつ、ポスト「京」重点課題1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とともに、AIを利用した計算手法の高度化・高速化に取り組みます。



### 4-3 創薬ビッグデータ統合システム 理化学研究所 計算科学研究センター HPC/AI駆動型医薬プラットフォーム部門 本間光貴

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します。



## 1. 分子病態解析

### 1-1 動的構造機能解析 横浜市立大学 大学院生命医科学研究科 池口満徳

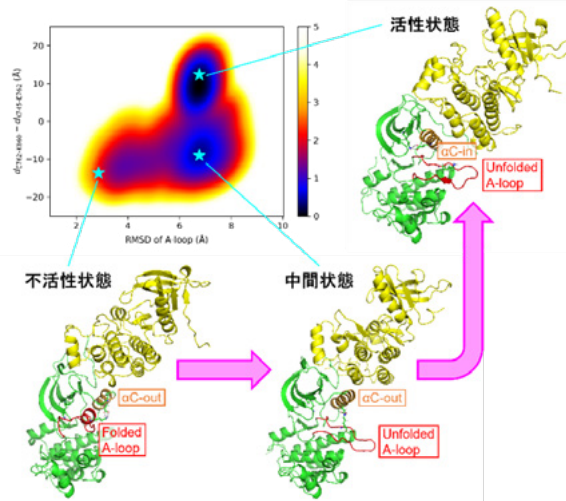
#### 大規模分子動力学計算を用いたEGFRダイマーの活性化ダイナミクス

上皮成長因子受容体 (EGFR) は、がんなどの創薬標的タンパク質であり、細胞外にEGFが結合するとダイマーを形成し活性化する。細胞内のキナーゼドメインでは、非対称なダイマーが形成され、レシーバー側のキナーゼが不活性状態から活性状態へ構造変化して活性化する。このダイマー形成による活性化メカニズムを明らかにするために、不活性状態と活性状態のダイマー構造をモデリングし、その間の構造変化をパスウェイ探索/MSM法による大規模分子動力学 (MD) シミュレーションによって探索した。まず、構造変化の経路について、最小自由エネルギー経路探索をSTRING法により実施した。次に、経路上の32個の代表構造について、それぞれ、420 nsのMDを10回ずつ実施し、計134.4  $\mu$ sのパスウェイ周辺をサンプリングしたトラジェクトリを統合し、マルコフ状態遷移モデル (MSM) を構築した。これらの大規模MD計算は、「富岳」にてMD計算ソフトGENESISを用いて実施した。

MSMから見積もった自由エネルギーマップから、活性状態が最安定であり、不活性状態と活性状態の間に準安定な中間状態が存在することが分かった。また、構造変化の解析により、ダイマー界面の構造変化がレシーバーキナーゼの活性化を引き起こす様子を明らかにできた。

上記の活性、不活性、中間状態のダイマー構造および膜外ドメイン構造を用いて、臨床等で確認された変異体220種類について、変異体をモデリングしてMD計算を実施した。各状態での400 ns以上のトラジェ

EGFR キナーゼダイマー活性化に関わる自由エネルギーマップと代表的な3状態



クトリから、平均二乗変位や原子間距離の変動等を野生型と比較した結果、初期状態や変異残基に依りて、野生型と振舞いに違いが見出される部位があることが分かった。

### 1-2 タンパク質活性予測 京都大学 大学院理学研究科 林 重彦

#### タンパク質活性予測

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法を用いて薬剤分子結合を解析する手法を開発し、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解析を行った。本手法の開発により、薬剤分子結合によるタンパク質の構造変化を長時間MDシミュレーションで考慮しながら、非経験的量子化学計算による共有結合形成などを含む化学反応のエナジエティクスを直接的な計算が可能となった。

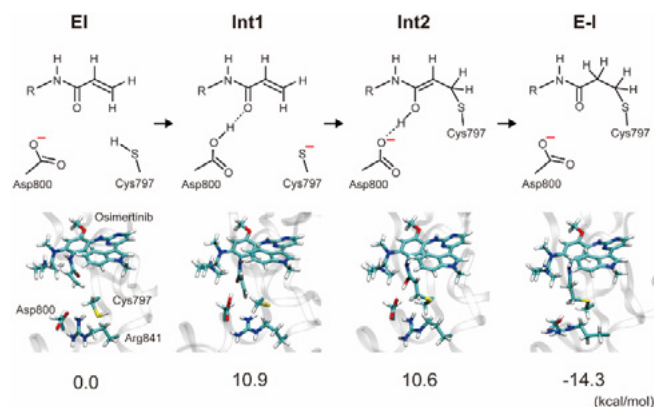
キナーゼのリン酸化反応解析に関しては、MEK1キナーゼのリン酸化によるATPase活性化状態の遷移状態構造を決定する手法を開発し、リン酸化による活性化の分子機構を明らかにした。その結果、activation loopへのリン酸化によるアロステリックな構造変化により、ATP加水分解反応活性部位の構造がコンパクトになり、ATP加水分解反応の活性化自由エネルギーが減少することを明らかにした。

また、EGFRキナーゼに対する共有結合型阻害剤分子であるosimertinibのチオマイケル付加共有結合形成反応について解析を行った(図)。その結果、チオマイケル付加反応の共有結合形成は非常に大きな結合自由エネルギーを与えることを見出し、持続性の高い薬効の分子メカニズムを明らかにした。また、特異な分子構造を有する反応中間状態の決定に成功し、効率的な共有結合形成反応過程を明らかにした。

さらに、HIVプロテアーゼのindinavir阻害剤分子結合に対する薬剤

耐性変異の影響を解析し、薬剤耐性を与える相互作用と構造変化を明らかにした。また、SARS-CoV-2メインプロテアーゼの複数の共有結合型阻害剤分子の結合を解析し、その共有結合メカニズムを明らかにした。

EGFR キナーゼに対する共有結合型阻害剤分子 osimertinib の結合反応経路



## 2. 薬剤反応性推定

### 2-1 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

#### 「富岳」を用いた大規模分子動力学計算によって、 薬剤応答性に対する遺伝子変異の影響を高速かつ高精度に推定

近年、患者個人のゲノムに基づいて最適な治療を提供する「個別化医療」が注目されているものの、患者のゲノム解析から同定される遺伝子変異の90%以上は臨床的・機能的意義が不明である。そこで我々は、がんの主要な原因として知られているEGFR, RET, ALKといったキナーゼ遺伝子を対象に、がんゲノムで観測された体細胞変異を計800種類抽出し、薬剤応答性に対する各変異の影響を分子動力学 (MD) シミュレーションによって推定した。具体的には、EGFR 阻害剤2種 (ゲフィチニブ、エルロチニブ)、RET 阻害剤2種 (セルベルカチニブ、プラルセチニブ)、及びALK 阻害剤3種 (クリゾチニブ、アレクチニブ、ロルラチニブ) を対象薬剤として、MP-CAFEE法やMutationFEP法といった自由エネルギー摂動法 (FEP法) を用いて、アミノ酸変異による薬剤の結合

自由エネルギー (結合親和性) の変化を算出した。更に、非小細胞肺癌患者から検出されたEGFR, RET, 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、キナーゼ活性化・薬剤耐性化の分子メカニズムを推定した (図1A)。更に、超高周波超音波の照射によって生体分子の動的プロセスを加速する新しいシミュレーション技術 (Hypersound-Perturbed MD Simulation) を開発し、従来型MDと比較してタンパク質と医薬品の結合速度を10-20倍に上昇させることに成功した (図1B)。将来、本手法をスーパーコンピュータ「富岳」に適用することで、「京」の300倍の速度向上が見込まれ、治療薬探索など、医薬品開発のスピードアップに貢献することが期待される。

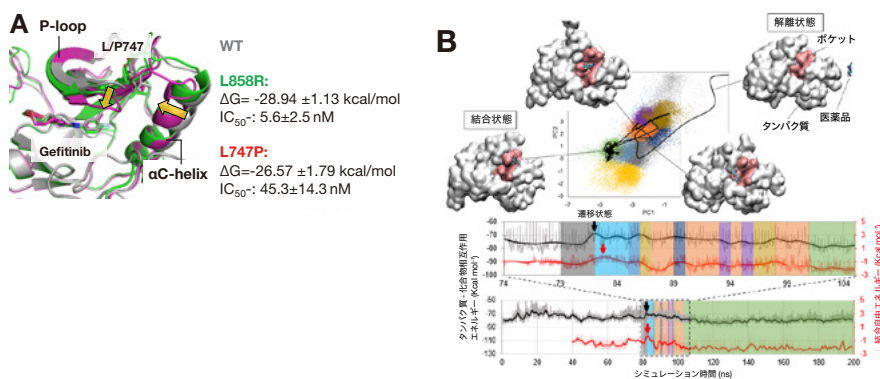


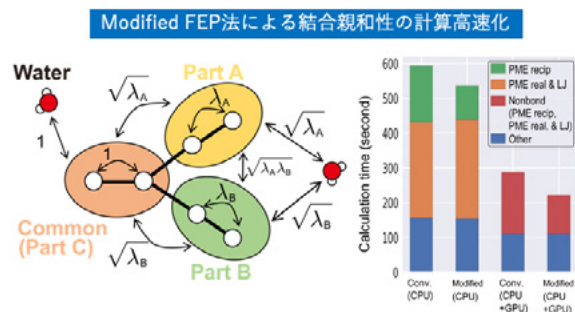
図1 (A) EGFR-Gefitinib 複合体のシミュレーション構造。EGFR 野生型 (灰色)、L858R 変異体 (緑)、L747P 変異体 (マゼンタ) の各々について1 $\mu$ sのMDシミュレーションを独立に3本実施し、得られた平均構造を示した。タンパク質、薬剤は、それぞれリボンモデル、スティックモデルで表示し、ALK 野生型・L858R 変異体と比較した際のL747R 変異体におけるP-loop及び $\alpha$ -C-helixの配向の変化を矢印で示した。また、MP-CAFEE法によって算出した結合自由エネルギー ( $\Delta G$ ) 及び実験的に測定した薬剤の50%阻害濃度 ( $IC_{50}$ ) をシミュレーション構造の右側に記載した。(B) Hypersound-Perturbed MDシミュレーションによって捉えたCDK2タンパク質のATPポケットに対する医薬品の結合パスウェイ (上) 超音波照射下のシミュレーションで捉えた全ての化合物結合ポーズを10色の点で表し、代表的な一つの結合経路 (黒線) 及び経路上の代表的な結合ポーズを示した。ここで緑色の結合ポーズが実験的に観測されている結合モードに最も近い。(下) 結合経路に沿ったCDK2-医薬品相互作用エネルギー (黒色) 及び結合自由エネルギー (赤色) の軌跡。結合経路上で最もエネルギーが高い遷移状態を矢印で示した。

### 2-2 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

キナーゼ・薬剤結合の分子機構をより詳しく理解するためには動的構造を考慮した分子シミュレーションが必要である。そのために薬剤とキナーゼ複合体の構造探索手法 (gREST) の開発と結合親和性計算 (FEP) の高速化を行い、分子動力学ソフトウェアGENESISに新しい機能として導入した。gRESTとFEPを組み合わせることで、複合体の結晶構造が得られない場合にも、信頼性の高いポーズ推定を行うことができる。さらにその構造を用いることで、結合親和性の予測精度も向上させることができた。また、「富岳」に高度化された新しいバージョンのGENESISを2022年7月にフリーウェアとして公開し、アカデミアだけでなく産業界の利用も可能とした。

さらに我々は、Src, RET, Ablなどのキナーゼをターゲットとして、複数の阻害剤との結合経路と結合親和性の予測を行った。異なる阻害剤を用いた場合も最安定構造にむかう結合経路において類似な分子メカニズムが働いていることが明らかになってきた一方で、阻害剤の分子サイズと構造柔軟性の違いによる独自性も明らかになってきた。さらに、従来のインシリコ創薬では全く考慮されていなかった細胞内分子混雑環境の影響をシミュレーションと実験の両面で解析した。この解析では、cSrc kinase、

#### Modified FEP法による結合親和性の計算高速化



阻害剤PP1、複数のBSAを混雑物質として加えた異なる蛋白質濃度の混合系を複数用意して、それぞれ10マイクロ秒以上の長時間の分子動力学計算を実行した。その結果、BSA存在下と希薄溶液中では結合親和性が異なるだけでなく、異なる結合経路が使われている可能性が予測された。この変化は混雑物とcSrc kinaseの間に働く弱い分子間相互作用によって生じた蛋白質構造揺らぎの変化に起因していた。



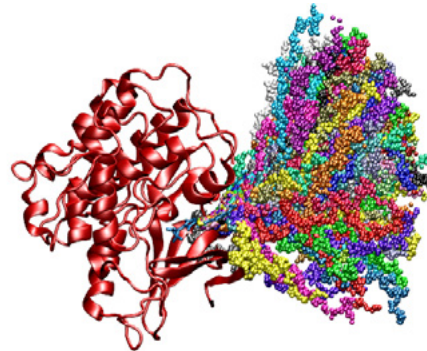
## 2. 薬剤反応性推定

## 2-3 結合速度論解析 東京工業大学 生命理工学院 北尾彰朗

## キナーゼと阻害剤の結合・解離メカニズム

本研究では、これまで開発してきたPaCS-MD/MSM法を用いて、キナーゼと阻害剤などの複合体の解離シミュレーションを行い、結合や解離の過程を詳しく調べ、結合のメカニズムを検討した。タンパク質-タンパク質複合体としてCDK2-CKSHS1、キナーゼ-阻害剤複合体としてCDK2と親和性の異なる2種類のアミノピラゾールと2種類のオキシインドール、計4種類のリガンドを対象として解離シミュレーションを行った。PaCS-MDを用いると、いずれの複合体に対しても、1回の解離に必要なシミュレーション時間は約150 ns程度であった。解離速度定数から予測される結合状態の滞留時間が0.9 - 40秒であることを考えると、PaCS-MDは様々なタンパク質複合体の解離シミュレーションに有効であることが分かった。CDK2の界面残基重心に対するCKSHS1の界面残基もしくはリガンド重心の相対位置を反応座標としてMSM解析(3D)を行った結果、計算した全ての複合体の結合自由エネルギーを正確に再現した。速度定数に関しては実験値から大きく外れた結果が得られたが、タンパク質-タンパク質複合体の場合は界面を形成する残基ペアの距離、タンパク質-リガンド複合体の場合はリガンドと近接残基の原子間距離を反応座標として導入してMSM解析を行うことにより、解離速度定数の改善に成功した。結合速度定数に関しては、大きな改善が得られ

CDK2からのアミノピラゾールの解離経路



なかったが、結合シミュレーションの導入により改善されることが期待される。本研究で用いた、PaCS-MDと残基間や原子間距離を反応座標としたMSM解析を行うアプローチは、キナーゼ-ペプチド複合体など、他の複合体にも容易に適用可能であり、今後様々な複合体に適用することにより、多くの成果を上げることが期待される。

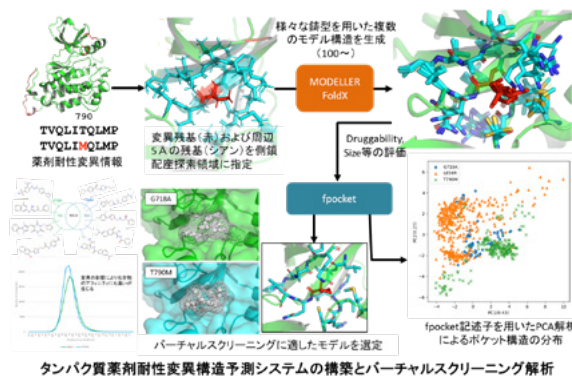
## 3. 薬剤分子設計

## 3-1 低分子デザイン 筑波大学 医学医療系 広川貴次

## 薬剤耐性化合物デザインに向けた変異構造予測システムの構築とバーチャルスクリーニング解析

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目指したタンパク質薬剤耐性変異構造予測システムの構築とバーチャルスクリーニング解析を行った。モデリング構築は、各変異に対し200構造を構築し、それぞれの構造に対し、リガンド結合部位をfpocketプログラムによりポケット内の物理化学的特性を18の記述子により表現した。ベンチマークとして、これらの記述子を変数とした主成分分析により3変異間でのポケット構造の性質を俯瞰した結果、変異体間で共通のポケット空間と同時に、各変異モデルに特異的な分布を有していることが明らかになった。今後、ポケット構造の分布が類似する変異体間では同じ薬剤の適用可能性を予測できるのではと考察している。

バーチャルスクリーニングのベンチマーク(G718AとT790M)では、ドッキングスコアの分布から、T790Mに対するヒット化合物がドッキングスコアが低く、変異体ごとにドラッグビリティが異なることが示唆された。また90,123化合物が両変異体に対して共有にヒットしており、これらは、変異にロバスタな回避性の性質を有する化合物である可能性が高いことを示した。同時に、それぞれの変異体でのみヒットした化合物は、それぞれの変異体特異的なフォーカスライブラリーとして活用できることも考えられる。



現在、ClinVer 変異データベースに対するEGFRを対象とした網羅的モデリングを行っており、変異154種について、61,600モデルを構築した。今後は、各変異体モデルに対し、化合物ライブラリースクリーニングを実施し、変異体間でのヒット化合物の比較、ドッキングスコア分布の比較等を行い、変異回避性の高いリガンドの特性を検討する予定である。

## 3. 薬剤分子設計

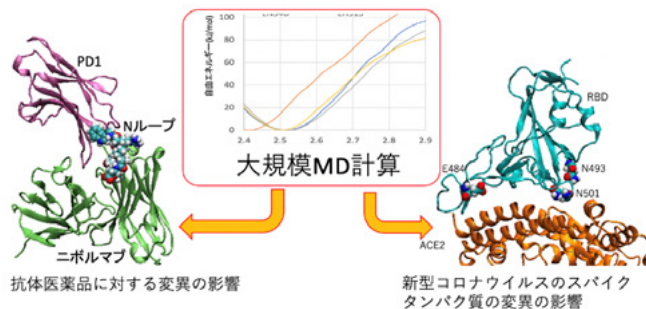
## 3-2 抗体医薬 東京大学 先端科学技術研究センター 山下雄史

## 抗体医薬品の解析・設計・開発支援

免疫系で重要な役割を持つ抗体は体内に侵入した異物（抗原）を見つけて強く結合することができるタンパク質である。この性質を活かすことで、抗体を病気の治療をする薬として使うことができる。抗体を基盤とする医薬品（抗体医薬品）は、がんなどの治療に使われているものもあるが、今後も多くの抗体医薬品が開発され、多くの難病治療に道を開くと考えられている。我々は、抗体医薬品をより効率的に開発するために、計算機シミュレーションを活用することを考えている。従来、医薬品開発には使えないと思われていた計算技術が、「富岳」を始めとするスーパーコンピュータの登場により応用可能な時代になった。高精度なシミュレーション技術により、効率的・効果的に抗体医薬品開発ができるようになっていきたいと思います。

我々が取り組んでいる課題は、抗体をどのように改変するとより望ましい性質になるのかというものである。抗体がしっかりと抗原に結合するようになると良い薬の候補になる。具体的には、がん免疫療法で用いられる抗体医薬品ニボルマブをどのように改変するとより結合力の高い抗体になるのかをMDシミュレーションを用いて設計した。また、抗原タンパク質PD-1にどのような変異が入ると相互作用が低下し、それに対してどのような変異をニボルマブに加えると親和性が回復するのかも計算機上で

## 大規模MD計算によるタンパク質間相互作用の解析



デモンストレーションした。また、機械学習や簡略モデルを用いて、抗原抗体相互作用を解析し改変設計を支援する手法も開発した（例えば、JPCB, 2021）。

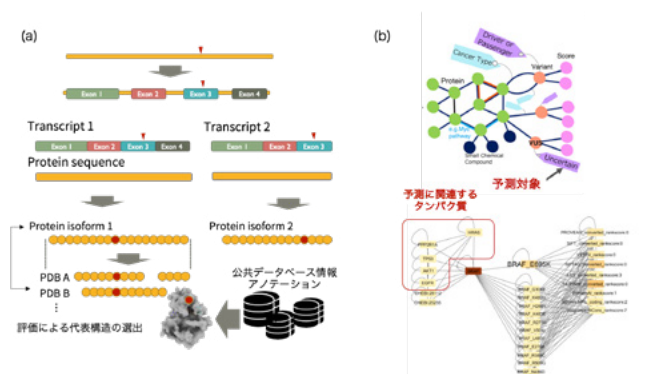
抗体研究を通して培った高精度シミュレーション技術を、実際の創薬支援（例えば、Protein Expr. Purif., 2022）や新型コロナウイルスのスパイクタンパク質の解析にも応用している。

## 4. AI・データサイエンス・社会実装基盤

## 4-1 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

## 疾患バリエーションとタンパク質構造の架け橋

ゲノム医療において、個々人のゲノム情報に基づく治療方針の決定は、疾患関連性をはじめとした既報情報との紐付けにより行われる。京都大学では、日本人の疾患関連バリエーション情報を集積するためのデータベース MGeNDを開発してきた。しかし、蓄積されているバリエーションの多くは、分子機能への影響や明確な疾患との関連が同定されていない。バリエーションによる機序を明らかにするためには、ゲノムの世界だけではなく、タンパク質立体構造レベルでの解釈が重要となる。そのため、我々は代表的なヒトタンパク質のアミノ酸配列とそれに基づく代表立体構造データを独自に定義し、MGeNDや世界的に広く用いられている ClinVar・COSMIC に集積されているバリエーションを対象に、タンパク質立体構造へのマッピングと、公共データベースから得られる多様な情報の紐付けを可能とするアノテーションパイプラインを構築した。さらに、疾患関連遺伝子を大規模なタンパク質ネットワークにマッピングし、グラフ深層学習の一種である Graph convolutional network (GCN) を用いて、疾患関連性を予測するモデル開発を行った。本モデルでは、周辺タンパク質との関連性を考慮してバリエーションの影響を予測し、可視化手法を用いて予測根拠となる関連タンパク質の同定が可能である。ベンチマークセットを用いた検証では、高い予測精度を示すことができた。さらなる改良により、機序不明なバリエーションにおけるメカニズムの解明に貢献できるものと考えている。そして、アノテーションされたバリエーション情報や GCN による予測結果、富岳



(a) バリエーションのゲノムレベルからタンパク質立体構造へのマッピング  
(b) GCNを用いた疾患関連性予測と予測根拠の可視化。タンパク質ネットワークを用いた知識グラフを構築し、GCNを用いることで、分子間の関係性を考慮した学習と予測が可能。

プロジェクトの他チームにより実施されたシミュレーション結果を公開するためのデータベースの開発を行ってきた。本データベース公開により、世界中の臨床家や研究者と情報を共有することができ、Precision Medicineの発展に資するものと期待される。

## 4. AI・データサイエンス・社会実装基盤

## 4-2 AI - シミュレーション融合 横浜市立大学 大学院生命医科学研究科 寺山 慧

AI-シミュレーション融合に向けて  
——AI・機械学習による分子設計や計算高速化が可能に

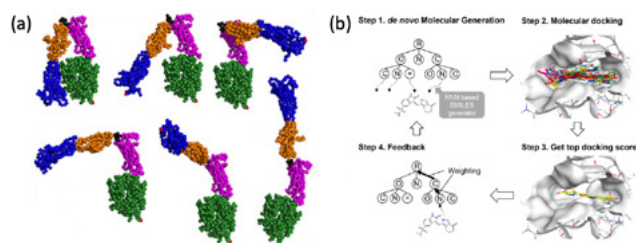
分子動力学計算 (MD 計算) をはじめとする様々なシミュレーション技法は分子の運動や機能を解析・予測する強力なツールである。しかし、一般にこれらのシミュレーションは多くの計算時間がかかり、「富岳」のようなスパコンでも計算の効率化は避けて通れない課題である。そこで AI や機械学習、様々な最適化手法を利用することで計算を高速にする手法が開発されている。また AI による薬などの分子の自動設計手法も近年盛んに研究されている。ここでは、バイズ最適化を用いた新規疎視化シミュレーション手法 Adaptive ENM とドッキングシミュレーションを利用した分子設計手法 SBMolGen を紹介する。

Adaptive ENM はアミノ酸1つを単位(玉)として扱う疎視化シミュレーションを高度化した手法である。玉同士の間働く力を適切に設定することで分子の動きを再現することが可能だが、設定パラメータの候補が多く、設定は簡単ではない。そこでバイズ最適化という手法を用いることで、出来るだけ少ない試行回数で適切なパラメータを探索する枠組みを導入し、実際に大きな構造変化のサンプリングに成功した(図(a)、文献1)。

SBMolGen は、強化学習に基づく分子生成手法 ChemTS とドッキングシミュレーションを連携させることで、ターゲットタンパク質のポケットに強く結合する分子を自動的に設計する手法である(図(b))。ChemTS

が生成した候補分子を、ドッキングシミュレーションでタンパク質との相互作用を見積もり、そのスコアを ChemTS に反映させることでより相互作用の強い分子が徐々に生成される。数万個の分子を生成しながら探索することでポケットと強く相互作用する分子を発見している(文献2)。

このように AI・機械学習とシミュレーションが融合した手法は、今後計算科学やインシリコ創薬では不可欠な技法になると期待される。



(a) Adaptive ENM によりサンプリングされた Integrin  $\alpha V$  の大きな構造変化  
(b) SBMolGen による分子設計の概略図。ドッキングシミュレーションと連携することでターゲットタンパク質のポケットに強く結合する分子を設計できる。

文献1 R. Kanada, K. Terayama, A. Tokuhisa, S. Matsumoto, Y. Okuno, "Enhanced Conformational Sampling with an Adaptive Coarse-Grained Elastic Network Model Using Short-Time All-Atom Molecular Dynamics," *Journal of Chemical Theory and Computation*, Vol.18, No.4, pp.2062–2074, 2022.

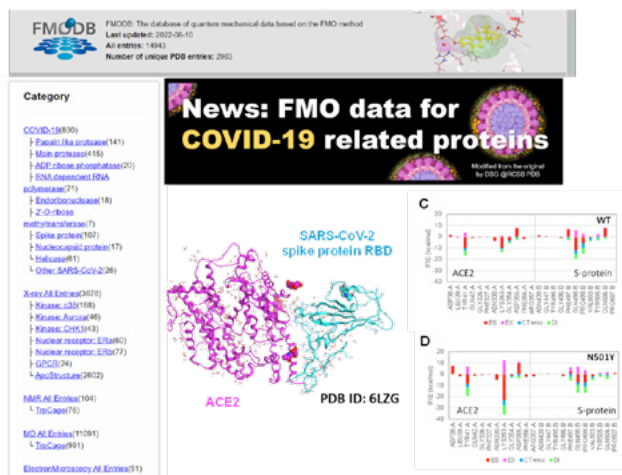
文献2 B. Ma, K. Terayama, S. Matsumoto, Y. Isaka, Y. Sasakura, H. Iwata, M. Araki, and Y. Okuno, "Structure-Based de Novo Molecular Generator Combined with Artificial Intelligence and Docking Simulations," *Journal of Chemical Information and Modeling*, Vol. 61, No.7, pp.3304–3313, 2021.

## 4-3 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推進プログラム 本間光貴

## AI構築のための量子化学計算値データベースと自動計算プロトコルの整備

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、AI のハイパーパラメータ最適化を実施できるプラットフォームを構築し、ライフインテリジェンスコンソーシアムを通して製薬企業に公開した。シミュレーションについては、AI の学習用データベースの整備を行い、FMO 創薬コンソーシアムと連携して、FMO 自動計算プロトコルの開発を行った。開発したプロトコルを用いて、COVID-19 関連タンパク質に対する FMO 計算を網羅的に実施し、10 種類以上のタンパク質で合計 830 件の FMO データを公開した。これらのデータから、COVID-19 変異株の S protein の ACE2 に対する親和性を定量的に解析し、感染力を検証した。また、S protein と抗体の親和性予測を行い、高い相関 ( $r^2=0.57$ ) を示した。FMO-AI 力場の開発を目指して、Q10, TrpCage, Bromodomain, CDK2 の分子動力学計算によって得られたそれぞれ 1 万個程度の構造に対して FMO 計算を行い、FMO DB に登録した。これらは、FMO 電荷 AI の開発に貢献した。構造生物学的手法によって決定したタンパク質構造を量子化学レベルで最適化し、より高精度のタンパク質構造データを作成するために、FMO 法による高速な構造最適化手法 (FDPD 法) を開発した。Pim1 の活性予測では、MM 最適化構造、QM-MM 最適化構造に比較して、FDPD 法によってタンパク質全体の QM 最適化を行った構造は、MM ( $r^2=0.58$ )、QM-MM ( $r^2=0.67$ ) に比べて、より高い予測精度 ( $r^2=0.81$ ) を示した。

## 世界初のタンパク質の量子化学計算値データベース (FMO DB)



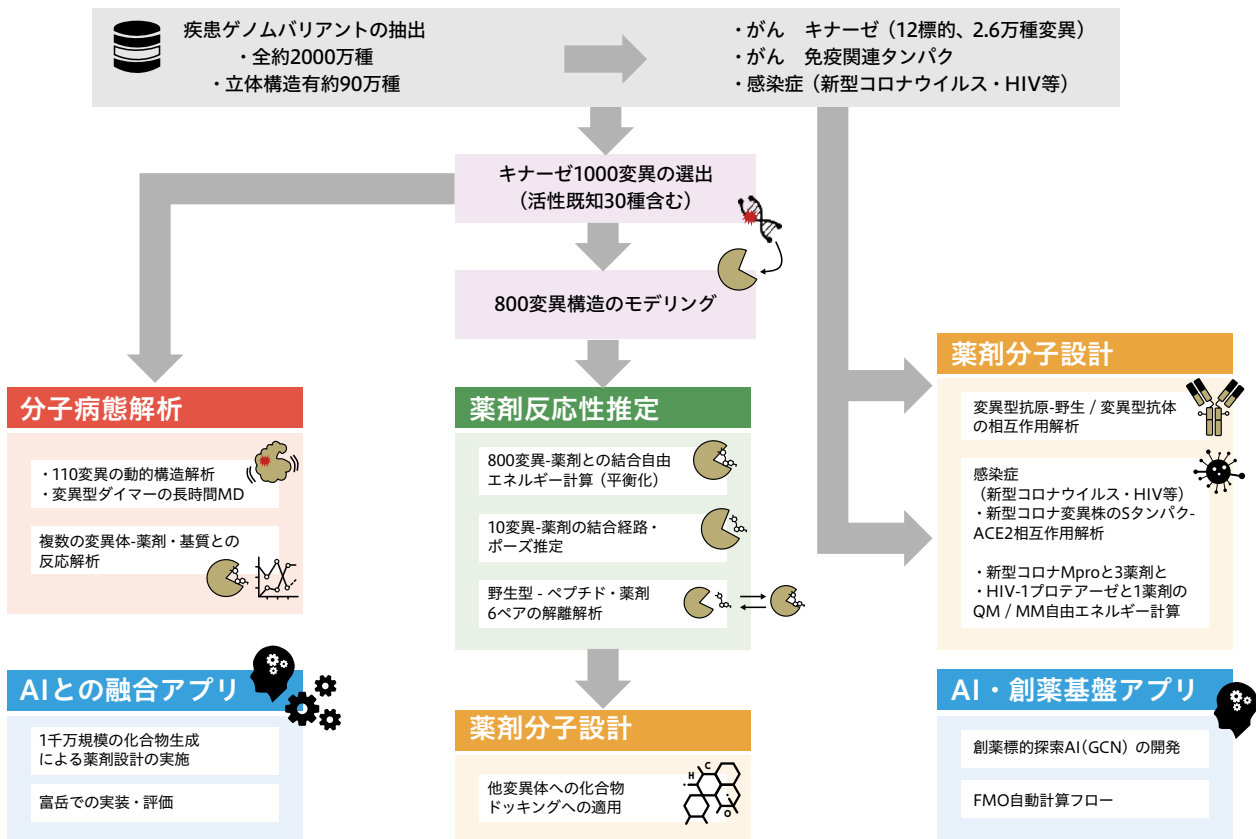
# 創薬ビッグデータ 統合システム

すでに「京」で実効的な実績があるアプリケーションが「富岳」において演算能力を最大限に活かせるようになり、「富岳」に対してチューンナップされたMD計算ソフトおよび創薬計算手法を用いて、創薬計算フロー（スクリーニングからリード最適化）にそって連結した統合システムを開発しています。



ポータルサイトによる全アプリ公開  
全アプリ公開用ポータルサイト開発

世界初の疾患ゲノム構造DBとして公開  
DBの完成と1000変異の計算結果の登録



## アプリケーション一覧

### 分子病態解析

#### ▶ QM/MM RWFE-SCF (Quantum Mechanical/Molecular Mechanical ReWeighting Free Energy-Self Consistent Field)

タンパク質活性部位に対する非経験的量子化学計算による構造最適化とタンパク質環境の長時間分子動力学シミュレーションを組み合わせたハイブリッド法。タンパク質の大きな構造変化を考慮しながら薬剤分子結合部位の構造や化学反応エネルギー、変異導入による結合自由エネルギー変化などを高精度に決定することができる。

## 分子病態解析

## ▶ MP-CAFEE (Massively Parallel Computation of Absolute binding Free Energy with well-Equilibrated states)

標的タンパク質と低分子量有機化合物の結合親和性(結合自由エネルギー)を大規模分子動力学シミュレーションによって精密に計算する。医薬品候補が複数ある場合に、結合親和性計算を通して薬効の高い化合物を予測することが出来る。低分子量有機化合物が介在したタンパク質活性化・制御の分子メカニズムの解明に役立つ。

## ▶ MutationFEP

タンパク質上のアミノ酸変異に起因する薬剤の結合自由エネルギーの変化量を計算する。分子標的薬の薬剤感受性に対するタンパク質変異の影響(薬剤耐性度)を推定することが出来る。

## ▶ HyperSound-perturbed Molecular Dynamics simulation (HS-MD)

超高周波超音波を照射した条件下において、生体分子の分子動力学シミュ

レーションを行う。タンパク質の立体構造変化やタンパク質と医薬品の結合といった生体分子の動的プロセスを効率的に捉えることが出来る。

## ▶ GENESIS (Generalized-Ensemble Simulation System)

マルチスケール分子動力学シミュレーションを実行する統合ソフトウェア。分子モデルとして、粗視化モデル、全原子モデル、QM/MMモデルの利用が可能。シミュレーション手法としては、大規模並列MD計算、各種の拡張アンサンブル計算(REMD, REUS, MREM, GaMD, GaREUS, gREST, gREST/REUS)、自由エネルギー摂動計算法(FEP)、データ駆動シミュレーション、ターゲットMD計算等が導入されている。また、各種のトラジェクトリ解析プログラムも整備されている。

## ▶ PaCS-MD/MSM (Parallel Cascade Selection Molecular Dynamics Simulation in Combination with the Markov State Model)

標的タンパク質複合体(低分子、蛋白質、ペプチド、核酸)の結合親和性(結合自由エネルギー)、速度定数を分子動力学シミュレーションによって精密に計算する。複合体形成/解離機構の解明に役立つ。

## 薬剤分子設計

## ▶ FUJI (glyco) 力場

糖鎖修飾されたタンパク質の高精度モデル。糖鎖修飾されたタンパク質のシミュレーションを可能にする。

## ▶ UMB/mFBMD

抗原と抗体の解離座標に沿った自由エネルギー曲線の高精度計算をする。

## ▶ 薬剤耐性変異構造データベース

EGFR薬剤耐性変異体のモデリング構造データベースとフォーカスライブラリー

## AIとの融合アプリ

## ▶ ChemTSv2

強化学習と深層学習を組み合わせることで望みの特性・物性を持つ分子を探索する手法ChemTSの後継版。ChemTSに比べて高速化・利便性が向上している。また、様々な目的(創薬・材料開発等)に合わせて容易に評価関数が設定・実装可能になっている。

## ▶ ChemLAQA (Conformation search using Look Ahead based on Quadratic Approximation (LAQA))

グレイボックス最適化(Gray-box optimization)と分子シミュレーションに基づく単分子の効率的な構造最適化ライブラリー。SDFやSMILES等の分子情報のみから、密度汎関数法(DFT等)や半経験的分子軌道法(xTB等)などのシミュレーションと組み合わせることで、効率的に安定構造の探索が可能。

## ▶ PHYSBO (optimization tool for PHYSics based on Bayesian Optimization)

効率的なブラックボックス最適化を行うベイズ最適化の高速・スケーラブルなPythonライブラリー。シミュレーションや実験の効率的なパラメータ探索などに幅広く利用可能。COMBO(COMMON Bayesian Optimization)の後継である。

## AI・創薬基盤アプリ

## ▶ PathoGN

ゲノム変化(バリエーション)の病原性(Pathogenicity)を予測する機械学習手法。分子間の関係性を点と辺からなるグラフとして表現し、グラフに対する深層学習手法を用いることで、分子間相互作用を考慮した予測を可能とする。

## ▶ FMO自動計算フロー

FMO法による量子化学計算を自動化し、前処理、設定、計算実施、解析までを行うフロー。COVID-19関連タンパク質の網羅的な計算、AI力場用のデータの収集、MDによる動的挙動解析との連携などに利用できる。また、将来的には、Web上でのオンデマンドサービスに応用可能。

## タンパク質の柔らかさを予測するAI

2021.2.5

巨大かつ複雑な生体高分子の機能メカニズム解明に期待

生体高分子の立体構造は分子認識や酵素活性など個々のタンパク質が担う機能の基盤であり、運動性はタンパク質の機能を厳密に制御していることが知られています。こうした生体高分子の運動性を知ることは、機能メカニズムの正確な理解につながり、ひいては病態解明や医薬品開発を行う上で必要不可欠です。

近年発展が著しいクライオ電子顕微鏡 (cryo-EM) は、単粒子解析を通じて未知であった生体高分子の立体構造を原子～近原子分解能で次々と明らかにしている一方、cryo-EMで解析対象となる生体高分子は一般に巨大かつ複雑であるため、それらの運動性情報 (柔らかさ) を得ることは技術的に困難でした。

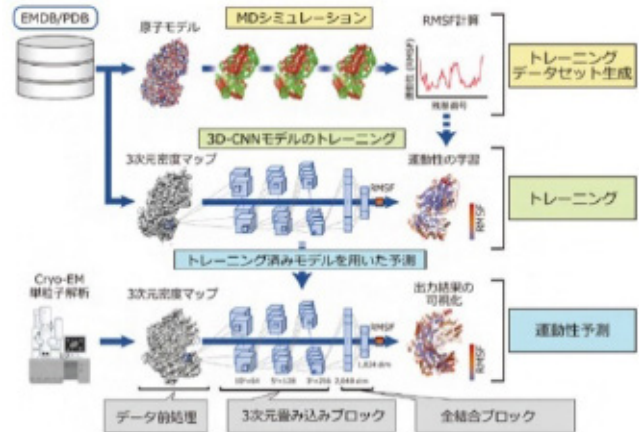
今回、共同研究グループは、このようなタンパク質の運動性を、cryo-EMの計測データのみから原子レベルで簡便・迅速に推定できる画期的なAI「Dynamics Extraction From cryo-EM Map; DEFMap」を開発し、深層学習 (AI)・分子動力学計算 (シミュレーション)・cryo-EMデータ (実験) を組み合わせることで生体高分子の運動性情報を得る手法を世界で初めて構築しました。

これによって、これまで解析が困難であった生体高分子の運動性が明らかになり、機能メカニズムに対する新たな知見の発見につながると考えられます。

また、本手法は超巨大なウイルス粒子などにも適用が可能です。本研究

で構築した学習済みモデルはリポジトリ上で公開されており、あらゆる研究者が3次元密度マップから簡便に運動性を抽出・確認できます。今後、スーパーコンピュータ「富岳」などの利用を通じて大規模な動的情報を蓄積することで、本手法のさらなる精度の向上が見込まれ、今後の生命科学、医学、薬学分野の研究開発の加速に大きく貢献すると期待できます。

### DEFMapの全体的なワークフロー



**論文発表** Shigeyuki Matsumoto, Shoichi Ishida, Mitsugu Araki, Takayuki Kato, Kei Terayama, Yasushi Okuno, "Extraction of protein dynamics information from cryo-EM maps using deep learning", *Nature Machine Intelligence*, 10.1038/s42256-020-00290-y

**プレスリリース** 理化学研究所、京都大学、横浜市立大学、大阪大学  
[https://www.riken.jp/press/2021/20210205\\_1/index.html](https://www.riken.jp/press/2021/20210205_1/index.html)

## 混雑した細胞内で薬はどう効くのか

2021.7.9

生体内環境を考慮した酵素と阻害剤の結合シミュレーション

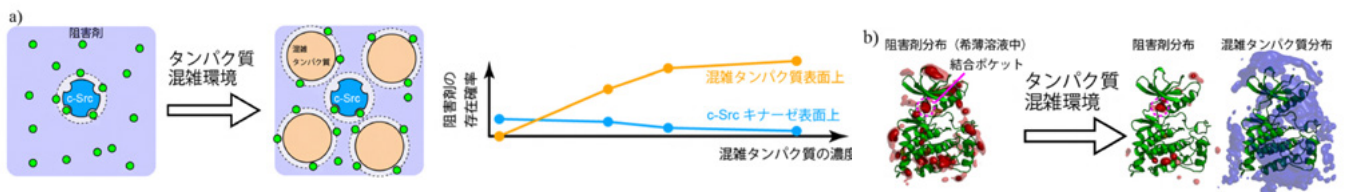
創薬の基本戦略の一つは、疾患に関わるタンパク質 (酵素) の活性を阻害する薬剤基質 (阻害剤) の探索です。酵素と結合して作用する阻害剤が、さまざまなタンパク質や核酸、代謝物などが高濃度に存在する細胞内環境で、どのように酵素と結合するのかはよく分かっていませんでした。

今回、理化学研究所生命機能科学研究センター、ミシガン州立大学生化学および分子生物学部、前橋工科大学工学部生命情報学科の国際共同研究グループは、スーパーコンピュータ「京」と米国のMD専用計算機「Anton2」を用いて、がんに関わる酵素 c-Src キナーゼとその阻害剤であるATP競合性阻害剤、およびそれらを囲むタンパク質 (混雑タンパク質) のウシ血清アルブミン (Bovine Serum Albumin, BSA) との相互作用をMDシミュレーションで解析しました。その結果、混雑タンパク質が阻害剤を吸着することで、酵素周囲の阻害剤の実効濃度が減少し、酵素活性阻害の効率が低下することが分かりました。さらに、

阻害剤が酵素に接近する初期過程も変化し、希薄水溶液中とは異なる経路をたどって阻害剤が結合ポケットに挿入されることを発見しました。

計算機を用いた効率的な薬剤開発の試みは近年非常に注目されており、多くのシミュレーション研究が行われるようになりました。しかし、そのほとんどは希薄水溶液中と同じ条件で計算されており、より細胞内環境に近いと考えられるタンパク質混雑環境の影響を考慮したシミュレーション研究は本研究が世界的にも初めての試みといえます。新規薬剤の開発にかかる時間がさらに長期化している現在、本成果は、タンパク質混雑環境を考慮した新たな分子設計の可能性を示すものとして期待できます。さらに、本研究で用いたソフトウェアGENESISは、スーパーコンピュータ「富岳」において高並列化効率でMD計算が実行できるようにチューニングされており、フリーソフトウェアとして公開されています。「富岳」などの優れた計算資源を有効に用いることで、細胞内環境をより深く理解し、生体内環境と近い条件での分子設計を実現できる日も近づいていると期待されています。

### 溶液の混雑化がもたらす c-Src キナーゼと阻害剤の結合の変化



**論文発表** Kento Kasahara, Suyong Re, Grzegorz Nawrocki, Hiraku Oshima, Chiemi Mishima-Tsumagari, Yukako Miyata-Yabuki, Mutsuko Kukimoto-Niino, Isseki Yu, Mikako Shirouzu, Michael Feig, Yuji Sugita, "Reduced Efficacy of a Src Kinase Inhibitor in Crowded Protein Solution", *Nature Communications*, 10.1038/s41467-021-24349-5

**プレスリリース** 理化学研究所  
[https://www.riken.jp/press/2021/20210709\\_2/index.html](https://www.riken.jp/press/2021/20210709_2/index.html)

## 生体膜カルシウムイオン輸送の分子機構を解明

2021.9.28

ATPによるタンパク質の輸送機構に新しい知見

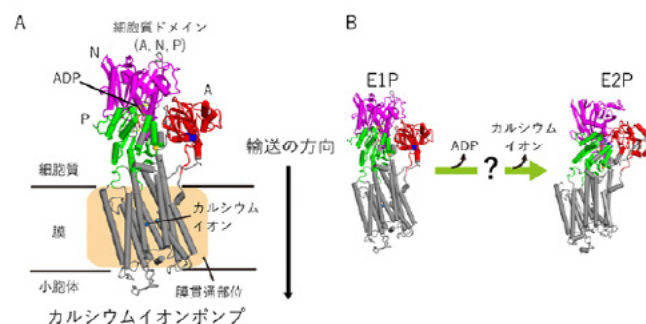
生体内ではさまざまなイオンポンプが働いており、その機能不全が重篤な疾患の原因になります。イオン輸送の分子機構を詳細に理解することは創薬・医学の発展に貢献するものと期待されています。カルシウムイオン輸送が複数のステップで起こること、カルシウムイオンポンプ (SERCA) の構造変化が重要な役割を果たすことは、X線結晶構造解析などの実験研究から提唱されてきましたが、どのようにして構造変化が起きるのかを測定することは、実験的に困難だったため、詳細な分子機構は明らかにされていませんでした。

本共同研究チームは、スーパーコンピュータ「Oakforest-PACS」と「京」を用いて、カルシウムイオンポンプ内に存在するカルシウムイオンが小胞体へ放出される過程について、SERCAに結合したカルシウムイオンが小胞体内へ排出される E1P 状態から E2P 状態へのステップに着目し、共同研究チームが開発している分子動力学計算ソフトウェア GENESIS を用いてそれぞれの状態における分子動力学計算を行い、カルシウムイオン輸送の分子機構を明らかにしました。さらに、ストリング法を GENESIS に導入したことで、単体の分子動力学計算では不可能な比較的大きな膜タンパク質の E1P 状態と E2P 状態間の経路の解析が可能となり、この過程におけるタンパク質の構造変化と周りの影響を計算するこ

とで、膜分子やタンパク質内の過渡的な相互作用がこの過程の中間状態を安定化していることを発見しました。

現在は「富岳」の高い演算性能を利用してさまざまなイオンポンプのイオン輸送機構の解明を目指した研究を行っており、関連する疾患の解明や創薬に貢献していきます。

### カルシウムイオンポンプ (SERCA) の構造と E1P 状態から E2P 状態への変化



**論文発表** Chigusa Kobayashi, Yasuhiro Matsunaga, Jaewoon Jung, Yuji Sugita, "Structural and energetic analysis of metastable intermediate states in the E1P-E2P transition of Ca<sup>2+</sup>-ATPase", *Proceeding of the National Academy of Sciences of the United States of America*, 10.1073/pnas.2105507118

**プレスリリース** 理化学研究所  
[https://www.riken.jp/press/2021/20210928\\_1/index.html#wrapper](https://www.riken.jp/press/2021/20210928_1/index.html#wrapper)

## 世界初！ 微小管がメカノセンサーであることを実証

2021.10.11

微小管の構造変化がモータータンパク質のダイナミクスを変調させることを解明

細胞の分化、発生、疾患などにおいて、力学的なストレスは、筋肉や骨に影響を与えたり、脳へのストレスにより記憶や認知障害、アルツハイマー病などの発症につながる恐れがあります。このような力学的ストレスを感知する細胞内のセンサーであるメカノセンサーの役割を、細胞の形態形成や細胞内物質輸送などにおいて重要な役割を果たしている「微小管」も果たしているのではないかと仮説を支持する研究結果が相次いで報告されてきました。

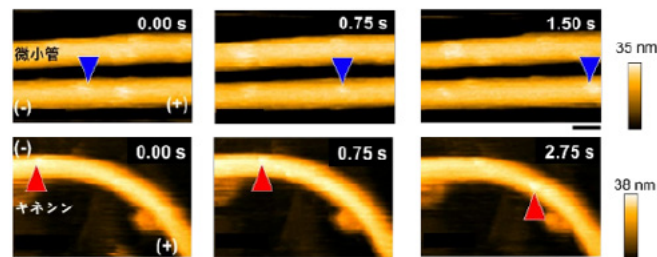
微小管は、細胞内では細胞骨格として機能するだけでなく、細胞活動に必要な物質を輸送するレールとしての役割も果たしています。物質輸送にはモータータンパク質であるキネシンやダイニンなどが関わり、微小管上を移動することで物質が運ばれます。このように様々な細胞活動に携わる微小管は、常に力学的なストレスに晒されることになり、さらに最も剛直な細胞骨格でもあるため、力学ストレスの影響を真っ先に受けることになります。

そのため微小管は、力学的な情報を生化学的な情報へと変換するメカノセンサーとしても機能しているのではないかと考えられてきました。研究グループは、モータータンパク質であるキネシンの運動性を高解像度の観察システムである高速原子間力顕微鏡で解析し、微小管のメカノセンサーとしての機能を分子レベルで解明しました。その結果、微小管に沿って運動するキネシンは、微小管の屈曲等の構造的な変形により運動性を変化させるということを明らかにしました。また全原子分子動力学シミュレーション

研究によって、運動速度を変化させる機構は、微小管の微小な構造変化がキネシンと微小管の結合親和性を変化させることによるものであることも明らかにしました。

これらの結果は、微小管が力学ストレスを感知しモータータンパク質の運動性を変調するメカノセンサーとして機能するという直接的な証拠であり、世界初の報告になります。

細胞を取り巻く力学環境を研究対象としたバイオメカニクスやメカノバイオロジーなどの学術分野、細胞内物質輸送の障害に起因する神経疾患研究、さらに力学センサーなどの開発を目指す材料科学分野への波及効果が期待されています。



屈曲した微小管のレールを走るモータータンパク質キネシン 1 分子の高速原子間力顕微鏡イメージ。屈曲した微小管上ではキネシンは減速走行する

**論文発表** Syeda Rubaiya Nasrin, Christian Ganser, 西川聖二, Arif Md. Rashedul Kabir, 佐田和己, 山下雄史, 池口満徳, 内橋貴之, Henry Hess, 角五 彰, "Deformation of microtubules regulates translocation dynamics of kinesin." (微小管の変形がキネシンの並進運動ダイナミクスを変調する), DOI: 10.1126/sciadv.abf221

**プレスリリース** 北海道大学、自然科学研究機構、東京大学、横浜市立大学、名古屋大学  
[https://www.hpci-office.jp/materials/press\\_20211014-1.pdf](https://www.hpci-office.jp/materials/press_20211014-1.pdf)

TOPICS 1

「富岳」による新型コロナウイルスの治療薬候補同定

「富岳」を使ったシミュレーションにより、他の病気の治療に使われている既存の医薬品の中から、新型コロナウイルスに効果のある治療薬の候補を探索・同定することを目的とした研究

ウイルスタンパク質に  
強く結合する医薬品を探索

現在、新型コロナウイルスの感染拡大に伴い、世界中でワクチンや治療薬の研究開発が進められています。

そもそも新型コロナウイルスなどのウイルスの場合、我々の細胞表面のレセプタータンパク質と呼ばれるタンパク質にウイルスのタンパク質が結合することで、細胞内に取り込まれます。そして、細胞内に取り込まれたウイルスは、細胞内でウイルスRNA（遺伝物質）を放出し、このウイルスRNAがコピーされると同時に、細胞の機能を乗取り、ウイルスタンパク質を合成することで、ウイルス粒子を生成し増殖していくのです。

これまでの研究から、新型コロナウイルスのウイルス粒子の生成に関与しているウイルスタンパク質は、複数個あることがわかっています。そのため、現在、新型コロナウイルスの治療薬の研究開発では、いくつかのウイルスタンパク質に的を絞り、そのウイルスタンパク質に強固に結合することで、ウイルス粒子の生成と増殖を抑える薬剤の探索が進められています。

しかしながら、ウイルスタンパク質と薬剤がどれくらいの強さで結合するかを実験によって測定することは困難です。そこで、実験に代わる有用な手段として注目されているのが、コンピュータ・シミュレーションです。シミュレーションにより、ウイルスタンパク質と薬剤がどれくらいの強さで結合するかを精密に計算することができます。

本研究においては、新型コロナウイルス粒子の増殖に深く関与している「メインプロテアーゼ」と呼ばれるウイルスタンパク質に的を絞り、「富岳」を用いたシミュレーションにより、2000種類以上の既存の医薬品の中から、メインプロテアーゼに対して強い結合性を示す医薬品の探索と同定を行いました。

「京」でも難しかった  
シミュレーションに「富岳」で挑戦

実は、「富岳」の前世代のスーパーコンピュータ「京」でも、ウイルスタンパク質と薬剤がどのくらいの強さで結合するかを、シミュレーションにより計算することは可能でした。しかしながら、「京」での計算では、ウイルスタンパク質と作用することがわかっている薬剤を使って、結合した状態を出発点にシミュレーションしていました。タンパク質は複雑な3次元構造をした分子で、構成する原子の数も非常に多いため、「京」をもってしても、薬剤がウイルスタンパク質にどのように結合するかをシミュレーションすることは、非常に難しかったのです。また、数千種類の薬剤について、限られた時間内でシミュレーションするには、「京」では計算容量が足りませんでした。

一方、新型コロナウイルスの治療薬を「探索」ということは、作用するもの・作用しないものを含めて、どの薬剤がより強く結合するかを見つけ出すことです。そこで、今回、「富岳」を使ったシミュレーションでは、分子動力学シミュレーションと呼ばれる精緻な計算手法を用いることで、メインプロテアーゼと既存の2128種類の医薬品との相互作用をそれぞれ再現し、可視化しました（図1）。これは「富岳」がもつ高い計算性能だからこ

そでできることであり、これにより、医薬品の相対的な評価が可能になりました。また、既存の医薬品を対象とした理由は、すでに医薬品として承認されていることから副作用などのリスクが少なく実用化までの道のりが短いためです。しかもこれらの中には、ウイルス治療薬に限らず、さまざまな病気の治療薬が含まれているため、2000種類以上といった大きな母数から探索ができるのです。

シミュレーションの結果、結合力が強いと予想される数十種類の医薬品を約10日間で割り出すことに成功しました。今後、ソフトウェアの調整が進めば、約2日間での計算も可能になると考えています。

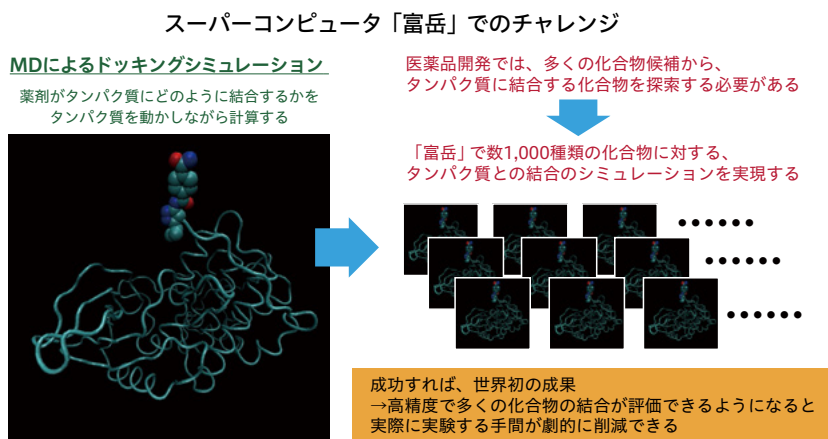


図1 スーパーコンピューター「富岳」でのチャレンジ 新型コロナウイルスの治療薬の探索においては、既存の医薬品の中から、どの医薬品がより強く結合するかを見つけ出すことが求められる。そこで、「富岳」を使った分子動力学シミュレーションでは、メインプロテアーゼと2128種類の医薬品がそれぞれどのように相互作用するかを可視化していった。



## 「富岳」で達成した、 分子動力学シミュレーションでの世界初の成果

タンパク質を構成する原子や分子は、周囲にある水分子などと相互作用をしながら、常に少しずつ動いています。このような原子や分子の動きをシミュレーションするために使われる手法が、分子動力学シミュレーションです。

この手法では、まず、観測データなどを基に、原子の最初の配置を決めます。そして、1個の原子に働く他の原子の力を計算します。原子同士の間には、化学結合の力や静電気力などがあります。次に、力を受けた原子がどのように運動するかをニュートンの運動方程式に基づき計算します。この計算を何度も繰り返すことで、時間の経過とともに原子の配置がどのように変化していくかを再現できます。ただし、タンパク質の場合、何万個もの原子1個1個について、この計算を同時に行う必要があるため、非常に高い計算能力が求められます。それに対し、今回、世界トップの計算性能を誇る「富岳」が威力を発揮しました。分子動力学シミュレーションを用いて、数千規模の化合物（薬剤）とタンパク質の作用を明らかにしたのは世界初の成果です。

## 結合部位だけでなく表面全体に結合する 薬剤の挙動を世界で初めてとらえる

実際、今回のシミュレーションでは、すでに知られているメインプロテアーゼの結合部位（活性ポケット、以後、ポケット）に、どの医薬品がはまり込むかを数値計算していきました。ただし、ポケットにはまり込むからといって結合力が強いとは限らないため、ポケットへの滞在時間を算出しました。

また、シミュレーションを進める中で、興味深い発見がありました。それは、ポケットだけでなく、メインプロテアーゼの表面全体に長く滞在する医薬品があるということです。そこで、今

回実施した2128種類の医薬品について、ポケットへの滞在時間と、表面全体への滞在時間の両方を算出し、縦軸をポケットへの滞在時間、横軸を表面全体への滞在時間として、グラフに点を書き入れていきました（図2）。図2のオレンジの線は、滞在時間上位100位までとそれ以下との境界線です。

その結果、多くの医薬品がポケットにも表面全体にも長い時間滞在しない一方で、数十種類の医薬品が、長い時間滞在する、つまり、強い結合力を示すことがわかりました。また、何の予断や予備知識をもつことなく、単純に「富岳」によるシミュレーション結果のみに基づき、2128種類もの医薬品の結合力の強さの違いを相対的に比べることができたことで、より有用な医薬品を効率良く同定できる可能性が高まりました。

## 数十種類の医薬品を算出、 うち12種類は臨床研究・治験が進行中

実際、シミュレーション結果から結合力が高いと予想された数十種類の医薬品のうちの12種類は、有望な新型コロナウイルス治療薬候補として、現在、海外で臨床研究や治験が実施されているものであることがわかりました。これは、「富岳」によるシミュレーションの精度や信頼性の高さを裏付けるものです。

中でも以前から有望視されていた「ニコロサミド」と「ニタゾキサニド」の2つが、シミュレーション結果のみから、メインプロテアーゼのポケットにも表面全体にも長く滞在する医薬品として示されたことは、注目に値します（図3）。また、結合力が高いと予想された医薬品の中には、日本の製薬メーカーのものも含まれていました。

今回の成果を踏まえ、今後、ライセンスを有する製薬会社や医学研究者と臨床研究や治験について協議していくほか、メインプロテアーゼ以外のタンパク質についても有望な医薬品の探索と同定を進めていく計画です。それにより、新型コロナウイルスの1日も早い収束に貢献したいと考えています。

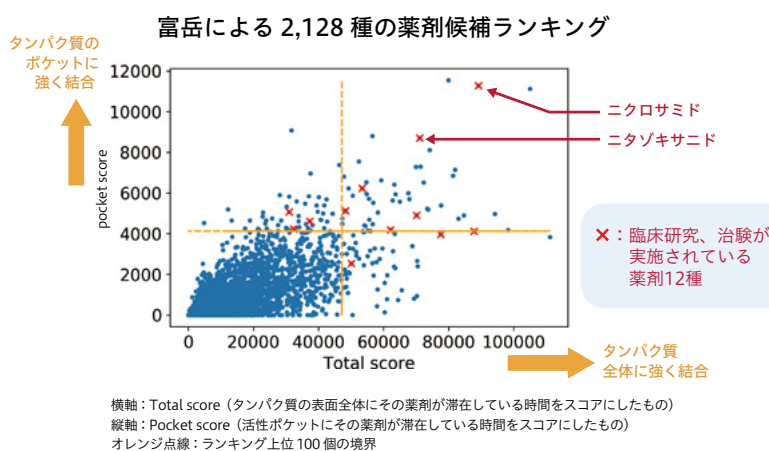


図2 富岳による2,128種の薬剤候補ランキング 今回実施した2,128種類の医薬品について、結合部位（ポケット）への滞在時間と、表面全体への滞在時間の両方を算出し、縦軸をポケットへの滞在時間、横軸を表面全体への滞在時間として、グラフに点を書き入れていった。その結果、数十種類の医薬品が、強い結合力を示すことが明らかとなった。しかも、そのうちの12種類が有望な新型コロナウイルスの治療薬候補として、現在、海外で臨床研究や治験が実施されているものであることがわかった。

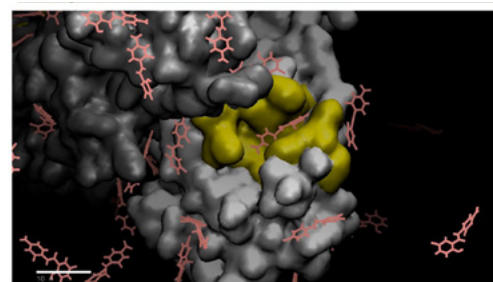


図3 Main proteaseとNiclosamideとの結合シミュレーション ニコロサミド（ピンク色の化合物）が新型コロナウイルスのメインプロテアーゼ（灰色の部分全体）の結合部位（ポケット、黄色の部分）に長く滞在している様子。今回、ポケットだけでなく表面全体にも長く滞在していることが初めてわかった。

## TOPICS 2

## COVID-19の治療薬候補を既存薬の中から計算で探し出す

私たちの生活を一変させた新型コロナウイルス感染症 (COVID-19)。

世界中の製薬会社や研究機関がその治療薬の開発に取り組んでいます。

奥野さんは、スーパーコンピュータ「富岳」を中核とする HPCI(革新的ハイパフォーマンス・コンピューティング・インフラ)を利用して精密なシミュレーションを行い、2128種類もの既存薬(すでに他の病気の治療に使われている薬)の中から

COVID-19の治療薬として有望な候補を見つけることに成功しました。

\*HPCI マガジン「富岳百景」(高度情報科学技術研究機構)に掲載されたものです。

### 精密なシミュレーションから生まれた 「驚くほどよい結果」

COVID-19の被害を抑え、経済・社会活動を回復するために、治療薬やワクチンの開発が急ピッチで進められています。治療薬については、抗ウイルス薬を中心とした既存薬の中から新型コロナウイルスに効果を示すものを見つけ出そうという動きがさかんです。既存薬は治験で安全性が確認されているため、開発期間の短縮が期待できるからです。しかし、世界中で細胞実験から治験まで様々なレベルの研究が行われているにもかかわらず、まだ効果的な治療薬の発見には至っていません。

こうした状況の中、期待を集めているのがスーパーコンピュータを活用したシミュレーションです。「実験よりもずっと短い時間でたくさんの薬を調べることができる上に、薬の作用の仕方までわかるからです」と奥野さんは説明します。奥野さんは、理化学研究所などが開発・整備中の新型コロナウイルス対策のための「富岳」の優先的な試行的利用と、HPCI コロナウイルス感染症対応臨時公募を利用し、「富岳」と、筑波大学の「Cygnus」を使って、2020年4月に既存薬の中からCOVID-19の治療薬候補を探索する研究に着手しました。

「富岳」を用いた研究課題は、新型コロナウイルスが細胞内で増殖する際に重要な「メインプロテアーゼ (NSP5)」というタンパク質をターゲットとし、2,128種類もの既存薬の中から治療

薬の候補を探そうというものです。「富岳」の中にこのタンパク質と既存薬分子のモデルをつくり、「分子動力学シミュレーション」という手法でタンパク質と薬分子のふるまいを計算しました。「この計算により、タンパク質の活性ポケット\*1が開いたり閉じたりする様子や、薬分子が離れたところからやってきて活性ポケットや他の部分にくっついたり離れたりする様子を、精密に再現することができました (図1)」。

そして、薬分子がメインプロテアーゼの活性ポケットに滞在している時間などを指標として既存薬をふるいにかけることで、数十種類の候補を選び出すことに成功しました (図2)。「これらの薬は、メインプロテアーゼに結合する能力が高いと予想され、新型コロナウイルスに薬効を示す可能性があります。シミュレーションを行う前は、このような薬がほとんど見つからないとか、逆に、たくさん見つかりすぎてしまうことを心配していましたが、妥当な数の候補が見つかりました。さらに、このうちの12種類は海外で新型コロナウイルス治療への臨床研究や治験が行われています。自分でも驚くほど、よい結果が出ました」と奥野さん。臨床研究や治験に進んだ薬は効果が確認されているわけで

\*1 タンパク質の中には「くぼみ」をもつものがあり、そこに決まった分子が結合することで働き始める。このくぼみを「活性ポケット」と呼ぶ。決まった分子の代わりに別の分子が活性ポケットに結合すると、タンパク質の動きを抑えられる可能性があるため、そのような分子を探すことが薬の開発の第一歩となる。

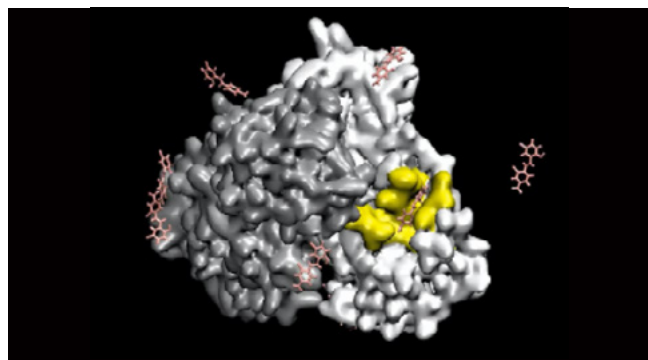


図1 「富岳」で行った分子動力学シミュレーションの一例 新型コロナウイルスの増殖に重要なメインプロテアーゼというタンパク質(灰色)と、ニコロサミドという既存薬(ピンク)のふるまいをシミュレーションした動画の一場面。活性ポケット(黄色)に1個の薬分子が入り込んでおり、タンパク質から離れている薬分子や、活性ポケットとは別の場所にくっついている薬分子もある。このようなシミュレーションを2,128種の既存薬について行った。

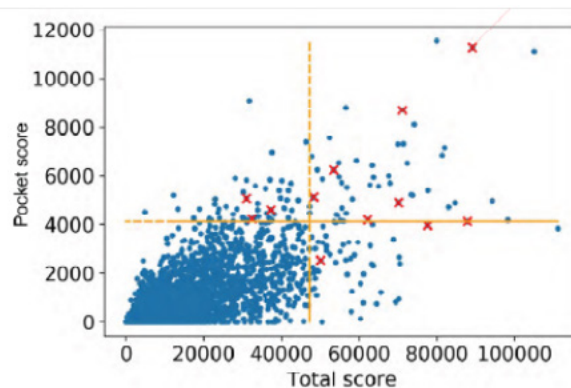


図2 治療薬候補の絞り込み 図1のようなシミュレーションで、薬がタンパク質の活性ポケットに滞在する時間(縦軸)と、タンパク質の表面のどこかに滞在する時間(横軸)を調べた。グラフ上1個の点が、1種類の薬を表し、×は海外で臨床試験または治験が行われている薬を示す。このグラフでは右上ほどタンパク質との結合能力が高いと予測され、黄色い線で区切った右上の部分にある数十種類の薬を、治療薬の候補と判定した。

すから、そうした薬がシミュレーションでも候補に選ばれたということは、シミュレーションの信頼性の高さを示しています。

12種類の中でも特に結合能力が高いと予想された「ニコロサミド」という薬は、日本では未承認ですが、海外では寄生虫の駆除薬として使われており、安価で安全性も確認されています。今回の結果を受けて、臨床研究に弾みがつくことが期待されます。

### 計算法の工夫で「富岳」の 計算能力を生かす

奥野さんたちが「富岳」での計算を開始したのは4月末。準備期間は1ヵ月もありませんでしたが、計算は順調に進み、6月末には治療薬候補を得ることができました。こんな離れ業ができたのは、『京』で培ったノウハウがあった上に、『富岳』がとても使いやすい計算機だったからです。

奥野さんたちは、「京」でも薬の候補を選ぶためのシミュレーションを行っていましたが、「京」の計算能力では、今回のような精密なシミュレーションを行う場合、調べられる薬の数は数十種類ほどでした。「創薬の最初の段階では『多数の中から選ぶ』ことが重要であり、実験では非常に多くの候補を調べるのですが、『京』ではまだそのプロセスをシミュレーションで実践するには届いていませんでした。しかし、今回、『富岳』で2,000種類以上の薬を計算できたことで、実践に大きく近づいたと思います。精密な分子動力学シミュレーションにより2,000種類以上の既存薬の中から候補を選び出せたことは、世界初の成果であり、計算技術としても、学術的にも大きな意義をもっています。しかも、この計算は、「富岳」のごく一部である5,500ノードを使用しただけで達成されたのです。

この成功の背景には、「富岳」の計算能力に加えて、計算上の工夫や努力がいくつもありました。その1つは、計算の際に薬分子の濃度を通常の数百倍に高め、タンパク質との相互作用が起りやすいようにしたことです。計算時間を短縮するための工

夫ですが、濃度が高すぎるなど条件によっては計算ができなくなってしまうため、慎重な条件設定が必要でした。また、薬が活性ポケットに滞在する時間が実験値と整合するかも確かめました。さらに、「富岳」では膨大な計算結果が得られるため、その解析方法も同時に開発する必要がありました。

### Cygnusのよさを生かして 別のタンパク質もシミュレーション

筑波大学の「Cygnus」では、「富岳」で選び出された治療薬候補が、新型コロナウイルスがつくる別のタンパク質(NSP16)に結合するかどうかを調べました。NSP16の阻害剤(NSP16に結合して、その働きを抑える薬)は、メインプロテアーゼの阻害剤に比べると開発途上であり、既存薬の可能性をシミュレーションで探ることは挑戦的な試みです。

「Cygnus」は、人工知能の研究でもよく使われるGPU<sup>\*2</sup>の最新版を豊富に搭載しており、分子動力学シミュレーションを高速に行うことが可能です。『Cygnus』はすでに正式運用中でしたので、計算に取りかかるまでに短時間で計算機に合わせた準備ができ、安定した計算ができました。

「富岳」での計算と同様、薬分子の濃度を高めて分子動力学シミュレーションを行い(図3)、その結果から、薬分子がNSP16の活性ポケットに結合する能力を予想します。計算結果は解析中ですが、タンパク質が異なれば結合しやすい薬も変わると考えられ、どのような解析結果が出るかが注目されます。

\*2 「Graphics Processing Unit」の略で、画像処理に適した演算装置を指す。コンピュータの演算装置としては、CPU(中央演算装置)が搭載される場合が多いが、計算の種類によっては、GPUのほうが計算が速い場合がある。

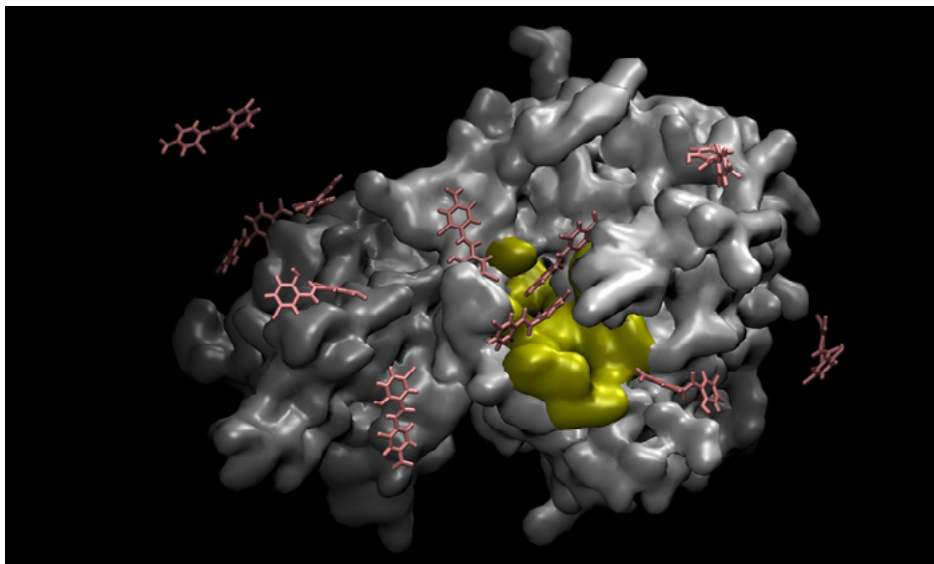


図3 Cygnusで行った分子動力学シミュレーションの一例。メインプロテアーゼをターゲットとして得られた治療薬候補であるニコロサミドと、NSP16との相互作用を、「富岳」を使用したときと同様の手法で計算した。

# プロジェクトの 統合的推進

プロジェクト内の分担機関間の研究者や研究開発担当者との連携を密に取りながら、プレジジョンメディスンの加速を目指し、研究開発と一体となって、分野振興や研究成果の普及など、分野を超えた取り組みの推進を行っています。

## 外部機関との連携協力

### 医療機関

(実データを創出・管理する主体との連携)

国立がん研究センター  
がん研究会 がん化学療法センター  
京都大学医学部付属病院 がんセンター  
慶応義塾大学病院臨床研究推進センター  
東京医科歯科大学医学部付属病院  
愛知県がんセンター  
佐賀大学医学部付属病院がんセンター  
国立国際医療研究センター

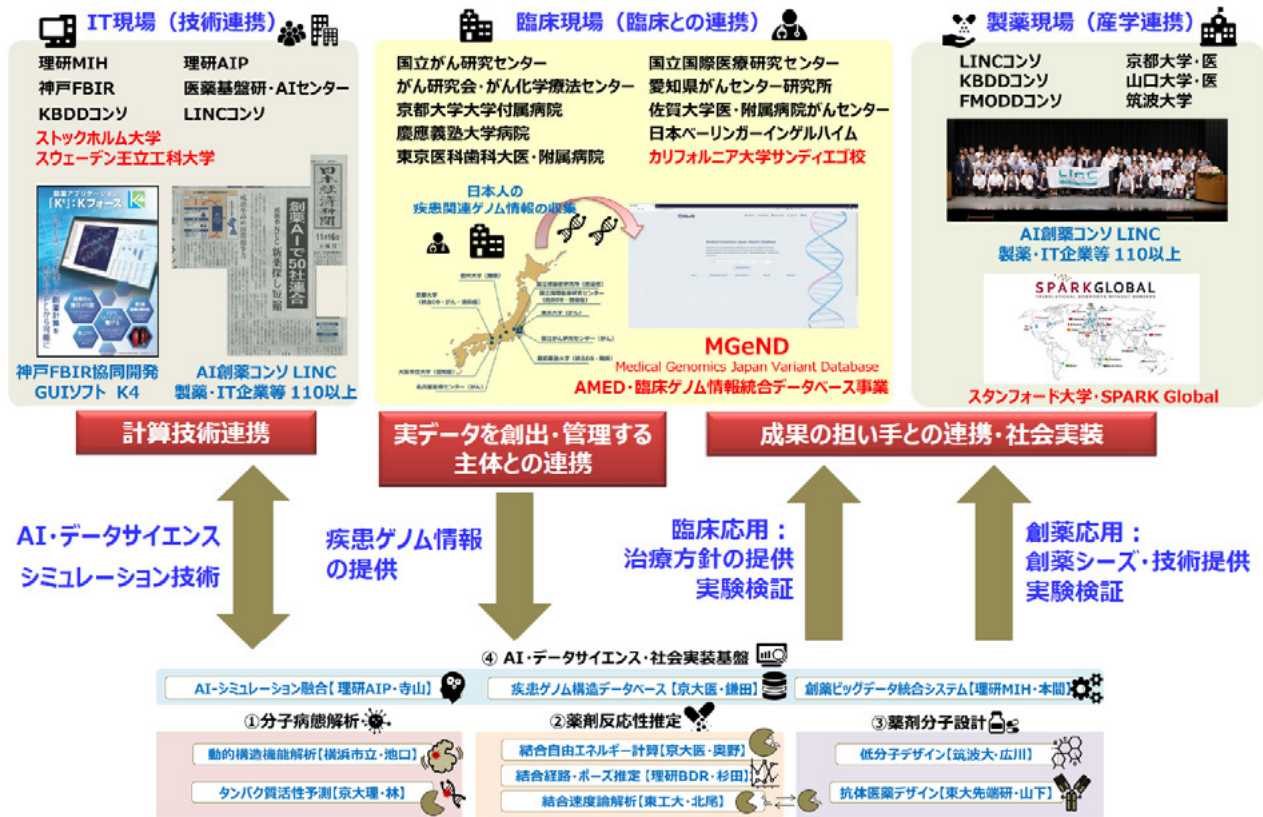
### 計算技術連携

■ AI・データサイエンス連携  
津田 宏治(理化学研究所 革新知能統合研究(AIP)センター)  
水口 賢司(医薬基盤健康栄養研究所)  
ライフ インテリジェンス コンソーシアム(LINC)  
■ シミュレーション連携  
KBDD(K supercomputer-Based Drug Discovery project by Biogrid pharma consortium)  
神戸医療産業都市推進機構  
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

### 国の他の研究プロジェクトや行政組織、民間企業など

(成果の担い手との連携)

■ 臨床応用  
間野 博行(国立がん研究センター)  
河野 隆志(国立がん研究センター)  
片山 量平(がん研究会)  
日本ペーリンガーインゲルハイム  
カリフォルニア大学サンディエゴ校  
■ 創薬応用  
KBDD(K supercomputer-Based Drug Discovery project by Biogrid pharma consortium)  
ライフ インテリジェンス コンソーシアム(LINC)  
FMO創薬コンソーシアム(FMODD)  
神戸医療産業都市推進機構  
柳田 素子(京都大学医学研究科 医学専攻内科学講座 腎臓内科学)  
清木 誠(山口大学医学部医学科 システムズ再生・病態医学)  
小柳 智義(筑波大学つくば臨床医学研究開発機構)  
計算技術連携



## 研究成果の普及

「富岳」による最先端の科学的成果創出や成果の社会実装を強力に推進することを目指し、研究開発の成果を学会のシンポジウムやセミナー開催、ポスター発表などの活動を通じて情報発信しています。

### ワークショップ・ セミナー

### スーパーコンピュータ「富岳」創薬活用ワークショップ

本プログラムでは、2021年3月9日からのスーパーコンピュータ「富岳」共用開始を受け、産業分野、連携する臨床の先生方を交え、研究内容の紹介を中心に、創薬・ゲノム医療への応用を視野に入れた年数回の意見交換を行っています。

### GENESIS 講習会

GENESISは独自の高速化、並列化手法を持つため、効率の良い計算手法が利用でき、富岳などのスーパーコンピュータやGPU 計算機、インテル計算機などに対応しています。このソフトウェアをより多くの方に使っていただくため講習会を開催しています。

### 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

発行 「富岳」成果創出加速プログラム ②国民の生命・財産を守る取組の強化の領域 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

理化学研究所 計算科学研究センター HPC/AI 駆動型医薬プラットフォーム部門 <http://mddpm.riken.jp>

制作協力: 京都通信社 デザイン: 中曽根デザイン