

審 査 の 結 果 の 要 旨

氏名 海東 和麻

本論文は「化学－生物情報を協奏的に用いる創薬システム」と題し、遺伝子発現プロファイル、特殊な化学構造、薬物代謝情報などの利用により医薬品の候補として期待される化学構造式を発生させる手法、並びに薬物代謝酵素に対する代謝部位（Site of Metabolism: SoM）予測に基づく代謝物の構造式出力手法構築に関する研究をまとめたものであり、5つの章から構成されている。

第1章では、研究背景と研究目的を述べている。まず、医薬品開発の現状と課題に始まり、創薬におけるヒット化合物、リード化合物の導出とその最適化について述べている。特に、ヒトにおいて主要な生体異物代謝酵素である Cytochrome P450 (CYP) の重要性と、薬物動態の改善に寄与しうる特殊な化学構造について詳説している。続いて、創薬化学分野で近年盛んに研究が行われている構造生成器、SoM 予測モデルを中心とする機械学習手法、そしてオミクスデータの利用について、それぞれ詳説している。これらの背景を踏まえ、分子設計において考慮すべき要素として、1. 設計した分子は包括的な生物応答情報が反映されている必要があること、2. 設計した分子は機能最適化を志向した構造展開がなされる必要があること、3. 設計した分子の代謝物を化学的解釈可能な形で予測する必要があることの3点を挙げている。その上で、これらの要素を満たす“化学－生物情報を協奏的に用いる創薬システムの開発”を本論文の目的と設定し、その達成のための具体的な課題として、遺伝子発現プロファイルを入力とする構造生成器の開発、中心骨格を固定するハイブリッド型構造生成器の開発、新規記述子と構造変換システムによる代謝物予測手法の開発を掲げている。

第2章では、遺伝子発現プロファイルを入力とした新規化学構造を発生する構造生成器について論じている。遺伝子発現プロファイル解析と Variational Autoencoder (VAE) を融合させた構造生成器 **TR**anscriptome-based **I**nference and generation **O**f Molecules with desired **PHE**notypes by machine learning (TRIOMPHE) の概要と構造生成の結果について詳述している。TRIOMPHE は、標的タンパク質の発現量に摂動を与えることで得られた標的摂動プロファイルと、化合物添加により得られたプロファイルとの相関係数に基づいて発現類似化合物を選択し、VAE に入力することで新規化合物の構造式を出力している。VAE 構築に用いる化合物データの表現法として SELFIES を用いることで、新規化合物を効率的に発生させ得ることを示している。ケーススタディとして、既存のリガンドが存在する標的タンパク質の遺伝子発現プロファイルから新規化合物を発生させた事例を示しており、既往手法よりも標的タンパク質に対するリ

ガンドとして妥当な構造が得られることを述べている。更に、オーファンタンパク質に対するリガンドの設計事例についても説明している。

第3章では、中心骨格を保持しつつ、ビシクロ[1.1.1]ペンタンなど、これまで創薬化学において十分な検討がなされてこなかった特殊な構造を含む化合物を網羅的に発生させる構造生成器 **Exhaustive Molecular library Production In a scaffold-REtained manner (EMPIRE)** について論じている。EMPIRE の特徴が、現実的に存在しうる構造式を取得しやすいが、事前にフラグメントを組み合わせるルールを設計する必要があるビルディングブロック型構造生成器と、膨大なルール設計は不要であるが、学習に膨大なデータセットを必要とする深層学習型構造生成器とを組み合わせることで、互いの補完が達成されることにあることを説明している。また、化学構造の主骨格を保持する既往手法と比較を行い、提案手法である EMPIRE がビシクロ[1.1.1]ペンタンやキュバンなどの特殊な化学構造やホウ素原子、ケイ素原子を含む構造式を網羅的に発生可能であることを示している。

第4章では、化合物の部分電荷を用いた新規記述子の利用による、予測性能と解釈性を担保した CYP 基質分子の SoM 及び代謝物予測システムについて論じている。構造式から算出される部分電荷である Gasteiger Charge と、構造式中のトポロジカル的性質を組み合わせた新規記述子を考案し、SoM 予測に有用であることを示している。また、SoM 予測モデルの予測結果に基づき、代謝物の構造式を予測するシステムである **MEtabolite Rule-based Converter by Index (MERCI)** を構築し、広範な代謝反応に由来する代謝物の構造式を正確に出力可能であることを明示している。

第5章では、各章に記載された成果を総括し、それを踏まえた課題と今後の展望について論じている。提案手法を組み合わせることで、遺伝子発現プロファイルを入力とし、特殊な構造を含む構造式の発生と代謝物予測を行った事例を示している。また、提案手法の実践的創薬適用を念頭に置き、周辺部で整備すべき方法論や、創薬システムに用いるデータベースの整備、システム更新体制等について明示している。創薬システムの展望として、希少疾患や感染症に対する治療薬創製、材料開発を念頭に置いた産業応用、多様な創薬モダリティへの展開を論じている。

以上、本論文は、生物応答情報として遺伝子発現プロファイルを利用した構造発生と特殊な構造を含む構造展開、さらには生成構造の代謝物評価を行うシステム開発に関する一連の成果をまとめたものであり、創薬における課題を解決するための新規手法となる得るものである。一連の研究成果により、多様な標的分子に対するヒット化合物の創製、性質改善を志向した特殊な構造を含むリード化合物の創製、並びに毒性代謝物を含む代謝物の予測が可能となると期待され、本論文の成果は、化学システム工学、データ駆動型化学、ひいては創薬研究の発展に大きく貢献するものと判断される。

よって本論文は博士（工学）の学位請求論文として合格と認められる。