Precise Molecular-Feature Analysis (PMA) of UHPLC/Q-TOF MS for Metabolite Profiling in Synthetic Biology

The major goal of synthetic biology is to generate desired valuable substances with a good conversion from substrates to products by optimization of genetic and regulatory processes and pathways within cells. The secondary metabolites of microorganisms are biosynthesized in genetically encoded pathways, which involve multiple genes. These biosynthetic genes are usually clustered with regulatory and resistance genes in microbe genomes. Under laboratory cultivation conditions, most of these gene clusters are usually not expressed. Therefore, the potentially valuable “silent” pathways represent an unexploited reservoir of new secondary metabolites in microbial genome mining approaches. We use three strategies to activate these pathways, including extended cultivation, native host engineering and heterologous host expression.

To meet the challenges of metabolite profiling in synthetic biology, we have developed Precise Molecular-feature Analysis (PMA) of UHPLC/Q-TOF MS. High resolution MS, isotope pattern, ion charges, molecular-feature algorithm and retention time are used together in PMA to get cleaned mass spectra for precise identification and to integrate MS peaks for comparative quantification. Process automation with large UHPLC/Q-TOF MS data sets is realized by scripting with VBS.

In non-targeted analysis of engineered native hosts in various media, precise molecular-feature MS libraries with all metabolites of wild-type samples cultivated in same conditions are first established. In the comparative metabolite profiling, new compounds from engineered native hosts, which might be biosynthesized by activated unknown gene clusters, can be precisely identified by efficient comparison with established reference MS libraries in a fully automatic way.

For the known biosynthetic pathways in heterologous host expression, precise molecular-feature MS libraries with targeted metabolites are simulated based on adduct type, high resolution MS, isotope pattern, ion charge state and peak width (FWHM). In following PMA of heterologous expression samples, these metabolites can be quickly found by direct comparison with simulated MS libraries in targeted analysis.

In the PMA process, MS molecular-feature algorithm is used as second dimension in metabolite separation, in addition to UHPLC. This two dimension strategy has significantly increased the separation performance for co-elutes in chromatogram and decreased the background noise influence in MS spectroscopy. Besides comparative metabolite profiling, PMA-guided quantification and preparative separation has also be implemented for reproducibility monitoring in the re-fermentation and extraction process, as well as for fractionation and purification in preparative scale to get targeted metabolites or new compounds in the natural product drug discovery driven by synthetic biology.