


|  |        |  |
|--|--------|--|
|  | 領域代表者  | 大阪大学・大学院工学研究科・教授<br>菊地 和也（きくち かずや）<br>研究者番号:70292951                         |
|  | 研究領域情報 | 領域番号：23A204<br>研究期間：2023年度～2027年度<br>キーワード：ケミカルバイオロジー、天然物化学、潜在空間、情報活用、有機合成化学 |

なぜこの研究を行おうと思ったのか（研究の背景・目的）

● 研究の全体像

天然物と合成化合物ライブラリーという二つの化合物リソースを活用した生物活性分子の発見・同定は、化学と生物学の融合分野であるケミカルバイオロジー研究推進の駆動力となってきた。本領域では、これらに続く第3のリソースを提案する。この第3のリソースは、**天然物の生物活性データを基に深層学習技術によって構築される化合物潜在空間（Latent Chemical Space）**からデジタルデータとして得られ、強力な有機合成技術で実空間に具現化されるものである。天然物と情報学研究との融合により生まれる化合物潜在空間は、**データ駆動型ケミカルバイオロジー研究**というパラダイムシフトを起こし、生物活性分子設計に変革をもたらす。この実現に向け、ケミカルバイオロジー、情報科学、有機合成の3班構成による「サイバー生物活性分子デザインラボ」を始動する。この第3のリソースから作り出された化合物を起点として、新しい生体機能解明や医薬・農業シーズに結び付く画期的生物活性分子を高効率に開発できる**生物活性分子デザイン法の新学理構築**を目指す。

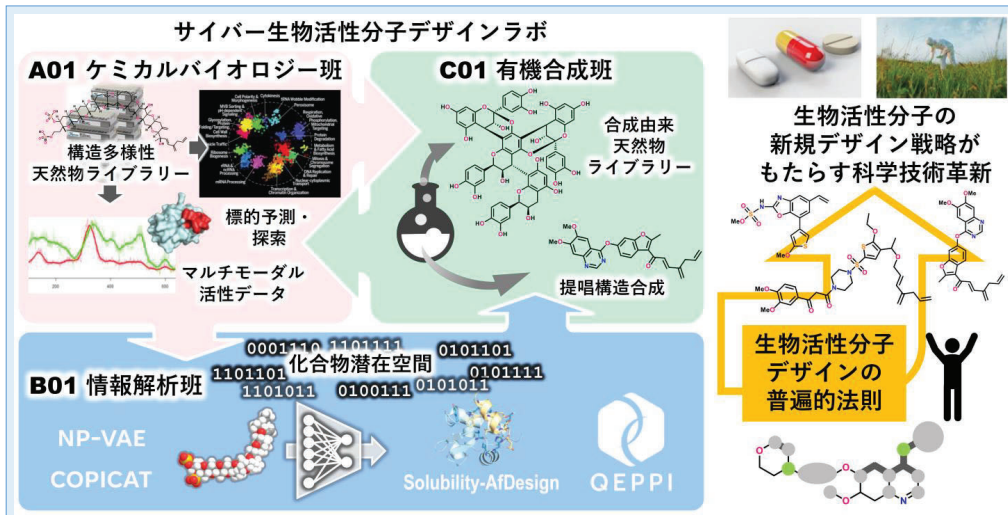


図1 本領域で推進するサイバー生物活性分子デザインラボ

● 研究の背景

化合物の化学構造と生物活性の関係を構造活性相関（Structure-Activity Relationship, SAR）という。現代のケミカルバイオロジー研究や創薬研究では、まず化合物ライブラリーに含まれる膨大な分子群からランダムな探索研究によってヒット化合物を取得し、経験的な構造改変によるSARを明らかにすることによって、より強力で特異的な生物活性分子を取得するアプローチが定法として用いられている。しかし、ライブラリー中の化合物数を増やしても、ランダムに機能分子を見出すこの探索法では、多様な分子の構造デザインには

限界があった。そこで、強力な生物活性分子を論理的・効率的にデザインするためのシステマティックな戦略が強く求められるようになった。本領域では、複雑な天然物と深層学習等の情報学の融合により**化合物潜在空間**を構築し、実空間上では知られていなかった新たな生物活性分子をデザインすることによって第3のリソースを生み出し、ケミカルバイオロジー研究に変革をもたらす計画を立てた。

この研究によって何をどこまで明らかにしようとしているのか

本領域は、生物活性分子のデザインにおける複雑構造の単純化工程、及び類似活性をもつ新規骨格化合物探索、双方の革新的スピード化と省力化を実現するため、天然物化学を規範としたケミカルバイオロジーと情報学との融合研究を展開する。このため「サイバー生物活性分子デザインラボ」をバーチャルに構成し、適宜オープンリソースとしての公開を目指す。天然物の生物活性データを基に深層学習によって構築した**化合物潜在空間**は、第3の化合物リソースを生み出す。この分子デザインラボを活用することで、生物活性分子に必要な構造パターンを明らかにし、生物活性分子デザインの普遍的法則を見出すことを目標としている。

新しい研究領域が興るには、目的意識を揃えと同時に、新しい研究手法の技術サポートが必須である。従来の天然物有機化学は「複雑な分子構造」が最重要視され、その機能向上の学理追求は数居の高い研究領域であった。この観点から本領域では、若手研究者を中心とする共同研究、情報交換を通じて、情報学との融合研究による分子デザインの基盤を成す学理へ意識を向けると同時に研究手法に関する知識を共有し、領域全体のボトムアップを図る。

**A班（ケミカルバイオロジー）**：我が国の伝統ある天然物リソースから、タンパク質間相互作用、タンパク質-RNA相互作用、等の分子間相互作用に干渉して生物活性を引き起こす分子を選抜し、作用機構や標的分子に関するデータを網羅的に収集する。その誘導体のSAR情報や各種オミクスデータを領域研究に提供する。さらに、潜在空間が生み出す化合物の生物活性を評価し、「第3の化合物リソース」を活用した新たなケミカルバイオロジーを始動する。すなわち、潜在空間と実空間を活性データで繋ぐ役割を果たす。

**B班（情報解析）**：A班が提供する構造と活性のマルチモーダルなデータを、B班による情報解析により数値化し、班員が独自に開発した複数のアルゴリズムの入力データとして用いることで潜在空間に反映させ、活性類似性予測の有効性と汎用性を強化する。具体的には、A班の解析データから分子間相互作用予測法COPICATによって知識抽出を行い、タンパク質間相互作用標的活性を予測する低分子評価関数QEPPI、及びAlphaFold2等で提案されたペプチド配列デザイン問題を解く深層学習Solubility-AfDesignを活用することで、NP-VAEによる潜在空間の拡充を図る。潜在空間によりデザインされた分子構造はA班とC班に提示される。新規活性物質候補の創出によりA班とC班による活性評価と合成の連携プロセスを促進するとともに、応用情報学としても独創性の高い手法展開によって、ケモインフォマティクス分野の活性化を目指す。

**C班（有機合成）**：化合物潜在空間によって予測された新規構造の生物活性分子候補は、実空間にてC班が独自の化学技術を駆使して合成し、A班による活性評価を経て、B班の予測との「答え合わせ」を行う。候補分子の化学構造は多様かつ複雑であると予想されるため、C班が有する独自かつ多彩な合成技術を適切に選択して対応する。またこのスキームを加速するために、保有する基盤合成技術の高度化を並行して行う。この一連のサイクルを通じて潜在空間の予測精度を向上させ、最終的に創出した生物活性分子をケミカルバイオロジー研究用のツールや創薬シーズとして発展させる。

具体的には、A班が取得する生物活性データのインプット、B班が有する複数のアルゴリズムによる情報解析、このアウトプットとして化合物潜在空間からデザインされる生物活性分子の化学構造の閲覧に対応する一連のシステムを構築する。このようにデータのインプット、解析、アウトプットをオンライン上で完結することで、非同期での班間連携が可能になり、化合物潜在空間の開拓に資するデータの蓄積と閲覧、共同研究の進捗管理などをシームレスに参画者全員が共有できる。

