

○榊原康文^{1,2}・新井真理子²・白井文幸³・吉田稔³・八代田陽子³（北里大未来工¹・慶應大理工²・理研 CSRS³）

創薬では膨大な化合物空間の中から目的の化合物構造を効率よく見つける必要がある。近年、この課題を解決するために深層生成モデルが注目されているが、既存の多くの手法は化合物の特性や既存薬との類似性に基づいて生成を行い、標的タンパク質との相互作用情報を考慮していない。そのため、特定の疾患に関連するタンパク質と高い特異性をもって相互作用する化合物を効率的に生成することには依然として課題が残る。

本研究では、この課題を克服するために、タンパク質特異的な化合物生成モデルを開発した（図1）。モデルの基盤には Transformer を採用し、エンコーダには大規模データで事前学習されたタンパク質言語モデル ESM-2 (Lin *et al.*, 2022) を用い、デコーダには榊原らが開発した分子生成モデル FRATTVAE (Inukai *et al.*, 2024) を導入した。ESM-2 を活用することで、標的タンパク質を効率よく潜在変数ベクトルへと変換し、タンパク質間の類似性や機能情報を学習することが可能となる。一方、FRATTVAE は化合物をフラグメント単位に分解した木構造として扱うことにより、高精度かつ効率的な分子生成が可能である。

性能評価として、本モデルと既存のタンパク質特異的な生成モデル AlphaDrug (Hao *et al.*, 2022) との比較実験を行った。その結果、ドッキングシミュレーションにおける結合スコア (Docking Score) において、本モデルは AlphaDrug のスコアを大きく上回った。また、高親和性化合物の割合も大幅に改善されており、生成される化合物の特異性・親和性の両面で優位性を示した。さらに、AlphaDrug では MCTS などの追加探索アルゴリズムが必要であるのに対し、本モデルでは生成と探索を一貫して行うことが可能で、生成・探索の処理速度を約 100 倍に向上させた。物性値条件を明示的に指定可能であるため、目的に応じた化合物の特性制御も容易であることが確認された。具体的応用として、抗がん剤のターゲットとして注目されるタンキラーゼに対する新規阻害剤の設計を行った。本モデルに対して既知のタンキラーゼ阻害剤のデータを用いて fine-tuning を施し、生成された化合物構造の有効性を検証するために、AlphaFold3 (Abramson *et al.*, 2024) を用いたドッキングシミュレーションを実施した。結果として、生成された新規化合物はいずれもタンキラーゼの触媒部位に存在するアデノシンポケットへの特異的結合を示した。

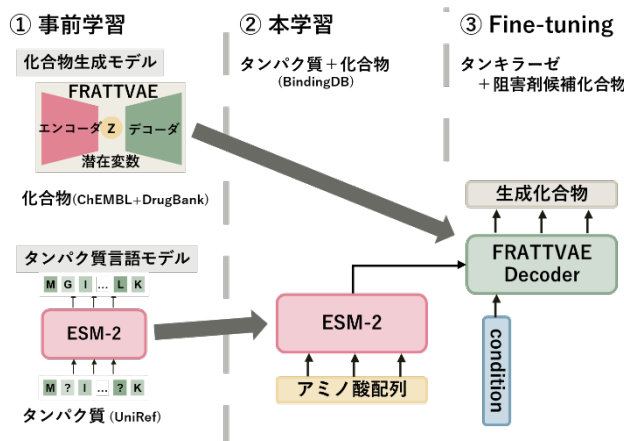


図1：タンパク質特異的な化合物生成モデル

1) T. Ochiai, *et al.*, *Commun. Chem.* **2023**, 6, 249.

2) Y. Ohnuki, M. Akiyama, Y. Sakakibara, *J Cheminform.* **2024**, 16, 103.

PROFILE

榊原康文（北里大学未来工学部 教授）

東京工業大学大学院修了後、富士通研究所、慶應義塾大学理工学部を経て、北里大学未来工学部教授、現在に至る。博士（理学）。バイオインフォマティクスを専門とし、バーチャルスクリーニングによる創薬、生成 AI を用いたゲノム合成、深層学習を適用した医療画像解析などの研究に従事。通産省第五世代コンピュータプロジェクトの研究を出発点とし、人工知能を生命科学の分野に応用する研究に黎明期から携わってきた。（Email: sakakibara.yasubumi@kitasato-u.ac.jp）