

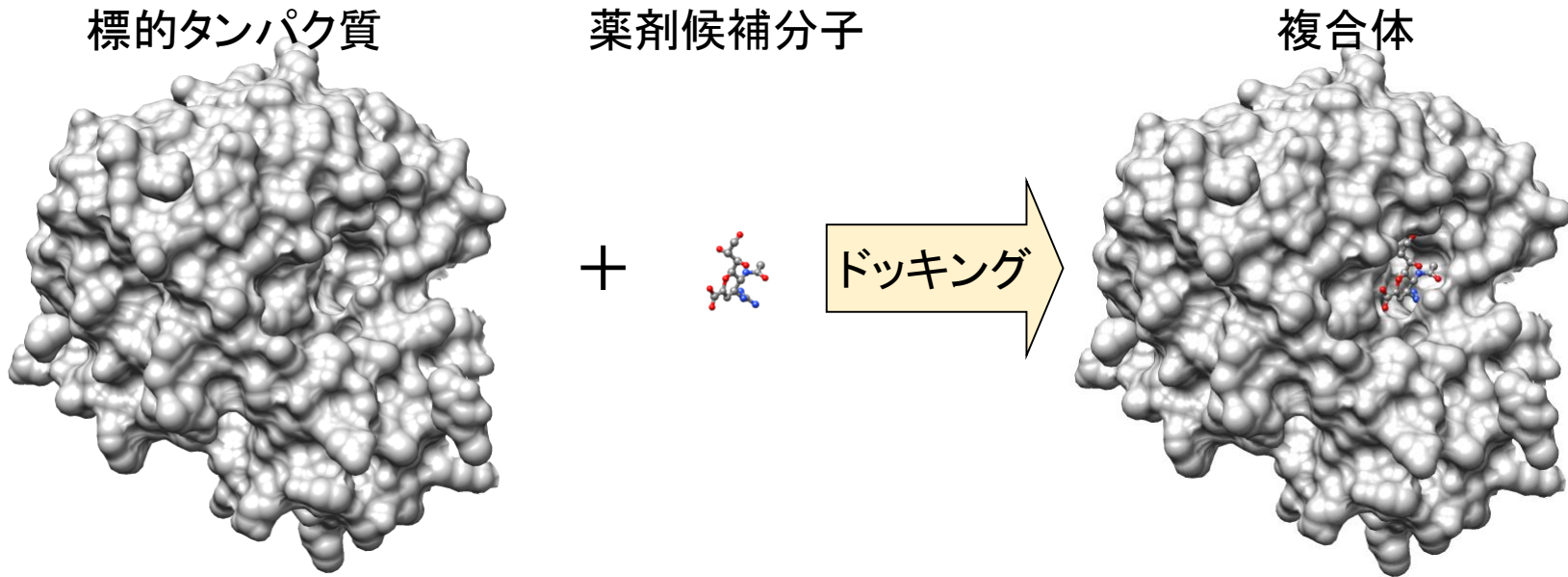
オートエンコーダとGANを用いた タンパク質-薬剤候補分子の 高精度複合体構造モデリング

東京大学大学院農学生命科学研究科

清水謙多郎

研究の背景

- タンパク質と薬剤の候補となる化合物やペプチド（以下、候補分子と呼ぶ）の高精度のドッキングは、バーチャルスクリーニングおよび結合親和性の高い薬剤分子の設計に極めて重要



バーチャルスクリーニング: タンパク質に結合する化合物を探索すること
ドッキング: 分子と分子の相互作用に基づき複合体構造を予測すること

研究の背景

- 結合親和性予測の研究は盛んに行われており、対象によっては高い精度の予測結果が得られている。

[Schneider M., et al. Bioinformatics, Advance Access Publication, July 2019.](#)

- しかし、複合体構造モデリングの精度は一般に低く、例えば、Glide、GOLD、AutoDock Vina、DE Fredなど主なツールで、すべて結晶構造とのRMSD値が平均5Å程度、RMSD値が2Å以下のものが20%以下であったという報告もある。

[Jeffrey R., et al. bioRxiv, Nov. 2018.](#)

- 精度が要求される場合は、分子動力学法(MD)が用いられるが、多大な時間を要する。また、時間をかけても天然の複合体構造が得られるような有効なサンプリングが容易でないことは古くから指摘されている。

[Basciu, A., et al. bioRxiv, Oct. 2018.](#)

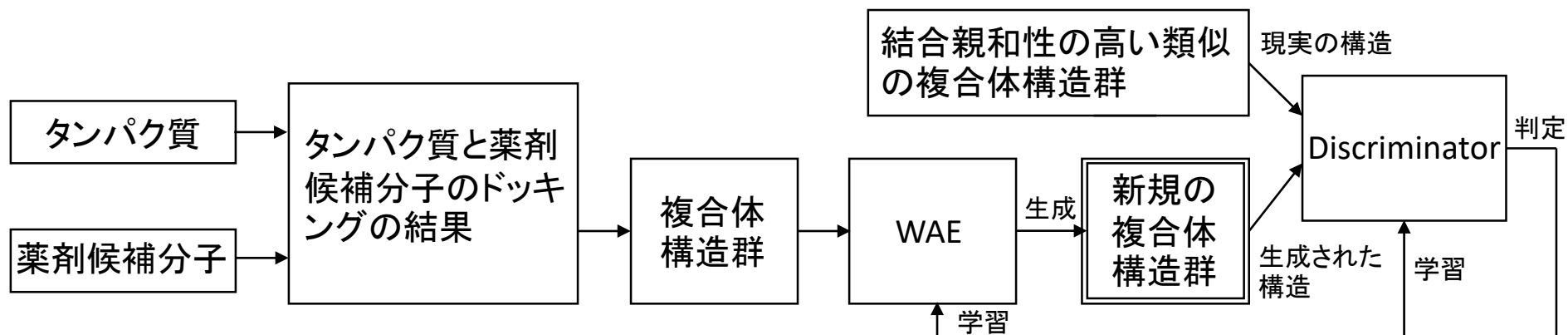
研究目的

- 深層学習を用いて、タンパク質-薬剤候補分子の複合体を高精度でモデリングする手法を開発する
- 類似の候補分子で結合親和性の高い複合体構造が得られている場合は、その情報を学習に取り込むことにより、効果的なモデリングを行う

研究内容

- ドッキングで得られた複合体構造群をもとに、Wasserstein Auto-Encoder (WAE) を用いて、これらを補完する新しい立体配座を生成する
 - WAEの潜在空間で、結合親和性の高い複合体構造を効率的に判別できるようにする
- 敵対性生成ネットワーク (Generative Adversarial Networks, GAN) の機構を用いて、類似の候補分子が結合した既知の複合体構造を学習に組み込む
 - 実験によって得られている複合体構造を活用して、結合親和性の高い複合体構造を効果的に生成できるようにする

研究内容



達成目標

- MDのように多大な計算時間を必要とせず、従来のドッキングより精度の高いタンパク質－薬剤候補分子の複合体構造のモデリングを実現する
- 具体的なタンパク質に対して、薬剤化合物が結合していない構造に対してドッキングを行い、本手法により、予測精度が向上することを確認する
- WAEのみを用いた複合体モデリングは令和3年3月、類似の候補分子の情報を利用したモデリングは令和3年10月までに実施

予想される効果

- タンパク質と薬剤候補分子とのドッキングにおいて、従来の手法では再現することが難しい、親和性の高いドッキング構造(複合体構造)の生成を可能にするもので、薬剤のバーチャルスクリーニングの精度向上、作用効果の高い薬剤分子の設計に大きく寄与する
- 薬剤分子の設計では、実験等である程度の結合親和性を有することが確認されている候補分子をもとにして、さらに結合親和性の高い候補分子を探索するのが一般的である。このような場合は、本研究で開発する、類似の複合体構造を学習させる機構は、予測精度の向上に大いに貢献するものと期待される
- 薬剤の対象も、低分子化合物に限らず、ペプチドやタンパク質にも適用することが可能である
- 本手法は、新規酵素の設計や農薬の設計、機能性食品の開発にも適用できるものであり、一般性がある