

氏 名	オ 吳 ミン ア 敏 兒
学位(専攻分野)	博 士 (薬 学)
学 位 記 番 号	薬 博 第 609 号
学位授与の日付	平 成 19 年 3 月 23 日
学位授与の要件	学 位 規 則 第 4 条 第 1 項 該 当
研 究 科 ・ 専 攻	薬 学 研 究 科 創 薬 科 学 専 攻
学 位 論 文 題 目	Prediction of biodegradation pathways using knowledge base of enzymatic reactions (酵素反応の知識ベースを用いた環境物質分解経路予測)
論文調査委員	(主 査) 教 授 金 久 實 教 授 馬 見 塚 拓 教 授 辻 本 豪 三

論 文 内 容 の 要 旨

代謝パスウェイは、生体内化合物と酵素反応で構成された反応経路グラフであり、それらのうち生化学的実験を経て既知となった部分パスウェイは、たとえば KEGG PATHWAY に代表されるパスウェイデータベースに収録されている。しかしながら、登録されるデータは、基本的に生化学的実験の成果に頼っている部分が大きく、データ量やその片寄りに関する問題を避けて通ることが出来ない。たとえば薬物代謝のように、細胞にとって外来因子と見られる化合物が如何に代謝され無害化されていくのか、あるいは無害化されずに毒性を発揮するのかというメカニズムについては、ほとんど未解明のままである。また、植物の二次代謝産物のように、利用される化合物の構造は分かっているのにそれらを繋ぐ酵素反応が未知のまま残されているものもある。これらのうち、特に生命系維持にとって問題となりうる外来化合物群について、生分解パスウェイの解析を行ったり反応経路の予測法を開発したりすることは、創薬における薬物の安全性評価のほか、環境物質の浄化などを考察する上で、きわめて重要である。このような目的の下、本論文では、代謝経路を構成する酵素反応の反応メカニズムを原子レベルで表記する手法の開発とその網羅的解析を行ったうえで、それに基づく代謝反応経路予測法を開発した。また、代謝経路が未知のいくつかの環境因子に対してこの手法を適用し、予測性能の評価を行った。

まず、酵素反応の反応メカニズムに関しては、大きく次の3つの手順に従って、原子レベルでの情報の抽出を行った。

- 1) 酵素反応前後において関連している化合物ペアの間で化合物構造を比較する。
- 2) 原子アライメントの結果から、反応の前後で変化していない部分 (Match 領域)、付加または脱離反応によって変化している部分 (Difference 領域)、および Match 領域と Difference 領域の境界にあって反応の中心的な舞台と考えられる部分 (Reaction center)、の3つの部分構造情報を抽出する。
- 3) 上記の3つ部分構造情報から、それぞれ、M 原子、D 原子、および R 原子と呼ぶ原子情報を導出し、酵素反応の原子変換パターン (RDM パターン) を構築する。

このように抽出された RDM パターン情報は、基質特異性さえ考慮しなければ、酵素反応の反応メカニズムの本質的な部分を捉えていると考えられ、従来の解析からも、EC 番号分類 (酵素反応分類) とのよい相関があることが指摘されている。この情報は、RPAIR データベースとして KEGG LIGAND データベースに収録されており、現段階で総計7,093エントリー (化合物ペアとして) である。これを、KEGG PATHWAY データベースに出現する4,256酵素反応に関して網羅的に解析した結果、反応メカニズムとして1,901個の原子変換パターンを得ることが出来た。さらに、この1,901個の RDM パターンの出現を11種類に分類される代謝パスウェイのカテゴリー毎に解析すると、1,576個の RDM パターンは、11カテゴリーのどこかでのみ観察される特徴的な反応パターンであることが分かった。また、生分解経路に特徴的に出現する RDM パターンも観察された。特に、ジオキシゲナーゼによって触媒される芳香環のジヒドロキシル化反応、芳香環から伸びるアルキル鎖を、アルコール、さらにカルボキシル基へと変換する一連の酸化反応が、特徴的であった。このように、特徴的な反応メカニズムが特定の代謝マップ毎に繰り返し利用される傾向があることを見出すことが出来たが、以前の解析から、代謝

マップに登場する化合物については、より構造の類似したものが反応パスウェイの道のり的にもより近くに局在する傾向があることが指摘されている。今回の結果は、そのような傾向が反応レベルでも保持されていることを示すものであり、また、その傾向は生分解経路などで顕著である。

代謝反応経路の予測法は、特に環境物質の分解経路予測に焦点を絞って行った。本論文では、KEGG PATHWAY データベースに収録されている“xenobiotics biodegradation and metabolism” カテゴリーの RDM パターンと化合物構造セットを利用している。予測の流れは、

- 1) まず、クエリー化合物構造を、化合物構造比較法を用いて化合物構造セットから検索する。
- 2) 検索結果のうち構造類似性スコアが高く、かつ RDM パターン周辺がきちんとアライメントされており、さらに、本論文で独自に決めた生分解性スコアの優位なものを抽出する。
- 3) データセットから抽出した RDM パターン（既知の反応パターン）をクエリー化合物に適用し、新しい化合物構造をグラフ理論的に創出する。

というものである。手順のうち構造類似性の高いものを採るのは、酵素反応における酵素の基質特異性を反映させるためである。このプロセスによって、酵素反応における基質特異性（構造類似性）と反応特異性（RDM パターン）の両方の情報を取り入れることができる。また、新規に創出された化合物構造を、手順 1) における新たなクエリー化合物とすることによって、酵素による代謝反応パスウェイを次々に再構築することが可能である。本論文では、この予測法を用いて、環境物質の一つ、1, 2, 3, 4-Tetrachlorobenzene の分解経路の予測を試みた。その結果、米国ミネソタ大学で構築されている生分解経路データベース UM-BBD に登録されている 1, 2, 3, 4-Tetrachlorobenzene の生分解経路（生化学的な知識）と完全に一致していた。さらに、UM-BBD にのみ登録されている 120 個の化合物構造について、それぞれ上記の分解経路予測を行った結果、構造類似性スコアが比較的高い分子がデータセットに存在する場合には、予測率がきわめて良好であることが分かった。一方で、利用できる化合物構造、酵素反応情報が限定的であるため、特殊な化合物構造や、稀にしか使われない酵素反応に関しては、正解率は低かった。

本論文では、酵素反応変換パターンのデータベースの構築、それに基づいた生分解経路の特徴的な反応パターンの導出、そして、生分解経路の代謝反応経路予測法の構築と評価を行った。このような予測手法の開発は、ポストゲノム時代におけるパスウェイ解析において、新薬の候補化合物を選出する際に有用であるばかりか、主にバクテリアに担われている環境浄化メカニズムを原子レベルで解明する際にも有効であると考えている。

論文審査の結果の要旨

化合物の代謝経路情報は KEGG PATHWAY などに代表されるパスウェイデータベースに収録されてきたが、一方で、登録されているデータに関する問題も指摘されてきた。その多くは、データ量そのものが少ないこと、また、登録されるデータ（代謝反応経路）に片寄りがあり細胞内代謝を網羅的に収集できていないことなどである。実際、薬物代謝のような外来化合物が生体内で代謝され常在化合物へと無害化されていく経路の多くが未知であったり、植物の二次代謝経路に見られるような、代謝を担う酵素反応（酵素遺伝子）そのものが未知であるなど、代謝経路を網羅的に解析する際には支障となる場合が多々ある。本論文において、申請者は、特に、細胞外由来の化合物群が細胞内で代謝されるべき経路、すなわち化合物の代謝運命を動的に予測することを目的としている。このような、生命系維持にとって問題となりうる外来化合物群について、生分解パスウェイの解析を行ったり反応経路の予測法を開発することは、創薬における薬物の最適化や安全性評価のほか、環境物質の浄化などを考察する上で、きわめて重要である。本論文において、申請者は、代謝パスウェイデータベースに登場する全酵素反応の反応メカニズムをグラフィ的に表現する手法の開発と、それを用いた網羅的解析を行ったうえで、その解析結果を基に、薬物の生分解経路を予測する手法を開発した。さらに、米国ミネソタ大学で構築されている生分解経路データベース UM-BBD に登録されているが、KEGG には未登録の化合物に対して生分解経路予測を実施し、本手法の有効性を示すことに成功している。ここでは、反応メカニズムの解析、反応経路予測法の開発、UM-BBD に対する適用と有用性評価、の 3 点に関して、評価を行った。

まず、酵素反応の反応メカニズムについて、申請者は、原子レベルでの情報を抽出した KEGG RPAIR データベースの

解析を行った。このデータベースでは、1) 酵素反応前後において原子の由来関係を保持している化合物ペアを定義し、2) それらの間で原子アライメントを取ることによって、反応前後での原子の変化パターンを抽出する。このとき、変化していない (M)、付加または脱離反応によって変化している (D)、またはその両者の境界にある反応中心 (R)、の3つの部分構造情報を抽出し、酵素反応の原子変換パターン (RDM パターン) を構築している。このような、グラフを用いて反応変化の情報を抽出する手法は、すでに多くの研究がなされているところであるが、反応中心の周辺領域のみを高度なグラフ計算に基づいて高精度に抽出し、かつ、必要に応じて適宜手作業による修正を行っているところが重要である。これは、全酵素反応に対して、関与する反応メカニズムをアノテーションする作業に他ならず、その結果得られたデータベースは、非常に有用なものとなることが期待される。実際、このように抽出された RDM パターン情報は、従来の解析によっても、EC 番号分類とのよい相関を示すことが判明している。この情報を、KEGG PATHWAY データベースに出現する全酵素反応に関して解析した結果、1,901個の RDM パターンうち1,576個は、11カテゴリーのどれかでのみ観察される特徴的な反応パターンであることも分かった。特に、生分解経路の反応は特徴的であり、他の代謝経路との違いが浮き彫りにされている。このような、代謝経路を反応のメカニズムの視点から分類・考察する解析は、過去に例がなく、非常に興味深いものといえる。

代謝経路の予測法について、本論文では、特に環境物質の分解経路予測に焦点を絞って行っている。具体的には KEGG PATHWAY データベースに収録されている“xenobiotics biodegradation and metabolism”カテゴリーの RDM パターンと化合物構造を利用し、まず、クエリー化合物と構造類似性の高い化合物の探索を行う。探索結果から、さらに、RDM パターンとなるべき部分構造の周辺領域がきちんと原子アライメントされている（すなわち RDM が定義可能である）ものを選出する。その後、選ばれた化合物が持っている RDM パターン（すなわち既知の反応パターン）をクエリー化合物のグラフ構造にも適用し、新しい化合物構造グラフを生成する。以上の手順のうち、前半部の構造類似性の高いものを採用する部分は、酵素反応における酵素の基質特異性を反映させるためであり、RDM パターンが酵素反応のメカニズムの情報のみしが入り入れられていないという欠点を補うためのものである。したがって、この前段階の処理は、十分に合理的であるということが期待される。また、最終的に生成された化合物構造を新たなクエリー構造とすることによって、代謝反応パスウェイの再構築（パスウェイの予測）が可能になる。もちろん、部分グラフの変換パターンの適用によるパスウェイ予測は、現在のパスウェイ予測の主流となっているものであるが、本手法には、他の手法とは異なり、全酵素反応を網羅的に収集し、そのために、適用できる範囲が広いという利点がある。

最後に本論文で提案されている手法の有用性についてであるが、申請者が本予測法を用いて、環境物質の一つ1, 2, 3, 4-Tetrachlorobenzeneの分解経路の予測を試みた結果、米国ミネソタ大学で構築されている生分解経路データベース UM-BBD に登録されている生分解経路と、完全に一致していることが判明した。UM-BBD に登録されているパスウェイは生化学的に決定された知識であり、本手法において申請者が構築した一連の手順の妥当性が、一定の評価を得るに値するものであることが示されている。さらに、UM-BBD にのみ登録されており、本論文の執筆段階で KEGG には登録されていなかった120個の化合物構造について、本手法による分解経路予測を行っている。その結果、構造類似性スコアが比較的高い分子がデータセットに存在する場合には、極めて高い予測率を示すことがわかり、一方、構造類似性の高い化合物が見つからない場合には、予測自体を行うことが困難であることが示された。予測不可能なクエリー化合物が存在するということは、本手法を適用する上での限界を示すものであり、その点に関しては、まだまだ改善の余地が多いものと思われる。ただし、現在データベースに登録されている化合物構造、および酵素反応情報が限られているためであると考えられることもでき、その場合、将来的には本手法の予測率については向上していくことが期待できるものであることも考慮されるべきである。これこそが、データベースを基に知識を抽出するという、本手法の利点を端的に表すものであると考えられることも出来る。

以上のように、本論文で提案された代謝経路の解析手法、および予測手法は、ポストゲノム時代におけるパスウェイ解析において、特に、新薬の候補化合物を選出する際に有用であることが期待されるものである。よって、本論文は博士（薬学）の論文として価値のあるものと認める。さらに、平成19年2月22日に論文内容とそれに関連した事項について試問を行った結果、合格と認定した。