

量子化学計算と機械学習を用いた化学反応予測システムの開発

○藤波 美起登¹, 清野 淳司², 中井 浩巳^{1,4}¹早大先進理工,²早大理工研,³JST-CREST,⁴京大 ESICB

m-fujinami@fuji.waseda.jp

緒言: 化学において反応物から得られる生成物を正確に予測することは重要な課題のひとつである。これまで、情報化学の分野を中心にコンピュータにより化学反応を予測するシステムが多数開発されてきた。特に近年、化合物のグラフ情報を中心的な記述子とした機械学習により反応を予測する ReactionPredictor^[1]が開発され、その有効性が示された。しかし、予測精度の課題からこれらのシステムが商業的に普及するには至っていない。本研究では予測精度の向上に向けて、非経験的な量子化学計算から得られる様々な情報を記述子とした機械学習により反応を予測する、新たな反応予測システムの基盤を構築した。

理論と実装: 本手法では量子化学計算の結果を記述子として用いるために、局在化した軌道の情報を記述する自然混成軌道 (natural hybrid orbital; NHO) を算出する。そして化学反応を「電子を供与する NHO から電子を受容する NHO への電子移動」と定義し、2段階の処理により反応予測を行う。第一段階では NHO と原子に関する記述子を用いて、反応しうる NHO のスクリーニングを行う。第二段階では反応しうると判断された NHO の組み合わせから、ランキング形式で相互作用しやすい NHO の組を予測する。反応データベースには有機化学教科書^[2]内の 1110 の 2 電子移動反応を用いた。すべての記述子は B3LYP の 6-31++G** (H-Ar)、SDD (K-) レベルの計算により得た。機械学習手法にはニューラルネットワークを用いた。Figure 1 に本手法の概要を示す。

結果と考察: Table 1 に本手法 (Present) と ReactionPredictor による反応部位スクリーニングの精度を示す。Table 1 より、Present は ReactionPredictor と比べて少ない記述子数で同等の精度のスクリーニングを実現した。Table 2 に Present と ReactionPredictor による軌道相互作用の予測精度を示す。Table 2 より Present は 84.6%の精度で正しい軌道相互作用を予測した。この予測精度は反応データベース中のデータ数が増加することでさらに向上することが他の数値検証結果から示唆された。発表当日は学習する反応データ数と予測精度の関係および本手法によるラジカル反応などの予測結果も示し、本手法の性質、汎用性を議論する予定である。

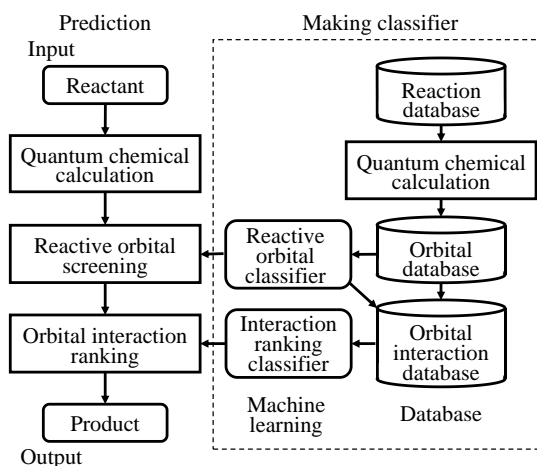


Figure 1. Flowchart of the present scheme

Table 1. Results of reactive site screening

	Number of reaction data	Number of descriptors	True negative rate ^{*1}		False negative rate ^{*2}	
			donor %	acceptor %	donor %	acceptor %
ReactionPredictor	5551	1500	84.9	68.8	1.9	1.3
Present	1110	24	82.0	75.3	1.0	0.7

*1 反応しない部位への正しい予測の割合 *2 反応する部位への誤った予測の割合

Table 2. Results of orbital interaction ranking

	Number of reaction data	Number of descriptors	Accuracy of interaction prediction ^{*3} %
ReactionPredictor	5551	1500	98.5
Present	1110	48	84.6

*3 正しい軌道相互作用が上位 5 位以内に予測される割合

[1] M. A. Kayala and P. Baldi, *J. Chem. Inf. Model.*, **52**, 2526 (2012).

[2] M. Jones Jr. 著, 奈良坂紘一他監訳, ジョーンズ有機化学, 第 3 版, 東京化学同人 (2006).