



日本学術振興会・産学協力委員会

R022 量子構造生物学委員会

第4回 研究会 資料

2021年12月23日（木）

13:30～15:45

オンライン開催

**R022 量子構造生物学委員会
第4回研究会 プログラム**

2021年12月23日（木）
オンライン開催

13:30 – 13:35

はじめに

委員長・三木邦夫（京都大学）

講 演

司会：副委員長・中川敦史（大阪大学）

13:35 – 14:35 （講演 1）

計算生物学の創薬応用

医薬基盤・健康・栄養研究所／大阪大学・蛋白質研究所
水 口 賢 司

14:35 – 14:45 休 憩

14:45 – 15:45 （講演 2）

サイトカイン受容体を活性化する海洋天然物由来タンパク質の構造的考察

東北大学・大学院生命科学研究科
田 中 良 和

計算生物学の創薬応用

Applications of Computational Biology to Drug Discovery

国立研究開発法人 医薬基盤・健康・栄養研究所 AI 健康・医薬研究センター／

大阪大学 蛋白質研究所

水 口 賢 司

Kenji Mizuguchi

Artificial Intelligence Center for Health and Biomedical Research,
National Institutes of Biomedical Innovation, Health and Nutrition /

Institute for Protein Research, Osaka University

<https://mizuguchilab.org/>

一般に、コンピュータを用いたモデリングを構成する基盤技術として、1)データ統合とデータベース構築、2)データ駆動型モデリング、3)原理ベースのモデリング、の3つが重要な位置を占める（図1）[1]。このうち、2番目のデータ駆動型モデリングは人工知能（Artificial Intelligence, AI）技術など、現在のコンピュータ解析の中心として様々な分野で活用が進んでいる。本来のAIは人間が行う高度な判断をコンピュータで実現するための広範な技術、学問分野を指すが、現代のAIの中心は、機械学習と呼ばれる「データから規則や知識を抽出する技術・手法」であり、データ駆動型モデリングの中核を成すと言える。

計算システム生物学

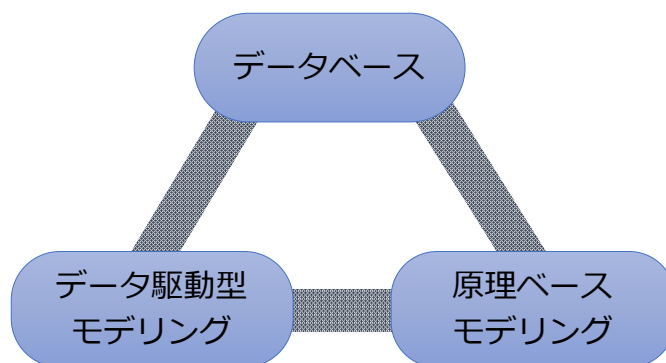


図1. コンピュータモデリングの基盤技術

例えば、生命科学の分野においては、タンパク質のアミノ酸配列のみから立体構造を高精度に予測できるAIプログラムとして、DeepMind社のAlphaFold2[2]が注目を集めている。また、化合物の化学構造のみから各種活性を予測するプロ

グラムも創薬研究の分野で一般に用いられている。

一方、ここで原理ベースのモデリングと呼ぶのは、個別のデータから規則を抽出するのではなく、より一般的な原理に基づいてモデルを構築する手法を指す。このカテゴリーには、物理学の法則に基づいてタンパク質など生体分子の動的構造を計算する分子動力学（Molecular Dynamics, MD）シミュレーションから、薬物動態の数値モデルまで、様々なアプローチが存在する。個別データが十分に得られない場合でも利用可能という強みがある一方、法則自身の適用範囲や近似の妥当性、MD 計算の場合一般に大きな計算資源を必要とするなどの制限もあり、データ駆動型モデリングとの相補的な利用により、有用性が発揮できると考えられる。

さらに、データ統合とデータベース構築が、データ駆動型及び原理ベースのモデリングのいずれに対しても基礎となり、これら 3 つの要素の中で最も重要な位置にあることを強調したい。一般に、コンピュータフレンドリーな形に整理されたデータをどれだけ利用できるかが、AI 開発を含むコンピュータモデリングの成否に大きな影響を与える。例えば、AlphaFold2 の成功は、Protein Data Bank (PDB) というタンパク質立体構造の公共レポジトリなくしては語り得ない。しかし、創薬研究を含む生命科学分野においては、十分に整理されていないデータが散在する現状があり、コンピュータ解析の適用を阻む原因になっている。

創薬研究への計算生物学、バイオインフォマティクス技術の応用においても、様々な問題について、異なる種類のモデリングアプローチとデータ統合との間の相互連間が見出される。多くの場合に創薬研究の出発点であるターゲット探索においては、我々の開発した創薬支援統合データウェアハウス TargetMine (<http://targetmine.mizuguchilab.org>; [3]) の活用により、ターゲット候補の絞り込みや妥当性の評価を行うことができる。また、AlphaFold2 の創薬応用はこれから多数現れると期待されるが、我々は以前から、進化の原理に基づく構造予測プログラム FUGUE ([4]) や PDB の複合体構造を学習データにした機械学習モデルによるタンパク質間相互作用部位予測プログラム PSIVER([5])を用

いて、新規ターゲット候補の評価に役立ててきた。また、ターゲットを設定した後のスクリーニング段階において、機械学習による活性予測は広く行われているものの、ペプチドや糖鎖など、実験データの蓄積が十分でない領域については、インフォマティクス解析に加えて MD シミュレーションを実行することで、新たな知見を獲得することができる [6]。

さらに、前臨床試験に向けては、薬物動態や毒性など、医薬品として満足すべき性質を適切に評価する必要がある。アカデミア創薬においては、企業が行っているように様々な薬物動態評価系や毒性評価系を常に稼働させておき、必要なデータを取得することは難しい。そこで我々は、創薬研究初期における薬物動態、心毒性、肝毒性の予測を目的としたインシリコの統合解析プラットフォーム（データベース、モデリング機能、予測機能から成るシステム）の構築を推進してきた(<https://www.id3inst.org/>)。薬物動態モデリングにおいては、必要な実験結果は公共のデータソースからある程度入手可能だが、実験条件が統一されていない、化合物の記述が不正確などの理由により、高精度の機械学習モデルの作成が進んでこなかった。我々は、マニュアルキュレーションによって質を高めた訓練データセットを利用することで、薬物動態パラメータを予測する機械学習モデルの精度を向上させられることを示した [7]。また薬物動態

態を幅広くカバーする合計 17 個の予測モデルを構築し、特に血液脳関門で重要な働きをするトランスポーターである P-glycoprotein の化合物輸送能を考慮することで、中枢移行性予測の精度向上に成功した (図 2) [8]。さらに、製薬企業 7 社と

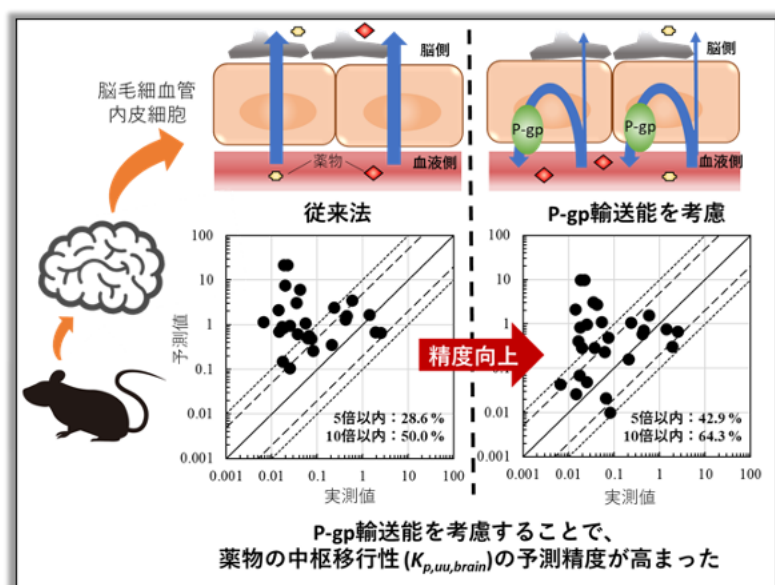


図 2. P-glycoprotein 輸送能を考慮した中枢移行性予測モデル

の企業連携を構築し[9]、これまで共有が極めて困難だった社内データを公共データに加えることで、データを質・量ともに拡充できることを示した。

このような具体例から、データ統合における共通の課題が浮かび上がってくる。これらの解決には、地道なマニュアルキュレーションの試みとともに、データを共有する仕組み作りや、データ産生のための AI の活用が必要と考えられる。

-
1. 夏目やよい、水口賢司、計算システム生物学による創薬：分子、構造からネットワークへ。日本薬理学雑誌, 149: 91-95 (2017).
 2. Tunyasuvunakool K, *et al.*, Highly accurate protein structure prediction for the human proteome. *Nature*, 596: 590-596 (2021).
 3. Chen YA, Tripathi LP, Fujiwara T, Kameyama T, Itoh MN, Mizuguchi K. The TargetMine Data Warehouse: Enhancement and Updates. *Front Genetics*, 10: 934 (2019).
 4. Shi J, Blundell TL, Mizuguchi K. FUGUE: sequence-structure homology recognition using environment- specific substitution tables and structure-dependent gap penalties. *J. Mol. Biol.*, 310: 243-57 (2001).
 5. Murakami, Y. and K. Mizuguchi, Applying the Naive Bayes classifier with kernel density estimation to the prediction of protein-protein interaction sites. *Bioinformatics*, 26: 1841-1848 (2010).
 6. Re S, Mizuguchi K. Glycan Cluster Shielding and Antibody Epitopes on Lassa Virus Envelop Protein. *J. Phys. Chem. B* 125: 2089-2097 (2021).
 7. Esaki T, Watanabe R, Kawashima H, Ohashi R, Natsume-Kitatani Y, Nagao C, Mizuguchi K. Data Curation can Improve the Prediction Accuracy of Metabolic Intrinsic Clearance. *Mol. Inform.*, 38: 1800086 (2019).
 8. Watanabe R., Esaki T., Ohashi R., Kuroda M., Kawashima H., Komura H., Natsume-Kitatani Y., Mizuguchi K., Development of an In Silico Prediction Model for P-glycoprotein Efflux Potential in Brain Capillary Endothelial Cells toward the Prediction of Brain Penetration. *J. Med. Chem.*, 64: 2725-2738 (2021).
 9. Komura H., Watanabe R., Kawashima H., Ohashi R., Kuroda M., Sato T., Honma T., Mizuguchi K., A public-private partnership to enrich the development of in silico predictive models for pharmacokinetic and cardiotoxic properties. *Drug Discovery Today*, 26: 1275-1283 (2021).

サイトカイン受容体を活性化する
海洋天然物由来タンパク質の構造的考察

Structural insights into proteins from marine natural products
that activate cytokine receptors

東北大学大学院生命科学研究科

田 中 良 和

Yoshikazu Tanaka

Graduate School of Life Sciences, Tohoku University

新たな機能の生理活性物質を探索する際、有機合成できない複雑な構造の化合物を含む生物由来の分子ライブラリーは有力な探索源となり得る。中でも、ホヤや海綿などの海洋生物からの抽出物は生理活性を有する分子の宝庫であり、事実、抗がん剤をはじめとしたさまざまな医薬品が海洋生物由来ライブラリーから取得されている [1,2]。本研究では、海綿由来の分子ライブラリーから発見された、トロンボポエチン (TPO) 様の生理活性を有するタンパク質分子 ThC [3]に焦点を当てて研究を行なった [4]。

TPO は、血球前駆細胞に存在する TPO 受容体に結合し、血小板への分化を誘導するサイトカインである。先行研究にて、TPO 受容体に相互作用し、TPO と同様の生理活性を有するタンパク質 ThC が海綿抽出物から発見された [3]。我々は、ThC による TPO 受容体の活性化メカニズムを明らかにすることを目指し、種々の生化学的および細胞生物学的解析を行った [4]。はじめに、エドマン分解、質量分析および X 線結晶構造解析により、ThC のアミノ酸配列を決定した。明らかになった配列には TPO との相同性は認められず、したがって、ThC が TPO とは異なる機構により TPO 受容体を活性化していることが示された。一方、ThC は既知の複数のタンパク質と非常に類似した構造を有しており、これらの類似タンパク質の分子特性を手掛かりに、ThC がカルシウム依存性の糖結合活性を持つホモ二量体レクチンであることを明らかにした。等温滴定型熱量測定により、ThC はリガンド分子と結合解離定数 K_d が数 μM 程度の相互作用

用をすることが明らかになった。さらに、ThC とリガンドの複合体の X 線結晶構造解析も実施し、リガンドの結合様式についての知見を得た（図）。TPO 受容体を発現させた哺乳類細胞を用いた実験からは、ThC が結合する糖鎖修飾部位が明らかになり、また、ThC によるシグナル伝達は受容体の細胞内部への移行を誘発せず、長時間継続することがわかった。ThC と TPO の間に強い相乗効果があることから、ThC は細胞表面の糖鎖に結合して受容体を強制的に二量体化することでシグナル伝達を誘導することが示唆された。これらの結果から、トロンボポエチン受容体を活性化するさまざまな経路が存在することが示唆された。

糖鎖修飾を介したサイトカインの活性化については、カルレティキュリン（糖タンパク質のフォールディングを介助するタンパク質）が変異すると糖鎖修飾を介して TPO 受容体を活性化するようになり、これにより骨髄増殖腫瘍が引き起こされることが報告されている。これまで、細胞外部からのアゴニストが存在しなかったため、そのメカニズムを明らかにすることができなかった。本研究により、ThC は TPO 受容体を活性化するアゴニストとして機能することがわかった。今後、ThC をアゴニストとして用いることで、サイトカイン受容体の活性化機構の理解が一層深まることが期待される。

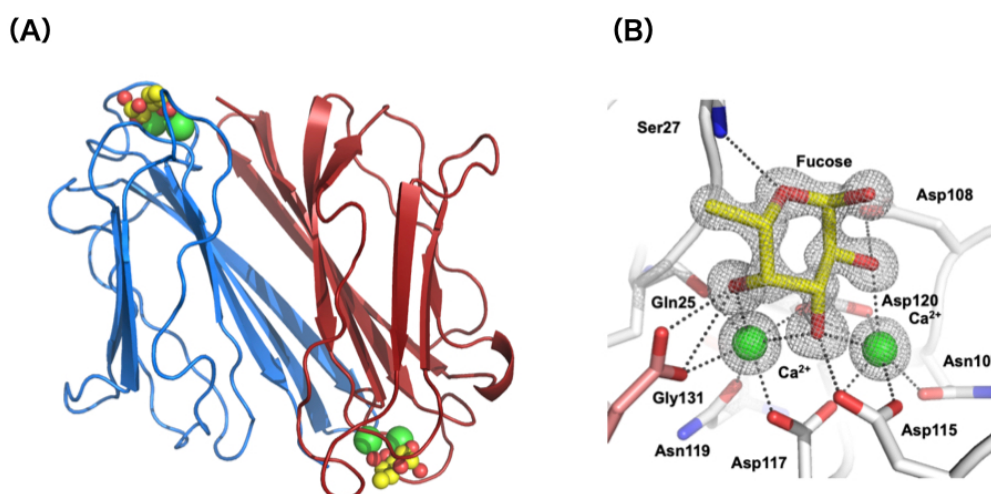


図. ThC とリガンド（フコース）との複合体の結晶構造
(A) 全体構造、(B) リガンド結合部位の構造

-
1. Sakai, R., Jares-Erijman, E., Manzanares, I., Elipe, M. V. S., and Rinehart, K. L. Ecteinasidins: Putative Biosynthetic Precursors and Absolute Stereochemistry. *J. Am. Chem. Soc.*, 118, 9017–9023 (1996).
 2. Sakai, R., Rinehart, K. L., Guan, Y., and Wang, A. H. Additional antitumor ecteinasidins from a Caribbean tunicate: crystal structures and activities in vivo. *Proc. Natl. Acad. Sci. USA*, 89, 11456-11460 (1992).
 3. Watari, H., Nakajima, H., Atsuumi, W., Nakamura, T., Nanya, T., Ise, Y., Sakai, R. A novel sponge-derived protein thrombocortecin is a new agonist for thrombopoietin receptor. *Comp. Biochem. and Physiol. C: Toxicol. Pharmacol.*, 221, 82-88 (2019).
 4. Watari, H., Kageyama, H., Masubuchi, N., Nakajima, H., Onodera, K., Focia, P. J., Oshiro, T., Matsui, T., Koder, Y., Ogawa, T., Yokoyama, T., Hirayama, M., Hori, K., Freymann, D. M., Komatsu, N., Araki, M., Tanaka, Y., Sakai, R. Hidden pathway for cytokine receptor activation: Structural insights into a marine sponge-derived lectin that activates the thrombopoietin receptor via recognition of the fucose moiety. *BioRxiv*, doi: <https://doi.org/10.1101/2021.10.29.466502>