

G1-1 天然変性領域の構造アンサンブル解析とアミノ酸変異等の影響解析

独自のツールで、天然変性領域の構造アンサンブル解析、アミノ酸変異、翻訳後修飾等の影響解析を支援

[1] 支援担当者

所属		①量子技術研究開発機構 量子生命科学領域
氏名		①河野 秀俊
AMED 事業	ユニット／領域名 課題名	インシリコユニット モデリングとシミュレーションによる生体高分子の構造機能解析
	代表機関 代表者	量子技術研究開発機構 河野 秀俊
支援技術のキーワード		アミノ酸変異、翻訳後修飾、天然変性、モデリング、分子動力学計算

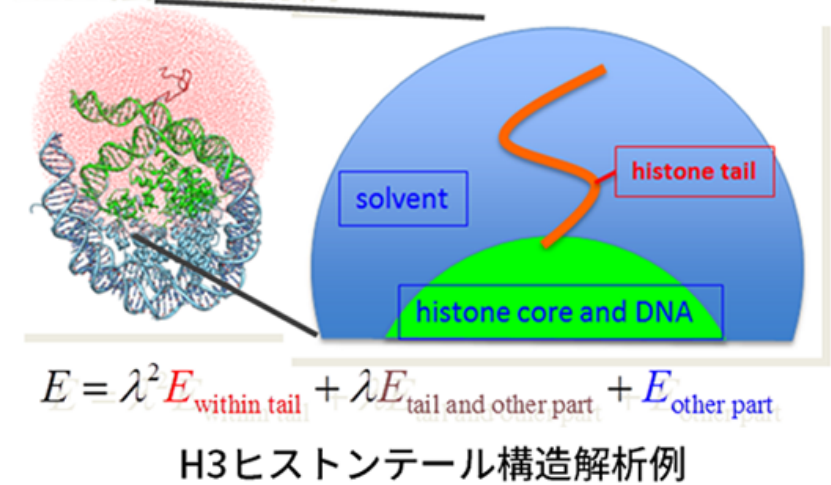
[2] 支援技術の概要

- 水中の全原子分子動力学（MD）シミュレーションによる動的挙動の詳細な解析
- LDA-ITER (Liner Discriminant Analysis with Iteration) 解析による、複数のMDデータの同時解析、MDデータの違いを生み出している要因残基の同定
- 拡張サンプリング法（ALSD法）による天然変性領域の構造アンサンブルの取得とアンサンブルの特徴解析
- タンパク質-DNA相互作用解析とDNA配列に依存した構造特性の解析

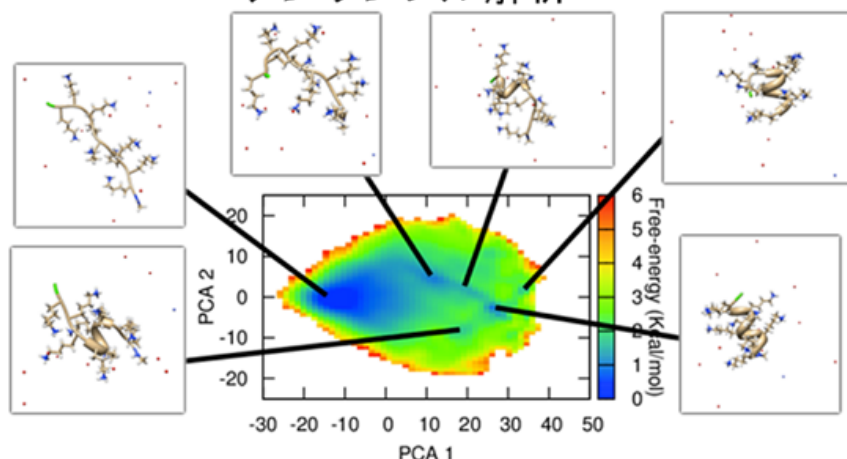
[3] 支援技術の利用例

- アミノ酸変異や翻訳後修飾の動的挙動への影響解析
- リガンドありなしでの動的挙動の解析及びリガンド結合、乖離のパスウェイ解析
- 複数のMD結果から動的挙動の違いを生み出している残基の推定
- 天然変性領域の構造アンサンブル特徴解析
- 転写因子のアミノ酸置換とDNA結合能変化の解析
- タンパク質複合体のモデリングと動的解析

ALSD法の適用例



アンサンブル解析

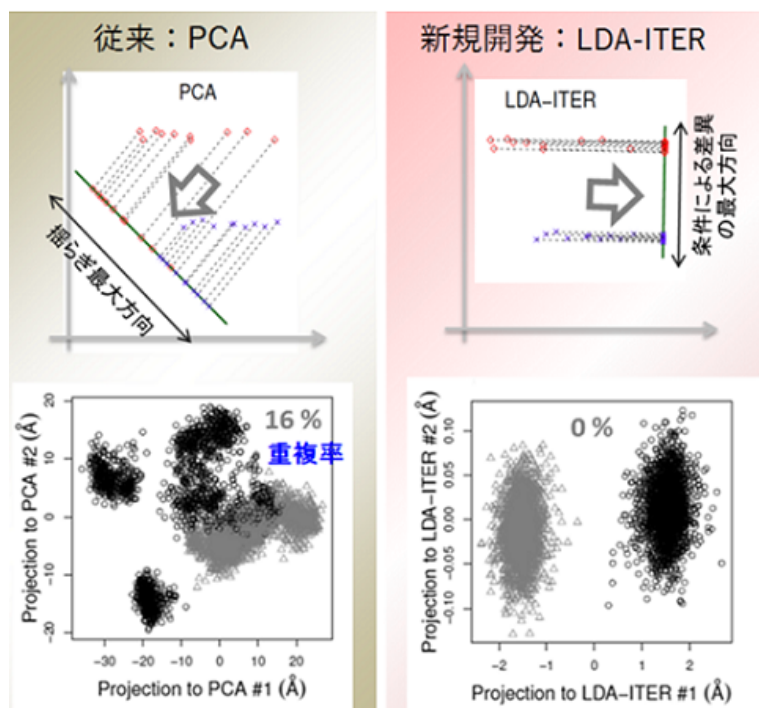


[4] 支援担当者の研究概要

分子シミュレーション、分子モデリング、生化学実験を通して、分子の機能発現メカニズムを調べている。特に、アミノ酸置換やエピソード的な変化が、分子構造の動特性に与える影響や分子間相互作用に与える影響を中心に研究を展開している。

そのため、新たな分子シミュレーション方法の開発やシミュレーションデータの解析方法を開発し、支援に供している。本事業でも活用する、我々が開発した Adaptive Lambda Square Dynamics (ALSD) 法を用いて、天然変性領域のような特定の安定構造を取らない分子の構造アンサンブルを取得し、そのアンサンブルの特徴解析から、分子の構造多様性や複合体形成時の構造を探っている（上図参照）。

また、異なる条件で実行されたMDデータの違いを解析できる反復 Linear Discriminant Analysis (LDA-ITER)法を用いて、変異などによる動特性への影響を調べている。この解析方法は、アミノ酸変異やリガンド有無などの条件の異なるMDシミュレーション結果から、条件の違いに起因する違いを見つける方法である（下図）。タンパク質など生体高分子の構造揺らぎは、一般的に異方性が高い。そのため、良く用いられているPCA法では、条件の違いによる差異が分子本来の構造揺らぎに隠されがちである。この方法では、条件の同じMDデータ間の差異を最小化し、条件の異なるMDデータ間の差異を最大化するようにデータを見る（射影する）ことによって、条件の違いに起因する変化を見つけ出すことができる。同時に、その変化の要因となった残基を同定することができる。



G1-2 低分解能実験データを用いた超分子の立体構造モデリング

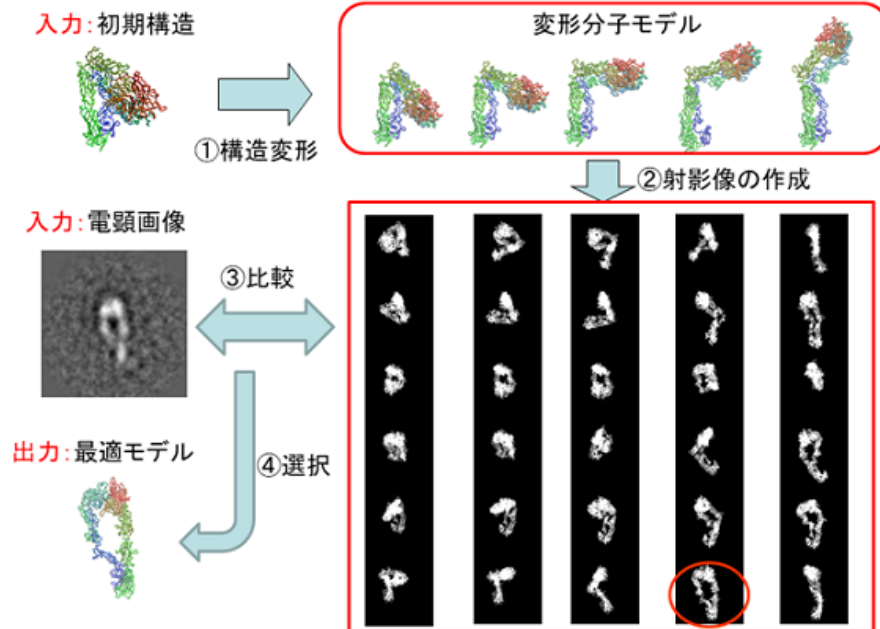
電子顕微鏡画像や量子ビーム散乱データなどの低分解能実験データを再現する超分子立体構造のモデリング

[1] 支援担当者

所属		①量子技術研究開発機構 量子生命科学領域
氏名		①松本 淳
AMED 事業	ユニット／領域名 課題名	インシリコユニット モデリングとシミュレーションによる生体高分子の構造機能解析
	代表機関 代表者	量子技術研究開発機構 河野 秀俊
支援技術のキーワード		モデリング、シミュレーション、電子顕微鏡、低分解能実験データ、超分子

[2] 支援技術の概要

電子顕微鏡画像や量子ビーム散乱データなどの低分解能実験データをもとに、計算機を用いて、超分子の立体構造モデリングを行う。



電子顕微鏡画像をもとにした3次元分子モデル構築の計算の流れ

[3] 支援技術の利用例

- 電子顕微鏡画像や量子ビーム散乱データなどの低分解能実験データからの3次元分子モデルの構築。
- 多数の電子顕微鏡画像から3次元分子モデルを構築することによる構造分布情報の取得。
- 構造分布情報の解析による構造変化経路の推定。
- 構築した多数の3次元分子モデルを用いて、超分子の構造変化の様子を捉えた動画の作成。

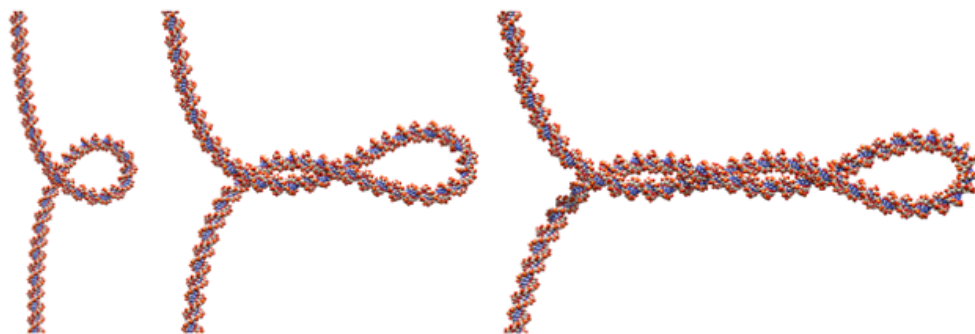
[4] 支援担当者の研究概要

電子顕微鏡画像や量子ビーム散乱データなどの低分解能実験データをもとにした分子モデル構築手法の開発を行っている。電子顕微鏡画像による単粒子構造解析の進歩は著しいが、多様な構造をとる分子は画像分類や平均化が困難なため、その3次元分子モデル構造の構築は極めて困難である。そこで、1枚のネガティブステイン電子顕微鏡画像をもとにして3次元分子モデルを構築する方法を開発した。そのほか、小角X線散乱（SAXS）や小角中性子線散乱（SANS）などの量子ビーム散乱データ、それに3DEM構造情報をもとにした3次元分子モデル構築を行っている。

また、実験データを再現するには、初期モデル構造を大きく変化させる必要がある場合があり、そのため、大規模構造変形のための計算手法の開発も行っている。



分子モデル（左）から作成した疑似電顕画像（中央：単純な射影画像、右：ネガティブステインモデル）。単純な射影画像では、2つの分子モデルを区別できないが、ネガティブステインモデルでは区別できる。



DNAの大規模構造変形

G2-1 分子シミュレーションによるタンパク質の機能解析

立体構造予測、分子動力学シミュレーション、量子化学計算など分子シミュレーションに基づくタンパク質の機能解析

[1] 支援担当者

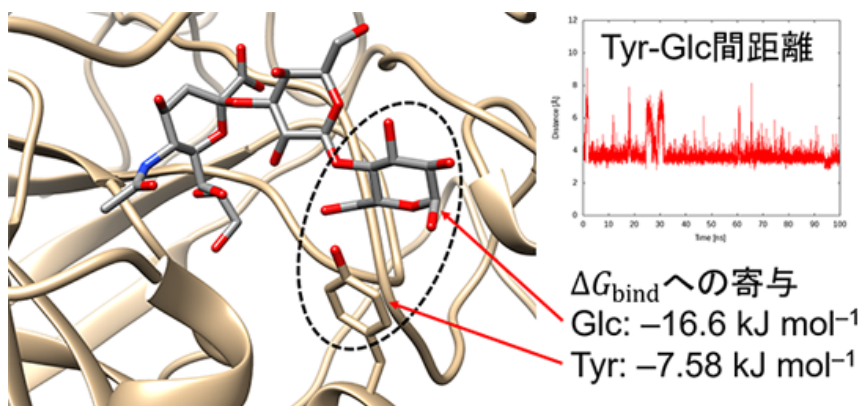
所属		①東京大学 大学院情報学環 ②東京大学 大学院農学生命科学研究科
氏名		①寺田 透 ②森脇 由隆
AMED 事業	ユニット／領域名 課題名	インシリコユニット タンパク質の複合体構造・分子間相互作用・ダイナミクス・機能の予測と解析
	代表機関 代表者	東京大学 寺田 透
支援技術のキーワード		比較モデリング、ドッキング、分子動力学シミュレーション、量子化学計算

[2] 支援技術の概要

比較モデリング法を用いてタンパク質単体の立体構造予測を、ドッキング法を用いて、タンパク質－タンパク質および、タンパク質－低分子化合物（リガンド）の複合体構造予測を行う。分子動力学（MD）シミュレーションを用いて、これら予測構造の精密化と安定性の評価を行う。また、MDシミュレーションを用いて、分子内・分子間相互作用解析、自由エネルギー計算、ダイナミクス解析などを行う。さらに、量子化学計算を用いて、反応メカニズム解析やMDシミュレーションで使用する力場パラメータの計算を行う。

[3] 支援技術の利用例

- Gタンパク質共役受容体とGタンパク質の立体構造予測、受容体2量体とGタンパク質の複合体構造予測、及び相互作用部位予測
- 同じリガンドへの親和性が異なるタンパク質群に対する、親和性の違いの原因となる残基の推定
- 2種類のヘム輸送タンパク質からなる輸送中間体の構造予測と、量子化学計算に基づく配位構造解析
- 量子化学計算に基づく、酵素反応における基質特異性のメカニズムの解明
- リガンドのタンパク質への結合自由エネルギー ΔG_{bind} に対する、タンパク質残基及びリガンド部分構造の寄与の計算（下図）



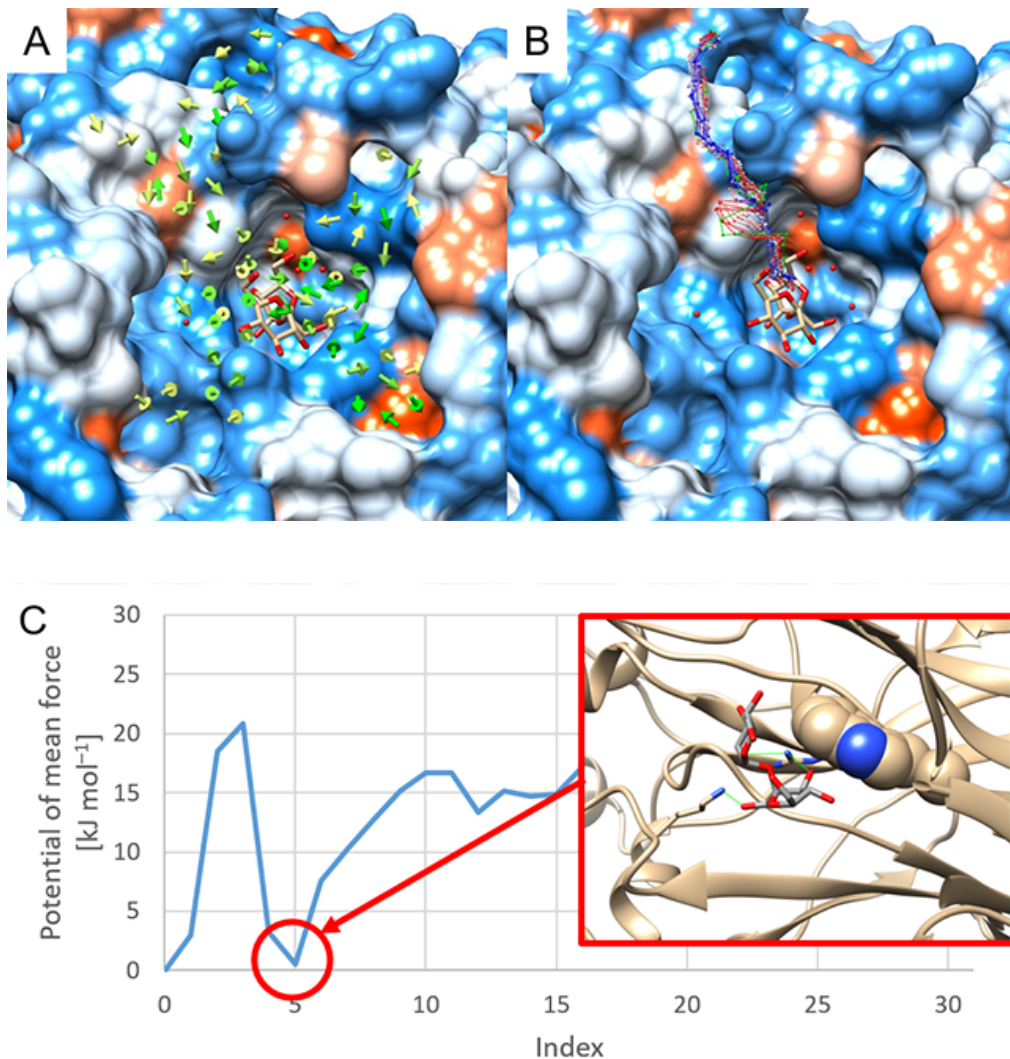
おたふく風邪感染に関わるウイルスタンパク質と細胞表面にある受容体糖鎖との間の相互作用解析

[4] 支援担当者の研究概要

タンパク質が機能する際の「動き」に注目し、この動きを生み出すメカニズムを、分子シミュレーションを用いて原子レベルで明らかにするための研究を行っている。

リガンドがタンパク質に結合する過程を明らかにするためにに行った研究を下図に示す。リガンドの結合過程を実験で追跡するのは困難であるため、MDシミュレーションによる結合過程の再現に期待が集まっている。しかし、リガンド結合の時間スケールはミリ秒程度で、通常の全原子モデルを用いたMDで追跡可能な上限（マイクロ秒程度）を大幅に超えるため、非常に速くリガンドが結合する系を除いて適用は難しい。そこで、私たちは計算コストの低い粗視化モデルを用いて結合経路を網羅的に探索した後、頻繁に現れる経路を、全原子モデルを用いて精密化する、マルチスケールアプローチを採用している。

Aは粗視化モデルを用いたMDシミュレーションの結果得られたリガンドの流れを示しており、リガンドが特定の経路を経由してリガンド結合ポケットに入ることを示している。ここで得られた経路の1つを、全原子モデルを用いて最適化し（B）、この経路に沿った自由エネルギープロファイルを計算した（C）。ここから、本来の結合状態の手前に準安定状態が存在し、疎水相互作用によって安定化されていることが示された（C）。



G2-2 タンパク質の相互作用予測、相互作用部位予測

立体構造未知のタンパク質に対する、アミノ酸配列を入力とした相互作用予測および相互作用部位予測

[1] 支援担当者

所属		①東京大学 大学院農学生命科学研究科 ②東京大学 大学院情報学環
氏名		①清水 謙多郎 ②寺田 透
AMED 事業	ユニット／領域名 課題名	インシリコユニット タンパク質の複合体構造・分子間相互作用・ダイナミクス・機能の予測と解析
	代表機関 代表者	東京大学 寺田 透
支援技術のキーワード		相互作用予測、相互作用部位予測、タンパク質機能予測、機械学習

[2] 支援技術の概要

立体構造未知のタンパク質に対して、アミノ酸配列のみを入力として、タンパク質が他の分子（タンパク質、核酸、糖、脂質、金属など）と相互作用するかどうか、相互作用する場合はどの部位で相互作用するかを、機械学習を用いて予測する。立体構造が既知の場合は、立体構造を元にしたリガンド結合部位（空間配置）予測を行う。

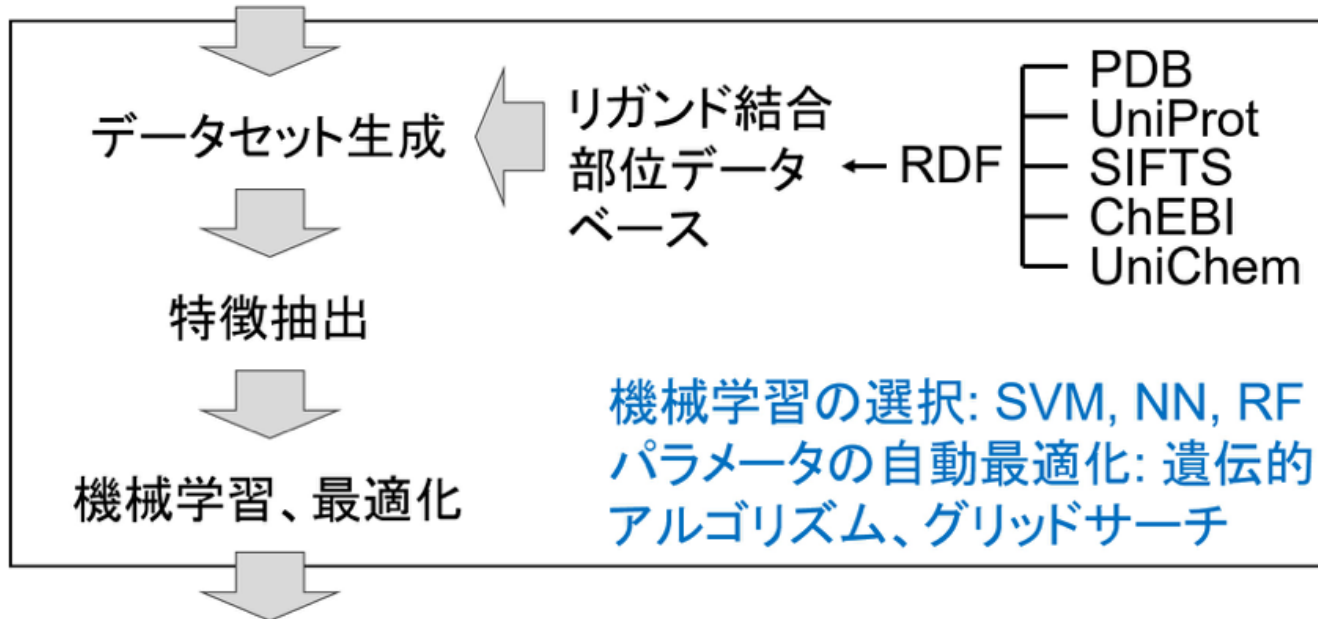
[3] 支援技術の利用例

- 構造がわからないタンパク質に対する機能予測（構造既知の場合はより高い精度で予測）
- タンパク質－タンパク質、タンパク質－リガンドのドッキングの前段階としての利用
- 計算結果と実験データとを比較しながら、手法やパラメータを調整し、精度の向上を図る
- 簡易版はWebで公開

[4] 支援担当者の研究概要

- 任意のリガンドの結合部位予測
リガンド結合部位予測ツールを自動生成し、さらに学習パラメータを自動的に最適化。
Komiyama *et al. Bioinformatics*, **32**, 901-907 (2015).

タンパク質の配列、リガンド



結合部位予測ツールの自動生成

- タンパク質-タンパク質相互作用予測
ホモロジーと機械学習を統合した高精度の相互作用予測システムを開発中。また、チロシンキナーゼに特化し、ドメイン情報と機械学習を統合した予測システムを開発し、AUC値0.996を達成。
Yarimizu, et al. *Adv. Bioinformatics*, **2015**, 528097 (2015).
- Tail-anchorタンパク質の予測
Tail-anchorタンパク質の膜貫通領域を含む配列パターンを分類し、HMMで学習、AUC値0.963を達成。
Shigemitsu et al. *BMC Bioinformatics*, **17**, 378 (2016).
- タンパク質構造をもとにしたリガンド結合部位（空間配置）予測
エネルギー最小化と配列の保存性の両方を考慮
Tsuiikawa et al. *J. Struct. Funct. Genomics*, **17**, 39-49 (2016).

性能評価の値は正解のリガンドをカバーする割合	結合空間予測 (bound)		結合残基予測 (unbound)	
	Top 1	Top 3	Top 1	Top 3
本手法	74.0	88.0	73.9	85.6
Q-SiteFinder	52.0	71.9	53.4	74.0

- 糖鎖結合部位の予測
酸性糖と非酸性糖に分けて特徴抽出し、さらにそれらを統合した予測プログラムを構築。
Banno *et al. Comput. Biol. Chem.* **66**, 36-43 (2017).
- 脂質結合タンパク質の分類予測
特徴抽出の手法について、プロファイルを拡張した新手法を提案し、lipid binding, lipid degradation, lipid metabolism, lipid synthesis, lipid transport, lipoproteinなどの分類を高精度（全体のAUC値0.98を達成）で予測。
Ueki *et al. Res. J. Life Sci. Bioinform. Pharm. Chem. Sci.* **1**, 223-236 (2016).
- 天然変性タンパク質のMoRFsの同定
天然変性領域のアミノ酸配列からMoRFs（molecular recognition features）を同定するシステムを開発し、世界最高性能で予測。
Fang, *et al. J. Bioinform. Comput. Biol.* **17**, 1950004 (2019).
- その他の結合部位予測
対象の分子に応じた特徴抽出を行い、タンパク質のDNA結合部位予測、金属結合部位予測などの手法を開発。

G3-1 バイオマーカーおよび治療法開発を加速するデータ駆動型モデリング

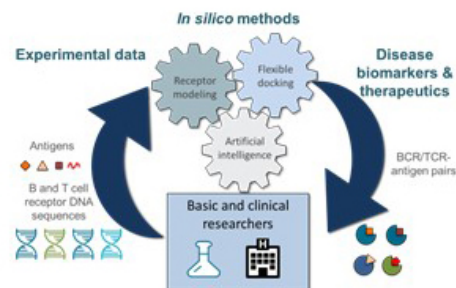
Functional analysis of B and T cell receptor sequence data, RNA- and DNA-binding protein modeling, multiple sequence alignment

[1] 支援担当者

所属		①大阪大学 微生物病研究所遺伝情報実験センターゲノム情報解析分野
氏名		①Daron Standley, 川口 美恵、都 雅子
AMED 事業	ユニット／領域名 課題名	インシリコユニット バイオマーカーおよび治療法開発を加速するデータ駆動型モデリング
	代表機関 代表者	大阪大学 Daron M. Standley
支援技術のキーワード		antibody, antigen, sequencing, structure, RNA

[2] 支援技術の概要

We are developing various bioinformatics techniques to support the discovery of immune-based biomarkers and therapeutics. To this end, we can perform high-throughput receptor modeling, flexible docking and functional classification of results using AI.



Overview of functional BCR and TCR discovery

[3] 支援技術の利用例

- High-throughput BCR and TCR modeling
- Structural clustering of BCR/TCR models
- Prediction of nucleotide binding sites
- Flexible nucleotide-protein docking
- Multiple sequence & structure alignment

[4] 支援担当者の研究概要

The research in our laboratory is concentrated in three main areas: (1) Analysis of B cell receptors (BCRs) and T cell receptors (TCRs) from large-scale sequence data; (2) modeling protein-nucleotide interactions, especially those that mediate immune responses; (3) multiple sequence and structural alignment.

Analysis of BCRs and TCRs from large-scale sequence data. BCRs and TCRs undergo molecular evolution within the lifetime of the host organism in response to antigen exposure. They also represent the fastest-growing class of therapeutic molecules in a wide range of disease areas. We are thus focusing on developing methods that can harness the Big Data from BCR and TCR sequencing studies to develop biomarkers and therapeutics. We make use of recent technology developed in our lab for modeling BCRs and TCRs to atomic resolution in a short amount of time (e.g., seconds).

Modeling protein-nucleotide interactions that mediate immune responses. Immune cells regulate the strength and duration of responses to external stimuli both at transcriptional and post-transcriptional levels. Of these, the post-transcriptional regulatory pathways that act via RNA-binding proteins (RBPs) have only recently been appreciated. In order to identify novel RBPs and their modes of action, we have developed a class of tools that can flexibly dock single-stranded RNA and DNA molecules to proteins using only sequence information (for both nucleotide and protein as input).

Multiple sequence and structural alignment. We develop the MAFFT multiple sequence and structure alignment package, which is one of the most popular and versatile alignment methods available.

G4-1 X線小角散乱などの実験データと分子動力学シミュレーションの連携研究

構造生物学実験と連携しつつ、スーパーコンピュータ等を利用した分子動力学シミュレーションの支援研究

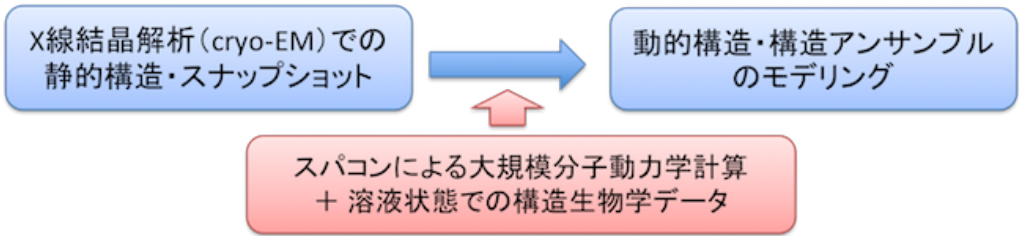
[1] 支援担当者

所属		①横浜市立大学 大学院生命医科学研究科
氏名		①池口 満徳
AMED 事業	ユニット／領域名 課題名	インシリコユニット 分子動力学計算による構造生物学データを活用した構造ダイナミクス研究
	代表機関 代表者	横浜市立大学 池口 満徳
支援技術のキーワード		分子動力学シミュレーション、分子モデリング、X線小角散乱、NMR

[2] 支援技術の概要

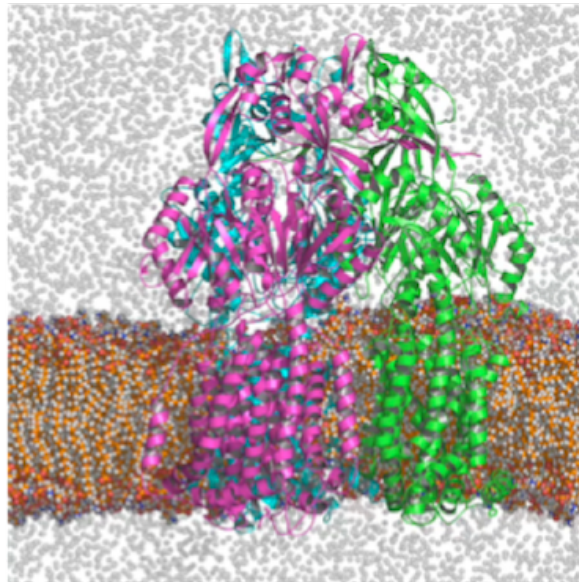
下記のようなケースについて、X線小角散乱（SAXS）やNMRなどの構造生物学実験と連携しつつ、スーパーコンピュータ等を利用した分子動力学シミュレーションの支援研究を行う。

- 溶液構造（結晶構造と異なる）のモデリング
- 分子モーター等の立体構造変化過程
- 他のリガンドの結合構造や非結合型構造のモデリング
- ディスオーダー領域を含むタンパク質のアンサンブル構造
- 部分ドメイン構造から、全長のタンパク質の構造モデリング
- リガンド結合による構造変化を誘起、アロステリック効果
- 結合サイトのダイナミクス、リガンドの結合能や選択性の予測
- タンパク質間相互作用のホットスポットの特定
- 機能的タンパク質のアミノ酸変異の設計



[3] 支援技術の利用例

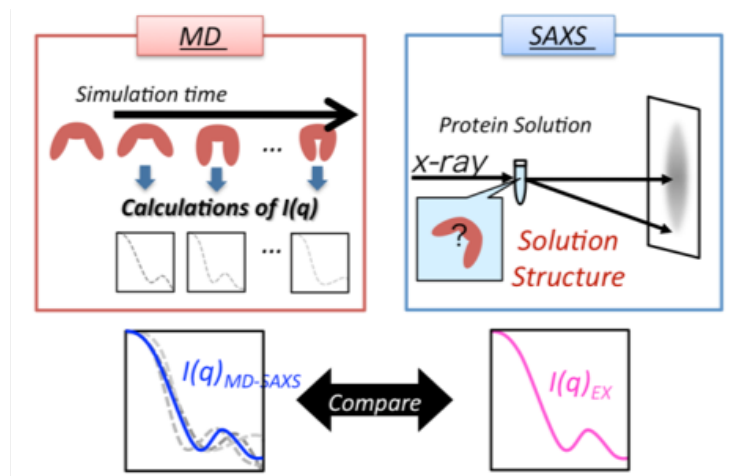
- 多剤排出トランスポーターAcrBのX線結晶構造をもとに、その機能発現時の動的な構造変化を分子動力学シミュレーションにより解析。
- ビタミンD受容体リガンド結合ドメインについて、X線小角散乱実験のデータをもとに、分子動力学シミュレーションにより、その溶液構造を解析。
- ヘテロクロマチン形成因子HP1クロモドメインのヒストンテイル認識について、非常に柔軟なディスオーダー領域を含む構造アンサンブルを、分子動力学シミュレーションにより解析。



多剤排出トランスポーターAcrBの
分子動力学シミュレーション

[4] 支援担当者の研究概要

スーパーコンピュータを活用し、タンパク質などの生体分子が機能を発現している際の動的構造について、分子動力学シミュレーションによる研究を行っている。X線結晶構造解析などで得られる構造はスナップショットであり、そのタンパク質が実際に働いている際には、動的に構造変化しつつ機能を果たしていることが多い。そのような動的な構造変化について、物理化学的相互作用に基づいた分子動力学シミュレーションを用いて解析している。さらには、分子動力学シミュレーションとX線小角散乱（SAXS）やNMRなどの構造生物学実験との連携により、溶液状態のタンパク質の構造解析を行っている。



分子動力学(MD)シミュレーションとX線小角散乱(SAXS)
連携による溶液構造解析

G5-1 タンパク質の高次構造情報を利用した創薬等研究加速に向けたバイオインフォマティクス研究

バイオインフォマティクス技術を駆使したタンパク質立体構造・相互作用・リガンド結合部位予測とこれに基づくタンパク質機能解明

[1] 支援担当者

所属		①産業技術総合研究所 人工知能研究センター
氏名		①富井 健太郎、土屋 裕子、山守 優、池田 修己
AMED 事業	ユニット／領域名 課題名	インシリコユニット タンパク質の高次構造情報を利用した創薬等研究加速に向けたバイオインフォマティクス研究
	代表機関 代表者	産業技術総合研究所 富井 健太郎
支援技術のキーワード		タンパク質類似性検索、タンパク質構造および相互作用部位予測、リガンド結合部位の類似性検索、リガンド結合部位予測、クライオ電顕データに基づく原子モデル構築

[2] 支援技術の概要

当グループでは、これまでに開発したタンパク質の立体構造予測や機能推定のための手法（1、2）およびリガンド結合部位の類似性比較のためのデータベース（3、4）等を利用したタンパク質立体構造・相互作用予測を実施し、これらを通してタンパク質機能解析の支援を行う。

1. プロファイル比較による高感度・高精度なタンパク質の配列類似性検索法（FORTE／DELTA-FORTE）
2. 遠縁タンパク質検索に適した高感度アミノ酸置換行列（MIQS）と、これを実装した配列比較法（Ssearch with MIQS）
3. タンパク質リガンド結合部位の類似性比較および検索データベース（PoSSuM）
4. 既存の低分子医薬品に特化した結合部位の類似性比較データベース（PoSSuMds）

上記に加え、下記支援を行う。

5. クライオ電子顕微鏡あるいはSAXSデータに基づくタンパク質全原子モデル構築
6. アポトーシス実行因子Caspase-3の基質タンパク質と切断部位の予測（ScreenCap3）
7. ミトコンドリアターゲティング配列と切断部位の予測（MitoFates）
8. タンパク質結晶における生物学的単位推定（Piac0）
9. 表面残基置換による結晶改善提案

[3] 支援技術の利用例

当グループでは下記に示す例のように、タンパク質の立体構造情報に基づく様々な解析を通してタンパク質機能の解明を目指す。

1. タンパク質の配列類似性検索法FORTE／DELTA-FORTEを用いた構造既知のタンパク質との網羅的な類似性検索と、これに基づくタンパク質のドメイン推定、タンパク質単体および複合体構造予測
2. リガンド結合部位データベースPoSSuMによる、既知タンパク質構造やFORTE等による予測モデル上の潜在的リガンド結合部位の同定、さらにその潜在的結合部位や、テンプレートであるリガンド結合タンパク質の機能およびリガンド情報等に基づくクエリータンパク質の機能推定
3. 分子動力学シミュレーションおよびドッキングシミュレーションから得られるタンパク質間相互作用の情報や、タンパク質表面物性等に基づく、タンパク質結合実験における変異アミノ酸の提案、およびタンパク質結合部位の予測
4. モデル生物以外のタンパク質の構造モデル構築と、モデルから得られる洞察に基づく、タンパク質の相互作用や機能発現の可能性の推定
5. （特に低解像度）クライオ電子顕微鏡あるいはSAXSデータに基づくタンパク質全原子モデル構築
6. 種々のバイオインフォマティクス技術による配列・構造・機能・進化等一連の解析を利用した研究推進
7. バイオインフォマティクス技術を用いたタンパク質結晶改善

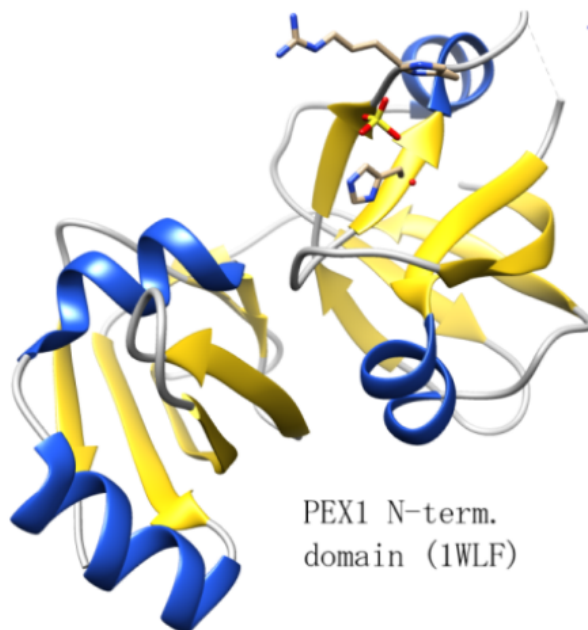
[4] 支援担当者の研究概要

タンパク質の配列情報や立体構造情報等に関する大規模データを利用した、立体構造および複合体構造予測法、タンパク質間相互作用（部位）予測法、リガンド結合部位予測法、cleavage site予測法等の開発と提供を行っている。近年、技術の発展に伴い、配列情報や立体構造情報、発現情報等は急増の一途をたどっている。こうした大規模データを効果的に取り扱い、そこから生物学的知識を発見・抽出するための計算科学的手法の開発を行う一方、開発手法をはじめとするバイオインフォマティクス技術を応用した、実際の生物学的課題の解決に役立てている。これまでに開発した高感度タンパク質類似性検索技術FORTE（図1）を用いて、従来の配列類似性検索手法では検出できなかったII型酵素ペルオキシソーム因子PEX1のN末端ドメイン（図2）や赤痢アメーバにユニークな酵素輸送受容体CPBFを構成するPPC様ドメイン等の同定および構造予測を行った。またクライオ電子顕微鏡データに基づいて、FORTEを使ったモデリングと分子動力学を使ったフレキシブルフィッティングによる構造解析を行い、ミトコンドリア外膜のトランスロケータタンパク質複合体が基質となるタンパク質を取り込む機構を明らかにする等の成果に貢献した（図3）。この他にも大規模データを利用したリガンド結合（候補）部位類似性検索のためのデータベース

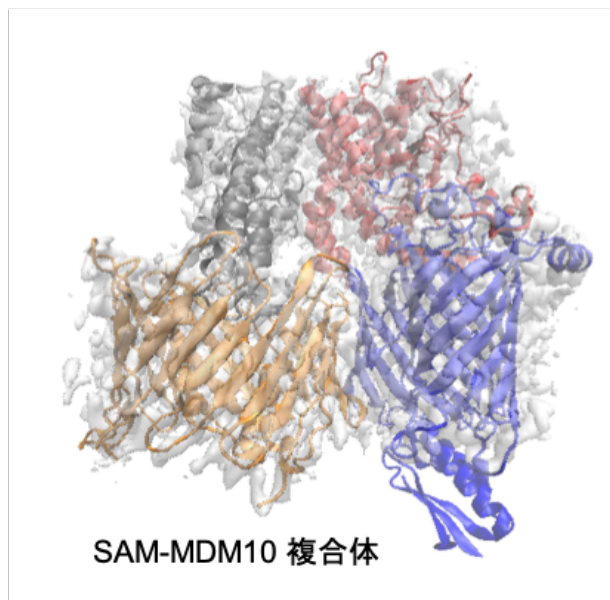
（<http://possum.cbrc.jp/PoSSuM/>、図4）、タンパク質-化合物間相互作用予測のための深層学習用データセット計算システム、抗体-抗原相互作用に特化したデータベース（パラトープ／エピトープ候補部位類似性検索）の構築を行っている。



(図1)



(図2)



(図3)

PoSSuM

Pocket Similarity Search using Multi-Sketches

Home Search K Search P Drug Search Dataset & Help About Us

About PoSSuM

Today, vast amounts of protein-small molecule binding sites are available in the Protein Data Bank (PDB). Comprehensive comparisons are computationally demanding, but will be useful for prediction of protein functions and drug discovery. We proposed a tremendously fast algorithm called "SketchSort" that makes it possible to enumerate similar pairs in a huge number of binding sites. We applied the proposed method to all-pair similarity searches for **9.7 million known and potential binding sites**, and discovered over 163 million similar pairs of binding sites.

Summary of Query

PDB ID	2G8A
Ref. Code	ADP
Chain ID	A
Residue No.	1-100
Protein Name	Structural 2-phosphoglycerate 3-phosphoglycerate hydrolase 1
UnProt ID	Q5G2A2
UnProt ID	Q5G2A2
EC No.	EC3.6.1.13
CATH code	3.40.30.300
SCOP code	3.4.2.1.2
Target DB	Known ligand binding sites
Sequence redundancy	All unorthodox
Cross similarity cut-off	0.10
Aligned length cut-off	7
Maximum number of hits	100
Annotations	EC:3.6.1.13 (Q5G2A2)

ADP

Search Result

Number of hits: 500

#	Type	PDB ID	Ref. Code	Chain ID	Residue No.	Cross. value	EC value	Aligned length	EC No.	CATH code	SCOP code	Target DB	Sequence redundancy	Cross similarity cut-off	Aligned length cut-off	Maximum number of hits	Annotations
1	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)
2	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)
3	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)
4	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)
5	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)
6	K	2G8A	ADP	A	1-100	0.100	EC3.6.1.13	7	EC3.6.1.13	3.40.30.300	3.4.2.1.2	Known	All unorthodox	0.10	7	100	EC:3.6.1.13 (Q5G2A2)

(図4)

G6-1 ナチュラルドラッグ（薬草由来）のネットワーク構造創薬（NSBDD）

生薬-代謝物DB（KNApSack family）をベースとして、天然生理活性物質からの創薬研究を総合的に支援

[1] 支援担当者

所属		①奈良先端科学技術大学院大学 情報科学研究科 ②長浜バイオ大学 バイオサイエンス学部 ③名古屋大学 情報文化学部
氏名		①金谷 重彦 ②白井 剛 ③太田 元規
AMED 事業	ユニット／領域名 課題名	インシリコユニット 生薬データベースの高度化と構造創薬への応用
	代表機関 代表者	奈良先端科学技術大学院大学 金谷 重彦
支援技術のキーワード		天然生理活性物質、リガンドスクリーニング、天然変性タンパク質、タンパク質複合体モデリング

[2] 支援技術の概要

- 生薬-代謝物DB（KNApSack family）による天然生理活性物質の探索・解析(図1)
- 天然変性タンパク質DB（IDEAL）によるタンパク質機能・相互作用解析
- リガンド構造比較法（COMPLIG）によるドラッグ（生理活性物質）構造比較
- リガンドスクリーニング法（VS-APPLE）によるドラッグとターゲットの複合体構造予測
- これらの技術を組み合わせたナチュラルドラッグ（薬草由来の天然物化合物）のネットワーク構造創薬の支援(図2)

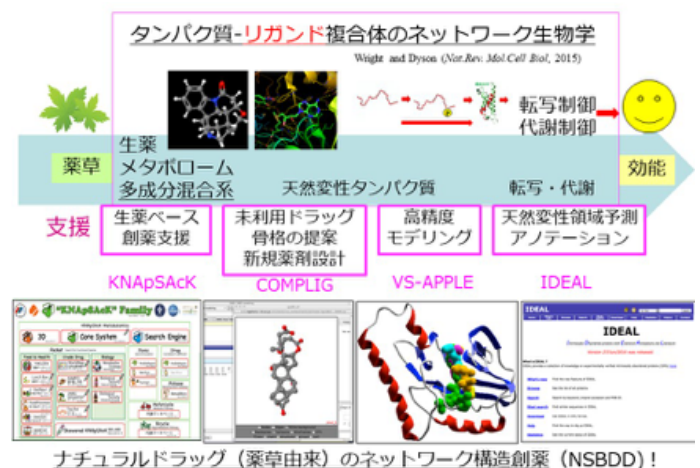
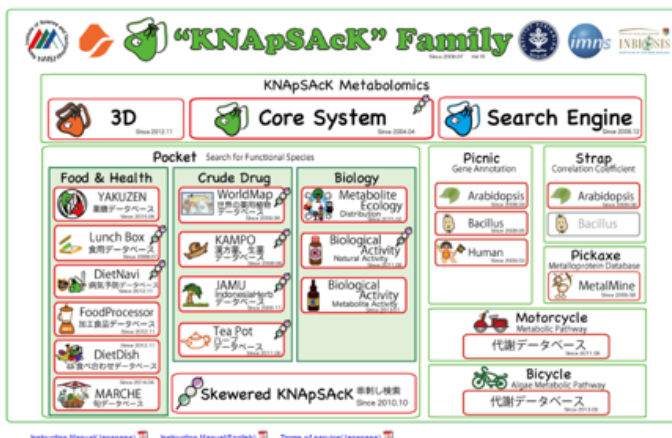


図2 支援の概要

[3] 支援技術の利用例

KNaPSAcKに収録された2072個の2次代謝物を、COMPLIGを用いて分子構造により分類し、140種の生理活性との相関を解析したところ、983の統計有意な構造-活性相関群が特定された (Mol.Inf,2014, 33, 790-801. Best Paper Award 2014受賞) (図3)

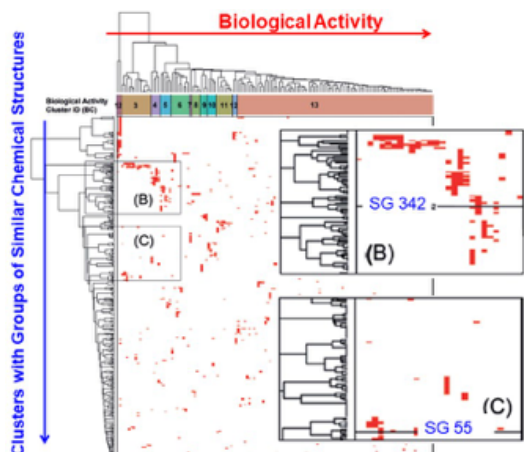


図3 二次代謝物の類似構造と活性情報との関係づけ

[4] 支援担当者の研究概要

- 11万組の生物種-天然物、1万組の生理活性天然物（生薬）-効能のデータを収録した生薬・メタボロームの世界標準データベース KNaPSAcK familyを開発し、系統的な天然生理活性物質の探索・解析を可能にした（植物生理学会2014論文賞）(図3)。
- タンパク質の天然変性領域（IDR）中にある相互作用部位の登録数世界一の天然変性タンパク質データベースIDEALを開発し、IDRの機能アノテーションを行った(図4)。
- 低分子立体構造のグラフマッチによる高速なフレキシブル構造比較法COMPLIGを開発し、KNaPSAcKに収録された生理活性天然物・承認薬・ヒト代謝物の構造分類を行った(図5)。
- リガンド重ね合わせを利用した独自のリガンドスクリーニング法VS-APPLEを開発し、ドラッグスクリーニングコンテストに参加し高い性能を示した（IT創薬グランプリ2016受賞）。

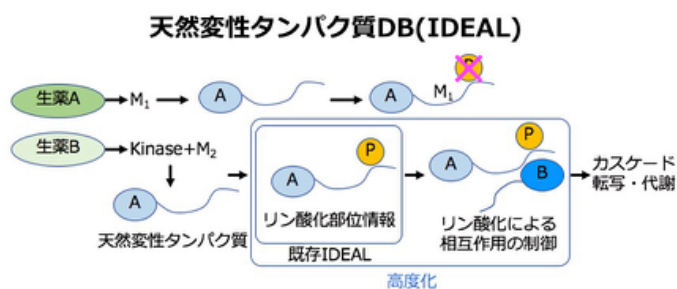


図4 翻訳後修飾による相互作用制御のデータベース化

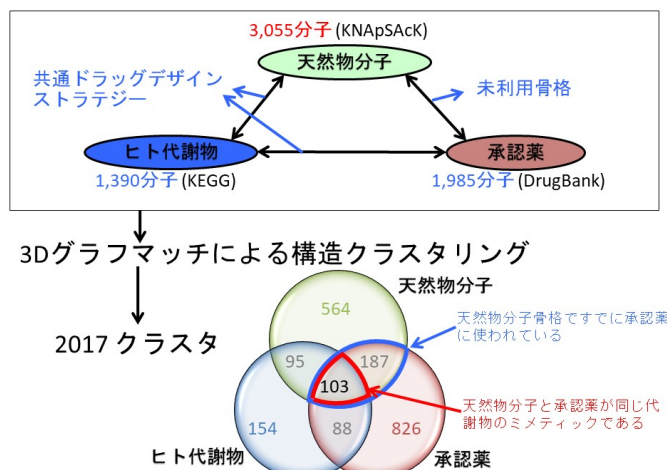


図5 天然物分子(生薬成分)・ヒト代謝物・承認薬の網羅的構造比較

G7-1 スーパーコンピュータTSUBAME3の計算機利用とインシリコスクリーニング支援

スーパーコンピュータ利用支援、天然変性領域予測支援、大規模PPI予測支援、分子動力学シミュレーションの実施・解析支援、機械学習に関する支援

[1] 支援担当者

所属		①東京工業大学 科学技術創成研究院 ②東京工業大学 情報理工学院
氏名		①関嶋 政和 ②秋山 泰、石田 貴士、大上 雅史
AMED 事業	ユニット／領域名 課題名	インシリコユニット 大規模分子シミュレーションによるインシリコスクリーニング支援と構造インフォマティクス技術の高度化
	代表機関 代表者	東京工業大学 関嶋 政和
支援技術のキーワード		スーパーコンピュータ、インシリコスクリーニング、蛋白質間相互作用解析、天然変性領域予測、機械学習

[2] 支援技術の概要

私たちは東工大のスーパーコンピュータTSUBAMEの利用支援及び、TSUBAMEを用いることで可能になる大規模分子シミュレーションと構造インフォマティクス技術によるインシリコスクリーニング支援を行っています。

東工大のスーパーコンピュータTSUBAMEは平成18年より運用中のクラスタ型スーパーコンピュータです。計算による科学技術研究の促進を目的として設置され、現在では東工大内のみに限らず学外の研究機関・民間企業の方などに幅広く使われております。平成18年に導入されたTSUBAME1.0 は2度の更新の後、平成22年にTSUBAME2.0に総入れ替えとなり、そして平成25年9月にはTSUBAME2.5へと大幅更新されました。平成29年8月に運用を開始したTSUBAME3.0は、Green500 List（省エネ性能の世界スパコンランキング）において世界1位を達成した上、AI分野で重要な半精度演算性能が47.2ペタフロップスと高く、人工知能分野における需要急増への対応が期待されています。TSUBAME3.0の倍精度の理論演算性能は12.15ペタフロップス（1秒間に12,150兆回の浮動小数点演算が可能）と、スーパーコンピュータ「京」を上回る世界最高レベルの性能となります。最新GPUの採用により、性能および電力効率の向上、ストレージの高速化および大容量化が実現し、計算ノードに搭載されるNVMe対応高速SSDの合算容量は1.08 PBと容量、速度ともに強化され、ビッグデータアプリケーションの処理速度を大幅に加速できます（以上、TSUBAMEを提供する東工大GSICのweb siteから要約）。我々は、このTSUBAME3.0の利用をインシリコユニットを中心に支援しています。

また、TSUBAMEを用いることで可能になる大規模分子シミュレーションを中心とした構造インフォマティクス技術によって、インシリコスクリーニング支援を行っています。具体的には、我々がこれまでに開発した高精度化合物スクリーニング法SIEVE-Scoreやファーマコフォアモデリングの経験、世界最高水準のタンパク質天然変性領域予測システムPrDOS-CNF、大規模並列計算によるタンパク質間相互作用予測MEGADOCKを最大限に活かし、（1）インシリコスクリーニングによるタンパク質と低分子化合物との相互作用様式の推定支援、（2）タンパク質の天然変性領域予測支援、（3）タンパク質間相互作用解析手法に基づく結合候補タンパク質探索と相互作用様式予測支援の3項目の支援を行っています



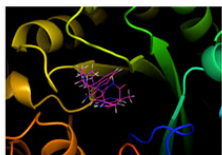
[3] 支援技術の利用例

TSUBAME3.0上において分子シミュレーションや機械学習に必要な計算機資源、及び各種ソフトウェア（AMBER, Gaussian, GAMESS, LAMMPS, Caffe）を提供しています。

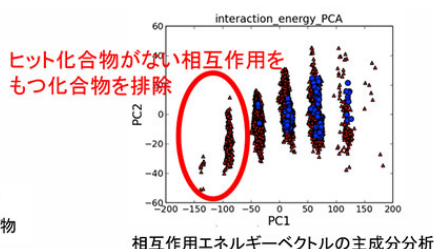
また、TSUBAMEを用いたin silico計算により、熱揺らぎを考慮したドッキングシミュレーションを実施し、バイオアッセイを行うことで阻害活性がある化合物の獲得に成功しています。

[4] 支援担当者の研究概要

代表者の関嶋は、in silicoのシミュレーションとin vitroの生化学実験を組み合わせた効率的な創薬手法を開発し、20のヒット化合物と、2つの抗原原虫活性がありかつ細胞毒性が低い化合物を獲得し、さらに2化合物に対して初めてスペルミジン合成酵素とプロレシン結合部位に化合物が結合した共結晶の立体構造解析(PDBID: 5B1S)に成功しました（生命医薬情報学連合大会 2014 研究奨励賞, R. Yoshino et al., Sci.Rep., 2017）。インシリコスクリーニング、タンパク質の天然変性領域予測、タンパク質間相互作用解析手法に基づく結合候補タンパク質探索と相互作用様式予測等に関する支援を実施しています。



ドッキングのスコアが高いが正しい相互作用をもたない化合物



また、スーパーコンピュータと機械学習を用いた創薬手法の研究を行っており、蛋白質・リガンド間の相互作用をエネルギーベクトルとして機械学習した手法(N. Yasuo and M. Sekijima, JCIIM, 2019)においては、Journal of Chemical Information and Modeling誌のMachine Learning Special IssueにおいてSupplementary Cover Artに選ばれた。また、ディープラーニングを用いたヒット化合物予測においては、特許出願も行っている。

Table of Contents

March 25, 2019
Volume 59, Issue 3
Pages 945-1282

About the Cover:
The construction process of an interaction energy vector. The interaction energy vector consists of hydrogen bond, van der Waals, and Coulomb interaction energy terms for each residue in the grid of protein-ligand docking, and is used as the feature vector for machine-learning-based virtual screening.
[View the article](#)

In this issue:

- > MACHINE LEARNING IN DRUG DISCOVERY: EDITORIAL
- > PERSPECTIVES
- > CHEMICAL INFORMATION
- > COMPUTATIONAL CHEMISTRY
- > COMPUTATIONAL BIOCHEMISTRY
- > PHARMACEUTICAL MODELING
- > BIOINFORMATICS
- > MATERIALS

[Download Hi-Res Cover](#)

ACS AuthorChoice
More about this program

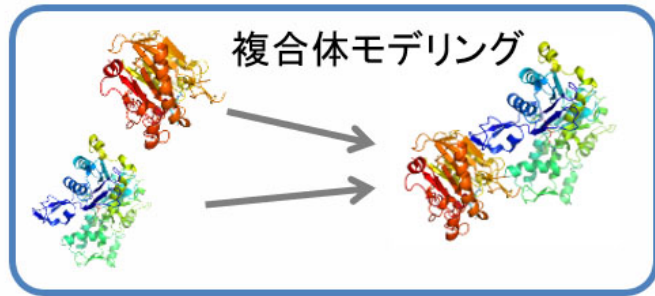
ACS Editors' Choice
More about this program

<https://pubs.acs.org/toc/jcisd8/59/3> より

天然変性領域は結晶化の妨げとなるため、構造生物学研究者にとって悩みのタネとなっている。分担者の石田が開発した天然変性領域予測システムPrDOSによりタンパク質のアミノ酸配列から予測を行う。本手法は最新の天然変性領域予測に関する国際コンペティション(CASP10 disorder)にて、全28グループ中、最高精度を達成している。

...EPEKLEEANKDDPERE... ➡ ...EPEKLEEANKDDPERE...

創薬標的としてタンパク質間相互作用が注目を集めているが、相互作用情報の蓄積が不十分なことと複合体構造の決定の難しさが研究上の障壁になっている。結合タンパク質探索と複合体モデリングを実施するソフトウェア分担者の秋山、大上らによって開発されたMEGADOCKにより、標的タンパク質に対して結合するタンパク質の探索と、複合体モデリングを実施している。



G8-1 インシリコスクリーニング及びFMO法による相互作用解析

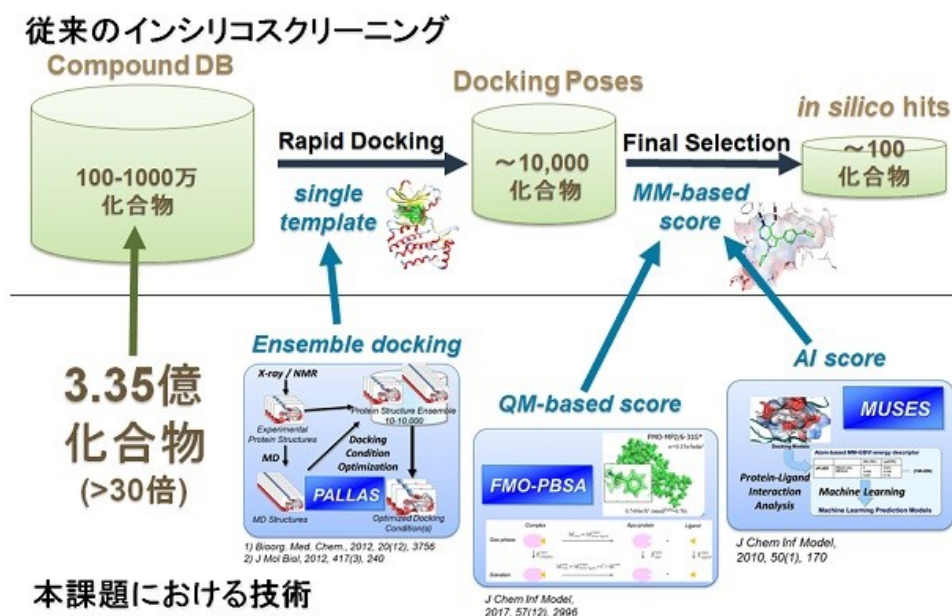
創薬や生命機能解明を目指した化合物のインシリコスクリーニング、及びFMO法による相互作用解析

[1] 支援担当者

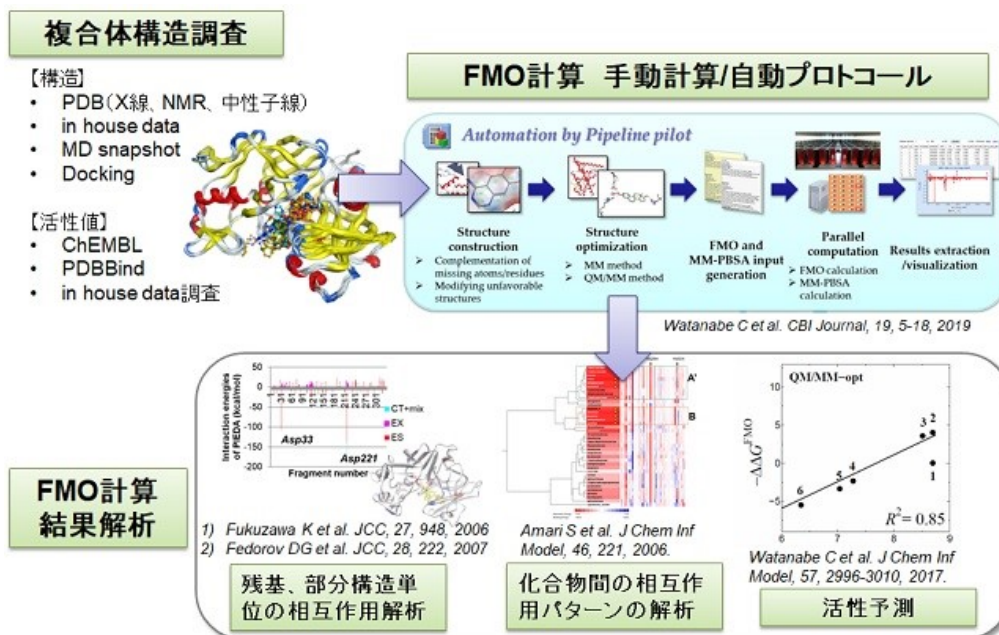
所属		①理化学研究所 生命機能科学研究センター ②星薬科大学 薬学部
氏名		①本間 光貴 ②福澤 薫
AMED 事業	ユニット／領域名 課題名	インシリコユニット 構造インフォマティクスとFMO計算を融合したインシリコスクリーニング支援研究
	代表機関 代表者	理化学研究所 本間 光貴
支援技術のキーワード		インシリコスクリーニング、FMO法

[2] 支援技術の概要

- PALLAS, MUSESによるタンパク質-リガンド間ドッキングに基づいたインシリコスクリーニング支援
- LAILAPSによるリガンド情報に基づいた多面的な類似構造検索及び機械学習予測



- 有望なヒット化合物の活性、及びADMET（体内動態・毒性）プロファイルを改善するための設計支援
- FMO計算による結晶構造の相互作用解析
- FMO計算による結晶構造の精密化



[3] 支援技術の利用例

以下を代表例とする25個を超えるターゲットに対するインシリコスクリーニングとヒットした化合物の最適化設計

- HCK（急性骨髄性白血病、前臨床段階）
- AdipoR（糖尿病、特許取得段階）
- ALK2（希少疾患FOP、特許取得段階）
- DOCK1/2（がん、自己免疫疾患、特許取得段階）

以下を代表例とするFMO計算による活性予測、及び1500を超える複合体の相互作用解析

- エストロゲン受容体リガンドの定量的な活性予測
- Pim1阻害剤のFMO+MM-PBSAによる定量的な活性予測
- レニン阻害剤の定量的な活性予測

[4] 支援担当者の研究概要

理化学研究所の本間研究室では、2007年からターゲットタンパク研究プログラム、創薬等PFなどのプロジェクトにおいて、実用的なインシリコスクリーニング手法の開発と、開発した技術を応用した創薬研究を行っている。

開発している技術の特徴としては、PDBにあるタンパク質構造データとChEMBL等のリガンドの構造活性相関データに基づいて、シミュレーションとインフォマティクスの両面を活用した実用的なものであり、ドッキング条件の半自動最適化システムPALLAS, 相互作用記述子による機械学習予測システムMUSES, 多面的なりガンド検索システムLAILAPSなどがある。

25を超えるターゲットに対してインシリコスクリーニングを実施しており、創薬シードやツール化合物を発見した。その中でいくつかの有望な創薬シードに関しては、特許取得段階や前臨床段階まで化合物を最適化することに成功している。

今回の創薬等LSでは、上記のインシリコスクリーニングの基盤に加えて、タンパク質のような大きな系の量子化学計算を可能にするフラグメント分子軌道法（FMO法）を創薬に活用する基盤の構築を進めている。すでに実用化しているFMO法による相互作用解析に加えて、FMO-PBSAなどの新規な結合親和性予測手法の開発や、世界初のタンパク質の量子化学計算値データベース（FMOデータベース）の公開を行った。

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



3/21 9:30より、薬学会にて公開記念シンポジウムを開催！
100人程度の参加者が来訪しました。

2019年
2月に公開

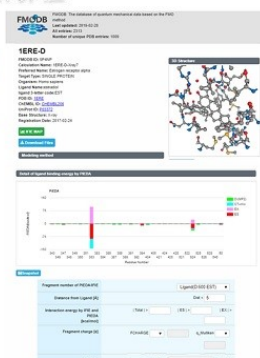


現時点で2,300個※の
FMO計算結果が登録
済み
20,000個のMDスナッ
プショットに対するFMO計
算も登録中

創薬分子設計

FMO
力場AI

分子認識メカニズムの
解明



※ユニークなPDB ID数は約1000種類

G9-1 分子モデリング及びシミュレーションを活用したインシリコ創薬支援

タンパク質構造に基づくインシリコスクリーニング、リガンド設計に向けた分子動力学シミュレーション、タンパク質構造モデリング

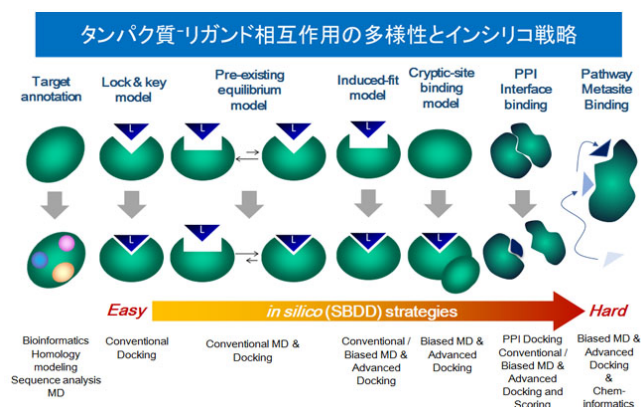
[1] 支援担当者

所属	①筑波大学 医学医療系生命医科学域	
氏名	①広川 貴次	
AMED 事業	ユニット／領域名 課題名	インシリコユニット 分子モデリングおよびシミュレーションを活用したインシリコ創薬支援
	代表機関 代表者	筑波大学 広川 貴次
支援技術のキーワード		インシリコスクリーニング、MDシミュレーション、ドッキングシミュレーション、タンパク質立体構造予測

[2] 支援技術の概要

複合要素技術を必要とするインシリコ創薬支援

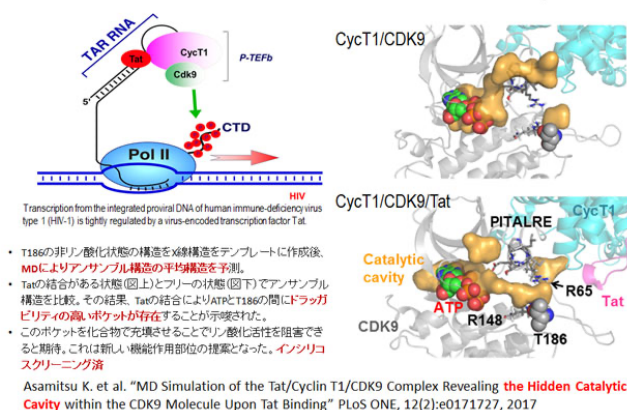
- ・ ポスト（プレ）ドッキングを必要とするような、難易度の高いドッキング計算、複合体構造モデリング
- ・ Gタンパク質共役型受容体他、創薬標的タンパク質に特化した分子モデリング
- ・ 分子シミュレーションとケモインフォマティクスを融合したタンパク質-リガンド相互作用解析
- ・ タンパク質の動的性質を考慮したインシリコスクリーニング
- ・ 様々な分子モデリング、シミュレーションの要素技術を統合した高度なインシリコ創薬



[3] 支援技術の利用例

- 活性天然物、ペプチドとタンパク質の複合体予測及び相互作用解析による作用機序の理解
- タンパク質の動的性質を考慮した活性部位予測及びヒット化合物からのリード化合物分子設計指針の提供
- Gタンパク質共役受容体のホモロジーモデリング、活性部位予測、リガンド結合予測及び作用機序予測
- 活性ペプチドの低分子化設計
- タンパク質-RNA複合体予測とアプタマー創薬支援
- タンパク質の動的性質を考慮した、構造活性相関解析

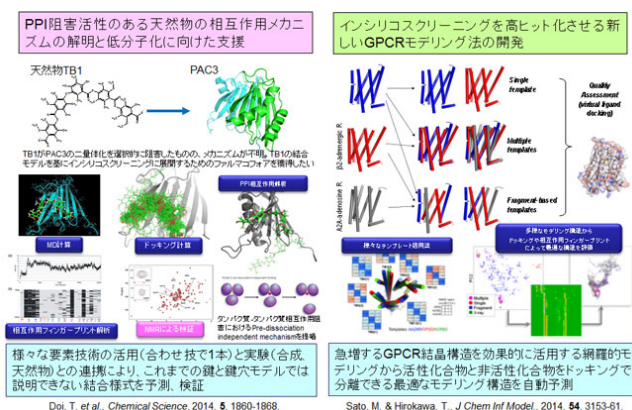
タンパク質間相互作用によるHidden Cavity標的



[4] 支援担当者の研究概要

近年、構造生物学データを起点とした創薬支援研究が注目されているが、構造生物学データの中には、特定の条件や環境に依存した構造情報であるために、そのままでは創薬へ適用が難しいものがある。分子モデリングや分子シミュレーションは、このような問題を補完できる技術であり、構造生物データと融合させることで、より高度な創薬支援研究が実現可能となる。そのためには、創薬標的タンパク質を中心に、ホモロジーモデリング、ドッキング計算（タンパク質-タンパク質、タンパク質-低分子、タンパク質-核酸、核酸-低分子）、分子動力学計算等の要素技術に基づいた分子モデリング及びシミュレーションを統合的に活用した実用性の高いインシリコ創薬の支援研究と高度化研究が必要だと考えている。

我々は、インシリコスクリーニングに特化したGPCRモデリングツール（Sato M. and Hirokawa T, J. Chem. Inf. Model., 2014）や要素技術を統合化する戦略で、創薬標的となるタンパク質-タンパク質、タンパク質-低分子、タンパク質-核酸、核酸-低分子の分子モデリングに取り組んでいる。代表的な成果として、2014年に発表した、タンパク質-タンパク質相互作用阻害天然物の作用メカニズムを分子モデリング、合成、NMR測定との連携により解明した成果が挙げられる（Doi T et al., Chemical Science, 2014）。この研究課題では、これまでの鍵と鍵穴モデルでは説明できない天然物特有の結合様式を様々な要素技術を活用することで予測が可能となり、実験（合成、天然物）との連携により、予測モデルが正しいことが検証された。



また、2015年の実験情報と統合化インシリコ戦略による活性化Parkin構造の予測（Yamano K, J. Biol. Chem., 2015）では、後に結晶構造にきわめて類似した予測構造を提案したことが明らかになり、本成果は、2015年、JBC Best paper of the weekに選定され、国際的な関心の高さが示された。

最近の成果として、筑波大学医学医療系の野口雅之研究室との共同研究において初期肺腺癌における標的タンパク質間相互作用阻害薬探索において活性化化合物の同定に成功した（Shiba-Ishii A et al., Clin Cancer Res, 2019）。実施内容としては、標的タンパク質であるSFNとSKP1それぞれの立体構造から、タンパク質-タンパク質ドッキング計算でSFN-SKP1複合体を予測した。次に、SFNタンパク質内のSKP1結合部位周辺に存在するポケットのドラッグabilityをコンピュータで評価し、低分子結合が可能であると判断した。この結果に基づき、SKP1と競合する医薬品候補となりうる小分子化合物を、既存薬データベースDrugBankからインシリコスクリーニングし、候補化合物の同定を行った。候補化合物は、共同研究先の筑波大学で様々な生化学実験で評価が行われ、その結果、2つの既存薬において特に強い阻害活性を持ち、さらにヌードマウスを用いた腫瘍抑制効果を有することが確認できた。現在、ドラッグリポジショニングとしての可能性を検討している。様々な構造データベースとインシリコ技術を統合的に活用した創薬支援の成果といえる。

タンパク質間相互作用阻害剤の探索

