The rapid design-test-redesign cycles of modern drug discovery and the demand for fast model (re)building whenever data becomes available have given rise to a trend to develop computational algorithms for automatic model generation. Automatic modelling processes allow computational scientists to explore large numbers of modelling approaches very efficiently and make QSAR/QSPR model building accessible to non-experts.

This poster was displayed at MedChem ADMET Eurpoe, 2008.

In this poster we will present an automatic model generation process for building QSAR models. The stages of the process that ensure models are built and validated within a rigorous framework are:

- Splitting data into training, validation and test sets (by cluster analysis)
- Descriptor calculation and filtering (2D SMARTS descriptors, whole molecular properties and user's imported descriptors)
- Application of modelling techniques (PLS, Radial Basis Functions with genetic algorithm, Gaussian Processes (GP))
- Selection of the best model (performance on the validation set is used as criterion) and evaluating it on the test set

This algorithm is implemented in the StarDrop environment for decision support within drug discovery and is referred to as the Auto-Modeler.

This poster is also available as a PDF file.