



2022年4月21日

分野： 自然科学系

キーワード： 化学反応、遷移状態、分子動力学シミュレーション、人工知能、説明可能な AI

AI が化学反応の行方を説明してくれる！ —コンピュータシミュレーションに対して説明を与える人工知能の応用—

【研究成果のポイント】

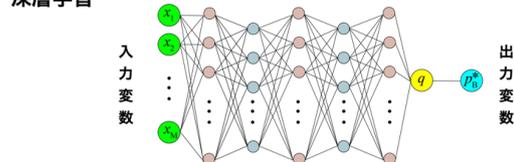
- ◆ コンピュータシミュレーションと人工知能の融合により、化学反応の遷移状態^{*1} を特定し反応の成否を決める要因を突き止める方法論を開発
- ◆ 反応物と生成物の中間にある遷移状態は不安定であり、実際にどの分子構造が遷移状態であるかを理解することは困難であったが、「説明可能な AI」^{*2} と呼ばれる人工知能の技術 XAI により情報を引き出すことに成功
- ◆ 今回の研究により化学反応による物質創成へ人工知能を応用する理論的方法を深化させることが期待

❖ 概要

大阪大学大学院基礎工学研究科の菊辻 卓真さん（博士後期課程修了）、森 勇介さん（博士前期課程修了）が金 鋼准教授、松林 伸幸教授と共に、分子科学研究所の岡崎 圭一准教授と九州大学の森 俊文准教授と協働し、「説明可能な AI」と呼ばれる人工知能の技術 XAI により、化学反応において活性化した状態の遷移状態にある分子構造を説明できることを世界で初めて明らかにしました。

一般に化学反応は反応前にある反応物と反応後の生成物を区別し、その中間に遷移状態と呼ばれる活性化状態を取ります。遷移状態を正確に理解することは、反応にどれだけの時間がかかるのかを正確に予測することに直結するだけでなく、より高い効率で生成物の獲得を目指す反応工学において重要であり、これまでも遷移状態に関する理論が数多く提案されてきました。しかしながら、遷移状態は極めて不安定な状態であり、実際にどの分子構造が遷移状態に対応するのかを理解することは難しいとされてきました。

深層学習



生成物へ行く確率 P_B^*

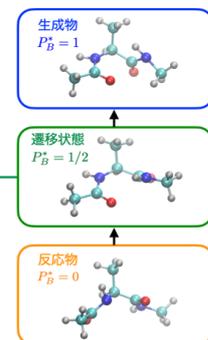
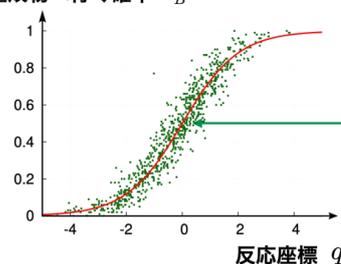


図 1：(上) 分子構造座標を入力変数、生成物へ行く確率 P_B^* を出力変数として深層学習し、反応座標 q を予測する。(下) 確率 P_B^* のデータ(緑点)と反応座標 q の関係を示す。赤線が深層学習の目的関数である。確率 $P_B^* = 1/2$ である点が遷移状態にある分子構造に対応している。このデータに「説明可能な AI」を用いることでどの入力変数が寄与しているかを明らかにした。

特に、コンピュータシミュレーションによって分子構造のデータが多数収集されるなかで、どれが遷移状態なのかを特定し、なぜ反応物へ戻らず生成物の方向へ反応が進むのかを説明することが求められていました。今回、共同研究グループは、タンパク質の構造変化（異性化反応）についてのコンピュータシミュレーションと「説明可能な AI」と呼ばれる人工知能の技術 XAI を融合する研究を展開し、遷移状態となる分子構造を特定するだけでなく、化学反応の成否を決める要因が何かを説明できることを明らかにしました。これにより、人工知能を用いることで、遷移状態を的確に予測し、所望の反応物を高効率に選択する理論的方法を構築しました。本研究成果は、米国物理学協会が発行する Journal of Chemical Physics 誌の 2022 年 4 月 21 日（木）出版号に掲載されます。（報道解禁設定はありません）

❖ 研究の背景

一般に化学反応では、反応物から生成物への反応と、逆に生成物から反応物への反応とが同時に進行しており、全体として反応物と生成物の割合が釣り合った平衡状態が達成されます。ここで、どのように化学反応が進行するかを理解するためには、反応を活性化させる中間状態としての遷移状態を理解することが鍵となり、これまで遷移状態理論と呼ばれる理論体系が開発されてきました。しかしながら、遷移状態は非常に不安定な状態であることから、遷移状態にある分子構造そのものを観測できる頻度が極めて少ないという課題がありました。このことは、化学反応だけでなく、タンパク質の折り畳みなど、安定な分子構造間の遷移過程のメカニズムを理解する上でも解決されるべき重要な問題であると考えられています。

❖ 研究の内容

今回、共同研究グループは、分子動力学シミュレーションと呼ばれるコンピュータシミュレーションに、「説明可能な AI」と呼ばれる人工知能の技術 XAI を融合させることにより、タンパク質の基本要素であるポリペプチド鎖^{※3}の主鎖二面角に関する異性化反応における遷移状態を特定し、さらに反応をもたらすメカニズムを説明できることを実証しました。

研究ではまず、分子動力学シミュレーションにより、ポリペプチド鎖の分子構造を多数収集します。そして、それぞれの分子構造を出発点として、どの程度生成物へ行くかを「確率」として定量化します。つまり、反応物にあれば確率 0、生成物にあれば確率 1、中間にある遷移状態にあれば確率 1/2 になるはずですが。

次に、深層学習^{※4}を用いて、出発点にある分子構造の座標を入力変数、生成物へ行く確率を出力変数として両者の関係を学習させます（図 1 上）。これにより、反応物から遷移状態を経由し生成物へ至る、化学反応の経路に沿った「反応座標」^{※5}を予測します。ただし、予測された反応座標は分子構造座標の非線形結合により表現される抽象的なものであり、このままでは反応のメカニズムを理解することができません。

そこで、「説明可能な AI」と呼ばれる人工知能の技術により、この課題を解決することを着想しました。具体的には、様々な構造における上記の「確率」データに対して、入力変数である分子構造座標のうち、どの変数が反応座標へ寄与しているのかを定量化しました。その結果、遷移状態（確率が 1/2 となる構造）において重要な変数を特定し、異性化反応をもたらすメカニズムを理解することに成功しました（図 1 下）。

❖ 本研究成果が社会に与える影響（本研究成果の意義）

本研究成果は、コンピュータシミュレーションと人工知能を融合させることにより、化学反応の遷移状態を説明できることを実証した例になります。遷移状態にある分子構造を的確に予測することは、特定の反応経路を高い選択性で選べるようになることに直結します。また、遷移状態を効率良く収集し、さらに説明可能なモデルを与えることができれば、タンパク質の折り畳みや酵素反応など多数の安定状態の間を複雑に構造変化するメカニズムを理解するのにも応用できます。化学反応による物質創成に関係する様々な問題へ今回の理論的方法を応用すれば、エネルギー、環境問題だけでなく医薬創成へも貢献できることが期待されます。

❖ 特記事項

本研究成果は、米国物理学協会が発行する Journal of Chemical Physics 誌の 2022 年 4 月 21 日(木) 出版号に掲載されます。(報道解禁設定はありません)

タイトル：“Explaining reaction coordinates of alanine dipeptide isomerization obtained from deep neural networks using Explainable Artificial Intelligence (XAI)”

著者名：Takuma Kikutsuji, Yusuke Mori, Kei-ichi Okazaki, Toshifumi Mori, Kang Kim, and Nobuyuki Matubayasi

DOI：https://doi.org/10.1063/5.0087310

なお、本研究は、科学研究費補助金基盤研究(B)、基盤研究(C)、新学術領域研究「水圏機能材料：環境に調和・応答するマテリアル構築学の創成」公募研究の支援を受けて行われました。また、本研究のコンピュータシミュレーションには、自然科学研究機構岡崎共通研究施設・計算科学研究センターと大阪大学サイバーメディアセンターのスーパーコンピュータを用いました。

❖ 用語説明

※1 遷移状態

化学反応において反応物から生成物へ遷移する過程の途中にある不安定な活性化状態のことを言う。遷移状態理論においてはポテンシャルエネルギー曲面の鞍点に相当する。

※2 説明可能な AI

深層学習の発展により高度な画像、テキスト、音声認識ができるようになったが、深層学習はブラックボックス型の学習モデルであり、なぜ内部でそのように予測したのか理解しにくいのが課題である。そこで学習結果に説明性を与える「説明可能な人工知能 (Explainable AI: XAI)」の技術が注目を浴びている。本研究では、XAI のうち局所的な説明性を与える LIME(Local Interpretable Model-agnostic Explanations)と SHAP(SHapley Additive exPlanations)の 2 種類の手法を用いて遷移状態を予測したデータに対する入力変数の寄与率を定量化した。

※3 ポリペプチド鎖

アミノ酸は重合しポリペプチド鎖を形成し、タンパク質の 1 次構造となる配列を構成する。ポリペプチド鎖には立体障害によって主鎖二面角の取りうる範囲が決まっている。主鎖二面角に関する異性化反応の遷移状態を知ることはタンパク質折り畳みの理解につながる基本的課題である。

※4 深層学習

生物のシナプス結合による神経回路網を模した機械学習モデルを人工ニューラルネットワークと言う。さらに、入力層と出力層のあいだに多数の中間層を用意し、データに含まれる特徴を各層に学習させるモデルのことを深層学習もしくはディープラーニングと言う。

※5 反応座標

化学反応の反応経路に沿って表現される、分子間の結合長や結合角などによる幾何学的な座標のことを言う。

❖ SDGs 目標



❖ 参考 URL

金 鋼 准教授 研究者総覧 URL <https://rd.iai.osaka-u.ac.jp/ja/02b4b79d5b8a2a8a.html>

松林 伸幸 教授 研究者総覧 URL <https://rd.iai.osaka-u.ac.jp/ja/f75999b6aa244763.html>

❖ 本件に関する問い合わせ先

大阪大学 大学院基礎工学研究科 准教授 金 鋼 (きん こう)

TEL : 06-6850-6276

E-mail: kk@cheng.es.osaka-u.ac.jp

自然科学研究機構 分子科学研究所・計算科学研究センター 准教授 岡崎 圭一 (おかざき けいいち)

TEL : 0564-55-7468

E-mail: keokazaki@ims.ac.jp

九州大学 先導物質化学研究所 准教授 森 俊文 (もり としふみ)

TEL : 092-583-7800

E-mail: toshi_mori@cm.kyushu-u.ac.jp