



# 機械学習を活用した有機分子の結晶構造予測

谷口卓也 Takuya TANIGUCHI

筆者は、機械学習を活用して有機結晶の構造相転移を予測することを目的とし、JSTの戦略的創造研究推進事業 ACT-X「トランススケール」領域の研究課題を進めてきた。構造相転移は、分子レベルの構造や配列の変化が結晶のマクロな機能創出につながる点に特徴があるトランススケール現象であり、その発現を予測することでマテリアル創製への貢献を目指している。構造相転移を発見するための第一ステップとしては、分子構造のみから結晶構造を予測する必要があり、本稿では機械学習を結晶構造予測に活用した筆者の最近の研究事例を中心に紹介する。

## 有機分子の結晶構造予測と計算の難しさ

有機分子の結晶構造は、医薬品の溶解性や安定性、あるいは有機半導体の電気伝導性などの物性を直接左右するため、その構造を事前に正確に予測することは研究開発において極めて重要である。しかし、有機結晶はファンデルワールス力や水素結合といった弱い分子間相互作用によって安定化されており、わずかなエネルギー差で多数の結晶構造（多形）を取り得る。このため、実験で得られる結晶構造を理論的に再現・予測することは容易ではない。

有機分子の結晶構造予測 (Crystal Structure Prediction : CSP) は、分子構造のみを入力として、現実で現れる安定な結晶構造を計算によって求める手法である。CSP は一般に、結晶構造候補を生成する「構造探索」と、それらを格子エネルギー（結晶中の分子間相互作用エネルギーの総和）の評価によって選別する「構造緩和・評価」の2つのステップから構成される。構造探索には準ランダム法や遺伝的アルゴリズムなどが用いられてきたが、これらの手法では低密度構造や熱力学的に不安定な構造も多数生成されてしまう。

このような「不要な構造候補の大量生成」が、CSP

の計算コストを押し上げる要因の1つである。近年、ニューラルネットワークポテンシャルの登場によりエネルギー評価の計算負荷は大きく低減されつつあるが、探索段階での非効率性の問題は、依然として残されている。

## 機械学習による探索設計の導入：SPaDe-CSP

このような背景の下、筆者は「探索すべき構造空間をいかに削減するか」がCSPの課題であると考えた。そこで開発したのが、機械学習を用いて結晶構造探索の初期条件を設計するCSPワークフロー「SPaDe-CSP」である<sup>1)</sup>。SPaDe (スペード) とは、Space group and Packing Density predictor の略称であり、分子構造を入力として、結晶の空間群と分子の充填密度をそれぞれ独立に予測する2つの機械学習モデルから構成されている (図1)。これらのモデルは、ケンブリッジ結晶構造データベースを基に学習されており、「どのような分子が、どのような空間群や密度の結晶を形成しやす

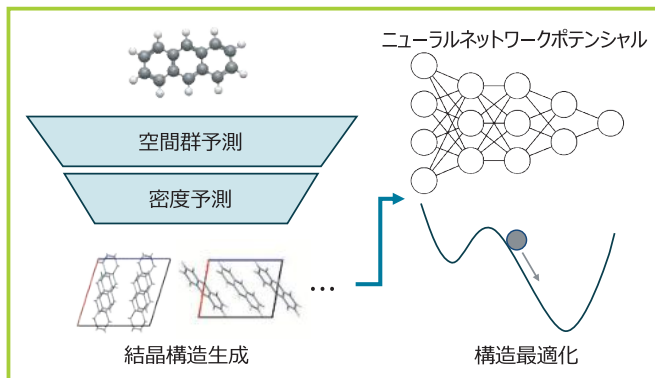


図1 SPaDe-CSPの計算ワークフロー

たにぐち・たくや

早稲田大学データ科学センター 准教授(任期付)  
【経歴】2019年早稲田大学先進理工学研究科先進理工学専攻一貫制博士課程修了、博士(工学)。同年同大学データ科学センター講師(任期付)、21年より現職。【専門】マテリアルズサイエンス、結晶化学、データ科学。【趣味】ウイスキー、麻雀。

E-mail: takuya.taniguchi@aoni.waseda.jp



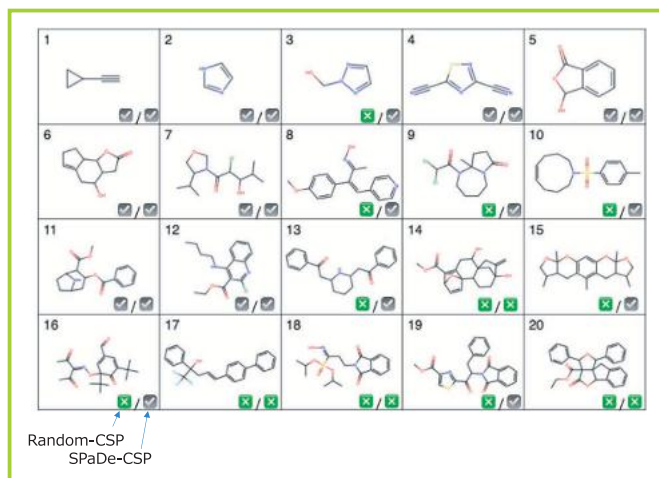


図2 20種類の有機分子の結晶構造予測と計算手法の違いによる成否

いか」を予測する。機械学習によって現実的な空間群と密度を事前に推定し、低密度で不安定な構造候補を初期段階で排除することで探索範囲を削減する。

SPaDe-CSPの有効性を検証するため、構造の複雑さが異なる20種類の有機分子を対象として結晶構造予測を行った。機械学習によって絞り込まれた条件下で構造探索を行い、得られた構造候補に対しては、ニューラルネットワークポテンシャルを用いて構造最適化およびエネルギー評価を実施した。その結果、20種類中16種類(80%)において、実験的に観測されている結晶構造を正しく再現することに成功した(図2)。この成功率は、空間群や密度に関する事前情報を用いずに探索を行った場合(Random-CSP)と比べて約2倍の成功率である。本結果は、機械学習による探索空間の事前設計が、CSPの成功率を高めるだけでなく、計算効率の向上にも大きく寄与することを示している。具体的には、空間群予測により探索対象の空間群候補を32種類から平均7~8種類に絞り込み、さらに密度予測によって実験値に近い格子パラメータを優先的にサンプリングすることで、構造緩和にかかる計算時間を約30%短縮できた<sup>1)</sup>。

この研究で提案したアプローチは、機械学習で結晶構造の探索範囲を削減することで、CSPをより実用的な材料設計ツールへと近付けるものである。一方で、分子量が大きく内部自由度を多く持つ柔軟な分子では、結晶化の際に多様な分子配座を取り得るため、探索空

間再び拡大するという課題が残されている。また、本研究では絶対零度(0 K)でのエネルギー安定性に基づいて構造を評価しているが、実際の材料開発では有限温度での安定性や構造相転移の有無が重要となる。

さらに、近年では有機結晶を対象とした生成AIの研究も進展している。敵対的生成ネットワーク(GAN)や拡散モデルを用いて、有機分子を結晶格子中に配置し、結晶構造そのものを生成する手法が提案されており、従来のCSPにおける構造探索とは異なるアプローチとして注目されている<sup>2,3)</sup>。これらの結晶生成AIが、新しい有機固体材料の発見にどの程度有効であるかについては現時点では十分に検証されていないものの、無機結晶分野にやや遅れながらも、有機結晶における生成AI研究は今後さらに発展していくと考えられる(筆者も貢献していきたい)。

### ACT-Xで得られたもの

最後に、ACT-Xを通して得られたものについて述べる。まず、「トランススケール」を共通のキーワードとして、様々な研究分野に取り組む研究者とのつながりを構築できた点が重要である。領域に参画する若手研究者とのネットワークに加え、アドバイザーを務める先生方との交流を通じて、多角的な視点から議論を行う機会が得られた。実際、領域会議でのディスカッションを通して共同研究にまで発展し、論文発表することができた<sup>4)</sup>。この研究者ネットワークは、今後の研究活動においても重要な基盤となると考えている。また、ライフイベント支援の充実も、本研究を進める上で大きな助けとなった。筆者は2人の子供を育児しているが、JSTのライフイベント支援制度を活用して研究補助者を雇用できたことにより、研究を効率的に推進することが可能となった。ACT-X研究者の多くが同年代であることから、研究とライフイベント、特に子育てとの両立に関する情報や経験を共有できた点も重要であり、研究環境の持続的な構築に寄与したと考えている。

1) T. Taniguchi, R. Fukasawa, *Digit. Discov.* **2025**, *4*, 3270.

2) Z. Ye et al., *The Innovation* **2024**, *5*, 2.

3) E. Jin et al., *arXiv* **2025**, 2512.06987.

4) C. Sato et al., *J. Am. Chem. Soc.* **2024**, *146*, 22699.