

機械学習と分子シミュレーションを用いた有機半導体の精緻な集合体構造予測

(北里大院理¹・東大院新領域²・物材機構³・JST CREST⁴・科学大物質理工⁵・北里大未来工⁶・神奈川県産総研⁷) ○關 拓和¹・篠崎 雄大¹・佐藤 俊輔¹・伊藤 良将¹・竹谷 純一^{2,3,4}・岡本 敏宏^{4,5}・渡辺 豪^{1,4,6,7}

Accurate Prediction of Assembled Structures of Organic Semiconductors Using Machine Learning and Molecular Simulation (¹Graduate School of Science, Kitasato University, ²Graduate School of Frontier Sciences, University of Tokyo, ³NIMS, ⁴CREST, JST, ⁵School of Materials and Chemical Technology, Institute of Science Tokyo ⁶School of Frontier Engineering, ⁷KISTEC) ○Takuya Seki,¹ Yudai Shinozaki,¹ Shunsuke Sato,¹ Ryosuke Ito,¹ Jun Takeya,^{2,3,4} Toshihiro Okamoto,^{4,5} Go Watanabe^{1,4,6,7}

The performance of organic semiconductors (OSCs) is significantly influenced by their crystal structures. Traditionally, molecular design has depended on the expertise and intuition of the researchers and does not make effectively use of computational methods. In this study, we introduce a machine learning model that can predict whether a molecule will form a herringbone (HB) packing structure. Furthermore, the combination of molecular simulations with the machine learning model has the potential to predict the crystal structure of OSCs. Our model, trained on the crystal structures of 210 previously reported OSC molecules and utilizing MACCS keys molecular descriptor with the LightGBM algorithm, achieved an accuracy of 91.0%. Analysis of SHAP values revealed key substructures that contribute to the formation of HB packing. By combining of this machine learning model with a novel crystal structure prediction method based on molecular simulations, it was possible to determine the crystal structures of a molecule.

Keywords : *Crystal Structure Prediction; Machine Learning; Molecular Dynamics; Organic Semiconductor*

有機半導体の高機能化には、結晶構造を考慮した分子設計が重要である。しかし現状、分子設計は実験研究者の知識や経験則に依存しており、計算科学的手法の効果的な活用が課題となっている。そこで本研究では、機械学習による有機半導体の π 系骨格の二次元パッキング構造の予測と分子シミュレーションを組み合わせた結晶構造予測¹⁾を目指した。既報の有機半導体 210 分子の結晶構造から、分子構造と二次元パッキング構造のライブラリを作成して、ヘリングボーン (HB) 様式の充填構造の予測モデルを構築した。複数のモデルを検証した結果、分子記述子 MACCS keys と機械学習モデル LightGBM を用いることで、正解率 91.0% で HB パッキングを予測できた。構築したモデルについて SHAP 値を算出することで、HB パッキング発現に寄与する部分構造を同定した。また、我々が提案した分子力学計算と分子動力学計算の組み合わせによる結晶構造予測手法²⁾と構築した機械学習モデルを用いることで結晶構造を一意に決定できる可能性を見出したが、その詳細は当日報告する。

1) T. Seki, G. Watanabe *et al.*, *ChemRxiv* (2024).

2) 渡辺 豪 他, 特開 2024-004296, 2022-0628.