

固体混合物画像を利用した機械学習によるエナンチオマー割合診断

(北大工¹・北大院工²・北大 WPI-ICReDD³・理研 AIP⁴) ○佐野 太一¹・白倉 逸人²・井手 雄紀³・Sheng Hu³・瀧川 一学^{3,4}・猪熊 泰英^{2,3}

Diagnosis of Enantiomer Ratio in Solid Mixture Image using Machine Learning (¹Fac. Eng., Hokkaido Univ., ²Grad. Sch. Eng., Hokkaido Univ., ³WPI-ICReDD, Hokkaido Univ., ⁴AIP, RIKEN) ○Taichi Sano,¹ Hayato Shirakura,² Yuki Ide,³ Sheng Hu,³ Ichigaku Takigawa,^{3,4} Yasuhide Inokuma^{2,3}

Here, we developed an image-based machine learning driven enantiomer ratio (*er*) prediction system of solid chiral compound mixtures. In our prototype system optimized for D- and L-tartaric acids, pictures of solid mixtures taken in a fixed format using 200 mg of samples were required as training data for model making. In this research, we decreased sample amount to *ca.* 1 mg and achieved prediction of *er* of (*R*)- and (*S*)-binaphthol (BINOL) mixtures with a mean averaged error (MAE) of 0.061 and R^2 of 0.94.

Keywords : Chiral Compound; Enantiomer Ratio; Axial Chirality; Binaphthol; Machine Learning

キラル化合物のエナンチオマー比率を分析することは、有機合成反応や医薬品の開発研究において重要なプロセスである。前発表者の研究において我々は、機械学習により固体混合物の画像から混合割合を予測するシステムの構築を行った。本研究では、このシステムをキラル化合物のエナンチオマー比率予測に応用した。ショ糖と食塩の混合物に対して最適化した機械学習システムを D-,L-酒石酸の混合物に適用すると、平均誤差(MAE) 0.064 でエナンチオマー割合の予測ができた。しかし、この手法では適切な機械学習モデルの構築に、1枚あたり合計 200 mg もの光学的に純粋なサンプルを使って撮影した画像が 300 枚以上必要であるという欠点があった。そこで、用いる混合物((*R*)-, (*S*)-BINOL)の総量を 1 mg に減量した上で、そのジクロロメタン溶液を乾固させた状態で撮影した画像を訓練データとする機械学習を検討した (Fig. 1a)。

訓練データの画像が 210 枚の段階ではエナンチオマー割合の予測精度が MAE = 0.11, $R^2 = 0.77$ であったが、それを 1,000 枚まで増やすことで MAE = 0.061, $R^2 = 0.94$ まで向上することが分かった (Fig. 1b)。この結果を同程度の精度を出したプロトタイプの機械学習(MAE = 0.057, $R^2 = 0.92$)と比べると、訓練データは3倍以上必要になるものの、画像撮影のフォーマットを揃える必要がなく、使用するサンプル量も 200 分の 1 となったことから、機械学習モデルの構築に要する時間が大幅に短縮された。さらに、この予測システムはラセミ化反応における反応溶液を乾固させてすぐに分析可能であることから、迅速なエナンチオマー比率の分析に応用可能であると考えられる。

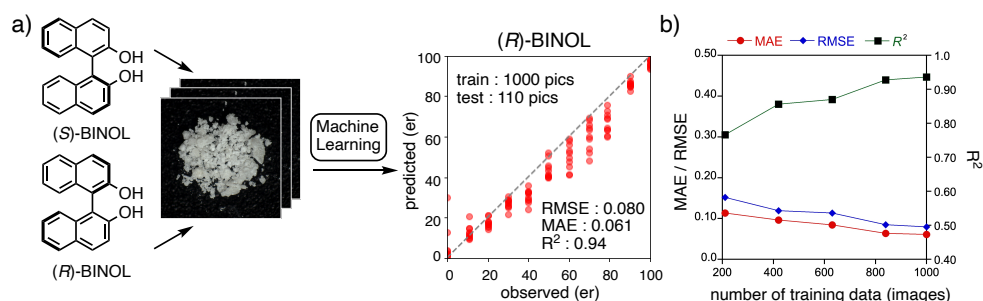


Fig. 1. a) BINOL のエナンチオマー割合診断 b) 訓練画像枚数に基づいた予測精度変化