An Introduction to Cheminformatics

Hugh Purdy

February 07, 2020

Abstract

Cheminformatics is generally defined as the application of computational and informatics approaches to address chemical problems. The field is often understood to be especially concerned with the representation and manipulation of chemical structural information and other molecular descriptors. Efficient methods for storing and searching such information enable approaches for the large-scale screening of chemical databases, the development of structure-based activity/property models, and other problems concerning chemical "information".

In this talk, I will give a high-level introduction to some of the central concepts from cheminformatics, including chemical identifiers, structure representations, structure/substructure searching, "similarity" searching, and using structural information as the input for quantitative structure-activity relationship (QSAR) models. Emphasis will be placed on 2D chemical information, though considerations for studying 3D information may be discussed.