EDA-EMERGE Specialized Course 6 & Local Course 1:

"Advanced course and practical exercise on Chemometrics, LC-LTQ Orbitrap MSn and computer tools for chemical structure generation and elucidation" &

"Introduction to R and multivariate statistic"

Venue: UFZ, Leipzig, Germany

Organizer: Dr. Werner Brack

Date: 23.10 to 25.10.2013

Time: 9 h00 – 18h30, 9 h00 – 18h30, 9h00 – 11h30

Course description

The SC6 was held together with LC1 in a three day EDA-EMERGE training course block on Chemometrics, LC-LTQ Orbitrap MSn and computer tools for chemical structure generation and elucidation (SC6) and on using R for multivariate statistic (LC1). The course focused on teaching, desktop analytical tools, concepts for effect directed analysis and chemical structure elucidation in water quality monitoring. This included a 3h introduction to R software and statistics (Dr. Steffen Neumann, Leibniz Institute Halle), 6h of lectures and tutorials in chemometrics (Prof. Jan Christensen, University of Copenhagen) and 15h of lectures, coursework and tutorials on chemical structure elucidation (Dr. Martin Krauss, UFZ Leipzig and Dr. Emma Schymanski, EAWAG Switzerland). At least 6h of self-study was necessary to revise the literature material provided by the course instructors, in addition to preparing for the tutorials.

The training was practice orientated and intensive, involving lectures and tutorials by international experts using state of the art computer based tools, real field data and LC-MS spectra of chemical pollutants. The lectures were followed by tutorials where students worked independently and then in groups using software packages such as Latentix and R. Furthermore, the participants were introduced to the basics of multivariate statistics, Latentix, R and how to apply these tools in chemical structure elucidation of complex LC data.

This amounted to a minimum total academic involvement of 30 hours (1.0 ECTS) for the participants.

AGENDA

Tuesday, 23.10.2012		
Time	**Title**	**Lecturer**
9:00-11:00	Introduction to R, part I	Steffen Neumann
11:00-11:20	Coffee Break	
11:20-13:00	Introduction to R, part II	Steffen Neumann
13:00-14:00	Lunch Break	
14:00-15:40	Multivariate Statistics, part I	Jan Christensen
15:40-16:00	Coffee Break	
16:00-18:30	Multivariate Statistics, part I	Jan Christensen

Wednesday, 24.10.2012 – Advanced LC-MS course		
Time	**Title**	**Lecturer**
9:00-9:30	Introduction to the workshop (recapitulation of previous content in summer workshop)	Martin Krauss
9:30-10:45	Molecular formula calculation (7 golden rules, adducts, isotopes, MS/MS; SIRIUS, MOLGEN-MS/MS, MZmine, online program) & Practical exercises	Martin Krauss
10:45-11:00	Coffee Break	
11:00-12:15	*In silico* fragmentation (MetFrag, MassFrontier); Canidate retrieval and fragments; exact mass versus formula search & Practical exercises	Emma Schymanski
12:15-13:15	Lunch Break	
13:15-14:15	Retention and behavior prediction (log KOW, CHI, correlations, retention index; compare models versus candidate elimination) & Practical exercises	Martin Krauss
14:15-15:00	Toxicity behavior and prediction (structural alerts, bioclipse, openTox, chemprop) & Practical exercises	Emma Schymanski
15:00-15:15	Coffee Break	
15:15-15:45	What if`s – the hard cases (not in ChemSpider, databases, no matches; structure generation; TPs: special cases)	Martin Krauss, Emma Schymanski
15:45-18:30	Extended practical exercices and discussion	All

Thursday, 25.12.2012		
Time	**Title**	**Lecturer**
09:00-11:30	Individual tutorials in PC-cabinet	All
11:30	End of training event	

COURSE CONTENT

- Multivariate Fingerprinting
    - Analytical data in environmental sciences
    - Data structure
    - Multivariate data
    - PCA visualization
- Contaminant Fingerprinting
    - Pattern recognition
    - What is a fingerprint
    - Tiered approaches for oil fingerprinting
    - Building a database
    - Retention time alignment
    - Normalization
    - Chemometrix
- Source identification of polycyclic aromatic hydrocarbons (PAHs) from multiple sources
    - Chemical analysis
    - Data pre-processing
    - Data structure
    - Chemical interpretation
- Molecular formula generation
    - High resolution mass spectrometry
    - Molecular formula from accurate mass
    - Heuristic filtering of molecular formulas
- Overview of compound and spectral databases
- Compound database and mass spectral database searching
- Candidate selection via *in silico* fragmentation
    - Basics
    - MetFrag
    - MetFusion
- Advanced analytical strategies
    - LC-HR-MS structure elucidation strategy
    - Using ionization behavior or deuterium exchange
- Molgen MS/MS
- Where to start identification?

    - Screening
    - Targets, suspects, non-targets
    - Peak picking
    - Making the “big list of compounds smaller
    - Related suspect screening
    - Detecting isotopes and adducts)
- Structure elucidation beyond compound databases – the trickier cases
    - Size of databases versus compound spaces
    - Structure generation and unknown spectra
    - Filtering strategies for structure elucidation
    - Examples