

REINVENT4インターフェイスの紹介

REINVENT 4^{1,2)} は、低分子の de novo デザイン、フラグメント伸長 / 連結、分子最適化などが行える生成 AI モデルを搭載したオープンソースのソフトウェアです。REINVENT4 インターフェイス (I/F) では、MOE の GUI から REINVENT 4 の入力ファイルの生成、計算の実行、出力ファイルの可視化を簡単に行えます。さらに、強化学習においては、MOE の記述子、QSAR モデル、ファーマコフォアモデル、ドッキングスコアなどを評価関数の要素として使用できます。REINVENT4 I/F は、MOE のアドオンプログラム共有サイトである SVL Exchange³⁾ から入手できます。本稿では、REINVENT 4 の概要と REINVENT4 I/F の機能や使用例について紹介します。

REINVENT 4 の概要

REINVENT 4 は、低分子の de novo デザイン (REINVENT)⁴⁾、フラグメント伸長 (LibiNVENT)⁵⁾、フラグメント連結 (LinkiNVENT)⁶⁾、分子最適化 (Mol2Mol)⁷⁾ などの生成 AI モデルと、それらを最適化する学習手法を統合したオープンソースのソフトウェアです。学習済みのモデルが搭載されており、すぐに低分子設計に活用できます。REINVENT 4 には、強化学習や転移学習を行う機能があり、膨大な化合物空間から、所望の特性や部分構造を持つ分子群を効率よく生成できます。その他には、分子生成だけを行うサンプリング機能や評価関数の適用だけを行うスコアリング機能があります。

REINVENT 4 の強化学習においては、生成する分子の多様性を高めるための Diversity filter や、スコアの高い分子を記憶して学習に再利用する Inception など、学習を効率よく進めるための機能も含まれています。また、評価関数には RDKit で計算できる分子量、極性表面積、logP の推算値などの記述子、合成可能性スコア、望まない SMARTS パターンに基づくアラート、フィンガープリントの類似度、Chemprop⁸⁾ モデルによる予測値などを組み合わせ使用できます。さらに、外部プログラムをこれらと同様に評価関数の要素として組み込む仕組みも備わっています。弊社では、この仕組みを使用して MOE のさまざまな機能を評価関数の要素として利用できるように実装した REINVENT4 I/F を開発しました。

REINVENT4 I/Fの機能

MOE の REINVENT4 I/F では、柔軟に設定可能な GUI とともに以下の機能を提供します。

MOE の評価関数の要素

強化学習における評価関数の要素として、MOE の記述子、QSAR モデル、フィンガープリントモデル、コンセンサスモデル、ファーマコフォアモデル、ドッキングスコア、特定の原子群との相互作用エネルギーが使用できます。これらは、既存の評価関数の要素と組み合わせたり、それらと同様にシグモイド関数や階段関数などで規格化したりできます。

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

実行できます。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ます (図 1)。

入力 SMILES の生成

フラグメント伸長においては、MOE ウィンドウ上で基準となる母核構造の伸長位置の原子を選択して、計算に必要な SMILES を生成できます。フラグメント連結でも同様に MOE ウィンドウ上から対象の置換基の連結点の原子を指定することで SMILES を生成できます。

MDB の入出力対応

転移学習やスコアリング機能などでは SMILES ファイルだけでなく、MOE の MDB ファイルを入力に指定できます。

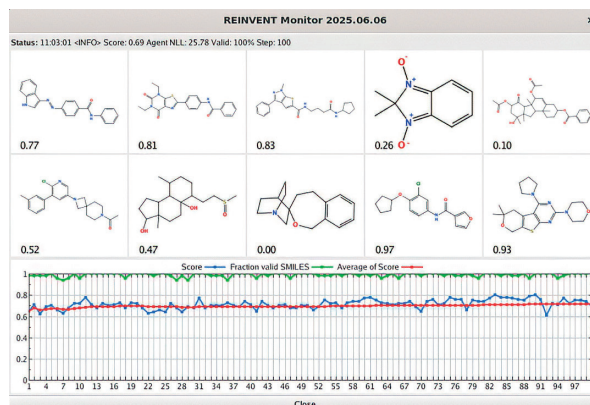


図 1. 強化学習における REINVENT4 I/F の監視機能。上部には、各計算ステップで生成された化合物と評価値が表示され、下部には横軸を計算ステップとした妥当な SMILES の割合 (緑)、評価値の平均 (青色) とその累積の平均 (赤) がプロットされている。

REINVENT4 I/Fの使用例

REINVENT 4 では低分子の de novo デザイン、フラグメント伸長 / 連結、分子最適化などの強化学習が行えますが、以下ではそれらの中でフラグメント伸長とフラグメント連結の REINVENT4 I/F における使用例を紹介します。

フラグメント伸長

フラグメント伸長の機能では、母核構造を保持しながら、指定した位置に置換基を導入した多様な分子を生成します。サンプルとして用いた分子系は、上皮成長因子受容体 (EGFR) と阻害剤のエルロチニブの共結晶構造 (PDB

ID: 1M17) です。計算時には、エルロチニブの二つのエーテル鎖を削除し、一方の水素を置換基導入位置として指定し、Asp776 との新たなイオン相互作用の形成を目標としました (図 2)。評価関数には、MOE の合成可能性スコア、ドッキングスコア、Asp776 とのイオン相互作用エネルギーを組み合わせました。以上の設定で、100 ステップの強化学習を行った結果、ドッキングスコアと Asp776 とのイオン相互作用エネルギーが良好な分子が得られました (図 3)。

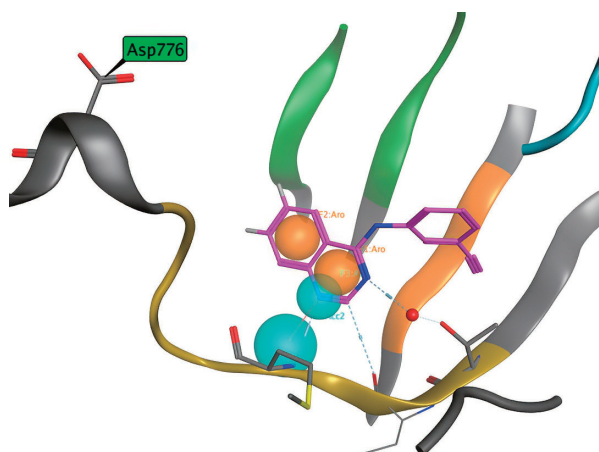


図 2. フラグメント伸長の強化学習で使用した EGFR と阻害剤の母核構造。芳香環フィーチャー (橙色の球) と水素結合アクセプター (水色の球) ファーマコフォアモデルはドッキング計算の拘束条件として使用。

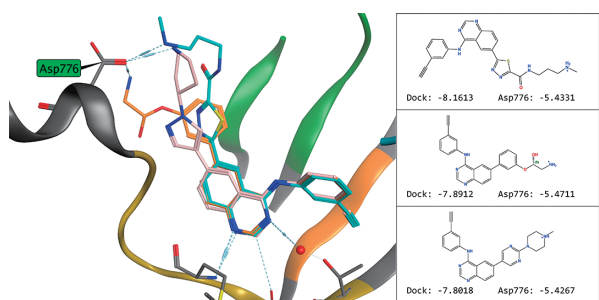


図 3. フラグメント伸長の強化学習で得られた分子のドッキング構造 (左) と構造式 (右) (Dock: ドッキングスコア, Asp776: Asp776 とのイオン相互作用エネルギー (kcal/mol))

■ フラグメント連結

フラグメント連結の機能では、指定した二つのフラグメントを連結するリンカーを導入した多様な分子を生成します。サンプルとして用いた分子系は、カゼインキナーゼ 2 とその阻害剤の共結晶構造 (PDB ID: 5CU4) です (図 4)。この阻害剤の両端の部分構造 (図 4 の橙色の部位) は、それぞれ単独のフラグメントとして共結晶構造 (PDB ID: 5CSV, 5CSH) があります。そこで、これらの二つのフラグメントを連結するリンカーを含む分子を探索することとしました。その際に既存の阻害剤が持つ Asn118 との水素結合を保つようなリンカーを探索することとしました。評

価関数には、MOE の druglike 記述子、合成可能性スコア、ドッキングスコア、Asn118 との水素結合エネルギーを組み合わせました。以上の設定で、100 ステップの強化学習を行った結果、ドッキングスコアと Asn118 との水素結合エネルギーが良好な構造が得られました (図 5)。

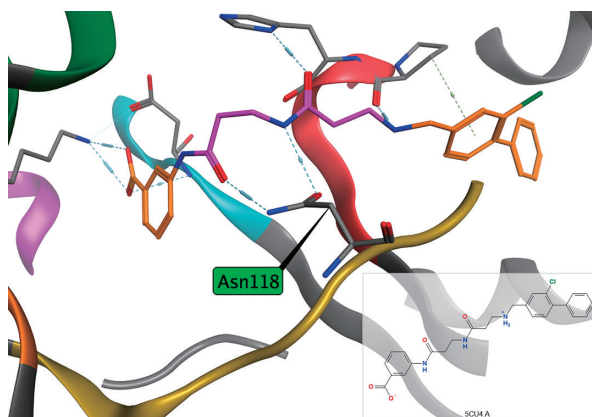


図 4. フラグメント連結の強化学習で使ったカゼインキナーゼ 2 と阻害剤の共結晶構造 (PDB ID: 5CU4)。阻害剤の両端のフラグメント (橙色) を計算で使用。

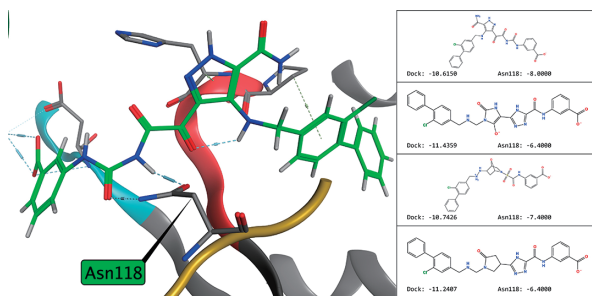


図 5. フラグメント連結の強化学習で得られた分子のドッキング構造 (左) と構造式 (右) (Dock: ドッキングスコア, Asn118: Asn118 との水素結合エネルギー (kcal/mol))

■ REINVENT4 I/F の入手先

REINVENT4 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 から無償で入手できます。

- 1) Loeffler, H. H.; et al. *J. Cheminf.* **2024**, *16*, 20.
- 2) MolecularAI. *REINVENT 4*. <https://github.com/MolecularAI/REINVENT4>
- 3) CCG SVL Exchange | MOE REINVENT4 Interface https://svl.chemcomp.com/#Reinvent_Interface
- 4) Blaschke, T.; et al. *J. Chem. Inf. Model.* **2020**, *60* (12), 5918–5922.
- 5) Fialková, V.; et al. *J. Chem. Inf. Model.* **2022**, *62* (9), 2046–2063.
- 6) Guo, J.; et al. *Digital Discovery* **2023**, *2*, 392–408.
- 7) He, J.; et al. *J. Cheminf.* **2022**, *14*, 18.
- 8) Heid, E.; et al. *J. Chem. Inf. Model.* **2023**, *64* (1), 9–17.