

## 計算化学と機械学習を駆使した機能性材料の理解と設計

(慶大理工<sup>1</sup>) ○畑中 美穂<sup>1</sup>

Understanding and rational design of functional materials based on computational chemistry and machine learning (<sup>1</sup>*Department of Chemistry, Faculty of Science and Technology, Keio University*) ○Miho Hatanaka<sup>1</sup>

Machine learning models that calculate the physical properties from information about the materials can be used to predict the physical properties of untested materials, *i.e.* in silico high-throughput screening. However, due to the small data set of metal complex materials, it is necessary to develop reasonable data collection systems and numerical representations. In this talk, I will introduce our recent works in two functional metal complex materials: the palladium complex catalysts and the cerium photofunctional complexes.

**Keywords :** *Metal Complexes, Feature Engineering, Global Reaction Route Mapping*

材料の情報から物性を計算する機械学習モデルは、量子化学計算よりも高速に未実験の材料の物性を予測できるため、*in silico* ハイスループットスクリーニングを可能にし、材料開発加速に貢献できると期待されている。高速、高効率にスクリーニングを行うためには、候補材料の列挙と物性の計算を自動で行うシステムが不可欠である。有機化合物のスクリーニングを行う場合、候補分子を列挙できれば、分子の文字列表現 (SMILES) から立体構造を生成する手法が確立されているため、スクリーニングの報告例は多い。これに対し、金属錯体の場合、(A)文字列表現から立体構造を生成することも、(B)量子化学計算を用いて高精度に物性を予測することも難しいために、報告例は非常に少ない。我々は、この問題(A)を解決するために、ケンブリッジ結晶構造データベース(CSD)<sup>1</sup>を元に金属錯体の立体構造のデータベースを構築し、機械学習による物性予測モデルの構築や、より良い機能性材料の発掘に取り組んでいる。本講演では、セリウム錯体のフォトレドックス触媒とホスフィン配位子を持つパラジウム錯体触媒に関する我々の研究を紹介する。

### 【セリウム-フォトレドックス触媒】

近年、希土類の1つであるセリウム三価陽イオン( $\text{Ce}^{3+}$ )の錯体が、既存のフォトレドックス触媒よりも高い還元力を持ち、活性化が困難な基質の反応にも適用できることが報告されている。<sup>2)</sup> そのため、 $\text{Ce}^{3+}$ 錯体の光物性の理解の深化や、励起エネルギー・寿命の高速な計算によるスクリーニングが求められている。 $\text{Ce}^{3+}$ の電子配置は $[\text{Kr}]4d^{10}4f^15s^25p^6$ であり、基底状態を含む7個の状態( $^2F$ )が擬縮退するため、第7励起状態がフォトレドックス触媒能に関わる。そこで、 $\text{Ce}^{3+}$ 錯体の第7励起状態の励起エネルギーを、基底状態計算で得られる情報から予測する機械学習モデルを構築し、励起エネルギーに寄与する因子の抽出を目指した。

まず、CSDの中から684種のランタノイド単核錯体の結晶データを集め、錯体部分のみを抽出し、ランタノイドをCeに置換した。得られた錯体構造をPM6-D3/Sparkle

レベルで構造最適化し、TD-DFT法を用いて励起エネルギーを計算した。この時、Ceには4s4p4d4f5s5p5d6s電子のみ露わに計算するsmall-core RECPを用いた。次に、Ceの5s5p5d6s電子のみを露わに計算するlarge-core RECPを用いた基底状態計算(DFT計算)を行い、得られたパラメタを予測モデルの説明変数に用いた。

Ce<sup>3+</sup>の第7励起状態は、4f-5d遷移の性質をもつため、Ce<sup>3+</sup>錯体も同様の性質をもつと予測していた。しかし、実際には図1に示すように、Ce錯体の第7励起状態の性質は、錯体に大きく依存しており、4f-5d励起、 $\pi$ - $\pi^*$ 励起、電荷移動型励起など様々な性質をもつ錯体が混在していた。このことから、励起の種類に応じて重要な因子が異なると予想できる。そこで、データの分類と、分類されたデータ内の線形回帰を行うLinearTree法<sup>3)</sup>を用いた機械学習モデルを構築した。類似したモデルであるランダムフォレストモデルでは分岐先のデータの目的変数の平均値を出力するが、LinearTreeモデルは分岐内で線形回帰モデルを作成する。LinearTreeモデルの精度は検証データに対して相関係数 $R^2$ は0.81、RMSEは0.33 eVと頑健であり、ランダムフォレストモデルより優れていた。

#### 【パラジウム-ホスフィン触媒】

パラジウム-ホスフィン触媒は、有機金属錯体触媒の中でも幅広く用いられている触媒の一つであるが、反応や基質の種類に応じて適切なホスフィン配位子が変わるため、高速なスクリーニングが求められている。しかし、異なる分子骨格を持つ触媒に対して共通して利用可能な特徴量が少ないため、類似構造を持つ触媒にのみ適用可能なモデルしか作れていない。そこで、遷移金属錯体触媒を用いる反応が、数種類の素反応の組み合わせで表されることに着目し、特徴量の一つとして素反応の活性化障壁を用いることを提案する。実際に、反応経路自動探索(GRRM)<sup>4)</sup>の一つである人工力誘起反応(AFIR)法を用いることで、モデル反応物の酸化的付加、1,1-挿入、1,2-挿入、トランスメタル化、配位などの反応エネルギーと活性化障壁のデータを収集するシステムを構築した。発表では、これらの特徴量を用いて、フッ化アシルの還元的分子変換反応の触媒能(収率・選択率)を予測するモデルについて議論する。

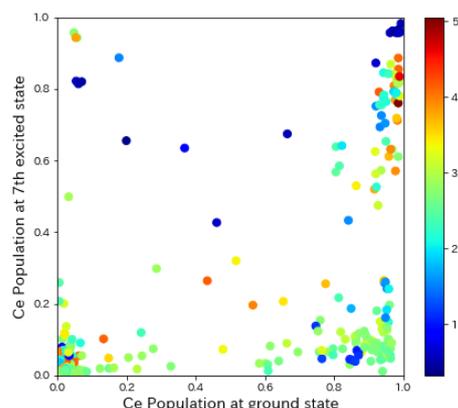


図1：基底状態と第7励起状態におけるCeのポピュレーション。色は励起エネルギー[eV]を表す。

1) C. R. Groom, I. J. Bruno, M. P. Lightfoot, S. C. Ward, *Acta Cryst.* **2016**, B72, 171. 2) Y. Qiao, E. J. Schelter, *Acc. Chem. Res.*, **2018**, 51, 2926. 3) H. Zhang, D. Nettleton, and Z. Zhu, **2019**, <https://arxiv.org/abs/1904.10416>. 4) S. Maeda, K. Ohno, K. Morokuma, *Phys. Chem. Chem. Phys.*, **2013**, 15, 3683.