使用分子动力学模拟和卷积神经网络快速预测液相酸催化反应速率

Alex K.Chew, SHengli Jiang, Weiqi Zhang, Victor M. Zavala, Reid C. Van Lehn

摘要：用于将生物质制造为高价值化学品的液相酸催化反应的反应速率对溶剂的成分非常敏感。而确定合适的混合溶剂在理论和实验上都是较为困难的。我们研究表明，三维卷积神经网络（3D-CNN）可以利用经典的分子动力学模拟数据生成的反应物-溶剂环境复杂构型，来准确预测生物质的Brønsted酸催化反应速率。我们开发了一种三维卷积神经网络（称为SolventNet），使用七种生物质衍生氧化物在水-共溶剂体系内的反应速率实验值和相关的分子动力学模拟数据来训练模型，以达到预测酸催化反应速率的目的。研究表明，SolventNet可以预测更多的反应物和溶剂系统的反应速率，比之前的模拟方法快一个数量级。这种将机器学习和分子动力学模拟结合的方法能快速、高通量地筛选溶剂体系，并且确定改进的生物质转化条件。

1. 引言

木质纤维素生物质得催化转化过程，是从可再生原料中获取交通运输所需燃料、高价值化学品等产品的一种很有前景的方法。生物质衍生分子转化通常由液相、酸催化的反应来促进（如图1.a例），而这些反应经常因为在水中受阻，故需要酸催化。