

審査の結果の要旨

氏名 宮尾 知幸

本論文は「A Study on Chemical Structure Generation Based on Inverse Quantitative Structure-Property Relationship/Quantitative Structure-Activity Relationship」(和文: 定量的構造物性相関/定量的構造活性相関モデルの逆解析を利用した化学構造創出に関する研究)と題し、定量的構造物性相関 (QSPR) /定量的構造物性相関 (QSAR) モデルを逆解析 (inverse QSPR/QSAR) することで、特定の物性や活性を有すると期待される化学構造を、コンピュータ上で効率的に生成するシステム構築に関する研究結果をまとめたものである。本論文は英文で書かれ、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が明示されている。QSPR/QSAR を利用した分子設計の背景として原子団寄与法を利用した古典的手法の説明から、仮想的スクリーニング、化学構造をゼロから構築する *de novo design* の説明などのコンピュータを利用した分子設計の背景が説明された後に、研究の主題である inverse QSPR/QSAR に関する既往の研究が概説される。既往の研究の課題として、1. 逆解析においてモデルの適用範囲 (AD) が考慮されていないこと、2. 線形モデル (MLR) 使用による低い予測精度、3. 限定的な記述子の使用、4. モデルに依存しない包括的な適用範囲 (universal AD) が考慮されていないこと、の4点が挙げられている。それぞれの問題点のために、inverse QSPR/QSAR の利用が制限されている現状を説明した後に、本研究の目的として、上記4つの課題を克服し、inverse QSPR/QSAR を利用した実用に足る化学構造生成システムを開発することが明示されている。

第2章では、化学構造生成機に用いられるアルゴリズムが詳述されている。課題3を克服するための手法として、単調変化記述子 (MCD) を化学構造生成に組み合わせることが提案されている。MCDのみを使用した場合であっても、包括的な記述子セットを使用した場合と比較して遜色のない予測精度を持つモデルが構築できることが、ケーススタディを通して示されている。また、課題4に対応するために、リングシステムと原子フラグメントを効率的に組み合わせることで化学グラフを構築する新規アルゴリズムが疑似コードを用いて説明されている。提案アルゴリズムを利用した化学構造生成速度は、既存の構造生成機による生成速度と比較して高速であることがケーススタディを通して示されている。また、構造生成に関連するアルゴリズムとして、多様な構造を生成するアルゴリズム、網羅的に構造生成をおこなった場合に生成する構造数の期待値を計算するアルゴリズム、構

造生成中に MCD を高速に計算するアルゴリズム、の 3 つのアルゴリズムが提案され、疑似コードによる説明とケーススタディによる実例が示されている。

第 3 章では、目的変数 (y) の値に対応する記述子 (\mathbf{x}) の情報を逆解析により取得する新規手法が説明されている。提案手法は、モデルの適用範囲を考慮していないこと (課題 1) と MLR による低い予測精度 (課題 2) を克服することを目的としている。提案手法は、混合正規分布により \mathbf{x} の事前分布を仮定し、それぞれの正規分布に対してガウス誤差を仮定した線形重回帰モデル (GMMs/cMLR) を構築することで、 y の値を所与とした際の \mathbf{x} の事後分布を解析的に導出する。この事後分布がモデルの適用範囲を考慮した指標になることが、水溶解度を y とした inverse QSPR 解析により示されている。また、回帰モデル構築手法として、GMMs/cMLR が MLR よりも優れていることが、シミュレーションデータを用いた解析と、Alpha1B, 1D, 2A, 2C アドレナリン受容体に対しての QSAR 解析から示されている。

第 4 章では、提案する inverse QSPR/QSAR を利用した化学構造生成システムが説明されている。このシステムは、第 2 章で説明した構造生成機と第 3 章で説明した逆解析の 2 つの部分をつなぐことにより構成される。両手法をつなぐためには、逆解析により取得した y の値を所与とした \mathbf{x} の事後分布の高密度領域を構造生成機に入力するための制約条件へと変換する必要がある。そのための手法として、事後分布である混合正規分布の一つの正規分布の中心を構造生成の制約条件とすることが提案されている。特に、離散値のみを持つ記述子に対しても高密度領域を規定するために、標的とした正規分布の主軸方向へ格子点を探索することが提案されている。提案システム、提案手法を分子設計に適用した一例として、トロンビンを標的とした inverse QSAR 解析が実施され、その結果についても詳細な検討がなされている。提案したシステムを利用することにより、モデルの適用範囲を満たし、予測親和性 (pK_i) が望ましい多数の新規構造を提案しており、提案システムの実用性が示されている。

第 5 章では、各章に記述された成果が総括されている。研究成果とともに課題についても明示されている。また、本論文のインシリコ分子設計における位置づけが、創薬、材料設計の二つの視点から論じられている。

以上述べたとおり、本論文は inverse QSPR/QSAR により、特定の物性または活性をもつ化学構造を生成するシステム開発についての十分な成果をまとめたものであり、既往の課題を克服するための新規手法の提案、それに付随する検討が詳述されている。一連の研究成果は、実験等で取得されたデータを有効活用することで分子設計を行うという、データ駆動型研究分野の進展を促すものと認められ、ケモインフォマティクスおよび化学システム工学の進展に貢献するものであると判断される。

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。