Semi-quantitative LC/ESI/MS analysis using predictive models of ESI ionization efficiencies

Jaanus Liigand[a], Piia Liigand [a], Mari Ojakivi [a], Karl Kaupmees [a], Anneli Kruve [a,b]

[a] University of Tartu, Institute of Chemistry, Ravila 14a, 50411 Tartu, Estonia, kruvelab.com, jaanus.liigand@ut.ee

[b] Free University of Berlin, Institute of Chemistry and Biochemistry, Takustr. 3, 14195 Berlin, Gemany

Overview

Purpose: Enabling standard substance free semi-quantitation in LC/ESI/MS via ionization efficiency (IE) predictions.

Methods: IEs of 400 compounds were measured in flow injection mode in ESI positive and negative mode in 30 different mobile phase compositions. Different machine learning approaches were used to develop models to predict IEs.

Results: Regularized random forest regression models for ESI negative and ESI positive mode were developed. Concentration misprediction for ESI+ <5- and for ESI- <2.7-fold.

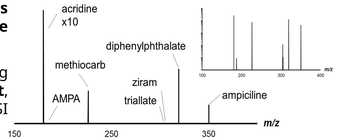
Introduction

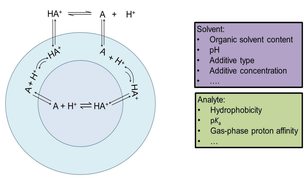
Non-target and suspect screening

- Thousands of compounds discovered and screened
- Mostly qualitative information available
- Quantitative information needed
- ESI ionization efficiency is compound and solvent dependent
- Standard substances are often not available
- Concentrations are very low; it is often impossible to separate and quantify pure compounds.
- Signals are affected by matrix effect;

Electrospray Ionization Efficiency (IE)

- shows **how efficiently** analyte **ions from liquid phase** form **gas phase ions**
- allows to study **processes** occuring along **ESI plume** and in **droplet**, giving more insight into ESI **mechanism**
- depends on **physico-chemical properties** of **analyte** and **eluent**
- relating analyte and eluent physico-chemical parameters to *IE* gives the power to predict *IE*
- ability to **predict IE** allows to carry out **standard substance free quantification** with any needed analyte in any preferable medium





Solvent:
- Organic solvent content
- pH
- Additive type
- Additive concentration
- ....

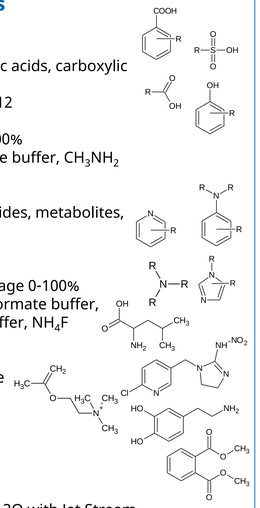
Analyte:
- Hydrophobicity
- pK_a
- Gas-phase proton affinity
- ...

Methods

Studied Compounds and Eluents

Test set

- Negative mode
  - 62 substituted phenols and benzoic acids, carboxylic and sulfonic acids
    - log*P* = -3.3 ... 6.3; pK_a = -4.5 ... 12
  - 10 eluent compositions
    - Acetonitrile percentage 20 – 100%
    - Additives: NH_3, HCOOH, acetate buffer, CH_3NH_2
    - pH = 2.7 ... 10.7
- Positive mode
  - 333 small molecules (drugs, pesticides, metabolites, *N*-bases, *O*-bases)
    - log*P* = -4.1 ... 7.7
  - 21 eluent compositions
    - Acetonitrile/methanol percentage 0-100%
    - Additives: HCOOH, TFA, NH_3, formate buffer, acetate buffer, bicarbonate buffer, NH_4F
    - pH = 2.0 ... 10.7
- Flow injection flow rate 0.2 mL/min
  - Agilent XCT iontrap with ESI source
  - $\log IE_1 = \log \frac{slope_{compound1}}{slope_{compound2}}$
- **COSMO-RS** and **PaDEL** descriptors



Validation set

- Negative mode
  - Agilent 1290 UHPLC + Agilent 6495 3Q with Jet Stream thermal focusing ESI source
  - 16 compounds
  - flow injection
    - 65/35 5 mM acetate buffer pH = 6.2
    - 30/70 5 mM acetate buffer pH = 3.6
  - Chromatographic separation with gradient elution
- Positive mode
  - Thermo Fisher Scientific LTQ IonTrap with ESI source
  - 18 compounds
  - flow injection
    - 10/90 10 mM formic acid/ acetonitrile

Results

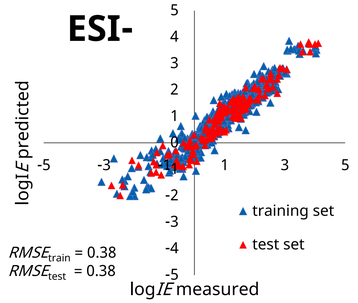


Figure 1: Fit obtained with random forest regression for the log*IE* values over all tested mobile phases in negative mode. Number of data points is 697

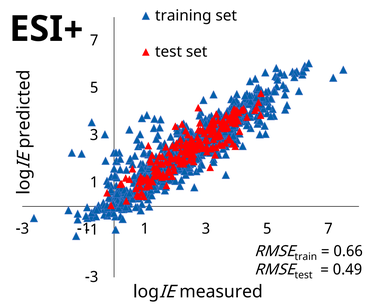


Figure 2: Fit obtained with random forest regression for the log*IE* values over all tested mobile phases in positive mode. Number of data points is 1166

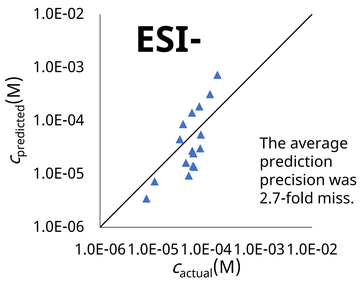


Figure 3: Correlation between the predicted concentrations and measured concentrations for the 16 acids in LC mode on Agilent 6495 3Q in ESI negative.

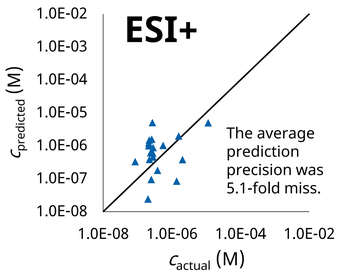


Figure 4: Correlation between the predicted concentrations and measured concentrations for the 18 drugs in flow injection mode on Iontrap (Thermo Fisher Scientific) in ESI positive. Eluent 90/10 acetonitrile/10 mM HCOOH.

How to implement this approach

Calibration compounds (6)

SMILES codes of compounds of interest

Concentration of calibration compounds

MS[1] signals (area or height) of compounds of interest and calibration compounds

Solvent and LC conditions:
- Gradient program
- Organic modifier %
- Additive type and *c*

Concentration prediction

How to choose?
- Ideally (6):
  - Otherwise (min 2)
- Covering a wide range of chemical space
  - Low, high, medium log*IE*
  - (log*P* and pK_a)
- **More calibration compounds = higher prediction accuracy and more reliable results**

Concentrations?
- In (linear) dynamic range
- Ideally 5 levels
  - If MS signal is stable enough also doable on 1 concentration level

MS conditions?
- MS[1]
- Use same settings you want to quantify with

Solvent?
- Use same solvent/ matrix/ gradient you want to quantify with

References

- Kruve and Kaupmees *Anal Chem* 89(9), 2017
- Liigand et al *Anal Chem* 89(11), 2017
- Liigand et al *J Am Soc Mass Spectrom* 28(3), 2017
- Kruve et al *J Am Soc Mass Spectrom* 28(5), 2017
- Liigand et al *J Am Soc Mass Spectrom* 26(11), 2015
- Liigand et al *J Am Soc Mass Spectrom* 25(11), 2014

Acknowledgement

This work has been supported by PUT 34 from Estonian Research Council, by the institutional funding IUT20-14 (TLOKT140141) from the Ministry of Education and Research of Estonia and by smart specialization doctoral stipend. The authors thank Graduate School of Functional Materials and Technologies for receiving funding from the European Regional Development Fund in University of Tartu, Estonia

European Union
European Regional Development Fund

Investing in your future