# Chemometric modelling of thermal properties

**Michal Dabros** (michal.dabros@hefr.ch)

Institute of Chemical Technologies, HES-SO – Haute école d'ingénierie et d'architecture Fribourg

Bd de Pérolles 80, CH-1700 Fribourg, Switzerland

**Anikó Udvarhelyi** (aniko.udvarhelyi@novartis.com)

PHAD GDD / Technical R&D GDD

Novartis Campus, CH-4056 Basel, Switzerland

## Abstract

Safety assessment studies constitute a crucial step in chemical process development. Reaction calorimetry, a technique of measuring the energy released or absorbed during chemical reactions, is an important tool for process safety evaluation and reaction screening. However, in certain applications, for instance in the thermal exploration of potentially highly exothermic compounds, it would be beneficial to predict the thermal behaviour as a function of the compounds' structure.

In this work, we present a novel approach of modeling the thermal behaviour of cyclic sulfamidates reacting with strong bases. The reaction enthalpy of a collection of five-membered cyclic sulfamidates was measured using a custom-made ml-scale reaction calorimeter. In order to link the measured reaction enthalpy to the compound structures, we sought to build a mathematical model using structural descriptors. We found that it was necessary to use descriptors derived from quantum-chemical calculations to ensure that relevant electronic effects were captured by the descriptors. Tailored to the problem at hand, we calculated Fukui indices on certain atoms, HOMO-LUMOs, reaction free energy estimates and bond dissociation energies with density functional theory (DFT) and semi-empirical DFT methods. The selected descriptors were expected to be linked to the reactivity and thermal behaviour of the sulfamidate ring.

A Partial Least Squares (PLS) model was then built, linking each compound's descriptors to its value of reaction enthalpy ($\Delta T_r$). The model was evaluated using the leave-one-out method to obtain a standard error of cross-validation (SEV). Next, the PLS model was optimized using Genetic Algorithms (GA), an optimization routine programmed to minimize the cross-validation error by selecting an optimal set of descriptors to be used by the model. With the optimized set of descriptors, the value of SEV was reduced to less than half of its original value, making the model more reliable for predicting the thermal behaviour of other compounds in the sulfamidate family.