



構造活性相関部会・ニュースレター <1 April, 2023>

SAR News ***No.44***

「目次」

//// Cutting Edge ////

深層学習と木探索に基づく de novo 分子設計

横浜市立大学 石田 祥一、吉澤 竜哉、寺山 慧 … 1

MOE REINVENT I/F の紹介

株式会社モルシス 木村 嘉朗 … 8

//// SAR Presentation Award ////

受賞コメントおよび講演要旨 … 16

//// Activities ////

<報告>

第 50 回構造活性相関シンポジウム 開催報告

名城大学 小田 彰史 … 37

<会告>

構造活性フォーラム 2023 会告

株式会社アグロデザイン・スタジオ 西ヶ谷 有輝 … 40

編集後記 … 41

//// Cutting Edge ////

深層学習と木探索に基づく de novo 分子設計

横浜市立大学 石田祥一、吉澤竜哉、寺山慧

1. はじめに

「望みの機能を持つ化合物が知りたい」——分子を扱っている研究者であれば誰もが望んでいることではないだろうか。所望の性質を持つ分子をゼロから(de novo)考えようとしたときに、その候補となる化合物の集合（化合物空間）の大きさは低分子有機化合物のみであっても10の60乗を超えると推定されている[1]。世界最大級の化合物データベースであるCASデータベースに収載の化合物数が10の8乗オーダーであること[2]を考えると、闇雲に望みの化合物を探しても見つけることがほぼ不可能であることは想像に難くないだろう。近年ではこの膨大な化合物空間から望みの化合物を探し出すために、深層学習を用いた様々な分子設計手法が提案されてきている[3-12]。これらの手法は「望みの機能」、例えばある標的タンパク質への活性等、を指定するとその機能を持つ「分子構造」を出力する「逆問題」を解くように設計されている[13]。本稿では、深層学習と木探索を組み合わせた de novo 分子設計手法である ChemTS[14]を取り上げ、阻害活性予測モデルとファーマコフォアモデルがどのように de novo 分子設計に活用できるかについて紹介していく。

2. de novo 分子設計手法 ChemTS とは

ChemTS[14]とは2017年にTsudaらが開発した de novo 分子設計手法であり、深層学習手法である再帰型ニューラルネットワーク (Recurrent Neural Network: RNN) [15]と強化学習の理論基盤を持った探索アルゴリズムであるモンテカルロ木探索 (Monte Carlo Tree Search: MCTS) [16]で構成されている。MCTSは効率的な化合物の探索のために用いられており、図1で示した選択・拡張・ロールアウト・バックプロパゲーションの4ステップを繰り返すことによって探索が進行していく。ChemTSでは分子はsimplified molecular input line entry system (SMILES)文字列[17]で表現されており、探索木の各ノードにはSMILES文字列から切り出されたトークン1つ(C, [C@H], =等)が割り当てられる。選択ステップでは現時点で探索すべきノード(葉ノード)についてツリー方策を元に選択する。拡張ステップでは選択された葉ノードから未探索のノードを追加する。ロールアウトステップでは追加された未探索ノードから完全なSMILES文字列になるようシミュレーションを実施し、得られた分子がどの程度所望の性質を持っているかを計算し報酬を求める。最後にバックプロパゲーションステップにて、探索木で辿ってきた各ノードの報酬と訪問回数を更新する。このサイクルは毎回探索木の根ノードから開始し、指定時間や設計分子数などの指定した閾値に達するまで実行される。

Monte-Carlo tree search (MCTS) の概要

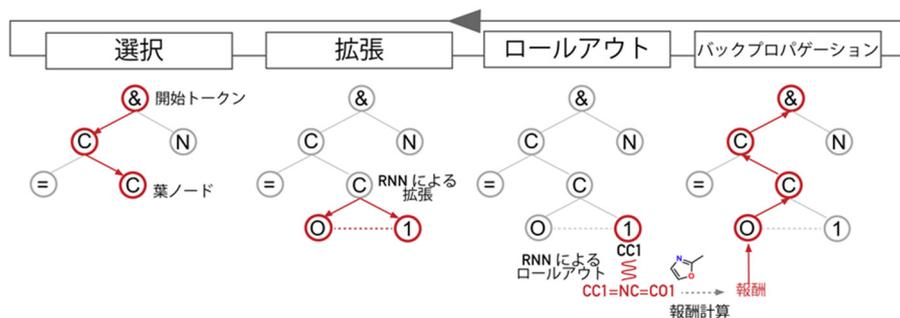


図1. 分子生成を行う Monte-Carlo tree search (MCTS)の概要。文献[25]より改変して引用。

ここで MCTS を分子設計に応用する際に課題になるのが、拡張ステップで「どのように未探索のノードを追加するか」とロールアウトステップで「どのように不完全な SMILES 文字列から完全な SMILES 文字列を生成するか」である。この課題は RNN を用いた SMILES 文字列生成モデルを構築することで解決できる。ChemTS では ZINC[18]から取得した約 25 万分子の SMILES 文字列を学習させた RNN モデルを用意し、開始トークン(&)または一連のトークンを入力した際に次にくるトークンの確率分布を得られるようにしている。拡張ステップでは RNN で予測された次にくるトークンの確率分布を元に未探索ノードを追加していく。ロールアウトステップでは未探索ノードおよび選択ステップで辿ってきた各ノードを入力にして終了トークン(¥n)が選択されるまでトークンを生成し SMILES 文字列を構築する。

これまでに Tsuda らの研究グループは、ChemTS を用いて主に材料分野において様々な性質を持つ分子を設計および実際に合成・評価をして ChemTS の分子設計手法としての性能を示してきており、具体的には、望みの吸収波長を持つ分子[19]、電荷を保存する分子（エレクトレットポリマー）[20]、蛍光分子[21]などが挙げられる。これまでに様々な分子設計手法が提案されてきてはいるが、実際に合成可能かつ所望の性質を持った分子を生成可能であることを実証してきた手法は多くはないため、新たに分子生成を試してみたいユーザーにとって ChemTS は有望な選択肢の一つになるだろう。創薬分野においても、in silico での評価のみの成果ではあるが Structure-based Drug Design[22]や Ligand-based Drug Design[23]に ChemTS を応用し有望と考えられる医薬品候補を生成することに成功している。

ここまでを読んでいただいた読者の中には実際に ChemTS を使ってみてみたいと思った読者もいるのではないだろうか。ChemTS は GitHub にて誰でも利用可能な形式で公開[24]されてはいたが、実際に利用しようとするアルゴリズムに精通していないとユーザー毎のユースケースに合わせた分子生成を実施するのが難しいのが実情であった。そこで筆者らは ChemTS を誰でも「手軽に」「任意の性質」に対しての分子生成が実施できるようにするため、ChemTSv2 の開発を進めている[25]。ChemTSv2 ではユーザーは「分子から所望の性質を計算し報酬関数を設定することのみ」に集中すればよく、探索パラメータや生成数等の設計条件は「事前に用意された設定ファイル」で指定できるため、ChemTS のアルゴリズムに精通していなくても簡単に ChemTS による分子設計を試すことができる。ChemTS で上記の作業を実施するためにはソースコード全体を理解した上で該当箇所を探し直接改変する必要があったが、ChemTSv2 ではそれぞれの作業をソースコードとは独立したファイルで用意・設定する枠組みであるため、ChemTS を用いた分子生成モデルの開発・検討が容易となる。次章からは阻害活性予測モデルとファーマコフォアモデルを用いた ChemTSv2 による分子設計について紹介していく。本稿で使用した報酬関数スクリプトおよび設定ファイルは GitHub 上に公開[26]しているため、実際に ChemTSv2 による分子設計を試してみたい方はそちらも参照されたい。また、実際どのように報酬関数を作成すればよいかについては文献[25]および GitHub 上の README[26]を併せて参照されたい。

3. ChemTS の in silico 創薬への応用

3.1 阻害活性予測モデルを活用した ChemTS による分子設計

医薬品候補化合物を探索する際には多岐にわたる項目を最適化すること（多目的最適化）が求められる。主な項目としては、標的とするタンパク質への主活性や選択性、化合物の体内動態・毒性 (ADMET) などが挙げられるだろう。このような多目的最適化問題に対して分子設計 AI を用いて取り組む一例として、ChemTSv2 を用いた epidermal growth factor receptor (EGFR) 選択的阻害剤設計について紹介していく[23]。ChemTSv2 を含む多くの分子設計手法において、複数の目的関数を同時に最適化したい場合には荷重和法などのスカラー化手法を用いて単一目的化する必要がある。例えば、ある物理化学的性質 A と B を同時に最適化しようとした場合、A と B を計算するための関数 $A(x)$ と $B(x)$ およびそれぞれの重み w_A ・ w_B を適切に組み合わせた関数 $C(x) = w_A * A(x) + w_B * B(x)$ 等を定義する必要がある。今回の EGFR 選択的阻害剤設計では、重み付き幾何平均で表される Dscore[27]を用いて単一目的化を行うこととした。各目的関数から算出される値は 0 から 1 の範囲になるようスケールした上で Dscore の計算に用いた。

EGFR 選択的阻害剤を設計するための最適化項目としては、阻害活性・薬物動態・その他化合物の性質の 3 カテゴリー計 18 項目を設定した。阻害活性カテゴリーに関して、EGFR および 8

種の EGFR 類縁チロシンキナーゼ (オフターゲット) Abelson tyrosine-protein kinase (ABL)、eprhin type-B receptor 4 (EPHB4)、receptor protein-tyrosine kinase erbB-2 (ERBB2)、fibroblast growth factor receptor 1 (FGFR1)、lymphocyte-specific tyrosine-protein kinase (LCK)、platelet-derived growth factor receptor beta (PDGFR β)、proto-oncogene tyrosine-protein kinase (SRC)、vascular endothelial growth factor receptor 2 (VEGFR2) への阻害活性 (pIC₅₀) 計 9 項目を設定した。薬物動態カテゴリーに関しては水溶性・膜透過性・代謝安定性の 3 項目、その他カテゴリーに関しては synthetic accessibility score (SAscore)[28]・quantitative estimate of drug-likeness (QED)[29]・急性毒性[30]・分子量フィルター[31]・Tox alert フィルター[32][33]・ChEMBL structure フィルター (構造妥当性評価) [31]の 6 項目を設定した。これら 18 項目の値を算出するために、阻害活性・薬物動態・急性毒性に関しては機械学習モデルである Light Gradient Boosting Machine (LightGBM) [34]を用いて予測モデルを構築し、残りの項目については RDKit[35]を用いた。それぞれの詳述は文献[23]を参照されたい。

実験設定に関して、選択性の考慮が有効であるかを (a) EGFR のみ考慮した分子設計と (b) EGFR およびその他チロシンキナーゼを考慮した分子設計を比較することで検証した。どちらの場合も薬物動態およびその他の項目に関しては考慮している。また、強化学習における「探索 (exploration) と活用 (exploitation)」の活用を優先させるため、探索パラメータ c を 0.2 に設定した。ChemTS における「活用を優先 (今回の設定のように探索パラメータ $c = 0.2$)」するとは、探索過程で最適と思われる分子が見つかったとそれを優先して探索を進めることを指す。一方「探索を優先する (例えば探索パラメータ $c = 1.0$)」とは、最適と思われる分子以外にも未探索の分子を幅広く探索することを優先することを指す。生成する分子数は 20 万とした。

図 2 に EGFR 選択的阻害剤設計における各項目の移動平均 (図 2a) および設計された化合物例 (図 2b) を示す。EGFR の阻害活性のみを考慮した場合、EGFR への阻害活性が上昇するにつれてオフターゲットへの阻害活性も上昇していることが図 2a からわかる。オフターゲットへの阻害活性を最小化するように考慮した場合、EGFR への阻害活性を上昇させながらもオフターゲットへの阻害活性を抑えることができていた (図 2b)。EGFR への阻害活性と 8 オフターゲットのうち最大の阻害活性の差 (Δ pIC₅₀) の観点でみると、EGFR のみの考慮の場合は移動平均で 0.5

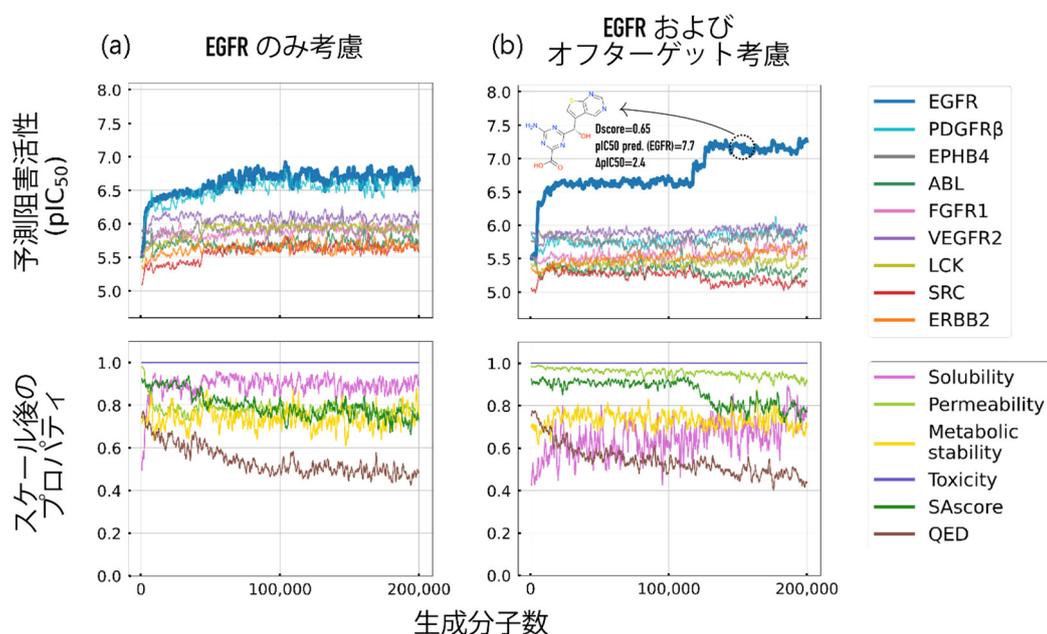


図 2. EGFR 選択的阻害剤の設計における各目的関数の最適化過程。(a) EGFR のみを考慮した場合、(b) EGFR およびオフターゲットの阻害活性を考慮した場合での最適化過程を表す。どちらも薬物動態・その他カテゴリーは考慮している。(b)では代表的な化合物を例としてあげている。(下段)創薬に好適を 1、不適を 0 としてスケールした各種プロパティ。文献[23]の図を改変し引用。

程度の差であったところ、オフターゲットの考慮をした場合に移動平均で 1.0 程度の差になり選択性が向上していた。Dscore の高い化合物を個別に確認していくと、図 2b に示す実際の化合物例では EGFR への阻害活性が 7.7 であり ΔpIC_{50} が 2.4 と予測値ベースではあるが十分な選択性を示す化合物が得られていたことがわかった。薬物動態およびその他のカテゴリーの項目に関しても、最適化の程度に差はあれど概ね許容範囲の値となっていた。詳述な議論に関しては文献 [23] を参照されたい。以上の結果より、ChemTSv2 と Dscore を組み合わせることで、18 という多い目的関数を同時最適化する必要があったとしても分子設計が実施可能であるということを示せたであろう。

3.2 ファーマコフォアモデルを活用した ChemTS による分子設計

次に EGFR 既知リガンドからファーマコフォアモデルを定義し、その条件を満たす分子設計について ChemTSv2 を用いて試してみる。ファーマコフォアモデルとは物理化学的特性である水素結合ドナー/アクセプターや疎水性/芳香族性などの特性基の 3 次元的配置が定義されたモデルのことを指す。ファーマコフォアモデルを用いた解析は、*in silico* スクリーニング手法のうち Ligand-based Drug Design としてよく利用されている。以降実際の試行実験について解説する。

まず TeachOpenCADD[36] の Talktorial009[37] を参考に、EGFR 既知リガンド 4 つ (図 3a) から RDKit を用いてファーマコフォアモデル (図 3b) を構築した。主な構築の手順としては、①EGFR 既知リガンド 4 つそれぞれについて RDKit の BaseFeatures.fdef に定義されている特性基を 3 次元配置と共に検索、②全てのリガンドが持つ特性基とその配置を k-means 法により特定、③結晶構造の情報も参考にしてファーマコフォアモデルを定義、である。今回は検索する特性基は Donor・Acceptor・Aromatic・Hydrophobic の 4 つに限定し、それぞれの特性基の許容半径は 0.5, 0.5, 1.0, 1.0 Å とした。また、目的関数としては「定義した特性基を満たす割合」と「特性基の 3 次元的配置における偏差平方和 (SSD)」の 2 つとした。報酬関数としてはその 2 つの目的関数の加重平均を算出することで単一目的化した。重みは「定義した特性基を満たす割合」:「特性基の 3 次元的配置における SSD」が 1:4 とし、特性基の 3 次元的配置を満たすような分子設計を重視する設定とした。ファーマコフォア解析の方法に関しては RDKit の User Group Meeting2016 の資料を参考にした [38]。探索パラメータ c は活用重視の 0.2 とし、分子生成数は 3 万に設定した。また不適切な化合物を探索段階から除外するために、ChemTSv2 で事前に定義してある分子フィルターのうち SAscore が 4.5 以下・ラジカル分子を除外というフィルターを設定した。ChemTSv2 では、これら以外の分子フィルターとして Lipinski の rule of five や PAINS に関するフィルター

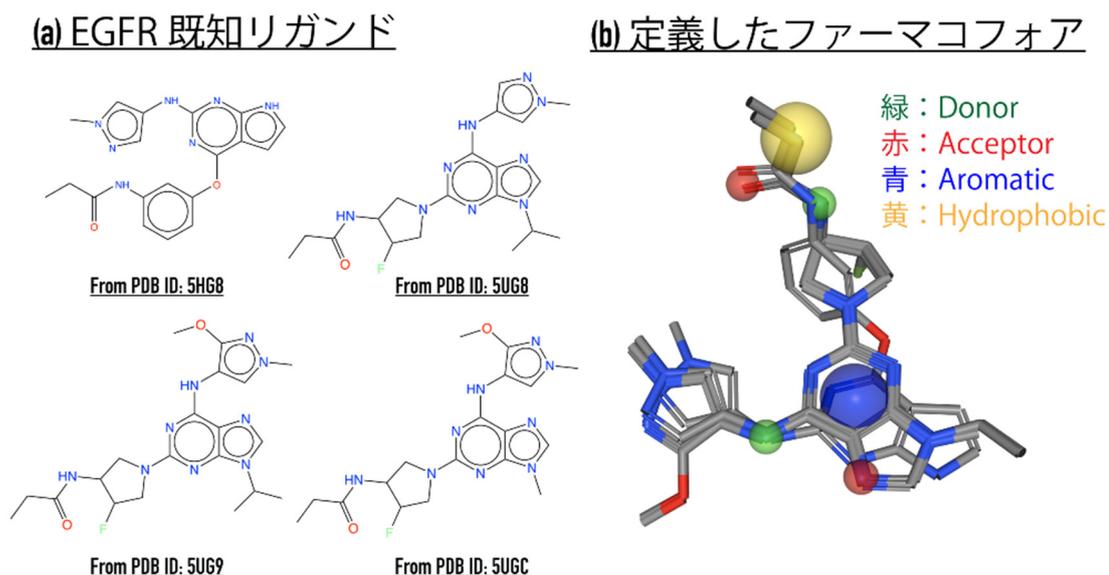


図 3. 既知 EGFR リガンド 4 種とそれより構築したファーマコフォアモデル

等も用意しており、設定ファイルにて簡単に指定することができる（詳述は GitHub の README[26]参照）。

図 4 は分子生成過程における SSD の移動平均を示したものである。探索過程において特に探索初期段階では探索中の分子がファーマコフォアの特徴基の条件を満たせず SSD が算出できないため、SSD が算出できていない箇所を線形補間した後に移動平均を算出している。探索終盤に近づくにつれて SSD の値が下がっていることから、うまく最適化が進んでいるといえる。加えて、定義した各特徴基の数を満たす割合は探索序盤から終盤にかけて 1 となる分子が生成されていたことから、ChemTS が定義されたファーマコフォアモデルの条件を満たす分子を設計できていることが示唆される。次に設計された分子を確認したところ、定義したファーマコフォアモデルの条件は満たしているものの、EGFR 阻害剤設計ができたというには不十分な結果であることがわかった。その中で特に改善すべき点としては、キナーゼ阻害剤の特徴である Aromatic, Donor, Acceptor が平面上に並んでいるという特徴は再現できていなかったこと、SSD を最小にするために芳香環が歪んでいたことが挙げられる。よりキナーゼ阻害剤として有望な分子を設計したい場合は、今回は削ってしまった特徴基を追加する等ファーマコフォアモデルを再構築する、歪みエネルギーをペナルティ項として加える、生成分子数を増やす、探索パラメータ c を調整するなどが考えられるので、もし興味のある読者は公開している報酬関数と設定ファイルを用いて挑戦して頂ければ幸いである。

偏差平方和 (SSD) の最適化過程

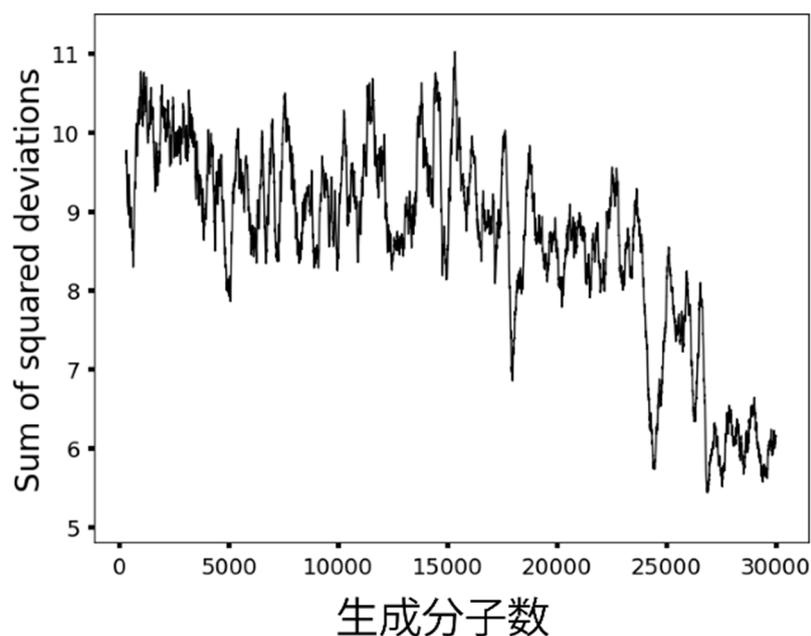


図 4. ファーマコフォアモデルの条件を満たす分子設計の最適化過程。

4. おわりに

本稿では、深層学習と木探索を組み合わせた *de novo* 分子設計手法である ChemTS について概説し、我々が現在開発している ChemTS の改良版である ChemTSv2 による阻害活性予測モデルとファーマコフォアモデルを用いた *de novo* 分子設計について紹介した。ここ 10 年ほどで人工知能分野のブレイクスルーもあり、非常に多くの *de novo* 分子設計手法が提案されてきている。そして分子設計 AI のさらなる発展のためには、実際にメディシナルケミストを含む化学者がそれら分子設計 AI を利用し評価を開発者へフィードバックすることが必要不可欠であろう。まだ分子設計 AI の能力に懐疑的な方は多いかと思うが、本稿で示したように ChemTS は設定された設計条件をきちんと満たすような分子設計が可能である。本稿が創薬を含めた様々な分子設計

に、ChemTS を含めた de novo 分子設計手法をどのように活用できるかを検討するきっかけになれば幸いである。

謝辞

本稿を執筆する機会をいただきました SAR News 編集委員長である幸瞳先生（理化学研究所）に感謝いたします。

参考文献

- [1] Kirkpatrick, P., Ellis, C. Chemical space. *Nature*, **432**, 823 (2004).
- [2] CAS REGISTRY®, <https://www.cas.org/cas-data/cas-registry> (accessed on 7 March, 2023).
- [3] Olivecrona, M., Blaschke, T., Engkvist, O., et al. Molecular de-novo design through deep reinforcement learning. *J. Cheminf.*, **9** (2017).
- [4] Kusner, M. J., Paige, B., Hernández-Lobato, J. M. Grammar variational autoencoder. *Proceedings of the 34th International Conference on Machine Learning*, **70**, 1945–1954 (2017).
- [5] Kadurin, A., Nikolenko, S., Khrabrov, K., et al. druGAN: An advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico. *Mol. Pharm.*, **14**, 3098–3104 (2017).
- [6] Gómez-Bombarelli, R., Wei, J., Duvenaud, D., et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Cent. Sci.*, **4**, 268–276 (2018).
- [7] Segler, M. H., Kogej, T., Tyrchan, C., et al. Generating focused molecule libraries for drug discovery with recurrent neural networks. *ACS Cent. Sci.*, **4**, 120–131 (2018).
- [8] Guimaraes, G. L., Sanchez-Lengeling, B., Outeiral, C., et al. Objective-reinforced generative adversarial networks (ORGAN) for sequence generation models. *arXiv.*, (2018). <https://doi.org/10.48550/arXiv.1705.10843>
- [9] Jin, W., Barzilay, R., Jaakkola, T. Junction tree variational autoencoder for molecular graph generation. *Proceedings of the 35th International Conference on Machine Learning*, **80**, 2323–2332 (2018).
- [10] Shi, C., Xu, M., Zhu, Z., et al. GraphAF: a flow-based autoregressive model for molecular graph generation. *International Conference on Learning Representations* (2020). <https://arxiv.org/abs/2001.09382>
- [11] Bagal, V., Aggarwal, R., Vinod, P. K., et al. MolGPT: molecular generation using a transformer-decoder model. *J. Chem. Inf. Model.*, **62**, 2064–2076 (2022).
- [12] Schneuing, A., Du, Y., Harris, C., et al. Structure-based drug design with equivariant diffusion models. *arXiv.*, (2022). <https://doi.org/10.48550/arXiv.2210.13695>
- [13] Bilodeau, C., Jin, W., Jaakkola, T., et al. Generative models for molecular discovery: Recent advances and challenges. *WIREs Comput. Mol. Sci.*, **12** (2022).
- [14] Yang, X., Zhang, J., Yoshizoe, K., et al. ChemTS: an efficient python library for de novo molecular generation. *Sci. Technol. Adv. Mater.*, **18**, 972–976 (2017).
- [15] Rumelhart, D. E., McClelland, J. L. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition: Foundations*, pp 318–362 (1987).
- [16] Coulom, R. Efficient selectivity and backup operators in Monte-Carlo tree search. *5th International Conference on Computer and Games* (2006).
- [17] Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**, 31–36 (1988).
- [18] Sterling, T., Irwin, J. J., ZINC 15—Ligand discovery for everyone. *J. Chem. Inf. Model.* **55**, 2324–2337 (2015).
- [19] Sumita, M., Yang, X., Ishihara, S., et al. Hunting for organic molecules with artificial intelligence: molecules optimized for desired excitation energies. *ACS Cent. Sci.*, **4**, 1126–1133 (2018).
- [20] Zhang, Y., Zhang, J., Suzuki, K., et al. Discovery of polymer electret material via de novo molecule generation and functional group enrichment analysis. *Appl. Phys. Lett.*, **118**, 223904 (2021).
- [21] Sumita, M., Terayama, K., Suzuki, N., et al. De novo creation of a naked eye-detectable fluorescent molecule based on quantum chemical computation and machine learning. *Sci. Adv.*, **8** (2022).
- [22] Ma, B., Terayama, K., Matsumoto, S., et al. Structure-based de novo molecular generator combined with artificial intelligence and docking simulations. *J. Chem. Inf. Model.*, **61**, 3304–3313 (2021).

- [23] Yoshizawa, T., Ishida, S., Sato, T., et al. Selective inhibitor design for kinase homologs using multiobjective Monte Carlo tree search. *J. Chem. Inf. Model.*, **62**, 5351–5360 (2022).
- [24] <https://github.com/tsudalab/ChemTS>
- [25] Ishida, S., Aasawat, T., Sumita, M., et al. ChemTSv2: democratizing functional molecular design using de novo molecule generator. *ChemRxiv.*, (2023). <https://doi.org/10.26434/chemrxiv-2023-mb880-v2>
- [26] <https://github.com/molecule-generator-collection/ChemTSv2>
- [27] Cummins, D. J., Bell, M. A. Integrating everything: the molecule selection toolkit, a system for compound prioritization in drug discovery. *J. Med. Chem.* **59**, 6999–7010 (2016).
- [28] Ertl, P., Schuffenhauer, A. Estimation of synthetic accessibility score of drug-like molecules based on molecular complexity and fragment contributions. *J. Cheminf.* **1**, (2009).
- [29] Bickerton, G. R., Paolini, G. V., Besnard, J., et al. Quantifying the chemical beauty of drugs. *Nat. Chem.*, **4**, 90–98 (2012).
- [30] Zhu, H., Martin, T. M., Ye, L., et al. Quantitative structure-activity relationship modeling of rat acute toxicity by oral exposure. *Chem. Res. Toxicol.* **22**, 1913–1921 (2009).
- [31] Winter, R., Montanari, F., Steffen, A., et al. Efficient multi-objective molecular optimization in a continuous latent space. *Chem. Sci.* **10**, 8016–8024 (2019).
- [32] Sushko, I., Salmina, E., Potemkin, V. A., et al. ToxAlerts: a web server of structural alerts for toxic chemicals and compounds with potential adverse reactions. *J. Chem. Inf. Model.* **52**, 2310–2316 (2012).
- [33] Non MedChem-Friendly SMARTS. <https://www.surechembl.org/knowledgebase/169485-non-medchem-friendly-smarts> (accessed on 7 March, 2023).
- [34] Ke, G., Meng, Q., Finley, T., et al. LightGBM: a highly efficient gradient boosting decision tree. *Adv. Neural Inf. Process Syst.*, **30**, 3146–3154 (2017).
- [35] Landrum, G. RDKit: Open-source cheminformatics. <https://www.rdkit.org>.
- [36] Sydow, D., Morger, A., Driller, M., & Volkamer, A. TeachOpenCADD: a teaching platform for computer-aided drug design using open source packages and data. *J. Cheminf.*, **11** (2019).
- [37] https://projects.volkamerlab.org/teachopencadd/talktorials/T009_compound_ensemble_pharmacophores.html
- [38] https://github.com/rdkit/UGM_2016/blob/master/Notebooks/Stiefl_RDKitPh4FullPublication.ipynb

//// Cutting Edge ////

MOE REINVENT I/F の紹介

株式会社モルシス 木村嘉朗

1. はじめに

Molecular Operating Environment (MOE) [1]は、創薬研究・生命科学研究のためのアプリケーションを統合した計算化学プラットフォームである。Chemical Computing Group (CCG) 社が独自に開発した Scientific Vector Language (SVL) を搭載しており、この開発環境を使用すれば、さまざまなサードパーティ製ソフトウェアとの連携が可能である。MOE の REINVENT インターフェース (MOE REINVENT I/F) はその活用例の一つである。低分子の生成モデルである REINVENT[2]は、近年幅広い分野で活用されているディープ・ニューラルネットワークの技術を創薬研究に取り入れた応用事例として注目されている。MOE REINVENT I/F では、MOE 上から REINVENT を実行するための使いやすい GUI と、強化学習の評価関数として使用できる MOE のさまざまな機能を提供する。MOE REINVENT I/F は、MOE のユーザーであれば、MOE のアドオンプログラム共有サイトである SVL Exchange[3]から無償で入手できる。本稿では、REINVENT の概要と MOE REINVENT I/F の機能や使用例について紹介する。

2. REINVENT の概要

REINVENT は、プログラミング言語 Python で実装された Recurrent Neural Network (RNN) に基づく SMILES 生成モデルである。SMILES は、分子の化学構造の表記法の一つで、図 1 のように分子を文字列で表現する。REINVENT では、化合物ライブラリーを用いて教師あり学習されたモデルを、フィンガープリントによる類似度や活性予測モデルなどの評価関数を取り入れた強化学習によって、所望の特性を持つような化学構造の SMILES を生成するようにさらに調整できる。

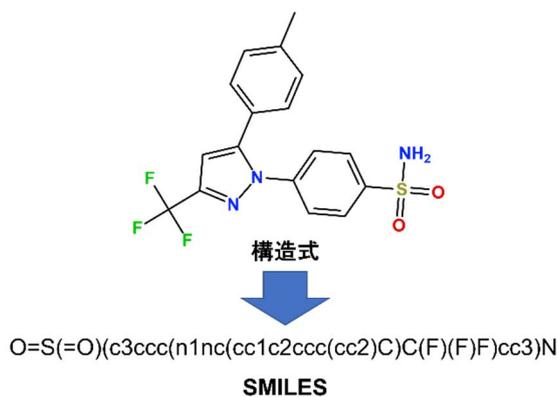


図 1. 構造式 (上) とその SMILES (下) の例

2.1 RNN の教師あり学習

RNN とは、時系列データを扱えるように拡張したニューラルネットワークである。主に、自然言語処理の分野で用いられている。RNN では、一つ前の隠れ層の出力が次の入力の一部となり、次の出力に影響を与える。図 2 は、RNN の教師あり学習の様子を示したものである。例えば、 X_0 から X_4 で構成された配列を教師あり学習する場合を考える。ここでは、 X_0 は開始の記号で X_4 は終了の記号である。 X_0 を最初の入力として与えると、出力として次の要素の確率分布が得られる。特に $P(X_1|X_0)$ は、 X_0 が得られた条件下での正解 X_1 が得られる確率を示している。こ

こでの隠れ層の出力（図2の右向きの矢印）が X_1 とともに次の入力として用いられる。この計算ステップを最後の X_4 の確率分布が得られるまで繰り返す。教師あり学習では、正解の要素の確率の対数にマイナスをつけたものを損失関数として定義し、これを最小化するように RNN の重み (W) やバイアス (b) といったパラメーターを最適化する。

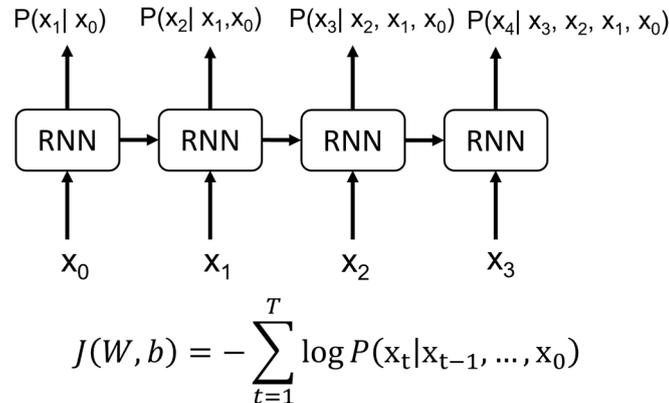


図 2. RNN の教師あり学習の概念図（上）と損失関数（下）

REINVENT においては、ここで示した配列が SMILES（正確には SMILES の先頭に開始の記号、末尾に終了の記号を加えた文字列）、配列の要素が SMILES のトークンに対応する。SMILES のトークンは、SMILES の構成要素の中で意味のある最小の文字列である。例えば、C（炭素）や Br（臭素）などの元素名、環構造を示す数字ラベル、分岐を示す丸括弧などがある。REINVENT の RNN のアーキテクチャには Gated Recurrent Unit (GRU) が用いられている。GRU は、Long Short-Term Memory (LSTM) と比較して構造が簡素であり、計算コストを抑えつつ同程度の効果が期待できる[4]。

2.2 RNN による配列データの生成

RNN による配列データの生成では、図 3 のように教師あり学習を行ったモデルにまず開始の記号 X_0 を入力する。次に RNN の出力として次の配列の要素の確率分布が得られるので、それに基づいて配列の要素をサンプリングする。そして、その配列の要素と RNN の隠れ層の出力を入力として、同様に次の要素をサンプリングする。この過程を終了の記号（図 3 においては、 X_4 ）が得られるまで繰り返し行われる。REINVENT においては、上記のサブセクションと同様に、ここで示した配列が SMILES、配列の要素が SMILES のトークンに対応する。

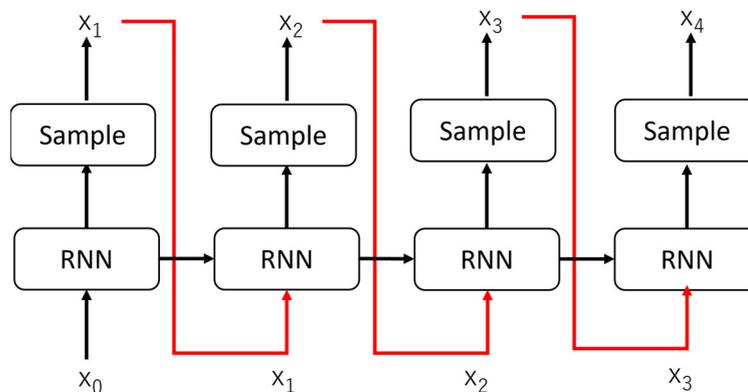


図 3. RNN の配列データの生成の概念図

2.3 REINVENT の構成と強化学習

REINVENT では、妥当な SMILES を生成するモデルである Prior Network が提供されている。この Prior Network は、医薬品候補化合物のアッセイデータベースの ChEMBL22[5]から SMILES を抽出して、これを基に上記の教師あり学習を行ったモデルである (図 4)。このモデルはおよそ 90%以上の高い確率で表現として正しい SMILES を生成できる。また、ユーザー自身が独自の SMILES を用いて一からモデルを構築したり、あるいは転移学習を行ったりできる。ここで言う転移学習とは、既存のモデルを基にさらに別のデータセットで教師あり学習を行い改良することである。強化学習では、Prior Network が Agent Network (強化学習によって調整されたモデル) の初期モデルとして用いられる。強化学習の各計算ステップでは、Agent Network によって SMILES 群を生成し、それらのスコアを計算する。既存機能では、いくつかのフィンガープリントによる類似度や活性予測モデルを評価関数に使用できるが、MOE REINVENT I/F では、評価関数としてさらに MOE の機能が使用できる。次にこれらのスコアと Prior Network から計算した SMILES 群の生成確率の対数から、理想的な生成確率の対数を求める。さらにこれと Agent Network による SMILES 生成確率の対数の差の二乗を損失関数として定義する。そして、この損失関数を最小化するように Agent Network の重み (W) やバイアス (b) といったパラメーターが更新される。このような計算ステップを繰り返し行うことで、Agent Network は高いスコアを持つ SMILES を生成できるように調整される。

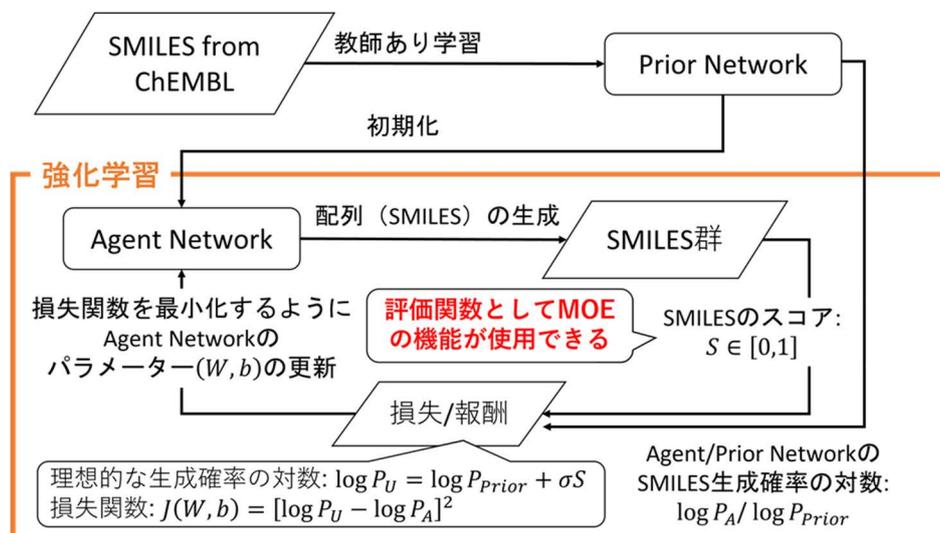


図 4. REINVENT の強化学習の概念図

3. MOE REINVENT I/F

MOE の REINVENT インターフェースである MOE REINVENT I/F では、使いやすい GUI とともにさまざまな機能を提供する。それらの中でも特に重要な機能は、強化学習に対する MOE の評価関数の提供である。図 5 には、強化学習における REINVENT と MOE とのやり取りとそれぞれが行う処理内容を簡単に示している。まず、REINVENT の生成モデルからサンプリングされた SMILES 群が、MOE に渡される。MOE では、SMILES 群を分子構造群に変換し、さらに評価関数が必要とする分子の情報に応じて、3 次元構造に変換したり、配座解析を実施したりする。その後、それらの構造を基に評価関数によりスコアを計算する。評価関数には、分子量や水素結合ドナー/アクセプター数などのさまざまな記述子、特定の部分構造の有無、活性/物性予測モデル、フィンガープリントモデル (フィンガープリントによる類似度)、ファーマコフォアモデル、ドッキングシミュレーションによるスコアが使用できる。さらに上記に列挙したモデルを自由に組み合わせたモデル (コンセンサスモデル) も使用できる。そして、これらの評価関数で計算されたスコアが REINVENT に渡される。REINVENT では、スコアから損失関数が計算され、

それに応じて生成モデルのパラメーターが更新される。この一連の流れが強化学習の 1 ステップとなる。

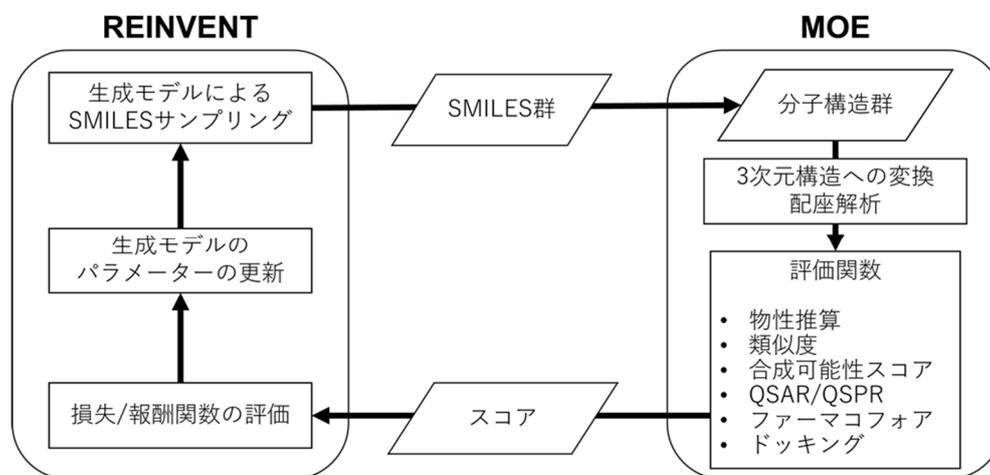


図 5. REINVENT と MOE を組み合わせた強化学習の概念図

MOE REINVENT I/F では、その他に以下の機能も提供する。

- 強化学習の実行と監視： GUI からさまざまなパラメーターを設定し、強化学習を実行できる。実行している間、生成された SMILES の構造とそのスコアを監視できる (図 6)。
- サンプリングした化合物の評価： 生成モデルでサンプリングした SMILES の分子構造の 3 次元化や評価関数の適用が行える。
- Prior Network の作成と教師あり学習の GUI： SMILES ファイルだけでなく、MOE 独自の分子データベースファイル (MDB) や SD ファイルを入力データとして Prior Network の作成と教師あり学習が GUI から行える。

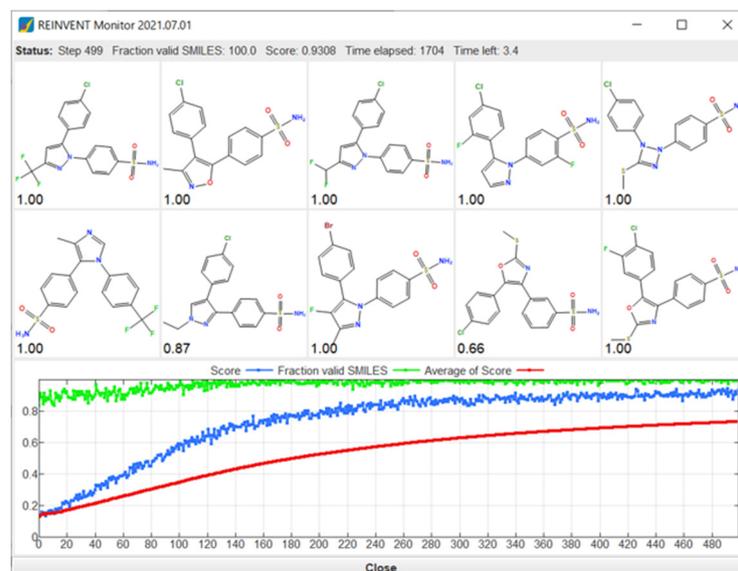


図 6. REINVENT の強化学習における MOE REINVENT I/F の監視機能。上部には、各計算ステップで生成された化合物とスコアが表示され、下部には横軸を計算ステップとした妥当な SMILES の割合 (緑)、スコアの平均 (青色) と累積の平均 (赤) がプロットされている。

4. MOE REINVENT I/F の使用例

ここでは、上皮成長因子受容体 (EGFR) の阻害剤を題材として、MOE REINVENT I/F の3つの使用例を紹介する。

4.1 ファーマコフォアモデルを評価関数に使用

図7左のように、EGFR とその阻害剤エルロチニブの共結晶構造 (PDB ID: 1M17) を基にして、水素結合アクセプターとその投影点と芳香環中心の特性球を持つファーマコフォアモデルを構築した。それを評価関数として500ステップの強化学習を行った。ここでは、MOE REINVENT I/F が Agent Network で生成された SMILES 群の分子構造の3次元化、配座解析、ファーマコフォア検索、スコアの計算を担っており、強化学習の各ステップで実施される。その結果得られた化合物群の一部が図7右に示されている。エルロチニブと同様の縮合環を持つ構造が得られているのが分かる。なお、実際の出力ではファーマコフォアモデルを満たす3次元構造として得られる。

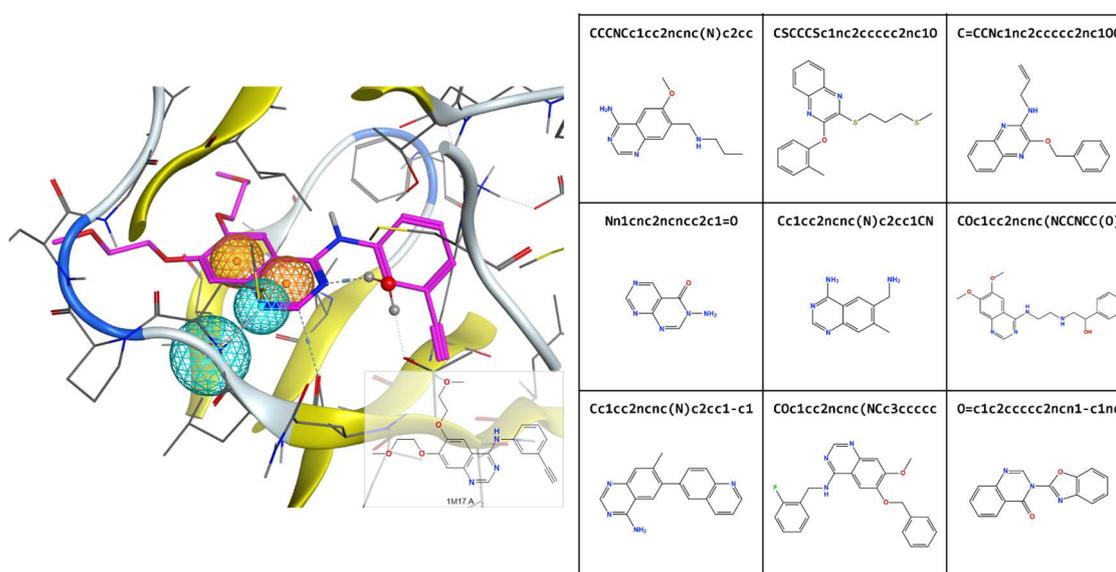


図7. 強化学習の評価関数に使用したファーマコフォアモデル (左) と生成された化合物群 (右)。ファーマコフォアモデルは水素結合アクセプター (水色) とその投影点 (水色) と芳香環中心 (橙色) の特性球を含む。

4.2 コンセンサスモデルを評価関数に使用

コンセンサスモデルは QSAR モデル、フィンガープリントモデル、ファーマコフォアモデルなどを組み合わせたモデルである。コンセンサスモデルの評価関数は、それを構成するモデルのスコアの変数に四則演算などを組み合わせて自由に定義できる。ここでは、MOE のファーマコフォアベースのフィンガープリント GpiDAPH3 で EGFR 阻害活性の高い化合物群と比較したときの類似度と、血液脳関門の透過性を予測する線形モデルを組み合わせたコンセンサスモデルを使用した。このモデルの評価関数は、血液脳関門の透過性が高く、既知の EGFR 阻害剤と似ている程、高くなるように定義した。得られた結果では、既知の EGFR 阻害剤と同様な縮合ヘテロ環を部分構造にもつ化合物が得られているのが分かる (図8)。

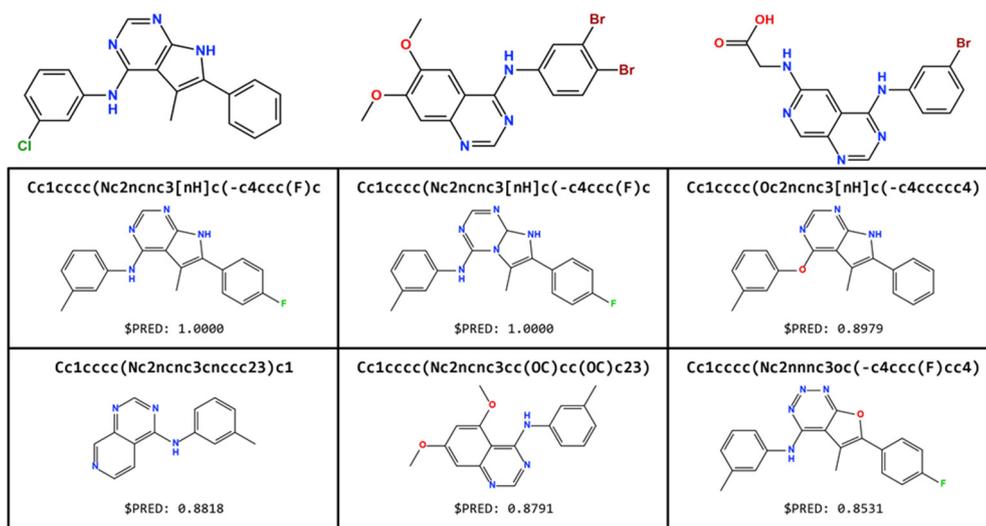


図 8. 既知の EGFR 阻害剤（上）と強化学習にコンセンサスモデルを用いて得られた化合物群（下）。コンセンサスモデルでは、EGFR 阻害剤のフィンガープリントモデルと血液脳関門の透過性を予測する線形モデルを組み合わせている。

2.3 強化学習の前に学習効率を高めるために転移学習を使用

既存の Prior Network では、エルロチニブにあるようなフェニルアセチレンを含む構造が得られにくい。これは、教師あり学習に用いた ChEMBL の SMILES 群にフェニルアセチレンがほとんど含まれないためである。また、そのような Prior Network で強化学習を行っても同様にそのような構造は得られにくい。そのような場合には追加の教師あり学習が有効である。ここでは、フェニルアセチレンを含む構造を MOE のリードライクデータベースや試薬ベンダーの化合物カタログから検索し、得られた化合物を用いて既存の Prior Network をさらに教師あり学習を行って改良した。その後の強化学習では、EGFR とその阻害剤エルロチニブの共結晶構造（PDB ID: 1M17）のエルロチニブのフェニルアセチレンをテンプレートとしたドッキングシミュレーションによるスコアを評価関数に用いた。その結果、フェニルアセチレンを含むさまざまな化合物のドッキング構造が得られた（図 9）。

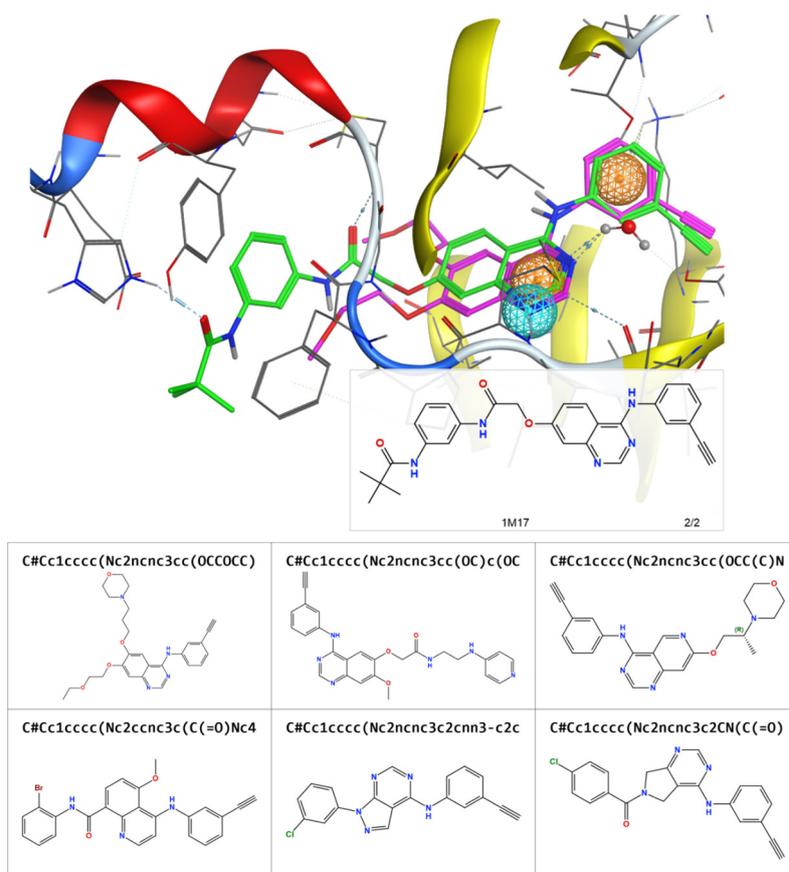


図9. EGFRの共結晶(上)と強化学習で得られた化合物群(下)。上の図には、ドッキングシミュレーションに使用したファーマコフォアモデル(水素結合アクセプター(水色)と芳香環中心(橙色)の特性球)と強化学習で得られた化合物(緑)の一つも合わせて表示している。

5. おわりに

今回は SMILES 生成モデルの REINVENT の概要とその MOE インターフェースである MOE REINVENT I/F の機能と使用例を紹介した。ここで紹介した使用例は、計算化学が専門でない方でも GUI から簡単に行えるものである。冒頭にも記したが、MOE REINVENT I/F は、MOE ユーザーであれば、MOE のアドオンプログラム共有サイトである SVL Exchange[3]から無償で入手できる。

REINVENT の他にも興味深い生成モデルやツールが GitHub の Molecular AI[6]で公開されている。例えば、Lib-INVENT[7]では母核構造を維持しつつ反応ルールに基づいた化合物生成が可能であり、GraphINVENT[8]ではグラフニューラルネットワークに基づく化合物生成が可能である。また、DockStream[9]を使用すれば、MOE REINVENT I/F と同様に REINVENT といくつかのドッキングシミュレーションソフトウェアを評価関数に用いることができる。

参考文献

- [1] <https://www.chemcomp.com/Products.htm>
- [2] Olivecrona, M., Blaschke, T., Engkvist, O., Chen, H. Molecular de-novo design through deep reinforcement learning. *J. Cheminform.*, **9**, 48 (2017). doi:10.1186/s13321-017-0235-x
- [3] <https://svl.chemcomp.com/>
- [4] Chung, J., Gulcehre, C., Cho, K., Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv*:**1412.3555** (2014). doi:10.48550/arXiv.1412.3555

- [5] Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., Overington, J. P. ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res.*, **40**, D1100-D1107 (2012). doi:10.1093/nar/gkr777
- [6] <https://github.com/MolecularAI>
- [7] Fialková, V., Zhao, J., Papadopoulos, K., Engkvist, O., Bjerrum, E. J., Kogej, T., Patronov, A. LibINVENT: reaction-based generative scaffold decoration for in silico library design. *J. Chem. Inf. Model.*, **62**, 2046-2063 (2022). doi:10.1021/acs.jcim.1c00469
- [8] Mercado, R., Rastemo, T., Lindelöf, E., Klambauer, G., Engkvist, O., Chen, H., Bjerrum, E. J. Graph networks for molecular design. *Mach. Learn.: Sci. Technol.*, **2**, 025023 (2021). doi:10.1088/2632-2153/abcf91
- [9] <https://github.com/MolecularAI/DockStream>

//// SAR Presentation Award ////

SAR Presentation Award について

「SAR Presentation Award」は、構造活性相関シンポジウムにおける若手研究者の発表を奨励し、構造活性相関研究の発展を促進するため、2010 年度に創設された。当初は応募制として審査対象講演の募集を行った。2012 年度からは、正式名称を「構造活性相関シンポジウム優秀発表賞」（英語表記 SAR Presentation Award）と定めた。

2022 年度 SAR Presentation Award について

2022 年度は、第 50 回構造活性相関シンポジウム（オンライン開催）における 45 歳以下の発表者（日本薬学会会員または受賞後に日本薬学会に入会いただける方）による一般講演（口頭発表・ポスター発表）を選考対象とすることとした。また同様の選考方法で構造活性相関シンポジウムの第 50 回を記念して実行委員会特別賞を合わせて選定した。

2022 年度 SAR Presentation Award 受賞者（演題番号順）

口頭発表	:	吉澤 竜哉	(横浜市立大学生命医科学研究科)
口頭発表	:	原田 隆平	(筑波大学計算科学研究センター)
ポスター発表	:	高谷 大輔	(大阪大学大学院薬学研究科)
ポスター発表	:	市丸 嘉	(湘南医療大学薬学部)
ポスター発表	:	星野 小百合	(横浜市立大学理学部)

第 50 回構造活性相関シンポジウム実行委員会特別賞

口頭発表	:	田中 良樹	((株) アグロデザイン・スタジオ)
ポスター発表	:	渡邊 友里江	(昭和大学薬学部)

受賞者の選考について

2022 年 11 月 10 - 11 日にオンラインにて各審査員から提出頂いた審査票を集計し、口頭発表 2 名、ポスター発表 3 名を受賞者として選出した。口頭発表の審査は点数方式、ポスター発表の審査は 3 演題選出するという方式で行った。審査員が共同演者となっている演題は選出不可としている。後日受賞者には、賞状と副賞を贈呈した。なお、審査にあたっての審査項目は下記の通りである。

口頭発表審査項目

- 講演要旨: 講演要旨は発表内容を反映して適切に作成されているか。
- 講演資料: スライドは専門領域の異なる参加者にもわかりやすく、見やすく、かつ発表時間に見合って適切に作成されているか。
- プレゼンテーション: 発表時に参加者にわかりやすく説明しているか。

- d) 研究の目的: 研究の背景と目的、先行研究との関係、研究の新規性あるいは有用性が明確になっているか。
- e) 研究成果: 価値のある成果が得られているか。
- f) 質疑応答: 質問等に対し、的確な応答・議論がなされたか。活発な討論がなされたか。
- g) 将来性: 研究内容について、将来の発展が期待できるか。

審査員

第 50 回構造活性相関シンポジウムに参加した 2022 年度常任幹事および幹事

<受賞者コメント>

KO-03

氏名 吉澤 竜哉 (よしざわ たつや)
所属 横浜市立大学 生命医科学研究科
演題 強化学習に基づく分子生成手法を用いたキナーゼ選択的阻害剤の設計

この度は名誉ある賞を頂戴し、大変光栄に存じます。ご評価くださいました先生方、並びに日本薬学会構造活性相関部会の先生方に厚く御礼申し上げます。発表では、研究内容に対して多様な観点からご意見をいただき、新たな気づきや学びを得ることができました。また議論を通じて研究に対する意欲も高まり、非常に意義深い時間となりました。

本研究では、強化学習に基づく分子構造生成手法を用いて、チロシンキナーゼを標的とした選択的阻害剤の設計を試みました。その結果、9種類のチロシンキナーゼに対する阻害活性、3種類の薬物動態に関わる性質、その他6種類の医薬品らしさに関わる性質からなる18種類の性質を最適化しながら、阻害剤候補化合物を設計することに成功しました。本研究が、分子構造生成手法開発の更なる発展に向けて貢献できることを期待しています。

最後に、本研究を遂行するにあたりご指導賜りました、横浜市立大学の寺山慧准教授、石田祥一特任助教、共同研究でお世話になりました先生方に深く感謝申し上げます。

強化学習に基づく分子生成手法を用いた キナーゼ選択的阻害剤の設計

○吉澤竜哉¹、石田祥一¹、佐藤朋広²、大田雅照²、
本間光貴²、寺山慧¹

(¹横浜市大・生命医、²理研)

E-mail: terayama@yokohama-cu.ac.jp

1. 背景と目的

医薬品候補化合物の探索は、標的タンパク質に対する選択性、薬物動態や毒性などの様々な基準を同時に満たすことが求められる。この探索は多目的最適化問題とみなすことができ、近年ではAI技術に基づく分子生成手法による分子構造の最適化が多数試みられている[1,2]。しかしそれらの研究では、創薬で現実的には考慮すべき標的タンパク質のホモログに対する阻害活性を十分に考慮できておらず、分子生成モデルで選択的阻害剤の設計が可能かは依然として不明であった。そこで本研究では、標的タンパク質及びそのホモログに対する阻害活性、薬物動態、毒性などが同時に最適化された選択的阻害剤を設計することを目的として、多目的最適化問題に適用可能な強化学習に基づく分子生成手法を開発した。本手法を用いて、9種類のチロシンキナーゼを標的とした選択的阻害剤の設計を試みた。

2. 方法

2.1. 分子生成手法

分子生成手法として、強化学習の一種であるMonte-Carlo tree searchとrecurrent neural networkで構成される、ChemTS[3]を使用した。ChemTSは、予め定義した報酬関数を最大化するように、分子の構造を生成する手法である。報酬関数として、Dscore[4]を導入した。Dscoreは、多目的最適化のための計算方法の一つであり、各最適化項目の重み付き幾何平均で表される。各最適化項目の値は、値の望ましさに応じて0から1に変換され、報酬計算に用いられる。

2.2. 最適化項目

最適化項目には、阻害活性、薬物動態、その他の性質の3カテゴリー、合計18項目を設定した。阻害活性の最適化項目として、epidermal growth factor receptor (EGFR)に対する阻害活性、8種類のEGFR類縁チロシンキナーゼ (receptor protein-tyrosine kinase erbB-2 (ERBB2)、Abelson tyrosine-protein kinase (ABL)、proto-oncogene

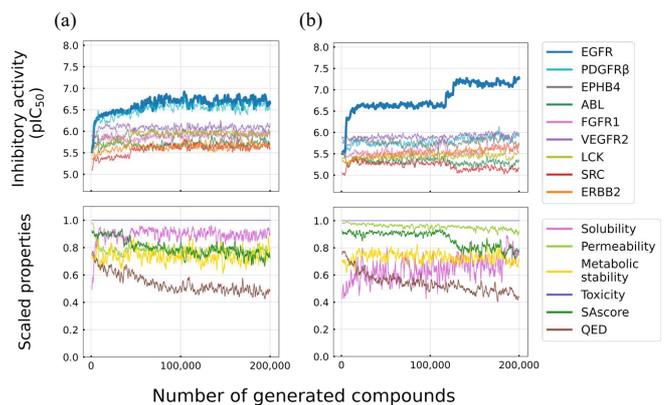


図 1. EGFR 阻害剤の設計における各最適化項目の値の推移 (a) Off-target タンパク質への阻害活性を考慮しない条件、(b) Off-target タンパク質への阻害活性を考慮した条件。上段は9種類のチロシンキナーゼに対する阻害活性、下段は薬物動態及びその他の性質の0~1に変換された値を示している。

tyrosine-protein kinase (SRC)、lymphocyte-specific tyrosine-protein kinase (LCK)、platelet-derived growth factor receptor beta (PDGFR β)、vascular endothelial growth factor receptor 2 (VEGFR2)、fibroblast growth factor receptor 1 (FGFR1)、ephrin type-B receptor 4 (EPHB4)に対する阻害活性の計9項目を設定した。薬物動態として、水溶性、膜透過性、代謝安定性の3項目を設定した。その他の性質として、synthetic accessibility score (SAscore)[5]、quantitative estimate of drug-likeness (QED) score[6]、molecular weight filter[1]、急性毒性、ToxAlerts filter[7,8]、ChEMBL structure filter[1]の計6項目を設定した。ここで、阻害活性、薬物動態、急性毒性の13項目については、LightGBMによる予測値を使用した。その他の5項目の値は、RDKitを用いて算出した。

2.3. 条件設定

本手法を用いて、9種類のチロシンキナーゼに対する選択的阻害剤の設計を試みた。選択性を考慮するため、標的タンパク質以外の8種類のタンパク質を off-target タンパク質とし、阻害活性を最

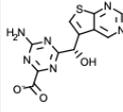
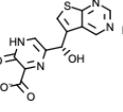
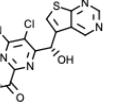
小化するように条件を設定した。また、off-target タンパク質を考慮しない条件も設定し、off-target タンパク質を考慮する条件との比較を行った。生成分子数は、全ての条件で 20 万個とした。

3. 結果とまとめ

図 1 に、EGFR 阻害剤の設計における各最適化項目の値の移動平均を示している。Off-target タンパク質を考慮しない条件においては、EGFR に対する阻害活性が上昇すると同時に、off-target タンパク質への阻害活性も上昇していた (図 1(a))。Off-target タンパク質を考慮した条件では、EGFR への阻害活性が上昇し、off-target タンパク質への阻害活性は上昇せず、10 倍以上の選択性が得られた (図 1(b))。阻害活性以外の項目の値は、off-target を考慮した条件の方が僅かに悪かったが、多くの項目が 0.8 程度、あるいはそれ以上で収束していた。以上の結果より、本手法が選択性を考慮した多目的最適化に有効であることが示唆された。

表 1 に、off-target を考慮した EGFR 阻害剤の設計における Dscore が高い生成分子の例、及び各最適化項目の予測値を示している。これらの分子は、EGFR への阻害活性 (pIC₅₀) が 7.6 以上と、十分な阻害活性を持つと予測された。また、EGFR への阻害活性と off-target タンパク質への阻害活性の最大値との差 (Δ_{BA} , pIC₅₀) は 2.1 以上であり、十分な選択性を持つと予測された。薬物動態及びその他の性質については、膜透過性が低いと予測された分子もあるものの、全体的に許容範囲内であった。これらの評価は予測値に基づく評価であ

表 1. EGFR 選択的阻害剤の設計における生成分子例と各最適化項目の予測値

生成分子			
Dscore	0.65	0.62	0.62
EGFR (pIC ₅₀)	7.7	7.6	7.7
Δ_{BA} (pIC ₅₀)	2.4	2.2	2.1
水溶性 (log(mol/L))	-2.4	-2.2	-2.4
膜透過性 (log(μ m/s))	0.99	0.50	0.99
代謝安定性 (1 hour remaining (%))	64	62	61
急性毒性 (LD ₅₀)	2.5	2.6	2.7
SAscore	4.0	4.0	4.1
QED	0.61	0.65	0.69

るため、予測モデルの適用範囲に関する解析を実施し評価の妥当性も併せて検証した。以上より、本手法によりある程度妥当であると考えられる有望な阻害剤候補を設計できたことがわかった。

EGFR 以外の 8 種類のチロシンキナーゼそれぞれを標的とした選択的阻害剤の設計では、全ての条件で標的タンパク質に対する結合親和性が上昇していた。一方で、選択性の大きさは条件ごとに異なっていた。PDGFR β 、EPHB4、ABL を標的とした条件では、EGFR を標的とした条件と同様に、10 倍以上の選択性が得られた。また、FGFR1、VEGFR2 を標的とした条件においても、10 倍程度の選択性が得られた。しかし、LCK、SRC、ERBB2 を標的とした条件では、選択性が得られなかった。

本研究では、多目的最適化に適用可能な分子生成手法を開発し、9 種類のチロシンキナーゼに対する選択的阻害剤の設計を試みた。その結果、18 種類の最適化項目のうちほとんどの項目を最適化することに成功し、9 種類のうち 6 種類のキナーゼに対して、有望な阻害剤候補の生成に成功した。一方で、本手法における生成分子の評価は、予測値を含んでいる点に留意する必要がある。今後は、分子生成モデルで設計された分子の信頼性をより向上させるために、予測モデルの適用範囲や不確実性を考慮した手法の開発を実施していく。

4. 謝辞

本研究は、AMED〈産学連携による次世代創薬 AI 開発 (DAIIA)〉の課題番号 JP22nk0101111、AMED BINDS の課題番号 JP22ama121023、文部科学省「富岳」成果創出加速プログラム「MD 駆動型プレジジョンメディスン」の支援を受けて実施されたものである。

5. 参考文献

- [1] Winter, R., *et al.*, *Chem. Sci.*, 10, 8016–8024. (2019)
- [2] Perron, Q., *et al.*, *J. Comput. Chem.*, 43, 692–703. (2022)
- [3] Yang, X., *et al.*, *Sci. Technol. Adv. Mater.*, 18, 972–976. (2017)
- [4] Cummins, D. J., *et al.*, *J. Med. Chem.*, 59, 6999–7010. (2016)
- [5] Ertl, P., *et al.*, *J. Cheminform.*, 1. (2009).
- [6] Bickerton, G. R., *et al.*, *Nat. Chem.*, 4, 90–98. (2012)
- [7] Sushko, I., *et al.*, *J. Chem. Inf. Model.*, 52, 2310–2316. (2012)
- [8] Non MedChem-Friendly SMARTS. <https://www.surechembl.org/knowledgebase/169485-non-medchem-friendly-smarts>, 2022; (accessed October 10, 2022).

<受賞者コメント>

KO-10

氏名 : 原田 隆平 (はらだ りゅうへい)
 所属 : 筑波大学計算科学研究センター
 演題 : 化合物の膜透過プロセスを紐解く自由エネルギー計算手法の開発

この度は、第50回構造活性相関シンポジウム優秀発表賞(口頭)という名誉ある賞を頂戴し、大変光栄に存じます。ご評価いただいた先生方、日本薬学会構造活性相関部会の関係者の方々、シンポジウム実行委員の方々に厚く御礼申し上げます。

本発表では、薬剤の生体膜透過性を評価する計算手法の開発について報告しました。生体膜を介して生体外から生体内へ薬剤を取り込み、疾患原因となる標的分子を活性阻害するプロセスにおいて、膜透過性の評価は重要です。分子動力学計算(MD)は、膜透過に伴う薬剤の構造変化を高い解像度で追跡できますが、現実的な計算コストで膜透過プロセスを抽出することは困難です。何故ならば、膜透過に要する時間スケールと比較し、MDが追跡できる時間スケールが極めて短いからです。MDが到達できない時間スケールで誘起される生命現象は生物学的レアイベントとみなされ、膜透過プロセスも該当します。そこで本研究では、膜透過プロセス抽出を可能とする計算手法を提案しました。具体的には、我々が開発を進めているレアイベントサンプリング法から、Parallel Cascade Selection MD (PaCS-MD) と Outlier Flooding Method (OFLOOD) を組み合わせたハイブリッドサンプリング法を提案し、膜透過性が既知である薬剤に対して膜透過プロセスを抽出しました。我々のレアイベントサンプリング法では、膜透過確率が高い条件の良い複数の初期構造から短時間MDを集散的に繰り返す「分散型MD」に基づき、膜透過プロセスを効率的に抽出します。最終的に、ハイブリッドサンプリング法から得られたトラジェクトリを用いてマルコフ状態モデルを構築し、膜透過に伴う自由エネルギーを計算することができました。自由エネルギーを参照することにより、薬剤が生体膜に対してどのような配向で膜透過するのか、膜透過の律速となる膜中心部でどのような構造変化をするのかなど、膜透過メカニズムの解明に有益な情報が得られます。さらに、マルコフ状態モデルから得られた固有値を用いて薬剤の膜透過係数を見積もったところ、実験値と高い相関($R^2 \sim 0.87$)を示しました。以上より、本手法は膜透過プロセスの解明や膜透過係数の見積もりに有用であることが示されました。将来展望として、本手法を環状ペプチドに適用し、合理的な薬剤デザインに貢献していければと考えています。

最後に、共同研究者としてお世話になりました筑波大学計算科学研究センターの森田陸離研究員、重田育照教授に感謝申し上げます。

化合物の膜透過プロセスを紐解く自由エネルギー計算手法の開発

○原田隆平¹, 森田陸離¹, 重田育照¹

(¹筑波大・計セ)

E-mail: ryuhei@ccs.tsukuba.ac.jp

1. 背景と目的

生体膜を介して生体外から生体内へ化合物を取り込み、疾患原因となる標的分子を活性阻害するプロセスにおいて、膜透過性の評価は重要である。分子動力学計算 (MD) は、膜透過に伴う化合物の構造変化を詳細に追跡できるが、通常の MD により現実的な計算コストで膜透過を抽出することは難しい。何故ならば、膜透過に要する時間スケールと比較して MD が追跡可能な時間スケールが極めて短いからである。通常の MD が到達不可能な時間スケールで誘起される生命現象は生物学的レアイベントとみなされ、膜透過プロセスも該当する。到達時間スケールの限界による探索不足を打破するため、構造サンプリング法が提案されてきた。本研究では、発表者らが開発を進めている構造サンプリング法 (レアイベントサンプリング法) [1-2] から、Parallel Cascade Selection MD (PaCS-MD) [3] と Outlier Flooding Method (OFLOOD) [4] を併用したハイブリッドサンプリング法 [5] に基づき、膜透過プロセスを抽出する。最終的に、得られるトラジェクトリを用いて膜透過に伴う自由エネルギープロファイルと膜透過係数を見積もり、膜透過性を議論する。[6]

2. 方法

提案手法は 4 つのステップからなる。第 1 ステップでは、PaCS-MD を適用して化合物の膜透過プロセスを抽出する。PaCS-MD は遷移確率が高い初期構造を選択して MD をリスタートするサイクルを繰り返すことにより、意図する遷移経路を抽出する。具体的には、化合物を脂質二重膜の上方へ配置し (図 1)、下方向へ向かう膜透過プロセスを抽出した。膜透過する可能性が高い初期構造を各サイクルで選択するため、化合物の重心座標 (z_{COM}) を指標とし、値が小さい初期構造を選択して MD をリスタートするサイクルを繰り返した。サイクルを繰り返すごとに z_{COM} は小さくなり、十分サイクルが経過すれば上方から下方へ向かう化合物の膜透過プロセスを抽出できる。しかし、1 回だけの PaCS-MD ではとりうる膜透過経路を効率的に探索することは難しい。そこで第 2 ステップにおいて、OFLOOD を併用したハイブリッドサンプリング法に基づき、広い構造空間における膜

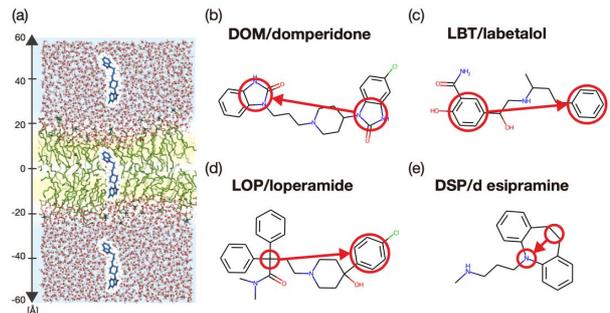


図-1 膜透過プロセス抽出に用いたシステム

透過プロセスの探索を試みた。

第 2 ステップでは、PaCS-MD により生成された初期膜透過プロセスをもとに周辺に存在するとりえる膜透過経路を探索するため、OFLOOD を併用した。OFLOOD は MD が生成するトラジェクトリを低次元の構造空間へ射影したのち、クラスタリングにより密な領域と疎な領域に分類し、疎な領域に属する MD スナップショットを初期構造に選択して MD をリスタートするサイクルを繰り返す。あえて確率的に出現確率が低い疎な分布 (はずれ値) に該当する構造群に着目して MD をリスタートすることで、出現確率が低い構造遷移を効率的に促進できる。

第 3 ステップでは、PaCS-MD と OFLOOD で探索した構造空間において膜透過に伴う化合物の自由エネルギープロファイルを計算する。具体的には、探索空間をカバーする複数の代表構造を選択し、長時間 MD を実行して生成されるトラジェクトリをもとにマルコフ状態モデルを構築する。信頼性の高いマルコフ状態モデルが構築できれば定常分布が求まるため各状態の存在確率が分かり、自動的に膜透過プロセスに伴う自由エネルギープロファイルが計算できる。

第 4 ステップでは、マルコフ状態モデルが提供する固有値を用いて時定数を見積り、膜透過係数を計算する。具体的には、マルコフ状態モデルを構築する際に見積もられる遷移行列を対角化して得られる固有値をもとに時定数を見積り、膜透過係数を算出する。

本手法の妥当性を検証するため、先行研究 [7] で調べられている 4 つの化合物 (図 1: DOM, LBT, LOP, DSP) に対して脂質二重膜における膜透過プロセスを抽出した。最終的に、膜透過に伴う自由エネルギープロファイル及び膜透過係数を計算し、実験値との相関関係を調べた。

3. 結果とまとめ

図 2 は PaCS-MD を適用して抽出した 4 つの化合物の膜透過プロセスを示す。PaCS-MD において、サイクルあたり 10 初期構造、初期構造あたり 100 ps の MD を実行し、200 サイクルまで繰り返した。図 1 が示すように、サイクルが進むにつれて化合物の重心座標が小さくなっていき、脂質二重膜の上方から下方へ化合物が膜透過していく様子がうかがえる。

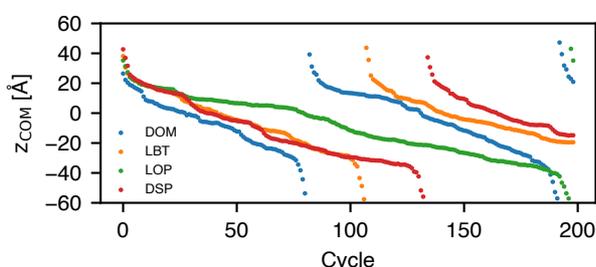


図-2 PaCS-MD による膜透過プロセスの抽出

次に、PaCS-MD により生成した初期膜透過プロセスをもとに OFLOOD を適用して広域的な構造空間を探索した。ここで、 z_{COM} に追加して先行研究 [7] で提案された化合物の膜表面に対する配向を記述する変数 (Δz) を考慮し、2 次元構造空間を定義した。サイクルあたり 50 個のはずれ値を初期構造に選択し、初期構造あたり 100 ps の MD を実行して 100 サイクルまで構造探索を繰り返した。十分サイクルを繰り返すことにより、2 次元構造空間を広く探索することができた。

更に、PaCS-MD と OFLOOD により探索した 2 次元構造空間から代表構造を選択し、OFLOOD に

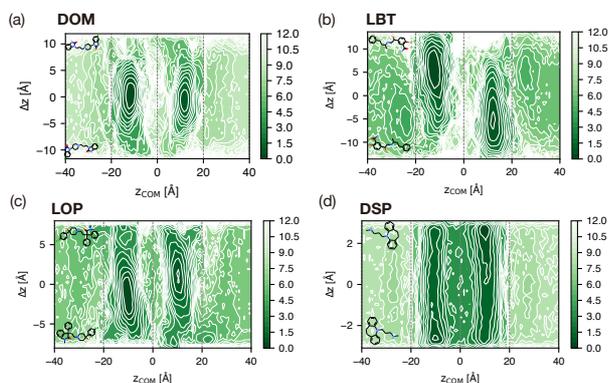


図-3 膜透過に伴う自由エネルギー

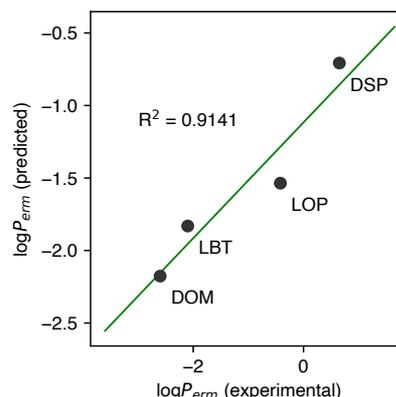


図-4 実験値と計算値の相関 (膜透過係数)

よりプロダクションランを実行した。具体的には、OFLOOD のサイクルあたり 50 個のはずれ値を初期構造に選択し、初期構造あたり 50 ns の長時間 MD を実行し、5 サイクルまで繰り返した。図 3 は長時間 MD により生成されたトラジェクトリからマルコフ状態モデルを構築して計算した 2 次元構造空間における化合物の膜透過に関する自由エネルギープロファイルである。図 3 より、化合物が膜透過する経路を考察することができる。

最終的に、化合物に対して構築したマルコフ状態の固有値から膜透過係数を算出した。図 4 は、化合物に対する膜透過係数の計算値と実験値との相関プロット ($R^2 \sim 0.91$) である。図 4 から明らかなように、定量的な一致は難しいものの、定性的に高い相関を示した。

以上より、本手法を適用することで化合物の膜透過に伴う自由エネルギープロファイル及び膜透過係数を定性的に算出できる。計算値の定量的な議論に関しては、計算の精度向上を目指すことにより将来的に改善していく予定である。

4. 謝辞

本研究は、筑波大学計算科学研究センターが推進する学際共同利用に基づき実施された。

5. 参考文献

- [1] R. Harada, T. Baba, Y. Takano, and Y. Shigeta, *Phys. Chem. Chem. Phys.*, **17**, 6155 (2015)
- [2] R. Harada, *Bull. Chem. Soc. Jpn.*, **91**, 1436 (2018)
- [3] R. Harada and A. Kitao, *J. Chem. Phys.* **139**, 035103 (2013)
- [4] R. Harada, T. Nakamura, Y. Takano, and Y. Shigeta, *J. Comput. Chem.* **36**, 97 (2014)
- [5] R. Harada and Y. Shigeta, *J. Chem. Theory. Comput.*, **15**, 680 (2019)
- [6] R. Harada, R. Morita, Y. Shigeta, *J. Chem. Inf. Model.*, **63**, 259 (2023)
- [7] M. Badaoui et al., *J. Phys. Chem. B*, **122**, 11571 (2018)

<受賞者コメント>

KP-03

氏名 高谷 大輔 (タカヤ ダイスケ)
所属 大阪大学大学院薬学研究科量子生命情報薬学分野
演題 インシリコスクリーニング(LBDD)のための最近のフィンガープリントの性能検討

この度、第 50 回薬学会構造活性相関シンポジウム SAR Presentation Award (ポスター) を賜りまして、大変光栄に存じます。

理論創薬分野において、新規活性化合物等を探索する目的で行われるインシリコスクリーニングでは、現在でも Ligand-based drug design (LBDD)及び Structure-based drug design (SBDD)が基盤技術として頻繁に用いられています。これらは研究スタート時点で利用できる情報によって使い分けられますが、今回は前者のアプローチに分類される類縁化合物探索法で使用される化合物特徴をビット列で表現したフィンガープリント(FP)のオープンソース実装を中心にその性能及び傾向を検討し、その FP 選択がスクリーニングの精度に及ぼす影響を調べた研究となります。ベンチマークセットとしては 81 ターゲットからなる DEKOIS 2.0、インシリコスクリーニングの評価指標には ROC や EF 等を用い、結果として、現在デファクトスタンダードである ECFP の RDKit 実装である Morgan や、MinHash を用いるのが良いということが確認されました。一方で、CDK で実装された FP を用いた場合でも性能が良くなるケースがしばしば確認されました。また、EF20 等では、様々な FP でスクリーニング性能向上が見られたことから、実際の *in vitro* スクリーニングのスループット量に応じ、様々な FP の検討が実際にヒット化合物を得る事につながる事が示されました。一方で、事前情報としてどのような情報を与えられた場合に、適切な FP を選択できるかが課題として残っております。今後、機械学習手法などを組み合わせてそのような課題を解決できるよう、研究室メンバーと協力し、さらなる検討を進めていきたいと考えております。

最後に、昨年続き厳しい社会情勢の中、リモート開催によって研究発表の機会をいただきましたこと、実行委員会の先生方に心より感謝申し上げます。

インシリコスクリーニング(LBDD)のための最近のフィンガープリントの性能検討

○高谷大輔¹、田雨時¹、福澤薫¹
(¹ 阪大院・薬)

E-mail: takaya-d@phs.osaka-u.ac.jp

1. 背景と目的

計算機を用いた創薬研究は企業やアカデミアを問わず幅広く普及している。研究環境に応じて計算機資源やソフトウェアライセンス数の規模は異なるが、世界中の計算化学者によるオープンソースのソフトウェア開発が盛んに行われ使用されている。本研究では、新規のインシリコスクリーニング基盤構築にあたり、利用できるライブラリ等の性能検討を行なった。

このうち Ligand-based drug design (LBDD) に分類される手法は、既存の活性化合物との類似した化合物は同等程度の活性が見込めるという仮説に従い、類似性を指標としてデータベース探索等を行い、類縁化合物をスクリーニングする手法である。化合物間の類似性指標である Tanimoto 係数 (Tc)¹ などの算出に使用される化合物特徴のビット列表現データをフィンガープリント (FP) と呼ぶ。使用した FP により、Tc 等の値が異なり検出される化合物の傾向が変化するため、FP 選択には注意が必要であることが知られている。また近年、多くの FP が提案され、オープンソースによる実装も公開されている。本検討では今後自身が LBDD を実施する時、どのような FP を選択すべきかの指針を考える上で、オープン実装の FP を中心にその性能及び傾向を調査した。

2. 方法

本検討では活性化合物とデコイ構造の判別能力に注目した。データセットは DEKOIS2.0² を用いた。これは予め前処理された活性既知化合物及びデコイ構造からなるデータセットであり、合計で 81 ターゲット存在する。RDKit³ (2022.03.03; python 3.9) 及び CDK⁴ (version 2.8; openjdk 11.0.15) のライブラリから利用可能な FP 計算関数などを調査し、それらを用いることでベンチマーク中の化合物の FP を計算した。CDK の FP 計算モジュールの呼び出しは JPype⁵ ライブラリを経由して行った。リガンドベースのスクリーニング性能の指標として Enrichment Factor 及び ROC などの代表的な指標を用いた。これらの性能は RDKit の組み込み関数を使用して計算した。

3. 結果とまとめ

本研究で採用したライブラリに存在する FP 計算関数を API リファレンス等から調査し、使用可能な FP の一覧を調べた。(表 1)

表 1: 検討した FP 名、または FP 計算関数の一部

RDKit	CDK	
MACCS FP	Fingerprinter	LINGO FP
Avalon FP	ExtendedFingerprinter	KlekotaRoth FP
Morgan FP	graph only FP	ShortestPathFingerprinter
MinHash FP	MACCS FP	AtomSignature-based FP
AtomPair FP	PubChem FP	SubstructureFingerprinter
Topological Torsion FP	Esate FP	Atom Pair 2D FP
Extended reduced graph	Hybridization Fingerprinter	

現時点の結果では、概ね ECFP⁶ 相当である Morgan FP の性能が良好だった。これは既存の報告と合致する点が多く妥当であると思われるが、検討の詳細は当日のポスターで報告する。

4. 謝辞

本研究の一部は AMED BINDS (JP22ama121030) の活動の一環として行われた。また多くのインターネット上のサンプルコードを参考にした。

5. 参考文献

- Performance of Similarity Measures in 2D Fragment-Based Similarity Searching: Comparison of Structural Descriptors and Similarity Coefficients. Chen, X. et al. *J. Chem. Inf. Comput. Sci.* 2002, 42, 6, 1407-1414
- Evaluation and Optimization of Virtual Screening Workflows with DEKOIS 2.0 – A Public Library of Challenging Docking Benchmark Sets. Bauer, Matthias R. et al. *J. Chem. Inf. Model.* 2013, 53, 6, 1447-1462,
- RDKit: Open-source cheminformatics. <https://www.rdkit.org>
- Chemistry Development Kit. <https://cdk.github.io/>
- JPype. <https://jpype.readthedocs.io/en/latest/>
- Extended-Connectivity Fingerprints. Rogers, D. et al. *J. Chem. Inf. Model.* 2010, 50, 5, 742-75

<受賞者コメント>

KP-06

氏名 市丸 嘉 (いちまる よしみ)

所属 湘南医療大学薬学部

演題 エポキシドを導入したインディールビン誘導体 Epox/Ind の抗がん活性に関する構造活性相関

この度は、第 50 回構造活性相関シンポジウム SAR Presentation Award を賜り、大変光栄に存じます。審査員の先生方、ならびに日本薬学会構造活性相関部会の先生方に厚く御礼申し上げます。

本研究では、様々な生物活性が知られる天然由来化合物「インディールビン」の抗がん活性を増強するためにオキシランを導入した「Epox/Ind」と、その臭素置換誘導体の構造活性相関について検討し、新規抗がん剤候補化合物として Epox/5-Br-Ind を見出しました。また、Epox/5-Br-Ind は Epox/Ind と比較して、細胞溶解液中でのオキシランの安定性が向上していることが、細胞傷害活性の増強に寄与していることも明らかにしました。本研究の活性実験は whole-cell レベルで、分子レベルでの作用機序解明が今後の課題ですが、今回の受賞を契機に、我々のインディールビン化合物ライブラリーが多くの方々の興味を引き、医薬品として多くの方々に役立つことを願い、これからも研究に励んでいきます。

最後に、本研究を進めるにあたりご指導を賜りました、日本大学薬学部宮入伸一教授ならびに、研究に関わってくれた卒業生諸氏に深く感謝申し上げます。

エポキシドを導入したインディルビン誘導体 EpoX/Ind の抗がん活性に関する構造活性相関

○市丸嘉¹、加藤紘一¹、栗原正明¹、宮入伸一²
(¹ 湘南医療大・薬、² 日本大・薬)

E-mail: yoshimi.ichimaru@sums.ac.jp

1. 背景と目的

藍の生薬に豊富に含まれるビスインドール化合物「インディルビン(1)」は、抗がん活性をはじめ、さまざまな医薬品様作用が知られている[1]。これまでに多くのインディルビン誘導体が開発されており、特に3'位-オキシム化誘導体 Indox(2)は高い抗がん活性を示すことが明らかにされている。インディルビン誘導体の生物活性は、ATP拮抗型のキナーゼ阻害活性に由来すると考えられており、サイクリン依存性キナーゼ(CDK)[2]やグリコーゲン合成酵素-3β(GSK-3β)[3]のATP結合部位にインディルビン誘導体が結合したX線結晶構造が報告されている。我々は、キナーゼのATP結合部位近傍に存在するアミノ酸残基との共有結合形成を期待して、Indoxにオキシラン(エポキシド)を導入したEpoX/Ind(3)を新たに開発した[4]。さらに、Indoxのインドール水素を臭素に置換することで抗がん活性が向上することに着目し[1]、本研究では、EpoX/Indの5, 5', 6, 6', 7位に臭素を1つないし2つ置換したEpoX/Br-Ind誘導体とその類縁化合物を合成し、がん細胞傷害活性に着目した構造活性相関の検討と、抗がん活性に影響を与える因子の検討に取り組んだ。

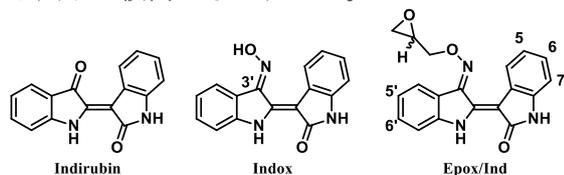
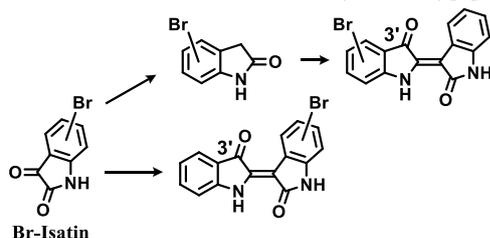


図1 EpoX/Ind誘導体の構造式

2. 方法と結果

臭素置換インディルビン誘導体は、イサチン誘導体を出発原料として合成した(Scheme)[5]。



Scheme Br-Indirubinの合成[5]

合成したIndirubin誘導体は、ヒドロキシルアミン塩酸塩と反応させてIndox誘導体に、更にエピプロモヒドリンと反応させてEpoX/Ind誘導体へと導き、計42の誘導体を合成した。

合成した誘導体は、ヒト肝がん細胞HepG2に対する細胞傷害活性を、alarBlue® assayで評価した。結果、5位に臭素が置換したEpoX/Ind誘導体(EpoX/5-Br-Ind)が最も強い活性を示した。HepG2の50%増殖阻害濃度(IC₅₀値)を抜粋してTableに示した。

Table インディルビン誘導体の細胞傷害活性

IC ₅₀ (μM)	Indirubin	Indox	EpoX/Ind
none Br	> 100	15	1.7
5-Br	> 100	6.9	0.62
6-Br	> 100	5.3	>100
7-Br	> 100	4.8	37
5'-Br	> 100	11	3.3
6'-Br	> 100	14	4.5
5,6-Br ₂	> 100	4.4	1.9

また、EpoX/Indは、オキシランが加水分解すると細胞傷害活性が低下することが分かっているため[4]、HepG2細胞溶解液中での加水分解速度(v)を検証した。結果、EpoX/5-Br-Indの基質濃度が高くなると、加水分解速度が低下することが分かり(図2)、EpoX/5-Br-Indの高い細胞傷害活性には、オキシランの安定性が関与していることが示唆された。追加の検証によって、EpoX/5-Br-Indがエポキシドヒドrolラーゼの自殺基質として作用することが示唆された。

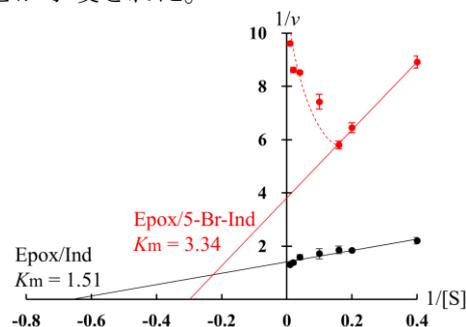


図2 Lineweaver-Burk plot

さらに、Epoxy/Ind の臭素の置換位置が、サイクリン依存性キナーゼ-2(CDK-2)[2]の ATP 結合部位親和性に及ぼす影響について GOLD を用いたドッキングシミュレーションを行った。リガンドは、Epoxy/Ind の単結晶 X 線結晶構造(図 3)をもとに量子科学計算で構造最適化した。結果、Epoxy/5-Br-Ind と Epoxy/6-Br-Ind のように細胞傷害活性が大きく異なる誘導体間(0.67 μM vs >100 μM)でも、GOLD ドッキングスコアに差はなく(69.31 vs 69.01)、Epoxy/Ind 誘導体の抗がん活性を、CDK-2 への親和性で説明することはできなかった。

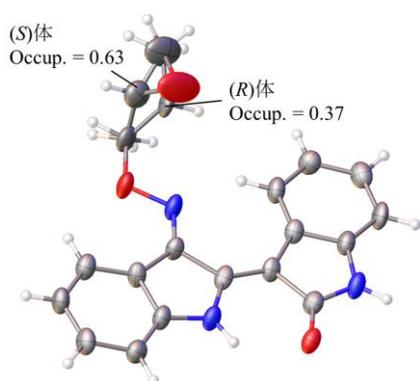


図 3 Epoxy/Ind の X 線結晶構造

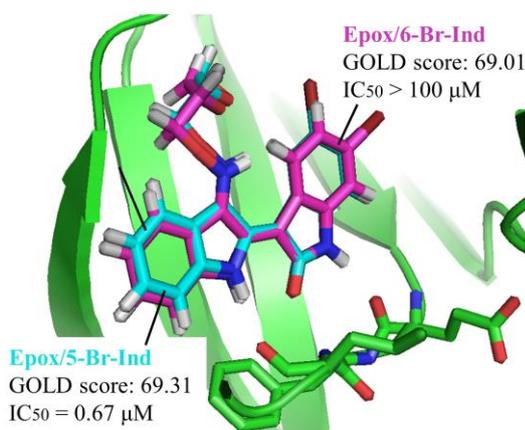


図 4 ドッキングシミュレーション

3. 考察とまとめ

本研究では天然物由来のビスインドール化合物インディルビンの抗がん活性に着目し、標的タンパク質のアミノ酸残基との共有結合形成を期待してオキシランを導入した Epoxy/Ind と、その臭素置換体を含む計 42 化合物について、がん細胞傷害活性を検討した。結果、最も活性の高い Epoxy/5-Br-Ind を見出し、5 位臭素置換によって Epoxy/Ind の生体内オキシランの安定性が向上することを明らかにした。

CDK-2(1E9H) の ATP 結合部位に対する Epoxy/Br-Ind 誘導体の親和性をシミュレーション

したところ、ドッキングスコアと IC₅₀ 値の間に相関は認められなかった。我々は細胞増殖に関連するキナーゼカスケードに対するインディルビン誘導体の活性を抗体アレイによって網羅的に解析し、これまでインディルビン誘導体の主な標的キナーゼと言われていた CDKs の他にも、様々なキナーゼに対してインディルビン誘導体が阻害活性を示すことを見出している[6]。このことから、ドッキングシミュレーションの結果は、Epoxy/Ind が CDKs でない標的キナーゼを阻害して、がん細胞の増殖を抑制していることを示唆している。今後、Epoxy/Ind の活性増強と、標的キナーゼの特定に取り組んでいく予定である。

4. 参考文献

1. H. Wang, Z. Wang, C. Wei, J. Wang, Y. Xu, G. Bai, Q. Yao, L. Zhang, Y. Chen, *Eur J Med Chem* **223** (2021) 113652.
2. T. G. Davies, P. Tunnah, L. Meijer, D. Marko, G. Eisenbrand, J.A. Endicott, M.E.M. Noble, *Structure* **9** (2001) 389.
3. J.A. Bertrand, S. Thieffine, A. Vulpetti, C. Cristiani, B. Valsasina, S. Knapp, H.M. Kalisz, M. Flocco, *J Mol Biol* **333** (2003) 393.
4. Y. Ichimaru, H. Saito, T. Uchiyama, K. Metori, K. Tabata, T. Suzuki, S. Miyairi, *Bioorg Med Chem Lett* **25** (2015) 1403.
5. Y. Ichimaru, T. Fujii, H. Saito, M. Sano, T. Uchiyama, S. Miyairi, *Bioorg Med Chem* **25** (2017) 4665.
6. Y. Ichimaru, M. Sano, I. Kajiwara, T. Tobe, H. Yoshioka, K. Hayashi, H. Ijichi, S. Miyairi, *Transl Oncol* **12** (2019) 1574.

<受賞者コメント>

KO-01

氏名 星野 小百合 (ほしの さゆり)

所属 横浜市立大学理学部

演題 短時間 MD シミュレーションによるリガンド-タンパク質結合親和性推定手法の検討

この度は名誉ある賞を頂き、大変光栄に思います。審査員の先生方、並びに日本薬学会 構造活性相関部会の先生方に心より御礼申し上げます。本シンポジウムでは、沢山の専門家の方々と議論を通して多くの学びを得ることができました。

本研究は、短時間 MD シミュレーションを実施後、リガンド RMSD (Root Mean Square Deviation) やリガンド-タンパク質相互作用エネルギーといった安定性評価指標について網羅的に比較し、それらの精度を検証しました。さらに、計算コスト削減を目指し、強化学習手法の1つである *m*-Best Arm Identification (*m*-BAI)の枠組みを導入しました。結果、本研究におけるデータセットの範囲では、リガンド-タンパク質相互作用エネルギーが短時間での評価に有効であること、*m*-BAIを導入することで、計算量を削減した場合であっても精度を維持できることが分かりました。今後は、実際の創薬分野での問題に即したより大規模なデータセットを対象に検証と手法開発を展開し、さらなる精度向上と計算コストの削減を目指したいと考えています。

最後に、本研究を進めるにあたりご指導を賜りました、寺山慧准教授をはじめ、横浜市立大学生命情報科学研究室、共同研究者の皆様に、心より感謝申し上げます。

短時間 MD シミュレーションによるリガンド-タンパク質結合親和性推定手法の検討

○星野小百合¹、石田祥一²、河東田道夫³、隅田真人⁴、奥野恭史⁵、寺山慧^{1,2}

(¹横浜市大・理、²横浜市大院・生命医、³高度情報科学技術研究機構、⁴理研 AIP、⁵京大院・医)

E-mail: terayama@yokohama-cu.ac.jp

1. 背景と目的

リガンド-タンパク質間の結合親和性推定において、予測精度と計算コストはトレードオフの関係にあり、両立することは難しい。その両立を目指した推定手法の1つとして、ドッキング後に短時間分子動力学(MD)シミュレーションを実行し、安定性を評価する手法が知られている[1][2]。しかし、様々な安定性評価の指標が提案されている中で、どの指標が親和性推定に適するののかについては、議論の余地がある。[3][4]そこで、本研究では、先行研究[1]を踏襲して短時間 MD シミュレーションを実施し、リガンド RMSD やタンパク質-リガンド相互作用エネルギー等の安定性評価の指標について網羅的に比較し、それらの精度を検証する。さらに、計算コスト削減のための強化学習の導入についても検討する。

2. 方法

2-1. データセット

本研究では、先行研究[1]で用いられていたタンパク質のうち 6 種類のタンパク質(PDBID: 2of2, 1w7x, 2b8t, 1mv9, 1vso, 3h15)を使用した。また、リガンドに関しても先行研究[1]と同様に、タンパク質ごとに、DUD-E データセットより active および decoy それぞれ 5 種類ずつ、計 10 種類のリガンドを選択した。

2-2. ドッキングおよび MD シミュレーション

まず、MD シミュレーションの入力ファイルを得るためにドッキングシミュレーションを実行した。ドッキングソフトウェアは AutoDock Vina (v1.2.3)を使用した。次に、出力結果として得られた最安定な結合ポーズを入力として、平衡化 MD シミュレーションを行った後、100 ns のプロダクト MD シミュレーションを、初速度をランダムに変えて 20run ずつ実施した。その出力として MD トラジェクトリが得られ、これを安定性評価指標の検討に用いる。MD シミュレーションには GROMACS (2021.4)を使用した。

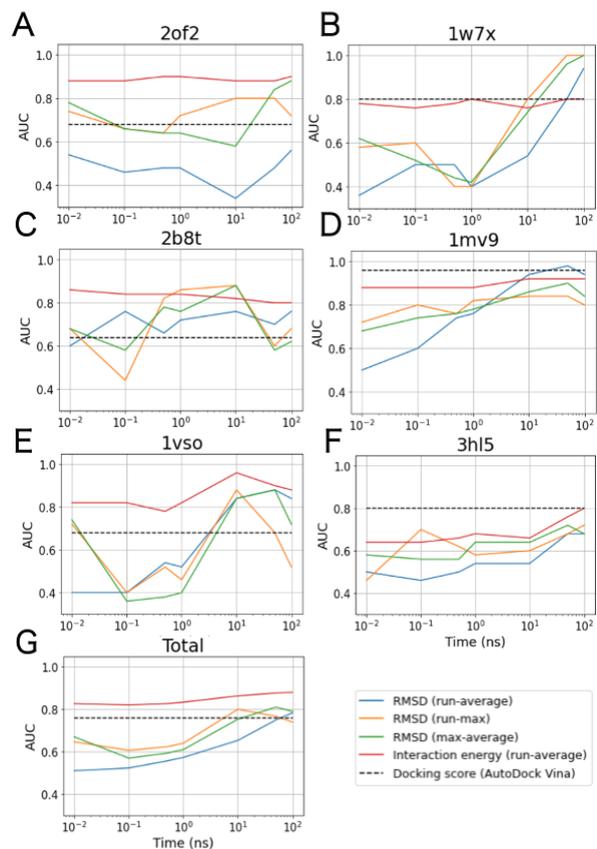


図-1 リガンド RMSD およびタンパク質-リガンド相互作用エネルギーによる安定性評価の結果

2-3. 安定性評価指標

MD シミュレーションによって得られたトラジェクトリから、安定性評価のための指標としてリガンド RMSD およびタンパク質-リガンド相互作用エネルギーを計算した。MD シミュレーション時間が与える影響を検討するために、計算開始から 10 ps, 100-ps, 500 ps, 1 ns, 10 ns, 50 ns, 100 ns それぞれの時間で上記指標を計算した。これらの指標を用いて、active と decoy の識別が正しく行われているか否かを ROC-AUC 値によって評価した。

リガンド RMSD については、リガンドごとに各 run の平均値からさらに平均した値(run-average)、各 run の最大値からさらに最大をとった値(run-max)、各時間ステップでの run 間の最大値からさらに最大をとった値(max-average)の 3 つの異なる値を採用した。

2-4. *m*-Best Arm Identification

各リガンドに対し複数回 MD シミュレーションを実施すると多大な計算コストがかかる。既に類似の設定で、最も良い候補を探索する際に、強化学習の一種である Best Arm Identification のアルゴリズムを用いることで、無駄な計算を省き計算高速化が可能であることが報告されている[5]。本研究でも、この枠組みは有効と期待されるが、ここでは複数(5 つの active リガンド)を効率よく探索する必要がある。そこで、上位 *m* 個を効率よく探索する枠組み *m*-Best Arm Identification のアルゴリズムである UGapE[6]を実装し、計算効率向上の検証を行った。

3. 結果

それぞれの系において、異なる指標を用いて評価をした際の結果を図 1 に示す。また、全タンパク質間での平均 ROC-AUC 値の推移を確認した(図-1G)。本データセットの範囲では、タンパク質-リガンド相互作用エネルギー(ROC-AUC: 0.88)、リガンド RMSD (ROC-AUC: 0.77)、ドッキングスコア(ROC-AUC: 0.76)の順に精度が高かった(図-1G)。リガンド RMSD については、値の算出方法の違いによる精度の変動は比較的小さく、10 ns で平均 0.73 の ROC-AUC 値となり、100 ns (ROC-AUC: 0.77)と比較して劣らない精度を示した。しかし、10 ns より短時間で評価した際にはその精度は低下した。一方、タンパク質-リガンド相互作用エネルギーについては、短時間(0.01 ns)であっても平均 0.83 の高い ROC-AUC 値が得られた(図-1G)。

次に、*m*-BAI を 1vso (100 ns)対象に適用した場合の結果を図 2 に示す。ここで、US (Uniform Sampling)は、各リガンドに対し同数の run を実施した場合の結果である。また、UGapE の *a* は探索パラメータであり、adaptive は探索パラメータを自動的に推定する手法[6]を用いている。図 2 の US より、各リガンドに対し十分な run 数を実施した場合は比較的高精度であるが、run 数を減らすことで ROC-AUC 値は 0.70 を下回っており、精度が低下していることが分かる。一方、Top5 (*m*=5)を効率的に探索する UGapE を適用した場合、必ずしも大きな計算効率向上はしていないものの、run 数を減らした際であっても、精度を維持できてい

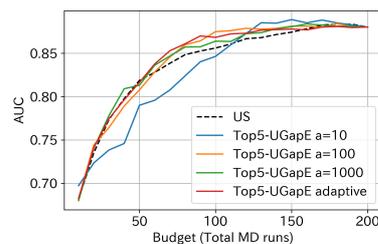


図-2 *m*-BAI を適用した際の安定性評価の結果

ることが確認できた(図 2)。

4. まとめと展望

本研究では、リガンド RMSD やタンパク質-リガンド相互作用エネルギーといった安定性評価指標を網羅的に比較することで、計算コストを抑えた有効な評価手法について検討した。その結果、短時間での評価においてはタンパク質-リガンド相互作用エネルギーが有効であることが分かった。さらに、*m*-BAI を適用させることで、run 数を減らした場合の精度向上が見込めることが分かった。本研究においては、decoy を 5 種類と限定していたが、今後は decoy の数を増やすことで、実際の創薬分野での問題に即した条件で手法を展開し、さらなる精度向上を目指す。

5. 謝辞

本研究は、AMED〈産学連携による次世代創薬 AI 開発 (DAIIA)〉の課題番号 JP22nk0101111、AMED BINDS の課題番号 JP22ama121023、文部科学省「富岳」成果創出加速プログラム「MD 駆動型プレジジョンメディシン」の支援を受けて実施されたものである。

6. 参考文献

1. H. Guterres and W. Im, *J. Chem. Inf. Model.*, 2020, 60(4), 2189-2198.
2. S. Kori *et al.*, *Bioorg. Med. Chem.*, 2021, 52(15), 116500
3. S. Mahmud *et al.*, *Brief. Bioinform.*, 2021, 22(2), 1402-1414
4. R. Kumari *et al.*, *J. Biomol. Struct. Dyn.*, 2022, 1-17
5. K. Terayama *et al.*, *Bioinformatics*, 2018, 34(5), 770-778
6. V. Gabillon *et al.*, *Advances in Neural Information Processing Systems*, 2011, 24, 2222-2230

<受賞者コメント>

KO-15

氏名 田中 良樹 (たなか よしき)
所属 株式会社アグロデザイン・スタジオ
演題 タンパク質立体構造情報を利用した新規 ALS 阻害型除草剤の探索

この度は、第 50 回構造活性相関シンポジウム実行委員会特別賞を頂き、誠にありがとうございます。ご評価下さいました審査員の先生方、並びに日本薬学会構造活性相関部会の先生方に厚く御礼申し上げます。

当社では、科学的エビデンスが明確な分子標的農薬を構造ベースでデザインし、安全性と薬効を両立する農薬の研究開発を行っています。本発表では、アセト乳酸合成酵素 ALS を標的とした新規除草剤の構造ベース創農薬の流れとその成果に関して紹介させていただきました。ALS 阻害化合物については、ハイスループット結晶構造解析と大規模ドッキングシミュレーションを組み合わせることで、既存薬とは化合物骨格が異なるヒット化合物を複数種類得ております。今後は、複合体構造情報を利用した合成展開を進めると同時に、安全性や土壌試験などを実施して製品化へ向けた研究開発を行ってまいります。

最後に、共同研究者である株式会社 Preferred Networks、理化学研究所放射光科学研究センター、高エネ機構構造生物学研究センターの皆様にご心より感謝申し上げます。

タンパク質立体構造情報を利用した新規 ALS 阻害型除草剤の探索

○田中良樹¹、氏原一哉²、坂井直樹³、山田悠介⁴、武本瑞貴²、森脇寛智¹、佐藤 匡史¹、JK Stanfield¹、平田邦生³、竹下浩平³、引田理英⁴、石谷隆一郎²、加藤龍一⁴、千田俊哉⁴、山本雅貴³、力丸健太郎²、西ヶ谷有輝¹

(¹株)アグロデザイン・スタジオ、²株Preferred Networks、³理化学研究所 放射光科学研究センター、⁴高エネ機構・構造生物学研究センター)

E-mail: nishigaya@agro.design

1. 背景と目的

近年、農薬は安全性基準の厳格化に起因する開発期間の長期化や開発費の増大が課題となっている。その解決法の一つとして分子標的農薬が有力である。この農薬は防除対象生物に特異的なタンパク質を標的として開発されるため、合理的な化合物設計が可能になる。さらに、防除対象に選択性が高く、ヒトや環境に対して低リスクの化合物となることが期待される。

本研究では、アセト乳酸合成酵素 (Acetolactate synthase : ALS) を標的とした新規除草剤の開発を目指している。ALS は植物の分岐鎖アミノ酸の生合成経路において重要な酵素であり、動物には存在しない。この酵素を標的とした農薬は、低薬量で高い除草作用を示すため、1970 年代からスルホニルウレア系除草剤をはじめ多数の薬剤が開発されている。一方で、抵抗性雑草が長期間の使用に伴い問題となっており、抵抗性雑草にも薬効を示す新規 ALS 阻害剤が望まれている。

2. 方法

抵抗性雑草に薬効を示す新たな ALS 阻害剤を創出するため、新規骨格を有する化合物の取得を指向し、タンパク質結晶構造解析とドッキングシミュレーションを利用した薬剤探索を行った。

まず、多数の ALS-既存薬の結合様式を改めて精査する必要があると考え、ALS タンパク質精製系を自社内で構築し、構造解析などの実験に使用することにした。精製したタンパク質について SPring-8 リガンドスクリーニングパイプラインと KEK スクリーニングシステムを使用し、40 種類以上存在する既存薬と ALS の複合体結晶構造の解析を試みた。ドッキングシミュレーション (PFN の AI 創薬プラットフォーム) の際には、精度向

上のため、結晶構造を再現するようにシミュレーション条件の検討を行った。その後、市販化合物約 1000 万種に対するドッキングシミュレーションを実施した。この結果得られた各化合物のドッキングスコアやアミノ酸残基との水素結合の情報などをもとに市販化合物を選抜し入手した。続いて、ALS 酵素活性アッセイと植物生育阻害試験にてリード化合物の選抜を行った。さらにヒットバリデーシンのための合成展開として、部分構造の置換等を施した化合物を合成し、初期ヒット化合物同様に活性試験を行った。

3. 結果とまとめ

新規・既知の ALS - 既存薬複合体結晶構造を利用したドッキングシミュレーション薬剤探索により、新規骨格のリード化合物創出に成功した。

まず複数の PDB 未登録 ALS - 既存薬複合体結晶構造の決定に成功した。次に新規・既知の結晶構造を利用し、ドッキング条件検討を進めた結果、結晶構造における化合物結合位置を再現する条件決定に成功した。決定した条件を用いて化合物選抜を実施し、268 化合物を購入した。そのうち 35 種に酵素阻害活性が認められ、その中に既存薬と異なる骨格を有する化合物が含まれていた。さらに植物生育試験での薬効と合成展開による酵素阻害活性の向上を確認した。また代表的な抵抗性変異型 ALS に対して、新規骨格化合物の阻害活性が低下しないことを確認している。

<受賞者コメント>

KO-15

氏名 渡邊 友里江 (わたなべ ゆりえ)

所属 昭和大学薬学部

演題 歯周病菌由来 PgDPP11 に対する、阻害剤 SH5 及び NPPB の in silico 相互作用解析

この度は、第 50 回構造活性相関シンポジウム実行委員会特別賞を賜り、誠にありがとうございました。発表にあたり、創薬研究をご専門とする先生方から、貴重なご意見をいただくことができました。昨今の厳しい状況が続く中で、発表の機会を与えてくださいました、シンポジウム実行委員の先生方、構造活性相関部会の先生方に心より感謝申し上げます。

当研究室ではこれまでに、歯周病原性細菌 *P. gingivalis* 由来 DPP11 (PgDPP11) を標的とした in silico スクリーニングを行い、ヒット化合物 (阻害剤 SH5 及び NPPB) を見出しています [Sci Rep 9, 13587 (2019).]。SH5 は PgDPP11 の同族タンパク質、多剤耐性日和見菌 *S. maltophilia* 由来 DPP7 (SmDPP7) に対しても阻害効果を発揮しますが、NPPB は PgDPP11 を選択的に阻害することが明らかにされています。そこで、本研究では、計算科学的手法を用いて、SH5 が PgDPP11 と SmDPP7 を共に阻害できる理由、及び NPPB が PgDPP11 に対して標的選択性を持つ理由を考察しました。計算の結果、NPPB は、SmDPP7 のリガンド結合部位を構成するアミノ酸残基との相互作用が弱く、安定な複合体構造を形成できない可能性が示唆されました。このことが、NPPB が PgDPP11 に対し、標的選択的である理由と考えられました。

本研究の遂行にあたり、岩手医科大学 阪本泰光先生、長岡技術科学大学 小笠原渉先生、北里大学 田中信忠先生から、多大なるご協力を賜りました。深く御礼申し上げます。最後に、有益なご助言とお力添えを賜りました、早川大地先生、合田浩明先生に心より感謝申し上げます。

歯周病菌由来 PgDPP11 に対する、阻害剤 SH5 及び NPPB の *in silico* 相互作用解析

○渡邊友里江¹、早川大地¹、合田浩明¹（昭和大薬¹）

E-mail: ywatanabe@pharm.showa-u.ac.jp

1. 背景と目的

歯周病は、口腔内の歯周病原性細菌の感染によって引き起こされる、慢性炎症性疾患である。近年、心血管障害[1]や糖尿病[2]、誤嚥性肺炎[3]、骨粗鬆症[4]といった全身性疾患との関連が明らかにされ、歯周病治療の重要性が注目されている。歯周病治療は歯磨きや生活習慣の改善といった一次予防が主体となっており、非侵襲的な薬物治療では、アジスロマイシン（ジスロマック®）などに効果が認められているが[5]、菌交代症や耐性菌の問題があるため繰り返し投与することの妥当性は得られていない。つまり、歯周病の進行と増悪を食い止める、効果的かつ選択的治療薬は現在までに存在しない。

歯周病原性細菌には数多くの種類が存在するが、その中でも、歯周病の増悪に最も大きく関与するのが、*Porphyromonas gingivalis* (*P. gingivalis*) である [6]。*P. gingivalis* は、糖非発酵性グラム陰性細菌の一種であり、炭素源・栄養源として炭水化物などの糖類ではなく、蛋白質やペプチドを利用する。*P. gingivalis* の炭素源・栄養源の獲得になくてはならない酵素が、*P. gingivalis* 由来 dipeptidase 11 (PgDPP11) である。それゆえ、PgDPP11 選択的阻害剤は、新規歯周病治療薬となる可能性があるとして期待されている。

近年、国際宇宙ステーション“きぼう”での微小重力下結晶化実験により、PgDPP11 の全体構造 (PDB ID: 4XZY) が、明らかにされた [7]。これまでに、当研究室はこの結晶構造を利用した *in silico* スクリーニングを実施し、岩手医科大学、長岡技術科学大学、北里大学との共同研究により PgDPP11 阻害活性を有するヒット化合物 (SH5 及びその誘導体 NPPB) を見出した [8] (図 1)。その後、SH5 に関しては、PgDPP11 との複合体構造 (PDB ID: 6JTC) が X 線結晶構造解析により決定されたが (図 2)、NPPB と PgDPP11 との複合体構造はまだ解かれていない。そこで本研究では、分子ドッキング計算及び分子動力学 (Molecular dynamics; MD) 計算により、PgDPP11 と NPPB との結合様式を解析し、PgDPP11 の SH5 及び NPPB の認識機構について考察することを目指した。

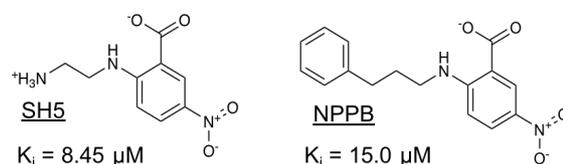


図 1. SH5(左)と NPPB(右)の化学構造式。PgDPP11 に対する阻害定数(K_i)をそれぞれの下に示す。

2. 方法

まず、結晶構造 (PDB ID: 6JTC) で観察される PgDPP11 と SH5 の結合様式を再現できる分子ドッキング計算手順を構築し、その構築した手順を用いて、PgDPP11 と NPPB の複合体モデル構造を構築した。これには、Small-Molecule Drug Discovery Suite 2019-3 (Schrödinger, LLC, New York, NY, 2019) を利用した。

次に、PgDPP11 と NPPB の複合体モデル構造と PgDPP11 と SH5 の複合体結晶構造を初期構造とし、それぞれの複合体について AMBER16 を用いて水溶液中での MD 計算を行った。定圧下 298K で 100 ns の MD を実施し、得られたトラジェクトリを AmberTools15 の cpptraj により解析した。結合自由エネルギー計算には、Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) 法に基づき、MMPBSA.py script [9] を用いた。'multiple-trajectory' approach とするため、PgDPP11 タンパク質単独、リガンド分子単独の MD をそれぞれ 150 ns、100 ns 実行し、後半 50 ns のトラジェクトリを結合自由エネルギー計算に利用した。さらに、 ΔG_{bind} をアミノ酸残基ごとにエネルギー分割し、PgDPP11-SH5 複合体、PgDPP11-NPPB 複合体で相互作用に重要なアミノ酸残基を比較した。

3. 結果と考察

分子ドッキング計算、及び MD 計算の結果、SH5 と NPPB で共通する、nitrobenzoic acid 部分は、どちらの複合体においても Arg337, Thr650, Asn670, Arg673 と水素結合相互作用を形成する可能性が示唆された (図 2)。しかし、MD トラジェクトリ解析の結果、PgDPP11-NPPB 複合体では、Arg337 との水素結合が PgDPP11-SH5 複合体よりも強いことが示された (水素結合頻度は前者が 93.6 % である一方、後者が 38.1 %)。また、SH5

のアミノ基は、基質ペプチド N 末端の認識に働く Asp672、Asn218 [7]と強い水素結合相互作用を形成することが示唆された(図 2)。一方、NPPB は SH5 のアミノ基に相当する部位がフェニル基に置換されているため、Asp672 や Asn218 との水素結合相互作用を形成せず、多様な相互作用を形成している可能性が示唆された。結合自由エネルギー計算の結果、PgDPP11-SH5 複合体、PgDPP11-NPPB 複合体の ΔG_{bind} は、それぞれ、 -22.6 kcal/mol 、 -19.6 kcal/mol であった。これらの値は阻害定数の傾向と一致した。さらに、エネルギー分割の結果、 ΔG_{bind} にもっとも寄与するのは、どちらの複合体においても Arg673 であることがわかった。しかし、その次以降に寄与が大きいアミノ酸は、PgDPP11-SH5 複合体では Trp219、Asn218、PgDPP11-NPPB 複合体では Thr650、Arg337 と、安定な複合体形成に働くアミノ酸残基には違いがある可能性が示唆された。

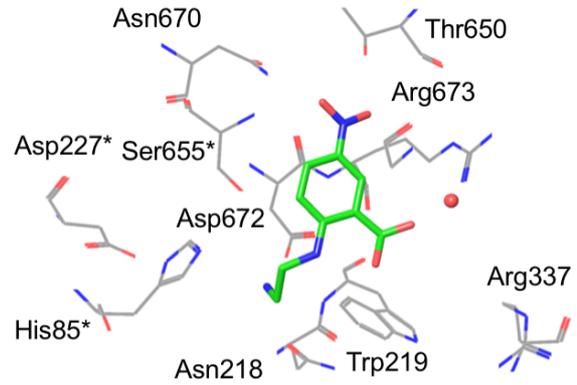
SH5 は PgDPP11 の同族タンパク質、日和見菌 *Stenotrophomonas maltophilia* 由来 DPP7 (SmDPP7) に対しても阻害活性を有することが明らかにされている [8]。しかし、NPPB は SmDPP7 に対して阻害活性を持たず、NPPB は PgDPP11 選択的である。NPPB が PgDPP11 に対し選択性をもつ理由を考察するため、ホモロジーモデリングと上述の計算手法を組み合わせて SmDPP7 についても複合体モデルの構築を現在進めている。

4. 参考文献

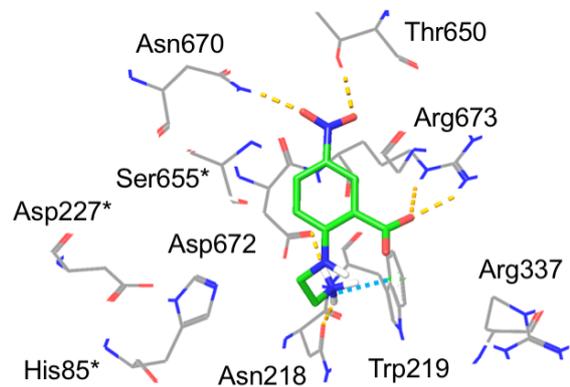
1. Humphrey LL et al, *J Gen Intern Med.* **23**, 2079-2086 (2008),
2. Preshaw PM et al, *Diabetologia* **55**, 21-31 (2012),
3. Paju S, Scannapieco FA, *Oral Dis.* **13**, 508-512 (2007),
4. Wang CJ, McCauley LK, *Curr Osteoporos Rep.* **14**, 284-291 (2016),
5. Oteo A et al, *J Clin Periodontol* **37**, 1005-1015 (2010),
6. Griffen AL et al, *J. Clin. Microbiol.* **36**, 3239-3242 (1998),
7. Sakamoto Y et al, *Sci Rep* **5**, 11151 (2015),
8. Sakamoto Y et al, *Sci Rep* **9**, 13587 (2019),
9. Miller BR III et al, *J. Chem. Theory Comput.* **8**, 3314-3321 (2012),

5. 謝辞

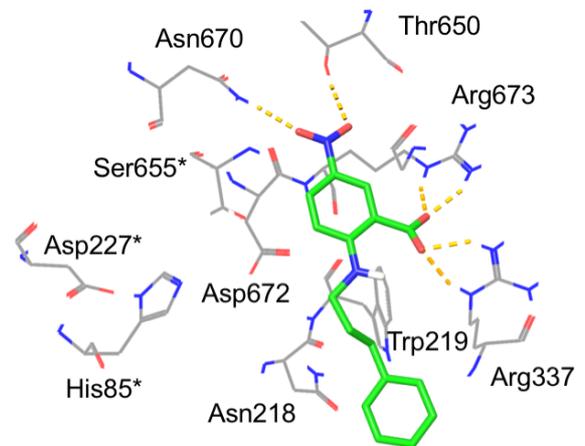
岩手医科大学 薬学部 阪本泰光先生、北里大学 薬学部 田中信忠先生、長岡技術科学大学 工学研究科 小笠原渉先生には、PgDPP11 の構造と機能の理解のため、多大なるご協力と有益なご助言を賜りました。深く御礼申し上げます。本研究は、科学研究費補助金(18K14886)の助成を受けたものです。



A) Crystal structure (PDB ID: 6JTC)



B) Optimized PgDPP11-SH5 complex



C) Optimized PgDPP11-NPPB complex

図 2. A : PgDPP11-SH5 複合体の結晶構造 (PDB ID: 6JTC) で観察された結合様式。B : PgDPP11-SH5 複合体の MD 計算の最終構造をエネルギー最小化した構造。C : PgDPP11-NPPB 複合体の MD 計算の最終構造をエネルギー最小化した構造。水素結合を黄色、cation- π 結合を水色の点線で示す。

//// Activities ////

第50回構造活性関連シンポジウム開催報告

日時：令和4年11月10日(木)・11日(金)

会場：オンライン開催

主催：日本薬学会構造活性関連部会

協賛：CBI学会・人工知能学会・日本化学会・日本コンピュータ化学会・日本生化学会・日本生物物理学会・日本農薬学会・日本バイオインフォマティクス学会・分子科学会・理論化学会・日本薬学会医薬化学部会

後援：日本農芸化学会

1975年5月10日に第1回が開催された構造活性関連シンポジウムも、今年度で50回目の節目を迎えた。それを記念する第50回構造活性関連シンポジウムは、構造活性関連部会の主催により2022年11月10～11日の2日間にわたって開催された(図1)。準備段階で新型コロナウイルス感染症の流行状況が予測できなかったため、前回・前々回と同様にオンラインによる開催とした。この3年弱の間にオンラインツールが広く使用されるようになり、オンライン学会の開催に向けてさまざまな選択肢が存在したが、今回は第50回記念・特別・招待講演と口頭発表のセッションにZoomを、ポスター発表及び懇親会にoViceを用いた。構造活性関連研究の裾野を広げるため、参加費について学生は無料、また日本薬学会会員、協賛・講演学会会員についても無料とした。その結果、参加登録237名(うち学生41名)、と多くの参加者があり、前回・前々回と同規模の開催となった。協賛学会会員として、従来よりご協賛いただいているCBI学会(52名)、日本農薬学会(18名)、日本化学会(8名)からの参加が多かったが、今回よりご協賛いただいた日本生物物理学会(6名)、日本バイオインフォマティクス学会(3名)、日本コンピュータ化学会(3名)等からも参加者があり、さまざまな分野の方にご参加いただくことができた。

今回のシンポジウムではin silico創薬全般に関する発表を、関連する実験研究も含めて募ったところ、一般口頭発表15件、ポスター発表27件の演題登録があった。また、第50回記念講演および特別講演を1件ずつ、招待講演を3件企画し、さらに2つのランチョンセミナーを用意した。

第50回を記念した記念講演では、日本薬学会構造活性関連部会長の本間光貴先生(理研)に「シミュレーションとインフォマティクスによる医薬品設計研究 再び定量的予測へ」と題して、構造活性関連研究のこれまでの歩みと今後の展望についてご講演をいただいた。また招待講演は狭義の構造活性関連研究にこだわらず、広い分野から著名な若手・ベテラン研究者の先生方にご講演いただいた。初日の招待講演では、日本薬学会学術貢献賞を受賞されたご経験のある栗原正明先生(湘南医療大薬)に「核内受容体に作用する薬物分子の創製」と題して、実験化学者の視点から見た創薬について、in silicoによる試みも交えてご紹介いただいた。また同じく初日に行われた高田匠先生(京大複合研)のご講演は「加齢に応じて進行する蛋白質中アミノ酸残基異質性化に関する分離分析手法の現状」と題して、新進気鋭の実験研究者として現在進行中の生体内非酵素反応研究について最新の成果をご講演いただいた。2日目の招待講演で

第50回
構造活性関連シンポジウム
2022年 11/10(木)・11(金) オンライン開催

第50回記念講演 本間光貴 (日本薬学会構造活性関連部会長)(理化学研究所)
「シミュレーションとAIを組み合わせた医薬品設計定量予測の重要性と今後の展開」

特別講演 立花裕樹 (塩野義製薬株式会社)
「新型コロナウイルス感染症治療薬S-217622の創製」

招待講演
伊藤育志 (東京大学大学院情報理工学系研究科)
「進化・前院とAI」
栗原正明 (湘南医療大学薬学部)
「核内受容体に作用する薬物分子の創製」
高田 匠 (京都大学複合分子科学研究所)
「加齢に応じて進行するアミノ酸残基異質性と分離分析手法の現状」

一般講演
口頭発表(Zoom)およびポスター発表(oVice)
発表申込・8月16日(火)～9月23日(金) 必着
講演要旨・10月14日(金) 必着
発表/参加登録予約申込の詳細は、ホームページをご覧ください。

参加登録費
【日本薬学会会員・協賛学会会員・後援学会会員】 無料
【学生(会員・学生会員問わず)】 無料
【上記以外】 1,000円 ※費用込のみ。参加取り消しによる返金ははじりません。

主催：日本薬学会構造活性関連部会
協賛：CBI学会、人工知能学会、日本化学会、日本コンピュータ化学会、日本生化学会、日本農薬学会、日本バイオインフォマティクス学会、日本生物物理学会、日本農芸化学会、日本薬学会医薬化学部会
後援：日本農芸化学会

問い合わせ先・第50回構造活性関連シンポジウム実行委員会
〒100-8555 東京都千代田区千代田1-8-1 理研 100-8555 100-8555
E-mail: sym@SARConline.jp u.ac.jp

図1 シンポジウムポスター

は人工知能の分野において著名な研究者である伊庭斉志先生（東大院情報理工）より「進化・創発と AI」をタイトルとして、最適化問題の解法や人工知能について、創発を中心的なキーワードとしてご紹介いただいた。また特別講演として、製薬企業における実際の COVID-19 治療薬の創薬についてのご講演、「COVID-19 経口治療薬 Ensitrelvir (S-217622) の創製」を立花裕樹先生（塩野義）に行っていた。

一般口頭発表では、*in silico* 創薬による化合物探索研究はもちろんのこと、タンパク質立体構造のモデル評価や相互作用解析、抗体による糖鎖認識機構の予測、ドッキングや自由エネルギー計算のための手法の開発など、多様な研究テーマの成果が紹介され、活発な質疑応答があった。本シンポジウムでは Zoom Meetings による講演・口頭発表を行ったため、質疑応答も座長の指示に従い質問者が直接質問をすることが可能となった。昨年までの Zoom Webinars による講演では Q&A に寄せられた質問を座長が代読する形式がとられており、議論が深まりにくいとして仲西前実行委員長が改善の必要性を指摘していた。今回はそれを受けての改善であったが、これによって対話としての質疑応答が実現できた。ただしこれが実現できたのは、3 年弱におよぶ COVID-19 禍でオンラインミーティングが一般化したこと（発言権の受け渡しに戸惑わなくなったこと）も理由に挙げられるかもしれない。

ポスター発表は oVice を用いてヴァーチャル空間内で実施した。昨年同様セッション直前に各発表の 1 分間のショートプレゼンテーション動画を視聴したのちに示説を行った。ヴァーチャル会場のレイアウトを図 2 に示すが、各ブースにタイトルおよび演者名を記載することで目指すブースがどこにあるかを明示した。ヴァーチャル空間を使用することで参加者が可視化され、あちこちのブースを渡り歩いたり通路で参加者同士が会話したりと、対面開催に近い雰囲気でもポスター発表を行うことができた。またポスターについては別途参加者限定サイト上にも掲示しており、発表内容をそちらでも確認できるようにした。オンラインでのポスター発表は内容を一瞥して目当ての研究を探すという行動がとりにくいが、サイト上での掲示によりそれが可能となり、議論の活性化や閲覧の効率化ができたのではないかと考えている。

以上の研究発表の他に、1 日目にランチョンセッション（12 時 25 分～13 時 20 分）を企画し、企業からのセミナーを 2 件実施した。こちらも Zoom Meetings を用いたが、トラブルなく実施することができ、また質疑についても Meetings の機能により参加者自身の声で行うことができた。

初日のセッション終了後には、oVice で情報交換会（従来の懇親会に該当する）を開催した。こちらでは発表の質疑応答の続きや情報交換、コロナ禍でなかなか会うことができない方とのご挨拶など、様々な目的で活用いただけたのではないと思う。ただし参加登録者 70 名のうち実際の参加者は 30 名弱に留まっており、オンラインツールの一般化に伴って目新しさも逡減したことから、過去 2 年間と比べると「オンライン懇親会」の意義は低下しているかもしれない。



図 2 oVice ポスター会場レイアウト(1 日目)

SAR Award は、審査希望のあった発表を対象に、部会幹事のオンライン投票の結果を集計し、受賞者を決定した。また第 50 回を記念して制定した実行委員会特別賞についても SAR Award と同様の方法で受賞者を選定した。以下に受賞者を紹介する。受賞者の皆様にはお祝い申し上げますとともに、今後のご活躍を期待します。また、審査していただいた先生方にも御礼申し上げます。

SAR Award (5 名) :

- ・ KO-03 吉澤竜哉 (横浜市立大学生命医科学研究科)
「強化学習に基づく分子生成手法を用いたキナーゼ選択的阻害剤の設計」
- ・ KO-10 原田隆平 (筑波大学計算科学研究センター)
「化合物の膜透過プロセスを紐解く自由エネルギー計算手法の開発」
- ・ KP-03 高谷大輔 (大阪大学大学院薬学研究科)
「インシリコスクリーニング(LBDD)のための最近のフィンガープリントの性能検討」
- ・ KP-06 市丸嘉 (湘南医療大学薬学部)
「エポキシドを導入したインディルビン誘導体 Epox/Ind の抗がん活性に関する構造活性相関」
- ・ KP-09 星野小百合 (横浜市立大学理学部)
「短時間 MD シミュレーションによるリガンド-タンパク質間結合親和性推定手法の検討」

第 50 回構造活性相関シンポジウム実行委員会特別賞 (2 名) :

- ・ KO-15 田中良樹 (アグロデザイン・スタジオ)
「タンパク質立体構造情報を利用した新規 ALS 阻害型除草剤の探索」
- ・ KP-15 渡邊友里江 (昭和大学薬学部)
「歯周病菌由来 PgDPP11 に対する、阻害剤 SH5 及び NPPB の *in silico* 相互作用解析」

オンライン学会が一般化したこともあり、今回のシンポジウムは大きなトラブルなく実施することができた。学会の対面開催が増えてきた今年度においては、逆にオンライン開催であったことが今回のシンポジウムの特色となったのではないかと考えている。開催にあたって多大なるご助言、情報、資料等をくださいました常任世話人・常任幹事の皆様、第 50 回シンポジウム実行委員の先生方に感謝いたします。また、これらをもとに本シンポジウムを企画、運営していただいた第 50 回シンポジウム実行委員およびテクニカルアドバイザーの先生方にも謝意を表します。

最後に、協賛・展示いただいた企業の皆様、協賛・後援いただいた学会の皆様、第 50 回記念講演・特別講演・招待講演をお引き受け下さいました先生方、発表者、座長、一般参加の皆様に厚く御礼申し上げます。

第50回構造活性相関シンポジウム
実行委員長 小田 彰史

//// Activities ////

<会告>

構造活性フォーラム 2023

「AlphaFold 時代における構造ベース創薬」(仮)

主催: 日本薬学会構造活性関連部会

協賛・後援: 未定

会期: 2023年6月(予定)

会場: Zoomによるオンライン開催

開催趣旨: AI創薬などの新しい創薬技術の発展が著しい現在において、タンパク質などの構造データをもとに理論的に低分子化合物などを設計する構造ベース創薬は、ますます重要になると考えられる。特にタンパク質立体構造予測プログラムである AlphaFold2 の登場によって、タンパク質構造解析手法の立ち位置を見直さざるを得なくなっている。そこで今回、実験的な構造解析技術とそのデータを活用した創薬に焦点をあててフォーラムを開催する。

プログラム、参加登録等の詳細は部会ホームページにてお知らせします。

構造活性フォーラム 2023 実行委員長
株式会社アグロデザイン・スタジオ 西ヶ谷有輝

部会役員人事

2023 年度 常任世話人

部会長	本間光貴 (理化学研究所)
副部会長	服部一成 (塩野義製薬 (株))
副部会長	竹田-志鷹 真由子 (北里大学 薬学部)
会計幹事	川下理日人 (近畿大学 理工学部)
庶務幹事	杉本 学 (熊本大学大学院 先端科学研究部)
広報幹事	加藤博明 (広島商船高等専門学校)
SAR News 編集長	幸 瞳 (理化学研究所)
ホームページ委員長	高木達也 (大阪大学大学院 薬学研究科)

構造活性相関部会の沿革と趣旨

1970年代の前半、医薬薬を含む生理活性物質の活性発現の分子機構、立体構造・電子構造の計算や活性データ処理に対するコンピュータの活用など、関連分野のめざましい発展にともなって、構造活性相関と分子設計に対する新しい方法論が世界的に台頭してきた。このような情勢に呼応するとともに、研究者の交流と情報交換、研究発表と方法論の普及の場を提供することを目的に設立されたのが本部会の前身の構造活性相関懇話会である。1975年5月京都において第1回の「懇話会」(シンポジウム)が旗揚げされ、1980年からは年1回の「構造活性相関シンポジウム」が関係諸学会の共催の下で定期的開催されるようになった。

1993年より同シンポジウムは日本薬学会医薬化学部会の主催の下、関係学会の共催を得て行なわれることとなった。構造活性相関懇話会は1995年にその名称を同研究会に改め、シンポジウム開催の実務担当グループとしての役割を果たすこととなった。2002年4月からは、日本薬学会の傘下組織の構造活性相関部会として再出発し、関連諸学会と密接な連携を保ちつつ、生理活性物質の構造活性相関に関する学術・研究の振興と推進に向けて活動している。現在それぞれ年1回のシンポジウムとフォーラムを開催するとともに、部会誌のSAR Newsを年2回発行し、関係領域の最新の情勢に関する啓蒙と広報活動を行っている。本部会の沿革と趣旨および最新の動向などの詳細に関してはホームページを参照頂きたい。

(<https://sar.pharm.or.jp/>)

編集後記

今号では「構造生成器」をテーマとして、4人の先生にご執筆をお願いしました。横浜市大の先生方には、先生たちのところで開発されている ChemTSv2 の仕組みと事例についてご紹介いただきました。株式会社モルシスの木村先生には、アストラゼネカ社で開発された REINVENT の仕組みと、REINVENT を CCG 社の MOE 上で使用できるインターフェースについてご紹介いただきました。奇しくも今号の準備中に OpenAI 社の ChatGPT が公開されるなど、generative model が大きく注目されています。化合物のデザインをモデルが担う場面は増えるのかもしれませんが、便利になる反面、その化合物構造がなぜ生成されたのか、あるいは取り扱っている化合物群の構造活性相関について、研究者自身が考察する力は変わらず求められると考えます。ご多忙の中、快くご執筆していただいた先生方に深く感謝申し上げます。

また今号から3名の先生方が編集委員に加わりました。新体制で進めてまいりますので、引き続きよろしくお願ひいたします。

(編集委員会)

SAR News No.44 2023年4月1日

発行:日本薬学会 構造活性相関部会長 本間光貴

SAR News 編集委員会

(委員長) 幸 瞳、浴本亨、遠藤智史、合田浩明、仲西 功、原田俊幸

*本誌の全ての記事、図表等の無断複写・転載を禁じます。