PIIA LIIGAND

Expanding and improving
methodology and applications
of ionization efficiency
measurements

DISSERTATIONES CHIMICAE UNIVERSITATIS TARTUENSIS

182

PIIA LIIGAND

Expanding and improving methodology and applications of ionization efficiency measurements

Institute of Chemistry, Faculty of Science and Technology, University of Tartu, Estonia

The dissertation is accepted for the commencement of the degree of Doctor of Philosophy in Chemistry on 11th June, 2019 by the Council of Institute of Chemistry, University of Tartu.

Supervisors: Assoc. Prof. Anneli Kruve-Viil, Institute of Chemistry, University of Tartu, Estonia; Department of Environmental Science and Analytical Chemistry, Stockholm University, Sweden

Dr. Karl Kaupmees, Institute of Chemistry, University of Tartu, Estonia

Opponent: Prof. Dr. Susan D. Richardson, Arthur Sease Williams Professor of Chemistry, Department of Chemistry and Biochemistry, University of South Carolina, USA

Commencement: August 15th, 2019, 10:15

European Union
European Regional
Development Fund

Investing
in your future

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

I.   Liigand, P.; Kaupmees, K.; Kruve, A. Ionization Efficiency of Doubly Charged Ions Formed from Polyprotic Acids in Electrospray Negative Mode. *J. Am. Soc. Mass Specrtom.* 2016, *27*, 1211–1218.

II. Liigand, P.; Kaupmees, K. Haav, K. Liigand, J.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. Think Negative: Finding the Best Electrospray Ionization/MS Mode for Your Analyte. *Anal. Chem.* 2017, *89*, 5665–5668.

III. Liigand, P.; Heering (Suu), A.; Kaupmees, K.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. The evolution of electrospray generated droplets is not affected by ionization mode. *J. Am. Soc. Mass Specrtom.* 2017, *28*, 2124–2131.

IV. Liigand, P.; Liigand, J.; Cuyckens, F.; Vreeken, R.J.; Kruve, A. Ionisation efficiencies can be predicted in complicated biological matrices: A proof of concept. *Anal. Chim. Acta* 2018, *1032*, 68–74.

V.   Liigand, P.; Kaupmees, K.; Kruve, A. Influence of the amino acid composition on the ionization efficiencies of small peptides. *J. Mass Spectrom.*, 2019, doi: 10.1002/jms.4348

VI. Liigand, P.; Liigand, J.; Kaupmees, K., Kruve, A. Revealing the hidden treasure: Making ionization efficiency values from literature directly comparable, submitted for publication in *Anal. Chem.*

Author's contribution

Paper I: Main person responsible for planning and writing the manuscript. Performed all the experimental work.

Paper II: Main person responsible for planning and writing the manuscript. Performed all the experimental work.

Paper III: Main person responsible for planning and writing the manuscript. Performed all the experimental work, except pH_{abs} and droplet size measurements.

Paper IV: Main person responsible for planning and writing the manuscript. Performed all the experimental work.

Paper V: Main person responsible for planning and writing the manuscript. Performed all the experimental work.

Paper VI: Main person responsible for planning and writing the manuscript. Performed all the experimental work.

ABBREVIATIONS

α	Degree of dissociation/ionization
BP	Becke-Perdew (density)functional
CEM	Chain emission model
COSMO-RS	Conductor-like Screening Model for Real Solvents
CRM	Charge residue model
CSF	Cerebrospinal fluid
DFT	Density functional theory
DIA	Direct infusion analysis
EDTA	Ethylenediaminetetraacetic acid
ESI / ESI+ / ESI-	Electrospray ionization /Positive mode / Negative mode
FIA	Flow-injection analysis
GB	Gas-phase basicity
HMDB	Human metabolome database
HPLC	High-performance liquid chromatography
IE	Ionization efficiency; the abbreviation is used for quantitative value
IEM	Ion evaporation model
LC	Liquid chromatography
LDA	Linear discriminant analysis
$\log IE$	Logarithm of ionization efficiency
$\quad \log IE^{\text{ESI+}}$	measured in positive mode ESI
$\quad \log IE^{\text{ESI-}}$	measured in negative mode ESI
$\log P$	Logarithm of octanol-water partition coefficient
LOO	Leave-one-out (validation)
m/z	Mass-to-charge ratio
MeCN	Acetonitrile
MS	Mass spectrometry or mass spectrometer
MV	Molecular volume
NMR	Nuclear magnetic resonance
pH	Negative logarithm of hydrogen ions in a solution
$\quad \text{pH}_{\text{abs}}^{\text{H}_2\text{O}}$	pH expressed on the absolute scale but shifted by a constant in order to make values directly comparable to the conventional aqueous pH values
$\text{p}K_{\text{a}}$	Negative logarithm of acid dissociation constant
$\text{p}K_{\text{b}}$	Negative logarithm of base dissociation constant
RF	Response factor
$\log RRF$	Logarithm of relative response factor
$\quad \log RRF_{\text{comparable}}$	Logarithm of relative response factor transformed to a uniform quantitatively comparable value
RI approximation	Resolution-of-identity approximation
RIE	Relative ionization efficiency
$s_{\text{consistency}}$	Consistency standard deviation

s_{MAE}	Mean absolute error
s_{RMSE}	Root-mean-square error
SMILES	Simplified molecular-input line-entry system
SPADNS	2-(4-Sulfophenylazo)chromotropic acid trisodium salt
TA	*trans*-3(3-pyridyl)acrylic acid
TM	Target mass
TZVP	Triple zeta valence + polarization (base)
WAPS / WANS	Weighted average positive/negative sigma

INTRODUCTION

Electrospray ionization (ESI) is a method used to produce gas phase ions in which high voltage is applied to a liquid. ESI is commonly used to couple liquid chromatography (LC) and mass spectrometry (MS). LC/ESI/MS enables to determine trace amounts of compounds and is thus an important tool in various fields, e.g. environmental analysis, drug discovery, metabolomics, proteomics, and food safety.

Although ESI/MS is widespread, one major caveat of using the technique is the need for standard substances because compounds have different ionization efficiencies (more than six orders of magnitude). Ionization efficiency is the efficiency of conversion of molecules from solution to gas-phase ions. This means that using solely mass spectrometric signal to quantify the analytes can lead to misestimating the concentrations millions of times. However, suitable standard substances are not always available or obtaining them might not be financially feasible. One solution is using predictions of ionization efficiencies. Ionization efficiencies depend on the setup (instrument, geometry of the ESI source), the compound itself (structure and properties such as acidity, hydrophobicity, etc.), and its surrounding environment (eluent composition and pH, other compounds in the droplet).

Ionization efficiencies have been studied by several groups and many different parameters have been found to influence the ionization efficiency of a compound. Previous studies have mostly been focusing on small singly charged analytes. Additionally, the studies so far have in most cases focused on certain compound classes and are unfortunately not quantitatively comparable. The conditions can vary a great deal in these studies, such as ESI ionization mode (positive or negative), instrument (ESI source geometry, ion optics), surrounding matrix (neat solvent vs complex matrices, e.g. blood), used compound classes and their properties, eluent composition (content of water, pH, buffer components). Therefore, the conclusions reached are controversial as the ESI process itself is complex and is shown to depend on all the aforementioned conditions.

The most significant shortcomings that hinder the universal applicability of ionization efficiency prediction models are the inability to account for multiply charged compounds, different matrices, ESI modes, and the possibility to quantitatively compare the results of different researchers and setups.

Therefore, the main goal of the thesis is to expand and improve ionization efficiency measurement methodology and applications. To achieve this goal, firstly, ionization efficiency measurement methodology is expanded to compounds which can obtain multiple charges and to measurements in more complex media, namely biological matrices. Secondly, the methodology is improved so that measurements conducted in different ESI modes (positive and negative) would be quantitatively comparable. Finally, and most importantly, an approach is developed to transform ionization efficiency values from various

sources and conditions into a uniform dataset. As a result, all the data measured from all over the world can be transformed into a quantitatively comparable set, thus enabling to obtain more quantitative insight into ESI mechanism, opening new ways to model the ESI process and to predict the ionization efficiencies more accurately.

1. REVIEW OF LITERATURE

1.1 ESI/MS analysis

Mass spectrometry (MS) with electrospray ionization (ESI[1]) source is a key technique in various research fields, ranging from food and environmental analysis to metabolomics and proteomics.[2,3] Commonly, the ESI source is preceded by liquid chromatography (LC) to first separate the compounds in a sample. The ability to analyze samples almost directly with direct infusion[4] or flow injection experiments[5] coupled with MS has provided a tremendous increase in sample throughput. This technique has proven useful for the analyses of a wide range of samples from human blood plasma[6] to historic wines[7] to ecological samples.[8]

Despite widespread application, current understanding of the ESI process is still limited.[2,9] Ionization efficiencies for compounds vary in the ESI process, meaning that for two compounds of equimolar concentrations ESI/MS does not produce equal signals in the MS, on the contrary, the signals may be vastly different (more than 6 orders of magnitude). For quantitative analysis, standard substances are required due to the large differences in ionization efficiencies observed in ESI/MS.[10–13] Unfortunately, standard substances are often not available for metabolites and degradation products; therefore, knowing or predicting the ionization efficiency of these compounds would be extremely useful for estimating their concentrations and therefore their relative importance. Positive mode ESI (ESI+) is generally preferred as more compounds are expected to ionize in this mode.[2,14] However, the major advantage of negative ion mode (ESI-) is the lower background noise.[2,14]

1.2 Ionization in the ESI source

ESI is a technique used to produce gas phase ions from solution in mass spectrometry in which a high voltage is applied to the liquid to create ESI plume, a spray of small droplets (Figure 1). ESI can be used for the analysis of a wide range of compound classes, from small molecules to polymers, nucleic acids, and proteins.[14]
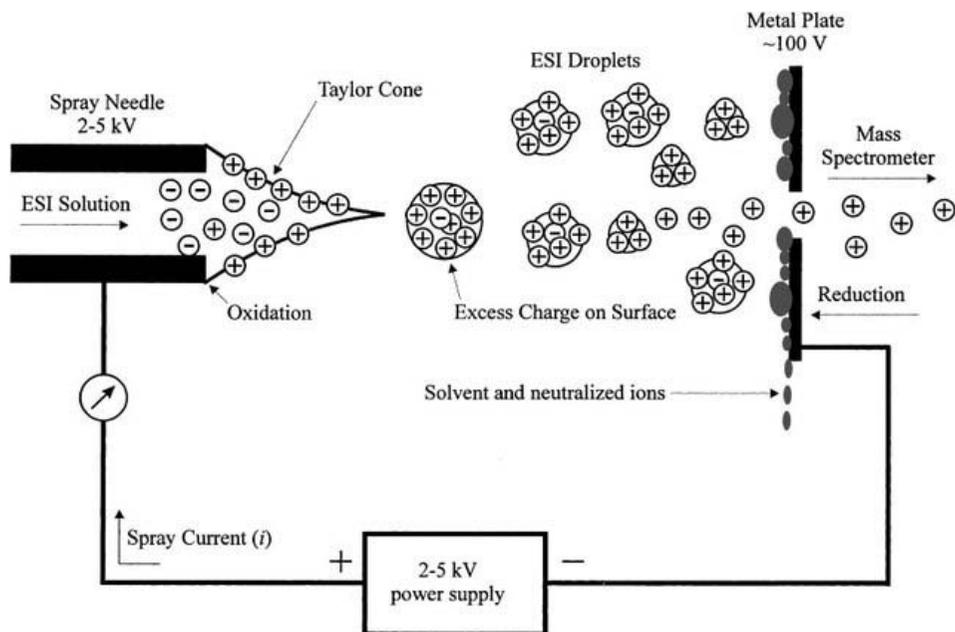
Figure 1. Schematic representation of electrospray ionization (ESI) process. Reprinted from ref [2].

1.2.1 Ionization mechanism

There are three main ion release mechanisms which have been introduced and are widely accepted:[2,15–17] ion evaporation model (IEM) for low molecular weight species,[18] charged residue model (CRM)[15,19] for large globular species and chain ejection model (CEM) for unfolded proteins and disordered polymers.[20,21] However, the borders of these mechanisms based on size, structure, and polarity of the analyte are unclear and it is highly probable that a molecule ionizes according to several models, where usually one is more dominant than the other.[22,23] Recently, Konermann et al.[23] showed that polypropylene glycol can ionize either via IEM or CRM depending on the size of the initial droplet. Interestingly, also the charge state of the gas phase ions depends on the ionization model.[23]

Ion evaporation model[18] applies to low molecular mass molecules and small inorganic ions.[17,24] This mechanism assumes that the increased charge density that results from solvent evaporation eventually causes Coulomb repulsion to overcome the liquid's surface tension, resulting in a release of ions from the droplet surface. The formed product is a small gas-phase cluster, consisting of the ion and a few solvent molecules which are lost by collisions with background gas molecules as the cluster moves towards mass spectrometer. Offspring droplets are formed from the outer layer of the parent droplet and thus the product droplets are enriched in species with the highest surface affinity.

Charged residue model (CRM), also known as Coulomb fission mechanism, assumes that the increased charge density due to solvent evaporation causes large droplets to divide into smaller and smaller droplets, which eventually only consist of a single ion. As the last solvent shell evaporates the charge of the droplet is transferred to the analyte. Ionization degree in the case of CRM does not depend on the charge of the analyte in the solution phase. The possible maximum amount of charges can be calculated from Rayleigh stability limit:[25]

$$z_R = \frac{8\pi}{e} \sqrt{\varepsilon_0 \gamma R^3} \tag{1}$$

where z_R is Rayleigh charge or the number of elementary charges at the surface of the droplets, R is the droplet radius, γ is the surface tension, ε_0 is the electrical permeability of vacuum and e is the elementary charge.

The most recently proposed ionization mechanism is chain ejection model (CEM). Long, unfolded proteins are hydrophobic in nature and are more likely to reside on the surface of the droplet. One terminus of the chain gets expelled from the droplet to the gas phase which is followed by a step-wise ejection of the rest of the chain and results in a separation of the chain from the droplet. This whole process is referred to as the chain ejection model. CEM has several similarities with IEM, e.g. the molecules with higher surface affinity are more likely to be ejected from the surface of the droplet to the gas phase and therefore have higher ionization efficiency.

1.2.2 Evolution of the ESI plume

In the interpretation of the ESI process, one important factor is the solvent composition, which is usually described in terms of initial composition since the actual composition in the plume is difficult to measure. However, it has been shown that solvent pH,[26–29] organic solvent content,[30–34] and droplet size[30,33,35,36] change along the plume. The complex dynamics of the ESI process make the transition of ions from solution-phase to gas-phase difficult to model. Atomistic molecular dynamics methods for such simulations have been used,[37–42] but are usually limited to the droplets containing up to a few thousand solvent molecules. Therefore, it is advantageous to directly perform optical spectroscopic measurements of physicochemical parameters of droplets and their changes in the electrospray plume in order to correlate the ions observed in the initial sample solution to those observed in the gas phase by the mass spectrometer.

More recently, several researchers[26,27,31–33,35,43] have developed laser-induced fluorescence strategies to probe properties of ESI plume, e.g. solvent fractionation, pH and temperature. They have correlated these changes with droplet size evolution along the ESI plume by mimicking electrospray sources (i.e. without actual MS measurements) and in combination with mass spectrometry.[28,30,34,36,44]

Solvent pH change of approximately 0.5 units along the ESI plume was observed by Girod et al.[28] The change is explained by evaporation of the solvent and thus the increasing concentration of acid (below pH 7) or base (above pH 7). For pH 7, the change in pH is different for positive and negative mode. This difference in the needle tip was first discussed by Zhou et al.[26] who showed that the pH at the needle tip and prior to spraying varies due to the electrochemical generation of excess OH^- ions (in negative-ion mode) or H^+ ions (in positive-ion mode). They and others have observed this phenomenon later as well.[27–29] Zhou et al.[31] studied changes in organic solvent content (acetone, acetonitrile, ethylene glycol, formic acid) and water binary mixtures and observed solvent percentage change up to 35% (percentage points, for acetone). Wang and Zenobi[32] and Hopkins et al.[33]showed that polarity increases with the decreasing droplet size due to solvent evaporation and water entrainment from the surrounding air. Girod et al.[30] and Liigand et al.[34] established that the biggest change (in percentage points) was observed at the edges of the spray plume and for more volatile organic solvents.

ESI+ and ESI- mode have seldom been profiled within the same study,[26–28,35] but even then the ESI spray parameters for positive and negative modes were different. Zhou et al.[26,27] and Girod et al.[28] have studied the pH change in water droplets and Wortmann et al.[35] have compared the droplet size for acetonitrile droplets for ESI+ and ESI-. However, usually, the mobile phase contains both water and organic solvent and the evolution of several parameters – pH, organic modifier content, droplet size – occurs simultaneously. Despite the fact that most ESI studies are carried out with acidic additives in the mobile phase, these studies have the lowest initial pH of 6.5.[28]

1.3 Ionization efficiency

Ionization efficiency (IE), the amount of ions generated from a specific compound in the ionization source, may vary from compound to compound by more than six orders of magnitude.[2,13,14,45–48] Different authors have used various terms to denote ionization efficiency such as (relative) response factor, molar response of the analyte, relative ion response, equimolar response factor. In ESI only a fraction of the analyte molecules are ionized in the ion source and only a part of the resulting gas-phase ions are successfully transmitted to the mass analyzer and are eventually detected.[49] Ionization efficiency in ESI is highly dependent on the solvent[2,14,50–54], ionization mode[2,14] as well as properties of the analyte itself.[10,11,13,46,55,56]

It has been found that solvent properties such as its surface tension, pH, additives and organic solvent as well as its percentage influence the ionization efficiency. Most often aqueous solutions of either acetonitrile or methanol are used as mobile phase in ESI. It has been shown that in many cases ESI response is higher in solutions with *higher organic modifier percentage*.[52] This may be due to more efficient desolvation of the droplets, allowing them to reach

Rayleigh limit faster and eventually generate smaller droplets more rapidly.[53] Tang and Kebarle[10] found that the signal of the analyte is affected by the *surface tension* of the solvent. It has been observed that different solvents have different influence on ionization in electrospray ionization mass spectrometry (ESI/MS) analysis.[2,14]

The *pH* of the mobile phase also influences ionization in the ESI plume.[54,57] It has been established that in general basic analytes provide higher sensitivity with more acidic mobile phase in ESI positive mode (ESI+) and acidic analytes with more basic mobile phase in ESI negative mode (ESI-).[14] It has also been shown that protonated forms of basic analytes can frequently be observed even if the pK_a (of the protonated analyte) is far below the solution's pH.[13,34,51] This phenomena, called wrong-way-round ionization was first observed by Mansoori et al.[58] and has been described by Zhou and Cook.[51] Similarly, deprotonated forms of analytes can be observed when ESI/MS analysis is performed with acidic solutions with pH lower than the analyte pK_a.[45,51] Therefore, it can only be concluded that pH is an important parameter influencing ionization.

1.3.1 Ionization efficiency in ESI positive mode

It has been found that more *hydrophobic* compounds tend to have higher ionization efficiencies. Cech et al.[11] found that in the case of selected tripeptides the non-polar surface area is affecting their signal in mass-spectrometer. They concluded that analytes need to have (1) large enough non-polar surface area to move to the surface of the droplet and (2) a structural element that allows charging (protonation, deprotonation, adduct formation, etc.). This tendency was also confirmed by Leito et al.[55] based on a study of esters and aromatic amines as well as by Cramer et al.[59] for drug-like molecules. Chalcraft et al.[56] modeled response factors for polar metabolites and found that important descriptors include molecular volume, octanol-water distribution coefficient and absolute mobility of the ion. Nguyen et al.[60] found a positive correlation between ESI signal and adjusted mass.

Another important factor of the analyte is the *acidity* of the compound.[13,46,54,61] Ehrmann et al.[46] tried to predict analyte signal in ESI/MS and observed the best predicting power while using the solution phase basicity of analyte (pK_b). The effect of gas phase proton affinities was smaller than previously stated for the studied compounds.[59,62] Oss et al.[13] calculated six physicochemical properties and found that ionization efficiency correlates best with pK_a value and *molecular volume* of the analyte. Hermans et al.[63] also observed a strong correlation between ESI response and molecular volume. Tang and Kebarle[10] found that the signal of the analyte is affected by the *surface tension* of the solvent and ion evaporation rate constant of the analyte.

1.3.2 Ionization efficiency in ESI negative mode

In negative ionization mode, similar parameters have been found to be important.[45,47,50,64,65] Huffman et al.[50] found that in negative mode compounds that are more *acidic* and *hydrophobic* ionize better. Similar trends were also observed by Alymatiri et al.[64] Henriksen et al.[47] found that for phenols and phenoxy alkanoic acids ionization depended on which *organic solvent* was used as well as on the hydrophobicity of the compound. They concluded that ionization was more efficient in methanol than in acetonitrile and that octanol-water partition coefficient of analyte (log$P$) was in better correlation with ESI/MS signal than the acidity of the analyte (pK_a). Ghosh and Jones[65] observed relative response factor increase with increasing acetonitrile content and increasing nonpolar surface area of the compound. Kruve et al.[45] discovered in the negative ionization mode that ionization efficiency can be best predicted by using a *degree of dissociation* (*α*) for the analyte and *charge delocalization* (*WAPS* parameter) which describes the delocalization of the charge in the anion.

1.3.3 Models to predict ionization efficiencies

Attempts have been made to predict ESI response based on the aforementioned physicochemical properties of the analyte by several groups[10,11,46,56,60,61,66–68] as well as by our own group.[13,45,55,69] The types of models have been various from simpler algorithms like multiple linear regression[13,45,56,59,63,67] to more sophisticated algorithms such as random forest[70] and artificial neural networks.[68]

Several research groups have made significant efforts to quantify the ionization efficiency and to reveal properties of the analyte, which are crucial in the ionization process. The correlations between ionization efficiency and evaporation rate,[71] log$P$,[72,73] hydrophobicity,[74] retention times of small peptides in reversed-phase LC,[75] non-polar surface area,[76] gas-phase proton affinity,[77,78] pK_a[73,79,80] and molecular surface area[73] have been observed. The obtained numerical models for predicting ionization efficiencies are vastly different and this has, until recently,[81] strongly limited the universal applicability of these predictions. Although considerable experimental support exists for a positive correlation between hydrophobicity and ESI ionization efficiency, there have been a number of studies[13,45,82] where the statistically significant correlation between the ESI/MS response and log$P$ has not been found. Some of the most crucial factors likely to contribute to the general confusion are a small number of compounds and eluent combinations included in the individual studies (up to 186).[83] However, for a physicochemical parameter to become statistically significant in the model, the value of the parameter needs to vary in a sufficiently wide range within the dataset. As the individual studies are limited to specific compound classes, it is highly likely that some parameters (log$P$, pK_a, gas phase basicity *GB*, etc.) will not become statistically significant simply

because all of the compounds were very similar from the perspective of this one variable.

One of the most reasonable solutions to overcome the confusion is to pool together ionization efficiency data from the literature for all available compounds resulting in a dataset of compounds with vastly different physico-chemical properties. This will allow:

(1) more accurate modeling, and a better understanding of the underlying processes;
(2) making informed decisions on optimum eluent and ESI polarity in addition to other possibilities, e.g. best derivatizing reagent choice;
(3) to validate existing ionization efficiency models;
(4) the possibility to study the phenomena of electrospray ionization with modern machine learning tools for which a large amount of data is essential to enhance the accuracy of the models.

The ionization efficiency data and findings discussed here have been used to develop a random forest model.[70]

1.4 Formation of multiply charged species

Many analytes (e.g. peptides) form multiply charged ions in ESI source. The multiply charged ions are very beneficial for compounds which form singly charged ions that have too high *m/z* values for most mass analyzers. Multiple charging and its reasons have been studied; however, much is still unclear in this field.

It has been shown that ionization efficiency of large multiply charged molecules depends on the distance between chargeable sites[84] and on the structure of the molecule (how well is the ionized form stabilized).[50] Nevertheless, it is still not fully clear how to predict the charge state and the ionization efficiency of the molecule in ESI/MS analyses. Wong et al.[85] developed a model for predicting the maximum possible number of charges of polyethylene glycol depending on the number of monomers and their affinity towards the ion that is merging with it. Later Schnier et al.[86] and Smith et al.[87] studied the number of protons that bind to the peptide in the gas-phase and found that the number of protons is very similar to the number of basic amino acid residues in the peptide. They also noticed that if another basic center is protonated in the close proximity of a basic residue, the basicity of that residue is decreased. The same tendency was also observed by Felitsyn et al.[88] who studied native proteins. Furthermore, the availability of basic sites has been shown to be important for determining the charge state of the analyte. [86,87,89,90]

It has been observed that the charge state of an analyte in solution does not always correlate with the charge state observed in MS analysis.[14] Iavarone et al. was the pioneer in studies of supercharging.[91] They observed an increase in the charge state for peptides if additives like diethylamine, 2-methoxyethanol, ethylene glycol, glycerol or 3-nitrobenzyl alcohol were added to the solution.[91-93]

Although the mechanism of supercharging is yet to be fully explained, the authors related this effect with the additives being less volatile than water and thus increasing surface tension. It has also been suggested that higher charge states are achieved by the denaturing effect of supercharging reagents.[94,95] Supercharging has also been observed in negative ion mode[96,97] where organic bases were used as supercharging reagents and a positive correlation was found between gas-phase basicities (*GB*) and charge state distributions. It was observed that in negative ionization mode the correlation was linear and in positive mode the correlation had a maximum value at *GB* of approximately 800 kJ/mol.[97]

There have been no studies done on ionization efficiency of multiply charged analytes to our knowledge, although its relevance has been pointed out.[11]

1.5 Applications of ionization efficiency

Currently, the only reasonable way to obtain quantitative information from ESI/MS is to use internal standards. If internal standards are not available, sometimes structurally similar compounds are used. In these cases, a set of compounds is measured and structurally most similar compound is used for quantification[98,99] despite leading to large errors.[100] For example, using 2-nitrophenol to quantify 4-nitrophenol can cause a serious (ca 40 times) misestimation of concentrations.[45] Therefore, it is important to account for the ionization efficiencies of different compounds.

Ionization efficiency values are useful for obtaining more accurate concentration estimations. Being able to predict ionization efficiency values aids in choosing the most optimal eluent composition, ionization mode, derivatization reagent and helps to estimates the limits of the used technique. For instance, knowing the ionization efficiency of a compound beforehand would allow choosing a solvent where the ionization efficiency is the highest, therefore allowing lower quantitation and detection limits. It saves time because it is not necessary to inject calibration graph solutions. Knowledge of ionization efficiency allows choosing the optimal concentration and would reduce cases, where the chosen concentration is either too low or the signal is saturated.

2. EXPERIMENTAL

Different solvents and IE measuring methods and equipment were used. See an overview of the used experimental setup in Table 1. For more details, please see the chapter below.

Table 1. Overview of used instruments, ESI polarities, sample introduction type (DIA – direct infusion, FIA – flow injection analysis) and used eluents. The instrumental setups are further explained in "Instrumentation" – "Ionization efficiency measurements" and eluent compositions are described in "Chemicals". In Paper VI data from literature was pooled and no experiments were carried out.

	Paper I	Paper II	Paper III	Paper IV	Paper V
Instrument	Agilent XCT	Agilent XCT and 3Q	Agilent 1Q	Thermo LTQ	Agilent XCT
ESI polarity	ESI−	ESI− and ESI+	ESI+ and ESI−	ESI−	ESI+
FIA/DIA	DIA	FIA	DIA	FIA	FIA
%MeCN	80%	80%	80% 50%	80%	80%
Water phase additive	0.1% formic acid; 0.1% ammonia	pH 4.00 (formic acid); 0.1% formic acid; 0.1% ammonia	Buffers pH 3.5 … 5.5 (0.5 pH unit increments)	0.1% ammonia	0.1% formic acid
Compounds	Indicators and dicarboxylic acids	*trans*-3(3-pyridyl)acrylic acid; amino acids, substituted benzoic acids and phenols, oligo-peptides, and poly-functional aromatic compounds	Nile Red, 5(6)-carboxy-2',7'-dichloro-fluorescein	Small pharma-ceutical compounds	Amino acids, short peptides
Laboratory[a]	UT	UT	UL	JJ	UT

[a]*UT – University of Tartu, UL – University Claude Bernard Lyon 1, JJ – Janssen Pharmaceutica NV.*

2.1 Ionization efficiency measurement procedures

As it is complicated to measure absolute IE values, the ionization efficiencies relative to an anchoring compound are measured. All IE values are expressed in logarithmic form as log*IE* values to make the data easier to present and to analyze. In ESI positive mode the log*IE* of methyl benzoate is taken as 0 value and in negative mode the log*IE* of benzoic acid is taken as 0. The log*IE* values of all the other compounds are expressed relative to these compounds. For measurements, benzoic acid in ESI− mode and tetraethylammonium in ESI+ mode were used as reference compounds, if not stated otherwise. The loga-rithmic ionization efficiency (log*IE*) of benzoic acid in 20/80 0.1% ammonia

solution/acetonitrile in ESI− mode has been previously taken as $0.^{45}$ The logarithmic ionization efficiency (logIE) of tetraethylammonium in 20/80 0.1% formic acid solution/acetonitrile in ESI+ mode has been previously measured as $3.95.^{13}$ Anchoring is used to give comparable values for all compounds. Ionization efficiencies were measured according to one of the two procedures: flow injection (FIA) and direct infusion (DIA) analysis.

2.1.1 Flow injection ionization efficiency measurement

This procedure uses autosampler to inject solutions of compounds one-by-one. The solution of anchor compound is injected at the beginning, middle and end of the sequence to ensure the repeatability of measurements throughout the sequence. Calibration graphs are constructed for both anchor compound as well as all the other compounds. logIE values are obtained from the slope of the corresponding calibration graphs of compound and reference compound:

$$\log IE(\text{A}) = \log \frac{slope(\text{A}) \cdot IC(\text{A})}{slope(\text{B}) \cdot IC(\text{B})} \tag{2}$$

where A is the compound for which the logIE is calculated and B is the reference compound.

The reproducibility of measurements is calculated as a consistency standard deviation ($s_{\text{consistency}}$).

$$s_{\text{consistency}} = 10^{consistency_{log}},$$
$$consistency_{log} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2}{n_1 + n_2 + \cdots + n_k - k}} \tag{3}$$

where n – number of replicate measurements per compound, s – standard deviation of logIE measurements for compound, k – number of compounds.

2.1.2 Direct infusion ionization efficiency measurements

This procedure uses a "T-piece" and syringe pumps to directly infuse the mixture of two compounds to MS. The concentration ratio is varied by changing the infusion rates of two pumps from 1.7 µL/min to 6.7 µL/min so that the sum of the infusion rates of two pumps would always be 8.4 µL/min. Using this procedure first relative ionization efficiency of a pair of compounds is measured:

$$RIE\frac{A_1}{A_2} = \frac{IE(A_1)}{IE(A_2)} = \frac{R(A_1) \cdot C(A_2) \cdot IC(A_1)}{R(A_2) \cdot C(A_1) \cdot IC(A_2)} \qquad (4)$$

where R_1 and R_2 are the responses monoisotopic peak of the compounds and C_1 and C_2 the respective concentrations of the compounds in the spray and IC_1 and IC_2 their corresponding isotope corrections. Only area of monoisotopic peak is integrated, all other isotopes are taken into account by IC, the isotope correction factor. For each compound pair, the RIE is measured on five concentration ratios and the obtained RIE values are averaged. First, the logIE values are temporarily assigned to compounds by minimizing the sum of squares (SS) of differences between measured logRIE values and the assigned logIE values[13] within one solvent:

$$SS = \sum_{k=1}^{n_m} \left\{ \log RIE_k(A_i, A_j) - \left[\log IE(A_i) - \log IE(A_j) \right] \right\}^2 \qquad (5)$$
$$\rightarrow \min$$

where n_m is the number of measurements and log$RIE_k(A_i, A_j)$ is the result of k-th measurement which has been conducted between compounds A_i and A_j. For this step, the logIE value of the reference compound (e.g. benzoic acid) is taken as zero.

Consistency of the scale is expressed as consistency standard deviation:

$$s_{\text{consistency}} = 10^{consistency_{log}},$$
$$consistency_{log} = \sqrt{\frac{SS}{n_m - n_c}} \qquad (6)$$

where n_c is the number of assigned logIE values (i.e. the number of analyzed substances) and n_m is the overall number of measurements.

2.1.3 Anchoring between media

In order to make logIE measurements in different media (e.g. between solvent and urine) comparable, it is necessary to measure one compound (anchor) in all the desired media. This approach was used to anchor the ionization efficiency values between (1) ESI+ and ESI−, (2) different solvents (Figure 2), and (3) solvent and biological matrices (Figure 3). The reference medium in ESI+ is 20/80 0.1% formic acid solution/acetonitrile and in ESI− it is 20/80 0.1% ammonia /acetonitrile. logIE values measured in other media are anchored to values measured in a reference solvent corresponding to the used ESI mode. The difference between logIE values of an anchor (B) in different media is found as:

$$\log IE(\text{B in M}) = \log \frac{R(\text{B in M}) \cdot C(\text{B in S})}{R(\text{B in S}) \cdot C(\text{B in M})}$$
$$= \log \frac{slope(\text{B in M})}{slope(\text{B in S})} \tag{7}$$

where S denotes the reference medium (solvent stated above) and M the medium of interest. The log*IE* measurements of anchor in different media should be carried out as close in time as possible to avoid drifts in instrument sensitivity. The final difference between media is calculated as an average of triplicate measurement carried out on different days.

The log*IE* values for each compound in the specific matrix are, thereafter, found as:

$$\log IE(\text{A in M}) = \log RIE(\text{A}) + \log IE(\text{B in M}) \tag{8}$$

where log*IE*(A in M) is the log*IE* value of compound A in medium M (anchored to reference solvent, numerical comparison between media possible), log*RIE*(A) is the log*IE* value of compound A relative to compound B in medium M (not anchored, numerical comparison not possible) and log*IE*(B in M) denotes the difference between log*IE* values of the anchor in reference solvent and in medium M.
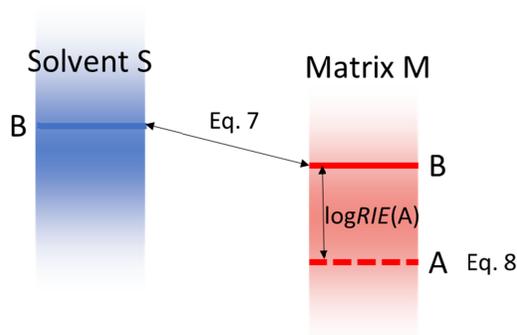
Figure 2. Anchoring between solvent (S) and matrix (M) by using compound B, which is measured in both media.
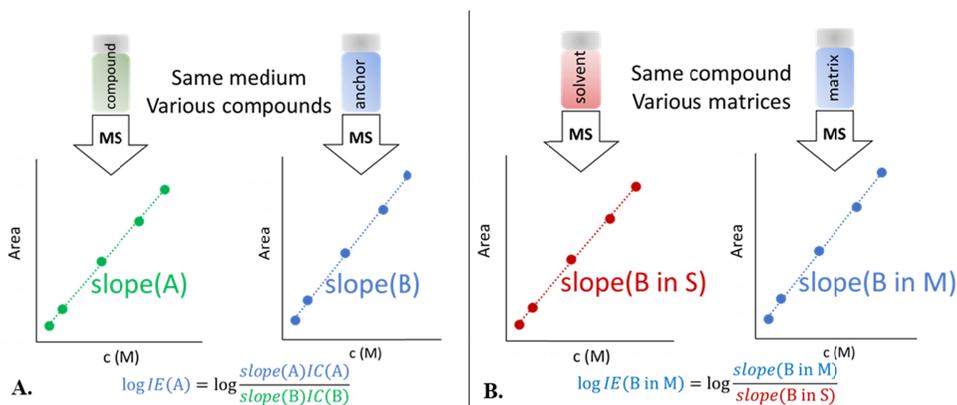
Figure 3. The process of log*IE* measurements within one medium (A) and the process of anchoring between different measurement media (B).

2.2 Instrumentation

2.2.1 Ionization efficiency measurements

Ionization efficiency measurements were carried out on mass-spectrometric setups:

(1) Agilent XCT ion trap mass spectrometer. Default settings for ESI source were used: nebulizer gas pressure 15 psi, drying gas flow rate 7 L/min, drying gas temperature 300 °C. The capillary voltage between MS and nebulizer was ±3500 V. All remaining ion transport parameters were determined by the Target Mass (TM) parameter, set by the operator. TM in the positive mode ESI was found according to ref.[69] and in the negative mode, the TM was set as closest value rounded to the nearest 50 of expected *m/z*.

(2) Agilent 6496 Triple Quadrupole mass spectrometer with Agilent Jet Stream source (AJS, Agilent Technologies, Santa Clara, CA, USA). Used AJS parameters were: nebulizer gas pressure 20 psi, drying gas flow rate 14 L/min, drying gas temperature 250 °C, sheath gas flow 11 L/min, sheath gas temperature 350 °C. The capillary voltage between MS and nebulizer was ±3000 V and nozzle voltage was ±1500 V.

(3) LTQ ion trap (Thermo-Fisher Scientific, San Jose, USA) mass spectrometer coupled with an Accela liquid chromatograph (Thermo Fisher Scientific, San Jose, USA). All measurements were carried out in the ESI negative MS scan mode. Sheath gas flow rate 35 psi, auxiliary gas flow 10 a.u., sweep gas flow rate 5 a.u., spray voltage -3.5 kV, and capillary temperature 275 °C were used. The flow rate was 0.2 mL/min.

2.2.2 Measurement of the degree of ionization using NMR and UV-Vis

Determination of degrees of ionization for a reference compound *trans*-3(3-pyridyl)acrylic acid was performed on 700 MHz NMR Bruker Avance II 700 NMR spectrometer. ^{1}H NMR spectra of *trans*-3(3-pyridyl)acrylic acid were registered in acetonitrile-d$_3$:H$_2$O (80/20 v/v) mixtures at aqueous pH 1.0, 3.0, 4.0 and 10.0. For NMR, the chemical shift of the protons was used to calculate the degrees of ionization for the reference compound in the solvents (based on the calibration with solutions with known degrees on ionization, see Appendix 1). Similar approach was used for UV-Vis spectrophotometric titration to verify the results. Double beam spectrophotometer Thermo Nicolet Evolution 300 was used. Scan speed was intelliscan mode (from 120 nm/min to 1200 nm/min, scan speed changes depending on how strongly absorbance depends on the wavelength), bandwidth was 1.5 nm and data registration interval was 1 nm. Results of the UV-Vis measurements confirmed the results obtained by ^{1}H NMR experiment. (See Appendix 1for details).

2.2.3 ESI plume pH, organic solvent and droplet size profiling

The experimental setup profiling the spray plume consists of an excitation laser and two different optical detection systems mounted on a moving stage (see Appendix 2 for the scheme). It is coupled with Agilent Single Quad 6100 mass spectrometer equipped with a modified Agilent Jet Stream ESI source (Agilent Technologies, Santa Clara, CA, USA) in order to allow the laser injection into the plume and the epifluorescence measurements. Used ESI parameters were: capillary voltage 3500/-3500 V, nebulizing gas pressure 15 psi, drying gas flow rate 7 L/min, drying gas temperature 300 °C, sheath gas flow rate 1 L/min and temperature 80 °C. Solutions were introduced in the ionization source at 50 µL·min^{-1} flow rate with a KDS100 syringe pump (KD Scientific, Holliston, MA, USA).

A continuous laser (λ = 473 nm) emitting in a single longitudinal mode was used to profile ESI plume. The output power of the laser was around 500 mW and its beam diameter is 1.5 mm (divergence 1 mrad). The laser is injected through the objective using two reflecting protected aluminum flat mirrors ($R >$ 90%). The laser beam is focused into the spray and the fluorescence is collected via an objective used in an epifluorescence configuration. Fluorescence spectra from ESI plume were recorded, point by point (pixel size 500 µm), by an ultra-compact spectrophotometer (B&WTek Inc., Newark, DE, USA). The plume was profiled for 15 mm starting from ESI needle tip, except for 50% acetonitrile solution where the plume was profiled for 13 mm due to poor solubility of the fluorescent probe.

First of all, the chromism of Nile Red (20 µM) was calibrated in acetonitrile/water binary solvent mixtures. Variation in the acetonitrile/water

ratio induces a shift of the maximum emission wavelength. The solvent composition of an unknown solution can be determined based on the λ_{max} using the calibration curve. Acetonitrile content was profiled in ESI+ and ESI− for solutions initially containing 80% and 50% acetonitrile (v/v) from fluorescent measurements of the solvatochromic dye Nile Red.

For pH measurements, stock solutions containing 0.1 mM of 5(6)-carboxy-2',7'-dichlorofluorescein were made in dimethyl sulfoxide. From the initial solution, 10 μM binary acetonitrile/water solutions were prepared, where the water phase pH (denoted in the text as pH) varied: 3.50, 4.00, 4.50, 5.03, 5.52, 6.04, and 6.51. Water phase solutions were prepared by first making 0.1% formic acid solution and then adjusting pH by adding the 25% ammonium hydroxide solution until the desired pH. Water phase pH values were measured with Hanna Instruments pH211 Microprocessor pH Meter equipped with a 4 mm-diameter microelectrode (Pt // 3.5 mol/L KCl+AgCl). The obtained water phase solutions were then mixed in different ratios with acetonitrile so that acetonitrile content in the calibration solutions was 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%. In order to establish calibration curves, fluorescence spectra of all prepared 63 mixtures were measured and the logarithm of the ratios of fluorescence emission intensities was calculated. pH was profiled for mobile phases with initial compositions of acetonitrile/aqueous solution with pH of 5.03 in the ratio of 80/20 and 50/50, and also for 80/20 mixture with an aqueous solution with a pH of 4.00. The corresponding $pH_{abs}^{H_2O}$[a] changes were measured by Dr. A. Heering and are listed in Appendix 2.[101] The change of the droplet size was studied for the same mobile phases.

The obtained fluorescence spectra are an average of droplets present in the size of the laser beam (~1 mm). So the pH and mobile phase composition measured correspond to an area of ~0.78 mm² in the ESI plume and not for a unique droplet.

The obtained raw data was processed using OriginPro 7.0 software. Profiles of pH and solvent composition in the ESI plume (Figure 13) were determined from calibrations in different solutions.

The temperature of the droplets was not profiled. However previous studies performed with methanol droplets have shown either a slight increase of temperature[36] or some decrease[102] along the plume. The temperature of the plume in the referred studies was found to be between 295 K and 307 K.

2.3 Physicochemical parameter calculations

COSMO-RS method[103] was used for calculating various parameters: aqueous pK_a, log$P$ (octanol-water), charge delocalization parameters (*WAPS*/*WANS* and

---

[a] The notion $pH_{abs}^{H_2O}$ means that pH is expressed on the absolute scale, but values are shifted by a constant in order to make the pHabs values directly comparable to the conventional aqueous pH values.

Klamt parameters). In Paper II ACE and JChem acidity and basicity calculator[104] was used to calculate pK_a values. Degree of dissociation α of the compounds was calculated from the computed pK_a values and the water phase pH. The calculated physicochemical parameters can be found in Appendices 3–5.

First, full geometry optimization and energy calculation was carried out at the DFT BP TZVP level with the RI approximation and applying the COSMO continuum solvation model for all compounds using Turbomole, version 6.4.[105] For most compounds several conformers corresponding to different local energy minima were found. All of these were taken into account by statistical weighing inherent in the COSMO-RS procedure. The default convergence criteria of Turbomole were used: wave function convergence max difference 10^{-6} Hartree, geometry convergence max gradient |dE/dxyz| 10^{-3} Hartree/Bohr. This first computation step yields for every conformer the following data: the geometry of the conformer, detailed data on the shape of molecular cavity, the polarization charge densities mapped onto the cavity surface, the total electronic energy of the species submerged into a virtual conductor ($\varepsilon = \infty$), and molecular surface area and volume. Molecular cavity refers to the cavity constructed for the particular conformer within the COSMO solvation theory ‑ constructed using smoothed spheres using atomic radii ~20% larger than van der Waals radii. This cavity was later used as the molecular volume. The cavity surface refers to the so-called sigma-surface ‑ polarization charge density on the molecular surface. For further information about the COSMO-RS theory see reference [103].

Secondly, the COSMO-RS calculation was carried out on all compounds using the above-listed data as input data with the COSMO*therm*, version C3.0, release 14.01.[106] COSMO-RS calculations take into account the interactions between species and the solvent/medium molecules, as well as between the solvent molecules themselves (implicit solvation model). The solvent composition is a required input parameter for COSMO-RS calculations and in these calculations, water was used as a solvent. Zero concentrations were used for the studied molecules. This way the interactions between the studied compounds and the solvent is taken into account but not the interaction between the molecules of studied compounds themselves. This situation corresponds well to the reality of very low concentrations used in the experiments and is common practice for carrying out such calculations. Both van der Waals interactions (electrostatic interactions: dipole-dipole, ion-dipole, etc. forces as well as dispersion forces) and hydrogen bonds (implicitly) are taken into account. These interactions are quantified via statistical counting and averaging of energies of pairwise interactions of molecular surface segments using polarization charge density maps of compounds created in the first step, taking into account the concentrations of the respective species in the solution. Terms accounting for vibrational contributions to the G_{tot} are also added in this step. This is done implicitly, as these are represented through the experimental data used for parameterization of the method. The energetics of these interactions are calculated at the 298 K, using statistical thermodynamics procedure whereby

also the conformers of all the interacting molecules are taken into account and statistically weighted based on their relative stabilities. This way, the entropy effect of the same species present in multiple conformers is also accounted for. As a result, G_{tot} value is found for every compound.

Degree of ionization is calculated from the pK_a value of the compound of interest and pH value of the eluent in ESI+ mode:

$$\alpha = 1 - \frac{1}{1 + \frac{10^{-pH}}{10^{-pKa}}} \tag{9}$$

Used pH value is that of the water phase and pK_a value is calculated with COSMO-RS in water. As it is difficult to account for processes in the spray that lead to the change in solvent properties and since the water content of the droplets in the ESI spray increase during evaporation of a more volatile organic phase, a simplification is usually done and the properties of analyte in water phase are used.[13]

And in ESI− mode:

$$\alpha = \frac{1}{1 + \frac{10^{-pH}}{10^{-pKa}}} \tag{10}$$

Charge delocalization parameter is calculated as weighted average positive sigma for anions (*WAPS*) and as weighted average negative sigma for cations (*WAPS*):[107]

$$WAPS = \frac{\int_{\sigma=0}^{\infty} \sigma \cdot p(\sigma) d\sigma}{A \int_{\sigma=0}^{\infty} p(\sigma) d\sigma} \tag{11a}$$

$$WANS = \frac{\int_{\sigma=0}^{-\infty} \sigma \cdot p(\sigma) d\sigma}{A \int_{\sigma=0}^{-\infty} p(\sigma) d\sigma} \tag{11b}$$

where σ is the polarization charge density on the surface of ion, $p(\sigma)$ is the probability function of σ and A is the surface area of the ion. The smaller the *WAPS/WANS* absolute value, the more delocalized the charge in the ion. It has been proposed that values above absolute value of 4.5 indicate ions with localized charge.[107]

2.4 Modeling ionization behavior

2.4.1 The model predicting multiple charging (Paper I)

To find out why certain substances give multiply charged species in the mass spectrum and why others do not, it is necessary to know the physicochemical properties of all of the substances. In addition to the properties of analytes, also the properties of the solvent need to be taken into account. It is known that some processes that lead to ionization of the analyte occur in the solvent phase and some in the gas phase.[2] Since it is difficult to account for processes in the spray that lead to the change in solvent properties a simplification is usually done and the properties of analyte in water phase are used.[13] In the ESI spray during droplet evaporation, water content increases as the more volatile organic component vaporizes. It is also significantly less complicated to measure water phase pH values. The same assumption is used in the current work. For model development, physicochemical properties obtained by COSMO-RS calculation were used. All statistical tests were carried out at 95% confidence level. Linear discriminant analysis (LDA) was carried out with statistical program R[108] using the package Mass.[109]

2.4.2 Ionization efficiency prediction models in matrices (Paper IV)

Based on the calculated physiochemical parameters and measured log*IE* values a predictive model was fit in each matrix. Multilinear regression analysis was used to obtain the model describing the relationship between log*IE* and physico-chemical properties. The general form of the equation was:

$$\log IE = coef_{WAPS} \cdot WAPS + coef_{\alpha} \cdot \alpha + intercept \qquad (12)$$

where the coefficients depend on the matrix.

For each model root-mean-square error (s_{RMSE}) was found to describe the differences between predicted log*IE* values and measured values.[110]

$$s_{\mathrm{RMSE}} = 10^{RMSE_{log}},$$
$$RMSE_{log} = \sqrt{\frac{\sum_{i=1}^{n}\left(\log IE_{predicted} - \log IE_{experimental}\right)^2}{n}} \qquad (13)$$

Additionally, the goodness-of-fit test was used to estimate the quality of the developed matrices.

$$F = \frac{\sum \left(\log IE_{predicted} - \overline{\log IE_{experimental}} \right)^2 / (n-1)}{\sum \left(\log IE_{experimental} - \log IE_{predicted} \right)^2 / (r-n)} \qquad (14)$$

Where n is the number of compounds and r is the number of concentration levels incorporated into the calibration graph and $\overline{\log IE_{experimental}}$ denotes the mean value of all the measured logIE values. From F-values the p-values were calculated using the degrees of freedom of the numerator and denominator. Higher p-values indicate higher explained variation in logIE values by the model.

In order to validate the obtained results, the cross-validation method 'leave-one-out' (LOO) approach was used. Cross-validation was preferred due to the need to estimate the applicability of the method over a wide range of logIE values. LOO approach means that each compound was left out from the model fitting process once; thereafter, the model was used to predict the logIE value of the compound not involved in the model development. After this, the process was repeated for another compound, so that each compound was left out once from the model development. In the case of conventional validation set approach, the logIE values could have been predicted only for 2 to 3 compounds, which would provide insufficient information about the model.

2.5 Transferring data from literature to uniform dataset (Paper IV)

Ionization efficiencies of all compounds were calculated using Quantem electrospray ionization efficiency prediction model in corresponding ESI mode developed by J. Liigand was used.[81] This model is based on PaDEL descriptors[111] of the compound and empirical eluent descriptors: viscosity,[112] surface tension,[113] polarity index[114] and water phase pH. For model development regularized random forest algorithm[115] from RRF: Regularized Random Forest library in R was used. For data treatment, in-house developed R-script was used. PaDEL descriptors have been calculated from SMILES notation using ChemDES online platform.[116] The predicted ionization efficiencies corresponded to universal ionization efficiency scale comparable to previous results.[13,45]

2.5.1 Accuracy estimation

The accuracy of prediction is described as a root-mean-square error, s_{RMSE} similarly as stated in Eq. 13, and as a mean absolute error, s_{MAE}:

$$s_{\text{MAE}} = 10^{MAE_{log}}, MAE_{log} = \frac{1}{n} \sum_{i=1}^{n} \left| \log RRF_{pred} - \log RRF_{comparable} \right| \qquad (15)$$

This means that if the ionization efficiency of compound A is predicted to be 100-times higher than the ionization efficiency of the methyl benzoate and s_{RMSE} is 2.2, the actual ionization efficiency of compound A would be 45 (= 100/2.2) to 220 times (= 100 · 2.2) higher than that of methyl benzoate ($\log IE = 2.00 \pm 0.34 = \log 100 \pm \log 2.2$).

2.6 Eluents

As eluent components acetonitrile (J.T.Baker, Deventer, Netherlands, HPLC grade), MilliQ water (Millipore Advantage A10 MILLIPORE GmbH, Molsheim, France), formic acid (Fluka, 98%, Buchs, Switzerland) and ammonium hydroxide (Lach:Ner, 25%, Czech Republic), buffer pH 7.00 (Fluka, Buchs, Switzerland) were used. Eluent compositions were described in Table 1.

2.7 Chemicals

Dimethyl sulfoxide (Sigma, Steinheim, Germany) was used to prepare stock solutions of fluorescent probes: fluorescent pH indicator 5(6)-carboxy-2',7'-dichlorofluorescein (Sigma, ≥ 95%) and solvatochromic Nile Red (Invitrogen, Cergy, Pontoise, France).

Ionization efficiencies were measured for 3-nitrophthalic acid, adipic acid, eosin B, bromophenol blue, bromothymol blue, phenol-2,4-disulfonic acid, phenolphthalein, cresol red, *m*-cresol purple, tiron, thymolphthalein, benzoic acid, salicylic acid, sorbic acid (Reakhim, Russia), bathocuproinedisulfonic acid, SPADNS (Chemapol, Czech Republic), bromocresol purple (Schering AG Berlin, Germany), bromocresol green, eosin Y (Sigma-Aldrich, USA), maleic acid, thymol blue (E. Merck Darmstadt, Germany), 5-sulfosalicylic acid (Lachner, Czech Republic), glutaric acid, pimelic acid, suberic acid (Aldrich, USA), phthalic acid, fumaric acid, isophthalic acid, itaconic acid, succinic acid, mesaconic acid, terephthalic acid (obtained as a kind gift from the Institute of Pharmacology, Tartu, Estonia), lincomycin hydrochloride, dodecanoic acid, fumaric acid (Sigma, Steinheim, Germany), warfarin (DuPont Pharma, Wilmington, DE, USA), naproxen (Synthex Research Center, Edinburgh, UK), taurocholic acid sodium salt hydrate (Acros Organics, Geel, Belgium), 3-[(trifluoromethyl)sulphanyl]benzoic acid (3-CF_3SO_2-benzoic acid, a kind gift from prof. L. M. Yagupolskii), tetraethylammonium perchlorate (Fluka, Buchs, Switzerland), *trans*-3(3-pyridyl)acrylic acid (Aldrich, St. Louis, USA). The ionization efficiencies of amino acids from the L-amino acid kit (Sigma, Germany) were measured: α-alanine, glycine, lysine, phenylalanine, proline, aspartic acid, threonine, asparagine, methionine, glutamic acid, glutamine, tyrosine, leucine, tryptophan, cysteine, histidine, valine, isoleucine, arginine, serine. In addition, the ionization efficiencies were measured for β-alanine

(Fluka, Switzerland) and oligopeptides from American Peptide Company: Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu, Tyr-Ile-His-Pro-Phe, Asn-Arg-Val-Tyr-Ile-His-Pro-Phe, Arg-Pro-Gly-Phe-Ser-Pro-Phe-Arg, Asp-Arg-Val-Tyr-Ile-His-Pro-Phe, Val-Tyr-Ile-His-Pro-Phe, Ile-His-Pro-Phe, Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$, Asp-Arg-Val-Tyr-Ile-His-Pro, Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg, Phe-Phe-Gly-Leu-Met-NH$_2$, Arg-Arg-Pro-Tyr-Ile-Leu, Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly, Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH$_2$, Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$, Gln-Phe-Phe-Gly-Leu-Met-NH$_2$, Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$; from KJ Ross-Petersen Aps, Denmark: Arg-Pro-Pro, Trp-Ala-Gly-Gly-Asn-Ala-Ser-Gly-Glu, Arg-Pro-Pro-Gly-Phe, Thr-Arg-Ser-Ala-Trp-NH$_2$, Arg-Pro-Pro-Gly-Phe-Ser-Pro-Leu, Gly-Lys-Pro-Ile-Pro-Asn-Pro-leu-Leu-Gly-Leu-Asp-Ser-Thr, Thr-Arg-Ser-Ala-Trp, Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met, Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu, Arg-Arg-leu-Ile-Glu-Asp-Ala-Glu-Tyr-Ala-Ala-Arg-Gly, Arg-Pro-Pro-Gly-Phe-Ser, Cys-Tyr-Phe-Gln-Asn-Cys, Lys-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Leu, Arg-Pro-Pro-Gly-Phe-Ser-Pro; from Bachem, Germany: Ac-Gly-Lys-OMe, Gly-βAla-βAla, Phe-Phe-Phe-Phe, Gly-Gly-Asp-Ala, Gly-Pro-Gly-Gly; and two were synthesized in house (purity confirmed by LC/MS): Gly-Gly-Gly-Phe-Phe-NH$_2$, Gly-Gly-Gly-NH$_2$.

Biological matrices liver tissue, brain tissue, urine, and blood from a healthy dog (beagle) were obtained from in-house sources at Janssen Pharmaceutica (Beerse, Belgium), plasma and cerebrospinal fluid (CSF) of a healthy dog (beagle) were obtained from Bioreclamation IVT, USA. For brain and liver tissue, 1 part of tissue was homogenized with 9 parts of MilliQ water to form tissue homogenates. Biological matrices were stored frozen at -20 ºC, except for blood which was used fresh (within 2 hours). For plasma and blood, K$_2$EDTA was used as an anticoagulant. A simple standard protein precipitation sample preparation was carried out: 50 µL of the stock solution of the compound was added to a mixture of 400 µL of acetonitrile and 50 µL of biological matrix: plasma, urine, whole blood, CSF, liver or brain tissue (1 part of tissue homogenized with 9 parts of water). This mixture was thoroughly mixed and centrifuged for 10 min at 13 000 g. The supernatant (injection volume 5 µL) was used for MS analysis.

In the direct infusion experiments sum of the flow rate was 8.4 µL/min and in flow injection analysis the flow rate was 0.2 mL/min. Concentrations of each compound were chosen so, that the signal would be in the linear range. During every measurement, the linearity of signal to concentration graph was checked and thereby it was verified that the signal was not saturated in any of the measurements.

3. RESULTS AND DISCUSSION

3.1 Incorporating multiply charging compounds

3.1.1 ESI− multiple charging on the example of indicators

As the ionization efficiencies of multiply charged ions have not been studied, the first aim was to study multiple charging in negative ionization mode in acidic and basic solvent (80/20 MeCN/0.1% formic acid or 0.1% ammonia, respectively). In the literature there is much ambiguity of the charge states of the compounds and, therefore, it is aimed to reveal the factors that affect the ability to yield multiple charging in ESI/MS based on analysis of small molecules in negative ESI ionization mode. It is aimed to propose a quantitative model that can be used to predict the charge state.

The ionization efficiency scale was compiled in both basic and acidic solvent (Table 1) and for both doubly and singly charged ions. Altogether 29 compounds were studied covering a wide range of acidity (aqueous pK_a values from -3 to 11) and a wide range of hydrophobicity ($logP$ from -0.5 to 7.5). Small molecules allow determining conditions and parameters affecting multiple charging based on a relatively simple system. This is a good starting point to move on to bigger systems such as peptides. 9 of the 29 studied compounds gave multiply charged ions in addition to singly charged ions in the mass spectrometer in both solvents (given in Table 2 with the corresponding $logIE$ values). Additionally, also phenol-2,4-disulfonic acid gave doubly charged ions, but due to spectral interferences, it was not possible to measure the corresponding $logIE$ value. Benzoic acid was taken as the reference substance, meaning that its $logIE$ value was set arbitrarily to be zero in basic solvent as defined earlier in literature.[45] Altogether 23 measurements in basic and 17 measurements in acidic solvent were made. In basic solvent, the ionization efficiency scale range was about 3 logarithmic units for doubly charged ions and 4 logarithmic units for singly charged ions. In acidic solvent, the corresponding values were 4 and 3 logarithmic units. For singly charged ions the ionization efficiency values did not statistically differ in two solvents according to the t-test (95% confidence level was used). For doubly charged ions the $logIE$ difference in two solvents was statistically significant for all compounds, except for tiron and SPADNS.

Table 2. log*IE* values in negative ionization mode for molecules that gave doubly charged species. In this chapter log*IE*$_1$ is the log*IE* of the molecule via formation of singly charged ion; log*IE*$_2$ of the molecule via formation of doubly charged ion.

	pK_{a1}	pK_{a2}	log$P$	Doubly charged					Singly charged		
				α_2 pH=10.74	α_2 pH=2.68	logIE_2 pH=10.74	logIE_2 pH=2.68	$WAPS$ ·10^{5a}	logIE_1 pH=10.74	logIE_1 pH=2.68	$WAPS$ ·10^5
Bromophenol blue	0.99	2.96	6.86	1.00	0.34	2.63	0.28	1.86	2.34	1.93	1.54
Bathocuproine-disulfonic acid	-1.70	5.05	-0.06	1.00	0.00	1.95	1.06	1.92	0.45	0.42	2.02
Bromocresol green	0.55	3.49	6.63	1.00	0.13	1.71	-0.15	1.83	1.45	2.01	1.62
Bromothymol blue	1.15	5.76	7.47	1.00	0.00	1.23	-0.25	1.70	2.95	2.16	1.44
Eosin Y	3.30	4.54	5.03	1.00	0.00	1.15	0.27	1.90	1.50	1.47	1.60
SPADNS	-2.32	-2.19	1.41	1.00	1.00	0.66	0.16	2.57	-0.20	-0.58	2.34
Sulphosalicylic acid	-1.53	2.31	2.68	1.00	0.70	-0.38	-2.54	2.60	0.22	0.22	2.38
Tiron	-2.66	-2.15	3.07	1.00	1.00	-0.62	-0.29	5.01	-0.68	-0.34	3.98
$S_{\text{consistency}}$						0.16	0.15		0.40	0.08	

It was observed that analytes with certain functional groups tend to give doubly charged ions. All substances that gave doubly charged ions had at least one sulfo group, also in most cases (except bathocuproinedisulfonic acid) one or several electronegative groups (hydroxyl or bromo) and were aromatic. Dicarboxylic acids did not give doubly charged species in the MS independent of the pK_a values and distance between carboxyl groups. The maximum distance between carboxylic groups in the case of dicarboxylic acids was about 10 Å in the case of suberic acid (octanedioic acid). In the case of tiron or sulfosalicylic acid, for example, which formed doubly charged ions, the distance between charged groups was about 5 Å.

Linear discriminant analysis (LDA) was carried out to find properties to predict whether an analyte gives or does not give doubly charged ions in negative ionization mode. LDA model was found using a training set composed of 18 randomly chosen compounds. The validation set composed of 11 substances.

Among the training set, there were 5 substances that gave doubly charged ions and 13 that did not give doubly charged ions (see Appenfix 3). Different combinations of molecular parameters calculated with COSMO*therm* were used (log*P*, *WAPS* parameters for anions and dianions, molecular area, pK_a values). Model was chosen so that it would give the highest prediction precision for the training set. This model was based on pK_{a2} and log*P* values:

$$F = -0.48 \cdot pK_{a2}(A) + 0.60 \cdot \log P \qquad (16)$$

If $F > 0$ then analyte A gives doubly charged ions,
If $F < 0$ then analyte A does not give doubly charged ions.

This model has a prediction accuracy of 94%. Only slightly worse prediction accuracy (89%) was obtained with a model that contained molecular volume or *WAPS* instead of log*P* value. This model applies in acidic solvent as well as in basic solvent since the same compounds gave doubly charged ions in both solvents, although with different ionization efficiency. From the LDA model, it can be seen that main properties that determine multiple charging are acidity and hydrophobicity. One important parameter is pK_{a2} that describes the dissociation of the second acidic group. However, several substances that according to pK_a value of the substance and pH of the solvent should be charged, did not give doubly charged ions (for example eosin B, phthalic acid, pimelic acid). Only analytes that had substantially lower pK_a than solvent pH gave doubly charged ions. Similar result was also obtained by Felitsyn et al.[88] who studied multiple charging of peptides. They showed that obtaining a second charge in the ESI process is hindered by electrostatic repulsion of charges and influenced by the hydrophobicity of the compound. Dissociation makes acids more hydrophilic that in turn makes them stay inside the droplet. This means that only substances with substantial hydrophobic character can move to the

surface of the droplet in the ESI process. As a result, only substances with reasonable acidity (in this study pK_{a2} value below 5.8) combined with remarkable hydrophobic character can give doubly charged ions in the ESI process.

This model was tested on a validation set containing 11 substances of which 4 gave doubly charged ions. The prediction precision was 82% for the validation set. LDA gave incorrect results for three substances (eosin B, bathocuproinedisulfonic acid and bromocresol purple). The false positive result for eosin B and bromocresol purple could be caused by additives in the substance that could suppress signal. Therefore, it is possible that eosin B could give doubly charged ions but its intensity was below the limit of detection. On the other hand, it is possible that some other parameter is affecting the formation of doubly charged species that has not been considered in this study. The $\log P$ value of bathocuproinedisulfonic acid is very similar to substances that did not give doubly charged species and probably this is the main reason the prediction is false negative. In the calculation of physicochemical properties of bathocuproinedisulfonic acid zwitterionic conformers were included and observed to be the most favored conformers in the liquid phase. It has been shown by Teesch et al.[117] that the structure of gas-phase ions and solvent phase ions are, however, different. Consequently, it can be presumed that $\log P$ is not the best parameter to describe the distribution of zwitterionic compounds such as bathocuproinedisulfonic acid on the surface of the droplet.

Parameters calculated with COSMO-RS method were correlated with measured $\log IE$ values in a basic solvent. The correlation was poor ($R^2 < 0.55$) with Klamt parameters (polarity, polarizability, hydrogen bond accepting capability, hydrogen bond donating capability; see Appendix 3 for details). It was not possible to correlate α_2 to ionization efficiency values obtained in the basic solvent because all the analytes were fully deprotonated and degree of dissociation therefore equal to one for all analytes. With the rest of the parameters, the correlation was good. The best correlation was observed between pK_{a1} and ionization efficiencies corresponding to singly charged analytes ($R^2 = 0.79$, the correlation was positive). Good correlation was also obtained with $WAPS$ and hydrogen bond donating capability (H_{don3}) for the neutral species (R^2 is 0.68 and 0.72 accordingly). For doubly charged ions good correlation with $WAPS$ parameter, molecular volume (MV) and hydrogen bond donating capability (H_{don3}) was observed. In acidic solvent the correlations were similar; however, there was no correlation between doubly charged analytes ionization efficiency and $WAPS$ parameter.

The span of the $\log IE$ scale and order of the substances in the scale was found to depend on solvent pH (see Figure 4). In acidic solvent, the range of the $\log IE$ scale for doubly charged species is one logarithmic unit wider than in basic solvent. Therefore, the pH is a suitable tool to increase or decrease MS sensitivity for doubly charged species. Generally, the ionization efficiencies for doubly charges species are about one logarithmic unit lower in the basic solvent than for acidic solvent. Also, the order of the $\log IE$ values of the substances is

dependent on solvent pH. The ionization efficiency values for singly charged species are less influenced by solvent pH.

In basic solvent, more doubly charged species are formed in the solution phase (pK_{a2} is substantially lower than solvent pH allowing the second protonation step) and the signal corresponding to doubly charged species and corresponding logIE values are significantly higher. In acidic solvent, the formation of doubly charged ions is not favored and more singly charged ions are formed. As a result, the logIE values of doubly charged ions are lower than in basic solvent. Only for SPADNS and tiron the logIE values in basic and acidic solvent do not change which can be explained by their very low pK_{a2} values (negative) and degrees of dissociation (α_2 value) of 1 in both solvents. In basic solvent, only bromothymol blue and eosin Y give higher logIE values for singly charged species than for doubly charged species. As mentioned before the changes between acidic and basic solvent for singly charged species are statistically insignificant according to t-test.

In both solvents, bromothymol blue has the highest logIE value for singly charged species. This is probably because it has the highest hydrophobicity and therefore the singly charged species will move to the surface of the droplet easily, even if not a lot of them are formed in solution. The lowest logIE value belongs to tiron, which is a relatively small molecule with high hydrophilic character compared to other studied compounds. Also, another similar small and hydrophilic molecule, sulfosalicylic acid, has low logIE value.
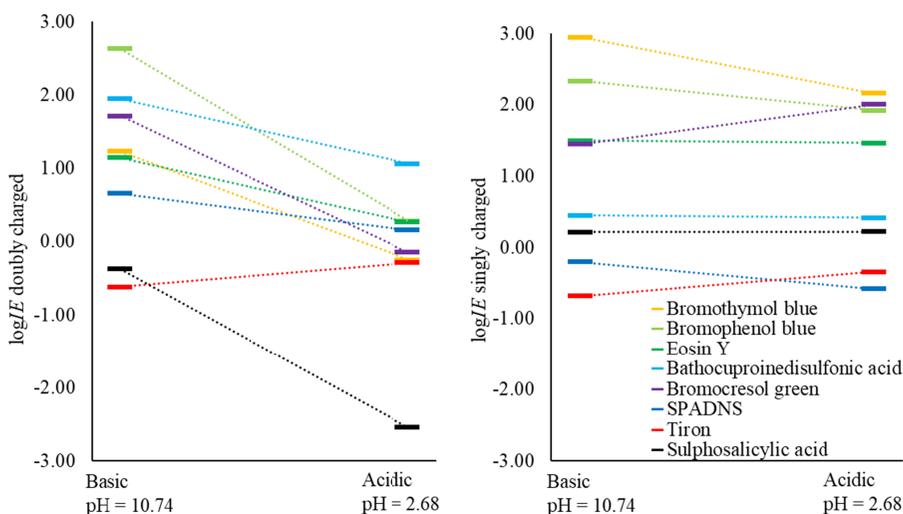
Figure 4. Ionization efficiencies of analytes in acidic and basic solvents, where logIE is the logIE of the molecule via formation of singly or doubly charged ion.

In acidic solvent higher $\log IE_1$ for singly charged species were observed for analytes that (1) have higher molecular volume and are more hydrophobic (higher $\log P$ value) and (2) give doubly charged species with lower $\log IE_2$ (bromothymol blue, bromocresol green and bromophenol blue). Among the previously mentioned compounds, the highest $\log IE_1$ value belongs to bromothymol blue, which also has the highest hydrophobicity. Substances with lower $\log IE$ values either give doubly charged ions with high $\log IE_2$ value or are relatively hydrophilic substances such as sulfosalicylic acid and tiron. This phenomenon can be explained as if the compound forms extensively doubly charged species; therefore, there is only a limited amount of compound left for forming singly charged species. Relatively large eosin Y has medium ionization efficiency and is different from others in respect that it does not contain sulfo groups.

To formulate a model for predicting $\log IE_1$ via formation of singly charged ions, different parameters were tested that had the best correlation with $\log IE_1$ (pK_{a1}, *WAPS,* and hydrogen bonding capacity for neutral substance) and also their combinations. The best parameter to describe ionization efficiencies ($\log IE_1$) was charge delocalization parameter *WAPS*. It can be concluded that charge delocalization is the most important parameter.

All in all, it can be seen that $\log IE_2$ values in basic and acidic solvents differ for doubly charged species, which shows that $\log IE$ values depend on pH. Solvent pH is, therefore, an important parameter to take into account to form a universal $\log IE$ prediction model and also to reach higher selectivity in the analysis. It can be seen from the previous discussion that the formation of singly and doubly charged species are strongly related processes.

It was of interest if a predictive model for $\log IE$ values can be obtained based on physicochemical properties of analyzed substances. Best quantitative model (both for basic and acidic solvent) for predicting singly charged analytes ionization efficiency was obtained by using *WAPS* values for singly charged ions:

$$\log IE_1 = \left(-1.14 \pm 0.23\right)\cdot WAPS + \left(3.36 \pm 0.51\right) \qquad (17)$$

Prediction precision of this model can be estimated by the root-mean-square error of 5.1 times (see explanation in 2.4.1 and 2.5.1) and square of correlation coefficient (R^2) of 0.64. Only substance to deviate significantly from the model was tiron. However, for doubly charged ions, the obtained models for predicting $\log IE$ values quantitatively gave unacceptable results.

3.1.2 ESI+ multiple charging on the example of amino acids and small peptides

Ionization efficiencies of 21 amino acids were measured in ESI positive mode (see Appendix 3). The range of the logIE scale for amino acids was approximately two orders of magnitude from logIE values 1.24 to 3.15. The ESI/MS response of amino acids has been also previously studied by Chalcraft et al.[56] and by Hermans et al.[63] For the common amino acids our results are in good correlation with the findings of Chalcraft et al. ($R^2 = 0.50$, see Figure 5) and Hermans et al. ($R^2 = 0.85$, see Figure 5). Compounds deviating the most from the ideal correlation line are Val and Ala, which are some of the most hydrophilic and smallest amino acids in both studies. Chalcraft et al. also found that hydrophobicity and molecular volume are influencing the relative response factors. Hermans et al. found a correlation between response factors of amino acids and molecular volume.
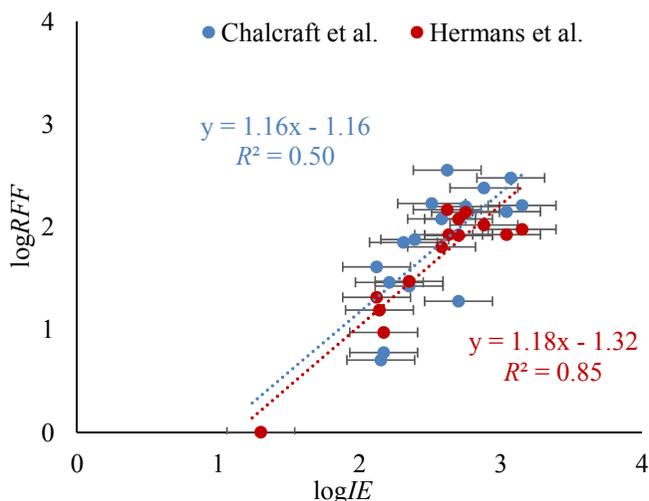
Figure 5. Correlation between logIE values of amino acids and logRFF values.[56,63]

Previously developed model [Paper II] was used to predict the ionization efficiency values of the amino acids and to confirm its applicability to amino acids (Figure 6 and Appendix 3). A good prediction power for predicting logIE values (root-mean-square error, s_{RMSE} = 3.5 times) was observed. Using physicochemical parameters corresponding to the non-zwitterionic conformers showed a significantly better fit to the experimental data than using parameters for zwitterionic conformer. Though zwitterionic conformers are most abundant in the solvent phase, it is expected that the relevant conformers in analyte-water clusters and in gas-phase are different (see Figure A3-1).[117–120]
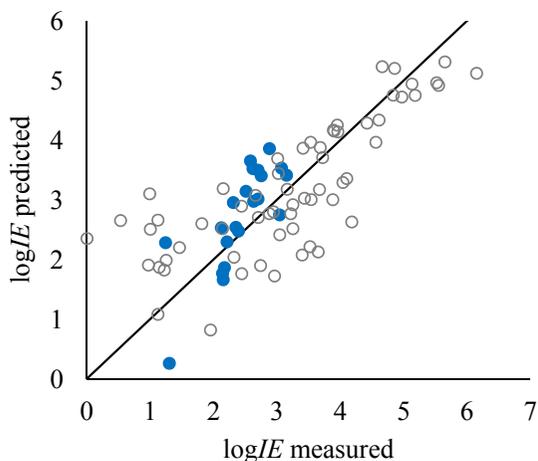
Figure 6. Correlation of measured and predicted log*IE* values of amino acids (in blue) calculated by the prediction model from Eq. 22 (log*IE* = (1.39 ± 0.29)·α + (0.36 ± 0.05)·*WANS* + (4.01 ± 0.27), [Paper II]); black line depicts the ideal correlation (slope = 1, intercept = 0).

Ionization efficiencies of 38 oligopeptides with varying length of amino acid residues ($n = 2 - 14$) were measured. The log*IE* values ranged over 4 orders of magnitude (log*IE* values from 1.93 to 5.61) and are shown in Table 3. Oligopeptides were divided into 7 groups with common amino acid chain. Several oligopeptides also formed multiply charged ions in ESI/MS (data shown in Table A3-3, Appendix 3). The extent of multiple charging and the intensity ratios of multiply charged ions varies between instruments (Figure 7). It was expected that the charge state distribution depends on the instrumental conditions used[23] and, therefore, log*IE* values for all charge states were summed together:

$$\sum \log IE = \log\left(\sum_{z=1}^{z_{max}} 10^{\log IE_z} \right) \qquad (18)$$

The values are presented in Table 3. The summed log*IE* values of oligopeptides (Σ log*IE*) had an acceptable correlation between different instruments and ESI sources ($R^2 = 0.44$) whereas using, for example, only singly charged ions yielded in no correlation ($R^2 = 0.001$).
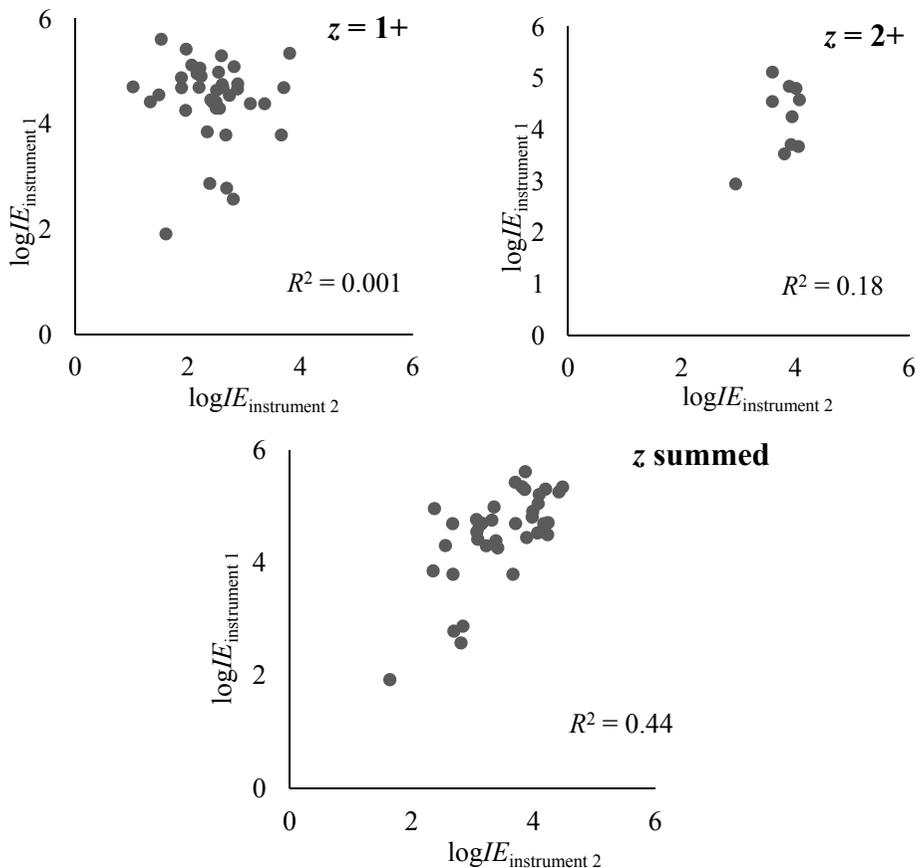
Figure 7. Correlation of log*IE* values obtained by measurements carried out with different instrumental setups for different charge states: singly charged ($z = +1$), doubly charged ($z = +2$) oligopeptides and for log*IE* values obtained by summing up all charge states (z summed). Instrument one is an ion trap instrument, instrument 1 is a triple quadrupole MS. The results of Instrument 1 are used in the discussion and in Table 3.

For predicting ionization efficiencies (see below), our hypothesis is that the log*IE* of the peptide is proportional to the sum of the measured ionization efficiencies of all of the amino acids (values Table A3-2, Appendix 3) that the oligopeptide with the length of *n* amino acid residues consists of:

$$\log IE_{peptide} \sim \log IE_{sum} = \log\left(\sum_{1}^{n} 10^{\log IE_{amino\ acid\ n}}\right) \qquad (19)$$

To test this hypothesis the correlation between the measured log*IE*$_{peptide}$ and the sum of the ionization efficiencies of amino acids log*IE*$_{sum}$ can be used. Here, this approach results in a good correlation between measured and predicted log*IE* values of oligopeptides ($R^2 = 0.70$).

Table 3. The log*IE* values of oligopeptides organized in groups with the same or similar amino acid backbone. Amino acids with high log*IE* values (log*IE* > 3; Leu, Phe, **His**) are marked in green, with low log*IE* values (log*IE* < 2; Gly, Cys) marked in red, and basic amino acids are marked in bold (**Arg**, **Lys, His**).

|		Σ log*IE*	log*IE*predicted	Group
1	Gln-Phe-Phe-Gly-Leu-Met-NH₂	5.34	3.66	A
2	**Lys**-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH₂	4.91	3.74	A
3	Gln-Gln-Phe-Phe-Gly-Leu-Met-NH₂	4.76	3.68	A
4	**Arg**-Pro-**Lys**-Pro-Gln-Gln-Phe-Phe-Gly	4.69	3.69	A
5	Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH₂	4.65	3.71	A
6	**Arg**-Pro-**Lys**-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met	4.55	3.83	A
7	Phe-Phe-Gly-Leu-Met-NH₂	4.38	3.63	A
8	Asp-**Arg**-Val-Tyr-Ile-**His**-Pro-Phe-**His**-Leu	5.34	3.70	B
9	Asp-**Arg**-Val-Tyr-Ile-**His**-Pro-Phe	5.30	3.70	B
10	Asn-**Arg**-Val-Tyr-Ile-**His**-Pro-Phe	5.25	3.70	B
11	Val-Tyr-Ile-**His**-Pro-Phe	5.20	3.61	B
12	Tyr-Ile-**His**-Pro-Phe	4.80	3.55	B
13	Asp-**Arg**-Val-Tyr-Ile-**His**-Pro	4.49	3.58	B
14	Ile-**His**-Pro-Phe	4.44	3.51	B
15	**Lys**-**Arg**-Pro-Pro-Gly-Phe-Ser-Pro-Leu	5.42	3.71	C
16	**Arg**-Pro-Pro-Gly-Phe-Ser-Pro-Leu	5.30	3.68	C
17	**Arg**-Pro-Gly-Phe-Ser-Pro-Phe-**Arg**	5.04	3.57	C
18	**Arg**-Pro-Pro-Gly-Phe-Ser-Pro	4.98	3.53	C
19	**Arg**-Pro-Pro-Gly-Phe-Ser	4.75	3.47	C
20	**Arg**-Pro-Pro-Gly-Phe	4.30	3.45	C
21	**Arg**-Pro-Pro	2.87	3.21	C
22	Thr-**Arg**-Ser-Ala-Trp-NH₂	4.54	3.23	D
23	Thr-**Arg**-Ser-Ala-Trp	4.41	3.23	D
24	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-**Arg**	4.70	3.49	E
25	Cys-Tyr-Phe-Gln-Asn-Cys	4.69	3.28	E
26	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-**Arg**-Gly-NH₂	4.26	3.49	E
27	Trp-Ala-Gly-Gly-Asn-Ala-Ser-Gly-Glu	4.96	3.10	F
28	Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu	4.30	3.11	F
29	Gly-Gly-Asp-Ala	2.78	2.54	F
30	Gly-Gly-Gly-Phe-Phe-NH₂	3.85	3.47	G
31	Gly-Gly-Gly-NH₂	1.93	1.78	G
32	Gly-**Lys**-Pro-Ile-Pro-Asn-Pro-Leu-Leu-Gly-Leu-Asp-Ser-Thr	5.61	3.85	-
33	**Arg**-**Arg**-Leu-Ile-Glu-Asp-Ala-Glu-Tyr-Ala-Ala-**Arg**-Gly	4.70	3.75	-
34	Phe-Phe-Phe-Phe	4.69	3.67	-
35	**Arg**-**Arg**-Pro-Tyr-Ile-Leu	4.52	3.75	-
36	Ac-Gly-**Lys**-OMe	3.79	2.60	-
37	Gly-Pro-Gly-Gly	3.79	2.69	-
38	Gly-β-Ala-β-Ala	2.57	2.48	-

3.1.2.1 Basic amino acids with higher hydrophobicity and larger volume have higher ionization efficiencies

It was observed that amino acids with higher hydrophobicity and larger volume (Leu, Phe, His) have higher ionization efficiency values. The same effect has been previously observed for other small molecules.[13,45] These compounds are probably closer to the surface in ESI droplets and it is easier for this type of compounds to enter the gas phase. Also, more basic amino acids (His, Arg, Lys) had higher log*IE* values. Basic amino acids are expected to have a net positive charge in the solution phase and are more easily ionized.[34] Acidic amino acids (Asp, Glu), as well as smaller and more hydrophilic amino acids (Gly, Cys, Ala), had lower ionization efficiencies. Small and hydrophilic amino acids are most probably situated in the interior of ESI droplets and the transition from the solvent phase to the gas phase is, therefore, more difficult. The significant similarities to previously observed trends [Paper II] allowed suggesting that ionization efficiency can be predicted for the amino acids with a similar model. In order to test this, the log*IE* values of all of the amino acids were predicted using a previously developed model [Paper II]. The correlation between the predicted and measured values was high (for amino acids $R^2 = 0.61$). The prediction error for ionization efficiencies of amino acids was on average $s_{RMSE} = 3.6$ times which is a great improvement compared to areas assuming equal ionization efficiencies and using peak areas for quantitation (s_{RMSE} up to 82 times for amino acids). The correlation graph is given in Figure 6, blue dots represent amino acids and grey points represent the data used for model development. This is very encouraging, as the training set consisted mostly of compounds with only one or two basic centers which are not zwitterionic in the solution phase.

3.1.2.2 Longer oligopeptides have higher ionization efficiencies

In general, oligopeptides, in general, had higher ionization efficiencies than amino acids and similar trends between IE and influencing parameters were observed. The analyzed oligopeptides were divided into seven groups, with a repeating amino acid residue backbone (see Table 3). Longer and bulkier oligopeptides and oligopeptides consisting of more hydrophobic amino acids had higher ionization efficiencies. Oligopeptides composing of smaller and hydrophilic amino acids tended to have lower ionization efficiencies. The longest oligopeptide had highest ionization efficiency (log*IE* = 5.61) and one of the smallest and most hydrophilic oligopeptides (Gly-Gly-Gly-NH$_2$) had the lowest ionization efficiency (log*IE* = 1.92). This means that the sensitivity difference between the best and the worst ionizing compound is ca 5000 times. It seems that from a length of 5-6 amino acids onwards the ionization efficiency values level off: the increase in oligopeptide length does not increase ionization efficiency significantly (see Figure 8). This plateau effect is expected when the surface activity of the peptide is high enough to cause almost all of the peptide

molecules to reside in the droplet's surface. This effect indicates that as the length of the peptide increases the effect of each added amino acids diminishes until it can be presumed that the log*IE* values are equal for longer oligopeptides.
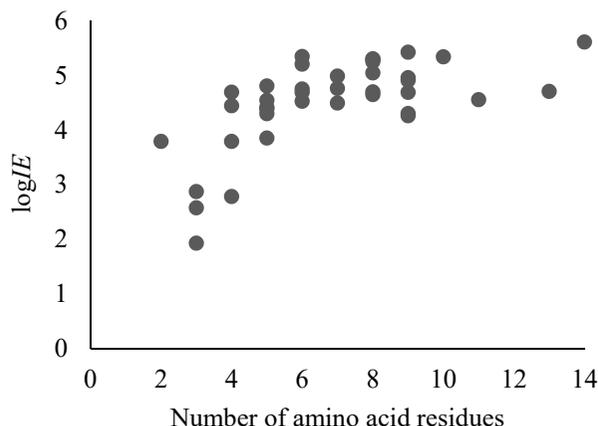
Figure 8. Ionization efficiencies of oligopeptides versus the number of amino acid residues they consist of.

3.1.2.3 The amino acid composition is more important than length

In addition to the number of amino acid residues incorporated, the nature of the oligopeptide, e.g., its hydrophobicity, presence of charged side chains, amino acid residues with high log*IE* values, are influencing the ionization efficiency of the peptide. For example, the oligopeptide with the smallest number of amino acid residues (two residues) in the current set was Ac-Gly-Lys-OMe (log*IE* = 3.79) which has significantly higher ionization efficiency than oligopeptide consisting of three Gly residues (log*IE* = 1.92). This is likely the case as Ac-Gly-Lys-OMe contains Lys residue which is charged at current pH and is significantly bulkier than Gly. Similarly, Gly-βAla-βAla (log*IE* = 2.57) has higher ionization efficiency than Gly-Gly-Gly-NH$_2$ probably due to two extra methyl groups in the side chains of both β-Ala residues (ionization efficiency is 4.5 times higher). The difference, however, is statistically insignificant. Additionally, oligopeptide Phe-Phe-Phe-Phe (log*IE* = 4.69) is quite short in comparison with other oligopeptides. However, as it consists of phenylalanine, which has one of the highest ionization efficiencies, this oligopeptide has a rather high ionization efficiency.

Oligopeptides Asp-Arg-Val-Tyr-Ile-His-Pro-Phe (log*IE* = 5.30) and Asp-Arg-Val-Tyr-Ile-His-Pro (log*IE* = 4.49) differ only by one additional Phe residue. Phe has high ionization efficiency and thus also increases the ionization efficiency of the oligopeptide (0.81 Δlog*IE* units). A similar effect of the importance of amino acid in the oligopeptide with high ionization efficiency occurs in the group C where oligopeptide Arg-Pro-Pro (log*IE* = 2.87) has 30

times lower ionization efficiency than Arg-Pro-Pro-Gly-Phe (log*IE* = 4.30). Also, in group E it can be seen that longer backbone does not necessarily mean higher log*IE* value (Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg, log*IE* = 4.70 vs Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH$_2$, log*IE* = 4.26). Based on this example, it can be proposed that the accessibility (low steric hindrance) of charged amino acid Arg is more beneficial than having a longer oligopeptide with less accessible charge center.

3.1.2.4 Carboxylic vs amide group

From group D it can be concluded that amidification of the C-terminus does not significantly influence the ionization efficiency. For these oligopeptides, there is a clear basicity center (Arg) and the difference between pK_a values and hydrophilicity for carboxylic acid and amide group is insufficient for having a significant influence. In addition, group B allows estimating the influence of Asp (polar acidic side chain) vs Asn (polar uncharged side chain) for oligopeptides Asp-Arg-Val-Tyr-Ile-His-Pro-Phe (log*IE* = 5.30) and Asn-Arg-Val-Tyr-Ile-His-Pro-Phe (log*IE* = 5.25). Again, there is no significant difference, probably due to the fact that both Asp and Asn are uncharged at this pH (in 0.1% formic acid, pH = 2.68). Both peptides already contain Arg residue which is charged; Arg also has high ionization efficiency. The same comparison can also be done in group F with oligopeptides Trp-Ala-Gly-Gly-Asn-Ala-Ser-Gly-Glu (log*IE* = 4.96) and Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu (log*IE* = 4.30) where the difference between Asp and Asn is slightly larger, probably due to the fact that neither oligopeptide contains charged side chains (no basic charge centers) nor do they contain amino acids with high ionization efficiencies.

3.1.2.5 Predicting ionization efficiencies of oligopeptides based on logIE values of amino acids

The similarities observed for the ionization behavior of amino acids and oligopeptides suggest that similar physicochemical parameters are significant for describing their ionization efficiencies. However, calculating physico-chemical parameters for oligopeptides is significantly more time consuming as both the molecular size and conformational space increase considerably with each added amino acid residue. Therefore, it was tested whether the ionization efficiency of an oligopeptide is correlated to the ionization efficiencies of the amino acids it consists of. As described above, the effect of modified (amidified) C-terminus is insignificant; therefore, there was no distinction between peptides with or without modified (amidified) C-terminus. Also, a single equation (Eq. 18) was applied to all oligopeptides independent of the number of charge states formed.

The measured log*IE* values and the sum of log*IE* values of amino acids were in a good correlation with the measured values (R^2 = 0.70, Figure 9). The way the summarizing (Eq. 18) takes into account that each following amino acid, independent on its position, has less effect on the overall ionization efficiency

of the oligopeptide. Still, the slope of the correlation graph is different from one (0.46 ± 0.05). This is expected, as though amino acids influence the ionization process of the oligopeptide the structure of the units changes as a result of amide bond formation and simple addition does not account for this. The slope and intercept values are expected to depend on the solvent and instrument used; therefore, to apply the summing up approach the slope and intercept of the correlation graph need to be applied. By doing so it is possible to estimate the logIE of an oligopeptide from its amino acid composition in our specific solvent-instrument system with an average error of 3.3 times.
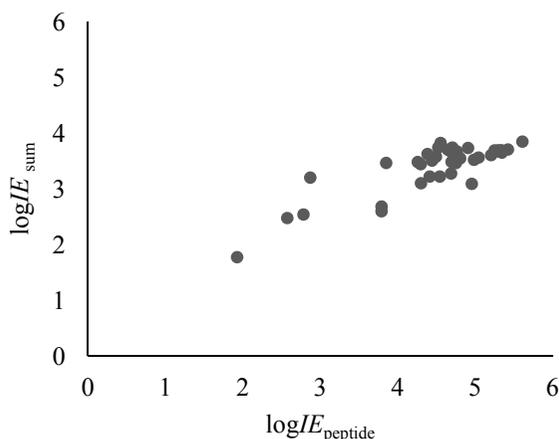
Figure 9. Correlation between measured and predicted ionization efficiencies of oligopeptides ($R^2 = 0.70$). Slope = 0.46 ± 0.05 and intercept = 1.33 ± 0.23; slope and intercept can be used to normalize predicted ionization efficiency values to a measurement setup.

The order of the amino acids may play a role in charge state distribution of the ions formed during the ionization process. As the order of the amino acids is expected to affect the sum of all ion formed less, in the abovementioned model the signals of all the charge states are summed.

However, as pointed out before (Figure 8) there is a leveling off of ionization efficiency values which means that from about a length of 5–6 amino acid residues the addition of another amino acid residue influences the ionization efficiency of an oligopeptide very little. It also means that from a certain length onwards the ionization efficiencies can be assumed to be equal.

3.2 Predicting ionization efficiencies in biological matrices

Based on the previously obtained promising results for ESI negative mode in various solvents[52,121] and instrumental setups[122] it was aimed to go one step further by predicting the ionization efficiencies for analysis in biological matrices. Therefore, the aim was to study whether ionization efficiencies in ESI

negative mode can be predicted in biological matrices (plasma, urine, whole blood, cerebrospinal fluid (CSF), liver and brain tissue homogenates). For this purpose, ionization efficiency values of 10 compounds, predominantly pharmaceuticals, were measured in different biological matrices with flow injection analyses. The ionization efficiency model was fitted in each matrix as well as in neat solvent (80/20 MeCN/0.1% ammonia) to compare with. The worst-case scenario, a simple protein precipitation sample preparation without any chromatographic separation of the analyte and matrix compounds is used, as a proof of concept that ionization efficiencies can be predicted under severe matrix effect conditions. The method is cross-validated by the 'leave-one-out' validation method.

The span of the logIE values (Table 4) measured within one biological matrix varied from 2.40 logIE units in the brain to 4.47 logIE units in blood. The narrowest span was observed in the neat solvent (ca 1.67 logIE units). The difference in spans demonstrates the extent of compressing or expanding the logIE scales by the matrix compounds.

The variation of logIE values of compounds between different matrices was significant: from 0.98 (warfarin, logIE_{solv} - logIE_{urine}, 1.07 - 0.09) to 3.09 (fumaric acid, logIE_{solv} - logIE_{blood}, -0.60 - (-3.69)) logIE units. This variation demonstrates that ionization efficiencies are considerably influenced by the matrix components. For example, in blood samples, the signal of warfarin is suppressed by 89.5% relative to the neat solvent. Additionally, it is observed that the variation in logIE values from one matrix to another was lower for compounds with higher ionization efficiencies (e.g. warfarin and taurocholic acid versus fumaric acid and sorbic acid).

In the literature,[123,124] it has been shown that matrix effect may vary with analyte concentration. The concentration dependence of matrix effect is a very delicate question and it may depend on the way matrix effect is calculated. Namely, differences are observed if calibration graph slopes or peak areas are used.[123] This is especially important if the matrix alters the linear range of the method or intercept values.[125,126] Here all measurements are carried out in the linear range and calibration graph slopes are used to calculate matrix effect. If the measurements are in the linear range, it does not matter which specific concentration is used, as the calculated slope is independent of the con-centrations in this range.

The order of the logIE values of the compounds remained roughly the same from one matrix to another. This means that compounds with higher logIE values in the neat solvent are also ionized better in the presence of matrix components. The same was confirmed by the correlation studies (see Figure 10). The logIE values in different biological matrices were in good correlation with the logIE values in the neat solvent; see Table 4. The highest correlation was observed between logIE values measured in urine and brain tissue homogenate with logIE values measured in the neat solvent ($R^2 = 0.87$ for both). The lowest correlation was observed between logIE values measured in CSF extract and logIE values measured in the neat solvent ($R^2 = 0.67$). The

correlation graphs are presented in Figure 10. Additionally, the IE values measured in the neat solvent are in good correlation with the previous measurements[121] carried out on a different instrument ($R^2 = 0.95$).

$$\log IE_{urine} = 1.80 \log IE_{solv} - 2.00$$
$$R^2 = 0.87$$

$$\log IE_{plasma} = 2.15 \log IE_{solv} - 1.74$$
$$R^2 = 0.82$$

$$\log IE_{blood} = 2.10 \log IE_{solv} - 1.75$$
$$R^2 = 0.71$$

$$\log IE_{CSF} = 1.42 \log IE_{solv} - 1.50$$
$$R^2 = 0.67$$

$$\log IE_{liver} = 1.40 \log IE_{solv} - 1.01$$
$$R^2 = 0.86$$

$$\log IE_{brain} = 1.37 \log IE_{solv} - 0.80$$
$$R^2 = 0.87$$
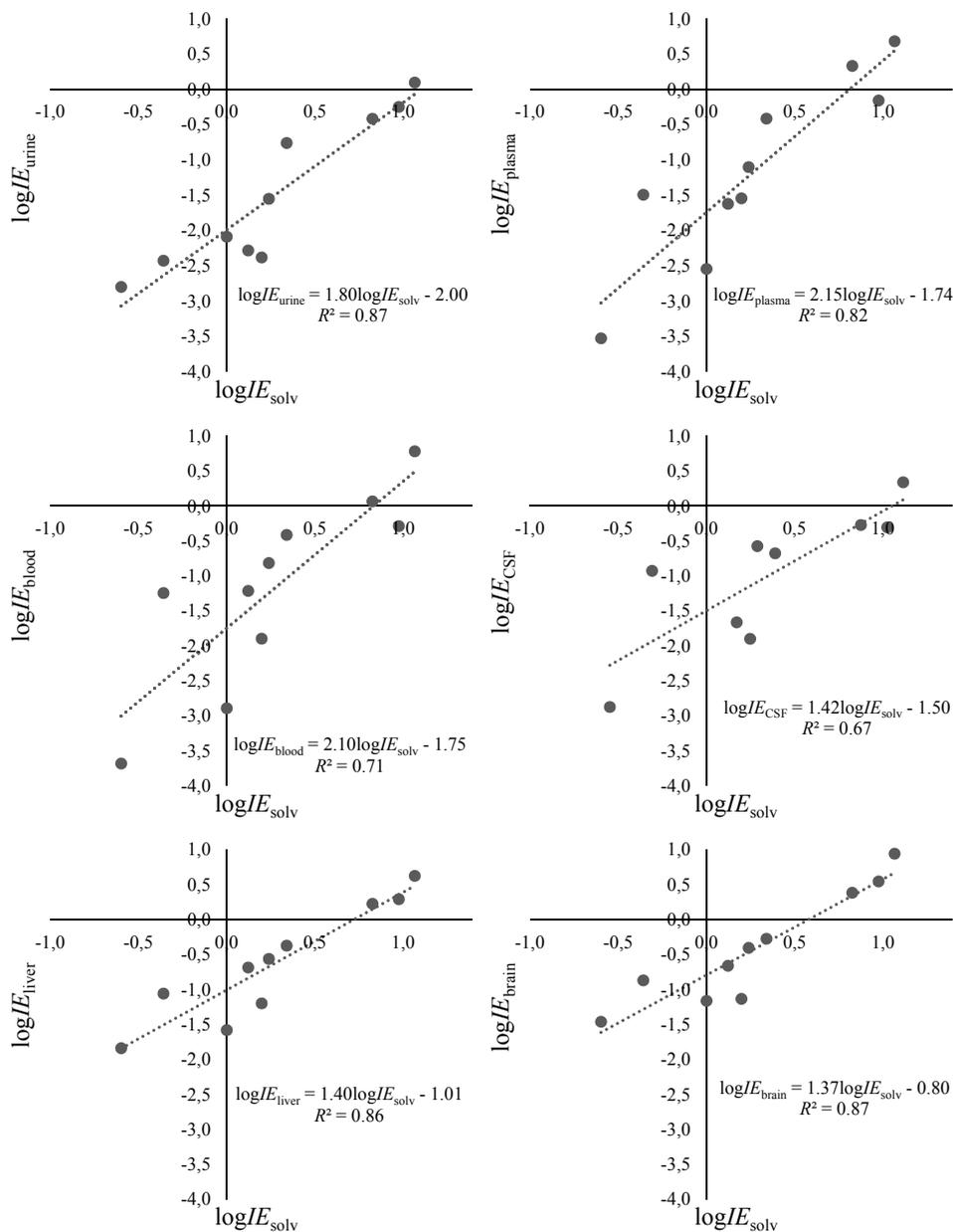
Figure 10. Correlations of logIE measurements in different biological matrices with measurements in 80/20 acetonitrile/0.1% ammonia aqueous solution ($\log IE_{solv}$).

Table 4. The ionization efficiencies ($\log IE$) measured in various matrices and neat solvent ($\log IE_{solv}$). $\log IE_{solv}{}^{*}$ values refer to the ionization efficiency values from ref.[121] that have been measured on a different instrument and in 20/80 0.1% ammonia/acetonitrile mobile phase. The reproducibility of measurements was calculated as a consistency standard deviation – $s_{consistency}$. NA^{a} – values not possible to measure; NA^{b} – compounds not measured before in the same solvent.

	$\log IE_{urine}$	$\log IE_{plasma}$	$\log IE_{blood}$	$\log IE_{CSF}$	$\log IE_{liver}$	$\log IE_{brain}$	$\log IE_{solv}$	$\log IE_{solv}{}^{*}$
warfarin	0.09	0.68	0.78	0.34	0.62	0.94	1.07	NA^{b}
taurocholic acid	-0.25	-0.16	-0.29	-0.31	0.29	0.54	0.97	NA^{b}
3-CF$_3$SO$_2$-benzoic acid	-0.42	0.33	0.06	-0.28	0.22	0.38	0.83	1.69
salicylic acid	-0.76	-0.42	-0.42	-0.68	-0.38	-0.28	0.34	0.39
dodecanoic acid	-1.55	-1.10	-0.82	-0.58	-0.57	-0.41	0.24	NA^{b}
benzoic acid	-2.09	-2.55	-2.90	NA^{a}	-1.58	-1.17	0.00	0.00
naproxen	-2.28	-1.62	-1.22	-1.67	-0.69	-0.66	0.12	NA^{b}
lincomycin	-2.38	-1.55	-1.90	-1.90	-1.20	-1.14	0.20	NA^{b}
sorbic acid	-2.43	-1.50	-1.25	-0.93	-1.06	-0.87	-0.36	-0.40
fumaric acid	-2.80	-3.53	-3.69	-2.88	-1.84	-1.46	-0.60	NA^{b}
$s_{consistency}$	0.33	0.15	0.25	0.14	0.10	0.19	0.12	
R^2 with neat solvent	0.87	0.82	0.71	0.67	0.86	1.37	-	0.95
Slope	1.80	2.15	2.10	1.42	1.40	1.37	-	1.75
Intercept	-2.00	-1.74	-1.75	-1.50	-1.01	-0.80	-	0.07

For all correlations, the intercept values were negative; this pinpoints that biological matrices suppress ionization for the studied compounds. For all biological matrices, the correlation graph slopes were significantly above 1. These two findings show that in general the signal of compounds with lower ionization efficiencies is suppressed more than the signal of compounds with higher ionization efficiencies. This is well in line with the surface excess charge model proposed by C.G. Enke.[24] According to this model, the ionization efficiency of a compound depends both on the compounds affinity towards droplet surface charge and on the co-eluting compounds affinity towards droplets surface charge. Compounds with lower affinity have lower ionization efficiencies, and additionally, are more easily outcompeted from the surface of the droplets by the matrix compounds.

Based on the correlation graphs the most complicated matrices were blood, plasma, and urine. For these matrices, the intercepts were the lowest and slopes the highest. Blood and plasma are known to cause severe ionization suppression even after protein precipitation due to the omnipresence of phospholipids.[127] Additionally, urine samples are known to have a high salt concentration which is not completely removed by the sample preparation. For example, Dams et al. [128] have observed ionization suppression of 85% even after using protein precipitation with acetonitrile as a sample preparation method. High salt concentrations are known to cause severe ionization suppression[129] due to analyte precipitation in ESI.[130]

A good correlation between $\log IE$ values measured in matrices and in the neat solvent hints that ionization efficiencies can be predicted in the matrices similarly to the already published predictions in the neat solvent.[13,45] In order to test this further, different physicochemical parameters were used for modeling. Previously,[45,121] it has been shown that $\log IE$ values in the neat solvent are best described by charge delocalization parameter, *WAPS,* and degree of ionization, *α. WAPS* values also had the highest correlation with $\log IE$ values measured in biological matrices (see calculated parameters in Appendix 4). These parameters were also used to fit the multilinear models for predicting the $\log IE$ values measured in biological matrices. The obtained models have the general form as Eq. 20 and the respective constants are described in Table 5. The obtained models possess good predictive power; the R^2 values ranged from 0.55 (urine) to 0.81 (liver). The obtained fits are graphically shown in Figure 11 (each color represents one matrix).

$$\log IE = coef_{WAPS} \cdot WAPS + coef_\alpha \cdot \alpha + intercept \qquad (20)$$

Table 5. Coefficients for predicting log*IE* values in different matrices in Eq. 20.

	coef_{WAPS}	coef_{α}	intercept	R^2	s_{RMSE} (times)	s_{RMSE} from LOO validation (times)	p Goodness-of-Fit
Solvent	-0.18 ± 0.05	1.18 ± 0.26		0.72	2.24	2.29	0.997
Urine	-0.31 ± 0.12	2.04 ± 1.01	-1.93 ± 0.93	0.55	6.31	6.79	0.988
Plasma	-0.45 ± 0.13	1.08 ± 0.68		0.77	7,94	10.47	0.993
Blood	-0.50 ± 0.13	1.29 ± 0.69		0.78	8,13	20.41	0.996
Liver	-0.32 ± 0.07	0.96 ± 0.37		0.81	3.09	6.61	0.999
Brain	-0.30 ± 0.07	1.08 ± 0.38		0.73	3.16	3.31	0.997
CSF	-0.34 ± 0.12	0.66 ± 0.63		0.71	6.76	15.49	0.977

The coefficients for *WAPS* in the model fitted for log*IE* values in urine, liver, blood and brain matrix are very similar and only in urine matrix, the intercept value became statistically significant. This can most likely be attributed to the relatively high salt content in urine as compared to the other matrices. Obviously, the salts have a much larger effect on the ion suppression than either the lipids, bile acids or proteins remaining after sample preparation in other matrices.
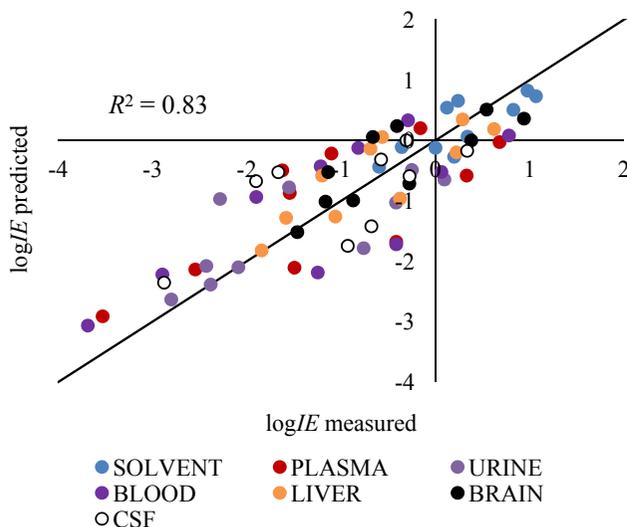
Figure 11. Correlation of all the measured log*IE* values and predicted log*IE* values in different matrices. Each dot represents one compound in one matrix, different colors indicate different matrices.

The accuracy of the models can also be described with the root-mean-square error of the models from the LOO validation, s_{RMSE} = 7.2 times. This value shows that on average the mismatch between the predicted and measured ionization efficiencies is lower than 8.3 times. Until now, in the absence of authentic standards, equal ionization efficiencies are assumed in all matrices. For example, if the ionization efficiencies for all compounds used in this study are assumed to be equal to the ionization efficiency of benzoic acid and peak areas are used to describe the abundance of the compounds present in the sample it would lead to an average error of 660 times. This means, that the proposed approach improves predicting ionization efficiency by almost two orders of magnitude.

Moreover, all experiments in this study were carried out in flow injection mode without any chromatographic separation. Therefore, the ionization efficiencies of all of the studied compounds are affected by all of the matrix compounds present after sample preparation. In the case of chromatographic

separation, all of the analyzed compounds would co-elute only with a fraction of matrix compounds and it is commonly expected that the matrix effect would significantly decrease. Nevertheless, each analyte would co-elute with different matrix compounds and, therefore, much more complicated effect on the ionization efficiencies could occur. The results obtained with flow injection analyses serve as a good starting point for developing a universal approach that would be compatible both with liquid chromatography and flow injection metabolomics. Additionally, as the push towards high throughput is ever increasing and this is driving metabolomics studies also towards flow injection analyses,[5,131] the current approach already is applicable.

3.3 Combining ESI polarities

Our group has extensively studied both ESI+ and ESI− mode and thousands of logIE values have been measured in both ESI modes. However, these are still two separate IE scales and comparison of values measured in can only be qualitative (order of the compounds, span). The quantitative comparison would be beneficial to better compare the two modes, but even more importantly it would allow the user to choose the best ionization mode for a specific compound or a set of compounds. In order to combine the two ionization modes, it is important to find a compound ionizing to a similar extent in both ESI modes (see Figure 12 for the outline of the process).
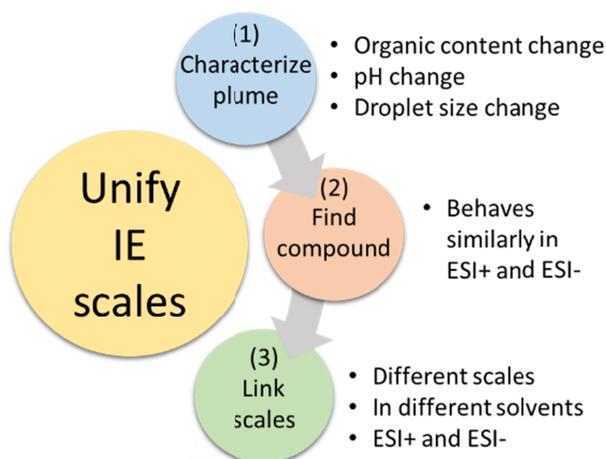
Figure 12. Process of unifying the different IE scales measured in different solvents and in ESI+ and ESI−.

3.3.1 Comparison of ESI+ and ESI− plume: pH, organic phase percentage, and droplet size

For the comparison, it is important to first quantify the changes occurring in the ESI plume: organic modifier percentage, pH as well as droplet size change of the "parent" ESI droplets with the diameter range of a few micrometers. For quantitative comparison of pH changes, the absolute pH is used which enables comparing different solvents and solvent compositions. Even more so, it will be estimated, whether or not the ESI polarity affects the magnitude of mobile phase properties' change in the ESI plume. This enables for the first time to compare the processes occurring in ESI source and aids in better understanding of the differences and similarities of the two modes. This is novel in three aspects: (1) the most acidic mobile phases are studied, (2) exact same set of ESI source parameters for both ESI+ and ESI− are used thereby enabling accurate and direct comparison between the two modes, (3) absolute pH is used, enabling to accurately compare mobile phases with different organic phase content and different aqueous pH.

Previous studies have measured either the aqueous phase pH or the pH of the mobile phase (a mixture of water phase and organic solvent) using calibration in the water phase. While this approach estimates the pH wit acceptable accuracy to model the ionization processes, it is expected to be not enough for comparison between ESI modes. Depending on the organic modifier, it can significantly influence the pH of the solvent mixture indicating that the conclusions drawn from the aqueous phase pH might be over-simplified. The conventional approach is improved upon by for the first time profiling the pH in the spray according to the absolute pH scale.[132] It is expressed as $pH_{abs}^{H_2O}$ values,[133] which enables direct comparison of the pH between different solvents and water (see Appendix 2). The acidity of solvent composition was probed in the aqueous pH (denoted as pH, refers to the "conventional" pH calibrated and measured in water) range from 3.21 to 5.16 using pH-chromic fluorescent dye as the molecular probe.

The results of acetonitrile content and $pH_{abs}^{H_2O}$ change in the ESI plume for all studied mobile phases are presented in Table 6. In all cases, the acetonitrile content decrease across the plume is 5.3 to 7.5% (percentage point), as acetonitrile is more readily vaporized during the evaporation process of the droplets. These changes in acetonitrile content in the ESI plume have been taken into account for the determination of the pH. This means for every point in the pH calibration curve the corresponding acetonitrile content percentage was used. For all mobile phases, the ESI+ mode has a somewhat smaller change in acetonitrile content, however, the difference is statistically insignificant.

Table 6. Change of acetonitrile (MeCN) content and $pH_{abs}^{H_2O}$ in the ESI plume ($y = 3$ mm), $x_{max} = 15$ mm from the needle tip, except for 50% MeCN initial content, where $x_{max} = 13$ mm. Water phase pH is brought in brackets for comparison.

Initial MeCN content	80%		80%		50%	
Initial pH (aq)	4.00		5.03		5.03	
Initial $pH_{abs}^{H_2O}$	7.64		8.43		7.05	
ESI mode	ESI+	ESI−	ESI+	ESI−	ESI+	ESI−
ΔMeCN content % (percentage point)	5.6%	5.7%	5.3%	7.4%	7.0%	7.5%
$pH_{abs}^{H_2O}$(pH) at $x = 0$ mm	7.47 (3.78)	7.81 (4.27)	8.20 (4.80)	8.49 (5.14)	6.83 (4.82)	7.19 (5.16)
$pH_{abs}^{H_2O}$ (pH) at x_{max}	6.66 (3.21)	7.19 (3.91)	7.52 (4.28)	7.82 (4.73)	6.28 (4.54)	6.62 (4.88)
$\Delta pH_{abs}^{H_2O}$ (ΔpH) in ESI plume	0.81 (0.57)	0.62 (0.36)	0.68 (0.52)	0.67 (0.41)	0.55 (0.28)	0.57 (0.28)

The $pH_{abs}^{H_2O}$ at the ESI needle tip ($x = 0$ mm) is different (confirmed by t-test) from the $pH_{abs}^{H_2O}$ of the initial bulk solution (Table 6). The $pH_{abs}^{H_2O}$ difference of 0.06 to 0.23 pH units is observed between the bulk mobile phase and mobile phase in the needle tip. The $pH_{abs}^{H_2O}$ change along the plume is visualized in Figure 13b. It can be observed that the $pH_{abs}^{H_2O}$ of the solution decreases steadily along the ESI plume (i.e. becomes more acidic). For 50% acetonitrile content in the initial solution, the change in the pH is lower than for 80% initial acetonitrile content. The droplets for 50/50 acetonitrile/pH 5.00 water phase are more acidic according to $pH_{abs}^{H_2O}$ than for 80/20 acetonitrile/pH 4.00. This demonstrates that aqueous pH alone is not suitable for describing the acidity of the mobile phase. The largest difference of the $pH_{abs}^{H_2O}$ for ESI+ and ESI− down the plume evolution was observed for mobile phase containing 80/20 acetonitrile/pH 4.00 and the change was up to 0.81 pH units.
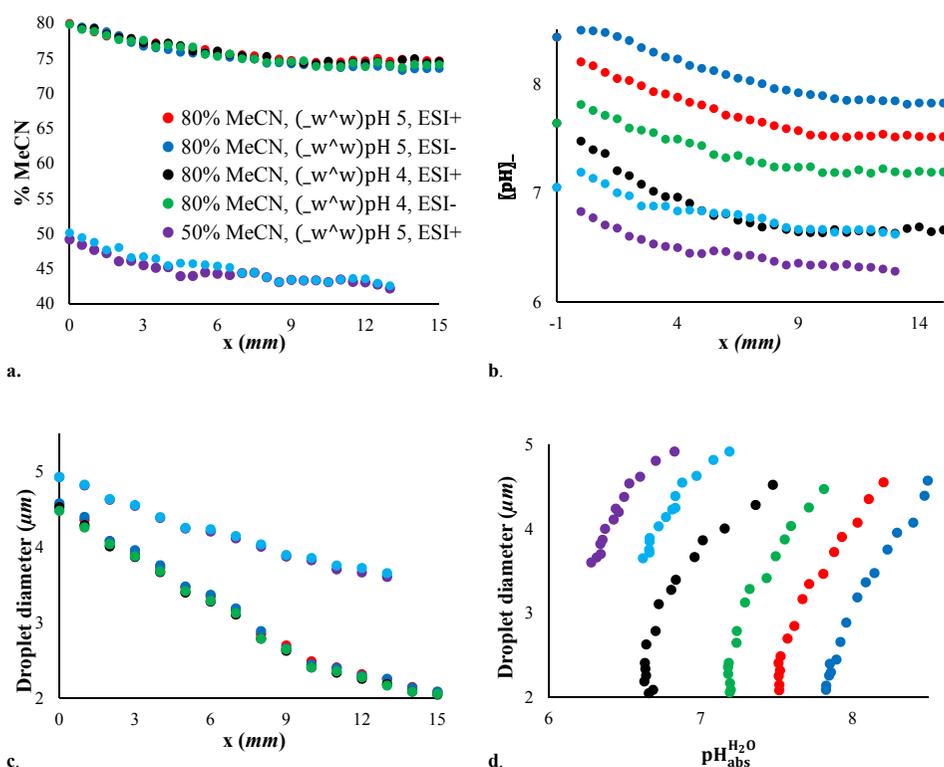
Figure 13. Change in (a) acetonitrile content (%MeCN) in the middle of the ESI plume ($y = 3$ mm), (b) $pH_{abs}^{H_2O}$ in the middle of the plume. pH in $x = -1$ mm is the $pH_{abs}^{H_2O}$ of the initial mobile phase, (c) average droplet diameter in the spray plume versus the axial distance x from the emitter tip (the average error is ±0.6 μm) (d) average droplet diameter versus the $pH_{abs}^{H_2O}$ in ESI plume. The legend is the same for a, b, c and d.

The droplet size was determined by Mie scattering measurements (by Dr. M. Girod in University Claude Bernard Lyon 1). The comparison of the projection of the signal intensity as a function of the scattering angle with the Mie theory allows determining the average droplet diameter at each point in the ESI plume, for the previously determined solvent compositions. The droplets shrink as they move downstream due to the evaporation process. Note that the reported droplet size is an average of the size of droplets within a pixel, which is an area of 500 µm x 500 µm. Moreover, the droplets constituting the ESI plume are poly-disperse. The droplet size distribution for each pixel is found to be Gaussian with a standard deviation σ ranging from 1% to 2.5%. The droplet diameter is decreasing along the plume from 4.5 µm to 2 µm for the acetonitrile/water 80/20 solvent mixture. For the 50/50 mixture, the initial droplet diameter is higher (4.9 µm) and the solvent evaporation is less efficient, leading to droplet diameter of 3.6 µm at $x = 13$ mm from the emitter tip. However, droplet size measurements demonstrate that droplet diameter change is independent of the ESI mode.

The droplets profiled in this study are the large parent droplets (in a micrometer scale) observed in the ESI plume. They are in different stages of development: (1) approaching Rayleigh limit,[25] (2) close to the Rayleigh limit and (3) already undergone Coulomb explosion and emitted a significant amount of excess charge via formation of smaller offspring droplets. Therefore, not all of the droplets are at the same stage and profiling gives us the average of these droplets.

The mobile phase becomes more acidic (compared to the initial acidity) at the needle tip ($x = 0$ mm) in the ESI+ mode, whereas in ESI− it becomes more basic. The acidities at this point are not yet influenced by the selective evaporation that occurs in the plume, as the measurements are made directly in the needle tip. This is also observable on Figure 13a, which shows the acetonitrile content along the plume starting from the needle tip. The change in $pH_{abs}^{H_2O}$ is caused by the electrochemical reactions taking place in the very end on the needle capillary. In ESI+ mode the electrochemical reaction creates an additional amount of protons from water or additives, while in ESI− hydroxyl ions are generated or protons are neutralized.[26–28,134,135]

The $pH_{abs}^{H_2O}$ change occurring between the bulk phase and droplets in the needle tip is significant. However the absolute value of $pH_{abs}^{H_2O}$ change is very similar for ESI+ (up to 0.23 units) and ESI− (up to 0.17 units). Therefore, the absolute differences in the changes of $pH_{abs}^{H_2O}$ occurring in the needle tip for ESI+ and ESI− are statistically insignificantly different. Based on this, it is impossible to decide over the efficiency of electrochemical reaction between ESI modes and mobile phases. Van Berkel et al.[134] have previously found and calculated that the change could be in the order of 2 or more pH units. This may result from the fact, that in their study pure water without any buffering agent was used.

It is worth mentioning though that the $pH_{abs}^{H_2O}$ differences are much smaller than those predicted by theoretical calculations.[42] A probable reason is that calculations can only be made based on the assumption of reached equilibria. However, it is important to note that it is not known whether the droplets formed in ESI are actually in their equilibrium. In previous studies it has been shown that some compounds may retain their liquid phase structures (charge location) due to kinetic trapping,[136] meaning that the drying of the droplets in ESI is too fast for the droplets to reach the actual equilibrium. Therefore, it is also likely that at least some of the droplets profiled in our work have not reached the chemical equilibria (including acid-base equilibrium).

For all studied mobile phases and in both modes it was observed that the solutions become more acidic along the plume. This is caused by more ready vaporization of acetonitrile and the resulting decrease of the content of acetonitrile in the droplets along the plume, causing an increase of water content. Three different factors operate here: (1) increase of formic acid acidity (decrease of pK_a) (the aqueous phase consists of 0.1% formic acid solution titrated with 25% ammonia solution to the desired pH), (2) increase of formic acid concentration in solution and (3) further increase of the share of the (more basic) water in the solvation sphere. While the last factor causes decrease of the solvated proton activity, the first two outweigh it, causing the droplet to become more acidic. The influence of solvent composition change on the basicity of ammonia is less important because (1) basicity of bases is less influenced by adding organic solvent[133] and (2) the used pH range is far from the pK_a of NH_4^+. The pH change in the plume is in general similar to that expected and also previously demonstrated.[28] Therefore indicating that tendencies found previously for neutral and basic mobile phases are also valid for acidic mobile phases.

The decrease of $pH_{abs}^{H_2O}$ along the plume is very similar in ESI+ and ESI− resulting in almost parallel evolution of lines within one mobile phase in Figure 13b. Therefore the $pH_{abs}^{H_2O}$ difference for ESI+ and ESI− generated by the electrochemical reaction in the needle tip remains essentially the same throughout the plume despite acetonitrile evaporation. Only in the case of the aqueous phase with a pH of 4.00, is the change in acidity greater in ESI+ than in ESI−, causing the $pH_{abs}^{H_2O}$ difference of 0.53 units at the maximum profiling length between different modes. If electrochemical reactions would be of the same efficiency for all mobile phases and the organic solvent evaporation rates would be indistinguishable, as observed from Figure 13a, it would be expected that more acidic mobile phase obeys less pH change than the less acidic (also keeping in mind, that the buffering capacity for this studied mobile phase is much higher).

Also, it was observed that $pH_{abs}^{H_2O}$ change is larger for the mobile phase containing 80% of acetonitrile than for mobile phase containing 50% of acetonitrile. The probable reasons could be: (1) the profiled length is longer

(15 mm versus 13 mm in the case of 50% acetonitrile) and (2) the change in the droplet size is larger (Figure 13c and d). The first possibility may be ruled out, as the plateau of the pH change has already been achieved at $x = 13$ mm. The droplet size change, however, is in almost linear relation with $pH_{abs}^{H_2O}$ (see Figure 13d). It is observed, that the droplet size changes more for mobile phase with 80% of acetonitrile than with 50% of acetonitrile. Acetonitrile evaporates more readily from solution with higher acetonitrile content:[30] both 50/50 and 80/20 mobile phases lose about 5% of acetonitrile; therefore, the total acetonitrile loss (and thereby the increase in water content leading to increased dissociation of formic acid) for the latter is higher. Solvent fractionation is less efficient when the initial water content is higher.

The method used allows us to profile the droplets down to 0.5 μm in diameter, therefore, indicating two mechanisms influencing the plume evolution. First, the drying of the droplets due to heated drying gas. In this process, the more volatile mobile phase component evaporates faster, but the changes are expected to be similar for the two ESI modes. Secondly, the fission of the charged nanodroplets from the parent droplets. This mechanism can have an effect on the composition of the remaining part of the parent droplet and is expected to influence ESI+ and ESI− mode differently. In ESI+ and ESI− ions of different type are removed from the parent droplets during ejection of nanodroplets, therefore, facilitating $pH_{abs}^{H_2O}$ change in different directions. Each fission of one nanodroplet removes 10 to 40% of the excess charge (depending on the solvent) from the parent droplet.[15,24,137,138] The general amount of excess charge can be estimated on the basis of pH change from bulk mobile phase to the mobile phase in the needle tip. Therefore, if the fission of nanodroplets would be a dominant process changing the composition of the parent droplets, it would be expected: (1) a significantly different $pH_{abs}^{H_2O}$ profiles for ESI+ and ESI− along the plume and (2) more significant changes in the droplets $pH_{abs}^{H_2O}$ along the plume.

These findings are important from two aspects. First, for modeling ESI process these findings demonstrate that desolvation due to drying dominates over the fission of nanodroplets in case of pneumatically assisted ESI sources. This means that similar models for ESI+ and ESI− mode can theoretically be constructed and this could also serve as a validation criterion for such models. Secondly, for practitioners, these findings show that qualitative findings related to the ESI process obtained on the ESI+ mode can almost directly be applied also in the ESI− mode. However, the importance of desolvation gas is huge and therefore quantitative findings are transferable only within one set of gas parameters between ESI+ and ESI−.

3.3.2 Linking ESI+ and ESI− IE scales

Linking ESI+ and ESI− is accomplished (as depicted in Figure 12) by (1) characterizing and comparing the ESI plume for ESI+ and ESI− (previous chapter, Paper III), (2) using a reference compound with both acidic and basic moieties with similar pK_a values, allowing ionization to similar extent in both modes from the same solution, and (3) measuring the links from ESI+ and ESI− anchoring compounds to the reference compound (in respective solvents) and unifying ESI+ and ESI− scales.

To make ESI+ and ESI− scales comparable, it is important to find a system where ionization efficiency value measured for a compound in ESI+ mode (log$IE^{\text{ESI+}}$) and ionization efficiency value measured for a compound in ESI− mode (log$IE^{\text{ESI-}}$) would become similar (steps (2) and (3)). It has previously been observed[45] that among other things, the ionization of compounds in the solvent and affinity towards droplet surface determines the extent of ionization in the ESI source. Therefore, it was necessary that basic and acidic groups of the reference compound would have similar ionization degrees in solution. Since the physicochemical changes in the ESI plume are similar in ESI+ and ESI−, the ionization in both modes occurs from fairly similar solutions. [Paper III] This significantly simplifies choosing a suitable solvent system for anchoring measurements.

Several compounds were tested and the closest degrees of ionization of the acidic and basic groups were observed for *trans*-3(3-pyridyl)acrylic acid (TA) in 80/20 acetonitrile/pH (aq) 4.00 (v/v) solution. The degree of ionization for the acidic group was 0.38 and for the basic group, it was 0.19. Since the pH in the plume region close to the MS entrance is somewhat lower (ca 0.8 for ESI+ and 0.1 units for ESI− [Paper III]) than the initial pH, the degrees of ionization are in fact even more similar and the ionization of TA in ESI+ and ESI− will also be similar. Here, focus is on the ionization of a specific group and not on the ionization of the compound as such, as some of its molecules may be as zwitterions in the solution. However, the simultaneous ionization of two groups is expected to affect ionization in both modes in the same way, as statistically the same fraction of molecules is present in the droplets as zwitterions in both modes.

The difference between anchor values ($\Delta_{\text{anchors}}^{\text{ESI+vs ESI-}}$) of ESI+ and ESI− can be derived (Figure 14):

$$\Delta_{\text{anchors}}^{\text{ESI+vs ESI-}} = \log IE_{Et_4N^+-MB}^{ESI+,0.1\%\,formic} + \log IE_{Et_4N^+}^{ESI+,0.1\%\,formic-pH4} \qquad (21)$$
$$+ \log IE_{TA-Et_4N^+}^{ESI+,pH4} - \log IE_{TA-BA}^{ESI-,pH4}$$
$$+ \log IE_{BA}^{ESI-,0.1\%ammonia-pH4}$$

where superscript denotes the ESI mode and media that the IE was measured in; subscript denotes the compounds that were measured.

The $\log IE_{\text{TA}-\text{Et}_4\text{N}^+}^{\text{ESI}+,\text{pH}\,4}$ of TA was measured to be -1.11 in ESI+ (anchor substance Et_4N^+) and $\log IE_{\text{TA}-\text{BA}}^{\text{ESI}-,\text{pH}\,4}$ was 0.73 in ESI− (anchor substance benzoic acid (BA)), both in 80/20 acetonitrile/pH 4.00 (v/v) solution. However, logIE scales in ESI+ and ESI− have been compiled in different solutions and to make the scales comparable, an anchor substance, BA in 80/20 acetonitrile/pH 4.00 (v/v) solution and 80/20 acetonitrile/0.1% ammonia (v/v) solution (ESI−), was infused. A difference ($\log IE_{\text{BA}}^{\text{ESI}-,0.1\%\text{ammonia}-\text{pH}\,4}$) of 0.90 logarithmic units was observed, meaning that the ionization efficiency of BA is 0.90 logarithmic units lower in 80/20 acetonitrile/pH 4.00. In ESI+ Et_4N^+ in 80/20 acetonitrile/0.1% formic acid (v/v) solution has an ionization efficiency of ($\log IE_{\text{Et}_4\text{N}^+-\text{MB}}^{\text{ESI}+,0.1\%\text{formic}}$) 3.95 relative to methyl benzoate.[13] It has been observed previously that the ionization efficiency of tetraethylammonium is not affected by changes in pH ($\log IE_{\text{Et}_4\text{N}^+}^{\text{ESI}+,0.1\%\text{formic}-\text{pH}\,4} = 0$).[34]

Substituting these values into Eq. 21 gives a difference of 3.95 – 1.11 – 0.73 + 0.90 = 3.01 logarithmic units between the IEs of anchors between ESI+ and ESI− (Figure 14). Knowing the difference in anchors allows direct comparison in logIE values between ESI+ and ESI− mode.

$$\log IE^{ESI+} = (1.39\pm0.23)\cdot\alpha + (0.36\pm0.05)\cdot WANS + (4.01\pm0.27) \quad (22)$$

$$\log IE^{ESI-} = (2.23\pm0.34)\cdot\alpha + (-0.51\pm0.04)\cdot WAPS + (1.04\pm0.34) \quad (23)$$
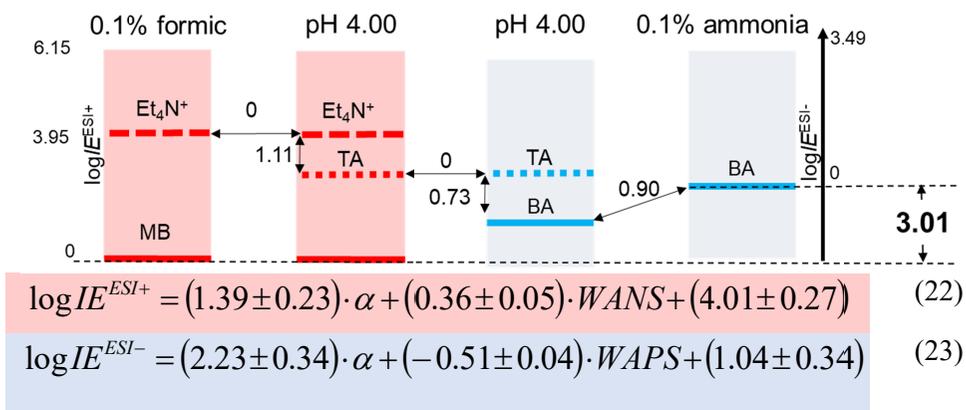
Figure 14. Comparison of logIE scales compiled in ESI+ (log$IE^{\text{ESI}+}$) and ESI− (log$IE^{\text{ESI}-}$).

To further verify this difference, the logIE prediction models were compared (Equations 22 and 23 in Figure 14). Previously developed a model (Eq. 23) for predicting IEs in ESI− (log$IE^{\text{ESI}-}$) was based on the degree of ionization and the charge delocalization in anions (*WAPS*).[45] In order to make the IE prediction models comparable, the model for logIE values on the ESI+ scale[13] was modified to contain a *WANS* parameter that describes charge delocalization in cations (see Figure 15 and Appendix 5,[139] and obtained Eq. 22 for finding

$\log IE^{\text{ESI+}}$. The ability to predict IEs is similar (s_{RMSE} = 7.4 times) to the previous version[13] of the model (s_{RMSE} = 7.2 times).
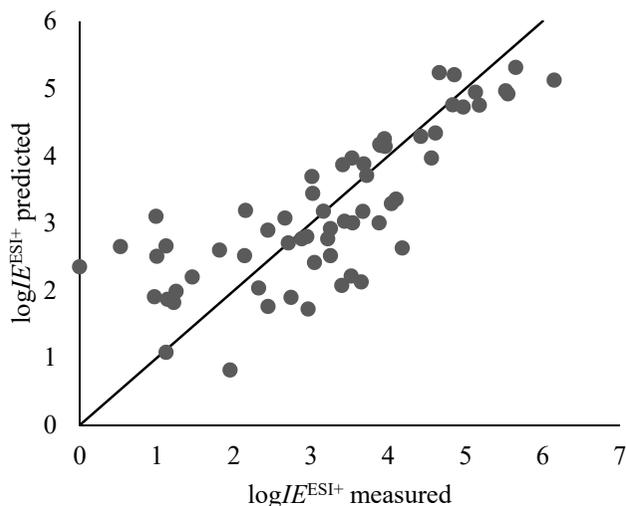
Figure 15. Predicted $\log IE^{\text{ESI+}}$ values calculated with the modified prediction model containing *WANS* plotted against measured $\log IE^{\text{ESI+}}$ values.

The intercept of such models carries information about the $\log IE$ (relative to the anchoring compound) of a hypothetical compound which is neutral (degree of ionization, α = 0) but has infinitely delocalized charge (*WANS* or *WAPS* = 0). Such compounds should ionize to the same extent in ESI+ and ESI−. Therefore, the numerical value of intercept carries information about the anchoring compound and the difference in intercepts allows comparing the anchoring values. The difference in the intercepts (2.97) in the $\log IE$ prediction models for ESI+ and ESI− (Figure 14) was statistically insignificantly different (*t*-test on 95% confidence level) from $\Delta_{\text{anchors}}^{\text{ESI+vs ESI−}}$ (3.01). This implies that the difference is not accidental and has been correctly assigned.

3.3.3 Comparison of ESI+ and ESI− mode for compounds ionizing in both modes

Now the ESI+ and ESI− have been made quantitatively comparable by the use of a common anchor. This knowledge was applied to a set of 33 compounds (Table 7), which ionize in both ESI+ (a range of 3.5 $\log IE$ units) and ESI− (a range of 3 $\log IE$ units) and found the difference between $\log IE^{\text{ESI+}}$ and $\log IE^{\text{ESI-}}$ by taking into account the difference in the anchors (Figure 16).

Table 7. Compounds used in the quantitative comparison of ESI+ and ESI− mode and pK_a, pK_b (pK_a of conjugated acid of a base), $\alpha^{\text{ESI+}}$, $\alpha^{\text{ESI-}}$, $\log IE^{\text{ESI+}}$, $\log IE^{\text{ESI-}}$, charge delocalization parameters (*WANS* and *WAPS*) values for these compounds.

Name	$\log IE^{\text{ESI+}}$	$\log IE^{\text{ESI-}}$	pK_a	pK_b	$\alpha^{\text{ESI+}}$	$\alpha^{\text{ESI-}}$	*WANS* $\cdot 10^5$	*WAPS* $\cdot 10^5$
2-pyridinepropionic acid	3.19	2.94	4.0	4.6	0.5	0.8	4.28	5.75
2-nitrobenzoic acid	1.29	2.43	2.3	-33.2	1.0	0.0	5.12	5.81
phthalic acid	1.91	3.33	3.2	-32.3	0.9	0.0	5.18	5.81
salicylic acid	1.73	3.37	2.8	-6.5	0.9	0.0	5.76	6.06
2-aminobenzoic acid	2.69	3.02	4.7	2.1	0.2	0.0	6.31	6.68
4-aminophenol	2.55	1.54	10.4	5.4	0.0	1.0	8.52	8.67
4-pyridinepropionic acid	3.37	2.93	4.0	5.1	0.5	0.9	4.69	5.81
2-aminophenol	2.80	2.39	10.3	4.5	0.0	0.8	8.07	8.27
terephthalic acid	1.87	3.12	3.6	-31.9	0.7	0.0	5.09	5.63
4-hydroxybenzoic acid	1.45	3.03	4.4	-6.1	0.3	0.0	5.93	6.80
4-aminobenzoic acid	2.61	2.81	4.7	2.7	0.2	0.0	6.75	6.96
2-cyanophenol	1.25	3.29	8.1	-6.3	0.0	0.0	5.61	6.3
benzoic acid	1.22	3.01	4.1	-31.5	0.4	0.0	6.06	7.1
Asn	2.12	2.52	4.0	6.5	0.5	1.0	7.96	8.27
Asp	2.21	2.52	4.0	6.5	0.5	1.0	8.59	8.52
Gln	2.31	2.43	4.1	7.3	0.4	1.0	6.79	7.35
Glu	2.39	2.7	3.9	7.5	0.6	1.0	7.46	8.66
His	3.04	2.32	3.8	7.3	0.6	1.0	7.39	7.23
Ile	2.75	2.92	4.8	7.6	0.1	1.0	5.54	3.39
Leu	3.15	2.89	4.8	7.5	0.1	1.0	5.49	3.79
Met	2.70	2.39	4.5	7.5	0.2	1.0	5.26	6.13
Phe	3.07	3.7	4.4	7.5	0.3	1.0	5.13	5.06
Ser	2.14	2.52	4.0	7.0	0.5	1.0	10.09	10.77
Thr	2.35	2.49	4.2	7.1	0.4	1.0	7.95	8.90
Trp	2.62	2.96	4.5	7.4	0.2	1.0	5.20	5.04
Tyr	2.51	3.31	4.0	7.5	0.5	1.0	6.26	6.00
Val	2.70	2.85	4.7	7.6	0.2	1.0	6.62	7.58
Phe-Phe-Phe-Phe	4.69	4.63	3.6	7.7	0.7	1.0		
Gly-Gly-Gly-Phe-Phe-NH$_2$	3.85	3.63	12.4	7.8	0.0	1.0		
Gly-Pro-Gly-Gly	3.79	2.94	3.2	7.8	0.9	1.0		
Gly-Gly-Asp-Ala	2.78	3.04	3.2	7.8	0.9	1.0		
Gly-β-Ala-β-Ala	2.57	2.65	3.8	8.1	0.6	1.0		
Gly-Gly-Gly-NH$_2$	1.91	2.11	12.6	7.8	0.0	1.0		

The difference in $\log IE$ values ($\log IE^{\text{ESI+}}$ - $\log IE^{\text{ESI-}}$) was compared against a constant of 0.3 logarithmic units, which refers to a two times difference in signal scale. A two-time increase is of significance for practitioners and is also statistically significant (based on repeatability of the measurements). Out of the investigated compounds, which all can be ionized in both modes, for six

compounds ESI+ is preferred (log*IE* difference > 0.3 log units), and for 15 compounds ESI− is preferred (log*IE* difference < -0.3). Twelve compounds ionize to a similar degree in ESI+ and ESI−. Thus, in approximately 46% of cases a compound is better ionized in ESI−, and indifferent for 36%.
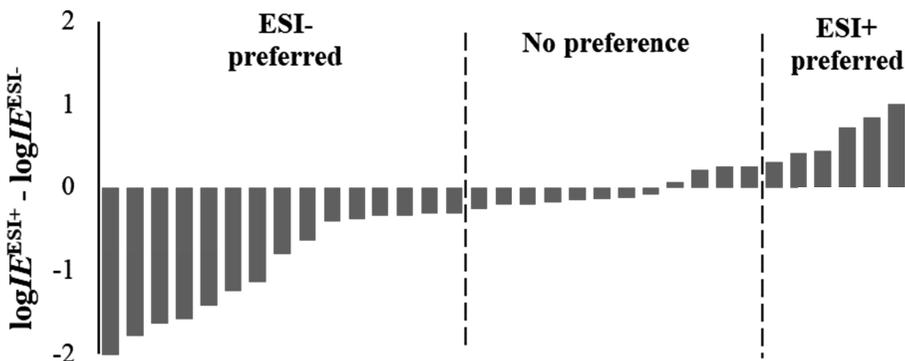
Figure 16. The difference in log*IE* values for compounds ionizing in both ESI+ and ESI− between the two modes.

Ionization efficiency is considerably enhanced (up to 100 times) in ESI− for compounds with low logIE^{ESI+}. More precisely, ESI− is preferred by compounds that are only oxygen bases (e.g. mostly carboxylic acids). Compounds being simultaneously both oxygen (carboxylic acid) and nitrogen bases fall into all three categories. For small peptides and amino acids, the differences between ESI+ and ESI− are mostly very small and either of the modes could be used. One of the exceptions is histidine, which contains a basic side chain that may account for a strong preference towards ESI+ mode (0.72 log*IE* units, 5.3 times). Aspartic acid and glutamic acid, amino acids with acidic side chain, show a weak preference towards ESI−. When interpreting these results, it is, of course, important to keep in mind that all the investigated compounds have acidic groups in their structure.

From previous studies[13,45] it is known that ionization efficiency in both modes is best described by the degree of ionization in solution (*α*) and charge delocalization (*WAPS* or *WANS*). For most compounds, the ionization degree in solution is sufficient to explain the ionization mode preference, including examples described above. For example, compounds that are oxygen bases (such as carboxylic acids) tend to be very weak bases but are at the same time medium strength acids. Therefore, the formation of anions is preferred for these compounds. For amino acids, both amino and carboxylic acid group are expected to be charged, and therefore only small differences between ESI+ and ESI− are expected.

Exceptions to this rule are the amino acids tyrosine and methionine. They are both fully ionized in ESI+ and ESI− solvent; however, tyrosine is better ionized

in ESI− and methionine is better ionized in ESI+. To explain this phenomenon, there is a need to look at the ionization process in more detail. There are two main requirements for a compound to become ionized: (1) it needs to become charged, and (2) it needs to partition to the surface of ESI droplet to be ejected to the gas phase. Previously differences in anion and cation surface affinities have been studied for protons and hydroxide ions[140] and significant differences have been observed. These differences are largely determined by ion-solvent interactions, which are different for cations and anions of the same analyte, arising from differences in charge delocalization, stereochemistry, and solvent properties. In the case of methionine, the charge in the cation is significantly better delocalized than in the anion, based on the comparison of *WAPS* and *WANS* (6.1 and 5.3, respectively, lower value means better charge delocalization). For tyrosine, the charge is slightly better delocalized in the anion (6.0 and 6.3 respectively). Therefore, the preferred ionization mode for tyrosine is opposite to methionine due to opposite charge delocalization of cation and anion. Additionally, tyrosine has an ionizable side-chain in ESI−.[141]

3.4 Mining data from literature and transforming it to enhance prediction model

As already discussed in the review of literature, there are several groups who have been involved in researching ionization efficiencies, but the biggest shortcoming of those studies is that they are mostly focused on a small set of compounds (usually structurally similar) and/or limited eluent system. In order to overcome this, it would be very beneficial if the data already published could be pooled together and made quantitatively comparable. This would allow for a large scale and wide chemical space covering analysis of ionization efficiencies and would allow making more impactful conclusions.

In order to do this (1) an approach to unify literature data to make it numerically comparable is proposed; (2) the consistency of the data presented in the literature is evaluated, and (3) guidelines for presenting relevant data in the future will be offered. To do this correlation of data from literature with Quantem prediction model recently proposed by our group will be used (Figure 17).[81] With this, the data from the literature will be unified and, for the first time, the values obtained from different sources will be compared.
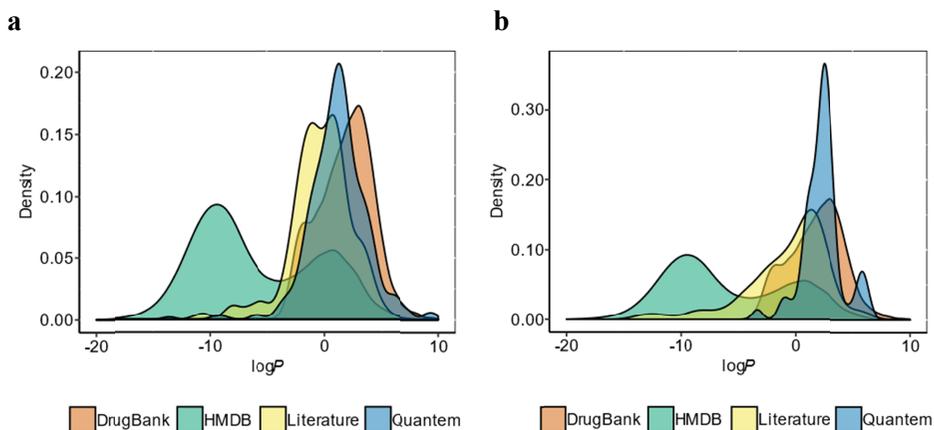
Figure 17. Comparison of the chemical space coverage based on log*P* values. a – ESI positive mode; b – ESI negative mode. The comparison is made with databases of DrugBank,[142] Human Metabolome Database (HMDB),[143] data used in model development (Quantem), and data pooled from literature. HMDB also includes compounds that have not been and cannot be measured with LC/MS.

3.4.1 Transformation of RF values into a uniform dataset

In literature, papers contained information in several forms: (1) relative response factors (*RRF*), (2) response factors (*RF*), (3) slopes of the calibration curve, and (4) data tables and/or charts with MS signal and concentration of the compound. All of this information was used to obtain logarithmic relative response factors (log*RRF*). The reference compound for each set of data was randomly chosen (if it was not already chosen in the original paper) so that this reference compound would have been present in all of the studied conditions (eluent, pH, equipment, etc.). In order to obtain meaningful results, at least 4 compounds per setup were required. Molar concentrations were used to make the data consistent.

All of the measurements from different publications were carried out relative to a different reference compound. This causes a systematic shift in the relative response factor values from the literature data. Additionally, measurements were carried out on different instruments and with different eluents. It is known from previous studies that ionization efficiencies measured on different instruments are correlated but may be compressed or expanded (slope values are significantly different from 1) relative to each other.[144] This arises from the differences in the ionization source, e.g. the heated ESI sources, such as the Jet Stream, which compress the ionization efficiency scale, thus the ionization efficiency values measured on this source are closer (similar) to oneanother.[144,145]

To make all the ionization efficiency values retrieved from the literature comparable, the values need to be transferred to the same reference system. In our previous studies, all of the IE values have been measured as relative values;

the *IE* of methyl benzoate and benzoic acid have been arbitrarily taken as IE = 1 (log*IE* = 0) in ESI positive and negative mode, respectively. Here, a transformation of the literature values was carried out using the predicted ionization efficiency values. For the transformation, half of the points were randomly selected (log*RRF* values from literature and corresponding predicted log*IE* values) and fitted a linear regression. The slope and intercept of the linear regression were used to transfer the rest of the log*RRF* values to the same scale (log*RRF*$_{comparable}$) with the predicted log*IE* values.

$$\log RRF_{comparable} = \frac{\log RRF - intercept}{slope} \qquad (24)$$

The transformation was carried out separately for each literature reference and setup used (for both different instruments and eluents). The transformation of the log*RRF* values is more reliable for datasets containing more compounds; therefore, papers with less than 4 data points per setup could not be included.
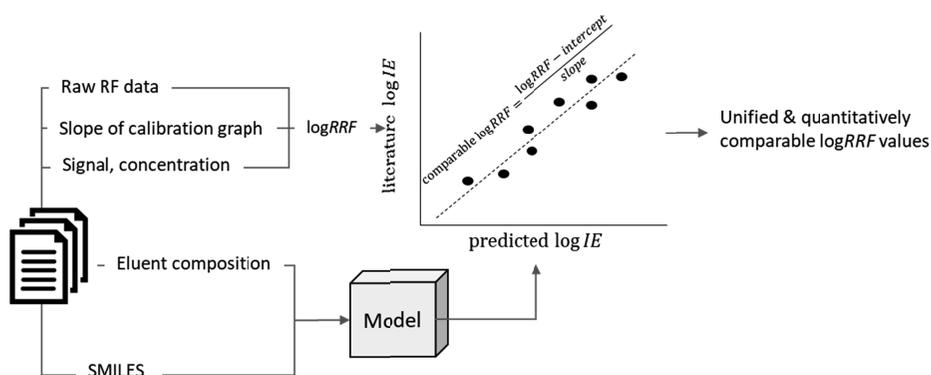
Figure 18. Process of pooling data from the literature, predicting ionization efficiency values and transforming literature log*RRF* values to quantitatively comparable values.

3.4.2 Uniform dataset characterization

The number of papers published in scientific journals is exponentially growing and so is the amount of data that could be used in interdisciplinary research. Even though open data is becoming a requirement for most leading scientific journals and most competitive funding schemes, an important question is if and how can the data be reused. While combining data for this paper, retrieving data from literature sources was not a straightforward task. Although FAIR data principles[146] are being acknowledged worldwide, there is a lack of standards on data presentation, a form of availability, and data curating.

Based on a thorough literature search, 57 papers were analyzed. Only 15 papers for ESI positive mode,[11,59,63,80,82,83,147–155] 7 papers for ESI negative mode[67,72,74,76,156–158] and two papers for both ESI positive and ESI negative

mode[78,98] contained information in a format that enabled retrieving and transforming the data as well as subsequent testing of the logIE prediction model. In the papers studied, most of the data was presented in figures and not in a tabulated form. This is probably due to the fact that visual data is easier to grasp. However, retrieving data from graphs is very difficult, if not impossible. Sometimes the numerical data was given in Supporting Information, but that was rather exceptional. In some cases, the values could be found, but it was very difficult to retrieve the metadata (setup, eluent, concentration). Also, the authors were contacted to clarify experimental details; however, often the authors did not reply. One clear solution to simplify retrieving data from literature would be to publish the raw data in a tabulated format: the conditions used, exact concentrations of the solutions, together with the signals obtained.

From the analyzed papers 634 RF values of 440 unique compounds in 64[a] different eluent compositions were collected in ESI positive mode. In ESI negative mode, 373 RF values for 161 unique compounds in 47[b] different eluent compositions were collected (see Appendix 6 and Figure 18). The measurements had been carried out on instruments from 7 different vendors by 22 research groups.

The data collected from the literature supports the assumption that pooling data together from different literature sources will allow diversification. From the distribution of log$P$[159] (Appendix 6) values for each set of incorporated compounds it can be seen, that for both ESI positive and ESI negative mode, each individual study has focused on a specific set of compounds and, therefore, covers only a narrow log$P$ range. Additionally, in some studies, only compounds with very narrow pK_a range were incorporated, but the distribution of pK_a[160,161] values from over all the studies is much more diverse (Figure 19).

a   **b**
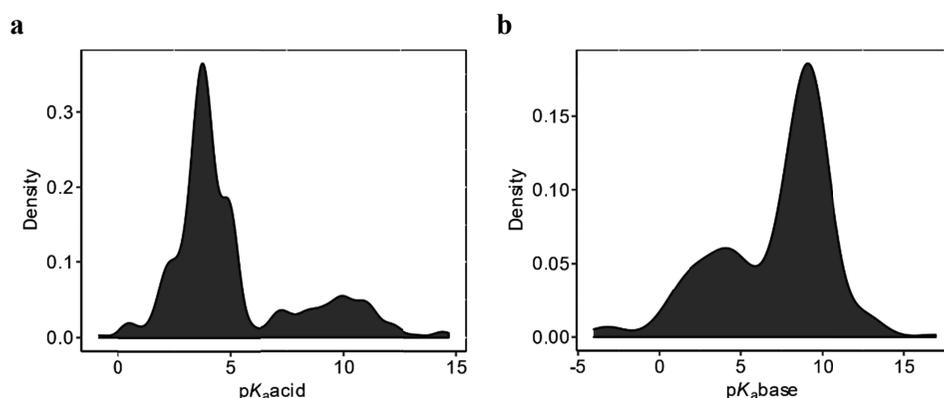
Figure 19. Distribution of aqueous pK_a values of studied compounds from literature a - acidic group pK_a,[160] b - pK_a[160] of the conjugate acid of basic group.

---

[b] Including eluent differences arising from chromatographic separation.

3.4.3 Comparing the log*RRF* values from the literature

Previously, it has been shown that the range of ionization efficiency values for the same set of compounds varies from instrument to instrument.[122] Additionally, the numerical values depend on the used reference compound. In order to overcome this variation, the transformation approach (described in the Experimental section) was used, which transfers the numerical *RF* values measured under different conditions to comparable log*RRF*$_{comparable}$ values (Figure 18). The transformation does not influence the order of compounds within a set or the relative differences between the compounds therein.

The correlation (Figure 20) between the log*RRF*$_{comparable}$ values obtained from publications and the values predicted by the log*IE* prediction model is high ($R^2 = 0.61$) for ESI positive mode. The s_{MAE} is less than 2.0 times ($s_{RMSE} = 2.3$ times). For 93% of compounds, the prediction error is less than 5 times. This accuracy is as good as for the prediction accuracy of the Quantem model development dataset (2.2 times).[81] Overall, the log*IE* prediction model correlates well with *RF* values obtained from literature in ESI positive mode and the prediction is good over various eluent compositions and different compound classes.

In ESI negative mode, the correlation between log*RRF*$_{comparable}$ values obtained from literature and log*IE* values predicted by the model is slightly lower ($R^2 = 0.50$, Figure 20). The s_{MAE} is 3.8 times ($s_{RMSE} = 5.9$ times). Similarly to positive mode, the model for negative mode was developed by using a broad set of compounds (log*P* from -6 to 9.5, molar masses 89 to 675 Da) and eluent conditions. However, the data retrieved from literature contained conditions not included in the model development, which is the main reason for the low correlation in the negative mode.

Based on the good correlation observed in Figure 20, especially for ESI positive mode, it is obvious that ionization efficiency values are comparable between publications. Our transformation approach shows good agreement even where originally contradictory findings have been observed. For example, Hermans et al.[63] observed a weak correlation between ionization efficiency and log*P* values. Ehrmann et al.[82], on the other hand, observed no clear relationship between ionization efficiency and log*P* values. After transformation, it can be seen that data from both papers shows good consistency with the ionization efficiency values predicted with Quantem model (s_{MAE} of 1.6 and 2.5 times, respectively). For the first time, the log*RRF*$_{comparable}$ approach allows the comparison of ionization efficiency values of amino acids, measured by Hermans et al.,[63] to the nucleobases, measured by Ehrmann at al.[82] It is observed that the values for nucleobases are very close to the values of the amino acids. For example, on the transferred log*RRF*$_{comparable}$ scale tryptophan has only 1.3 times higher response than cytosine. This is an example of how pooling data together with the proposed transformation approach allows obtaining a deeper understanding than each independent publication can.
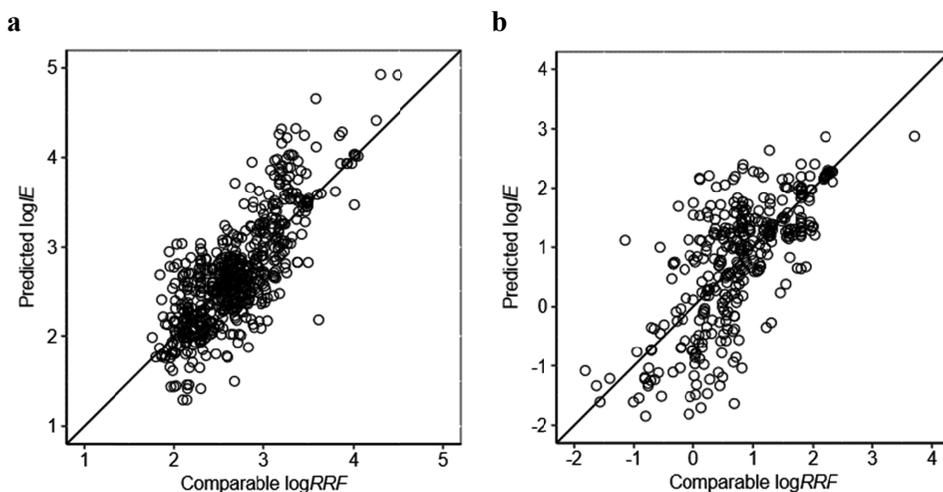
Figure 20. Correlations between predicted log*IE* and log*RRF*_{comparable} for a – ESI positive, b – ESI negative mode.

3.4.3.1 Different compound classes

To understand the scope of the developed model, the performance of the model in different compound classes needs to be assessed. For comparison in ESI positive mode, the compounds were divided into three sets: (1) hydrophobic oxygen bases like steroids, lipids, and fatty acids; (2) small molecules; and (3) peptides. As most of the studies had focused on the ionization of small molecules, set number (2) was the largest, containing 530 *RF* values (of which 353 unique compounds) from 8 studies. The peptide set was the smallest containing 43 *RF* values (31 unique compounds) from three studies. Interestingly, in all of the groups, the accuracy of ionization efficiency predictions was roughly the same: s_{RMSE} 2.4, 2.4 and 1.6 times; s_{MAE} 2.0, 2.0 and 1.4 times (Figure 21), respectively. Distinctive sets could not be formed in ESI negative mode.
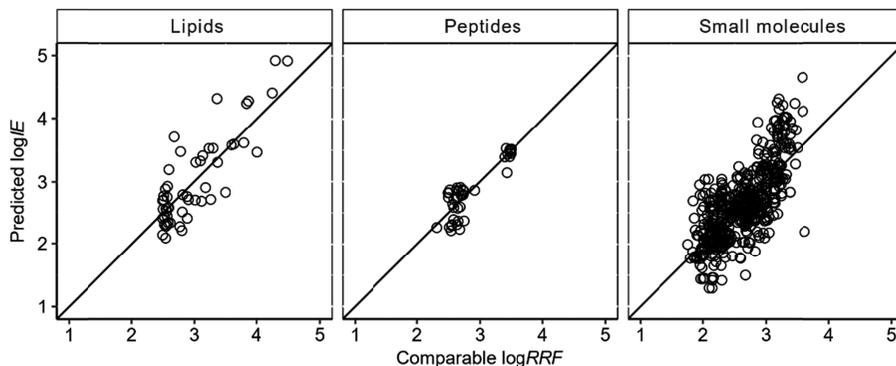
Figure 21. Comparison of predicted and measured values grouped by compound class in ESI positive mode. Lipids – Cifkova2012,[150] Byrdwell2013,[148] Alymatiri2015;[78] Peptides – CechEnke2000,[11] Leitner2007,[151] Raji2009,[152] Stavenhagen2013;[153] Small molecules – Yang2006,[155] Ehrmann2008,[162] Chalcraft2009,[149] Bedner2011,[147] Mandra2015,[80] Cramer2017,[59] Hermans2017,[63] Kiontke2017.[83]

3.4.3.2 Eluent pH

In ESI positive mode, data was divided into three groups based on aqueous phase pH of the eluent: acidic (pH < 5), neutral ($5 \leq$ pH < 8), and basic ($8 \leq$ pH). In all groups, the consistency between predicted and measured results is good (2.1 to 3.0 times); the points are close to the ideal fit line (Appendix 6). Comparing the s_{RMSE} values, slightly higher accuracy is observed for data measured with eluent containing water phase at acidic pH values (s_{RMSE} = 2.1 times). For neutral pH values, accuracy (s_{RMSE} = 2.3 times) is similar to the acidic subset; in basic pH values, the accuracy is unsurprisingly somewhat lower ($RMSE$ = 3.0 times) as basic water phase is in general not favorable for ESI positive mode.[9]

For ESI negative mode, the obtained accuracies are somewhat lower than in ESI positive (see Appendix 6). Unsurprisingly, the highest accuracy (s_{RMSE} = 2.1 times) is observed in basic pH, which is the most favorable for ESI negative mode.[9] The lowest accuracy is observed for neutral conditions (s_{RMSE} = 6.8 times). In acidic eluent, the predicted logIE values are slightly more accurate (s_{RMSE}= 2.8 times).

3.4.3.3 Organic modifier

In ESI positive mode, there were only two organic modifiers used: acetonitrile (MeCN) and methanol (MeOH). The s_{MAE} values are 1.8 times (s_{RMSE} = 2.2 times) and 2.0-times (s_{RMSE} = 2.5 times), respectively. This is not surprising as in the model development both organic modifiers were used as well.

Similarly to ESI positive mode, in ESI negative mode the model does not seem to have a preference. In addition to MeOH and MeCN, other, less common eluents were used: (1) acetone, for which the prediction model does not seem to work at all; (2) chloroform/MeOH mixture, for which there were very few data points with very similar IE values; (3) MeCN/isopropanol mixture, for which the model seems to work acceptably well (s_{MAE} = 1.6 times, s_{RMSE} = 1.8 times); and (4) tetrahydrofuran, for which there were too few data points, all in a very narrow range. Additionally, pure water was used, for which the prediction model does not work well (s_{MAE} = 5.6 times, s_{RMSE} = 7.7 times). Only MeCN and MeOH had been used in model development as organic modifiers and, therefore, the model is not expected to yield as good predictive power for eluents mentioned above.

3.4.3.4 Problematic issues

Analysis of the biggest outliers is useful to indicate the limitations of the approach. It is of interest if some specific groups of compounds are performing worse than others. Based on the distribution of the $\log P$ values (Appendix 6), the compounds from the study by Stavenhagen et al.[153] are the most hydrophilic ones. Correlating the $\log RRF_{comparable}$ values with predicted $\log IE$ values, the accuracy for the aforementioned compounds is good. Surprisingly, compounds from some publications which are chemically similar to the ones used in Quantem model development show large deviations. One reason could be that the similarity among the compounds in particular dataset is high, e.g. in the studies by Cramer et al.,[59] Mandra et al.[80] and Kiontke et al.[83] Even though the number of compounds is high, the physicochemical properties of the compounds are very similar.

The accuracy of the ionization efficiency transformation in ESI negative mode is not affected by the differences in the properties of the compounds used in literature and for Quantem model development. For example, compounds studied by Koivusalo et al.[163] were significantly more hydrophilic (see Appendix 6), but the prediction accuracy for these compounds is high (s_{MAE} = 1.1 times, s_{RMSE} = 1.1 times). The set from the study of Huffman et al.[74] showed lower prediction accuracy despite the significant overlap in the properties of the compounds used in model development. However, the compounds with low accuracy are measured either in a neat organic solvent or in the neat water phase. Such eluents, especially without a pH adjusting additive, are not favorable conditions for electrospray.[52]

3.4.3.5 Technical aspects that need to be considered before retrieving data from the literature

In order to make maximum use of the data collected by the community, it needs to be clear how data has been obtained. The lack of metadata does not allow

assessing the quality of the data. Used eluent, elution conditions, additive concentrations need to be unambiguously known. Furthermore, it has to be clear whether (1) the measurements have been carried out in the linear range and (2) the isotope peak distribution is accounted for.

Firstly, measurements are often carried out at the same concentration level for compounds with different ionization efficiencies without checking if the concentration is in the linear range. The signal linearity check is very important, as compounds measured under signal saturation conditions appear to have equal ionization efficiencies. This makes it impossible to determine causality between molecular properties and ionization efficiency.

Secondly, MS signal is divided into several peaks; especially for compounds containing chlorine, bromine, selenium, and sulfur. If the signals of two compounds with remarkably different isotope distribution are compared, the result is strongly affected by the isolation width used in measurements and in the data processing. Often, information about such practices is not available. Previously, we have used an isolation width of ≤1 Da and accounted for the presence of isotope peaks via isotope correction. This approach is recommended to others, as it is less influenced by the physical and chemical noise than simply increasing the isolation width.

SUMMARY

Compounds have several orders of magnitude different ionization efficiencies in ESI source leading to a misestimation of concentrations up to million times if just signals are used for quantification. Ionization efficiency prediction possesses high potential to overcome the lack of standard substances in LC/ESI/MS analysis. The aim of this thesis was to broaden the existing understanding of ESI in terms of analytes, matrices, instrument and developing a strategy to unify the vastly different and contradictory ionization efficiency measurements into one dataset. All of the gained knowledge throughout this work will allow the development of a universal model for predicting concentrations which is not limited to specific compound families, matrices or instrumental setups.

Firstly, the ionization efficiency measurement was expanded to multiply charged analytes on the example of indicator molecules and oligopeptides. It was shown that the ionization efficiency of a peptide based on the ionization efficiencies of amino acids it consists of. Also, it was shown that from peptide length of 5-6 amino acid residues there is a leveling of effect, which indicates that from that length onwards presuming equal ionization efficiencies does not produce a great error. For the first time, the gap between ionization efficiencies of small molecules and larger ones has been closed.

Secondly, as all the ionization efficiency measurements have been previously carried out in a neat solvent, it was seen that the ionization efficiency measurement approach can also be applied to more complex matrices. The ionization efficiencies of pharmaceutical compounds were measured in blood, plasma, urine, cerebrospinal fluid, and liver and brain tissue homogenates. It was seen that the overall ionization efficiency prediction error for these matrices is on average 8 times. This shows that the developed ionization efficiency approach is universal and can be applied to various matrices.

Thirdly, the two measurement modes, positive and negative ESI were combined with the help of anchor compound that ionizes to a similar extent in both modes. By this connection, ionization efficiencies measured on either mode are directly comparable to values measured in the other mode allowing to choose the best mode for a specific analyte. Interestingly, it was discovered that for a set of compounds ionizing in both modes, negative ionization mode seems to be the preferred one.

Lastly, all the knowledge gathered previously was combined to transform data from literature into one uniform dataset. This demonstrated that although contradictory conclusions may have been reached in individual publications, the overall consistency of data is good. Controversial results were probably reached due to a limited set of compounds and/or conditions studied in each publication. The developed transformation approach can be applied to other previous and future datasets to combine data into a uniform dataset. This will allow for quantitative comparison of data independent of location, eluent, and the instrument used. It will additionally allow gaining more insight into ESI mechanism

and developing more comprehensive model do predict ionization efficiencies more accurately thus enabling quantitation without standard substances.

The results of this thesis enabled to expand ionization efficiency measurement methodology multiply charged analytes and allowed for direct quantitative comparison of ionization efficiencies measured in different ESI modes, matrices and even in different labs. The gained knowledge lays a solid foundation to move forward with a universally applicable ionization efficiency prediction model that would allow standard substance free quantification as well as possible deeper understanding of the ESI process.

SUMMARY IN ESTONIAN

Elektropihustus ionisatsiooniefektiivsuste mõõtmismetoodika ja rakenduste laiendamine ja täiustamine

Elektropihustusionisatsioon-vedelikkromatograafia-massispektromeetria (ESI/LC/MS) on üks enim kasutatavaid meetodeid erinevate analüütide uurimiseks mitmetes valdkondades keskkonna analüüsidest ravimiarenduseni ja toiduohutuseni. Kuigi ESI/LC/MS võimaldab määrata väga madalaid ainete sisaldusi, on täpseteks mõõtmisteks vajalik standardainete kasutamine. See on tingitud sellest, et ainete ionisatsiooniefektiivsused erinevad miljoneid kordi, s.t. sama kontsentratsiooni kuid erineva ionisatsiooniefektiivsusega analüüdid võivad anda miljoneid kordi erineva signaali. Ionisatsiooniefektiivsus defineeritakse analüüdi molekulidest või ioonidest genereeritud gaasi faasi ioonide määrana. Ionisatsiooniefektiivsus sõltub aine struktuurist, kasutatavast eluendist, kasutatud ESI allika geomeetriast ning analüüdiga samal ajal elueeruvatest ühenditest. See tähendab, et kasutades vaid signaale ainete sisalduse määramiseks võime saada suurusjärke valesid tulemusi. Standardainete kasutamine pole alati võimalik, näiteks kui tehtud avastus on niivõrd uudne, standardainet pole võimalik sünteesida või selle soetamine on võimatu.

Ionisatsiooniefektiivsusi on uuritud mitmete erinevate teadusgruppide poolt kuid tihti on nendest tehtavad järeldused vastuolulised. Vastuolude põhjuseks võib olla, et tihti on ühe uuringu siseselt kasutatud väga sarnaseid ühendeid, ühendeid on vähe ning uuritud tingimused on väga piiratud. Laiema pildi mõistmiseks on vajalik erinevate teadusgruppide andmed ühendada ja uurida universaalseid trende.

Seega oli doktoritöö eesmärkideks: (1) laiendada ja täiendada ionisatsiooniefektiivsuste mõõtmise metoodika ka mitmelaengulisi ioone andvatele analüütidele ning keerulisematele maatriksitele, näiteks bioloogilised vedelikud; (2) ühendada ESI positiivses ja negatiivses režiimis saadud ionisatsiooniefektiivsuste väärtused numbriliselt võrreldavaks ja (3) arendada välja meetod, et koondada erinevad kirjandusandmed ning muuta need numbriliselt võrreldavaks andmekomplektiks.

Esmalt arendati väikeste indikaatormolekulide ja oligopeptiidide abil välja metoodika, et uurida mitmelaengulisi ioone moodustavate ühendite ionisatsiooniefektiivsusi. Näidati, et peptiidid ionisatsiooniefektiivsus on võrdeline teda moodustuvate aminohapete ionisatsiooniefektiivsuste summaga. Samuti selgus, et alates 5–6 aminohappe jääki sisaldavast oligopeptiidist toimub ionisatsiooniefektiivsuste nivelleerumine: sellisest pikkusest alates ei too võrdsete ionisatsiooniefektiivsuste eeldamine kaasa suuri vigu sisalduse määramisel.

Samuti laiendati ionisatsiooniefektiivsuste mõõtmise metoodikat keerulisematele bioloogilistele vedelikele. Ravimilaadsete ühendite ionisatsiooniefektiivsused mõõdeti veres, vereplasmas, seljaaju vedelikus, uriinis ning aju- ja maksakoe homogenaatides ning täheldati väga head korelatsiooni solvendis mõõdetud ionisatsiooniefektiivsuste väärtustega. Seetõttu oli võimalik ennustada ka keeru-

listes bioloogilistes vedelikes ühendite sisaldusi kasutades füüsikalis-keemilisi parameetrite kaudu ionisatsiooniefektiivsuste modelleerimist. Mudel võimaldab hinnata sisaldusi keskmise kaheksa kordse ennustusveaga.

Seejärel ühendati ionisatsiooniefektiivsuste skaalad ESI positiivses ja negatiivses režiimis kasutades ankurühendit, mis ioniseerub sarnasel määral mõlemas režiimis. Ühendamise tulemusena on mõlemas režiimis mõõdetud väärtused kvantitatiivselt võrreldavad ja võimaldavad valida konkreetse analüüdi jaoks sobiva ESI režiimi. Üllatuslikult selgus, et uuritud ainete jaoks, mis ioniseeruvad mõlemas režiimis olid ionisatsiooniefektiivsused kõrgemad negatiivses režiimis, mis on seni olnud vähem populaarne analüüsi režiim.

Kasutades kogu eelnevalt kogutud infot, arendati välja metoodika, et koondada andmeid kirjandusest ja ühendada need üheks andmekomplektiks. Ühendamine tõi välja asjaolu, et kuigi üksikute artiklite järeldused on vastuolulised, on andmete omavaheline kooskõla teadusgruppide vahel hea. Ilmselt olid vastuolulised järeldused tingitud piiratud tingimustest ja/või ainete omadustest. Arendatud metoodikat saab rakendada kõigile varasematele aga ka tulevikus saadavatele andmetele, et saavutada kvantitatiivselt võrreldavad tulemused.

Seega said täidetud kõik püstitatud eesmärgid ning doktoritöö tulemusena on võimalik kvantitatiivselt võrrelda erinevates laborites, erinevatel instrumentidel ja erinevates maatriksites mõõdetud ionisatsiooniefektiivsusi. Suuremad ja usaldusväärsemad andmehulgad võimaldavad arendada universaalse mudeli ionisatsiooniefektiivsuste ennustamiseks, mis omakorda vähendab vajadust standardainete järele LC/ESI/MS analüüsil. See omakorda võimaldab luua selgust ESI ionisatsioonimehhanismi osas.

REFERENCES

(1) Fenn, J. B. Electrospray Wings for Molecular Elephants (Nobel Lecture). *Angew. Chem. Int. Ed.* **2003**, *42* (33), 3871–3894. https://doi.org/10.1002/anie.200300605.

(2) Cech, N. B.; Enke, C. G. Practical Implications of Some Recent Studies in Electrospray Ionization Fundamentals. *Mass Spectrom. Rev.* **2001**, *20* (6), 362–387.

(3) Doerr, A.; Finkelstein, J.; Jarchum, I.; Goodman, C.; Dekker, B. The Nature Milestones in Mass Spectrometry. *Nat. Methods* **2015**, *12*, S1–S45.

(4) García-Sevillano, M. A.; García-Barrera, T.; Navarro, F.; Montero-Lobato, Z.; Gómez-Ariza, J. L. Shotgun Metabolomic Approach Based on Mass Spectrometry for Hepatic Mitochondria of Mice under Arsenic Exposure. *BioMetals* **2015**, *28* (2), 341–351. https://doi.org/10.1007/s10534-015-9837-9.

(5) Fuhrer, T.; Zamboni, N. High-Throughput Discovery Metabolomics. *Curr. Opin. Biotechnol.* **2015**, *31*, 73–78. https://doi.org/10.1016/j.copbio.2014.08.006.

(6) Schuhmann, K.; Almeida, R.; Baumert, M.; Herzog, R.; Bornstein, S. R.; Shevchenko, A. Shotgun Lipidomics on a LTQ Orbitrap Mass Spectrometer by Successive Switching between Acquisition Polarity Modes: Shotgun Lipidomics with Polarity Switching. *J. Mass Spectrom.* **2012**, *47* (1), 96–104. https://doi.org/10.1002/jms.2031.

(7) Jeandet, P.; Heinzmann, S. S.; Roullier-Gall, C.; Cilindre, C.; Aron, A.; Deville, M. A.; Moritz, F.; Karbowiak, T.; Demarville, D.; Brun, C.; et al. Chemical Messages in 170-Year-Old Champagne Bottles from the Baltic Sea: Revealing Tastes from the Past. *Proc. Natl. Acad. Sci.* **2015**, *112* (19), 5893–5898. https://doi.org/10.1073/pnas.1500783112.

(8) Norbert Hertkorn; Mourad Harir; Kaelin M. Cawley; Philippe Schmitt-Kopplin; Rudolf Jaffé. Molecular Characterization of Dissolved Organic Matter from Subtropical Wetlands: A Comparative Study through the Analysis of Optical Properties, NMR and FTICR/MS. *Biogeosciences* **2016**, *13* (8), 2257–2277. https://doi.org/10.5194/bg-13-2257-2016.

(9) Kostiainen, R.; Kauppila, T. J. Effect of Eluent on the Ionization Process in Liquid Chromatography–Mass Spectrometry. *J. Chromatogr. A* **2009**, *1216* (4), 685–699. https://doi.org/10.1016/j.chroma.2008.08.095.

(10) Tang, L.; Kebarle, P. Dependence of Ion Intensity in Electrospray Mass Spectrometry on the Concentration of the Analytes in the Electrosprayed Solution. *Anal. Chem.* **1993**, *65*, 3654–3668.

(11) Cech, N. B.; Enke, C. G. Relating Electrospray Ionization Response to Nonpolar Character of Small Peptides. *Anal. Chem.* **2000**, *72* (13), 2717–2723. https://doi.org/10.1021/ac9914869.

(12) Zhou, S.; Cook, K. D. A Mechanistic Study of Electrospray Mass Spectrometry: Charge Gradients Within Electrospray Droplets and Their Influence on Ion Response. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 206–214. https://doi.org/10.1016/S1044-0305(00)00213-0.

(13) Oss, M.; Kruve, A.; Herodes, K.; Leito, I. Electrospray Ionization Efficiency Scale of Organic Compounds. *Anal. Chem.* **2010**, *82* (7), 2865–2872. https://doi.org/10.1021/ac902856t.

(14) Cole, R. B. *Electrospray and MALDI Mass Spectrometry: Fundamentals, Instrumentation, Practicalities, and Biological Applications.*; Wiley: Hoboken, 2010.

(15) Kebarle, P.; Verkerk, U. H. Electrospray: From Ions in Solution to Ions in the Gas Phase, What We Know Now. *Mass Spectrom. Rev.* **2009**, *28* (6), 898–917. https://doi.org/10.1002/mas.20247.

(16) Kebarle, P. A Brief Overview of the Present Status of the Mechanisms Involved in Electrospray Mass Spectrometry. *J. Mass Spectrom.* **2000**, *35* (7), 804–817.

(17) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9. https://doi.org/10.1021/ac302789c.

(18) Iribarne, J. V. On the Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, *64* (6), 2287. https://doi.org/10.1063/1.432536.

(19) Dole, M. Molecular Beams of Macroions. *J. Chem. Phys.* **1968**, *49* (5), 2240. https://doi.org/10.1063/1.1670391.

(20) Ahadi, E.; Konermann, L. Modeling the Behavior of Coarse-Grained Polymer Chains in Charged Water Droplets: Implications for the Mechanism of Electrospray Ionization. *J. Phys. Chem. B* **2012**, *116* (1), 104–112. https://doi.org/10.1021/jp209344z.

(21) Konermann, L.; Rodriguez, A. D.; Liu, J. On the Formation of Highly Charged Gaseous Ions from Unfolded Proteins by Electrospray Ionization. *Anal. Chem.* **2012**, *84* (15), 6798–6804. https://doi.org/10.1021/ac301298g.

(22) Gamero-Castaño, M.; de la Mora, J. F. Mechanisms of Electrospray Ionization of Singly and Multiply Charged Salt Clusters. *Anal. Chim. Acta* **2000**, *406*, 67–91.

(23) Duez, Q.; Metwally, H.; Konermann, L. Electrospray Ionization of Polypropylene Glycol: Rayleigh-Charged Droplets, Competing Pathways, and Charge State-Dependent Conformations. *Anal. Chem.* **2018**, *90* (16), 9912–9920. https://doi.org/10.1021/acs.analchem.8b02115.

(24) Enke, C. G. A Predictive Model for Matrix and Analyte Effects in Electrospray Ionization of Singly-Charged Ionic Analytes. *Anal. Chem.* **1997**, *69* (23), 4885–4893. https://doi.org/10.1021/ac970095w.

(25) Taflin, D. C.; Ward, T. L.; Davis, E. J. Electrified Droplet Fission and the Rayleigh Limit. *Langmuir* **1989**, *5* (2), 376–384. https://doi.org/10.1021/la00086a016.

(26) Zhou, S.; Edwards, A. G.; Cook, K. D.; Van Berkel, G. J. Investigation of the Electrospray Plume by Laser-Induced Fluorescence Spectroscopy. *Anal. Chem.* **1999**, *71* (4), 769–776. https://doi.org/10.1021/ac981259r.

(27) Zhou, S.; Prebyl, B. S.; Cook, K. D. Profiling PH Changes in the Electrospray Plume. *Anal. Chem.* **2002**, *74* (19), 4885–4888. https://doi.org/10.1021/ac025960d.

(28) Girod, M.; Dagany, X.; Antoine, R.; Dugourd, P. Relation between Charge State Distributions of Peptide Anions and PH Changes in the Electrospray Plume. A Mass Spectrometry and Optical Spectroscopy Investigation. *Int. J. Mass Spectrom.* **2011**, *308* (1), 41–48. https://doi.org/10.1016/j.ijms.2011.07.020.

(29) Girod, M.; Antoine, R.; Dugourd, P.; Love, C.; Mordehai, A.; Stafford, G. Basic Vapor Exposure for Tuning the Charge State Distribution of Proteins in Negative Electrospray Ionization: Elucidation of Mechanisms by Fluorescence Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (7), 1221–1231. https://doi.org/10.1007/s13361-012-0375-1.

(30) Girod, M.; Dagany, X.; Boutou, V.; Broyer, M.; Antoine, R.; Dugourd, P.; Mordehai, A.; Love, C.; Werlich, M.; Fjeldsted, J.; et al. Profiling an Electrospray Plume by Laser-Induced Fluorescence and Fraunhofer Diffraction Combined to Mass Spectrometry: Influence of Size and Composition of Droplets on Charge-

State Distributions of Electrosprayed Proteins. *Phys. Chem. Chem. Phys.* **2012**, *14* (26), 9389–9396.

(31) Zhou, S.; Cook, K. D. Probing Solvent Fractionation in Electrospray Droplets with Laser-Induced Fluorescence of a Solvatochromic Dye. *Anal. Chem.* **2000**, *72* (5), 963–969. https://doi.org/10.1021/ac990912n.

(32) Wang, R.; Zenobi, R. Evolution of the Solvent Polarity in an Electrospray Plume. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (3), 378–385. https://doi.org/10.1016/j.jasms.2009.10.022.

(33) Hopkins, R. J.; Reid, J. P. Evaporation of Ethanol/Water Droplets: Examining the Temporal Evolution of Droplet Size, Composition and Temperature. *J. Phys. Chem. A* **2005**, *109* (35), 7923–7931. https://doi.org/10.1021/jp0516543.

(34) Liigand, J.; Kruve, A.; Leito, I.; Girod, M.; Antoine, R. Effect of Mobile Phase on Electrospray Ionization Efficiency. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (11), 1853–1861. https://doi.org/10.1007/s13361-014-0969-x.

(35) Wortmann, A.; Kistler-Momotova, A.; Zenobi, R.; Heine, M. C.; Wilhelm, O.; Pratsinis, S. E. Shrinking Droplets in Electrospray Ionization and Their Influence on Chemical Equilibria. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (3), 385–393. https://doi.org/10.1016/j.jasms.2006.10.010.

(36) Soleilhac, A.; Dagany, X.; Dugourd, P.; Girod, M.; Antoine, R. Correlating Droplet Size with Temperature Changes in Electrospray Source by Optical Methods. *Anal. Chem.* **2015**, *87* (16), 8210–8217. https://doi.org/10.1021/acs.analchem.5b00976.

(37) Oh, M. I.; Consta, S. Stability of a Transient Protein Complex in a Charged Aqueous Droplet with Variable PH. *J. Phys. Chem. Lett.* **2017**, *8* (1), 80–85. https://doi.org/10.1021/acs.jpclett.6b02319.

(38) Sharawy, M.; Consta, S. How Do Non-Covalent Complexes Dissociate in Droplets? A Case Study of the Desolvation of DsDNA from a Charged Aqueous Nanodrop. *Phys Chem Chem Phys* **2015**, *17* (38), 25550–25562. https://doi.org/10.1039/C5CP04331J.

(39) Ahadi, E.; Konermann, L. Molecular Dynamics Simulations of Electrosprayed Water Nanodroplets: Internal Potential Gradients, Location of Excess Charge Centers, and "Hopping" Protons. *J. Phys. Chem. B* **2009**, *113* (20), 7071–7080. https://doi.org/10.1021/jp810599f.

(40) Sharawy, M.; Consta, S. Characterization of "Star" Droplet Morphologies Induced by Charged Macromolecules. *J. Phys. Chem. A* **2016**, *120* (44), 8871–8880. https://doi.org/10.1021/acs.jpca.6b08486.

(41) Consta, S.; Oh, M. I.; Malevanets, A. New Mechanisms of Macroion-Induced Disintegration of Charged Droplets. *Chem. Phys. Lett.* **2016**, *663*, 1–12. https://doi.org/10.1016/j.cplett.2016.08.001.

(42) Malevanets, A.; Consta, S. Variation of Droplet Acidity during Evaporation. *J. Chem. Phys.* **2013**, *138* (18), 184312. https://doi.org/10.1063/1.4804303.

(43) Davis, D.; Portelius, E.; Zhu, Y.; Feigerle, C.; Cook, K. D. Profiling an Electro-spray Plume Using Surface-Enhanced Raman Spectroscopy. *Anal. Chem.* **2005**, *77* (24), 8151–8154. https://doi.org/10.1021/ac0511039.

(44) Girod, M.; Antoine, R.; Dugourd, P.; Love, C.; Mordehai, A.; Stafford, G. Basic Vapor Exposure for Tuning the Charge State Distribution of Proteins in Negative Electrospray Ionization: Elucidation of Mechanisms by Fluorescence Spectro-scopy. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (7), 1221–1231. https://doi.org/10.1007/s13361-012-0375-1.

(45) Kruve, A.; Kaupmees, K.; Liigand, J.; Leito, I. Negative Electrospray Ionization via Deprotonation: Predicting the Ionization Efficiency. *Anal. Chem.* **2014**, *86* (10), 4822–4830. https://doi.org/10.1021/ac404066v.

(46) Ehrmann, B. M.; Henriksen, T.; Cech, N. B. Relative Importance of Basicity in the Gas Phase and in Solution for Determining Selectivity in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 719–728. https://doi.org/10.1016/j.jasms.2008.01.003.

(47) Henriksen, T.; Juhler, R. K.; Svensmark, B.; Cech, N. B. The Relative Influences of Acidity and Polarity on Responsiveness of Small Organic Molecules to Analysis with Negative Ion Electrospray Ionization Mass Spectrometry (ESI-MS). *J. Am. Soc. Mass Spectrom.* **2005**, *16* (4), 446–455. https://doi.org/10.1016/j.jasms.2004.11.021.

(48) Raji, M. A.; Fryčák, P.; Temiyasathit, C.; Kim, S. B.; Mavromaras, G.; Ahn, J.-M.; Schug, K. A. Using Multivariate Statistical Methods to Model the Electrospray Ionization Response of GXG Tripeptides Based on Multiple Physicochemical Parameters. *Rapid Commun. Mass Spectrom.* **2009**, *23* (14), 2221–2232. https://doi.org/10.1002/rcm.4141.

(49) Tang, X.; Bruce, J. E.; Hill, H. H. Characterizing Electrospray Ionization Using Atmospheric Pressure Ion Mobility Spectrometry. *Anal. Chem.* **2006**, *78* (22), 7751–7760. https://doi.org/10.1021/ac0613380.

(50) Huffman, B. A.; Poltash, M. L.; Hughey, C. A. Effect of Polar Protic and Polar Aprotic Solvents on Negative-Ion Electrospray Ionization and Chromatographic Separation of Small Acidic Molecules. *Anal. Chem.* **2012**, *84* (22), 9942–9950. https://doi.org/10.1021/ac302397b.

(51) Zhou, S.; Cook, K. D. Protonation in Electrospray Mass Spectrometry: Wrong-Way-Round or Right-Way-Round? *J. Am. Soc. Mass Spectrom.* **2000**, *11* (11), 961–966.

(52) Kruve, A. Influence of Mobile Phase, Source Parameters and Source Type on Electrospray Ionization Efficiency in Negative Ion Mode: Influence of Mobile Phase in ESI/MS. *J. Mass Spectrom.* **2016**, *51* (8), 596–601. https://doi.org/10.1002/jms.3790.

(53) Zhou, S.; Hamburger, M. Effects of Solvent Composition of Molecular Ion Response in Electrospray Mass Spectrometry: Investigation of the Ionization Process. *Rapid Commun. Mass Spectrom.* **1995**, *9* (15), 1516–1521.

(54) Kiontke, A.; Oliveira-Birkmeier, A.; Opitz, A.; Birkemeyer, C. Electrospray Ionization Efficiency Is Dependent on Different Molecular Descriptors with Respect to Solvent PH and Instrumental Configuration. *PLOS ONE* **2016**, *11* (12), e0167502. https://doi.org/10.1371/journal.pone.0167502.

(55) Leito, I.; Herodes, K.; Huopolainen, M.; Virro, K.; Künnapas, A.; Kruve, A.; Tanner, R. Towards the Electrospray Ionization Mass Spectrometry Ionization Efficiency Scale of Organic Compounds. *Rapid Commun. Mass Spectrom.* **2008**, *22* (3), 379–384. https://doi.org/10.1002/rcm.3371.

(56) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Efficiency Without Chemical Standards. *Anal. Chem.* **2009**, *81* (7), 2506–2515. https://doi.org/10.1021/ac802272u.

(57) Liigand, J.; Laaniste, A.; Kruve, A. PH Effects on Electrospray Ionization Efficiency. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (3), 461–469. https://doi.org/10.1007/s13361-016-1563-1.

(58) Mansoori, B. A.; Volmer, D. A.; Boyd, R. K. 'Wrong-Way-Round' Electrospray Ionization of Amino Acids. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 11.

(59) Cramer, C. J.; Johnson, J. L.; Kamel, A. M. Prediction of Mass Spectral Response Factors from Predicted Chemometric Data for Druglike Molecules. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (2), 278–285. https://doi.org/10.1007/s13361-016-1536-4.

(60) Nguyen, T. B.; Nizkorodov, S. A.; Laskin, A.; Laskin, J. An Approach toward Quantification of Organic Compounds in Complex Environmental Samples Using High-Resolution Electrospray Ionization Mass Spectrometry. *Anal Methods* **2013**, *5* (1), 72–80. https://doi.org/10.1039/C2AY25682G.

(61) Mandra, V. J.; Kouskoura, M. G.; Markopoulou, C. K. Using the Partial Least Squares Method to Model the Electrospray Ionization Response Produced by Small Pharmaceutical Molecules in Positive Mode: Modelling Positive Electrospray Ionization Response. *Rapid Commun. Mass Spectrom.* **2015**, *29* (18), 1661–1675. https://doi.org/10.1002/rcm.7263.

(62) Amad, M. H.; Cech, N. B.; Jackson, G. S.; Enke, C. G. Importance of Gas-Phase Proton Affinities in Determining the Electrospray Ionization Response for Analytes and Solvents. *J. Mass Spectrom.* **2000**, *35*, 784–789.

(63) Hermans, J.; Ongay, S.; Markov, V.; Bischoff, R. Physicochemical Parameters Affecting the Electrospray Ionization Efficiency of Amino Acids after Acylation. *Anal. Chem.* **2017**, *89* (17), 9159–9166. https://doi.org/10.1021/acs.analchem.7b01899.

(64) Alymatiri, C. M.; Kouskoura, M. G.; Markopoulou, C. K. Decoding the Signal Response of Steroids in Electrospray Ionization Mode (ESI-MS). *Anal. Methods* **2015**, *7*, 10433–10444. https://doi.org/10.1039/C5AY02839F.

(65) Ghosh, B.; Jones, A. D. Dependence of Negative-Mode Electrospray Ionization Response Factors on Mobile Phase Composition and Molecular Structure for Newly-Authenticated Neutral Acylsucrose Metabolites. *The Analyst* **2015**, *140* (19), 6522–6531. https://doi.org/10.1039/C4AN02124J.

(66) Bökman, C. F.; Bylund, D.; Markides, K. E.; Sjöberg, P. J. R. Relating Chromatographic Retention and Electrophoretic Mobility to the Ion Distribution within Electrosprayed Droplets. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (3), 318–324. https://doi.org/10.1016/j.jasms.2005.11.006.

(67) Wu, L.; Wu, Y.; Shen, H.; Gong, P.; Cao, L.; Wang, G.; Hao, H. Quantitative Structure–Ion Intensity Relationship Strategy to the Prediction of Absolute Levels without Authentic Standards. *Anal. Chim. Acta* **2013**, *794*, 67–75. https://doi.org/10.1016/j.aca.2013.07.034.

(68) Golubović, J.; Birkemeyer, C.; Protić, A.; Otašević, B.; Zečević, M. Structure–Response Relationship in Electrospray Ionization-Mass Spectrometry of Sartans by Artificial Neural Networks. *J. Chromatogr. A* **2016**, *1438*, 123–132. https://doi.org/10.1016/j.chroma.2016.02.021.

(69) Kruve, A.; Kaupmees, K.; Liigand, J.; Oss, M.; Leito, I. Sodium Adduct Formation Efficiency in ESI Source: Sodium Adduct Formation Efficiency in ESI Source. *J. Mass Spectrom.* **2013**, *48* (6), 695–702. https://doi.org/10.1002/jms.3218.

(70) Liigand, J. Standard Substance Free Quantification for ESI/MS Analysis Based on Ionization Efficiency Scale, University of Tartu: Tartu, 2019.

(71) Kebarle, P.; Tang, L. FROM IONS IN SOLUTION TO IONS IN THE GAS PHASE. *Anal. Chem.* **1993**, *65* (22), 972A-986A. https://doi.org/10.1021/ac00070a715.

(72) Henriksen, T.; Juhler, R. K.; Svensmark, B.; Cech, N. B. The Relative Influences of Acidity and Polarity on Responsiveness of Small Organic Molecules to Analysis with Negative Ion Electrospray Ionization Mass Spectrometry (ESI-MS). *J. Am. Soc. Mass Spectrom.* **2005**, *16* (4), 446–455. https://doi.org/10.1016/j.jasms.2004.11.021.

(73) Golubović, J.; Birkemeyer, C.; Protić, A.; Otašević, B.; Zečević, M. Structure–Response Relationship in Electrospray Ionization-Mass Spectrometry of Sartans by Artificial Neural Networks. *J. Chromatogr. A* **2016**, *1438*, 123–132. https://doi.org/10.1016/j.chroma.2016.02.021.

(74) Huffman, B. A.; Poltash, M. L.; Hughey, C. A. Effect of Polar Protic and Polar Aprotic Solvents on Negative-Ion Electrospray Ionization and Chromatographic Separation of Small Acidic Molecules. *Anal. Chem.* **2012**, *84* (22), 9942–9950. https://doi.org/10.1021/ac302397b.

(75) Cech, N. B.; Krone, J. R.; Enke, C. G. Predicting Electrospray Response from Chromatographic Retention Time. *Anal. Chem.* **2001**, *73* (2), 208–213. https://doi.org/10.1021/ac0006019.

(76) Ghosh, B.; Jones, A. D. Dependence of Negative-Mode Electrospray Ionization Response Factors on Mobile Phase Composition and Molecular Structure for Newly-Authenticated Neutral Acylsucrose Metabolites. *The Analyst* **2015**, *140* (19), 6522–6531. https://doi.org/10.1039/C4AN02124J.

(77) Amad, M. H.; Cech, N. B.; Jackson, G. S.; Enke, C. G. Importance of Gas-Phase Proton Affinities in Determining the Electrospray Ionization Response for Analytes and Solvents. *J. Mass Spectrom.* **2000**, *35* (7), 784–789. https://doi.org/10.1002/1096-9888(200007)35:7<784::AID-JMS17>3.0.CO;2-Q.

(78) Alymatiri, C. M.; Kouskoura, M. G.; Markopoulou, C. K. Decoding the Signal Response of Steroids in Electrospray Ionization Mode (ESI-MS). *Anal Methods* **2015**, *7* (24), 10433–10444. https://doi.org/10.1039/C5AY02839F.

(79) Gioumouxouzis, C. I.; Kouskoura, M. G.; Markopoulou, C. K. Negative Electrospray Ionization Mode in Mass Spectrometry: A New Perspective via Modeling. *J. Chromatogr. B* **2015**, *998–999*, 97–105. https://doi.org/10.1016/j.jchromb.2015.06.009.

(80) Mandra, V. J.; Kouskoura, M. G.; Markopoulou, C. K. Using the Partial Least Squares Method to Model the Electrospray Ionization Response Produced by Small Pharmaceutical Molecules in Positive Mode: Modelling Positive Electrospray Ionization Response. *Rapid Commun. Mass Spectrom.* **2015**, *29* (18), 1661–1675. https://doi.org/10.1002/rcm.7263.

(81) Liigand, J.; Wang, T.; Kellogg, J. J.; Smedsgaard, J.; Cech, N. B.; Kruve, A. Quantifying the Unquantifiable: Estimating the Concentration of All Compounds Identified with Liquid Chromatography Electrospray Ionization Mass Spectrometry. *Nat. Methods Submitted for publishing*.

(82) Ehrmann, B. M.; Henriksen, T.; Cech, N. B. Relative Importance of Basicity in the Gas Phase and in Solution for Determining Selectivity in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 719–728. https://doi.org/10.1016/j.jasms.2008.01.003.

(83) Kiontke, A.; Oliveira-Birkmeier, A.; Opitz, A.; Birkemeyer, C. Electrospray Ionization Efficiency Is Dependent on Different Molecular Descriptors with Respect to Solvent PH and Instrumental Configuration. *PLOS ONE* **2016**, *11* (12), e0167502. https://doi.org/10.1371/journal.pone.0167502.

(84) Douglass, K. A.; Venter, A. R. Investigating the Role of Adducts in Protein Supercharging with Sulfolane. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (3), 489–497. https://doi.org/10.1007/s13361-011-0319-1.

(85) Wong, S. F.; Meng, C. K.; Fenn, J. B. Multiple Charging in Electrospray Ionization of Poly (Ethylene Glycols). *J. Phys. Chem.* **1988**, *92* (2), 546–550.

(86) Schnier, P. D.; Gross, D. S.; Williams, E. R. On the Maximum Charge State and Proton Transfer Reactivity of Peptide and Protein Ions Formed By Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1086–1097.

(87) Smith, R. D.; Loo, J. A.; Loo, R. R. O.; Busman, M.; Udseth, H. R. Principles and Practice of Electrospray Ionization- Mass Spectrometry for Large Polypeptides and Proteins. *Mass Spectrom. Rev.* **1991**, *10*, 359–451.

(88) Felitsyn, N.; Peschke, M.; Kebarle, P. Origin and Number of Charges Observed on Multiply-Protonated Native Proteins Produced by ESI. *Int. J. Mass Spectrom.* **2002**, *219* (1), 39–62.

(89) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. Effect of Reducing Disulfide-Containing Proteins on Electrospray Ionization Mass Spectra. *Anal. Chem.* **1990**, *62* (7), 693–698.

(90) Loo, J. A.; Udseth, H. R.; Smith, R. D. Solvent Effects on the Charge Distribution Observed with Electrospray Ionization-Mass Spectrometry of Large Molecules. *Biol. Mass Spectrom.* **1988**, *17* (5), 411–414.

(91) Iavarone, A. T.; Jurchen, J. C.; Williams, E. R. Effects of Solvent on the Maximum Charge State and Charge State Distribution of Protein Ions Produced by Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (11), 976–985.

(92) Iavarone, A. T.; Jurchen, J. C.; Williams, E. R. Supercharged Protein and Peptide Ions Formed by Electrospray Ionization. *Anal. Chem.* **2001**, *73* (7), 1455–1460. https://doi.org/10.1021/ac001251t.

(93) Iavarone, A. T.; Williams, E. R. Mechanism of Charging and Supercharging Molecules in Electrospray Ionization. *J. Am. Chem. Soc.* **2003**, *125* (8), 2319–2327. https://doi.org/10.1021/ja021202t.

(94) Sterling, H. J.; Daly, M. P.; Feld, G. K.; Thoren, K. L.; Kintzer, A. F.; Krantz, B. A.; Williams, E. R. Effects of Supercharging Reagents on Noncovalent Complex Structure in Electrospray Ionization from Aqueous Solutions. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (10), 1762–1774. https://doi.org/10.1016/j.jasms.2010.06.012.

(95) Sterling, H. J.; Kintzer, A. F.; Feld, G. K.; Cassou, C. A.; Krantz, B. A.; Williams, E. R. Supercharging Protein Complexes from Aqueous Solution Disrupts Their Native Conformations. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (2), 191–200. https://doi.org/10.1007/s13361-011-0301-y.

(96) Ganisl, B.; Taucher, M.; Riml, C.; Breuker, K. Charge as You like! Efficient Manipulation of Negative Ion Net Charge in Electrospray Ionization of Proteins and Nucleic Acids. *Eur. J. Mass Spectrom.* **2011**, *17*, 333–343.

(97) Douglass, K.; Venter, A. On the Role of a Direct Interaction between Protein Ions and Solvent Additives during Protein Supercharging by Electrospray Ionization Mass Spectrometry. *Eur. J. Mass Spectrom.* **2015**, *21* (3), 641. https://doi.org/10.1255/ejms.1360.

(98) Pieke, E. N.; Granby, K.; Trier, X.; Smedsgaard, J. A Framework to Estimate Concentrations of Potentially Unknown Substances by Semi-Quantification in Liquid Chromatography Electrospray Ionization Mass Spectrometry. *Anal. Chim. Acta* **2017**, *975*, 30–41. https://doi.org/10.1016/j.aca.2017.03.054.

(99) Hollender, J.; Schymanski, E. L.; Singer, H. P.; Ferguson, P. L. Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go? *Environ. Sci. Technol.* **2017**, *51* (20), 11505–11512. https://doi.org/10.1021/acs.est.7b02184.

(100) Hatsis, P.; Waters, N. J.; Argikar, U. A. Implications for Metabolite Quantification by Mass Spectrometry in the Absence of Authentic Standards. *Drug Metab. Dispos.* **2017**, *45* (5), 492–496. https://doi.org/10.1124/dmd.117.075259.

(101) Heering, A. Experimental Realization and Applications of the Unified Acidity Scale, University of Tartu: Tartu, 2017.

(102) Gibson, S. C.; Feigerle, C. S.; Cook, K. D. Fluorometric Measurement and Modeling of Droplet Temperature Changes in an Electrospray Plume. *Anal. Chem.* **2014**, *86* (1), 464–472. https://doi.org/10.1021/ac402364g.

(103) Klamt, A. *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design.*; Elsevier Science Ltd: Amsterdam, 2005.

(104) ACE and JChem acidity and basicity calculator https://epoch.uky.edu/ace/public/pKa.jsp.

(105) *TURBOMOLE V6.4, 2012, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; Available from Http://Www.Turbomole.Com.*

(106) Eckert, F.; Klamt, A. *COSMOtherm, Version C3.0, Release 13.01; COSMOlogic GmbH&CoKG, Leverkusen, Germany, 2013; Available from Http://Www.Cosmologic.De/.*; 2013.

(107) Kaupmees, K.; Kaljurand, I.; Leito, I. Influence of Water Content on the Acidities in Acetonitrile. Quantifying Charge Delocalization in Anions. *J. Phys. Chem. A* **2010**, *114* (43), 11788–11793. https://doi.org/10.1021/jp105670t.

(108) *R Core Team (2018). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.*

(109) *Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.*

(110) Miller, J. N.; Miller, J. C. *Statistics and Chemometrics for Analytical Chemistry*, 6. ed.; Prentice Hall: Harlow, 2010.

(111) Yap, C. W. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. *J. Comput. Chem.* **2011**, *32* (7), 1466–1474. https://doi.org/10.1002/jcc.21707.

(112) Snyder, L. R.; Kirkland, J. J.; Dolan, J. W. *Introduction to Modern Liquid Chromatography*, 3rd ed.; John Wiley & Sons, Inc.: New Jersey, 2009.

(113) Rudakov, O. B.; Belyaev, D. S.; Khorokhordina, E. A.; Podolina, E. A. Surface Tension of Binary Mobile Phases for Liquid Chromatography. *Russ. J. Phys. Chem. A* **2007**, *81* (3), 366–369. https://doi.org/10.1134/S0036024407030107.

(114) Katz, E.; Eksteen, R.; Schoenmakers, P.; Miller, N. *Handbook of HPLC*; Chromatographic science series; M. Dekker: New York, 1998.

(115) Deng, H.; Runger, G. Feature Selection via Regularized Trees. *ArXiv12011587 Cs Stat* **2012**.

(116) Dong, J.; Cao, D.-S.; Miao, H.-Y.; Liu, S.; Deng, B.-C.; Yun, Y.-H.; Wang, N.-N.; Lu, A.-P.; Zeng, W.-B.; Chen, A. F. ChemDes: An Integrated Web-Based Platform for Molecular Descriptor and Fingerprint Computation. *J. Cheminformatics* **2015**, *7* (1). https://doi.org/10.1186/s13321-015-0109-z.

(117) Teesch, L. M.; Orlando, R. C.; Adams, J. Location of the Alkali Metal Ion in Gas-Phase Peptide Complexes. *J. Am. Chem. Soc.* **1991**, *113* (10), 3668–3675.

(118) Kamariotis, A.; Boyarkin, O. V.; Mercier, S. R.; Beck, R. D.; Bush, M. F.; Williams, E. R.; Rizzo, T. R. Infrared Spectroscopy of Hydrated Amino Acids in the Gas Phase: Protonated and Lithiated Valine. *J. Am. Chem. Soc.* **2006**, *128* (3), 905–916. https://doi.org/10.1021/ja056079v.

(119) O'Brien, J. T.; Williams, E. R. Effects of Ions on Hydrogen-Bonding Water Networks in Large Aqueous Nanodrops. *J. Am. Chem. Soc.* **2012**, *134* (24), 10228–10236. https://doi.org/10.1021/ja303191r.

(120) Aikens, C. M.; Gordon, M. S. Incremental Solvation of Nonionized and Zwitterionic Glycine. *J. Am. Chem. Soc.* **2006**, *128* (39), 12835–12850. https://doi.org/10.1021/ja062842p.

(121) Kruve, A.; Kaupmees, K. Predicting ESI/MS Signal Change for Anions in Different Solvents. *Anal. Chem.* **2017**, *89* (9), 5079–5086. https://doi.org/10.1021/acs.analchem.7b00595.

(122) Liigand, J.; Kruve, A.; Liigand, P.; Laaniste, A.; Girod, M.; Antoine, R.; Leito, I. Transferability of the Electrospray Ionization Efficiency Scale between Different Instruments. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (11), 1923–1930. https://doi.org/10.1007/s13361-015-1219-6.

(123) Kruve, A.; Auling, R.; Herodes, K.; Leito, I. Study of Liquid Chromatography/Electrospray Ionization Mass Spectrometry Matrix Effect on the Example of Glyphosate Analysis from Cereals: Matrix Effect on the Example of Glyphosate Analysis. *Rapid Commun. Mass Spectrom.* **2011**, *25* (21), 3252–3258. https://doi.org/10.1002/rcm.5222.

(124) Furey, A.; Moriarty, M.; Bane, V.; Kinsella, B.; Lehane, M. Ion Suppression; A Critical Review on Causes, Evaluation, Prevention and Applications. *Talanta* **2013**, *115*, 104–122. https://doi.org/10.1016/j.talanta.2013.03.048.

(125) Manini, P.; Andreoli, R.; Mutti, A. Application of Liquid Chromatography–Mass Spectrometry to Biomonitoring of Exposure to Industrial Chemicals. *Toxicol. Lett.* **2006**, *162* (2–3), 202–210. https://doi.org/10.1016/j.toxlet.2005.09.023.

(126) Villagrasa, M.; Guillamón, M.; Eljarrat, E.; Barceló, D. Matrix Effect in Liquid Chromatography–Electrospray Ionization Mass Spectrometry Analysis of Benzoxazinoid Derivatives in Plant Material. *J. Chromatogr. A* **2007**, *1157* (1–2), 108–114. https://doi.org/10.1016/j.chroma.2007.04.040.

(127) Little, J.; Wempe, M.; Buchanan, C. Liquid Chromatography–Mass Spectrometry/Mass Spectrometry Method Development for Drug Metabolism Studies: Examining Lipid Matrix Ionization Effects in Plasma. *J. Chromatogr. B* **2006**, *833* (2), 219–230. https://doi.org/10.1016/j.jchromb.2006.02.011.

(128) Dams, R.; Huestis, M. A.; Lambert, W. E.; Murphy, C. M. Matrix Effect in Bio-Analysis of Illicit Drugs with LC-MS/MS: Influence of Ionization Type, Sample Preparation, and Biofluid. *J. Am. Soc. Mass Spectrom.* **2003**, *14* (11), 1290–1294. https://doi.org/10.1016/S1044-0305(03)00574-9.

(129) Lanckmans, K.; Van Eeckhaut, A.; Sarre, S.; Smolders, I.; Michotte, Y. Capillary and Nano-Liquid Chromatography–Tandem Mass Spectrometry for the Quantification of Small Molecules in Microdialysis Samples: Comparison with Microbore Dimensions. *J. Chromatogr. A* **2006**, *1131* (1–2), 166–175. https://doi.org/10.1016/j.chroma.2006.07.090.

(130) King, R.; Bonfiglio, R.; Fernandez-Metzler, C.; Miller-Stein, C.; Olah, T. Mechanistic Investigation of Ionization Suppression in Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (11), 942–950. https://doi.org/10.1016/S1044-0305(00)00163-X.

(131) Zampieri, M.; Sekar, K.; Zamboni, N.; Sauer, U. Frontiers of High-Throughput Metabolomics. *Curr. Opin. Chem. Biol.* **2017**, *36*, 15–23. https://doi.org/10.1016/j.cbpa.2016.12.006.

(132) Himmel, D.; Goll, S. K.; Leito, I.; Krossing, I. Anchor Points for the Unified Brønsted Acidity Scale: The RCCC Model for the Calculation of Standard Gibbs Energies of Proton Solvation in Eleven Representative Liquid Media. *Chem. - Eur. J.* **2011**, *17* (21), 5808–5826. https://doi.org/10.1002/chem.201003164.

(133) Suu, A.; Jalukse, L.; Liigand, J.; Kruve, A.; Himmel, D.; Krossing, I.; Rosés, M.; Leito, I. Unified PH Values of Liquid Chromatography Mobile Phases. *Anal. Chem.* **2015**, *87* (5), 2623–2630. https://doi.org/10.1021/ac504692m.

(134) Van Berkel, G. J.; Zhou, F.; Aronson, J. T. Changes in Bulk Solution PH Caused by the Inherent Controlled-Current Electrolytic Process of an Electrospray Ion Source. *Int. J. Mass Spectrom. Ion Process.* **1997**, *162*, 55–67.

(135) Mora, J. F. de la; Van Berkel, G. J.; Enke, C. G.; Cole, R. B.; Martinez-Sanchez, M.; Fenn, J. B. Electrochemical Processes in Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2000**, *35* (8), 939–952.

(136) Seo, J.; Warnke, S.; Gewinner, S.; Schöllkopf, W.; Bowers, M. T.; Pagel, K.; von Helden, G. The Impact of Environment and Resonance Effects on the Site of Protonation of Aminobenzoic Acid Derivatives. *Phys Chem Chem Phys* **2016**, *18* (36), 25474–25482. https://doi.org/10.1039/C6CP04941A.

(137) Kebarle, P.; Peschke, M. On the Mechanisms by Which the Charged Droplets Produced by Electrospray Lead to Gas Phase Ions. *Anal. Chim. Acta* **2000**, *406* (1), 11–35.

(138) Peschke, M.; Verkerk, U. H.; Kebarle, P. Features of the ESI Mechanism That Affect the Observation of Multiply Charged Noncovalent Protein Complexes and the Determination of the Association Constant by the Titration Method. *J. Am. Soc. Mass Spectrom.* **2004**, *15* (10), 1424–1434.

(139) Kaupmees, K.; Kaljurand, I.; Leito, I. Influence of Water Content on Basicities in Acetonitrile. *J. Solut. Chem.* **2014**, *43* (7), 1270–1281. https://doi.org/10.1007/s10953-014-0201-4.

(140) Tse, Y.-L. S.; Chen, C.; Lindberg, G. E.; Kumar, R.; Voth, G. A. Propensity of Hydrated Excess Protons and Hydroxide Anions for the Air–Water Interface. *J. Am. Chem. Soc.* **2015**, *137* (39), 12610–12616. https://doi.org/10.1021/jacs.5b07232.

(141) Tian, Z.; Kass, S. R. Does Electrospray Ionization Produce Gas-Phase or Liquid-Phase Structures? *J. Am. Chem. Soc.* **2008**, *130* (33), 10842–10843. https://doi.org/10.1021/ja802088u.

(142) Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; et al. DrugBank 5.0: A Major Update to the DrugBank Database for 2018. *Nucleic Acids Res.* **2018**, *46* (D1), D1074–D1082. https://doi.org/10.1093/nar/gkx1037.

(143) Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; et al. HMDB 4.0: The Human Metabolome Database for 2018. *Nucleic Acids Res.* **2018**, *46* (Database issue), D608–D617. https://doi.org/10.1093/nar/gkx1089.

(144) Liigand, J.; Kruve, A.; Liigand, P.; Laaniste, A.; Girod, M.; Antoine, R.; Leito, I. Transferability of the Electrospray Ionization Efficiency Scale between Different Instruments. *J. Am. Soc. Mass Spectrom.* **2015**, *26* (11), 1923–1930. https://doi.org/10.1007/s13361-015-1219-6.

(145) Periat, A.; Kohler, I.; Bugey, A.; Bieri, S.; Versace, F.; Staub, C.; Guillarme, D. Hydrophilic Interaction Chromatography versus Reversed Phase Liquid Chromatography Coupled to Mass Spectrometry: Effect of Electrospray Ionization Source Geometry on Sensitivity. *J. Chromatogr. A* **2014**, *1356*, 211–220. https://doi.org/10.1016/j.chroma.2014.06.066.

(146) Wilkinson, M. D.; Dumontier, M.; Aalbersberg, Ij. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; et al. The FAIR Guiding Principles for Scientific Data Management and Steward-ship. *Sci. Data* **2016**, *3*. https://doi.org/10.1038/sdata.2016.18.

(147) Bedner, M.; Duewer, D. L. Dynamic Calibration Approach for Determining Catechins and Gallic Acid in Green Tea Using LC–ESI/MS. *Anal. Chem.* **2011**, *83* (16), 6169–6176. https://doi.org/10.1021/ac200372d.

(148) Byrdwell, W. C. Quadruple Parallel Mass Spectrometry for Analysis of Vitamin D and Triacylglycerols in a Dietary Supplement. *J. Chromatogr. A* **2013**, *1320*, 48–65. https://doi.org/10.1016/j.chroma.2013.10.031.

(149) Chalcraft, K. R.; Lee, R.; Mills, C.; Britz-McKibbin, P. Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectro-metry: Predicting Ionization Efficiency Without Chemical Standards. *Anal. Chem.* **2009**, *81* (7), 2506–2515. https://doi.org/10.1021/ac802272u.

(150) Cífková, E.; Holčapek, M.; Lísa, M.; Ovčačíková, M.; Lyčka, A.; Lynen, F.; Sandra, P. Nontargeted Quantitation of Lipid Classes Using Hydrophilic Interaction Liquid Chromatography–Electrospray Ionization Mass Spectrometry with Single Internal Standard and Response Factor Approach. *Anal. Chem.* **2012**, *84* (22), 10064–10070. https://doi.org/10.1021/ac3024476.

(151) Leitner, A.; Emmert, J.; Boerner, K.; Lindner, W. Influence of Solvent Additive Composition on Chromatographic Separation and Sodium Adduct Formation of Peptides in HPLC–ESI MS. *Chromatographia* **2007**, *65* (11–12), 649–653. https://doi.org/10.1365/s10337-007-0219-5.

(152) Raji, M. A.; Fryčák, P.; Temiyasathit, C.; Kim, S. B.; Mavromaras, G.; Ahn, J.-M.; Schug, K. A. Using Multivariate Statistical Methods to Model the Electro-spray Ionization Response of GXG Tripeptides Based on Multiple Physico-chemical Parameters. *Rapid Commun. Mass Spectrom.* **2009**, *23* (14), 2221–2232. https://doi.org/10.1002/rcm.4141.

(153) Stavenhagen, K.; Hinneburg, H.; Thaysen-Andersen, M.; Hartmann, L.; Silva, D. V.; Fuchser, J.; Kaspar, S.; Rapp, E.; Seeberger, P. H.; Kolarich, D. Quantitative Mapping of Glycoprotein Micro-Heterogeneity and Macro-Heterogeneity: An Evaluation of Mass Spectrometry Signal Strengths Using Synthetic Peptides and Glycopeptides: Glycopeptide Ionisation Strength. *J. Mass Spectrom.* **2013**, *48* (6), 627–639. https://doi.org/10.1002/jms.3210.

(154) Zendong, Z.; Sibat, M.; Herrenknecht, C.; Hess, P.; McCarron, P. Relative Molar Response of Lipophilic Marine Algal Toxins in Liquid Chromatography/ Electrospray Ionization Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (17), 1453–1461. https://doi.org/10.1002/rcm.7918.

(155) Yang, W.-C.; Mirzaei, H.; Liu, X.; Regnier, F. E. Enhancement of Amino Acid Detection and Quantification by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (13), 4702–4708. https://doi.org/10.1021/ac0600510.

(156) Kamga, A. W.; Behar, F.; Hatcher, P. G. Quantitative Analysis of Long Chain Fatty Acids Present in a Type I Kerogen Using Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: Compared with BF3/MeOH Methylation/GC-FID. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (5), 880–890. https://doi.org/10.1007/s13361-014-0851-x.

(157) Lin, Y.; Liu, R.; Hu, F.; Liu, R.; Ruan, T.; Jiang, G. Simultaneous Qualitative and Quantitative Analysis of Fluoroalkyl Sulfonates in Riverine Water by Liquid Chromatography Coupled with Orbitrap High Resolution Mass Spectrometry. *J. Chromatogr. A* **2016**, *1435*, 66–74. https://doi.org/10.1016/j.chroma.2016.01.039.

(158) Tang, W.-T.; Fang, M.-F.; Liu, X.; Yue, M. Simultaneous Quantitative and Qualitative Analysis of Flavonoids from Ultraviolet-B Radiation in Leaves and Roots of *Scutellaria Baicalensis* Georgi Using LC-UV-ESI-Q/TOF/MS. *J. Anal. Methods Chem.* **2014**, *2014*, 1–9. https://doi.org/10.1155/2014/643879.

(159) Guha, R. Chemical Informatics Functionality in R. *J. Stat. Softw.* **2007**, *18* (1), 1–16. https://doi.org/10.18637/jss.v018.i05.

(160) Fraczkiewicz, R.; Lobell, M.; Göller, A. H.; Krenz, U.; Schoenneis, R.; Clark, R. D.; Hillisch, A. Best of Both Worlds: Combining Pharma Data and State of the Art Modeling Technology To Improve *in Silico* p K_a Prediction. *J. Chem. Inf. Model.* **2015**, *55* (2), 389–397. https://doi.org/10.1021/ci500585w.

(161) *ADMET Predictor, Version 9, Simulation Plus, Inc.*

(162) Ehrmann, B. M.; Henriksen, T.; Cech, N. B. Relative Importance of Basicity in the Gas Phase and in Solution for Determining Selectivity in Electrospray Ionization Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (5), 719–728. https://doi.org/10.1016/j.jasms.2008.01.003.

(163) Koivusalo, M.; Haimi, P.; Heikinheimo, L.; Kostiainen, R.; Somerharju, P. Quantitative Determination of Phospholipid Compositions by ESI-MS: Effects of Acyl Chain Length, Unsaturation, and Lipid Concentration on Instrument Response. *J. Lipid Res.* **2001**, *42* (4), 663–672.

(164) Buck, R.; Rondinini, S.; Covington, A.; Baucke, F.; Brett, C.; Camoes, M.; Milton, M.; Mussini, T.; Naumann, R.; Pratt, K.; et al. Measurement of PH. Definition, Standards, and Procedures. *Pure Appl. Chem.* **2002**, *74* (11), 2169–2200. https://doi.org/10.1351/pac200274112169.

(165) Rosés, M. Determination of the PH of Binary Mobile Phases for Reversed-Phase Liquid Chromatography. *J. Chromatogr. A* **2004**, *1037* (1–2), 283–298. https://doi.org/10.1016/j.chroma.2003.12.063.

(166) Mussini, T.; Covington, A. K.; Longhi, P.; Rondinini, S. Criteria for Standardization of pH Measurements in Organic Solvents and Water + Organic Solvent Mixtures of Moderate to High Permittivities. *Pure Appl. Chem.* **1985**, *57* (6), 865–876. https://doi.org/10.1351/pac198557060865.

ACKNOWLEDGMENTS

APPENDICES

Appendix 1

Determination of ionization degrees of
trans-3(3-pyridyl)acrylic acid

Chemical shift (δ) of protons is determined by the chemical environment of the proton, therefore chemical shifts reflect the structure of the compound and can be used to calculate the degree of ionization in solution. Determination of degrees of ionization was performed on 700 MHz NMR Bruker Avance II 700 NMR spectrometer. ^{1}H NMR spectra of *trans*-3(3-pyridyl)acrylic acid were registered in acetonitrile-d_3:H_2O (80/20 v/v) mixtures at aqueous pH 1.0, 3.0, 4.0 and 10.0. Degrees of ionization at different pH for the carboxylic group was calculated from chemical shift of E proton and for pyridinium group by averaging degrees of ionizations calculated from B and C protons (Figure A1-1 and Table A1-1).
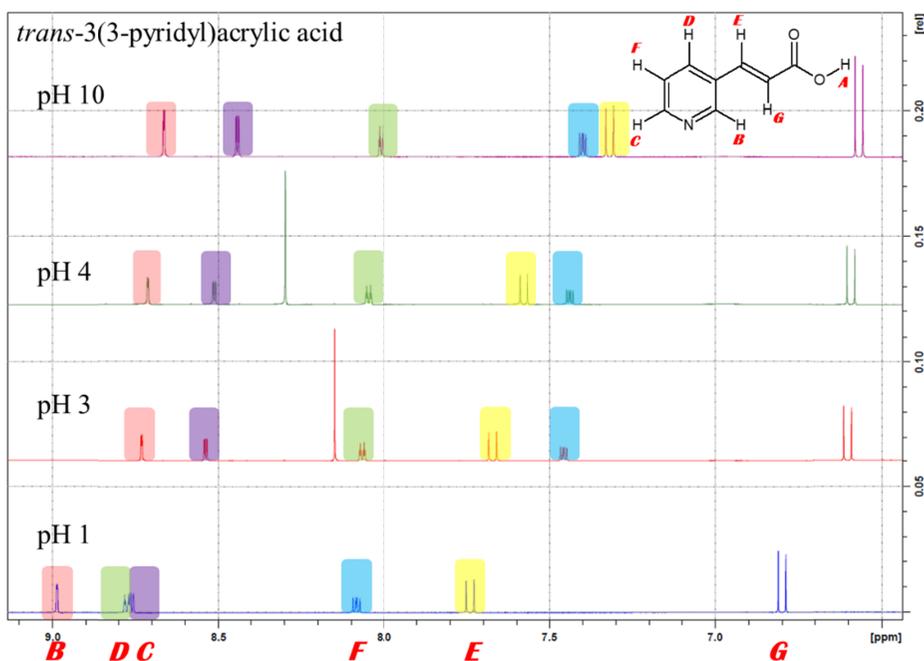
Figure A1-1. ^{1}H NMR spectra of *trans*-3(3-pyridyl)acrylic acid in acetonitrile-d_3:H_2O (80/20 v/v) mixtures at aqueous pH 1.0, 3.0, 4.0 and 10.0.

Table A1-1. Chemical shifts of protons (see reference of subscripts in Figure A1-1) at different pH values.

pH	δ_B	δ_D	δ_C	δ_F	δ_E	δ_G
10.0	8.664	8.010	8.443	7.399	7.319	6.567
4.0	8.713	8.048	8.513	7.439	7.579	6.592
3.0	8.732	8.069	8.539	7.457	7.673	6.601
1.0	8.987	8.776	8.760	8.085	7.740	6.800

Spectrophotometric titration was carried out for verification of the results. First, spectra in a solvent where the aqueous phase consisted of 0.1 M HCl (pH 1) solution was recorded – in that solution, both groups are protonated. After that, the spectra at different aqueous phase pH values were recorded until end-point was reached (no spectral changes detected – both groups deprotonated) using ammonium hydroxide as a titrant. As the spectra recorded at pH 3.91 is a complex mixture of spectra of multiple forms it is impossible to distinguish between the degrees of protonation for the acidic and basic site separately. The spectrum at pH 3.91 still enables to see that both of the groups are protonated to some extent confirming the NMR experiment result in a qualitative way. Double beam spectrophotometer Thermo Nicolet Evolution 300 was used. Scan speed was intelliscan mode (from 120 nm/min to 1200 nm/min, scan speed changes depending on how strongly absorbance depends on the wavelength), bandwidth was 1.5 nm and data registration interval was 1 nm. Results of the UV-Vis measurements confirmed the results obtained by ^{1}H NMR experiment.
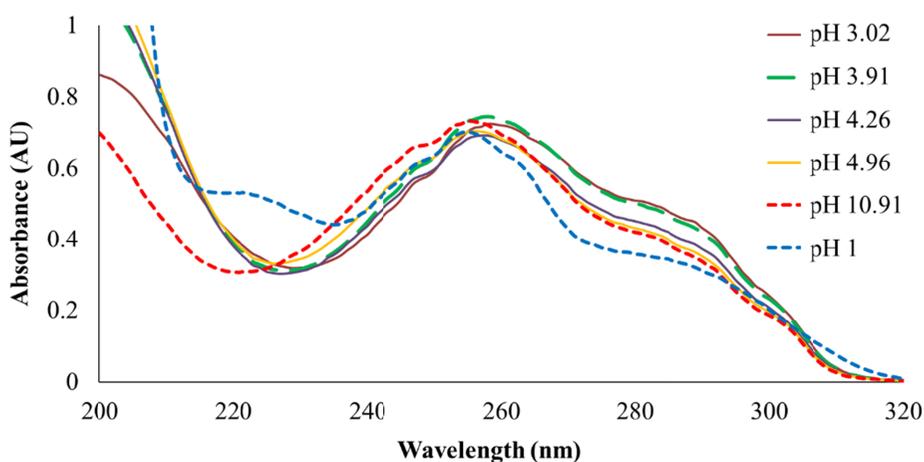
Figure A1-2. UV-Vis spectra of *trans*-3(3-pyridyl)acrylic acid in acetonitrile:H_2O (80/20 v/v) mixtures at aqueous pH 1.0 (0.1M HCl), 3.02, 4.26, 4.96 and 10.91 (0.1% NH_3).

Appendix 2

Absolute pH theory

The conventional pH is defined by IUPAC[164] as given in Eq. A2-1:

$$\text{pH} = -\log a_H \tag{A2-1}$$

where a_H is the relative activity in molal scale. This means that the zero point (standard state) of pH scale is activity of 1 mol/kg of H^+_{solv} *in the given solvent* (or solvent mixture). Therefore, every solvent or solvent mixture has its own pH scale and none of these scales are comparable due to unknown shifts of the zero points. To add confusion to these different pH scales there is the question of calibration, which leads to another possibility to define or name pH scales as $_s^s\text{pH}$ or $_w^s\text{pH}$, where subscript shows calibration and superscript measurement medium (s is solvent and w is water). This question is discussed in detail by Rosés.[165]

Recently Himmel et al. proposed a unified acidity scale that is based on the absolute standard chemical potential $\mu_{abs}(H^+)$ of the solvated proton [132] and where the zero point of the scale is the $\mu_{abs}(H^+)$ of the proton in the gas phase, which is arbitrarily set to 0 kJ mol^{-1}. Importantly, this zero point of the scale is universal to all possible media, thereby enabling comparison of the acidities of any given media on one scale. In any solvent, the chemical potential of the proton is decreased (becomes more negative) by solvation. The more negative is the proton's chemical potential the lower is its activity and consequently the acidity of the solution. The absolute acidity (absolute pH, pH_{abs}) can be calculated via the solvated proton's chemical potentials as follows:

$$\text{pH}_{abs} = -\frac{\mu_{abs}(H^+, \text{solv})}{RT \ln 10} \tag{A2-2}$$

where R is the molar gas constant, and T is the absolute temperature and $\mu_{abs}(H^+,\text{solv})$ is calculated as given in Eq. A2-3:

$$\mu_{abs}(H^+, \text{solv}) = \Delta_{\text{solv}}G^{\ominus}(H^+) - [\text{pH} \times RT \ln 10] \tag{A2-3}$$

where $\Delta_{\text{solv}}G^{\ominus}(H^+)$ is Gibbs energy of solvation and *pH* is the conventional pH. This approach is also fully universal in the sense that it does not set any limitations to the solvation sphere of the proton. At the same time, the properties of the solvation sphere (extent of solvation) of the proton are explicitly taken into account by the decrease of chemical potential.

As was said by IUPAC already in 1985[166] an "intersolvental" pH scale would be ultimately referenced to water due to the indisputable key role of water as a solvent. For the same reason absolute acidity is linked to water pH scale via Gibbs energy of solvation as follows:

$$\text{pH}_{abs}^{H_2O} = \text{pH}_{abs} + \frac{\Delta_{\text{solv}}G^{\ominus}(H^+, H_2O)}{RT \ln 10} \tag{A2-4}$$

The notion $\text{pH}_{abs}^{H_2O}$ means that pH is expressed on the absolute scale, but values are shifted by a constant in order to make the pH_{abs} values directly comparable to the conventional aqueous pH values (i.e. $\text{pH}_{abs}^{H_2O}$ value 7.00 refers to the acidity of the solution where the proton's chemical potential is as high as in aqueous solution with $_w^w\text{pH}$ 7.00). Thus the unified pH scale enables to express acidity of any media on a unified scale in the form of familiar aqueous pH ($_w^w\text{pH}$) values.

Absolute pH measurements

Absolute pH values (Table A2-1) were obtained by using a Metrohm 713 pH meter in differential potentiometry mode with two metal-coated glass electrodes (Laboratory of Glass Electrochemistry, St. Petersburg State University) as described in [133]. Measurements were made at (25 ± 1) °C. Consistency standard deviation of the results was 0.01 pH units and in total 52 measurements were made. Liquid junction potentials and uncertainties were calculated as in ref.[133] Absolute pH values were measured in bulk solutions and used as a reference for calibrating $\mathrm{pH}_{\mathrm{abs}}^{\mathrm{H_2O}}$ in the plume.

Table A2-1. The acidity of mobile phases expressed as $\mathrm{pH}_{\mathrm{abs}}^{\mathrm{H_2O}}$ values together with measurement uncertainties.

Mobile phase[a]	$\mathrm{pH}_{\mathrm{abs}}^{\mathrm{H_2O}}$	u_{RW}[b]	u_{C}[c]
MeCN/^w_wpH 5.50 80/20	8.59	0.01	0.14
MeCN^w_wpH 5.00 80/20	8.43	0.01	0.14
MeCN/^w_wpH 5.50 75/25	8.41	0.01	0.14
MeCN/^w_wpH 5.50 70/30	8.23	0.01	0.14
MeCN/^w_wpH 5.00 75/25	8.18	0.01	0.14
MeCN/^w_wpH 4.50 80/20	8.10	0.01	0.14
MeCN/^w_wpH 5.00 70/30	7.94	0.01	0.14
MeCN/^w_wpH 4.50 75/25	7.80	0.01	0.14
MeCN/^w_wpH 4.00 80/20	7.64	0.01	0.14
MeCN/^w_wpH 4.50 70/30	7.50	0.01	0.14
MeCN/^w_wpH 5.50 50/50	7.49	0.01	0.14
MeCN/^w_wpH 4.00 75/25	7.32	0.01	0.14
MeCN^w_wpH 5.50 45/55	7.30	0.01	0.14
MeCN/^w_wpH 3.50 80/20	7.18	0.01	0.14
^w_wpH 7 (diluted)	7.13	0.01	0.14
MeCN/^w_wpH 5.50 40/60	7.11	0.01	0.14
MeCN/^w_wpH 5.00 50/50	7.05	0.01	0.14
MeCN/^w_wpH 4.00 70/30	7.03	0.01	0.14
MeCN/^w_wpH 3.50 75/25	6.86	0.01	0.14
MeCN/^w_wpH 5.00 45/55	6.85	0.01	0.14
MeCN/^w_wpH 5.00 40/60	6.62	0.01	0.14
MeCN/^w_wpH 3.50 70/30	6.57	0.01	0.14
MeCN/^w_wpH 4.50 50/50	6.56	0.01	0.14
MeCN/^w_wpH 4.50 45/55	6.36	0.01	0.14
MeCN/^w_wpH 4.50 40/60	6.14	0.01	0.14

[a] The aqueous phase is 0.1% formic acid solution titrated with 25% ammonia solution to the desired ^w_wpH.
[b] Within-lab reproducibility estimates (u_{RW}) can be used to evaluate the internal consistency of the measured values.
[c] These combined standard uncertainty estimates (u_{C}) can be used to compare the acidities of the solutions in this scale with the acidities of aqueous solutions by conventional pH measurement.

Organic phase percentage and pH change in ESI plume
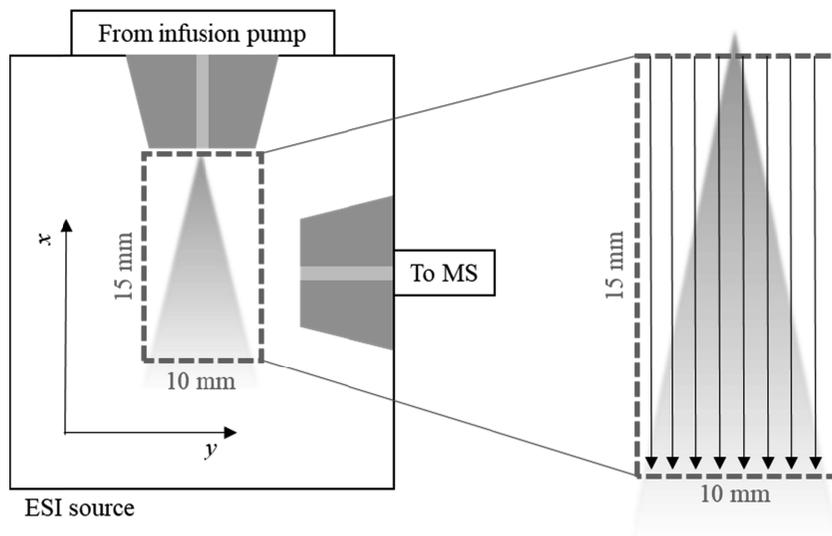
Figure A2-1. Setup for profiling changes in the ESI plume.

Appendix 3

Data compilation for indicator molecules

Table A3-1. Physicochemical properties calculated with COSMO-RS program, measured $\log IE$ values and LDA training set compounds (T) and validation set compounds (V). In this table $\log P$ - octanol–water coefficient for the neutral molecule.

	LDA	pK_{a1}	pK_{a2}	pH=10.74				pH=2.68				$\log P$	$WAPS \cdot 10^5$	
				α_2	α_1	$\log IE_2$	$\log IE_1$	α_2	α_1	$\log IE_2$	$\log IE_1$		anion	dianion
Bromothymol blue	T	1.15	5.76	1.00	0.00	1.23	2.95	0.00	0.97	-0.25	2.16	7.47	1.44	1.70
Bromophenol blue	T	0.99	2.96	1.00	0.00	2.63	2.34	0.34	0.65	0.28	1.93	6.86	1.54	1.86
Eosin Y	T	3.30	4.54	1.00	0.00	1.15	1.50	0.00	0.19	0.27	1.47	5.03	1.60	1.90
Bromocresol green	V	0.55	3.49	1.00	0.00	1.71	1.45	0.13	0.86	-0.15	2.01	6.63	1.62	1.83
Bathocuproinedisulfonic acid	V	-1.70	5.05	1.00	0.00	1.95	0.45	0.00	1.00	1.06	0.42	-0.06	2.02	1.92
Sulfosalicylic acid	T	-1.53	2.31	1.00	0.00	-0.38	0.22	0.70	0.30	-2.54	0.22	2.68	2.38	2.60
SPADNS	V	-2.32	-2.19	1.00	0.00	0.66	-0.20	1.00	0.00	0.16	-0.58	1.41	2.34	2.57
Tiron	V	-2.66	-2.15	1.00	0.00	-0.62	-0.68	1.00	0.00	-0.29	-0.34	3.07	3.98	5.01
Bromocresol purple	V	1.26	4.31	1.00	0.00			0.02	0.94			4.75	2.01	2.34
Eosin B	T	2.80	3.14	1.00	0.00			0.13	0.37			4.17	1.58	1.90
Thymolphthalein	T	10.04	10.98	0.32	0.56			0.00	0.00			5.79	1.66	1.98
Thymol blue	T	-0.89	7.54	1.00	0.00			0.00	1.00			4.51	1.70	2.01
Cresol red	T	1.08	6.31	1.00	0.00			0.00	0.98			4.10	2.19	2.59
m-cresol purple	T	-1.29	6.71	1.00	0.00			0.00	1.00			1.85	2.34	2.89
Phenolphthalein	T	9.28	9.04	0.98	0.02			0.00	0.00			2.66	2.60	3.31
Phenol-2,4-disulphonic acid*	T	-2.99	-1.91	1.00	0.00			1.00	0.00			2.10	4.28	5.25
Terephthalic acid	T	2.26	3.44	1.00	0.00			0.11	0.65			1.01	5.23	6.96
Suberic acid	T	4.91	4.88	1.00	0.00			0.00	0.01			1.12	5.24	6.05
Isophthalic acid	T	2.87	4.16	1.00	0.00			0.01	0.39			1.02	5.54	7.36
Pimelic acid	T	4.61	4.76	1.00	0.00			0.00	0.01			0.78	6.09	6.92
Phthalic acid	V	2.92	4.56	1.00	0.00			0.00	0.37			0.36	6.12	8.23
3-nitrophthalic acid	T	2.11	3.60	1.00	0.00			0.09	0.72			2.19	6.25	5.11

	LDA	pK_{a1}	pK_{a2}	pH=10.74				pH=2.68				$logP$	$WAPS·10^5$	
				$α2$	$α1$	$logIE_2$	$logIE_1$	$α1$	$α2$	$logIE_2$	$logIE_1$		anion	dianion
Adipic acid	T	4.98	4.98	1.00	0.00			0.00	0.00			0.18	7.04	8.10
Mesaconic acid	V	3.31	5.65	1.00	0.00			0.19	0.00			0.88	7.74	9.73
Itaconic acid	V	4.05	4.48	1.00	0.00			0.04	0.00			0.50	8.05	10.05
Glutaric acid	T	4.59	4.92	1.00	0.00			0.01	0.00			-0.35	8.22	9.65
Maleic acid	V	3.41	4.00	1.00	0.00			0.15	0.01			0.55	8.66	11.30
Succinic acid	V	4.27	4.89	1.00	0.00			0.02	0.00			-0.33	9.60	11.64
Fumaric acid	V	2.92	4.66	1.00	0.00			0.36	0.00			-0.26	9.64	12.69

Table A3-1 continues.

	Klamt parameters for neutrals						Klamt parameters for anions						Klamt parameters for dianions					
	Area	sig2	sig3	Hacc3	Hdon3	Volume	Area	sig2	sig3	Hacc3	Hdon3	Volume	Area	sig2	sig3	Hacc3	Hdon3	Volume
Bromothymol blue	441.15	278.47	-11.13	7.19	4.14	1279	440.70	670.25	642.39	5.08	2.88	1920	453.80	642.86	911.14	0.00	0.00	1298
Bromophenol blue	413.82	59.76	-14.01	1.97	1.21	210	410.24	89.30	66.52	0.95	0.58	209	412.11	123.81	153.43	0.00	0.00	209
Eosin Y	397.83	27.66	-3.45	0.95	0.59	101	397.34	41.93	35.34	0.45	0.27	101	398.37	64.11	95.61	0.00	0.00	101
Bromocresol green	412.07	134.63	-20.45	0.35	3.92	335	422.23	228.70	213.77	20.67	2.09	224	421.63	301.06	401.41	35.90	0.00	224
Bathocuproinedisulfonic acid	454.92	271.06	58.77	12.87	7.82	593	451.42	326.84	316.64	31.07	1.62	591	452.58	372.82	517.99	47.83	0.00	236
Sulphosalicylic acid	206.73	355.44	-184.82	28.36	19.70	451	204.79	280.44	232.24	7.69	5.13	268	202.51	450.45	700.83	2.82	1.75	266
SPADNS	418.45	329.71	-96.73	5.43	19.66	1214	416.26	381.30	184.66	23.28	11.29	2420	415.38	463.27	445.94	44.97	7.47	3416
Tiron	232.19	209.69	-152.21	0.28	19.87	570	229.07	222.22	114.73	11.56	7.62	256	227.08	321.48	427.90	38.51	1.64	304
Bromocresol purple	371.52	172.97	-63.80	0.73	8.73	657	371.32	264.37	216.29	24.74	4.73	654	368.82	325.34	452.86	42.24	0.00	663
Eosin B	400.82	106.99	-3.86	3.81	2.57	305	398.14	273.04	253.38	2.81	1.81	608	398.48	192.26	286.70	0.00	0.00	304
Thymolphthalein	419.42	363.62	-10.60	13.04	7.79	1354	415.90	395.03	460.73	3.83	2.23	899	420.38	141.66	236.78	0.00	0.00	222
Thymol blue	421.14	186.89	0.97	8.32	5.68	607	416.38	309.31	308.12	4.55	3.06	719	420.95	140.68	218.53	0.00	0.00	239
Cresol red	361.99	208.49	-28.73	7.83	4.79	543	362.56	283.80	292.18	3.09	1.95	457	362.83	294.46	447.27	0.00	0.00	367
m-cresol purple	346.67	319.84	9.07	13.55	9.28	737	343.29	601.01	605.96	8.45	5.66	1008	339.32	445.89	691.84	0.00	0.00	545
Phenolphthalein	319.33	69.28	-13.37	3.26	2.04	156	317.05	108.94	119.71	1.58	0.98	154	314.30	153.54	261.65	0.00	0.00	154
Phenol-2,4-disulphonic acid*	223.78	195.95	-128.31	16.29	11.74	248	221.27	405.62	252.07	13.42	9.43	394	218.96	505.70	698.92	3.78	2.54	341
Terephthalic acid	182.00	143.72	-62.20	10.66	7.34	189	180.90	188.87	224.34	3.83	2.56	149	178.59	221.53	417.77	0.00	0.00	110
Suberic acid	223.81	92.43	6.74	4.34	2.75	140	222.11	116.28	177.68	1.70	1.10	93	219.87	85.40	179.36	0.00	0.00	46
Isophthalic acid	182.90	167.07	-42.95	10.23	6.74	229	182.35	192.88	244.32	3.01	1.93	151	180.01	231.61	431.51	0.00	0.00	113
Pimelic acid	205.48	92.62	1.33	4.81	3.11	127	203.50	113.96	177.31	1.45	0.91	84	200.95	84.96	177.61	0.00	0.00	41
Phthalic acid	180.99	85.18	-28.11	5.72	3.89	117	179.80	202.84	272.10	3.06	1.95	153	178.23	251.42	495.16	0.00	0.00	114
3-nitrophthalic acid	209.17	271.61	-115.91	17.89	12.52	362	207.47	499.12	538.85	4.41	5.89	448	205.82	391.39	701.01	0.00	0.00	223

	Klamt parameters for neutrals						Klamt parameters for anions						Klamt parameters for dianions					
	Area	sig2	sig3	Hacc3	Hdon3	Volume	Area	sig2	sig3	Hacc3	Hdon3	Volume	Area	sig2	sig3	Hacc3	Hdon3	Volume
Adipic acid	184.09	94.17	-5.22	5.20	3.37	113	182.38	114.52	174.49	1.69	1.08	74	180.87	85.06	178.56	0.00	0.00	37
Mesaconic acid	155.39	108.81	-30.33	7.25	4.91	124	153.94	186.00	260.14	2.22	1.37	122	150.51	80.01	164.44	0.00	0.00	30
Itaconic acid	155.51	203.46	-55.89	13.06	8.74	220	153.10	407.01	584.34	5.89	3.78	248	151.37	248.76	514.30	0.00	0.00	92
Glutaric acid	164.88	94.27	-9.88	5.33	3.47	99	162.50	113.06	171.91	1.69	1.09	65	160.37	85.78	181.97	0.00	0.00	32
Maleic acid	141.30	87.82	-38.57	6.87	4.73	80	276.17	176.55	293.24	1.65	1.07	53	137.00	80.44	163.16	0.00	0.00	26
Succinic acid	145.04	92.07	-19.33	5.60	3.70	85	142.85	108.42	162.90	1.64	1.05	55	140.63	86.18	184.19	0.00	0.00	27
Fumaric acid	137.29	83.95	-31.57	5.74	3.95	80	136.09	100.26	130.90	1.65	1.06	79	134.38	85.77	177.84	0.00	0.00	26

* - gave doubly charged species but it was impossible to get quantitative results due to spectral interferences it was not possible to measure the corresponding logIE value.

Data compilation for amino acids

Table A3-2. Calculated physicochemical properties of amino acids. *WANS* (weighed average negative sigma) is a parameter describing charge delocalization (see Paper II for calculation equation); pK_a values are given for conjugated acid; the degree of ionization (α) is calculated from pK_a values (pH of 0.1% formic acid is 2.68).

Amino acid	log*IE* measured	including zwitterionic geometries				only non-zwitterionic geometries			
		log*IE* predicted	*WANS*·10⁵	α^{ESI+}	pK_a	log*IE* predicted	*WANS*·10⁵	α^{ESI+}	pK_a
Leu	3.15	2.04	-5.49	0.00	-0.16	3.41	-5.49	0.99	4.81
Phe	3.07	2.16	-5.13	0.00	-0.24	3.53	-5.13	0.99	4.54
His	3.04	2.68	-7.36	0.95	3.98	2.75	-7.36	1.00	5.59
Arg	2.88	3.86	-4.28	1.00	13.71	3.86	-4.28	1.00	11.55
Ile	2.75	2.02	-5.54	0.01	0.44	3.40	-5.54	1.00	5.63
Met	2.7	2.12	-5.26	0.00	-0.55	3.50	-5.26	0.99	4.94
Val	2.7	1.63	-6.62	0.01	0.50	3.01	-6.62	1.00	5.62
Pro	2.63	1.59	-6.73	0.00	-0.47	2.98	-6.73	1.00	5.90
Trp	2.62	2.14	-5.20	0.00	0.22	3.52	-5.20	1.00	5.42
Lys	2.58	3.65	-4.85	1.00	9.56	3.65	-4.85	1.00	10.68
Tyr	2.51	1.76	-6.26	0.00	-0.09	3.14	-6.26	1.00	5.27
Glu	2.39	1.57	-6.79	0.00	-0.14	2.48	-7.46	0.83	3.38
Thr	2.35	1.15	-7.94	0.00	-0.64	2.54	-7.94	1.00	5.98
Gln	2.31	2.71	-7.46	1.00	5.83	2.95	-6.79	1.00	5.83
Asp	2.21	0.92	-8.59	0.00	-1.33	2.30	-8.59	0.99	4.79
α-Ala	2.17	0.48	-9.82	0.01	0.44	1.86	-9.82	1.00	5.66
β-Ala	2.15	0.37	-10.39	0.07	1.58	1.66	-10.39	1.00	9.62
Ser	2.14	0.38	-10.09	0.00	0.05	1.77	-10.09	1.00	7.10
Asn	2.12	1.15	-7.96	0.00	-0.02	2.53	-7.96	1.00	5.71
Gly	1.3	-1.13	-14.28	0.00	0.30	0.26	-14.28	1.00	6.16
Cys	1.24	0.89	-8.66	0.00	-0.24	2.28	-8.66	1.00	6.74
s	0.24								

Figure A3-1a. Correlation between predicted and measured log*IE* values of amino acids. Predicted log*IE* values calculated using physicochemical descriptors for only non-zwitterionic geometries of amino acids.
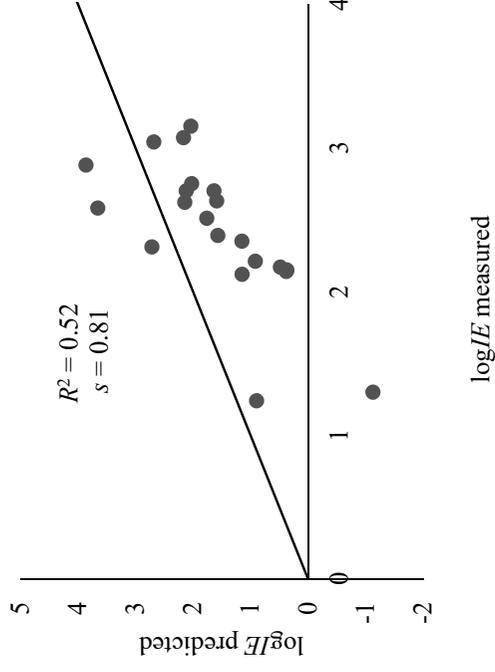
Figure A3-1b. Correlation between predicted and measured log*IE* values of amino acids. Predicted log*IE* values calculated using physicochemical descriptors for both zwitterionic and non-zwitterionic geometries of amino acids.

Data compilation for small peptides

Table A3-3 Measured sum of log*IE* values of different charge states (Σlog*IE*) and the sum on ionization efficiencies of amino acids in the specific oligopeptide (ΣAA log*IE*) and log*IE* values for different charge states of oligopeptides (log*IE*$^{\text{charge state}}$).

	Oligopeptide	Σlog*IE*	ΣAA log*IE*	log*IE*$^{1+}$	log*IE*$^{2+}$	log*IE*$^{3+}$
1	Gln-Phe-Phe-Gly-Leu-Met-NH$_2$	5.34	3.66	5.34		
2	Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$	4.91	3.74	4.91		
3	Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$	4.76	3.68	4.76		
4	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly	4.69	3.69	4.69		
5	Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH$_2$	4.65	3.71	4.64	2.94	
6	Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met	4.55	3.83	4.55		
7	Phe-Phe-Gly-Leu-Met-NH$_2$	4.38	3.63	4.38		
8	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu	5.34	3.70	4.42	5.11	4.79
9	Asp-Arg-Val-Tyr-Ile-His-Pro-Phe	5.30	3.70	5.12	4.84	
10	Asn-Arg-Val-Tyr-Ile-His-Pro-Phe	5.25	3.70	5.06	4.80	
11	Val-Tyr-Ile-His-Pro-Phe	5.20	3.61	5.09	4.57	
12	Tyr-Ile-His-Pro-Phe	4.80	3.55	4.66	4.25	
13	Asp-Arg-Val-Tyr-Ile-His-Pro	4.49	3.58	4.42	3.70	
14	Ile-His-Pro-Phe	4.44	3.51	4.39	3.52	
15	Lys-Arg-Pro-Pro-Gly-Phe-Ser-Pro-Leu	5.42	3.71	5.42		
16	Arg-Pro-Pro-Gly-Phe-Ser-Pro-Leu	5.30	3.68	5.30		
17	Arg-Pro-Gly-Phe-Ser-Pro-Phe-Arg	5.04	3.57	4.88	4.54	
18	Arg-Pro-Pro-Gly-Phe-Ser-Pro	4.98	3.53	4.98		
19	Arg-Pro-Pro-Gly-Phe-Ser	4.75	3.47	4.75		
20	Arg-Pro-Pro-Gly-Phe	4.30	3.45	4.30		
21	Arg-Pro-Pro	2.87	3.21	2.87		
22	Thr-Arg-Ser-Ala-Trp-NH$_2$	4.54	3.23	4.54		
23	Thr-Arg-Ser-Ala-Trp	4.41	3.23	4.41		
24	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg	4.70	3.49	4.70		

	Oligopeptide	$\Sigma\log IE$	ΣAA $\log IE$	$\log IE^{1+}$	$\log IE^{2+}$	$\log IE^{3+}$
25	Cys-Tyr-Phe-Gln-Asn-Cys	4.69	3.28	4.69		
26	Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH$_2$	4.26	3.49	4.26		
27	Trp-Ala-Gly-Gly-Asn-Ala-Ser-Gly-Glu	4.96	3.10	4.96		
28	Trp-Ala-Gly-Gly-Asp-Ala-Ser-Gly-Glu	4.30	3.11	4.30		
29	Gly-Gly-Asp-Ala	2.78	2.54	2.78		
30	Gly-Gly-Gly-Phe-Phe-NH$_2$	3.85	3.47	3.85		
31	Gly-Gly-Gly-NH$_2$	1.93	1.78	1.91		
32	Gly-Lys-Pro-Ile-Pro-Asn-Pro-Leu-Leu-Gly-Leu-Asp-Ser-Thr	5.61	3.85	5.61		
33	Arg-Arg-Leu-Ile-Glu-Asp-Ala-Glu-Tyr-Ala-Ala-Arg-Gly	4.70	3.75	4.70		
34	Phe-Phe-Phe-Phe	4.69	3.67	4.69		
35	Arg-Arg-Pro-Tyr-Ile-Leu	4.52	3.75	4.46	3.67	
36	Ac-Gly-Lys-OMe	3.79	2.60	3.79		
37	Gly-Pro-Gly-Gly	3.79	2.69	3.79		
38	Gly-β-Ala-β-Ala	2.57	2.48	2.57		
$s_{\text{consistency}}$		0.42				

Appendix 4

Ionization efficiency measurements in biological matrices

Table A4-1. Compounds used in measuring IEs in biological matrices, their pK_a, α, $\log IE$ values in solvent and in different biological matrices (cerebrospinal fluid (CSF), urine, plasma, blood, liver tissue, and brain tissue), charge delocalization parameters (*WAPS*) values for substances. C denotes the used molar concentration range for each compound. Warfarin is used to make the $\log IE$ values measured in different matrices numerically comparable.

| Name	$\log IE_{\mathrm{solv}}$	$\log IE_{\mathrm{plasma}}$	$\log IE_{\mathrm{urine}}$	$\log IE_{\mathrm{blood}}$	$\log IE_{\mathrm{liver}}$	$\log IE_{\mathrm{brain}}$	$\log IE_{\mathrm{CSF}}$	pK_a	α	$WAPS \cdot 10^5$	C (μM)
salicylic acid	0.34	-0.42	-0.76	-0.42	-0.38	-0.28	-0.68	3.33	1.00	6.06	6-62
benzoic acid	0.00	-2.55	-2.09	-2.90	-1.58	-1.17	NA	4.63	1.00	7.07	4-45
lincomycin	0.20	-1.55	-2.38	-1.90	-1.20	-1.14	-1.90	11.65	0.11	2.17	1-13
warfarin	1.07	0.68	0.09	0.78	0.62	0.94	0.34	4.63	1.00	2.43	2-23
naproxen	0.12	-1.62	-2.28	-1.22	-0.69	-0.66	-1.67	5.10	1.00	3.46	3-33
fumaric acid	-0.60	-3.53	-2.80	-3.69	-1.84	-1.46	-2.88	3.68	1.00	8.78	18-180
dodecanoic acid	0.24	-1.10	-1.55	-0.82	-0.57	-0.41	-0.58	4.96	1.00	2.85	4-42
taurocholic acid	0.97	-0.16	-0.25	-0.29	0.29	0.54	-0.31	-2.36	1.00	1.92	3-27
sorbic acid	-0.36	-1.50	-2.43	-1.25	-1.06	-0.87	-0.93	5.38	1.00	7.01	23-233
3-(CF$_3$SO$_2$)–benzoic acid	0.83	0.33	-0.42	0.06	0.22	0.38	-0.28	3.77	1.00	3.65	2-17
$s_{\mathrm{consistency}}$	0.12	0.15	0.28	0.25	0.10	0.19	0.14				
s (anchor)	0.21	0.33	0.01	0.1	0.05	0.08					

Appendix 5

Combining and quantitative comparison of ESI+ and ESI-ionization mode

Table A5-1. Charge delocalization values (*WANS*) calculated for compounds in ref [13].

Compound	$WANS \cdot 10^5$	Compound	$WANS \cdot 10^5$
$(C_4H_9N)_3$-P=N-$(C_4H_9N)_2$-P=N-C_6H_4-2-Cl	0.78	methomyl	4.14
tetrahexylammonium	0.26	benzophenone	3.03
$(C_4H_9N)_3$-P=N-C_6H_4-4-CF_3	1.35	aldicarb	3.45
$(C_4H_9N)_3$-P=N-C_6H_3-2,5-Cl_2	1.21	piperidine	6.17
$((CH_3)_2N)_3$-P=N-C_6H_5	1.81	aniline	8.23
tetrabutylammonium	1.28	2-methylpyridine	5.43
tetrapropylammonium	1.88	N-methylpiperidine	4.74
phenyl tetramethylguanidine	0.56	4-nitroaniline	6.41
tributylamine	1.80	pyridine	7.20
hexyl-methylimidazolium	0.48	dimethyl glutarate	3.45
diphenylguanidine	2.95	benzamide	5.86
tripropylamine	3.98	pyrrolidine	7.47
acridine	3.09	diethylamine	6.45
diphenylamine	3.88	phenylbenzoate	3.10
diphenyl phthalate	1.81	2-nitroaniline	6.24
1-naphthylamine	5.63	4-chloro-2-nitroaniline	5.47
DBU	3.50	tetramethylammonium	6.14
tetraethylammonium	3.19	dimethyl succinate	4.14
2,4,6-trimethylpyridine	3.46	guanidine	12.70
tetramethylguanidine	3.42	trimethylamine	7.76
methiocarb	2.80	dimethyl malonate	5.02
N,N-dimethylaniline	4.68	2-cyanophenol	5.61
ethyl-methylimidazolium	4.23	benzoic acid	6.06
triphenylamine	2.32	2,4-dinitroaniline	5.94
4-fluoro-3-nitroaniline	6.51	2,6-dimethoxy pyridine	4.03
dimethyl phthalate	2.80	ethylamine	11.97
triethylamine	3.98	3-chloropyridine	6.41
3-nitroaniline	6.47	2-methoxypyridine	5.21
benzylamine	6.58	2-chloropyridine	5.86
2,6-dimethylpyridine	4.26	ethyl benzoate	3.77
sulphanilamide	6.14	methyl benzoate	4.60
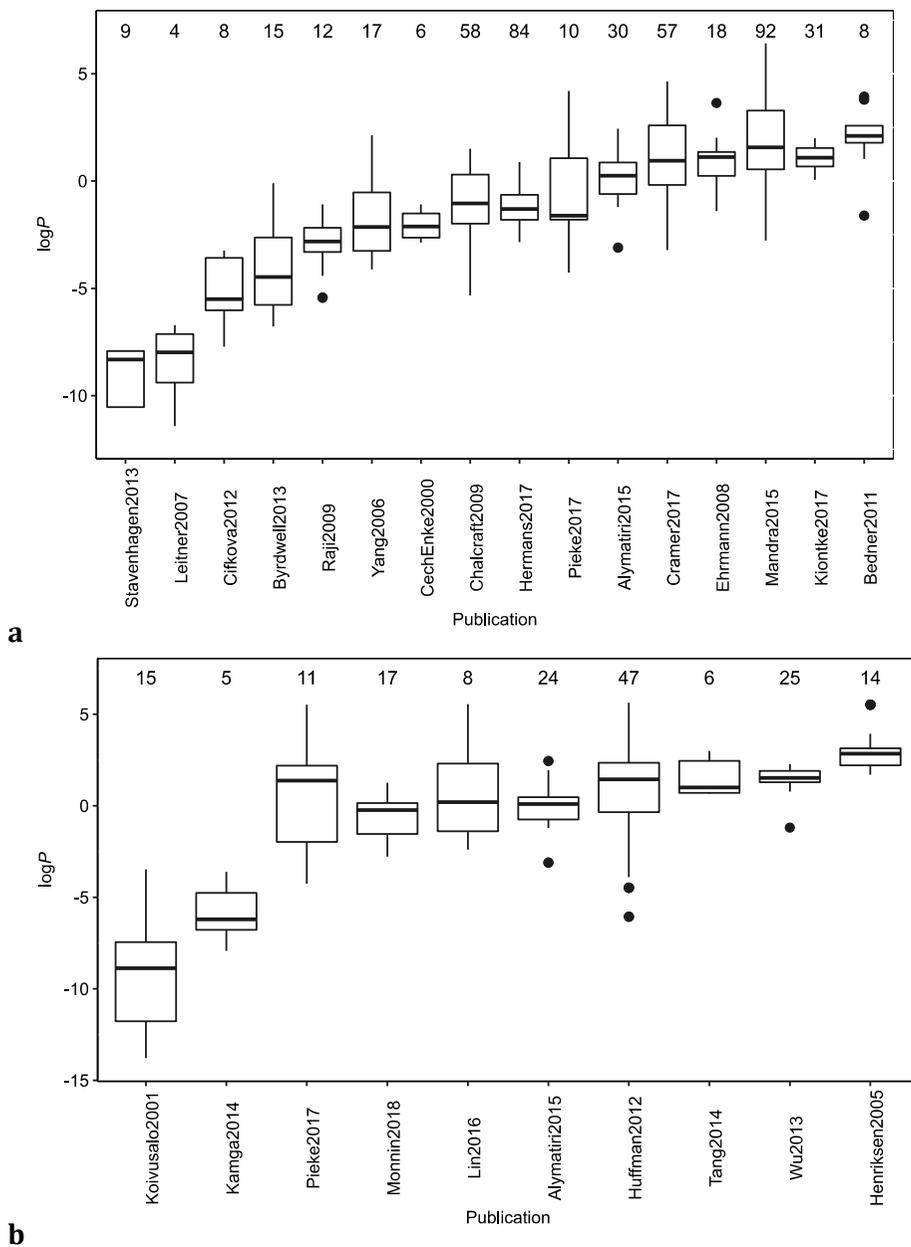
Figure A6-1. log*P* distributions of compounds studied in individual publications. a – ESI positive mode, b – ESI negative mode. The number above the boxplots represents the number of unique compounds in the corresponding set.
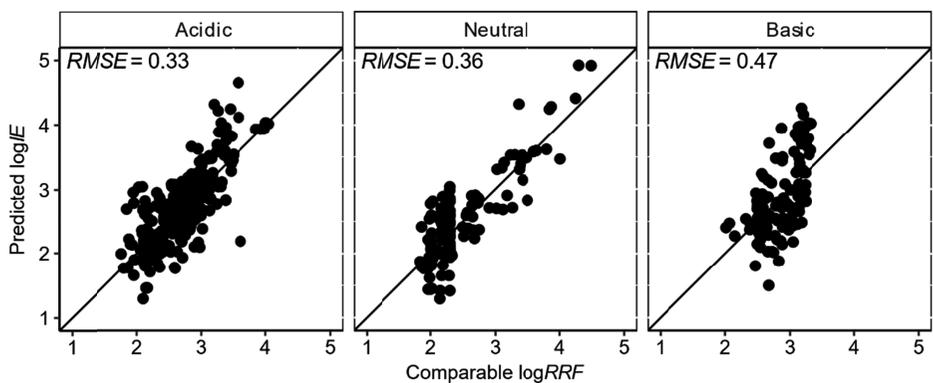
a
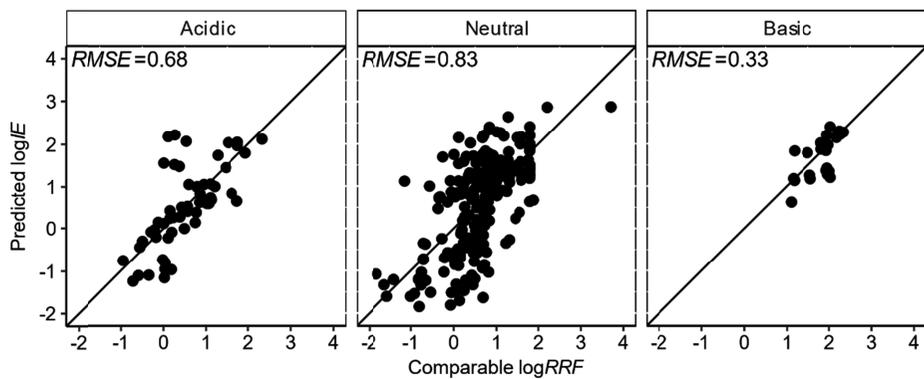
b

Figure A6-2. Comparison of predicted and measured values grouped by pH of the aqueous phase. a – ESI positive mode, b – ESI negative mode.
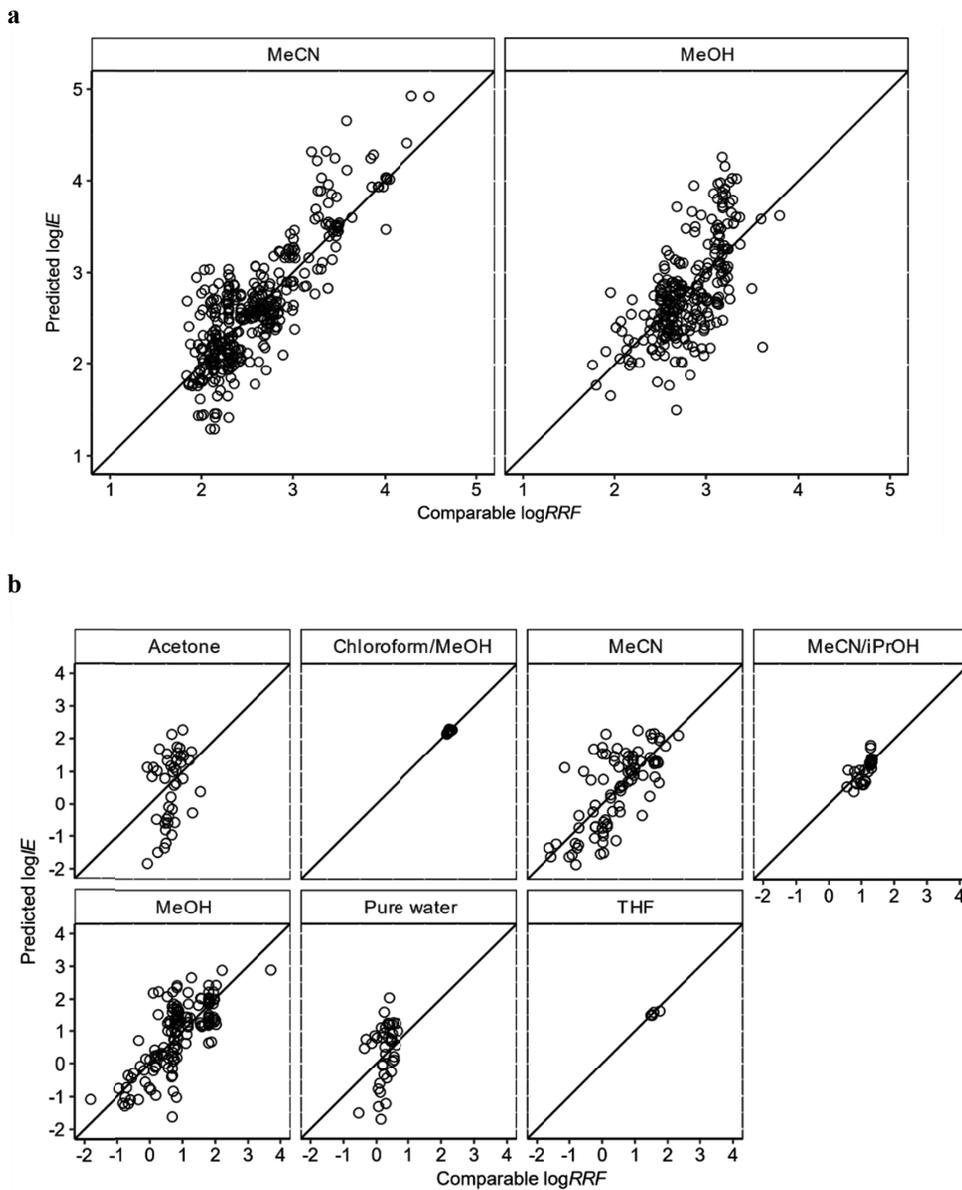
Figure A6-3. Comparison of predicted and measured values grouped by the organic modifier. a – ESI positive mode, b – ESI negative mode.
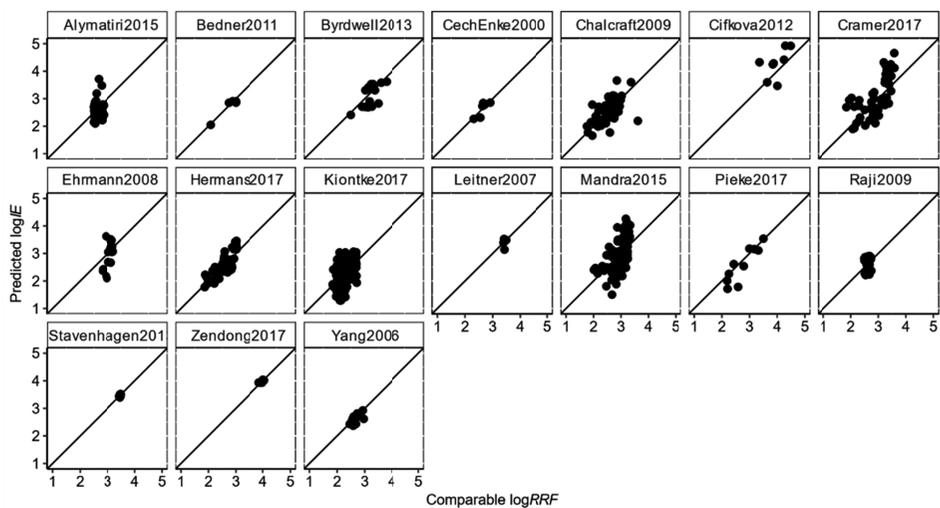
Figure A6-4. Comparison of predicted log*IE* and measured values from literature (comparable log*RRF*) in ESI positive mode.
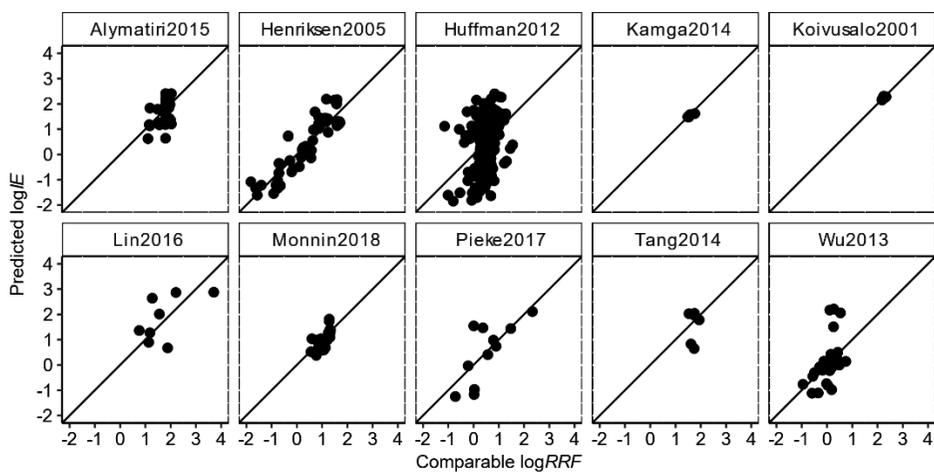
Figure A6-5. Comparison of predicted log*IE* and measured values from literature (comparable log*RRF*) in ESI negative mode.
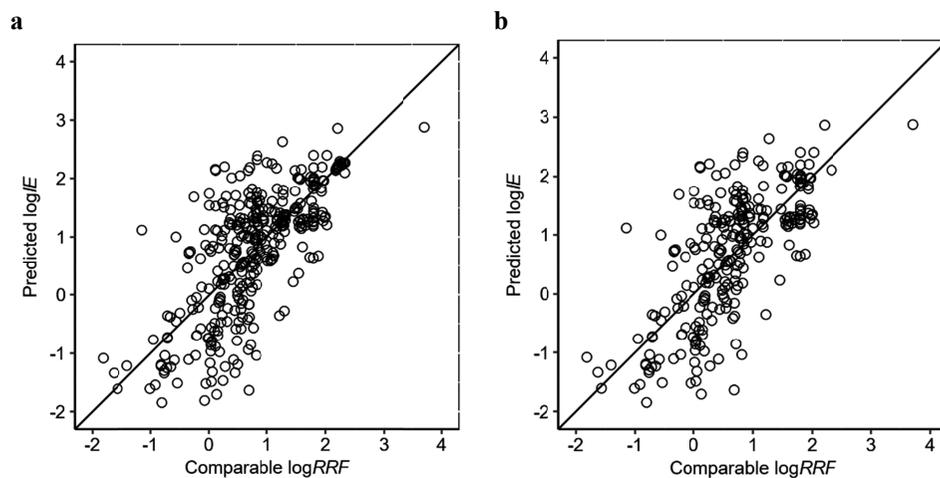
Figure A6-6. a – ESI negative mode all data; b – ESI negative mode with only MeOH, MeCN and water containing eluent compositions.

Table A6-1. Overview of data used in Paper VI.

Row Labels	Count of ESI+/ESI-
ESI-	**504**
AB Sciex	**31**
Zendong2017	<u>6</u>
60/40 acetonitrile/2 mM ammonium formate (both)	6
Wu2013	<u>25</u>
gradient methanol/0.025% acetic acid(aq)	25
Agilent	**218**
Huffman2012	<u>173</u>
acetone	45
acetonitrile	42
methanol	43
water	43
Monnin2018	<u>34</u>
gradient acetonitrile/isopropnaol(9/1)/0.02% acetic acid(aq)	17
gradient acetonitrile/isopropnaol(9/1)/10 mM ammonium acetate(aq)	17
Pieke2017	<u>11</u>
gradient acetonitrile/ammonium formate(aq, pH = 3.1)	11
Bruker	**26**
Kamga2014	<u>5</u>
tetrahydrofuran	5
Koivusalo2001	<u>15</u>
chloroform/methanol (1:2, 1% ammonia)	15
Tang2014	<u>6</u>
gradient acetonitrile/formic acid(aq, pH = 3)	6
Micromass	**54**
Henriksen2005	<u>54</u>
50/50 acetonitrile/water	13
50/50 methanol/water	13
acetonitrile	14
methanol	14
Shimadzu	**42**
Alymatiri2015	<u>42</u>
50/50 methanol/ammonia(aq, pH = 10)	21
methanol	21
Thermo	**8**
Lin2016	<u>8</u>
gradient methanol/1 mM ammonium acetate (both)	8
Waters	**125**
Ghosh2015	<u>125</u>
40/60 acetonitrile/10 mM ammonium formate(aq)	25
50/50 acetonitrile/10 mM ammonium formate(aq)	25
60/40 acetonitrile/10 mM ammonium formate(aq)	25
70/30 acetonitrile/10 mM ammonium formate(aq)	25
80/20 acetonitrile/10 mM ammonium formate(aq)	25

Row Labels	Count of ESI+/ESI-
ESI+	**641**
AB Sciex	**164**
Byrdwell2013	<u>15</u>
methanol	15
Kiontke2017	<u>124</u>
50/50 acetonitrile/20 mM formic acid(aq)	31
50/50 acetonitrile/water	31
80/20 acetonitrile/20 mM formic acid(aq)	31
80/20 acetonitrile/water	31
Zendong2017	<u>8</u>
60/40 acetonitrile/2 mM ammonium formate (both)	8
Yang2006	<u>17</u>
50/50 methanol/2 % acetic acid(aq)	17
Agilent	**80**
Bedner2011	<u>8</u>
gradient acetonitrile/0.1% formic acid(both)	8
Chalcraft2009	<u>58</u>
50/50 methanol/0.1% formic acid(aq)	58
Leitner2007	<u>4</u>
gradient acetonitrile/0.1% ammonium acetate(aq)	4
Pieke2017	<u>10</u>
gradient acetonitrile/ammonium formate(aq, pH = 3.1)	10
Bruker	**167**
Cífková2012	<u>8</u>
gradient acetonitrile/5 mM ammonium acetate(aq)	8
Hermans2017	<u>84</u>
gradient acetonitrile/0.1% formic acid (both)	84
Kiontke2017	<u>62</u>
50/50 acetonitrile/20 mM formic acid(aq)	31
50/50 acetonitrile/water	31
Stavenhagen2013	<u>13</u>
35/65 acetonitrile/0.1% formic acid(aq)	13
Shimadzu	**122**
Alymatiri2015	<u>30</u>
50/50 methanol/ammonia(aq, pH = 10)	30
Mandra2015	<u>92</u>
50/50 methanol/ammonia(aq, pH = 10)	92
Thermo	**108**
Alfaro2014	<u>2</u>
40/60 acetonitrile/0.1% formic acid (both)	2
Cech Enke 2000	<u>6</u>
50/50 methanol/0.5% acetic acid (both)	6
Cramer2017	<u>58</u>
50/50 acetonitrile/0.1% formic acid(aq)	58
Ehrmann2008	18
methanol 0.5% acetic acid	18
Raji2009	<u>24</u>
10/90 methanol/1 mM ammonium acetate(aq, pH = 6)	24
Grand Total	**1145**

113

PUBLICATIONS

CURRICULUM VITAE

Name: Piia Liigand (née Burk)
Date of birth: February 18, 1990
Citizenship: Estonia
E-mail: piia.liigand@ut.ee

Education:
2015–   University of Tartu, chemistry, doctoral studies
2013–2015   University of Tartu, chemistry, MSc, *cum laude*
2011–2013   University of Tartu, chemistry, BSc

Publications:

Liigand, P.; Liigand, J.; Kaupmees, K., Kruve, A. Revealing the hidden treasure: Making ionization efficiency values from literature directly comparable, submitted for publication in *Anal. Chem.*

Rebane, R.; Kruve, A.; Liigand, J.; Liigand, P.; Gornischeff, A.; Leito, I. Ionization efficiency ladders as tools for choosing ionization mode and solvent in LC/MS, submitted for publication in *Rapid Commun. Mass Spectrom.*

Liigand, P.; Kaupmees, K.; Kruve, A. Influence of the amino acid composition on the ionization efficiencies of small peptides. *J. Mass Spectrom.*, 2019, doi: 10.1002/jms.4348

Liigand, P.; Liigand, J.; Cuyckens, F.; Vreeken, R.J.; Kruve, A. Ionisation efficiencies can be predicted in complicated biological matrices: A proof of concept. *Anal. Chim. Acta* 2018, *1032*, 68–74.

Liigand, P.; Heering (Suu), A.; Kaupmees, K.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. The evolution of electrospray generated droplets is not affected by ionization mode. *J. Am. Soc. Mass Specrtom.* 2017, *28*, 2124–2131.

Liigand, P.; Kaupmees, K. Haav, K. Liigand, J.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. Think Negative: Finding the Best Electrospray Ionization/MS Mode for Your Analyte. *Anal. Chem.* 2017, *89*, 5665–5668.

Liigand, P.; Kaupmees, K.; Kruve, A. Ionization Efficiency of Doubly Charged Ions Formed from Polyprotic Acids in Electrospray Negative Mode. *J. Am. Soc. Mass Specrtom.* 2016, *27*, 1211–1218.

Rebane, R.; Kruve, A.; Liigand, P.; Liigand, J.; Herodes, K.; Leito, I. Establishing Atmospheric Pressure Chemical Ionization Efficiency Scale. *Anal. Chem.* 2016, *88*, 3435–3439.

Liigand, J.; Kruve, A.; Liigand, P.; Laaniste, A.; Girod, M.; Antoine, R.; Leito, I. Transferability of the Electrospray Ionization Efficiency Scale between Different Instruments. *J. Am. Soc. Mass Spectrom.*, 2015, *26*, 923–1930.

Vahur, S.; Teearu, A.; Haljasorg, T.; Burk, P.; Leito, I.; Kaljurand, I. Analysis of dammar resin with MALDI-FT-ICR-MS and APCI-FT-ICR-MS. *J. Mass Spectrom.*, 2012, *47*, 392–409.

Professional self-improvement and conferences:

2019   67th ASMS Conference on Mass Spectrometry and Allied Topics, poster presentation

2018   XXII International Mass Spectrometry Conference, poster presentation

2018   6 month internship at Multimodal Imaging Institute, Maastricht University, the Netherlands "DESI-MS analysis of peptides"

2018   66th ASMS Conference on Mass Spectrometry and Allied Topics, oral presentation

2017   6 month internship at Janssen Pharmaceutica NV, Belgium "Implementation of paper spray in generic quantitative analysis of drugs"

2016   XXI International Mass Spectrometry Conference, poster presentation

2014   3 month internship at University of Claude Bernard Lyon 1 "ESI spray profiling"

2014   XX International Mass Spectrometry Conference, poster presentation

Awards:

2016   Talveakadeemia NGO, third prize in oral presentation for "Ionisation efficiency and mechanism of ionization for multiply charged ions".

2015   Estonian Academy of Sciences, students award for master thesis: "Expanding the electrospray ionization efficiency scale in positive and negative mode ESI".

2013   National student award, 1. prize in the field of natural sciences and technology for bachelor thesis: "Analysis of colophony and dammar resin with APCI-FT-ICR-MS and ATR-FT-IR".

ELULOOKIRJELDUS

Nimi: Piia Liigand (née Burk)
Sünniaeg: 18. veebruar 1990
Kodakondsus: eestlane
E-mail: piia.liigand@ut.ee

Haridus:
2015– Tartu Ülikool, keemia, doktorantuur
2013–2015 Tartu Ülikool, keemia, MSc, *cum laude*
2011–2013 Tartu Ülikool, keemia, BSc

Publikatsioonid:

Liigand, P.; Liigand, J.; Kaupmees, K., Kruve, A. Revealing the hidden treasure: Making ionization efficiency values from literature directly comparable, submitted for publication in *Anal. Chem.*

Rebane, R.; Kruve, A.; Liigand, J.; Liigand, P.; Gornischeff, A.; Leito, I. Ionization efficiency ladders as tools for choosing ionization mode and solvent in LC/MS, submitted for publication in *Rapid Commun. Mass Spectrom.*

Liigand, P.; Kaupmees, K.; Kruve, A. Influence of the amino acid composition on the ionization efficiencies of small peptides. *J. Mass Spectrom.*, 2019, doi: 10.1002/jms.4348

Liigand, P.; Liigand, J.; Cuyckens, F.; Vreeken, R.J.; Kruve, A. Ionisation efficiencies can be predicted in complicated biological matrices: A proof of concept. *Anal. Chim. Acta* 2018, *1032*, 68–74.

Liigand, P.; Heering (Suu), A.; Kaupmees, K.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. The evolution of electrospray generated droplets is not affected by ionization mode. *J. Am. Soc. Mass Specrtom.* 2017, *28*, 2124–2131.

Liigand, P.; Kaupmees, K. Haav, K. Liigand, J.; Leito, I.; Girod, M.; Antoine, R.; Kruve, A. Think Negative: Finding the Best Electrospray Ionization/MS Mode for Your Analyte. *Anal. Chem.* 2017, *89*, 5665–5668.

Liigand, P.; Kaupmees, K.; Kruve, A. Ionization Efficiency of Doubly Charged Ions Formed from Polyprotic Acids in Electrospray Negative Mode. *J. Am. Soc. Mass Specrtom.* 2016, *27*, 1211–1218.

Rebane, R.; Kruve, A.; Liigand, P.; Liigand, J.; Herodes, K.; Leito, I. Establishing Atmospheric Pressure Chemical Ionization Efficiency Scale. *Anal. Chem.* 2016, *88*, 3435–3439.

Liigand, J.; Kruve, A.; Liigand, P.; Laaniste, A.; Girod, M.; Antoine, R.; Leito, I. Transferability of the Electrospray Ionization Efficiency Scale between Different Instruments. *J. Am. Soc. Mass Spectrom.*, 2015, *26*, 923–1930.

Vahur, S.; Teearu, A.; Haljasorg, T.; Burk, P.; Leito, I.; Kaljurand, I. Analysis of dammar resin with MALDI-FT-ICR-MS and APCI-FT-ICR-MS. *J. Mass Spectrom.*, 2012, *47*,392–409.

Enesetäiendus ja konverentsid:

2019   67th ASMS Conference on Mass Spectrometry and Allied Topics, poster ettekanne

2018   XXII International Mass Spectrometry Conference, poster ettekanne

2018   6-kuuline praktika Maastrichti Ülikool, "Desoprtsioon elektropihustus ionisatsiooni arendamine", juhendaja professor Ron M.A. Heeren

2018   66th ASMS Conference on Mass Spectrometry and Allied Topics, suuline ettekanne

2017   6-kuuline praktika ja koostöö Janssen Pharmaceutica NV-ga, Beerse, Belgia, juhendaja Rob J. Vreeken

2016   XXI International Mass Spectrometry Conference, poster ettekanne

2014   3-kuuline praktika University of Claude Bernard Lyon 1 "ESI pihuse uurimine"

2014   XX International Mass Spectrometry Conference, poster ettekanne

Auhinnad:

2016   Talveakadeemia MTÜ, 3. koht suulise ettekande eest "Mitmelaenguliste ioonide ionisatsiooni mehhanismi ja efektiivsuse uurimine ".

2015   Eesti Teaduste Akadeemia teadusauhind üliõpilastele magistritöö eest: "Elektropihustuse ionisatsiooniefektiivsuse skaala laiendamine positiivses ja negatiivses režiimis".

2013   Üliõpilaste teadustööde riiklik konkurss. 1. preemia loodusteaduste ja tehnika valdkonna preemia rakenduskõrgharidusõppe ja bakalaureuseõppe üliõpilaste astmes: "Kampoli ja dammaravaigu komponentide uurimine APCI-FT-ICR massispektromeetria ja ATR-FT-IR spektroskoopia abil".

DISSERTATIONES CHIMICAE
UNIVERSITATIS TARTUENSIS

1. **Toomas Tamm.** Quantum-chemical simulation of solvent effects. Tartu, 1993, 110 p.
2. **Peeter Burk.** Theoretical study of gas-phase acid-base equilibria. Tartu, 1994, 96 p.
3. **Victor Lobanov.** Quantitative structure-property relationships in large descriptor spaces. Tartu, 1995, 135 p.
4. **Vahur Mäemets.** The ^{17}O and ^{1}H nuclear magnetic resonance study of H_2O in individual solvents and its charged clusters in aqueous solutions of electrolytes. Tartu, 1997, 140 p.
5. **Andrus Metsala.** Microcanonical rate constant in nonequilibrium distribution of vibrational energy and in restricted intramolecular vibrational energy redistribution on the basis of slater's theory of unimolecular reactions. Tartu, 1997, 150 p.
6. **Uko Maran.** Quantum-mechanical study of potential energy surfaces in different environments. Tartu, 1997, 137 p.
7. **Alar Jänes.** Adsorption of organic compounds on antimony, bismuth and cadmium electrodes. Tartu, 1998, 219 p.
8. **Kaido Tammeveski.** Oxygen electroreduction on thin platinum films and the electrochemical detection of superoxide anion. Tartu, 1998, 139 p.
9. **Ivo Leito.** Studies of Brønsted acid-base equilibria in water and non-aqueous media. Tartu, 1998, 101 p.
10. **Jaan Leis.** Conformational dynamics and equilibria in amides. Tartu, 1998, 131 p.
11. **Toonika Rinken.** The modelling of amperometric biosensors based on oxidoreductases. Tartu, 2000, 108 p.
12. **Dmitri Panov.** Partially solvated Grignard reagents. Tartu, 2000, 64 p.
13. **Kaja Orupõld.** Treatment and analysis of phenolic wastewater with microorganisms. Tartu, 2000, 123 p.
14. **Jüri Ivask.** Ion Chromatographic determination of major anions and cations in polar ice core. Tartu, 2000, 85 p.
15. **Lauri Vares.** Stereoselective Synthesis of Tetrahydrofuran and Tetrahydropyran Derivatives by Use of Asymmetric Horner-Wadsworth-Emmons and Ring Closure Reactions. Tartu, 2000, 184 p.
16. **Martin Lepiku.** Kinetic aspects of dopamine D_2 receptor interactions with specific ligands. Tartu, 2000, 81 p.
17. **Katrin Sak.** Some aspects of ligand specificity of P2Y receptors. Tartu, 2000, 106 p.
18. **Vello Pällin.** The role of solvation in the formation of iotsitch complexes. Tartu, 2001, 95 p.
19. **Katrin Kollist.** Interactions between polycyclic aromatic compounds and humic substances. Tartu, 2001, 93 p.

20. **Ivar Koppel.** Quantum chemical study of acidity of strong and superstrong Brønsted acids. Tartu, 2001, 104 p.
21. **Viljar Pihl.** The study of the substituent and solvent effects on the acidity of OH and CH acids. Tartu, 2001, 132 p.
22. **Natalia Palm.** Specification of the minimum, sufficient and significant set of descriptors for general description of solvent effects. Tartu, 2001, 134 p.
23. **Sulev Sild.** QSPR/QSAR approaches for complex molecular systems. Tartu, 2001, 134 p.
24. **Ruslan Petrukhin.** Industrial applications of the quantitative structure-property relationships. Tartu, 2001, 162 p.
25. **Boris V. Rogovoy.** Synthesis of (benzotriazolyl)carboximidamides and their application in relations with *N*- and *S*-nucleophyles. Tartu, 2002, 84 p.
26. **Koit Herodes.** Solvent effects on UV-vis absorption spectra of some solvatochromic substances in binary solvent mixtures: the preferential solvation model. Tartu, 2002, 102 p.
27. **Anti Perkson.** Synthesis and characterisation of nanostructured carbon. Tartu, 2002, 152 p.
28. **Ivari Kaljurand.** Self-consistent acidity scales of neutral and cationic Brønsted acids in acetonitrile and tetrahydrofuran. Tartu, 2003, 108 p.
29. **Karmen Lust.** Adsorption of anions on bismuth single crystal electrodes. Tartu, 2003, 128 p.
30. **Mare Piirsalu.** Substituent, temperature and solvent effects on the alkaline hydrolysis of substituted phenyl and alkyl esters of benzoic acid. Tartu, 2003, 156 p.
31. **Meeri Sassian.** Reactions of partially solvated Grignard reagents. Tartu, 2003, 78 p.
32. **Tarmo Tamm.** Quantum chemical modelling of polypyrrole. Tartu, 2003. 100 p.
33. **Erik Teinemaa.** The environmental fate of the particulate matter and organic pollutants from an oil shale power plant. Tartu, 2003. 102 p.
34. **Jaana Tammiku-Taul.** Quantum chemical study of the properties of Grignard reagents. Tartu, 2003. 120 p.
35. **Andre Lomaka.** Biomedical applications of predictive computational chemistry. Tartu, 2003. 132 p.
36. **Kostyantyn Kirichenko.** Benzotriazole – Mediated Carbon–Carbon Bond Formation. Tartu, 2003. 132 p.
37. **Gunnar Nurk.** Adsorption kinetics of some organic compounds on bismuth single crystal electrodes. Tartu, 2003, 170 p.
38. **Mati Arulepp.** Electrochemical characteristics of porous carbon materials and electrical double layer capacitors. Tartu, 2003, 196 p.
39. **Dan Cornel Fara.** QSPR modeling of complexation and distribution of organic compounds. Tartu, 2004, 126 p.
40. **Riina Mahlapuu.** Signalling of galanin and amyloid precursor protein through adenylate cyclase. Tartu, 2004, 124 p.

41. **Mihkel Kerikmäe.** Some luminescent materials for dosimetric applications and physical research. Tartu, 2004, 143 p.

42. **Jaanus Kruusma.** Determination of some important trace metal ions in human blood. Tartu, 2004, 115 p.

43. **Urmas Johanson.** Investigations of the electrochemical properties of poly-pyrrole modified electrodes. Tartu, 2004, 91 p.

44. **Kaido Sillar.** Computational study of the acid sites in zeolite ZSM-5. Tartu, 2004, 80 p.

45. **Aldo Oras.** Kinetic aspects of dATPαS interaction with $P2Y_1$ receptor. Tartu, 2004, 75 p.

46. **Erik Mölder.** Measurement of the oxygen mass transfer through the air-water interface. Tartu, 2005, 73 p.

47. **Thomas Thomberg.** The kinetics of electroreduction of peroxodisulfate anion on cadmium (0001) single crystal electrode. Tartu, 2005, 95 p.

48. **Olavi Loog.** Aspects of condensations of carbonyl compounds and their imine analogues. Tartu, 2005, 83 p.

49. **Siim Salmar.** Effect of ultrasound on ester hydrolysis in aqueous ethanol. Tartu, 2006, 73 p.

50. **Ain Uustare.** Modulation of signal transduction of heptahelical receptors by other receptors and G proteins. Tartu, 2006, 121 p.

51. **Sergei Yurchenko.** Determination of some carcinogenic contaminants in food. Tartu, 2006, 143 p.

52. **Kaido Tämm.** QSPR modeling of some properties of organic compounds. Tartu, 2006, 67 p.

53. **Olga Tšubrik.** New methods in the synthesis of multisubstituted hydra-zines. Tartu. 2006, 183 p.

54. **Lilli Sooväli.** Spectrophotometric measurements and their uncertainty in chemical analysis and dissociation constant measurements. Tartu, 2006, 125 p.

55. **Eve Koort.** Uncertainty estimation of potentiometrically measured ph and pK_a values. Tartu, 2006, 139 p.

56. **Sergei Kopanchuk.** Regulation of ligand binding to melanocortin receptor subtypes. Tartu, 2006, 119 p.

57. **Silvar Kallip.** Surface structure of some bismuth and antimony single crystal electrodes. Tartu, 2006, 107 p.

58. **Kristjan Saal.** Surface silanization and its application in biomolecule coupling. Tartu, 2006, 77 p.

59. **Tanel Tätte.** High viscosity $Sn(OBu)_4$ oligomeric concentrates and their applications in technology. Tartu, 2006, 91 p.

60. **Dimitar Atanasov Dobchev**. Robust QSAR methods for the prediction of properties from molecular structure. Tartu, 2006, 118 p.

61. **Hannes Hagu**. Impact of ultrasound on hydrophobic interactions in solutions. Tartu, 2007, 81 p.

62. **Rutha Jäger.** Electroreduction of peroxodisulfate anion on bismuth electrodes. Tartu, 2007, 142 p.

63. **Kaido Viht.** Immobilizable bisubstrate-analogue inhibitors of basophilic protein kinases: development and application in biosensors. Tartu, 2007, 88 p.
64. **Eva-Ingrid Rõõm.** Acid-base equilibria in nonpolar media. Tartu, 2007, 156 p.
65. **Sven Tamp.** DFT study of the cesium cation containing complexes relevant to the cesium cation binding by the humic acids. Tartu, 2007, 102 p.
66. **Jaak Nerut.** Electroreduction of hexacyanoferrate(III) anion on Cadmium (0001) single crystal electrode. Tartu, 2007, 180 p.
67. **Lauri Jalukse.** Measurement uncertainty estimation in amperometric dissolved oxygen concentration measurement. Tartu, 2007, 112 p.
68. **Aime Lust.** Charge state of dopants and ordered clusters formation in CaF_2:Mn and CaF_2:Eu luminophors. Tartu, 2007, 100 p.
69. **Iiris Kahn**. Quantitative Structure-Activity Relationships of environmentally relevant properties. Tartu, 2007, 98 p.
70. **Mari Reinik.** Nitrates, nitrites, N-nitrosamines and polycyclic aromatic hydrocarbons in food: analytical methods, occurrence and dietary intake. Tartu, 2007, 172 p.
71. **Heili Kasuk.** Thermodynamic parameters and adsorption kinetics of organic compounds forming the compact adsorption layer at Bi single crystal electrodes. Tartu, 2007, 212 p.
72. **Erki Enkvist.** Synthesis of adenosine-peptide conjugates for biological applications. Tartu, 2007, 114 p.
73. **Svetoslav Hristov Slavov**. Biomedical applications of the QSAR approach. Tartu, 2007, 146 p.
74. **Eneli Härk.** Electroreduction of complex cations on electrochemically polished Bi(*hkl*) single crystal electrodes. Tartu, 2008, 158 p.
75. **Priit Möller.** Electrochemical characteristics of some cathodes for medium temperature solid oxide fuel cells, synthesized by solid state reaction technique. Tartu, 2008, 90 p.
76. **Signe Viggor.** Impact of biochemical parameters of genetically different pseudomonads at the degradation of phenolic compounds. Tartu, 2008, 122 p.
77. **Ave Sarapuu.** Electrochemical reduction of oxygen on quinone-modified carbon electrodes and on thin films of platinum and gold. Tartu, 2008, 134 p.
78. **Agnes Kütt.** Studies of acid-base equilibria in non-aqueous media. Tartu, 2008, 198 p.
79. **Rouvim Kadis.** Evaluation of measurement uncertainty in analytical chemistry: related concepts and some points of misinterpretation. Tartu, 2008, 118 p.
80. **Valter Reedo.** Elaboration of IVB group metal oxide structures and their possible applications. Tartu, 2008, 98 p.
81. **Aleksei Kuznetsov.** Allosteric effects in reactions catalyzed by the cAMP-dependent protein kinase catalytic subunit. Tartu, 2009, 133 p.

82. **Aleksei Bredihhin.** Use of mono- and polyanions in the synthesis of multisubstituted hydrazine derivatives. Tartu, 2009, 105 p.

83. **Anu Ploom.** Quantitative structure-reactivity analysis in organosilicon chemistry. Tartu, 2009, 99 p.

84. **Argo Vonk.** Determination of adenosine A_{2A}- and dopamine D_1 receptor-specific modulation of adenylate cyclase activity in rat striatum. Tartu, 2009, 129 p.

85. **Indrek Kivi.** Synthesis and electrochemical characterization of porous cathode materials for intermediate temperature solid oxide fuel cells. Tartu, 2009, 177 p.

86. **Jaanus Eskusson.** Synthesis and characterisation of diamond-like carbon thin films prepared by pulsed laser deposition method. Tartu, 2009, 117 p.

87. **Marko Lätt.** Carbide derived microporous carbon and electrical double layer capacitors. Tartu, 2009, 107 p.

88. **Vladimir Stepanov.** Slow conformational changes in dopamine transporter interaction with its ligands. Tartu, 2009, 103 p.

89. **Aleksander Trummal.** Computational Study of Structural and Solvent Effects on Acidities of Some Brønsted Acids. Tartu, 2009, 103 p.

90. **Eerold Vellemäe.** Applications of mischmetal in organic synthesis. Tartu, 2009, 93 p.

91. **Sven Parkel.** Ligand binding to 5-HT_{1A} receptors and its regulation by Mg^{2+} and Mn^{2+}. Tartu, 2010, 99 p.

92. **Signe Vahur.** Expanding the possibilities of ATR-FT-IR spectroscopy in determination of inorganic pigments. Tartu, 2010, 184 p.

93. **Tavo Romann**. Preparation and surface modification of bismuth thin film, porous, and microelectrodes. Tartu, 2010, 155 p.

94. **Nadežda Aleksejeva.** Electrocatalytic reduction of oxygen on carbon nanotube-based nanocomposite materials. Tartu, 2010, 147 p.

95. **Marko Kullapere.** Electrochemical properties of glassy carbon, nickel and gold electrodes modified with aryl groups. Tartu, 2010, 233 p.

96. **Liis Siinor.** Adsorption kinetics of ions at Bi single crystal planes from aqueous electrolyte solutions and room-temperature ionic liquids. Tartu, 2010, 101 p.

97. **Angela Vaasa.** Development of fluorescence-based kinetic and binding assays for characterization of protein kinases and their inhibitors. Tartu 2010, 101 p.

98. **Indrek Tulp.** Multivariate analysis of chemical and biological properties. Tartu 2010, 105 p.

99. **Aare Selberg.** Evaluation of environmental quality in Northern Estonia by the analysis of leachate. Tartu 2010, 117 p.

100. **Darja Lavõgina.** Development of protein kinase inhibitors based on adenosine analogue-oligoarginine conjugates. Tartu 2010, 248 p.

101. **Laura Herm.** Biochemistry of dopamine D_2 receptors and its association with motivated behaviour. Tartu 2010, 156 p.

102. **Terje Raudsepp.** Influence of dopant anions on the electrochemical properties of polypyrrole films. Tartu 2010, 112 p.

103. **Margus Marandi.** Electroformation of Polypyrrole Films: *In-situ* AFM and STM Study. Tartu 2011, 116 p.

104. **Kairi Kivirand.** Diamine oxidase-based biosensors: construction and working principles. Tartu, 2011, 140 p.

105. **Anneli Kruve.** Matrix effects in liquid-chromatography electrospray mass-spectrometry. Tartu, 2011, 156 p.

106. **Gary Urb.** Assessment of environmental impact of oil shale fly ash from PF and CFB combustion. Tartu, 2011, 108 p.

107. **Nikita Oskolkov.** A novel strategy for peptide-mediated cellular delivery and induction of endosomal escape. Tartu, 2011, 106 p.

108. **Dana Martin.** The QSPR/QSAR approach for the prediction of properties of fullerene derivatives. Tartu, 2011, 98 p.

109. **Säde Viirlaid.** Novel glutathione analogues and their antioxidant activity. Tartu, 2011, 106 p.

110. **Ülis Sõukand.** Simultaneous adsorption of Cd^{2+}, Ni^{2+}, and Pb^{2+} on peat. Tartu, 2011, 124 p.

111. **Lauri Lipping.** The acidity of strong and superstrong Brønsted acids, an outreach for the "limits of growth": a quantum chemical study. Tartu, 2011, 124 p.

112. **Heisi Kurig.** Electrical double-layer capacitors based on ionic liquids as electrolytes. Tartu, 2011, 146 p.

113. **Marje Kasari.** Bisubstrate luminescent probes, optical sensors and affinity adsorbents for measurement of active protein kinases in biological samples. Tartu, 2012, 126 p.

114. **Kalev Takkis.** Virtual screening of chemical databases for bioactive molecules. Tartu, 2012, 122 p.

115. **Ksenija Kisseljova**. Synthesis of aza-β^3-amino acid containing peptides and kinetic study of their phosphorylation by protein kinase A. Tartu, 2012, 104 p.

116. **Riin Rebane.** Advanced method development strategy for derivatization LC/ESI/MS. Tartu, 2012, 184 p.

117. **Vladislav Ivaništšev.** Double layer structure and adsorption kinetics of ions at metal electrodes in room temperature ionic liquids. Tartu, 2012, 128 p.

118. **Irja Helm.** High accuracy gravimetric Winkler method for determination of dissolved oxygen. Tartu, 2012, 139 p.

119. **Karin Kipper.** Fluoroalcohols as Components of LC-ESI-MS Eluents: Usage and Applications. Tartu, 2012, 164 p.

120. **Arno Ratas.** Energy storage and transfer in dosimetric luminescent materials. Tartu, 2012, 163 p.

121. **Reet Reinart-Okugbeni**. Assay systems for characterisation of subtype-selective binding and functional activity of ligands on dopamine receptors. Tartu, 2012, 159 p.

122. **Lauri Sikk.** Computational study of the Sonogashira cross-coupling reaction. Tartu, 2012, 81 p.
123. **Karita Raudkivi.** Neurochemical studies on inter-individual differences in affect-related behaviour of the laboratory rat. Tartu, 2012, 161 p.
124. **Indrek Saar.** Design of GalR2 subtype specific ligands: their role in depression-like behavior and feeding regulation. Tartu, 2013, 126 p.
125. **Ann Laheäär.** Electrochemical characterization of alkali metal salt based non-aqueous electrolytes for supercapacitors. Tartu, 2013, 127 p.
126. **Kerli Tõnurist.** Influence of electrospun separator materials properties on electrochemical performance of electrical double-layer capacitors. Tartu, 2013, 147 p.
127. **Kaija Põhako-Esko.** Novel organic and inorganic ionogels: preparation and characterization. Tartu, 2013, 124 p.
128. **Ivar Kruusenberg.** Electroreduction of oxygen on carbon nanomaterial-based catalysts. Tartu, 2013, 191 p.
129. **Sander Piiskop.** Kinetic effects of ultrasound in aqueous acetonitrile solutions. Tartu, 2013, 95 p.
130. **Ilona Faustova**. Regulatory role of L-type pyruvate kinase N-terminal domain. Tartu, 2013, 109 p.
131. **Kadi Tamm.** Synthesis and characterization of the micro-mesoporous anode materials and testing of the medium temperature solid oxide fuel cell single cells. Tartu, 2013, 138 p.
132. **Iva Bozhidarova Stoyanova-Slavova.** Validation of QSAR/QSPR for regulatory purposes. Tartu, 2013, 109 p.
133. **Vitali Grozovski**. Adsorption of organic molecules at single crystal electrodes studied by *in situ* STM method. Tartu, 2014, 146 p.
134. **Santa Veikšina**. Development of assay systems for characterisation of ligand binding properties to melanocortin 4 receptors. Tartu, 2014, 151 p.
135. **Jüri Liiv.** PVDF (polyvinylidene difluoride) as material for active element of twisting-ball displays. Tartu, 2014, 111 p.
136. **Kersti Vaarmets.** Electrochemical and physical characterization of pristine and activated molybdenum carbide-derived carbon electrodes for the oxygen electroreduction reaction. Tartu, 2014, 131 p.
137. **Lauri Tõntson**. Regulation of G-protein subtypes by receptors, guanine nucleotides and Mn^{2+}. Tartu, 2014, 105 p.
138. **Aiko Adamson.** Properties of amine-boranes and phosphorus analogues in the gas phase. Tartu, 2014, 78 p.
139. **Elo Kibena**. Electrochemical grafting of glassy carbon, gold, highly oriented pyrolytic graphite and chemical vapour deposition-grown graphene electrodes by diazonium reduction method. Tartu, 2014, 184 p.
140. **Teemu Näykki.** Novel Tools for Water Quality Monitoring – From Field to Laboratory. Tartu, 2014, 202 p.
141. **Karl Kaupmees.** Acidity and basicity in non-aqueous media: importance of solvent properties and purity. Tartu, 2014, 128 p.

142. **Oleg Lebedev**. Hydrazine polyanions: different strategies in the synthesis of heterocycles. Tartu, 2015, 118 p.
143. **Geven Piir.** Environmental risk assessment of chemicals using QSAR methods. Tartu, 2015, 123 p.
144. **Olga Mazina.** Development and application of the biosensor assay for measurements of cyclic adenosine monophosphate in studies of G protein-coupled receptor signalinga. Tartu, 2015, 116 p.
145. **Sandip Ashokrao Kadam.** Anion receptors: synthesis and accurate binding measurements. Tartu, 2015, 116 p.
146. **Indrek Tallo.** Synthesis and characterization of new micro-mesoporous carbide derived carbon materials for high energy and power density electrical double layer capacitors. Tartu, 2015, 148 p.
147. **Heiki Erikson.** Electrochemical reduction of oxygen on nanostructured palladium and gold catalysts. Tartu, 2015, 204 p.
148. **Erik Anderson.** *In situ* Scanning Tunnelling Microscopy studies of the interfacial structure between Bi(111) electrode and a room temperature ionic liquid. Tartu, 2015, 118 p.
149. **Girinath G. Pillai.** Computational Modelling of Diverse Chemical, Bio-chemical and Biomedical Properties. Tartu, 2015, 140 p.
150. **Piret Pikma.** Interfacial structure and adsorption of organic compounds at Cd(0001) and Sb(111) electrodes from ionic liquid and aqueous electrolytes: an *in situ* STM study. Tartu, 2015, 126 p.
151. **Ganesh babu Manoharan.** Combining chemical and genetic approaches for photoluminescence assays of protein kinases. Tartu, 2016, 126 p.
152. **Carolin Siimenson**. Electrochemical characterization of halide ion adsorption from liquid mixtures at Bi(111) and pyrolytic graphite electrode surface. Tartu, 2016, 110 p.
153. **Asko Laaniste.** Comparison and optimisation of novel mass spectrometry ionisation sources. Tartu, 2016, 156 p.
154. **Hanno Evard.** Estimating limit of detection for mass spectrometric analysis methods. Tartu, 2016, 224 p.
155. **Kadri Ligi.** Characterization and application of protein kinase-responsive organic probes with triplet-singlet energy transfer. Tartu, 2016, 122 p.
156. **Margarita Kagan.** Biosensing penicillins' residues in milk flows. Tartu, 2016, 130 p.
157. **Marie Kriisa.** Development of protein kinase-responsive photolumine-scent probes and cellular regulators of protein phosphorylation. Tartu, 2016, 106 p.
158. **Mihkel Vestli.** Ultrasonic spray pyrolysis deposited electrolyte layers for intermediate temperature solid oxide fuel cells. Tartu, 2016, 156 p.
159. **Silver Sepp**. Influence of porosity of the carbide-derived carbon on the properties of the composite electrocatalysts and characteristics of polymer electrolyte fuel cells. Tartu, 2016, 137 p.
160. **Kristjan Haav**. Quantitative relative equilibrium constant measurements in supramolecular chemistry. Tartu, 2017, 158 p.

161. **Anu Teearu**. Development of MALDI-FT-ICR-MS methodology for the analysis of resinous materials. Tartu, 2017, 205 p.

162. **Taavi Ivan**. Bifunctional inhibitors and photoluminescent probes for studies on protein complexes. Tartu, 2017, 140 p.

163. **Maarja-Liisa Oldekop**. Characterization of amino acid derivatization reagents for LC-MS analysis. Tartu, 2017, 147 p.

164. **Kristel Jukk**. Electrochemical reduction of oxygen on platinum- and palladium-based nanocatalysts. Tartu, 2017, 250 p.

165. **Siim Kukk**. Kinetic aspects of interaction between dopamine transporter and *N*-substituted nortropane derivatives. Tartu, 2017, 107 p.

166. **Birgit Viira**. Design and modelling in early drug development in targeting HIV-1 reverse transcriptase and Malaria. Tartu, 2017, 172 p.

167. **Rait Kivi**. Allostery in cAMP dependent protein kinase catalytic subunit. Tartu, 2017, 115 p.

168. **Agnes Heering**. Experimental realization and applications of the unified acidity scale. Tartu, 2017, 123 p.

169. **Delia Juronen**. Biosensing system for the rapid multiplex detection of mastitis-causing pathogens in milk. Tartu, 2018, 85 p.

170. **Hedi Rahnel.** ARC-inhibitors: from reliable biochemical assays to regulators of physiology of cells. Tartu, 2018, 176 p.

171. **Anton Ruzanov.** Computational investigation of the electrical double layer at metal–aqueous solution and metal–ionic liquid interfaces. Tartu, 2018, 129 p.

172. **Katrin Kestav.** Crystal Structure-Guided Development of Bisubstrate-Analogue Inhibitors of Mitotic Protein Kinase Haspin. Tartu, 2018, 166 p.

173. **Mihkel Ilisson.** Synthesis of novel heterocyclic hydrazine derivatives and their conjugates. Tartu, 2018, 101 p.

174. **Anni Allikalt.** Development of assay systems for studying ligand binding to dopamine receptors. Tartu, 2018, 160 p.

175. **Ove Oll.** Electrical double layer structure and energy storage characteristics of ionic liquid based capacitors. Tartu, 2018, 187 p.

176. **Rasmus Palm.** Carbon materials for energy storage applications. Tartu, 2018, 114 p.

177. **Jörgen Metsik.** Preparation and stability of poly(3,4-ethylenedioxythiophene) thin films for transparent electrode applications. Tartu, 2018, 111 p.

178. **Sofja Tšepelevitš.** Experimental studies and modeling of solute-solvent interactions. Tartu, 2018, 109 p.

179. **Märt Lõkov.** Basicity of some nitrogen, phosphorus and carbon bases in acetonitrile. Tartu, 2018, 104 p.

180. **Anton Mastitski**. Preparation of α-aza-amino acid precursors and related compounds by novel methods of reductive one-pot alkylation and direct alkylation. Tartu, 2018, 155 p.

181. **Jürgen Vahter**. Development of bisubstrate inhibitors for protein kinase CK2. Tartu, 2019, 186 p.