[L13] Machine learning applied to natural products: lessons from nature inspiring the design of new drugs.

Peter Ertl

Novartis Institutes for BioMedical Research, Basel, Switzerland

Natural products (NPs), i.e. molecules produced by living organisms, have evolved over a billion years long natural selection process to form optimal interactions with biological macromolecules. NPs are therefore a very useful source of inspiration for the design of new drugs. Application of machine learning techniques can help us to identify where this “bioactivity propensity” of NP molecules is encoded. In this presentation a cheminformatics analysis of structures of about 200,000 NPs is reported. First the physicochemical properties of NPs are compared with those of synthetic molecules. Then focus is given to various structural features of NPs (scaffolds, substituents, functional groups) to identify those that are typical for NPs and discriminate this class of molecules from synthetic molecules. A method to calculate the natural-product likeness developed at Novartis is also described. And finally application of the results to support various drug discovery techniques used in the pharmaceutical industry including scaffold hopping, design of NP-like libraries and fragment-based screening is described.