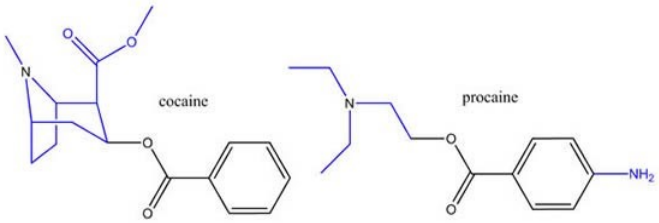
5.2: Similar-Structure, Similar-Property Principle

The Similar-Structure, Similar Property Principle is the fundamental assertion that similar molecules will also tend to exhibit similar properties. These properties can either be physical (e.g. boiling points) or biological (e.g. activity).

Example 1: Hexane and heptane should have similar boiling points and water solubility.

Example 2: Cocaine and procaine are both local anesthetics



Quantitative Structure-Property Relationships (QSPR) and Quantitative Structure-Activity Relationships (QSAR) use statistical models to relate a set of predictor values to a response variable. Molecules are described using a set of descriptors, and then mathematical relationships can be developed to explain observed properties. In QSPR and QSAR physico-chemical properties of theoretical descriptors of chemicals are used to predict either a physical property or a biological outcome.

In either case, a set of known molecules is used to create a **training set** that a statistical model can be derived from. These molecules have known properties or activities. An outside **test set** is used to validate the model. The test set consists of other molecules with known properties that are excluded from the training set. After the model is validated, it can be used to predict properties or activities of molecules that are outside the previous sets. One caveat- new test molecules cannot be sufficiently different from the ones used in previous sets.