

逆分子設計における探索パラメータの影響と結果分析

1 はじめに

本実験では、強化学習を用いた逆分子設計手法（REINVENT）において、探索パラメータである反復回数（Num_iterations）および1反復あたりの生成分子数（Num_molecules_per_iteration）が探索結果および分子の溶解度に与える影響を調査した。

2 実験設定

以下の設定で実験を行った：

- 反復回数（Num_iterations）：50
- 生成分子数/反復（Num_molecules_per_iteration）：非公開（想定として100前後）
- 目的関数: 溶解度（logS）

探索では、各反復ごとに新たな分子を生成し、その予測溶解度を評価、最良のスコアを持つ分子を保持して進化を行った。

3 結果

反復初期は $\log S \approx -3.1$ 程度の分子から始まり、探索を重ねるごとに改良された分子が出現し、最終的に $\log S \approx -1.91$ の分子が得られた。表??に上位5分子のSMILESと予測 $\log S$ を示す。

4 構造の傾向

上位分子には以下のような構造的傾向が見られた：

- 窒素を含む芳香族（ピリジン環など）
- -OH, -NH₂ など極性基
- S=O や S-S 結合など電子的に影響の大きい官能基

これらの構造は水溶性の向上に寄与すると考えられる。

Table 1: 予測 logS 上位 5 分子

順位	SMILES	logS
1	<chem>O=S(NC(O)(F)c1ccncc1)C(O)S1=CC=CN=C1</chem>	-1.9081
2	<chem>CN(C(=O)NC(N)(O)c1ccncc1)c1ccnnn1</chem>	-1.9397
3	<chem>NC(O)(NS(=O)C(O)S1=CC=CN=C1)c1ccncc1</chem>	-1.9421
4	<chem>CN(C(=N)NC(O)(O)c1cpncn1)S1=CN=CN=C1</chem>	-1.9443
5	<chem>CN(C(=O)NC(C)(N)c1ccncc1)c1cncnc1</chem>	-1.9448

5 考察

反復回数の影響

logS の推移を見ると、初期段階では大きく改善し、30 反復目以降はほぼ収束した。このことから、**50 反復は十分な探索回数**である一方、**より高い性能を目指すには 80 回以上への拡張も有効**であると予想される。

生成分子数の影響（仮定）

本実験では 1 反復あたりの生成分子数は不明であるが、もしこの数が少なければ多様性の確保が難しく、局所最適に陥る可能性が高まる。**今後は分子数を増やすことで初期多様性を向上させ、より有望な化合物探索が可能になると考えられる。**

6 結論

反復回数 50 回の設定で、初期よりも大幅に改良された高溶解度分子が得られた。探索の深化と多様性の両立が今後の課題であり、反復数・生成数の調整が最適化性能に与える影響は大きい。