UNIVERSITY OF LIÈGE

Faculty of Sciences

The use of Capillary Electrophoresis coupled to High Resolution Mass Spectrometry as an Analytical Tool for Biomolecule Analysis

Supervisor

Prof. Edwin De Pauw

Mass Spectrometry Laboratory, ULiège

Co-Supervisor

Prof. Gauthier Eppe

Mass Spectrometry Laboratory, ULiège

Academic year 2018-2019

Dissertation submitted by

Cédric Delvaux

for the fulfilment of the degree of

Doctor of Philosophy in Sciences

Jury

Supervisor	Prof. Edwin de Pauw	Université de Liège
Co-supervisor	Prof. Gauthier Eppe	Université de Liège
President	Prof. Bernard Leyh	Université de Liège
Secretary	Dr. Johann Far	Université de Liège
Member	Prof. Marianne Fillet	Université de Liège
Member	Prof. Ann Van Schepdael	Katholieke Universiteit Leuven
Member	Prof. Yannis François	Université de Strasbourg

Acknowledgments – Remerciements

Many people were involved during the completion of this thesis, either directly or indirectly. I received help from numerous people, in my professional environment of course, but also from my personal life. This section aims at expressing my gratitude to those who supported me during these sometimes-demanding years.

Of course, no thesis can be carried out and completed without an appropriate research environment. Consequently, I would like to start by acknowledging the University of Liège and the Mass Spectrometry Laboratory for the facility and the large instrument park, allowing me to get in touch with various instruments, techniques and mass spectrometry expertise. I would also like to thank the University of Liège for the different mandates as a teaching assistant, giving me the opportunity to perform my thesis while teaching chemistry at the same time. This combination was an excellent way for me to explore different professional aspects of chemistry (i.e. teaching and academic research) and allowed me to get further confidence for my future career. A thesis is also the occasion to travel and meet other research groups in (inter)national conferences. I would like to thank the University of Liège for giving me the opportunity to attend many great conferences and meetings. The fruitful discussions and inspiring ideas from these events were the basis of many results developed in this work. The "Patrimoine de l'Université de Liège" is also acknowledged for the financial support.

In the long list of people who should be acknowledged, I would like to start with Professor Edwin De Pauw for the opportunity to be part of his research group. As my promotor, he gave me the liberty of research that many of my coworkers envy and let me freely pursue the research topics I wanted to explore. Also, because of the many different aspects of research he develops at the Laboratory of Mass Spectrometry, fruitful collaborations naturally arise from the daily contact with other PhD students or post-doctoral researchers. Consequently, almost every developed chapter in this thesis relies on a collaborative work.

I would also like to express my gratitude to my jury members, especially Professor Ann Van Schepdael, Professor Marianne Fillet, and Professor Yannis François for accepting to review this manuscript and provide their valuable feedbacks. It is a great pleasure for me. Professor Bernard Leyh and Doctor Johann Far are also thanked for accepting to be respectively the President and Secretary of this thesis jury.

A special thanks must be directed to Dr. Johann Far, who taught me a lot about mass spectrometry, capillary electrophoresis, liquid chromatography and practical aspects of analytical chemistry in general. Since we started working together almost 10 years ago (I was then in my 3rd Bachelor year), I've never stopped learning from him. His pedagogy (although the methods are sometimes quite questionable), encouragements, technical assistance, and especially the constant support he provided me for reviewing presentations, posters, papers as well as discussions on experimental results and interpretation made this PhD thesis possible. The door of his office is always open for advice and help and I can never thank him enough for his decisive contribution to this work, but also during the last 10 years.

I would also like to express my gratefulness to Professor Bernard Joris from the Center of Protein Engineering at ULiège for the wonderful collaborations about the peptidoglycan characterization. It's always a pleasure for analysts to develop methods that can directly be used by other research groups to address real-life issues. Among his research group, I must specially thank Madeleine Boulanger and Marjorie Dauvin for our great collaborative work on these projects. Madeleine is also warmly thanked for helping me in the design of my thesis cover.

I would also thank Dr. Philippe Massonnet for involving me in his project about the characterization of disulfide bonds connectivity in peptides. As his thesis focused on the use of ion mobility for the study of such structured peptides, I could take advantage of his great knowledge about these systems in our collaboration. Christopher Kune is also acknowledged for the theoretical support and ion mobility experiments in this project... and be an active member of the "Jafar" team.

I would like to thank all the members of the Mass Spectrometry Laboratory for the past six years I have spent in their company. Special thanks go to my colleagues of -1 floor Elodie, Mathieu, Emeline, Caroline, Donatello, Julien and Mitch. The three interconnected offices were always a place for discussions, guidance, shared experience and support. My thesis would have been a lot less fun without my great "-1" colleagues. In particular, Elodie is specially thanked for helping me in various sample preparations, sample list configurations and thoroughly reviewing my thesis to the final version. Gabriel, Nicolas, Dominique, Max, Lisette, Nanou and Nancy are thanked for their technical assistance and access to the Q Exactive when necessary.

Concerning the teaching part of my thesis, I would like to warmly thank Stephane for the last 6 years at the Analytical Chemistry Laboratories. It was a real pleasure for me to be his colleague in this teaching activity. His constant availability and great care in the preparation of the different laboratory sessions were always appreciated. I will miss these Fridays… Dr. Cédric Malherbe is acknowledged as my predecessor for providing me great teaching contents and valuable advice.

Enfin, je termine ces remerciements en français pour dire à ma famille et à Jonathan à quel point ils ont été autant de piliers pour moi durant cette aventure, comme à chaque fois. Mes parents, Isabelle et Philippe, sont tout spécialement remerciés pour leur support constant durant mes 5 années d'études et 6 années de thèse, me poussant toujours vers l'avant et me permettant d'arriver où je suis aujourd'hui. Nennenne, en tant que marraine et grand-mère, est également remerciée d'avoir été cette personne si spéciale, éveillant en moi l'appétit d'apprendre et de comprendre et contribuant probablement à façonner la personne que je suis devenue. Ma petite sœur Valentine doit également figurer dans cette liste de personnes proches à remercier. Je suis très heureux de la compter parmi mes proches et la remercie pour ses encouragements « cactus ». Enfin, j'aimerais tout spécialement remercier Jonathan pour son soutien sans faille et sa grande compréhension dans les étapes de finalisation de cette thèse. Savoir que nous pouvons mutuellement compter l'un sur l'autre, peu importe les circonstances de la vie est une grande force, pour laquelle je ne peux lui être que profondément reconnaissant. Je le remercie pour tout ce que nous avons déjà accompli et me réjouis déjà de tout ce que nous accomplirons ensemble à l'avenir.

Abbreviations

BGE	BackGround Electrolyte
CE	Capillary Electrophoresis
CZE	Capillary Zone Electrophoresis
DNA	DeoxyriboNucleic Acid
DTIMS	Drift Tube Ion Mobility Spectrometry
EOF	ElectroOsmotic Flow
ESI	ElectroSpray Ionization
ETD	Electron Transfer Dissociation
FT-ICR	Fourier Transform Ion Cyclotron Resonance
FT-MS	Fourier Transform Mass Spectrometry
HCD	Higher energy Collision induced Dissociation
HPLC	High Performance Liquid Chromatography
HR-MS	High Resolution – Mass Spectrometry
ID	Inner Diameter
IM-MS	Ion Mobility – Mass Spectrometry
IMS	Ion mobility Spectrometry
IR	InfraRed
IRMPD	InfraRed MultiPhoton Dissociation
LIF	Laser-Induced Fluorescence
LC	Liquid Chromatography
m-A_2pm	*meso*-2,6-diaminopimelic acid
MALDI	Matrix Assisted Laser Desorption Ionisation
MRM	Multiple Reaction Monitoring
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
NACE	Non Aqueous Capillary Electrophoresis
nanoESI	nanoElectroSpray Ionization
OD	Outer Diameter
PGN	Peptidoglycan
PRM	Parallel Reaction Monitoring
PTM	Post-Translational Modification
RNA	RiboNucleic Acid
RP-HPLC	Reversed Phase - High Performance Liquid Chromatography
SIM	Single Ion Monitoring
SRM	Single Reaction Monitoring
TQ	Triple Quadrupole
TOF	Time Of Flight
TWIMS	Travelling Wave Ion Mobility
UPLC	Ultra Performance Liquid Chromatography
UV	UltraViolet
UVPD	UltraViolet PhotoDissociation

Table of contents

CHAPTER 1

Objectives and Strategies

The interest for Capillary Electrophoresis (CE) as a separation technique has recently reattracted research interest mainly due to the introduction of efficient interfaces for its hyphenation to Mass Spectrometry (MS). As a separation method, CE provides many advantageous analytical aspects including the variety of available electrophoretic modes, rapid analysis, small sample consumption and high separation efficiency. In addition, as the separation mechanism in CE is based on charge and "size" of the analytes, it provides an orthogonal separation to classical liquid chromatography (LC) separations based on partition coefficients between stationary and mobile phases.

The objective of this thesis is to take benefit of the above-mentioned advantages by using Capillary Electrophoresis (CE) coupled to High Resolution Mass Spectrometry (HR-MS) as an analytical tool for the analysis of various biomolecules. In particular, this work aims at deploying CE-MS when analytical challenges are addressed and requires its use, while other separation methods only provide partial information.

The first part of this thesis focuses on the implementation of CE-MS as a valuable analytical method for the analysis of peptides and muropeptides (i.e. glycopeptides associated with the bacterial cell wall). The increased characterization efficiency provided by the CE-MS methods is demonstrated in comparison to established LC(-MS) methods.

A second important part of this thesis was dedicated to the comparison between capillary electrophoresis and ion mobility separations. In this case, these two analytical techniques are both based on ion migration under the influence of an electric field either in the gas phase (ion mobility) or in solution (capillary electrophoresis). The comparison of these techniques first allows to assess their respective separation efficiencies but more importantly, to compare migration behaviors in both separation media. As it will be later detailed in this manuscript, the examination of the migration behaviors at equivalent charge states allows to address the retention of structural features from the solution to the gas phase.

The thesis is divided into seven chapters briefly introduced here below:

Chapter 2 reviews capillary electrophoresis and mass spectrometry and their respective analytical aspects. Various interfaces to hyphenate CE with MS are then reviewed and their advantages and drawbacks are critically discussed. To conclude, CE-MS is compared to classical analytical methods such as LC-MS or more recently IM-MS.

Chapter 3 and 4 demonstrate the analytical merits of CE-MS for the analysis and characterization of peptidoglycan fragments including muropeptides (Chapter 3) and cytoplasmic peptidoglycan-derived peptides (Chapter 4). This work was performed within the framework of a collaboration with the "Centre d'Ingénierie des Protéines" (CIP) or "Center for Protein Engineering" from the University of Liège. In particular, the ability to separate and detect species that are otherwise co-eluted by the classical reversed-phase chromatography method opens new opportunities for in-depth bacterial characterization. An absolute quantitation method was also developed based on an isotopically labelled standard.

Chapter 5 was dedicated to the comparison CE-MS and IM-MS for the separation of peptides bearing two intramolecular disulfide bonds in the gas phase and in solution, respectively. This work was supported by theoretical calculations for the determination of various physicochemical properties of the analytes including structural features in both migration media. Moreover, the comparison of the mobilities determined in the gas phase and in solution allowed to get insights into the influence of the electrospray desolvation process on the structural preservation of these species.

Chapter 6 focuses on the characterization of the electroosmotic flow (EOF) by two separate methods (neutral and charged markers). As this work was entirely performed on a low-sheath flow interface, the characterization of its specific suction effect was also performed and discussed.

Finally, Chapter 7 provides a summary of the different analytical methods and results obtained during this thesis and brings some future prospects to be examined.

CHAPTER 2

General Introduction:

The place of Capillary Electrophoresis coupled to Mass Spectrometry in the analytical toolbox

2.1. Capillary Electrophoresis

2.1.1. The discovery of Electrophoresis

Although electrophoresis has recently regained interest as a versatile analytical technique, its first description goes back to 1807 as Professors Strakhov and Reuss described for the first time the electrophoresis phenomenon on clay particles. In their experiment, they applied a constant electric field on clay particles dispersed in water and noticed the migration of these particles under the influence of the electric field[1]. Their work also led to the first description of electro-osmosis, which is discussed in detail in section "The Electro-Osmotic Flow (EOF)". Further work including experimental and theoretical descriptions of electrophoresis provided the basis for Tiselius's introduction of electrophoresis as a separation technique by the 1930s[2,3]. Tiselius was awarded the Nobel Prize in Chemistry in 1948 "for his research on electrophoresis and adsorption analysis, especially for his discoveries concerning the complex nature of the serum proteins."

2.1.2. From Tiselius' Apparatus to Zone Electrophoresis

Improvements of Tiselius' approach and in particular, the introduction of solids (such as paper) or gel matrices to separate compounds into discrete and stable bands (zones) eventually led to the widespread application of electrophoresis in biochemistry. By the 1960s, the constant improvement of gel electrophoresis methods allowed the separation of biologically relevant macromolecules including proteins and nucleic acids, assisting the rise of molecular biology[4–7]. Finally, the introduction of electrophoresis performed in open tubes (or capillaries) by Hjertén in 1967[8] circumvented various frequently encountered problems with gel electrophoresis (long analysis times, low efficiencies, and difficulties in detection and automation). Since then, the development in the manufacturing of narrow capillaries allowed CE to become a high-performance analytical technique.

2.1.3. The rise of capillary electrophoresis as a versatile separation method

The development of CE in the last decades has also led to the introduction of various modes of operations, each with a specific separation approach. The major modes of CE include Capillary Zone Electrophoresis (CZE), Capillary Gel Electrophoresis (CGE), Capillary IsoTachoPhoresis (CITP), Capillary ElectroChromatography (CEC), Micellar ElectroKinetic Chromatography (MEKC) and Capillary IsoElectric Focusing (CIEF). The following table summarizes these different modes with a brief overview of their specific separation mechanisms:

Table 2.1 - Modes of operations of capillary electrophoresis and description of the related separation mechanisms

Mode	Separation basis	Description
CZE	Free solution mobility	The capillary is filled with an appropriate Background Electrolyte (BGE) and the analytes are separated due to differences in their apparent electrophoretic mobilities. This mode of separation is further detailed in the next section.
CGE	Size	The capillary is filled with a polymer (e.g. polyacrylamide or agarose) which acts as a molecular sieve during the separation. Principally used for macromolecules such as DNA and SDS-saturated proteins.

CITP	Moving boundaries	The sample is introduced between a leading electrolyte with a high ionic mobility and a terminating electrolyte with a low ionic mobility. The analytes of interest have an intermediate ionic mobility. The application of the electric field results in a low electrical field in the leading electrolyte zone and a high electrical field in the terminating electrolyte. In fact, the electric field auto-adjusts in each zone to maintain a constant velocity with the lowest field across the zone with the highest mobility. At the end, the analytes migrate at the same velocity as separated zones.
CEC	Interaction with a stationary phase (not necessarily immobilized) under the influence of an electric field	The capillary is either coated with a stationary phase, or the BGE is spiked with solubilized or dispersed stationary phase. The EOF acts as the fluidic pump. The difference of migration is due to the repartition coefficients between the mobile and stationary phases and the electrophoretic mobilities of the analytes, mixing electrophoresis and chromatography in a single device. Typically used for chiral separations were the BGE is spiked with a chiral selector including crown ethers or cyclodextrins.
MEKC	Interaction with (migrating) micelles	The capillary is filled with a BGE containing surfactants at concentrations above the critical micelle concentration to form surfactant micelles. Differential interactions between the micelle and analytes is the basis for the separation.
CIEF	Isoelectric point	The capillary is filled with a mixture of analytes and ampholytes. A pH gradient is formed through the capillary using a basic solution at the cathode and an acidic solution at the anode. Focusing takes place as the charged ampholytes and analytes migrate until they reach the region of the gradient corresponding to their pI value. Once the focusing is complete, a mobilization step flushes the zones to the detection point (e.g. by application of pressure).

As can be seen from literature, CE-based methods are rapidly emerging in numerous analytical fields, with the implementation of robust and reliable methods for the analysis of a wide range of analytes. Thanks to its different operation modes, CE is progressively positioning as a central analytical approach for the analysis of major categories of biomolecules due to its versatility and the complementary information compared to other analytical approaches. For example, CE is now a key method for the separation of proteins thanks to CGE[9]. In addition, CGE is also becoming a reference method in the biopharmaceutical industry to support analytical characterization, process development, and quality control of therapeutic recombinant monoclonal antibodies (rMAbs)[10–13]. In parallel, CE has recently turned into a powerful approach for carbohydrate analyses[14] with high resolution separations of glycans combined to structural identification including positional isomers, as well as the type of linkage[15] (alpha or beta linkage). Forensic science has also been largely impacted by the recent developments of CE where the introduction of CE-based methods for the detection of various drugs in complex matrices constitutes an attractive alternative to HPLC and/or GC methodologies. In the forensic context, CE also plays an essential role in the separation of DNA and the development of methods for high throughput DNA typing[16,17]. For very large

systems, CE is able to analyze (very) large biological assemblies such as viruses, bacteria, fungi and whole cells[18]. Finally, the determination of binding constants is also possible thanks to "affinity CE"[19].

Despite the various available CE modes, this thesis focuses on the use of Capillary Zone Electrophoresis coupled to Mass Spectrometry (CZE-MS) for the development of analytical methods. Therefore, only this mode will be described in detail.

2.1.4. Capillary Zone Electrophoresis

Capillary zone electrophoresis (CZE) is the most widely used mode due to its simplicity of operation and versatility. CZE applies for the analysis of a wide range of polar and ionic species[20] including biomolecules such as amino acids[21], peptides and proteins[22] but also a wide range of enantiomers[23] (with the use of chiral selectors). For example, CZE can be applied on proteins for purity validation, screening charge variants[24,25] and conformational studies[26,27]. In CZE, the separation is based on the differential migrations of the analytes under the influence of the electric field to form discrete zones in the capillary. Both anionic and cationic analytes can be separated in a single run due to the electro-osmotic flow (see section "The Electro-Osmotic Flow (EOF)" for details). However, CZE does not separate the neutral analytes which co-elute with the EOF. The detection of the analytes is generally performed at the outlet of the separation capillary (as performed for CE-MS) or in a small detection window close to the outlet (as performed for CE-UV for example).

2.1.5. Theory of Capillary Electrophoresis

In capillary electrophoresis, the separation is based on the application of an electric field on a capillary filled with a background electrolyte (BGE). Due to the application of the electric field, two electro-driven phenomena take place within the separation capillary. First, the interaction of the electric field with the ionic analytes leads to their migration due to the electric force that is exerted on them. This phenomenon, referred as the "electrophoretic mobility", is detailed in the next section. In addition, a bulk solution displacement also takes place due to the application of the electric field. This phenomenon, referred as the "electroosmotic flow", originates from the diffuse double layer of ions at the capillary surface. The combination of these parameters is responsible for the differential migration of the analytes, and by extension, the detected migration time. In this context, the optimization of both parameters (i.e. electrophoretic mobility and electroosmotic flow) needs to be considered when developing analytical methods. The following sections detail each parameter separately.

2.1.5.1. The Electrophoretic Mobility

The term "Electrophoresis" is defined by IUPAC as "the motion of colloidal particles in an electric field". When electrophoresis is used as a separation method, the separation is based on differential analyte velocities under the influence of an electric field. The velocity of a particular ion is given by equation (1):

$$v = \mu_e . E \quad (1)$$

Where v is the velocity of the ion during the electrophoretic process ($m.s^{-1}$)

μ_e is the electrophoretic mobility of the ion ($m^2.V^{-1}.s^{-1}$)

E is the applied electric field ($V.m^{-1}$)

This equation leads to the definition of the mobility: "the mobility of a specific particle is defined as the ratio of its velocity to the magnitude of the electric field".

The applied electric field on the capillary is directly determined by the applied voltage during the separation and the length of the capillary. It is therefore identical (and most of the time constant) for all analytes. The electrophoretic mobility, however, is characteristic of the analyte. For a given analyte, its electrophoretic mobility is the direct balance between the electric force exerted by the electric field (equation 2) and the friction force exerted by the surrounding BGE (approximated for a spherical ion) (equation 3), which is the combination of the Stokes (static viscosity) and the Einstein (diffusion and Brownian motion) equations.

$$F_{electric} = q.E \quad (2)$$

$$F_{friction} = -6\pi.\eta.R_h.v \quad (3)$$

Where q is the charge of the analyte (C)

η is the viscosity of the background electrolyte ($kg.m^{-1}.s^{-1}$)

R_h is the ion's hydrodynamic radius (m)

Once the electric field is applied to the capillary, the ion reaches a steady velocity imposed by the balance between these forces. At this point, these forces apply in opposite directions on the ion, as described by equation 4:

$$q.E = 6\pi.\eta.R_h.v \rightarrow v = \frac{q}{6\pi.\eta.R_h}E \quad (4)$$

The comparison between equations (1) and (4) yields the analytical expression of the electrophoretic mobility (equation 5), also known as the Stokes-Einstein equation:

$$\mu_e = \frac{q}{6\pi.\eta.R_h} \quad (5)$$

This equation implies that:

- highly charged analytes migrate faster than less charged analytes (influence of q)
- compact analytes migrate faster than extended analytes (influence of R_h)
- all parameters impacting the viscosity of the migration BGE including temperature, addition of organic solvent, etc… directly influence the migration of all analytes at once

It is also important to highlight that for aqueous separations of biomolecules, the charge in solution is directly imposed by the different pK_a values of the ionizable residues contained in the structure. Therefore, the pH of the BGE has a tremendous importance in most analytical methods developed in this work.

2.1.5.2. The Electro-Osmotic Flow (EOF)

Besides the electrophoretic mobility of the analytes, the use of an electric field during electrophoresis generates another electro-driven phenomenon called electro-osmosis. Electro-osmosis consists in a flow of bulk solution inside the capillary upon the application of the electric field. This phenomenon originates from the interaction between the surface of the capillary and the BGE that leads to the formation of a diffuse double layer (accumulation of an excess of charges) in the solution in direct contact with the inner part of the capillary. In general, these charges are

caused by the ionization of groups at the surface by acid-base equilibria. For example, in bare fused silica capillaries, EOF is strongly correlated to pH, highlighting the influence of the ionization of the capillary's silanol groups ($SiOH \rightleftharpoons SiO^- + H^+$) in the EOF phenomenon. Alternatively, EOF was also described in non-ionizable materials such as PEEK or Teflon, indicating that the non-specific adsorption of ionic species at the surface can also generate EOF.

In fact, the presence of charges at the capillary surface (ionized silanol groups or adsorption of ions) results in an accumulation of counter-ions near the capillary wall to counterbalance these charges. This accumulation of ions can be viewed as two distinct regions. The closest region to the capillary wall, referred as the compact layer, consists in a region where counter-ions are tightly retained by the charges present on the capillary wall. The second part of the double layer constitutes the so-called "diffuse region" where counter-ions are still in excess but not sufficiently retained to prevent migration. Once the electric field is applied across the capillary, the counter-ions of the diffuse layer migrate toward the electrode of opposite charge. Because these counter-ions are hydrated, their migration is accompanied by their hydration spheres, which results in a direct movement of bulk solution. This bulk flow has a typical flat plug-like flow profile, i.e. a flow profile of equivalent velocity across the capillary section. Similarly to the electrophoretic velocity, the EOF velocity can be described by its velocity (equation 6):

$$v = \mu_{EOF} . E \quad (6)$$

Where v is the velocity of the electro-osmosis ($m.s^{-1}$)

μ_{EOF} is the electro-osmotic mobility ($m^2.V^{-1}.s^{-1}$)

E is the applied electric field ($V.m^{-1}$)

The electro-osmotic mobility is defined by several physico-chemical parameters that are directly linked to the BGE and the double layer. As a result, the electro-osmotic mobility can be expressed as:

$$\mu_{EOF} = \frac{\varepsilon . \zeta}{\eta} \quad (7)$$

Where ε is the dielectric constant of the BGE ($F.m^{-1}$or $kg.m.V^{-2}.s^{-2}$)

ζ is the zeta potential (V)

η is the viscosity of the BGE ($kg.m^{-1}.s^{-1}$)

As can be seen in equation (7), the magnitude of EOF can be regulated by adjusting the capillary surface potential (ζ) or BGE parameters (η and ε). Unfortunately, in most cases, a change in one of these parameters also affects the electrophoretic mobility of the analytes. The consequences of such changes are difficult to predict and need to be carefully assessed in regard with other relevant separation parameters. It is therefore important to consider the optimization of the conditions on both EOF and analyte mobilities to obtain efficient separations.

The dielectric constant and the viscosity of the BGE are directly determined by the BGE composition and can be readily adapted. In particular, the use of organic solvent in the BGE composition can heavily affect these parameters. In contrast, the zeta potential is rather difficult to adjust in a straightforward way but is usually performed by the covalent or absorption coating of polymers to the surface of the inner wall of the capillary.

The zeta potential is usually used for colloidal dispersions and can be defined as the electric potential between the shear plane within the double layer (i.e. the plane between the compact and the diffuse layer) relative to a point in the free solution. In the specific context of CE, the zeta potential is usually defined as the potential difference between the free BGE and the compact layer of BGE attached to the inner surface of the capillary. Therefore, the zeta potential is essentially linked to the charge density at the inner surface on the capillary wall, as presented in Figure 2.1:

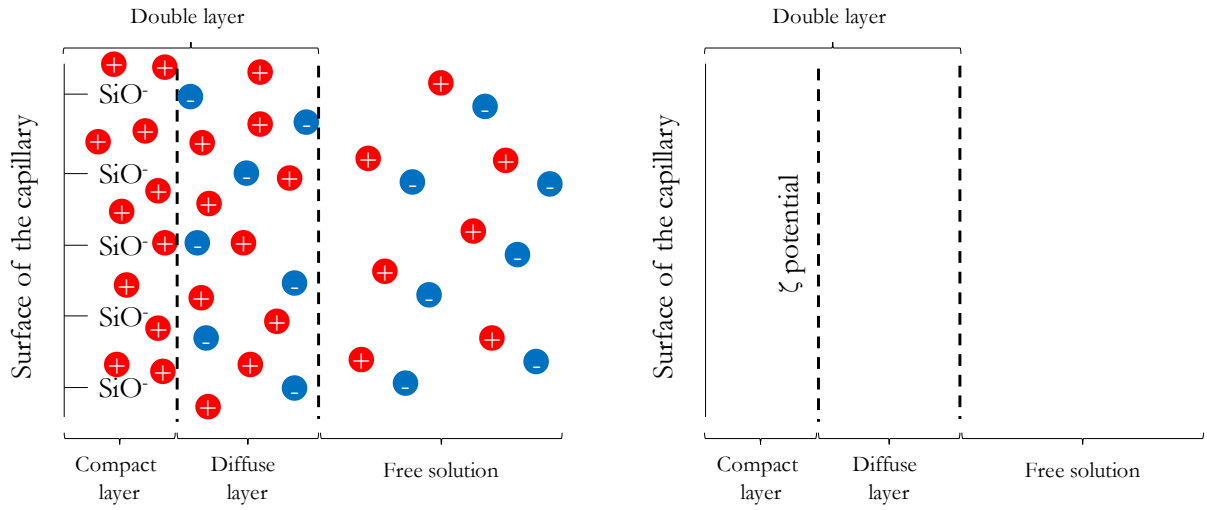


Figure 2.1 - Schematic representation of the double layer found at the contact point between the BGE and the capillary surface

In the case of bare fused silica capillaries, this charge is directly linked to the ionization of the silanol groups, with significantly greater EOF at high pH (silanols are present as SiO^-) than at low pH (silanols are present as SiOH). Figure 2.2 shows the influence of pH on the electro-osmotic mobility for different materials:

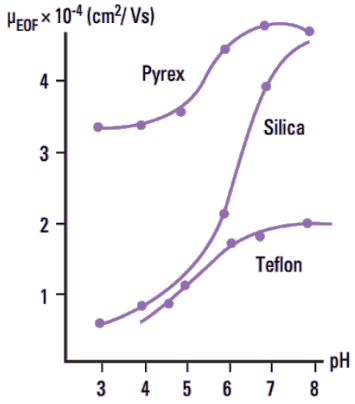


Figure 2.2 - Evolution of the mobility of EOF according to pH for Pyrex, Silica and Teflon capillaries (adapted from Lukacs and Jorgenson[28])

The zeta potential is also influenced by the charge density within the free portion of BGE in the capillary. In consequence, the ionic strength of the BGE directly impacts the EOF. Increasing the ionic strength results in a decreased zeta potential and consequently reduces the EOF.

From an analytical point of view, the EOF is often considered valuable but requires a strict control to avoid impairing the resolution of the analyte separation. As Jorgenson and Lukacs showed in early developments[29], capillary zone electrophoresis resolution can be compromised if the EOF mobility is much larger than the electrophoretic mobilities of the analytes. In such cases, a high

EOF value (generally for high pH in bare fused silica) results in very fast migration of the analytes, with a subsequent incomplete separation.

One of the major advantages of EOF is the ability to resolve all analytes in a single run, regardless of their charge, provided that their electrophoretic mobility is lower than the mobility of the EOF. As the effective mobility of an analyte (μ_{app}) is the superimposition of the intrinsic electrophoretic mobility and the EOF mobility (see next section, equation 8), the control over EOF is a parameter of importance for analytical method development. In bare fused silica capillaries, the capillary wall is globally negatively charged and the EOF is directed from the anode to the cathode, sweeping all analytes to the cathode. In that case, cations migrate at the highest velocity since their intrinsic electrophoretic mobilities are added to the EOF. Then, the neutrals are all carried out by the EOF and are consequently not separated. Finally, the anions have the lowest velocity since their electrophoretic mobilities are opposed to the EOF.

Among available options to control EOF, a change in the BGE physicochemical properties including concentration, pH or ionic strength, the temperature (affecting multiple parameters such as the viscosity, diffusion and reaction rates) and the use of an organic solvent are the main options. Another frequently applied strategy consists in modifying the capillary surface properties. To this end, the simplest way is to use capillaries of different chemical composition than the traditional bare fused silica such as Teflon[28]. However, the mainstream strategy relies on the use of dynamic or covalent coating in bare fused silica capillaries to achieve the desired EOF characteristics[30]. Figure 2.3 shows an example of the use of a cationic polymer to reverse the EOF compared to classical bare fused silica EOF:

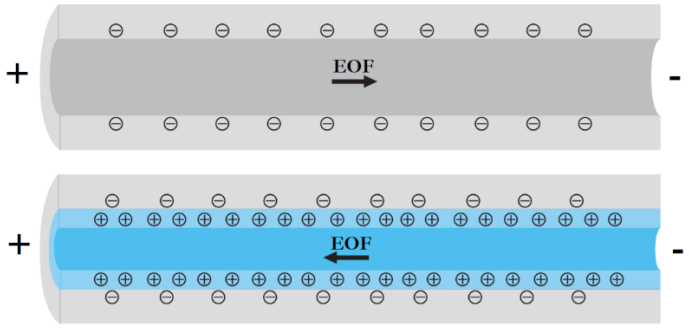


Figure 2.3 - Upper part: classical EOF direction using bare fused silica capillaries
Lower part: reversed EOF direction due to a positively charged coating

The use of coatings is of particular interest since it can reduce, neutralize or even reverse the direction of the EOF according to the nature of the coating. In addition, this strategy can also considerably decrease the adsorption of analytes (such as proteins) onto the capillary wall[31].

2.1.5.3. The apparent mobility

The combination of the electrophoretic mobility to the EOF mobility is referred as the apparent mobility (μ_{app}) as described by equation 8:

$$\mu_{app} = \mu_e + \mu_{EOF} \quad (8)$$

Consequently, the apparent mobility is the actual mobility of an analyte under precise experimental conditions. It can be experimentally derived from several experimental parameters, as explained in the following section.

2.1.6. The apparent mobility, the migration time and the electropherogram

Experimentally, the relevant parameter for the detection of an analyte during a CE experiment is its migration time. This migration time is defined as the time required for the analyte to migrate through the capillary from the inlet of the capillary to the detection point. According to the method of detection, the detection point is the outlet of the capillary (for example with CE-MS), the position of a contactless conductivity detector or a small window in the capillary for spectroscopic detection (for example with CE-UV).

The detection of the migration time can lead to the apparent mobility (μ_{app}) by equation 9:

$$\mu_{app} = \frac{l}{MT.E} = \frac{l.L}{MT.V} \quad (9)$$

Where μ_{app} is the apparent mobility of the analyte of interest ($m^2.V^{-1}.s^{-1}$)
l is the effective length of the capillary to the detector (m)
L is the total length of the capillary (m)
MT is the migration time of the analyte (s)
V is the applied voltage (V)
E is the applied electric field ($V.m^{-1}$)

As mentioned above, the apparent mobility is the combination of the electrophoretic mobility of the analyte and the EOF mobility. The measurement of the EOF mobility leads to the determination of the electrophoretic mobility: positive for cations, and negative for anions when the CE experiment is performed in normal polarity, i.e. the positive electrode is applied at the inlet of the capillary. The electrophoretic mobility can directly lead to physicochemical properties of the analyte related to its charge in solution and/or its hydrodynamic radius (equation 5). This highlights the need for the development of reliable EOF measurement methods.

Similar to the chromatogram in HPLC/UPLC, the graphical representation of the migration time through the run is referred as an "electropherogram" where the intensity of a specific analyte's signal (or the total electropherogram of the different species contained in the sample) is represented as a function of time. Depending on the detection method, the signal's units in the electropherogram are expressed as the detected units (intensity in a mass spectrum in cps of a specific m/z or absorbance unit AU for example). An example of electropherogram is given in Figure 2.4:

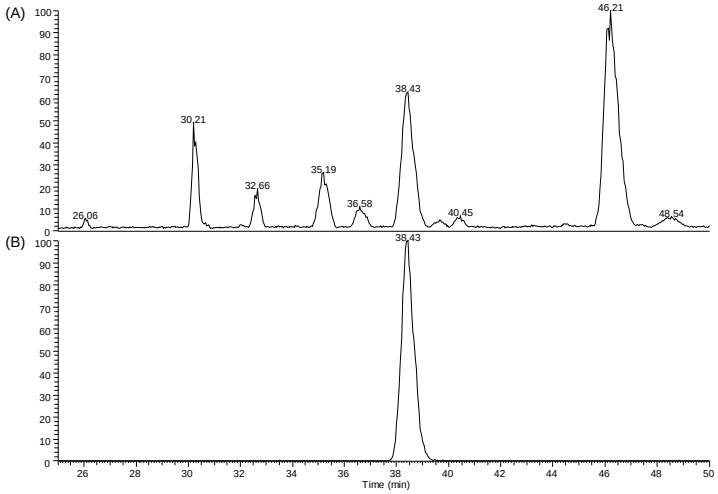


Figure 2.4 - (A) Total ion current electropherogram of different analytes included in a sample analyzed by CZE-MS (B) Extracted electropherogram of a particular analyte of a sample analyzed by CZE-MS

2.2. Mass Spectrometry

Mass spectrometry (MS) is defined as an analytical technique that ionizes chemical species contained in a sample and detects the corresponding ions based on their mass-to-charge ratios (m/z). The detection of these ions is represented by a mass spectrum where the different ion intensities (in relative intensity as presented in Figure 2.5 or in absolute counts per scan or cps) are plotted as a function of the mass-to-charge ratio, as presented in Figure 2.5.

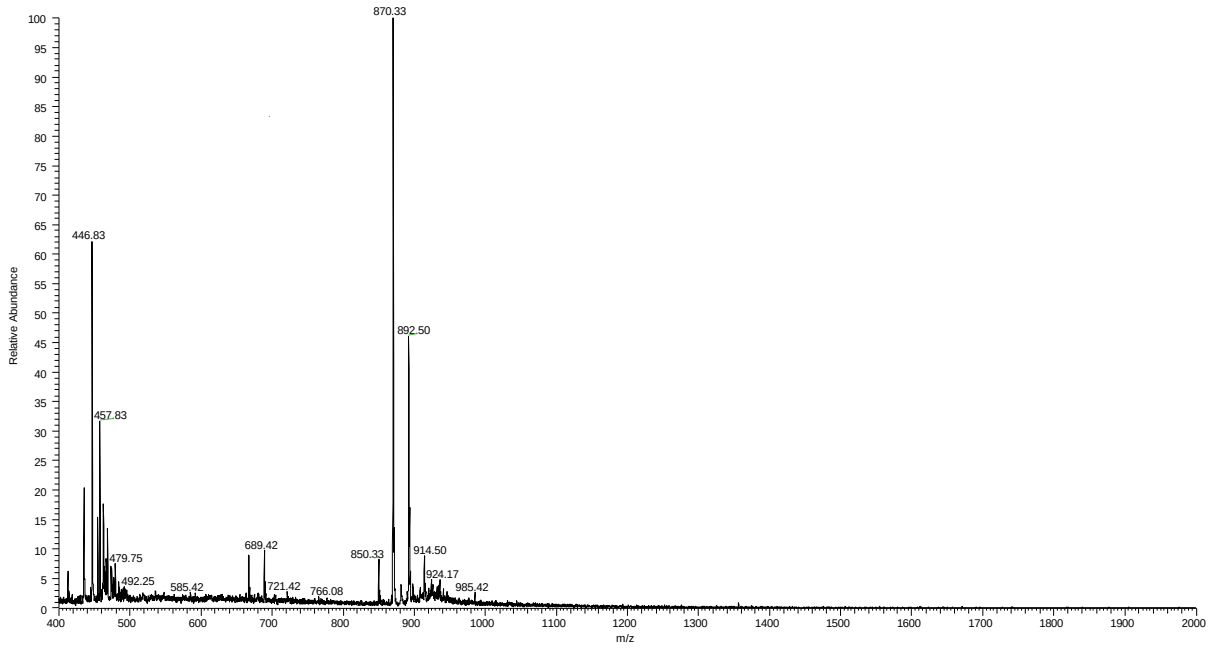


Figure 2.5 - Example of a mass spectrum with a mass range of 400 - 2,000

Because of its very high sensitivity, low limits of detection and versatility, MS has a central place in the analytical toolbox. Indeed, MS is used in a wide range of applications covering both analytical and physical chemistry. In particular, the field of MS has established as a vital bioanalytical tool in many fields such as functional genomics[32], proteomics[33], metabolomics[34], drug discovery[35], and chemical diagnostics[36]. At the earlier stage of MS development, MS was mainly applied for fundamental studies such as the study of gas phase reactions[37], the elucidation of molecular structural features[38] or the determination of thermodynamic properties of biomolecules[39].

Although MS is still developing rapidly, its introduction is not new. Indeed, J.J Thompson is often considered as the father of Mass Spectrometry since he introduced the mass spectrograph in 1897 (see Figure 2.6).

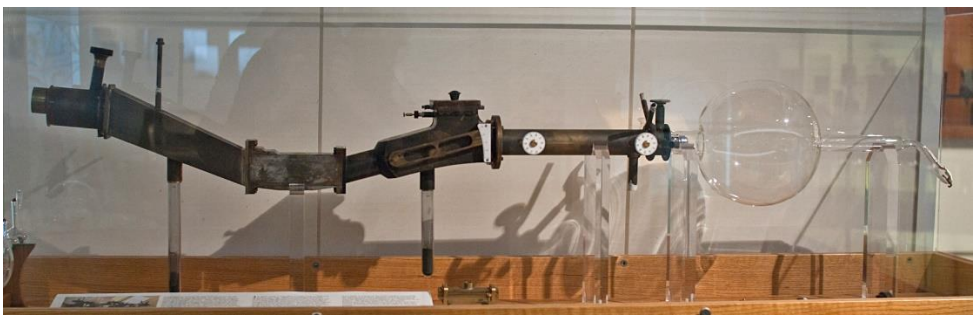


Figure 2.6 - Early design of a Mass Spectrograph by J.J. Thompson

MS has constantly progressed through various innovations and development, with the attribution of several MS-related Nobel Prizes. Some of these Nobel Prizes are listed below:

- **1922**: the Nobel Prize of Chemistry is awarded to Francis William Aston "for his discovery, by means of his mass spectrograph, of isotopes, in a large number of non-radioactive elements, and for his enunciation of the whole-number rule"
- **1989**: half of the Nobel Prize in Physics was awarded to Hans Dehmelt and Wolfgang Paul for the development of the ion trap technique in the 1950s and 1960s
- **2002**: the Nobel Prize in Chemistry was awarded to John Bennett Fenn "for the development of electrospray ionization (ESI)" and Koichi Tanaka "for the development of soft laser desorption (SLD) and their application to the ionization of biological macromolecules, especially proteins."

The general principle of Mass Spectrometry can mainly be represented by a three-step process:

1) ionization of the sample (ionization source)
2) m/z separation in time or space of the produced ions (mass analyzer)
3) ion detection and generation of the mass spectrum (detector and computer treatment)

These three different steps are detailed in the following sections.

2.2.1. Ionization of the Sample

In a typical MS procedure, the analytes contained in a solid, liquid or gaseous sample are converted into ions by the ionization source to enter into the mass spectrometer operating in vacuo. During this process, the analytes acquire one or several charge(s) by different mechanisms including electron ejection or capture, (de)protonation, formation of charged adducts or direct transfer to the gas phase of an originally charged species. Depending on the ionization mechanism and the nature of the analyte, the ionization process leads to the formation of either positively or negatively charged ions.

A wide variety of ionization procedures are described and are often classified into hard and soft ionization methods, depending on the amount of energy required for the ionization of the analyte. In general, hard ionization methods such as Electron Impact (EI) are accompanied by substantial fragmentation of the analytes upon ionization. In the extreme case of Inductively Coupled Plasma Mass Spectrometry (ICP-MS), the harsh ionization provided by the plasma atomizes and ionizes the molecular species, leading to the elemental analysis of the sample. Alternatively, soft ionization procedures such as electrospray ionization (ESI) and matrix assisted ionization by laser desorption (MALDI) allow the preservation of entire molecular species. In some cases, large non-covalent assemblies such as DNA quadruplexes[40] can even be transferred to the gas phase without dissociation.

In the present work, the electrospray ionization (ESI) was used throughout the study for its well-suited properties for the direct hyphenation between CE and MS. For this reason, this type of ionization is further detailed in the following section.

2.2.1.1. Electrospray ionization (ESI)

The electrospray ionization was first described in 1984 by Fenn[41] for the ionization of proteins. Fenn and co-workers developed the electrospray interface for mass spectrometry based on the work of Dole and coworkers[42] who generated gas-phase ions of macromolecules by spraying a solution from the tip of an capillary needle submitted to a potential difference of several kilovolts.

This ionization technique rapidly became one of the most widely used ionization technique due to its ability to perform qualitative and quantitative analyses on a wide variety of nonvolatile compounds ranging from small molecules to complex high mass biological structures. In this regard, ESI seems to be almost unlimited in the size of transferable molecules in the gas phase, as a study showed the ability of ESI to spray and ionize species up to 5 000 000 Da species with nearly 5000 charges ionized[43]. In addition, the ability to work at ambient pressure combined to the possible direct hyphenation to liquid phase separation techniques such as liquid chromatography or capillary electrophoresis made ESI the most commonly equipped ion source on commercial instruments.

Typically, in ESI-MS, a solution containing the analytes is infused at atmospheric pressure through a thin metallic needle. A high electric potential (generally a few kV) is applied to the needle while the sample flows in the micro to milliliters per minute range to be constantly sprayed at the tip of the metallic needle and form highly charged droplets (i.e. nebulization). In the positive ion mode, the application of the electric potential (in this case, the capillary is set at a positive potential) causes an accumulation of positively charged ions in the solution at the tip. When the appropriate field is reached, the liquid shapes as a "Taylor cone" with a subsequent continuous production of droplets enriched in positive ions. This phenomenon is illustrated in Figure 2.7:

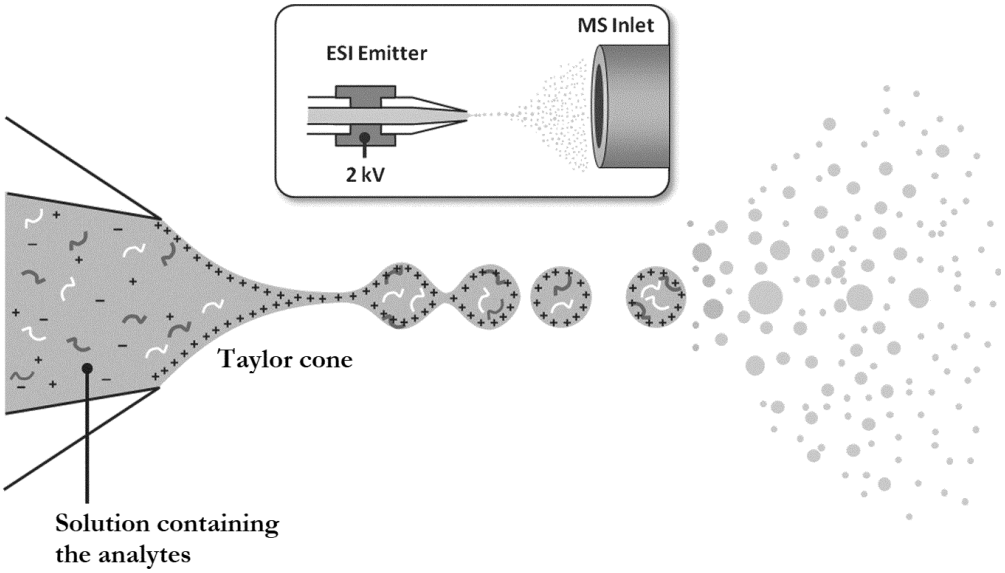


Figure 2.7 - Schematic representation of the Electrospray (ESI) process and the formation of the Taylor cone adapted from Hawkridge *et al.* [44]

Once the droplets are produced in the gas phase, solvent evaporation occurs with a subsequent decrease in the droplet's radius until the coulombic repulsions between the ions overcomes the surface tension of the droplet, threshold known as the "Rayleigh limit"[45–47]. This phenomenon leads to the coulomb fission of the droplets and subsequently, smaller droplets are produced. As evaporation carries on, the daughter droplets also undergo coulomb fissions themselves, leading to a fast reduction in both size and charge of the droplets according to Dole's model.

The transition between the first droplets to the single desolvated analyte is still a matter of debate. Classically, two mechanisms describe the production of desolvated ions in the gas phase from the droplets, as:

- Ion Evaporation Mechanism (IEM)[48–50], proposed by Iribarne and Thomson states that during the process of solvent evaporation (due to the high temperature inside the source

of the mass spectrometer) and before the Rayleigh limit is reached, solvated ions are ejected from the droplets due to the sufficiently high electric field at the surface of the droplet.

- Charged Residue Mechanism (CRM)[42,51], proposed by Dole: as explained above, a series of droplet fission events leads to a final droplet containing a single analyte molecule. The last solvent molecules evaporate until the ion is completely desolvated.

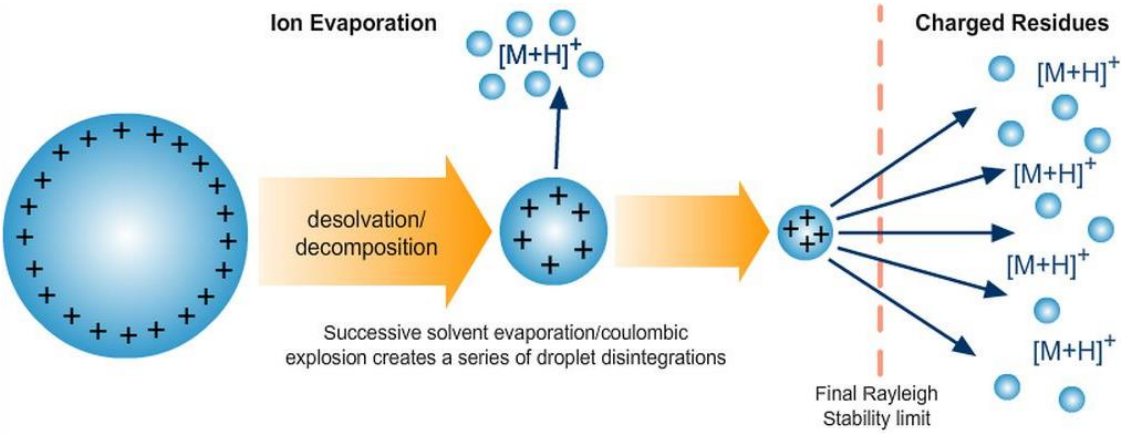


Figure 2.8 - Schematic illustration of the Ion Evaporation Mechanism and Charged Residue Mechanism leading to the production of desolvated ions during the ESI process (from www.chromacademy.com/Electrospray-Ionization-ESI-for-LC-MS)

It is generally assumed that small ions are preferably transferred into the gas phase by the ion evaporation mechanism[52–54] while larger ions tend to use the charged residue mechanism[55,56]. Other models including a combination of charged residue with field emission[57] or a chain ejection model[58] (CEM) were also introduced for disordered polymers.

Because extensive solvent evaporation is required in ESI, typical solvents for electrospray ionization are prepared by mixing aqueous solutions with volatile organic solvents such as acetonitrile or methanol. Practically, one needs to notice that, in order to close the electrical circuit within the ESI source, electrons are also involved in the process. Consequently, oxidation (in the positive ion mode) or reduction (in the negative ion mode) of species present in the Taylor Cone also takes place during the ESI process. The current linked to this phenomenon is in the order of a few µA. Typical ESI flowrates range from 1 to 100 $\mu L.min^{-1}$ but higher flowrates are sometimes required for effective hyphenation to HPLC. Most of the time, the ESI process needs to be assisted by a (heated) desolvation gas for efficient desolvation of ions.

The introduction of nanoelectrospray[59–61] (described in the following section), characterized by much lower flowrates, is considered as the most efficient form of electrospray in terms of ionization efficiency. Indeed, the very low flowrate ensures higher ionization efficiency without the need for desolvation gas assistance. Nonetheless this statement is still under debate due to the introduction of other limitations, especially in the field of "native mass spectrometry"[62].

2.2.1.2. Nanoelectrospray ionization (nanoESI)

This evolution of the electrospray technique was introduced by Mann and Wilm in 1996[60]. The term "nanoelectrospray" reflects the nanoliter per minute flowrate range and the size of the generated droplets, which is in the nanometer range instead of the classical micrometer range for ESI.

Indeed, in the conventional electrospray source, the typical flowrate is in the µL or sub-$mL.min^{-1}$ range. As a result, the initially formed droplets have an average diameter in the µm range[43,52,63]. In

contrast, the nanoelectrospray source uses a very small spraying orifice in combination with a very low flowrate. In its original form, the nanoelectrospray process directly dictates the flow and consequently, the flow operates without external pumping. The result is that the flow rate of such device is only a few dozen $nL.min^{-1}$, generating primary droplets with an average diameter lower than 200 nm, which yield gaseous ions in less fission events and desolvation.

The main advantage of nanoelectrospray is the improved ionization efficiency and ion transmission leading to an improvement in terms of the limit of detection. A recent work by El-Faramawy *et al.* showed[64] that the efficiency of the nanoelectrospray ionization followed an almost linear increase when the flowrate was lowered from 1,000 to 4 $nL.min^{-1}$, with a maximum efficiency of about 1%. In addition, the improved desolvation due to the very low flow rates alleviates the need for nebulizing gas.

Because the droplet fission processes are not critically different between ESI and nanoESI, the major difference originates from processes associated with the dispersion of the liquid into charged droplets and the size of the initial droplets. In general, the main reasons are:

- A time scale influence: as the larger droplets require at least one additional fission event to reach the necessary size for ion release to occur, the formation of ions happens later when starting with initially larger droplets
- A solvent evaporation influence: the solvent contained in the initial droplet need to sufficiently evaporate to reach the necessary charge density for fission to occur. This evaporation results in an increase in both analyte and salt in the droplet. As a result, more fissions mean higher salt concentration in the final droplets.
- A last but not least phenomenon is the use of nanoESI, especially with CE or nano-liquid chromatography has a higher sample-to-volume ratio, which affects the physicochemical properties of the liquid phase being sprayed.

As a result, nanoelectrospray generally provides a higher tolerance towards salt concentration[60,61] due to the minimization of ion suppression and matrix effects, which can be problematic using conventional ESI.

Finally, it should be stressed out that the development of nanoelectrospray is still an ongoing process, mainly due to the key role of the emitter for an efficient nanoelectrospray. Indeed, the emitter tip architecture heavily impacts the Taylor cone formation and optimal parameters for efficient nanoelectrospray need to be carefully selected.

2.2.2. *m/z* analysis

After the production of the gas-phase ions by the ion source, these ionic species are separated according to their mass-to-charge ratios (m/z) by the mass analyzer. As a consequence, multiply charged species (such as those produced by the ESI process for example) present m/z ratios that are fractions of their actual masses.

The ions are either space-resolved or time-resolved according to the hardware of the mass analyzer. There is a relatively large variety of mass analyzers, which are all based on the use of static or dynamic electric and/or magnetic fields. The characteristics of a specific mass analyzer lie in the way it uses these fields to measure the m/z ratio of ions. Naturally, each setup has advantages and drawbacks that need to be carefully considered for a specific application. Intrinsic properties including ion beam versus ion trapping modes, continuous versus pulsed analysis, low versus high kinetic energies are also important parameters that needs to be considered for a particular

application. For example, trapping systems are particularly suited for multiple stage tandem mass spectrometry (MS^n). Alternatively, high resolution mass spectrometers (such as FT-ICR and Orbitraps) provide exact masses and fine isotopic structures that are interesting for the characterization of unknown analytes[65,66]. The counterpart to the high mass resolving power is usually a lower ion transmission efficiency leading to a worse detection limit of the mass analyzer.

A great strength of MS lies in the constant development and improvement of mass analyzers. The best example of the continuous development is the introduction on the market of a new type of mass analyzer, the Orbitrap, in 2005. Additionally, the existing mass analyzers are constantly improved, leading to more reliable, better performing instruments. Another important trend in MS is the coupling between different mass analyzers into the same instrument to increase their versatility and allow tandem MS analysis to be performed. Some common hybrid instruments are the triple-quadrupole, the quadrupole time-of-flight or the ion trap–FT ion cyclotron resonance (ICR) mass spectrometer which all allow the MS analyst to obtain a mass spectrum resulting from the decomposition of an ion selected in the first analyzer.

In general, five major characteristics are used to assess the performance of a mass analyzer, which are:

- the mass range limit is the maximum m/z the mass analyzer can handle
- the dwell time and the scanning rate (or the analysis speed understood as the scan speed) is the rate at which the mass analyzer measures the m/z ratios over a particular mass range. It is expressed as the range of mass units per second or in mass units per millisecond. This property is critical when a separation method (such as CE and LC) is hyphenated to MS for an efficient sampling of the chromatographic or electrophoretic peaks. In this case, the scan speed needs to be fast enough to obtain a sufficient number of mass spectra per peak, which is expected to be equal or higher than 15 points per peak.
- the ion transmission efficiency is the fraction of ions reaching the detector compared to the total number of ions introduced inside the mass analyzer.
- the mass accuracy is the error (usually mean root square) of the m/z value measured by the mass analyzer typically provided in part per million (ppm). Practically, it is the difference between the measured experimental mass ($m_{experimental}$) and the theoretical mass ($m_{theoretical}$) divided by the theoretical mass as expressed below:

$$\text{mass accuracy} = \frac{(m_{experimental} - m_{theoretical})}{m_{theoretical}} \quad (10)$$

  The mass accuracy is essentially interlocked to the mass resolving power of the mass analyzer. High mass accuracy has important applications, especially for elemental composition for example.
- the resolving power is linked to the ability of a mass analyzer to distinguish close m/z values. Although various definitions of the resolving power exist, it is generally admitted that if Δm is the smallest mass difference for which two peaks with masses m and $m + \Delta m$ are resolved, the definition of the resolving power (R) is $R = m/\Delta m$. Alternatively, the resolving power can also be defined for an apparently isolated peak. In this case, the resolution is calculated using the peak width Δm at x % of valley of the peak height (which is generally taken to be 50 % and Δm is then referred as the "full width at half maximum" (FWHM). Low or high resolving power usually refers to a value that is equal or greater than about 10,000 (FWHM), respectively. However, there is no exact limit between these two terms. Nowadays, the terms of ultra-high resolving power and even extreme resolving

power were introduced due to the very last development of the mass analyzers such as the Orbitrap and the dynamically harmonized ion cyclotron resonance (ParaCell™).

The importance of resolution can be directly viewed in Figure 2.9, where different resolutions (provided by different mass analyzers) are directly compared on the same species:

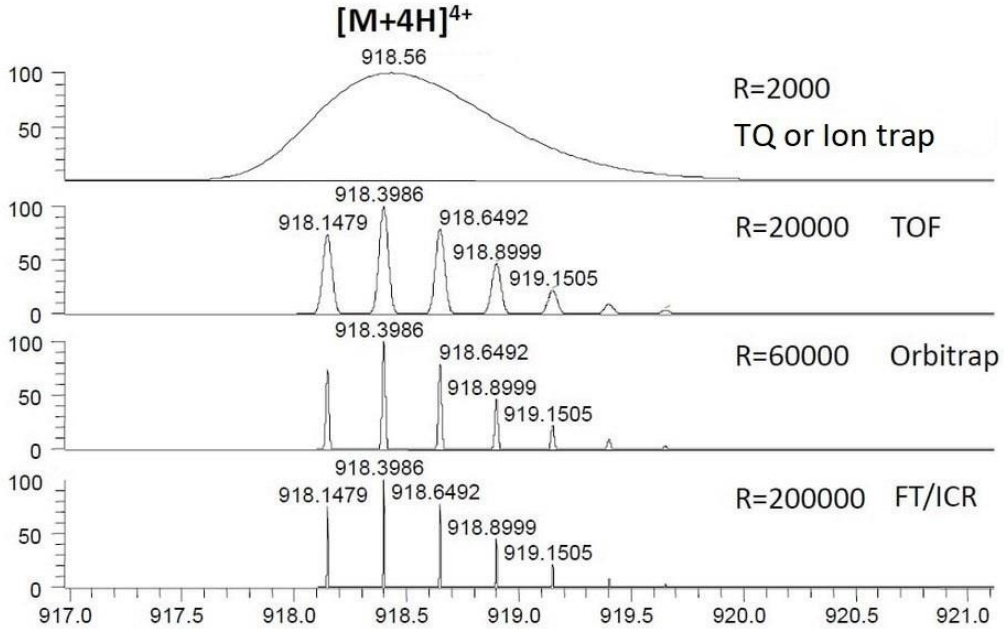


Figure 2.9 - Influence of the mass resolving power (FWHM) on the spectrum obtained for the same species on its 4+ charge state. From top to down: Triple Quadrupole (TQ) or ion trap spectrum at resolution 2,000; Time-Of-Flight (TOF) spectrum at resolution 20,000; Orbitrap spectrum at 60,000 at resolution and Fourier Transform – Ion Cyclotron Resonance (FT-ICR) spectrum at resolution 200,000 (from De Hoofman and Stroobant [67])

De Hoffmann and Stroobant reviewed the major characteristics of the different mass analyzers available in 2007, as demonstrated in Table 2.2:

Table 2.2 - Various available mass analyzers and their respective characteristic features (from De Hoofman and Stroobant [67])

	Quadrupole	Ion trap	TOF	TOF reflectron	Magnetic	FTICR	Orbitrap
Mass limit	4000 Th	6000 Th	>1000 000 Th	10 000 Th	20 000 Th	30 000 Th	50 000 Th
Resolution FWHM (*m/z* 1000)	2000	4000	5000	20 000	100 000	500 000	100 000
Accuracy	100 ppm	100 ppm	200 ppm	10 ppm	<10 ppm	<5 ppm	<5 ppm
Ion sampling	Continuous	Pulsed	Pulsed	Pulsed	Continuous	Pulsed	Pulsed
Pressure	10^{-5} Torr	10^{-3} Torr	10^{-6} Torr	10^{-6} Torr	10^{-6} Torr	10^{-10} Torr	10^{-10} Torr
Tandem mass spectrometry	Triple quadrupoles MS/MS fragments precursors neutral loss	— MS^n fragments	— —	PSD or TOF/TOF MS/MS fragments	Consecutive sectors MS/MS fragments precursors neutral loss	— MS^n fragments	— —
	Low-energy collision	Low-energy collision	—	Low- or high-energy collision	High-energy collision	Low-energy collision	—

Because this thesis mainly relied on the application of High-Resolution Mass Spectrometry (HR-MS), only the FT-ICR, Orbitrap and TOF mass analyzers were used throughout the work. For this reason, these particular mass analyzers are further detailed in the following sections.

2.2.2.1. The Fourier Transform - Ion Cyclotron Resonance (FT-ICR) Mass Analyzer

The Fourier Transform – Ion Cyclotron Resonance (FT-ICR) mass analyzer uses strong magnetic fields to confine ions within a cell referred as the "Ion Cyclotronic Resonance" cell or ICR cell. In this setup, the ion velocity is kept low in combination with a strong magnetic field to confine the ion on a circular trajectory, therefore referred as "ion cyclotron" or "Penning trap". The equations related to the ion's motion are described below. For a particular ion, a balance between a centripetal force (due to the magnetic field) and a centrifugal force (due to the angular velocity of the ion) is created in the ICR cell:

$$F_{centripetal} = q.v.B \quad (11)$$

$$F_{centrifugal} = \frac{m.v^2}{r} \quad (12)$$

Where q is the charge of the ion (C)
v is the velocity of the ion ($m.s^{-1}$)
B is the intensity of the magnetic field (T)
m is the mass of the ion (kg)
r is the radius of the circular trajectory (m)

The balance of these two forces (combination of equations 11 and 12) results in a stable trajectory of the ion:

$$q.v.B = \frac{m.v^2}{r} \text{ or } q.B = \frac{m.v}{r} \quad (13)$$

The frequency of the circular trajectory of the ion (f in Hz or s^{-1}) within the ICR cell is then given by equation 14:

$$f = \frac{v}{2\pi.r} \quad (14)$$

And finally, the angular velocity of the ion (ω in $rad.s^{-1}$) is given by equation 15:

$$\omega = 2\pi.f = \frac{v}{r} = \frac{q}{m}.B \quad (15)$$

Equation 15 shows the direct link between the angular velocity (ω) or the angular frequency (f) of a particular ion and its ratio B×(q/m). Even though the angular frequency and the angular moment are not directly dependent on the ion's velocity, the radius of its trajectory increases when its velocity increases. When the appropriate excitation energy frequency is applied within the ICR cell, the radius of its trajectory becomes larger than the cell's dimension and the ion is ejected out of the cell.

Practically, the ICR cell is typically a box (whose geometry is not necessarily cubic) of a few centimeters side which is inserted into a strong magnetic field. The magnetic field is produced by a superconductive magnet which typically yields values around several Tesla (T). For example, the instrument used in this study, the LTQ-FT Ultra Hybrid from Thermo Finnigan, provides a 7T magnetic field but magnets as powerful as 24 T were already tested by 1999[68].

In FT-ICR mass analyzers, the masses of the ions are deduced by the measurement of their respective angular frequencies. For a 3T magnetic field, the ion cyclotron frequency is 1.65 MHz at 28Da and 11.5 kHz at 4000Da, highlighting the wide range of frequencies that needs to be

measured. Once the frequencies are measured, they are converted into a mass spectrum by use of Fourier transforms or related techniques.

2.2.2.1.1. Ion Cyclotron Resonance

Once a particular ion is in the ICR cell, the irradiation of the ICR cell with an electromagnetic wave that has the same frequency than the ion results in a resonant absorption of this wave, where energy is transferred from the exciting wave to the ion. This energy transfer results in an increase of the ion's kinetic energy and thus, an increased radius of its trajectory.

The circulation of ions close to the walls of the ICR cell induces an alternative "image current" that can be measured. The principle of ICR cell is then to alternate excitation phases where ions with a specific mass are resonantly targeted by the exciting wave and detection phases where the image current of the excited ions is detected. Efficient detection requires the ions of a given mass to circulate as tight packets on their own orbits, which can be achieved by pulsed excitations.

2.2.2.1.2. Fourier Transform - Mass Spectrometry

The principle of this technique is to simultaneously excite all ions present in the ICR cell by a fast wave (approximatively 1µs) containing a large frequency range. As a result, all ions contained in the ICR cell are simultaneously excited and the increase in the radius of their trajectories bring them close to the detection walls in phase. Then, the resulting image current from the coherently excited trapped ions is detected by the detection plates and converted from the time domain to the frequency domain using Fourier transform. These frequencies are finally used to obtain the mass spectrum as depicted in Figure 2.10:

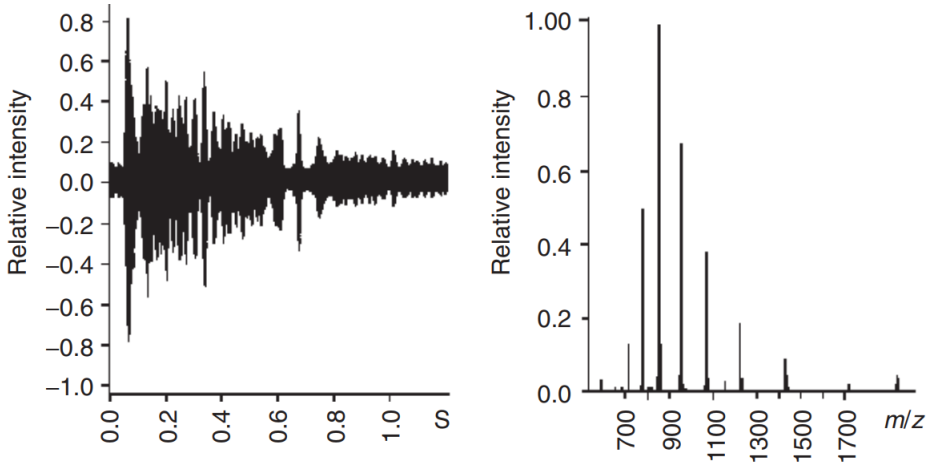


Figure 2.10 - Conversion of the frequencies (transient signal) detected in an ICR cell by Fourier transform to the corresponding mass spectrum from De Hoofman and Stroobant [67]

The excitation of all ions contained in the ICR under a magnetic field B_0 cell with a broad-band radiation with the same energy (provided by the same V_0 potential) for the same duration T_{exc} results in the coherent excitation of all ions to the same orbit. Consequently, the optimization of the excitation wave parameters allows generating the appropriate radius (close to the detection plates) for the efficient detection of the image current:

$$r = \frac{V_0 . T_{exc}}{B_0} \quad (16)$$

This equation, demonstrated by Marshall and co-workers[69], shows that the radius is indeed independent of the m/z ratio. However, the cyclotronic frequency of each ion is characteristic of its mass. Thus, broadband excitation using the same voltage at each frequency will propel all the ions onto the same radius, but at frequencies depending on their m/z ratios.

Because the technique is based on a Fourier transform, the resolution of this mass analyzer is linked to the observation time (relaxation time) of the ions within the ICR cell. In the gas phase, the relaxation time is highly dependent on the vacuum quality and on the stability of the ions. If the vacuum value is too low, collisions with residual gas molecules slow down the ions and their motion become incoherent.

To achieve high mass resolving power, a very high vacuum is required in the ICR cell (about 10^{-5} – 10^{-6}Pa), which is an important limitation to the technique. According to the mass resolving power applied for the analysis, FTMS can record one mass spectrum in a time window between one and a few dozen seconds. In extreme cases (such as resolution of 500,000 or 1,000,000 and above), the acquisition (dwell time) of one mass spectrum can last more than 1 minute. As a result, this feature is critical when hyphenation with a separation technique is considered.

The dynamic range of such instrument is generally limited due to coulombic repulsions between the ions but these instruments are sufficiently sensitive to detect about 10 ions in the ICR cell.

FT-ICR instruments play an important role in the study of gas phase reaction because they enable observation of ions over long time spans. It is also possible to selectively isolate species of interest through intense irradiation at resonance frequencies to only keep a particular m/z of interest for further gas phase investigation. As a result, the high-resolution study of gas phase reactions including slow fragmentation processes[70], equilibria between ionic species[71] and ion–molecule reactions[72] can be performed by FT-ICR instruments.

Finally, the possibility to select a particular ion of a given m/z in the ICR cell offers a great advantage for MS/MS and MS^n experiments. Because the detection is performed in a non-destructive way, fragmentation experiments can be performed on the same ions that were used for detection. As result, alternance of a phase where the m/z analysis is performed with a phase where a particular m/z is isolated and fragmented (by the use of a collision gas in another mass analyzer, electron capture or by UVPD or IR irradiation) can lead to high-resolution MS^n data with the same pool of ions from the source as the detection is not destructive.

2.2.2.2. The Orbitrap Mass Analyzer

Until recently, FT-MS was exclusively used to refer to FT-ICR mass spectrometry. The combination of the high stability and uniformity of the superconductive magnets with the high accuracy of frequency measurements turned FT-ICR as the most efficient MS-based technique in terms of mass resolving power and mass accuracy. The introduction of ESI and MALDI for the analysis of sometimes large biological molecules turned this ability into an increasing important feature. Furthermore, this trend was also accompanied by an increase in the complexity of the analyzed samples and performance of separation techniques prior to the MS analysis.

These ever-increasing demands for more efficient mass analyzers created the conditions for the appearance of a new type of mass analyzer: the Orbitrap. Similarly to the FT-ICR, the orbitrap is an electrostatic ion trap that uses the Fourier transform to obtain mass spectra. This analyzer was introduced by Makarov as patents[73] and scientific literature[74] by the end of the 90's, early 2000. Thermo commercialized the first hybrid instrument based on this mass analyzer in combination

with a linear ion trap in 2005. Note that the Time of Flight mass analyzers also considerably improved with mass resolving power increasing from 8,000 (FWHM) during the late 90's to more than 50,000 currently.

2.2.2.2.1. Description of the Orbitrap

Interestingly, this mass analyzer inherited some features from other types of mass analyzer such as the principle of image current detection from the FT-ICR, the use of ion trapping in precisely defined electrode structures from the ion trap and pulsed injection and the use of electrostatic fields from the TOF analyzers[75]. These combined features provide a powerful mass analyzer with various advantages over other designs including the unnecessity for a superconducting magnet like in FT-ICR or severe limitations on space charge characteristic of the ion traps[75].

The design of this mass analyzer is based on two electrically isolated outer electrodes shaped as half-barrels separated by a tiny gap and a central electrode shaped as a spindle. When a voltage is applied between the outer and the central electrodes, the resulting electric field is purely linear along the axis of the central electrode and the ions are confined in the trap because of the radial component of the field which strongly attracts them towards the central electrode. The injection of the ions (typically at high kinetic energy, generally a few keV) is either performed tangentially through the interstice between the two outer electrodes or off-axis through a hole in one of the outer electrode (see Figure 2.11). Once injected, the trajectories of the ions are bend by the radial electric field around the central electrode while the tangential velocity of the ions creates an opposing centrifugal force. As a result, the ions remain on a nearly circular trajectory inside the trap. Additionally, because of the particular design of the central electrode, the ions harmonically oscillate around the central electrode at a frequency depending on the m/z ratio.

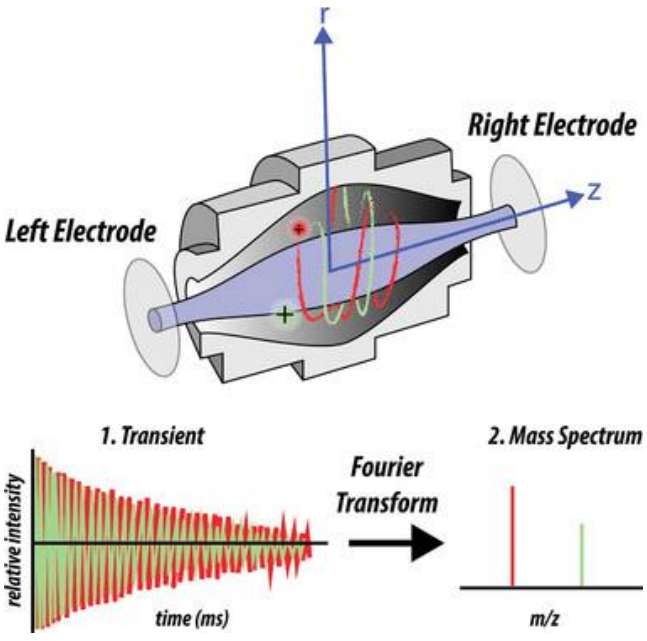


Figure 2.11 - Schematic representation of an Orbitrap mass analyzer and production of the corresponding mass spectrum from Savaryn *et al.* [76]

The image current of the axial oscillations of the ions is then measured by parts of the outer electrodes that act as detection plates. The equation connecting the detected frequencies to the m/z ratios of this mass analyzer is somehow more complex than for the FT-ICR but in essence, the radial frequency of the ions inside the Orbitrap is proportional to their m/q ratios and is

independent of the kinetic energy of the injected ions. Similar to the FT-ICR, the image current induced by the oscillations of the ions within the Orbitrap is converted into the corresponding mass spectrum by a Fourier transform.

Several improvements of the initial design were implemented to bring this analyzer to practice. First, the development of pulsed injection from an external ion accumulating device, referred as C-trap[77], was introduced in 2005 in combination of a linear ion trap to form the first commercial instrument. This technological feature was an important step for the implementation of the Orbitrap because it allows an effective decoupling of the orbitrap from the ion source or ion transmission device. Thanks to the C-trap, any selection device or fragmentation technique can be easily interfaced with the Orbitrap. A few years later, the addition of various fragmentation devices including the higher collision energy dissociation[78] (HCD) cell, electron transfer dissociation (ETD)[79] and UVPD capabilities marked additional evolutions for the Orbitrap-based instruments. By 2009, the introduction of a new type of ion guide (the so-called S-lens) improved the transfer efficiency of by 3-5fold increase for full scan spectra and almost 10-fold increase for the MS/MS mode. By 2011, the resolving power of the analyzer has been increased to 240 000 at m/z 200 for a 768ms analysis time[80]. Finally, in 2011, a combination of the Orbitrap with a quadrupole mass analyzer was launched under the name of Q Exactive series and became a popular instrument for proteomics, metabolomics and high throughput screening[81].

2.2.2.2.2. High-Resolution Orbitrap Mass Spectrometry for Quantitative Analysis

One of the most important application of MS is its use as a quantitative analytical tool, historically established by the use of (triple) quadrupole (TQ) instruments as the most widely used quantitative MS method[82]. In this tandem MS analysis, an ion of a particular mass is first selected (i.e. the precursor ion), fragmented and product ion(s) from the precursor ion(s) is/are selected for detection, which is referred as a transition. This method, commonly named "selected reaction monitoring" (SRM), is typically implemented on modern TQ instruments for quantitative MS, including for proteomic analysis[83]. For such complex samples, the required specificity is achieved through the detection and relative abundance of multiple specific diagnostic fragments of a precursor ion (multiple transitions, previously known as MRM or multiple reaction monitoring) for quantitative analysis.

Though effective, this MS strategy sometimes suffers from a lack of selectivity to reliably discriminate the targeted analytes from interferences, especially in highly complex samples. Also, experimental parameters such as the collision energies need to be optimized for each transition, which is can be tedious if numerous transitions have to be followed (in this case, the term "multiple reaction monitoring" or MRM is preferred). Finally, as each transition needs to be followed separately, the dwell time required to perform the MS analysis can compromise the efficient sampling of the electrophoretic/chromatographic peaks typically characterized by second to sub-minute peak widths.

In this context, high-resolution instruments (and noticeably the hybrid Quadrupole–Orbitrap instruments) provide a valuable additional dimension of selectivity, with the ability to distinguish ions based on high-resolution MS information. The multiplexing capability of such instruments ("parallel reaction monitoring" or PRM instead of MRM, or the SWATH technology) is also a valuable feature, as it significantly reduces the tandem MS dwell time by simultaneously scanning all product ions of a precursor ion with high resolution and high accuracy, as presented in Figure 2.12:

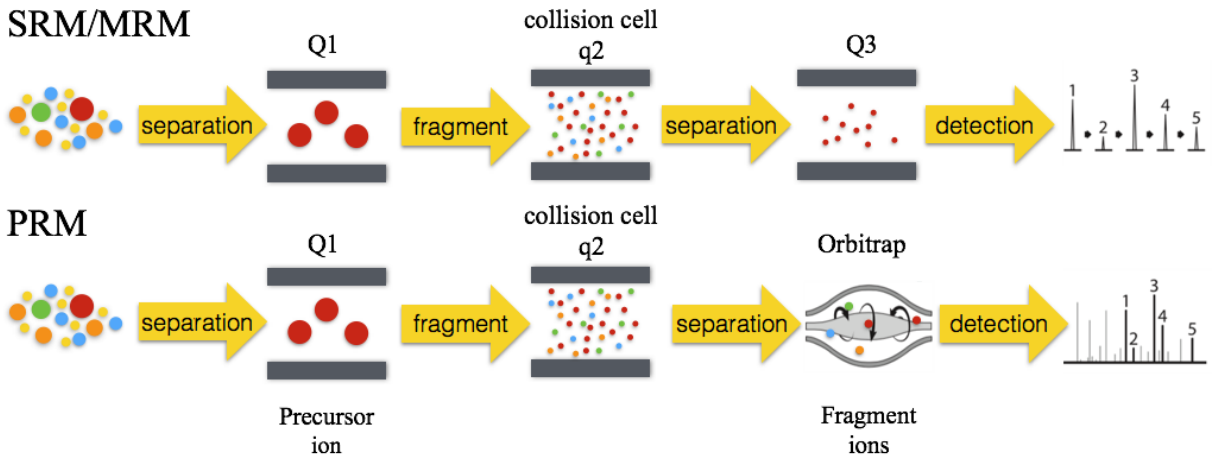


Figure 2.12 - Comparison of the triple quadrupole SRM/MRM strategy to the Quadrupole-Orbitrap PRM strategy adapted from www.creative-proteomics.com

In addition, Quadrupole–Orbitrap instruments often achieve similar analytical performance in terms of selectivity, dynamic range, sensitivity and limit of detection than the reference TQ-based SRM technique[84]. Besides, such instruments offer several advantages over the TQ-based strategy including the optional SRM optimization (therefore time saving when switching from one MS instrument to another), high-resolution MS data and easier method development. However, the number of transitions that can be followed is still limited in PRM according to the total dwell time of PRM experiments in regards of the chromatographic peak width.

For all these reasons, the MS data provided by FTMS-based instruments, including Orbitrap and the last and new FT-ICR technologies, will play an increasingly significant role in MS quantification methods for complex samples in a near future[85].

Finally, it should be noted that MS-based quantifications are generally divided into relative quantification or absolute quantification. The relative quantification is principally based on two categories of methods referred as "label-based" and "label free" methodologies. In the label-based method, species of interest (often proteins or peptides) are chemically modified with stable-isotope labels *in vitro* or *in vivo* for quantification while the label free does not require this modification step. Consequently, the label free method is generally simpler and more cost-effective as it significantly reduces the sample preparation compared to the label-based method. However, this approach requires a strict control over the experimental conditions to obtain relevant results. Finally, the absolute quantification relies on the introduction of an internal standard for the determination of the absolute amount of the targeted species. The golden standard currently relies on the isotopic dilution method, where a stable isotopically labelled internal standard (^{13}C, ^{15}N or ^{18}O for example) is used for quantification.

2.2.2.3. The Orbitrap vs FT-ICR Mass Analyzers

Because the Orbitrap was introduced several decades after the FT-ICR, the progress in several areas, including broadband phase correction[86] or increasing field strength provided by available superconductive magnets[87] make FT-ICR MS a serious concurrent to the Orbitrap MS.

Although there are major differences in the size and cost, these two mass analyzers have a series of common features.

First, the ions are trapped in ultra-high vacuum to allow (very) long free paths for the detection of the ions (sometimes up to hundreds of kilometers). Then, the mass spectra are generated by the image current of the ions and FT post-process of their frequencies yields to the mass spectra. As a result, the resolving power of both devices is directly proportional to the detection time (T_{det}) and inversely proportional to the period of main oscillations (T_{osci}).

The main difference between both analyzers however lies in the relationship between the period of main oscillations (T_{osci}) and m/z. In FT-ICR, intense magnetic fields are used and, as a result, T_{osci} is directly proportional to m/z and the resolving power (RP) is inversely proportional to m/z. However, in the Orbitrap, ions are confined by an electrostatic field in which T_{osci} and RP are respectively directly and inversely proportional to the square root of m/z. Consequently, when comparing FT-ICR and Orbitrap capabilities, there will be a critical m/z_{crit} below which the resolving power achieved for the same oscillation period T_{osci} will be higher for FT-ICR than the Orbitrap. Conversely, above the critical m/z_{crit}, the Orbitrap will show a higher resolving power than the FT-ICR at similar scan speed. Figure 2.13 shows this effect, where the critical m/z_{crit} is around 300 for the comparison of a 15T FT-ICR and a compact high field Orbitrap:

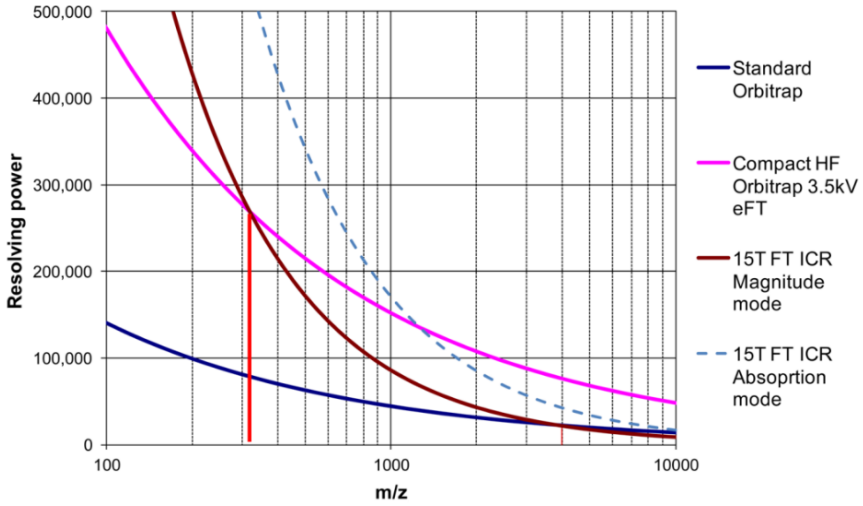


Figure 2.13 - Evolution of the resolving power as a function of m/z for various FT-ICR and Orbitrap designs from Zubarev and Makarov [75]

One should also note that the slower decrease in resolving power for Orbitrap in comparison with FT-ICR allows to get sufficient resolving power even at relatively high m/z [88].

Because only electric fields are required for the ion trapping, the size of Orbitrap instruments can be small, and commercial instrument can be manufactured as benchtop instruments. Interestingly, a compact size is also beneficial from an analytical point of view since bigger size means lower axial oscillation frequency and thus lower resolution at the same period of main oscillations. Thus, counterintuitively, larger Orbitrap instruments possess lower analytical properties than smaller Orbitrap instruments. Concerns about the charge capacity or space charge effect may rise when decreasing the size of the Orbitrap cell but the central electrode presents the great advantage of shielding the ions on different paths and thus greatly increase the effective charge capacity compared to other designs. As a result, this high charge capacity yields a high dynamic range, with reported values of 4 orders of magnitude in a single full-scan mass spectrum[89].

This situation contrasts with the FT-ICR, where a larger size of the trap increases the ion capacity by decreasing the space charge effect. Additionally, the constant strong magnetic field inside the trap ensures that the cyclotron frequency remains ICR-size independent.

One of the major advantages of FT-ICR compared to the Orbitrap is its ability to deal with low energy ions and its virtually unlimited trapping capabilities in time. This unique feature allows to subject the ions to various types of excitations (UV, IR, electron beam, fragmentation with a gas, …) while being trapped as well as gas-phase reactions study including ion-ion and ion-molecule reactions. These applications are not yet used with an Orbitrap analyzer because high-energy ions colliding with others chemical species are usually lost within a few seconds after injection.

2.2.2.4. Time-Of-Flight (TOF) Mass Analyzer

As its name implies, the Time-Of-Flight mass analyzer determines the m/z ratios of ions by measuring the time they require to drift in a free-field region (referred as the flight tube). In practice, the ions produced by the ion source are first accelerated towards the flight tube by a difference of potential providing the same kinetic energy to all ions. Because ions are characterized by their respective masses, they also present a distribution of velocities. Consequently, the electric potential energy E_{el} of an ion characterized by a mass m and total charge q $(= ze)$ accelerated by a potential V is converted into kinetic energy E_k as shown by equation 17.

$$\mathrm{E_{el}} = \mathrm{q.V} = \mathrm{z.e.V} = \frac{\mathrm{m.v^2}}{2} = \mathrm{E_k} \quad (17)$$

Once they leave the acceleration region and enter the field-free region, they are separated according to their velocities, before being detected at the other extremity of the flight tube. As a result, the m/z ratios can be determined by measuring the time that ions take to move through the field-free region as shown in equation 18.

$$\mathrm{v} = \sqrt{\frac{2.\mathrm{z.e.V}}{\mathrm{m}}} \quad (18)$$

As the velocity of the ion is constant in the field-free region, the time t required to travel through the flight tube characterized by a length L before reaching the detector is given by equation 19.

$$\mathrm{v} = \frac{\mathrm{L}}{\mathrm{t}} \quad \rightarrow \quad \mathrm{t} = \frac{\mathrm{L}}{\mathrm{v}} \quad (19)$$

As a result, the m/z ratios of ions can be determined by measuring their times of flight, as shown in equation 20.

$$\mathrm{t^2} = \frac{\mathrm{m}}{\mathrm{z}}.\left(\frac{\mathrm{L^2}}{2.\mathrm{e.V}}\right) \quad (20)$$

Several improvements of the TOF mass analyzer, including the orthogonal injection (conversion of a continuous flow of ions, such as provided by ESI, into pulsed packets of ions compatible with the TOF) or the introduction of the reflectron (an electrostatic reflector that improves mass resolving power by correcting the kinetic energy dispersion of the ions) progressively turned the initial design into a fast high resolution mass analyzer.

This mass analyzer provides several advantages. First, there is theoretically no limitation to the highest mass that can be handled making this mass analyzer the preferred choice when working with high-mass species. Another great advantage is the relative constant mass resolving power through the m/z mass range. Finally, the duty cycle is the lowest of all high-resolution mass analyzers (in the order of 100µs/scan). This feature is particularly important when dealing with separation techniques that provide very narrow peaks.

The typical drawbacks include the relatively high influence of temperature on the mass accuracy (roughly 50ppm per °C) that requires a strict control of temperature for this type of mass analyzer. Ideally, a lock-mass compound should be used when working with TOF instruments. Finally, the electronic hardware (including the detectors) working in relation with TOF mass analyzers needs to be very fast.

2.2.3. Ion Detection and Generation of a Mass Spectrum

Once the ions pass through the mass analyzer, they are detected and converted into a usable signal by the detector. The role of the detectors is to generate an electric current from the ion current according to their relative abundance. Several types of detectors are currently used. Some detectors such as Faraday cups directly measures the charge current that is produced when an ion hits its surface and is neutralized. Other designs (electron multipliers or electro-optical ion detectors) are based on the kinetic energy transfer due to the collisions of the incident ions on a surface that generates secondary electrons. These secondary electrons are then further amplified to produce a measurable current. Because of the small number of ions leaving the mass analyzer for detection, substantial amplification of the signal is usually required.

In this context, the two types of mass analyzers used in this thesis (i.e. the FT-ICR and the Orbitrap) present a characteristically different detection process than the other mass analyzers. Indeed, the detection of the ions is performed within the mass analyzer itself, where the image current produced by the trajectories of the ions and subsequent determination of their frequencies are used for detection.

Ideally, a detector should independently detect all ions with the same efficiency to avoid intensity bias in the mass spectra. However, in practice, the efficiency to detect ions generally decreases when the mass of the ion increases. The result is that limitations for high-mass ion detection occur. The recent developments in ionization sources and mass analyzers for the detection of very high mass compounds require to overcome these limitations. As a result, new types of detectors based on different physical principles such as the cryogenic detector[90] are under development.

2.3. Capillary Electrophoresis coupled to Mass Spectrometry (CE-MS)

During the last decades, Capillary Electrophoresis (CE) has turned into an increasingly powerful tool in analytical chemistry for the separation and identification of a large variety of compounds, including biomolecules such as amino acids[91], carbohydrates[92], nucleic acids[93], and proteins[94]. CE combines high separation efficiency with rapid analyses and presents a low operational cost due to the negligible volumes of consumed background electrolyte ("buffer") in comparison to those used in chromatography. Besides, the complementarity of CE in terms of available separation mechanisms with already existing separation methods (such as LC) turns it into a particularly versatile platform (see

Table 2.*1* for details). Finally, the minute amount of sample (usually in the tens to hundreds of nL range) injected per analysis is particularly suited for applications where samples are available in very limited amounts such as clinical or biomedical applications[95,96].

However, the small amounts of sample injected and analyzed by CE also require a very sensitive detection method. Generally, CE is hyphenated to UV-Vis, Laser-Induced Fluorescence (LIF) or contactless conductivity detectors. Though effective, these detectors often struggle with complex samples due to the slight variation of migration times that frequently occurs in CE. In addition, the relative lack of specificity and selectivity combined to the poor structural information[97,98] about the analyte impairs the confident identification capabilities of such analytical approaches.

Consequently, Mass Spectrometry (MS) has been increasingly hyphenated to CE due to its specificity, sensitivity, excellent detection limits, and outstanding identification performances mainly due to tandem MS with various activation techniques or ion mobility coupling. Indeed, the mass spectrometer can be viewed as a second separation dimension as it characterizes the mass-to-charge ratios of the analytes in the gas phase (m/z). As a result, the coupling between CE and MS is beneficial on both sides. On one hand, MS provides a highly sensitive and selective detector without impairing the separation efficiency provided by CE (if properly interfaced). On the other hand, a highly efficient separation prior to the MS analysis allows the minimization of ion suppression effects and facilitates the handling of complex samples where large dynamic range of concentration among analytes frequently happen. For these reasons, CE-MS is a particularly powerful analytical tool since the lack of selectivity encountered with classical detection methods or slight variations in migration time (or even co-migrating) may now be identified by different mass-to-charge (m/z) ratios. As a result, CE-MS has progressively emerged as a key tool in bioanalysis[99–105], especially in the fields of –omics. For example, a recent CE-MS proteomics study by Dovichi's group proved to be sufficiently sensitive for the detection of angiotensin down to the zemptomole level[106]. Additionally, the association of tandem mass spectrometry (MS/MS or MS^n) with computational tools provides unequivocal identification for the characterization and quantification of various biomolecules such as proteins[107].

From a practical point of view, the use of MS as the detection method of CE requires the hyphenation of a sufficiently fast MS analyzer (i.e. a mass analyzer where the scan speed is sufficiently fast) because CE can separate extremely narrow peaks, similar to those obtained using UPLC. As a result, TOF-MS is often implemented since it acquires MS spectra at very high speed (the duration between single TOF ion pulse and ion detection have a typical duration of one hundred of microseconds). However, when targeted quantitation analysis is required, the hyphenation with triple quadrupole (TQ) instrument is generally the method of choice, despite their lower scan speed[104]. Ion traps are also a valuable cost-effective association with the dual advantage of presenting a usually faster scan speed than TQ instrument and the ability to perform

MS^n experiments. Finally, approaches relying on high resolution MS such as FT-ICR and Orbitrap mass analyzers offers an unequalled mass accuracy improving the confidence of the detection/quantification in terms of ion identification.

The hyphenation of Capillary Electrophoresis to Mass Spectrometry is most often performed by the use of an ESI source[108–114]. ESI is the ionization method of choice primarily because the ionized or polar compounds that are classically well separated by CE can be efficiently transferred into gas-phase ions. Because most mass spectrometers present limitations regarding mass range, ESI is also convenient for the determination of high molecular mass molecules due to the production of multiply charged states.

Alternative ionization techniques were also investigated for the coupling of CE to MS. For example, CE can be coupled to inductively coupled plasma-mass spectrometry (ICP-MS) for elemental analysis, including the speciation of metals in biological samples[115,116]. Atmospheric Pressure Chemical Ionization (APCI) and, more recently, Atmospheric Pressure Photo Ionization (APPI) might also be coupled to CE[117,118]. In this case, the different ionization selectivity and higher tolerance than ESI towards the use of nonvolatile BGE allows the transposition of well-established CE method to CE-MS. This approach also eases the method development due to the possible use of BGE (including the non-MS friendly phosphate buffers[119]) that are usually excluded due to their incompatibility with ESI-MS. Although promising, these approaches are currently not widely implemented and online CE-MS systems are almost exclusively based on CE-ESI-MS at the moment.

Naturally, CE-ESI-MS hyphenation needs several requirements to be considered:

- compatibility of the BGE composition with ESI-MS detection
- possible interference of the CE current on the ESI process
- the electric CE circuit needs to be closed at or close to the ESI sprayer
- the typical (very) low and BGE-dependent CE effluents need to be properly handled to avoid loss of the separation efficiency provided by CE

In the literature, a large variety of background electrolytes (BGEs) can be found in classical CE but only a few are sufficiently volatile for ESI. Consequently, compatible aqueous CE-ESI-MS BGEs are generally restricted to formate (or formic acid), acetate (or acetic acid), carbonate and ammonium ions. However, the introduction of nonaqueous solvents, referred as NACE (Non-Aqueous Capillary Electrophoresis) brings new possibilities for CE-ESI-MS separations. The substitution of water by organic solvents is convenient for ESI-MS detection due to the high volatility of solvents and low ESI currents. Moreover, the use of organic solvents reduces the occurrence of undesired side electrochemical reactions at the ESI tip, and therefore stabilizes ESI current and decrease background noise in the MS spectrum. Of course, when switching from water to organic solvents, one must consider the subsequent modifications on various separation parameters such as the pK_a values of the analytes, permittivity, viscosity, zeta potential or conductivity of the BGE and the resulting modifications of the separation selectivity and/or efficiency. For example, Bonvin *et al.* compared aqueous and nonaqueous CE-MS for the analysis of several pharmaceutical compounds in urine samples and found a 20- to 100-fold improvement in sensitivity for NACE compared to traditional aqueous CE separation[120].

A critical point of the CE-ESI-MS coupling is to prevent the CE current from disturbing the ESI process. From an analytical point of view, the application of high CE currents by the means of a high electric field and a concentrated BGE is favored to perform fast and efficient separations.

Consequently, CE voltages (generally in the order of tens of kV) are typically one order of magnitude higher than those used in ESI (generally in the few kV) and an efficient separation of the electrical circuit of CE from the circuit of the ESI source is required. The most convenient and robust way to achieve this separation is to ground the ESI needle and apply the ESI voltage on the MS inlet to create an undisturbed electrical field (i.e. efficiently separated from the CE current) for the ionization in the source[121]. Unfortunately, this MS source hardware configuration is rather unusual among the manufacturers of ESI-MS instruments[122]. A possible arrangement is to connect the CE and ESI power supply using a virtual ground by the use of a dividing bridge resistor. In any case, the efficient separation of the ESI and CE currents relies on the architecture of the interface and is a critical parameter for the robust and reliable hyphenation of both techniques.

Finally, the specific parameters of CE in terms of flowrate (very low and BGE-dependent flowrates exiting the CE capillary) must be efficiently handled to avoid impairing the high separation efficiency of CE. In this context, the interface for coupling CE to ESI-MS has also a paramount importance on the overall analytical performances. The next section reviews the different approaches for the hyphenation of CE-MS and highlights the advantages and drawbacks of each setup.

2.3.1. CE-MS Interfaces

The development of an efficient interface for the hyphenation of CE to ESI-MS has long been difficult to perform in a robust and convenient way. Continuous development of CE-MS interfaces has led to the introduction of various setups, which can be classified into two major configurations: interface with or without the presence of a make-up liquid[123] (often referred as "sheath liquid"). The presence of sheath liquid has a significant impact on the type of ESI regime that will be employed. Indeed, interfaces with sheath liquid are often characterized by "high" flowrates, typically in the microliter range (between 1 and 1000µL.min^{-1}). These flowrates are compatible with a classical electrospray source and generally require the assistance of a nebulizing gas for an efficient desolvation. Interfaces without sheath liquid are characterized by much lower flowrates, generally in the nanoliter range. In this case, a nanospray interface can be used and the detrimental use of a nebulizing gas can be avoided. In general, sheath liquid interfaces have been reported to improve the overall stability of the CE-MS coupling and to be less restrictive in the BGE compatibility with ESI-MS[124] while sheathless interfaces generally provide higher sensitivity.

The flowrate used in the interface heavily affects the overall analytical sensitivity of the setup. In a recent example demonstrated by Busnel *et al.*, a decrease of the flowrate from 350nL.min^{-1} to about 10nL.min^{-1} increased the sensitivity for angiostatin detection by a factor 20[125], as demonstrated in Figure 2.14 :

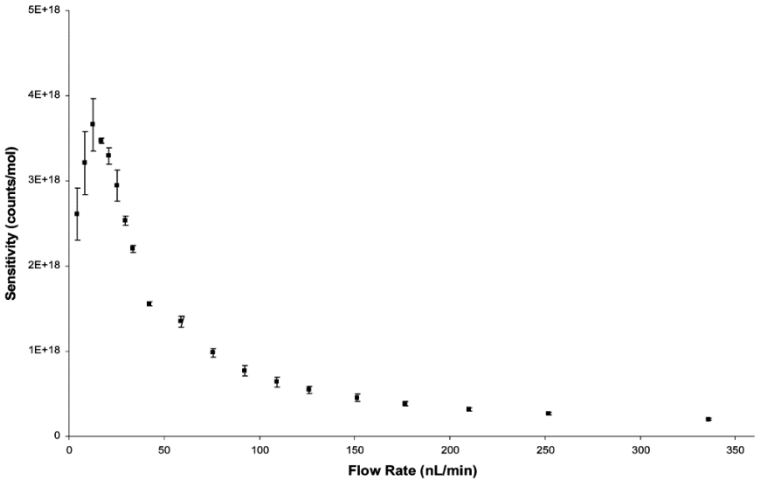


Figure 2.14 - Evolution of the peak intensity of Angiotensin I as a function of the CE flow rate (from Busnel *et al.*[125])

In the following sections, the various available interfaces are reviewed and critically compared.

2.3.1.1. Electrospray interfaces

The analytical performance of the electrospray interface is strongly correlated to the CE capillary internal diameter and the sheath liquid flowrate. Indeed, there is an optimal flowrate of sheath liquid for each internal diameter to provide the maximum ionization efficiency. In this configuration, the flowrate enables the formation of a Taylor cone that ideally matches the tip emitter. Working below the optimal flowrate cause the Taylor Cone to become unstable. Above the optimal flow rate, the enlarged Taylor cone emits larger droplets, also leading to a decreased ionization efficiency. Further increasing the flowrate leads to the "overflow regime" where the Taylor cone is also disturbed and do not optimally cover the CE capillary. In most cases, electrospray interfaces work with rather large i.d. emitters to handle the relatively high sheath liquid flowrates and a stable spray can be obtained (in general with the assistance of a gas), with a relative robustness to variations in flow rate, sheath liquid composition and sprayer position[126,127].

Among the interfaces to use a sheath liquid, the two major versions are the co-axial sheath-flow interface and the liquid junction interface. Each configuration is detailed in the following sections.

2.3.1.1.1. The co-axial sheath-flow interface

This type of interface was introduced in 1988 by the pioneering work of Smith *et al.*[128]. In the coaxial sheath liquid interface, a sheath liquid is delivered through a concentric tube that surrounds the CE capillary at relatively high flowrates to ensure the closure of the CE circuit and to generate the spray. In addition, a nebulization gas is also provided at high flowrate by a second surrounding tube to assist the nebulization. These two concentric tubes are then generally positioned in an orthogonal configuration toward the MS entrance. In this setup, the inner metallic tube containing the outlet of the CE capillary (generally stainless steel or platinum) acts as an electrode, as presented in Figure 2.15 :

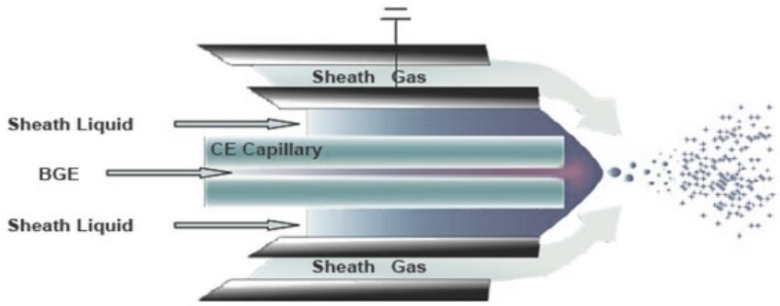


Figure 2.15 - Schematic representation of the co-axial sheath-flow interface (from Smith *et al.* [128])

In this type of interface, the nebulizing gas is one of the key parameters. The application of high gas pressures stabilizes the ESI current and provides an efficient desolvation. Unfortunately, the sheath gas leads to a loss of CE efficiency due to the hydrodynamic flow profiles (suction effect) that are generated inside the CE capillary[129]. Although detrimental for the analytical performances of this interface, this effect can be mitigated by decreasing the capillary internal diameter or increasing the length of the capillary. The composition of the sheath liquid is also of paramount importance. In most applications, a mixture of an organic solvent such as methanol or isopropanol and water provides a good compromise in terms of conductivity and volatility to achieve an optimal and stable ESI process[130]. Generally, an additional drying gas is used in the MS to further assist in the desolvation and evaporation of the ions[131].

This interface presents several advantages. First, CE and ESI can be handled almost independently, which allow both aspects to be optimized separately. Then, the constant supply of sheath liquid leads to a robust spray generation, which allows to work with a wide range of BGEs and separation methods characterized by a low (or even reverse) EOF[132].

However, this interface suffers two major disadvantages: the inevitable dilution of the analytes by the sheath liquid during the spray formation which decreases the overall sensitivity; and the suction effect inside the CE capillary. Indeed, the use of the sheath liquid and the nebulizing gas at relatively high pressures for the efficient desolvation of the ions causes suction effects in the CE capillary, with subsequent peak broadening[133,134] and a loss of separation efficiency due to a decrease in migration time[134].

Despite these problems, the sheath liquid interface can provide low detection limits (typically in the low fmol-range for peptides[122]) by performing an efficient optimization of several parameters including sheath-liquid composition and flowrate, position of the capillary tip, and various gas related parameters (dry gas temperature, nebulizing gas flow and pressure). Consequently, the co-axial sheath-flow has been successfully applied to many fields including peptidomics[135], environmental pollutants and food contamination analysis[136,137], glycomics[138], biomarker discovery and clinical diagnosis[96,139] and metabolomic studies[140,141].

2.3.1.1.2. The liquid junction interface

Another early design of interface was introduced by Henion and coworkers almost simultaneously to the co-axial sheath-flow interface : the liquid-junction interface[142,143]. In this setup, the end of the CE capillary is inserted in a liquid reservoir that is supplied with a sheath liquid to close the electric CE circuit and generate the spray. The reservoir also contains an electrode at the junction of the two capillaries for both electric circuits.

The general configuration is presented in Figure 2.16. The narrow space between the CE capillary and the ESI emitter (generally of 20–200μm) is filled with a make-up liquid that ensures the closure of both CE and ESI electrical circuits. During the process, the liquid and analytes exiting the CE capillary enters the ESI emitter and are directly sprayed, sometimes in the absence of nebulizing gas. This interface is however classified in the electrospray regime because of the relatively high flow rate that is typically used for this kind of interface.

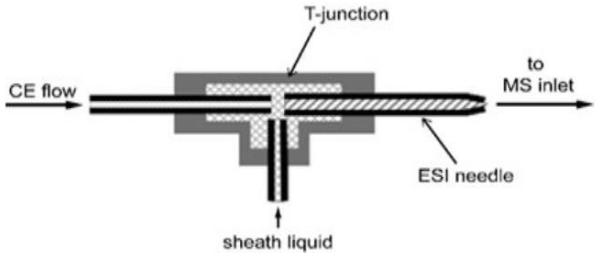


Figure 2.16 - Schematic representation of the liquid junction interface (from Maxwell *et al.* [144])

Similarly to the co-axial sheath-flow interface, the main advantage of this interface is that the CE and ESI processes can be operated independently due to the partial decoupling (both electrical and physical) of the CE system from the ESI emitter. Consequently, both ESI and CE parameters can be optimized separately.

However, this setup presents various (sometimes severe) drawbacks. First, the production of the junction and proper alignment between the CE capillary and the ESI emitter is difficult to achieve, mainly due to the difficulty to correctly position both parts in a reproducible way[145,146]. Also, electrochemical reactions occurring at the electrode (with subsequent formation of bubbles due to electrolysis) can eventually lead to the interruption of the electric field in the CE system. In some cases, an electroosmotic flow directed from the ESI emitter to the CE capillary takes place, requiring the application of a counter-pressure (typically 20–40 mbar) to the inlet capillary to counterbalance this effect. Finally, the dead volume at the junction between the capillaries causes peak broadening and, thus, loss of separation efficiency due to flow disturbance[147,148]. For all these reasons, the co-axial sheath-flow configuration has remained the preferred design in the electrospray interfaces.

2.3.1.2. Nanoelectrospray interfaces

As previously explained, when the flow rate generating the electrospray is decreased, a new spray regime referred as "nanoelectrospray"[60] occurs. Depending on the emitter internal diameter, it is generally admitted that the nanoelectrospray regime occurs between 1 and 1000nL.min^{-1}. The advantage of the nanoelectrospray resides in the production of smaller droplets compared to the electrospray. As a result, nanoelectrospray provides higher ionization efficiency, allows poorly volatile aqueous solutions to be sprayed and tolerates higher salt and/or contaminants concentrations (less suppression effects) than classical electrospray[61,149]. Also, in the nanospray interfaces, the typical emitter internal diameter is smaller than for classical electrospray and consequently, the delivered flow rates are significantly decreased. For an efficient Taylor cone formation, the distance between the tip and the MS entrance is also generally decreased, which enhances the ion transmission efficiency.

The ultimate goal of CE-MS interface is to perform the hyphenation without the help of any sheath liquid, mainly to avoid the subsequent dilution effect and taking advantage of the beneficial nanoelectrospray regime. As a result, many setups were introduced to create interfaces working in

the absence of sheath liquid, therefore referred as "sheathless" interfaces (i.e. without the assistance of a sheath liquid).

2.3.1.2.1. Preliminary approaches of the sheathless interface

Since the first attempts by Olivares and coworkers[150], the sheathless interfaces have turned into an increasingly attractive method for the coupling between CE and ESI-MS. As its name implies, a sheathless interface performs the hyphenation between CE and ESI-MS in the absence of sheath liquid. As a result, the ESI process takes place on the sole basis of the CE effluent, which must be optimized to simultaneously achieve suitable separation and sufficient ionization efficiencies. Compared to the electrospray interfaces (co-axial sheath-flow or liquid-junction), no dilution or dead volume effects occur and consequently, the CE separation efficiency is not altered. Furthermore, the flow rate in this type of interface (typically a few $nL.min^{-1}$) is directly imposed by the electrophoretic and electrospray processes, which allows the sheathless interface to operate in the beneficial nanoelectrospray regime. Consequently, this type of interface often provides better sensitivity compared to the sheath liquid interfaces, mainly because the analytes exiting the CE separation are not diluted[151].

In this approach, the major difficultly is the proper electrical closure of the CE and ESI circuits. To this end, the extremity of the CE capillary must be rendered conductive so that, when the solution exiting the CE capillary makes contact with this conductive part (which needs to be close to the ESI-MS source), the CE circuit is closed, and the spray is formed. As summarized in Figure 2.17, a relatively large variety of strategies have been developed to achieve this requirement, with various degrees of success.

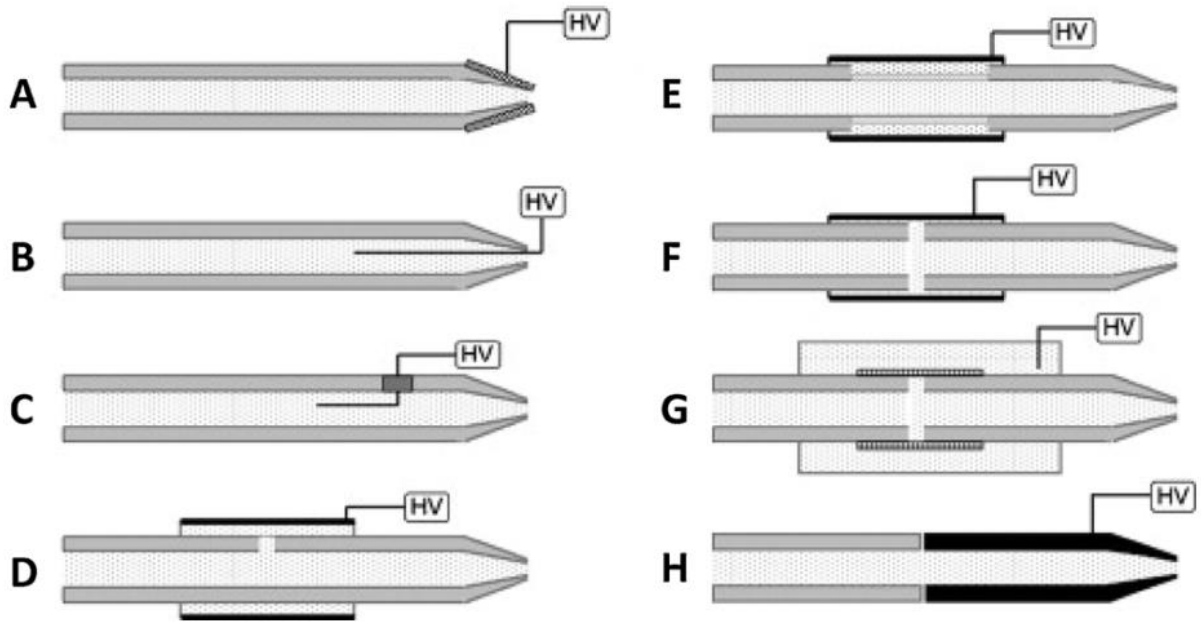


Figure 2.17 - Different strategies for the hyphenation of CE-ESI-MS by a sheathless interface : (A) conductive coating at the emitter tip, (B) wire inserted by the outlet of the emitter, (C) wire inserted through a drilled aperture in the emitter, (D) split-flow interface with a metal sheath, (E) porous capillary with a metal sleeve, (F) junction with metal sleeve, (G) microdialysis junction and (H) junction with a conductive emitter (from Bonvin *et al.*[152])

First, the mainstream strategy to produce a conductive portion at the end of the capillary relied on the use of a conductive coating on the capillary outlet with a metal (gold[153,154], silver[154,155], copper[156] or nickel[157]) or a conductive material such as or graphite[158–160] or polyaniline[161,162]. The major issue

with this approach is the rapid degradation of the conductive portion, leading to (very) limited lifetime due to the rapid degradation of this conductive portion under the necessary operational conditions for the spray formation. The insertion of a conductive wire into the outlet of the CE[163] capillary or by a tiny drilled aperture[164] was also considered but the frequent electrolytic reactions occurring close to the CE outlet generate bubbles, which prevents a robust operation due to unstable CE current. Moreover, the wire insertion into the CE capillary creates turbulence and the drilled aperture procedure is poorly reproducible.

Another approach, derived from the liquid junction interface, consists of the direct coupling of the CE capillary to a spray tip in the absence of make-up liquid. Various junction types to a conductive wire have been proposed including a metal sleeve[165,166] (Figure 2.17, part F) or a microdialysis tubing[167] (Figure 2.17, part G). The even more direct coupling of the CE capillary to a conductive spray emitter has also been introduced[168,169] (Figure 2.17, part H). The main advantage of this set-up is the easy replacement of the sprayer when needed, increasing the flexibility of this interface[169]. However, the production of a reliable junction is still difficult and time consuming. As for the liquid junction interface, peak broadening often occurs at the critical zone of junction due to misalignment and frequent incompatibility of the diameters of the two parts.

A few years ago, Moini *et al.* introduced "the split-flow interface" where the liquid exiting the CE capillary is split through a drilled aperture near the CE outlet to make contact with a coaxial metal sleeve[170] (Figure 2.17, part D). Similar to the introduction of a wire in the CE capillary, the major difficulty remains to drill an aperture in the CE capillary in a reproducible way to obtain the appropriate split ratio. Also, the aperture is difficult to produce on small i.d. capillaries (typically <30µm).

Continuous development on sheathless interfaces led to the introduction of a breaking new procedure, referred as "the etching process", to circumvent the problem of repeatability in the aperture production in the CE capillary. In this approach, a small portion of the CE capillary is etched with hydrofluoric acid to create pores in the capillary wall, which allow ion diffusion. To close the circuit, this etched portion of the capillary is often immersed in a buffer reservoir[171] or inserted in a metal sleeve containing a thin film of liquid[172]. The main disadvantage of this approach is that the exposed portion becomes fragile after the etching procedure and needs to be protected to avoid rupture. However, the very promising results of this interface paved the way for the introduction the "porous sheathless nanospray interface"[173] which is discussed in the next section.

In most of the presented configurations above, a gain in sensitivity is generally observed compared to the sheath liquid interfaces. However, the difficulty to manufacture these (sometimes complex) setups in a reproducible manner combined to the frequent instability of the spray due to the (very) low flow rates confined these approaches to "proof of concept" applications.

2.3.1.2.2. The porous sheathless nanospray interface

A major revolution in the sheathless area was developed by Moini and coworkers with the introduction of a sheathless nanospray interface based on the porous junction described above[173]. In this setup, the end of the CE capillary is etched with hydrofluoric acid to create a porous section at the very end of the separation capillary. The capillary is then inserted into a grounded metal needle until a small portion of the porous tip protrudes (~5–8 mm). The needle is then filled with a conductive liquid that diffuses through the pores and provide the electrical closure of the CE circuit. A great advantage of this setup is the easy and reproducible production of the conductive part compared to previously described setups. Also, this interface almost does not suffer from

electrolytic reactions at the end capillary. Based on this concept, Beckman Sciex Separations developed and commercialized an interface referred as "CESI (based on the combination of CE and ESI) – OptiMS". Figure 2.18 provides a schematic representation of this interface:

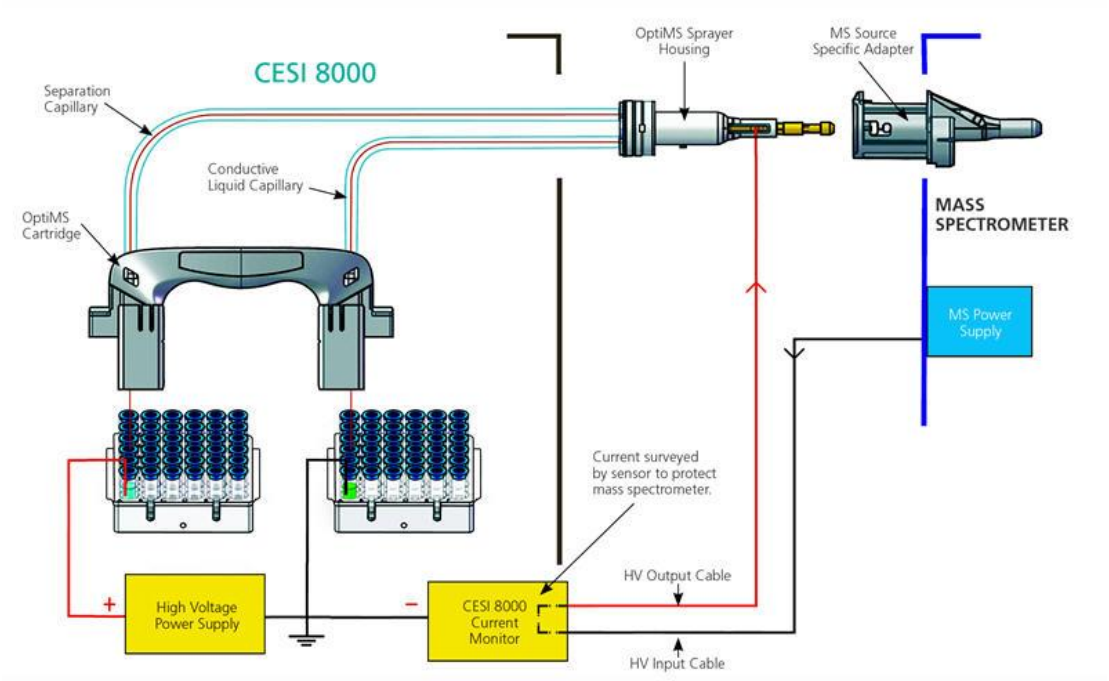


Figure 2.18 - Schematic representation of the "CESI-OptiMS "setup by Beckman Sciex Separations from sciex.com

This approach yields very promising results in various fields, including intact protein analysis[174], proteomics analysis and small pharmaceutical compounds[175]. For example, the gain of sensitivity of this interface compared to the sheath flow interface was demonstrated in a metabolomic study in human urine[176] where the sheath liquid configuration only detected 300 molecular features while the porous sheathless nanospray interface approximately detected 900 with good repeatability. This impressive increase of detected molecules could be attributed to the 8- to 30-fold better sensitivity compared to the classical sheath liquid interface. Additionally, further CE improvement (using isotachophoresis) with this setup increased the number of detected features to 1300.

Currently, the major limitation of this interface is the range of usable background electrolytes. As no sheath liquid is present in such interface, the BGE must provide efficient separation of the target analytes and simultaneously provide efficient and stable nanoESI spray.

2.3.1.3. Low sheath-flow interfaces

Recently, low sheath-flow interfaces (or low-dilution interfaces) were introduced to reduce the dilution and suction effects frequently occurring in the classical co-axial sheath-flow interface while allowing a higher flexibility regarding BGE composition. In this setup, the interface typically operates with a sheath liquid at low flow rates (generally lower than 1 $\mu L.min^{-1}$). This type of interface generally relies on the connection of the CE capillary to another capillary delivering the sheath liquid generally by the use of a stainless-steel tee union[177,178].

Another approach of the low sheath-flow interface consists in the "junction-at-the-tip" where the CE capillary is inserted into a needle providing the sheath liquid. In this configuration, the liquid exiting the CE capillary is mixed in the small volume between the end of the capillary and the needle tip. The low volume of sheath liquid between the CE capillary and the needle forms acts as

an outlet vial and the needle forms the terminal electrode to ensure a stable CE current independent of the flow[179–181].

The main advantage of this interface is to maintain a stable spray using sheath liquid at (very) low flow rate while minimizing the dilution and peak broadening effects. Also, as the flowrate and the dilution factor are decreased, the global sensitivity of the interface is enhanced.

2.3.1.3.1. The electrokinetically pumped sheath-flow interface

Based on the liquid junction design, continuous developments by the Dovichi's group led to the introduction of an automated CE-ESI-MS system with an electrokinetically pumped low sheath-flow nanospray interface[182]. In this configuration, the CE capillary is inserted into a glass emitter through a cross that provides a sheath liquid at nanoliter per minute flowrate. This very low sheath-flow is generated based on electroosmosis due to a potential applied to the sheath liquid reservoir. This design provides a highly stable extremely low flow, therefore eliminating the large dilution typical for this type of interface. Several generations of this device were already produced as demonstrated in Figure 2.19:

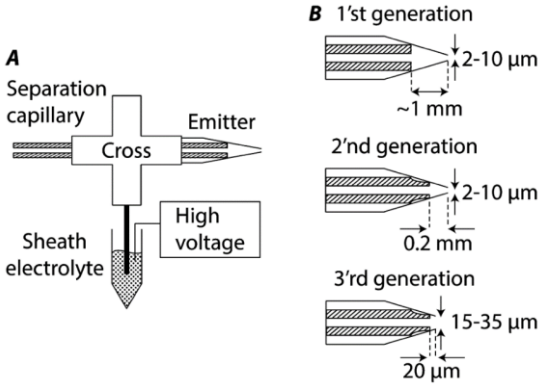


Figure 2.19 - Various generations of the electrokinetically pumped sheath-flow interface (from Sun *et al.*[183])

The great advantage of this interface is the high stability of the delivered sheath liquid. This setup could be used for more than 5000 minutes and still provide excellent results, as demonstrated in Figure 2.20:

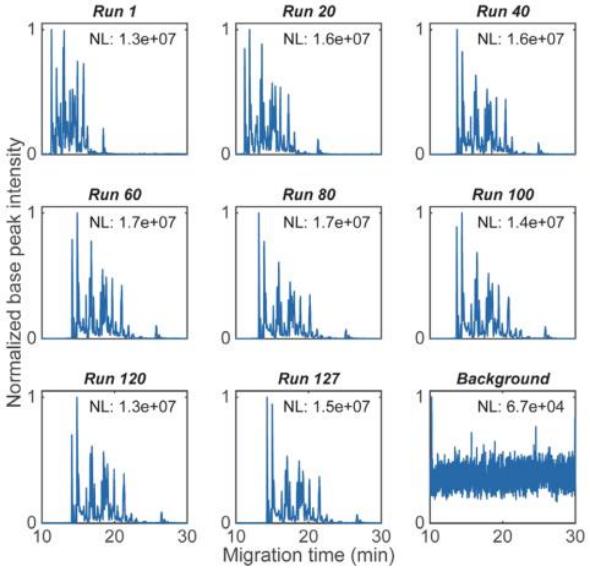


Figure 2.20 - Comparison of electropherograms obtained with the third generation of the electrokinetically pumped sheath-flow interface (from Sun *et al.*[183])

CMP Scientific commercialized this type of interface as the EMASS-II ion source. This CE-MS interface has already shown great promises for proteomic applications[182].

2.3.1.3.2. Comparison and evaluation of different recent interfaces

Recently, the comparison between the three most commonly used commercially available CE-MS interfaces was performed to evaluate their respective analytical performances. In this study, Höcker, Montealegre and Neusüß compared a modified version of the electrokinetically pumped interface, the porous sheathless interface and the classical co-axial sheath-flow interface for the analysis of various biomolecules including organic acids, light and heavy chains of a monoclonal antibody, and peptides[184]. The comparison was performed while maintaining the largest set of experimental parameters (such as instrument parameters, samples, time frame, mass range in positive and negative ion mode) constant to compare these different setups in the most objective strategy possible.

The comparison highlighted a similar increase in sensitivity of the low-flow interfaces (i.e. the electrokinetically pumped interface and the porous sheathless interface) compared to the co-axial sheath-flow interface from 13 for peptides up to 114 for the light chain of the monoclonal antibody. These results open exciting analytical developments about low-flow sheath liquid interfaces by providing similar analytical performances than the sheathless interface while being more tolerant about the BGE composition for method development.

2.3.1.3.3. The microfluidic "Analis" interface

Among the various available interfaces, this thesis exclusively shows results obtained on the so-called microfluidic "Analis" interface. This interface was developed in the framework of a collaboration between Analis (Suarlée, Belgium) and the Mass Spectrometry Laboratory (Wallonia FIRST Enterprise Project 6592). The development of this interface was conceived to combine as many attractive attributes as possible:

- the absence of nebulization gas to reduce both dilution and suction effects
- the use of sheath liquid in the nanospray regime to reduce ionization dependence on BGE composition but maintain high ionization efficiency. The use of SL also enables in-source reactions[185] or to deliver mass reference compounds for online MS calibration
- the versatility of the implementation on different MS instrument
- facilitated maintenance due to the easy handling of the interface

This interface can be sorted among the "low sheath-flow interfaces – junction at the tip" described above. Indeed, in this setup, the CE capillary is inserted in a conductive needle that is supplied with sheath liquid at a typical flowrate of 1μL.min^{-1} or below. In this configuration, a nanoESI source can be used and no desolvation gas is required. A general scheme of this interface is presented below:

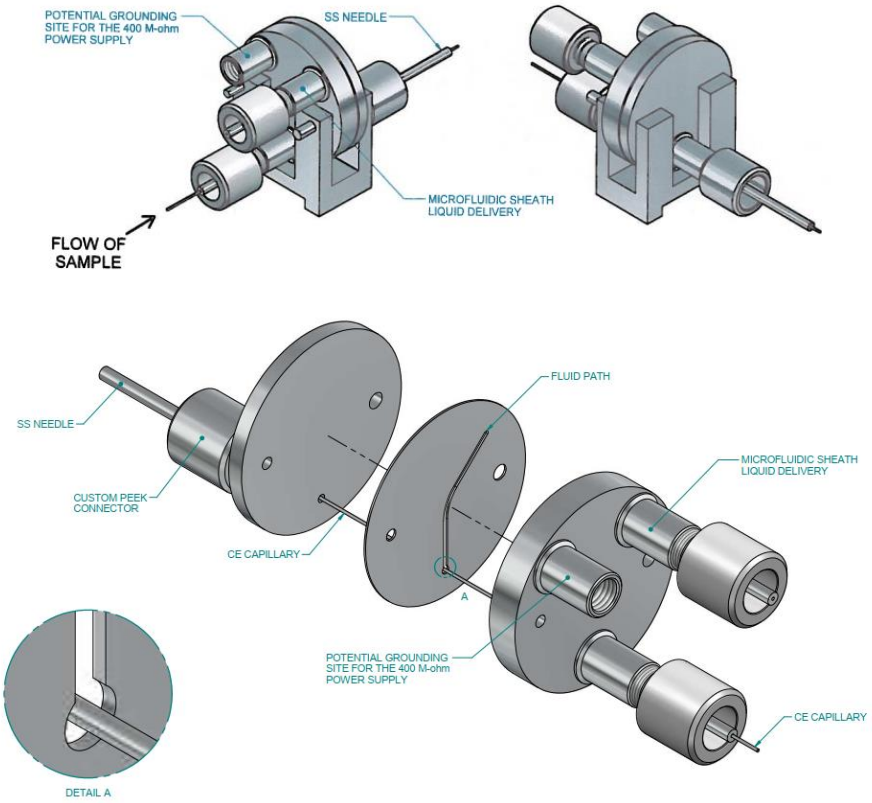


Figure 2.21 - Schematic representation of the microfluidic Analis interface

There are various advantages of this simple configuration. First, the use of sheath liquid enables the production of a stable spray, allowing to work with a wide range of BGEs and SL compositions. Also, the absence of nebulizing gas minimizes the loss of separation efficiency caused by suction effects.

This analytical performance of this interface was recently evaluated by Schappler's group[186] where the influence of the main operational parameters including sheath liquid composition and flow rate, working distance from MS inlet and ESI potential. They also modelled the mixing process of the CE effluent and the sheath liquid inside the Taylor cone to characterize its band-broadening effect. They found that its size (or cone type, see Figure 2.22 for details) was the critical feature to avoid band broadening effects.

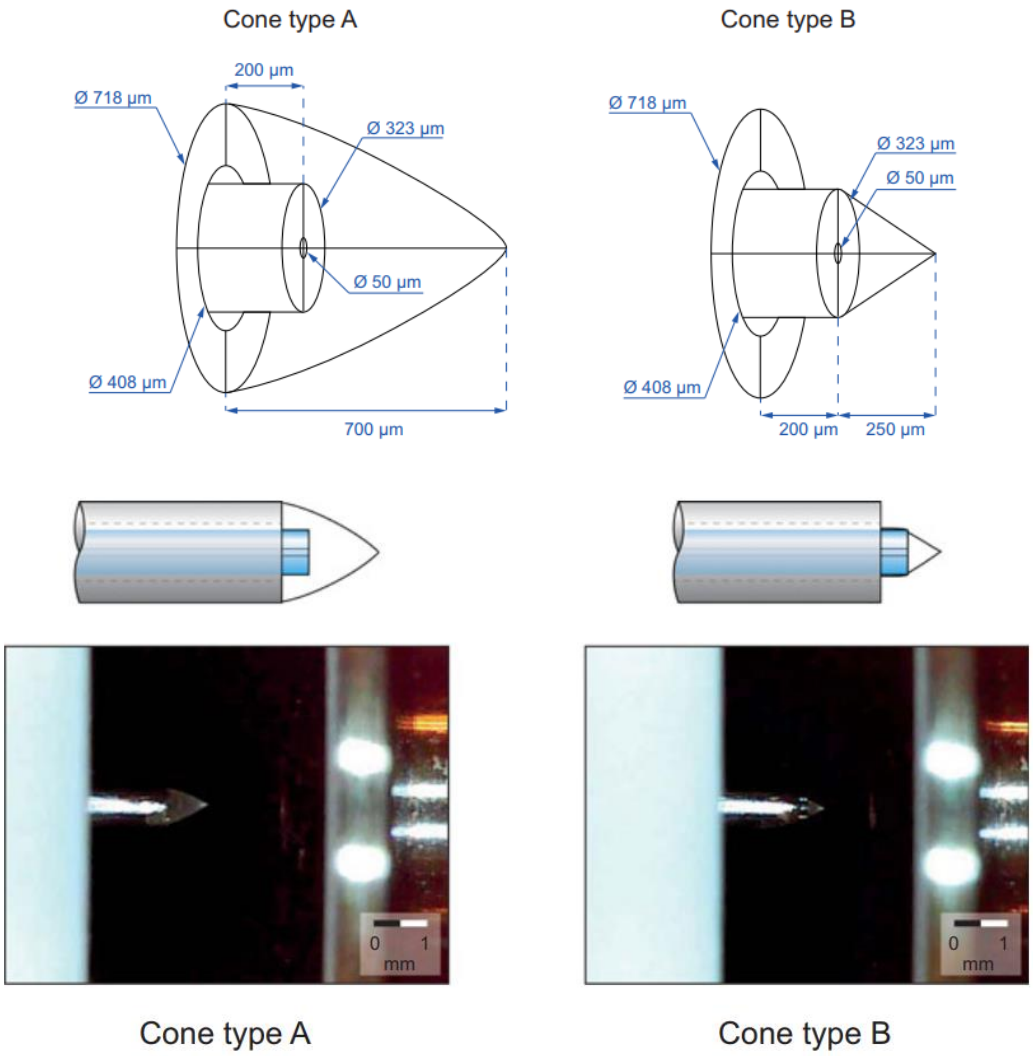


Figure 2.22 - Two types of Taylor cones produced by the microfluidic Analis interface from Gonzalez-Ruiz *et al.* [186]

In the type A cone, enlarged peak width and distorted shaped peaks were detected. On the contrary, narrow, well-defined peaks were obtained with the type B cone, as shown in Figure 2.23:

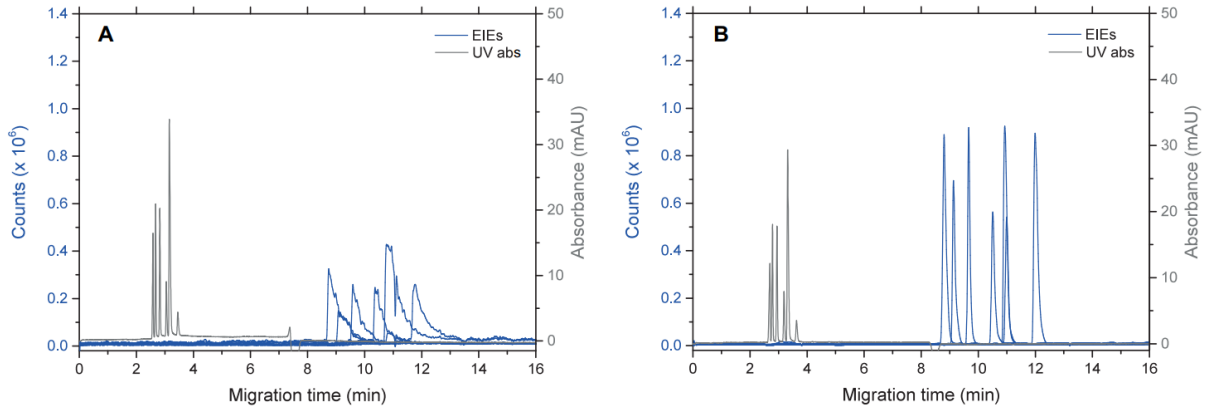


Figure 2.23 - Electropherograms obtained for the two different types of Taylor cones for a set of basic drugs with the microfluidic "Analis" interface from Gonzalez-Ruiz *et al.* [186]

This effect was explained by the residence time of the analyte in the Taylor cone. The estimations based on the theoretical calculations using the COSMOL model showed that the volumes were

estimated to be 125nL and 6.8nL for cone A and cone B respectively. With the calculated EOF and SL flowrate, the turnover time of the liquid contained in cones type A and B were roughly 10s and 0.5s, respectively. Consequently, cone A produces distorted and enlarged peaks because of the longer residence time in the Taylor cone, while cone B almost instantly sprays the analytes.

Finally, the interface was tested on a set of basic drugs analyzed by CE-TOF-MS and compared to the conventional pneumatically assisted sheath–flow interface commercially available from Agilent. The interface proved to be robust and provided repeatable results with limits of detection between 25 and 100 ppb, comparable to those obtained with the conventional sheath–flow interface. However, the low sheath–flow interface provided narrower peak widths (FWHM = 7–13 s) compared to the conventional SL interface (FWHM = 9–16 s). The results from the robustness and repeatability studies also showed that the interface can be operated for routine analyses.

2.4. CE-MS and other separation techniques

2.4.1. Comparison of CE-MS and LC-MS

The constant evolution of HPLC including improvements in column packing, instrumentation and various types of detectors available have turned HPLC a vital technology for the biological and pharmaceutical sciences. Nowadays, HPLC is routinely used in almost every scientific laboratory that requires qualitative and/or quantitative analyses[187–189]. The technique has proven to be a robust analytical tool for a wide range of applications, including for trace and ultra-trace analysis. The current trend in HPLC is in the use of capillary columns, where the main advantages are smaller sample sizes, low eluent consumption and easier interfacing with mass spectrometry (MS). Another great advantage of HPLC is its ability to be used for preparative or even process-scale separations[190].

HPLC is currently the preferred analytical approach for many applications. However, recent progress in CE, and notably the easier hyphenation to MS, offers an interesting alternative for many analytical issues. The historical struggle to efficiently couple CE to ESI-MS was due to the complicated closure of the electric circuits to efficiently provide both efficient electrophoretic separation and ESI ionization (see section 2.3 for details). However, the recent progress in hyphenation of CE with ESI-MS by the development of robust and sensitive interfaces combined to the versatility of CE modes of operation[108] allow CE-ESI-MS to be implemented in a growing number of research areas. Currently, the rapid development and implementation of CE-ESI-MS methods represents a promising alternative to classical LC methods. As a result, CE-ESI-MS has been successfully applied for the analysis of significant polar biomolecules such as peptides, proteins, nucleic acids, glycans, and drugs and related metabolites.

CE provides several advantages over LC. First, CE is particularly suited for the separation of charged analytes whereas the traditional reversed-phase LC separates analytes according to their interaction with the hydrophobic stationary phase. As a result, polar/charged species are often not well separated (or even not separated) by reversed-phase LC. For these problematic RPLC separations, CE brings an elegant solution. In addition, the implementation of chiral analysis by CE is also rapidly gaining attention due to the non-requirement of expensive chiral stationary phases like in LC[191].

On the other hand, despite its broad spectrum of applications, CE-MS currently suffers from a relatively bad reputation (probably due to its relative immaturity), impairing the establishment of the technique as a well-accepted alternative to LC-MS. In particular, the minute amounts of injected sample in CE can sometimes lead to insufficient sensitivity for real-world samples. In this regard, the implementation of sample preparation/ preconcentration techniques such homogenization, centrifugation, precipitation, solvent extraction, liquid–liquid extraction (LLE), solid-phase extraction (SPE), (micro)dialysis and ultrafiltration and stacking methods can significantly improve the sensitivity of CE–MS methods[192].

Finally, because the separation mechanisms are different in CE (differential migration according to the charge and the hydrodynamic radius of the analyte) and LC (repartition coefficient between a stationary phase and the mobile phase), they often provide complementary information. The following section exemplifies the comparison between HPLC-MS and CE-MS for proteomic analyses and demonstrates the complementarity of these two techniques.

2.4.1.1. Proteomic analysis

After the major breakthrough realized in the genome sequencing, and especially the complete sequencing of the human genome, the so-called "postgenomic era" research is currently further increasing our understanding of cellular pathways and networks by the development of proteomics[193]. However, unlike the relatively stable genome, the proteome is more complex due to post-translational modifications, introduction of sequence variants, etc… Besides, its composition evolves in time in multiple and inter-related responses to internal processes and external environmental stimuli. As a result, this complex and ever-changing analytical target poses an unprecedented analytical challenge. Unravelling the specific function of a particular protein and deducing complex molecular pathways involved in cellular responses can only be achieved by the use of rapid and quantitative analysis of the proteome.

The proteome is usually defined as "the set of expressed proteins in a given type of cell or organism, at a given time, under defined conditions". Proteomics is the study of the proteome and is now increasingly studied to unravel complex biological processes, including the detection of diseases, through the discovery of "biomarkers". A biomarker is "a characteristic entity that is objectively measured and evaluated as an indicator of a normal biological process, pathogenic process, or pharmacological response to a therapeutic intervention". According to the FDA[194], a "valid biomarker is measured by an analytical test system with well-established performance characteristics and for which there is an established framework or body of evidence that elucidates the physiological, toxicological, pharmacological, or clinical significance of the test result".

Current biomarker research is progressively becoming aware that the detection of only one or a few biomarker(s) for diagnostic purposes is vain. Instead, the introduction of the simultaneous measurement of the largest possible range of biomarkers (described as biomarkers patterns) should enable a statistically more robust detection of diseases (for example, patients with cancer and those without) by reducing the rate of false positive/false negative results. Once identified, the set of biomarkers can be used to diagnose a disease or to monitor the effects of a treatment. Consequently, it is anticipated that a more precise diagnosis will lead to more adapted therapies as well as to earlier intervention[195]. The major challenge is then the discovery of these biomarkers and most importantly, their effective validation[196].

From an analytical point of view, the study of the proteome is a serious analytical challenge because of the complexity and wide dynamic range of protein expression. Additionally, physicochemical properties of all the proteins contained a single sample can be diametrically opposed. The range of protein characteristics including solubility, hydrophilicity/hydrophobicity, isoelectric points (pI) values (pI expressed proteins can vary from 2 to 12), and a wide range of molecular weights requires that the analytical method used for such analysis should be of the highest separation efficiency possible but also being able to deal with the broadest range of molecular features possible.

As a result, the analytical procedure used in this context must be able to separate and detect as many analytes as possible, sufficiently robust to account for the intrinsic variation of biological samples (which often contain large quantities of interfering compounds and salts) and perform as uniformly as possible from analysis to analysis. It must also be as fast as possible (to allow high throughput) and relatively economical.

Proteomic analysis is typically performed on various biological samples. Among the available biological samples, body fluids are a target of choice as they serve for the transport of nutrients and elimination of metabolic products. Consequently, body fluids, such as serum, urine, and

cerebrospinal fluids are intensively used in so-called "non-invasive" methods for the detection of biomarkers. The body fluids analysis provides several advantages including less pain for the patient when collecting the sample and a higher accessibility for clinical tests. However, the requirements of tissue-specific biomarkers detection in body fluids is that 1) the target molecules is secreted and 2) the biomarker be identified as disease tissue-specific from a plethora of secreted biomolecules from various cell types and organs[197].

In early developments, the use of antibodies and the subsequent introduction of immunological assays was a first step for the discovery of biomarkers. These assays allow the detection and quantitation of specific target molecules (mainly peptides and proteins) in body fluids through standardized, automated high throughput tests, such as ELISAs. Although effective, the major drawbacks of this technique are the requirement of the prior knowledge of the target molecule and the production of a specific probe. These problems rapidly disqualified these approaches because of the much larger set of biomarkers that must be assessed and the need to identify unknown substances.

With the increasing progress in mass spectrometry, MS-based proteomics was expected to analyze thousands of peptides and/or proteins in one sample in a fast and efficient manner. The technique was promising a glorious future for proteomics, yet unanticipated limitations caused considerable disappointment in its early application:

- First, the complexity and wide dynamic range of peptides/proteins in body fluids (which for the example of serum can exceed 10^{12}) is simply impossible to handle by MS alone. As a result, the hyphenation to (very) high resolution separation methods to deal with this large dynamic range was necessary
- Then, the introduction of sample preparations (by the use of tryptic digest for example) to circumvent frequently encountered problems (such as LC columns clogging) resulted in an even greater complexity
- Finally, the amount of MS data generated by only one MS-based proteomic analysis is impossible to efficiently treat by hand, and the introduction of robust software for the automatic data interpretation was also mandatory

Considerable effort has been (and are still being) made to meet these different requirements. Currently, strategies relying on the combination of high-resolution separation techniques and advanced mass spectrometric detection represent the best approach to address the challenging analysis of the proteome.

Two major high-resolution separation techniques are used for the efficient peptide/protein separation: liquid-chromatography based technologies (including 2D-LC) and electrophoretic methods. Major approaches in LC-based proteomics include the use of reversed phase chromatography (RPLC, separating on hydrophobicity) and strong cation exchange chromatography (SCX, separating on charge). Among the electrophoretic methods, gel electrophoresis (including two-dimensional gel electrophoresis or 2-DE) with subsequent MS analysis is still commonly used for proteomics analysis. However, the introduction of capillary electrophoresis tends to decrease its application due to its various limitations. Among CE-based techniques, capillary zone electrophoresis (CZE) and capillary isoelectric focusing (CIEF) seems to have the greatest potential in MS-based proteomics[198]. Interestingly, the bioinformatic treatment of the data generated by CE-MS and LC-MS is comparable[199].

While LC is more commonly applied for MS-based proteomics, CE offers several advantages over LC-based method including:

- High resolution achieved with fast separations
- Use of inexpensive capillaries compared to LC-columns
- Stable and most of the time constant CE effluent. In particular, the absence of gradient which often results in change of the ionization parameters
- Ability to fast recondition the capillary by the use of NaOH instead of extensive washing steps

Despite its various advantages, CE still struggle to be implemented for proteomics probably because of its reputation. Before CE-MS hyphenation was efficient, CE was majorly coupled to UV detection, where specific analyte characterization was not yet possible. Furthermore, the limited amount of sample that can be injected using CE (generally a few tens to hundreds nL) further increased the problem of efficient detection. In contrast, LC allowed microliters to milliliters of sample to be injected and was easily interfaced to ESI-MS. As a result, the characterization of the analytes could be performed by its retention time (LC) and m/z ratio (MS). Not surprisingly, LC-MS methods (online and offline) became the reference methods for protein and peptide research. In addition, the relatively tedious hyphenation of CE to MS has contributed to the idea that CE-MS is difficult to manage and give rather unsatisfactory results.

However, the recent developments of efficient interfaces for CE-MS (see section 2.3) combined to the rapid development of MS in terms of detection limits make CZE-MS an increasingly powerful tool for proteomic analysis. Although the sensitivity of the sheath liquid approaches is generally lower than the sheathless approach, the increased stability of such system seems to be generally favored. Besides, appropriate optimization of the ionization parameters and a reduction of the flowrate of the sheath liquid[200] can lead to detection limits in the high-attomole range[122].

In CE, a particular attention should be taken about the adsorption of peptides and proteins to fused-silica surfaces. The interaction results from electrostatic interactions between positively charged peptide and protein surface regions and dissociated silanol groups, characteristic of the fused-silica surface. The most straightforward way to minimize these interactions is to reduce the capillary surface charge density by using an extremely acidic BGE. The use of volatile acids as formic or acetic acid were successfully used for peptide, but many proteins are either insoluble or unstable in such acidic buffers. An extremely basic BGE can also be considered where analytes have the same charge as the capillary, but stability and solubility problems remain. Though effective in some cases, these methods cannot be applied in situations where an intermediate pH is necessary to obtain the desired separation in CZE or simple incompatible with CIEF, where the pH must vary all along the capillary. Consequently, intense effort has been devoted to the development of capillary coatings[30] that enable efficient peptide and protein separations over a wide pH range where uncharged, hydrophilic coatings have shown great promise[108]. The polyacrylamide coating were introduced by Hjertén[201] in 1985 and still represent frequently used method for capillary surface modification, mainly due to the widespread use of polyacrylamide in slab-gel electrophoresis.

When comparing LC-MS and CE-MS approaches, the main advantage of LC seems to be its higher loading capacity. As a result, LC is a superior alternative if sensitivity is critical. On the downside, the separation of large molecules and analytes covering a wide range of size and hydrophobicity is more difficult using a LC approach. LC is also very sensitive to precipitates which results in clogging of the columns. 2D-LC systems tend to provide a considerable amount of information, but they are highly time consuming and consequently poorly adapted to routine clinical analysis. Because of

its various advantages, CE-MS can be used as a powerful tool for clinical diagnosis and biomarker discovery based on protein and/or peptide patterns[95].

However, it is important to highlight here that both techniques have advantages and drawbacks and they should be considered as complementary. The use of only one technique generally results in a partial detection of peptides/proteins compared to the combined results of the two techniques. A good example of this highly valuable complementarity is provided by Klein *et al.*[202]. In this study, the combination of CE-MS/MS and LC-MS/MS was showed to increase sequence coverage and the rate of identification of peptide sequences, as shown in Figure 2.24.

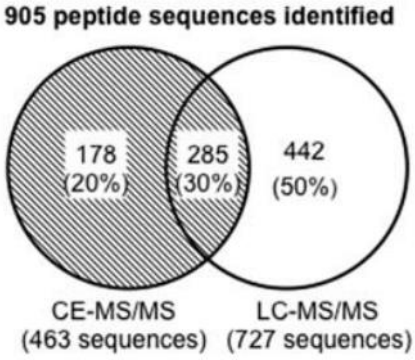


Figure 2.24 - Venn diagram depicting the number of unique peptides found in LC-, CE-MS/MS and in both methods of the urinary low-molecular weight proteome analysis from Klein *et al.* [202]

The results show that urine sample analysis by LC- and CE-MS/MS identified 905 unique peptide sequences, where 50% only identified with LC, 20% only with CE and 30% with both techniques. The higher LC coverage was attributed to the higher loading of sample onto the LC column. Interestingly, peptides identified in CE only were generally small and highly charged, explaining their non-detection by LC due to the inability to efficiently bind to the LC column. The authors conclude that LC-MS/MS and CE-MS/MS are highly complementary in identifying peptide sequences.

In another example by the group of Dovichi[203], the CZE-ESI-MS/MS analyses of proteins from *Mycobacterium marinum* (a bacterium causing opportunistic aqueous infections in humans) was compared to UPLC-ESI-MS/MS analyses. For the electrophoretic analysis, 334 peptides corresponding to 140 proteins were identified in 165 min. In comparison, the chromatographic approach identified 388 peptides corresponding to 134 proteins in 180min, as graphically demonstrated in Figure 2.25.

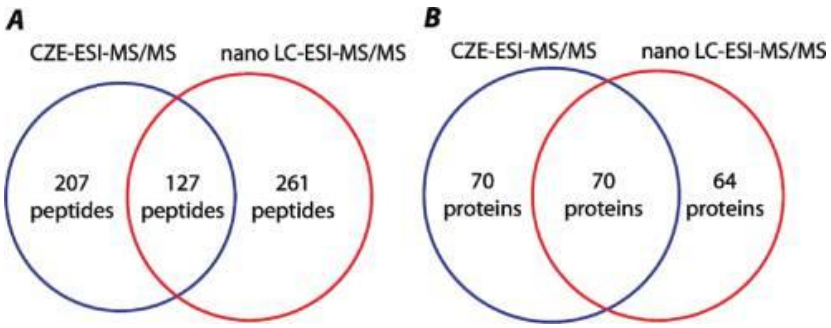


Figure 2.25 - Venn diagram of peptides (A) and proteins (B) identified by CZE-ESI-MS/MS and nano-LC-ESI-MS/MS of proteins from *Mycobacterium marinum* from Klein *et al.*[202]

Globally, 62% of peptides identified in CZE-ESI-MS/MS and 67% in UPLC-ESI-MS/MS were unique and the results show that CZE-ESI-MS/MS generally favors basic and hydrophilic peptides with low molecular masses. Most importantly, the introduction of CZE-ESI-MS/MS in the workflow increased the number of detected peptides and proteins by almost 35%.

Next to the complementarity with LC-MS, current development of CE-MS is also pushing the limits of required materials to perform proteomic analysis. In a recent study, Dovichi's group reported the use of femtogram amount of materials for proteomic analysis. The developments of such methods, with detection limits of 1 zmol, opens the way to the identification of proteins with relatively high abundance at the single-cell level[204].

In conclusion, CE-MS seems to be a powerful tool for proteomics analysis, especially when combined to LC-MS. Besides working on the protein/peptide itself, CE is also a powerful tool for the identification of PTMs such as phosphorylation[205] and glycosylation[206].

2.4.2. Comparison of CE-MS and IM-MS

Ion Mobility Spectrometry (IMS) is a separation technique where ions migrate through a buffer gas in the presence of an electric field. In that sense, IMS can be viewed as an electrophoresis in the gas phase where the separation is also governed by the size and shape of ions (see below for details). Since mass spectrometry is also a gas-phase analytical method of ions, the hyphenation of the two techniques, referred as "Ion Mobility – Mass Spectrometry" (IM-MS) provides an efficient analytical platform, where complementary information about the analytes can be obtained.

Although the coupling of IM-MS was introduced in 1962, it has only recently attracted attention, mainly because of the relatively recent commercialization of IM-MS instruments. The subsequent increase of research coupled to the introduction of various ion mobility modes has demonstrated the valuable hyphenation of these two techniques.

Currently, IM-MS is increasingly implemented in two main research areas. First, IMS can be used as an additional separation dimension to increase the analytical performance of the MS analysis. In particular, IMS can separate isobaric species (i.e. with identical m/z ratios), which cannot be distinguished by MS alone. The second main application of IMS is the structural assessment of ions in the gas phase[207].

2.4.2.1. Drift-time ion mobility spectrometry (DTIMS)

DTIMS was the first type of ion mobility to be introduced and is conceptually the easiest. In this setup, ions are introduced in a drift tube that is filled with a drift gas (typically helium) and a static uniform electric field is applied through the drift tube, inducing the migration of ions as shown in Figure 2.26:

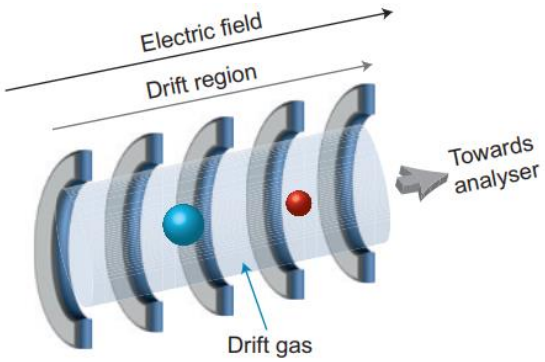


Figure 2.26 - Schematic representation of a drift tube IMS from Lanucara *et al.* [207]

As a result and similarly to the electrophoresis in solution, the force exerted by the electric field on the analyte ion is balanced by frictions with the buffer gas. Consequently, the analyte reaches a steady-state velocity inside the mobility cell referred as "v_d" and the required time of an ion to drift through the device is referred as the "drift time (t_d)". The ion mobility (K) is the ratio of the velocity of the ion (linked to its drift time (t_d) and the length of the mobility cell (L) and the electric field:

$$K = \frac{v_d}{E} = \frac{L}{t_d . E}$$

The Mason-Schamp equation[208], valuable in the case of low field limit (in terms of E/N), describes the relationship between the ion mobility and related physical parameters, including the charge (q) and the collision cross section (or CCS) of the analyte (Ω):

$$K = \frac{3}{16N}\left(\frac{2\pi}{\mu\, k_B T}\right)^{1/2}\frac{q}{\Omega}$$

where *K* is the ion mobility, *q* is the charge of the ion (i.e. z.e), *N* is the number density of the drift gas, *μ* is the reduced mass of the ion–neutral drift gas pair ($\mu = mM/(m+M)$, where *m* and *M* are the ion and gas-particle masses), k_B is the Boltzmann constant and *T* is the gas temperature in Kelvin.

In a first approximation, the CCS is a structural parameter of the analyte that can be viewed as the rotationally averaged projection of a molecule on a 2D plan (see Figure 2.27).

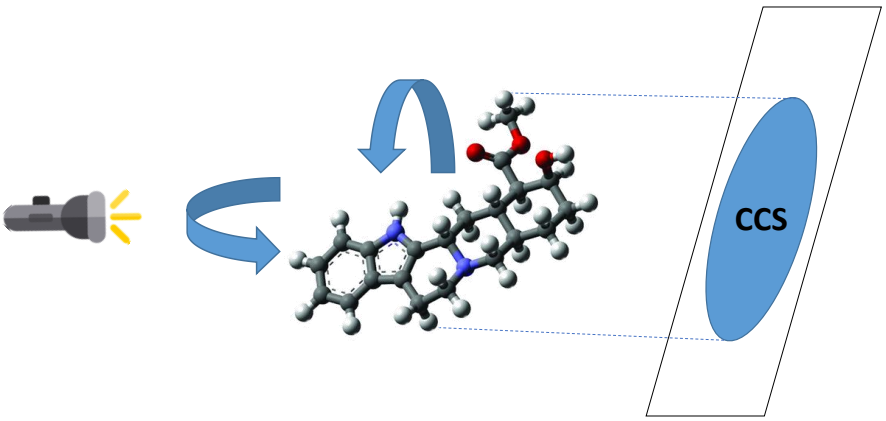


Figure 2.27 - Schematic representation of the rotationally averaged CCS

The Mason-Schamp equation implies that compact structures (lower CCS) migrate faster than more extended ones (larger CCS) due to fewer collisions with the buffer gas. In addition, ions with higher charge state also migrate faster than ions with lower charge state.

It should be mentioned that the proportionality between Ω and K is only valid at "low-field limit" conditions and below, i.e. when the ratio between the electric field strength and the buffer gas density is small[209].

2.4.2.2. Travelling-wave ion mobility spectrometry (TWIMS)

Compared to the drift tube of DTIMS, TWIMS is performed in a device constituted by a series of ring electrodes referred as a stacked ring ion guide (SRIG) where travelling voltage waves are applied[210]. Radio-frequency voltages of opposite phases are applied to adjacent electrodes to radially confine the ions inside the SRIG. On top of that, a repeating train of direct current (DC) pulsed voltage is applied to the successive electrodes of the SRIG to create a "travelling wave" that propels

the ions axially. Typically, a quick succession of travelling waves is used to induce the migration through the mobility cell. During their migration, ions can either "be carried" by the wave or "roll over" it. Consequently, ions presenting a low CCS value (i.e. more compact) are more efficiently transported by the successive waves because they undergo less roll-over events. As a result, they require less time to travel through the mobility cell and are characterized by a higher mobility value in $m^2.V^{-1}.s^{-1}$ (see Figure 2.28).

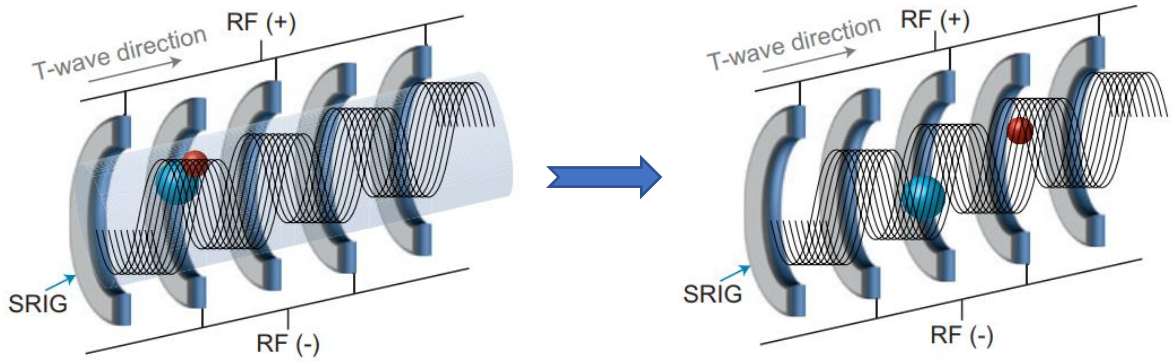


Figure 2.28 - Schematic representation of Travelling Wave Ion Mobility Spectrometry from Lanucara *et al.* [207]

In this device, the speed and the amplitude of the waves can be optimized to achieve efficient ion mobility separations. Because the TWIMS device is generally operated below the low-field limit, the determination of CCS is possible, although appropriate calibration procedure is required[211–214]. Indeed, unlike the DTIMS device, the direct relationship between Ω and K_0 is no longer valid because of the constantly changing electric field. Consequently, calibration of the drift time under defined conditions (gas nature, pressure, waves speed and height, …) is mandatory. In this regard, calibrants of similar physical and chemical features than the analytes are required.

2.4.2.3. IM-MS in analytical chemistry

There are several fields where IMS can bring a valuable contribution to analytical methods as an additional separation step. First, workflows where a classical separation technique (i.e. CE or LC separation) cannot be implemented, IMS can bring a fast, additional analytical dimension to the MS analysis. Such use of IMS has been demonstrated for MS imaging experiments. In MS imaging experiments, the aim is to localize the spatial distribution of analytes of interest such as drug metabolites or proteins in tissues and biological samples thanks to mass spectrometry[215–217]. The analytes are therefore directly ionized from a complex matrix and analyzed by MS. The introduction of IMS in this analytical procedure notably improves the identification confidence level of species of interest. For example, Stauber and co-workers incorporated TWIMS in complement to a MALDI-MS imaging workflow to separate and spatially localize isobaric compounds[218]. In contrast, the same workflow in absence of IMS provided a composite image where the two species were co-localized due to identical m/z ratios. Another example is the use of IMS in MALDI imaging is to get rid of matrix interference, allowing the accurate determination of the spatial distribution of the analyte of interest[219].

Another area where IMS has proven to be beneficial is in conjugation to other separation techniques. Typically, the analysis of complex samples relies on the combination of separation techniques such as LC or CE before MS analysis. MS is a suited detection method because the typical timescale of MS analysis is generally (much) lower than the timescale of the separation technique. Because IMS operates in the millisecond timescale, it can perfectly be inserted between the separation method and the MS analysis without impairing the analytical performance of the

workflow. In such workflow, TOF-MS is usually preferred as mass analyzer due to the well-fitting duty cycles (IMS: ms range; TOF-MS: µs range). Consequently, the analysis of complex samples can be improved by the incorporation of IMS without detriment to the already established complementarity of separation techniques and MS. The analytical performance of such approach is increased because of the complementary separation mechanism and additional peak capacity afforded by IMS. IMS can also be used to remove the background chemical noise in complex mixtures to improve the detection of low-abundant compounds or prevent false positive identification (e.g. pesticides[220])

Typically, the characteristic limits of detection achieved by MS are much lower than for most other analytical approach and LC-MS routinely detects low attomole amounts of analyte, even in complex mixtures[221]. Various groups showed that the incorporation of IMS in the analytical workflow critically improves the sensitivity of such methods. For example, the incorporation of IMS for non-targeted discovery proteomics by classical LC-MS showed a large increase (by 60% and above) in the number of identifications at both the peptide and protein levels using LC-TWIMS-MS compared to classical LC-MS alone[222,223]. In another example, Saba and co-workers demonstrated a greater than 10-fold improvement in the detection limits of peptides in complex mixtures when using LC-IM-MS compared to classical LC-MS[224].

Concerning the hyphenation of IMS to CE, early attempts of this coupling have been performed to replace the classically used UV detector by an IMS device[225]. Of course, since the MS information is not present in such a device, the characterization capabilities of this kind of approach is limited. Thus, similarly to the incorporation of IMS in the LC-MS workflow, the introduction of IMS in an already established CE-MS system was recently performed. In their study, Jooß *et al.*[226] showed that the combination of CZE and DTIMS coupled to MS applied for the separation and characterization of native and APTS-labeled *N*-glycans highly improved the separation efficiency of such complex samples. In fact, almost each peak resolved in CZE provided multiple peaks in IMS. This was explained by the presence of isomeric forms, including different linkages and/or gas-phase conformers. In another example, the introduction of IMS into a CE-MS workflow to reduce chemical noise and enhance detection limits was applied for the analysis of complex lipopolysaccharides[227].

In complement, the drift time can also be used as an additional identification-point (IP) to already established chromatographic and mass spectral data. In this context, IMS provides a higher degree of confidence in the overall identification process, therefore decreasing the likelihood of incorrect negative/positive results. A recent study by our group on pesticides[220] showed that drift time values can be used to increase the confidence in the identification of such analytes without extending the analysis time or modifying the sample preparation procedure. A statistical analysis of the results for 100 tested pesticides demonstrated that a 2% variation in drift time could be considered as the maximum acceptable criteria parameter for identification purpose, in agreement with existing criteria, for instance, for GC or LC retention time in European documents. A similar workflow[228] (UHPLC coupled to a TWIMS-QTOFMS) for pesticides screening in complex salmon feed matrices showed that CCS measurements present high intra- and inter-day repeatability (RSD <1%) and were almost unaffected by the complexity of the investigated matrices (CCS values ≤1.8%). The introduction of TWIMS in this workflow provided an extra-dimension, which resulted in increased peak capacity and selectivity in real samples. Thus, many false-positive results could easily be detected just by applying the maximum CCS tolerance of ±2%. CCS was therefore proposed as a valuable additional identification-point in the pesticides screening workflow.

Finally, because many current applications require high throughput sample analysis, IMS could play an increasingly important role due to its speed of analysis. Indeed, the typical duration of an IMS analysis lies in the millisecond timescale, which is much faster than traditional liquid phase separations (chromatographic and electrophoretic strategies). As a result, there should be an increasing implementation of IMS in analytical chemistry fields where rapid analysis is required (for example drug screening). IMS could also increase the speed of analysis by replacing more time-consuming separation techniques (typically HPLC, UPLC and/or CE), provided that sufficient separation between the species of interest is retained.

2.4.2.4. IM-MS for structural analysis

The other major application of IMS is the assessment of ions structural features in the gas phase. Although all modes of IMS separate analytes based on of their conformation, only the time-based mobility devices (DTIMS or TWIMS for example) allow a CCS determination. Characteristic structural information can be derived from the CCS value (which gives an indication of the ion's size and shape) and compared with data acquired with other structural techniques such as X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. IMS presents two main advantages compared to other biophysical techniques. First, only small amounts of sample (nanogram quantities) are required to perform an IM-MS analysis. Second, impurities are not an issue due to the ability to selectively isolate the species of interest from complex samples. As a result, samples of much lower purity can be analyzed by IMS compared to X-ray crystallography or NMR spectroscopy for structural characterization[229]. Table 2.3 summarizes the various advantages and disadvantages of major structural methods.

Table 2.3 - Comparative table of advantages and drawbacks of major structural techniques (adapted from Lanucara *et al.*[207])

	IM-MS	MS^n (*)	NMR spectroscopy	X-ray crystallography	Circular dichroism
Phase of analyte during analysis	**Gas**	**Gas**	**Liquid (or solid)**	**Solid**	**Liquid (gas/solid)**
Advantages	- Very sensitive - Can analyze mixtures - Fairly rapid for simple structures, with the most time-consuming parts being molecular dynamics simulations - Under carefully controlled ionization and acquisition conditions, can be used to determine structure from native conditions - Can determine stoichio-metry of complexes - Assesses 3D structure in dynamic motion	- Very sensitive - Can analyze mixtures - Fairly rapid, aided by mass spectra search engines and spectral databases - Can determine stoichio-metry of complexes - Products of electron mediated dissociation can be used to infer 3D structure	- Non-destructive - Can be very rapid when analyzing small molecules (<500 Da) - As most biologically relevant compounds are found in aqueous media, the ability to perform solution-state NMR spectroscopy means that structure can be analyzed under native-like conditions - May allow stereochemical information, bond angles and distances to be elucidated - Conventional solution state by (2D) NMR spectroscopy can measure 3D structure in dynamic motion	- Non-destructive - Can directly determine structural information at the atomic level - May allow stereochemical information, bond lengths and angles to be elucidated	- Non-destructive - Can measure exchanging structures (>picoseconds) - Rapid direct diagnostic test for certain structural features, for example relative helicity - Can determine native structural features - Allows determination of stereochemistry
Disadvantages	- Analyte must be able to be ionized - Relies on molecular dynamics simulations to indirectly determine precise structural information from CCS values; simulations become more challenging as molecules become larger - Destructive	- Analyte must be able to be ionized - Difficult to ascertain detailed 3D structure information directly from mass spectra - Destructive	- Difficulty to analyze mixtures of products; samples often have to be purified and concentrated, which may affect the structure of biological samples - Analysis of spectra becomes difficult for larger molecules	- Requires purified and crystallized material; may be time-consuming or impossible to generate crystals - Analyte may be damaged by the X-rays	- Cannot be used for mixtures. - Requires relatively concentrated (~0.5 $mg.mL^{-1}$) purified samples - Gives no specific structural information at the atomic level

(*)MS^n refers to multistage MS/MS experiments designed to record product ion spectra where "n" is the number of stages of tandem mass analysis[230].

2.4.2.5. Determination of ion's CCS: experimental CCS and theoretical calculations

In IMS, the analyte's structure can be assessed in the gas phase using the experimentally derived, rotationally averaged CCS values. These CCS values reflect the gas-phase conformation and shape of the analyzed ions under defined experimental conditions. In general, experimental CCS values are compared to theoretical CCS values derived from molecular dynamics simulations (MD), Molecular Mechanics (MM) or density Functional Theory (DFT). The structures used as inputs for calculations are sometimes derived from X-ray and/or NMR data[231–233]. Unfortunately, the choice of the calculation method (i.e. the force field that is used for the calculation) can heavily impact the theoretical results and needs to be carefully selected for a reliable interpretation of experimental IMS data. Also, this approach is difficult to apply with large structures (for example large proteins) because of the extremely demanding computational costs.

There are currently 3 main strategies for the calculation of theoretical CCS values (where helium is used as buffer gas):

1) the Projection Approximation (PA)[234] calculates the CCS value by averaging the geometric projection areas according to all possible orientations of the ion. The major advantage of this method is the relatively inexpensive computational cost. However, long-range interactions or several features of the scattering between the ions and the drift gas are not considered. Because of these limitations, CCS for ions over ~2 kDa are usually underestimated (typically 20%, until 30%) by this method and is therefore essentially used for predicting CCS of small molecules.
2) the Exact Hard-Sphere Scattering (EHSS)[235] calculates the CCS value by considering the collision integrals, that is the scattering and the collisions of the ion with the drift gas. Long-range interactions between the gas and the ions are however still ignored. This method generally provides a better approximation of the CCS value than the PA approach. In geometries with concave surfaces, CCS provided by the PA and EHSS methods show deviations as large as 20%. A recent improvement of this method, the Diffuse Hard-Sphere Scattering (DHSS), implements long-range interactions between the gas and the ion[236,237].
3) the Trajectory Method (TM)[238] is generally considered as the most reliable method as it takes into account collision effects and long-range interactions between the drift gas and the ion. As a result, this method is generally preferred for a more accurate description of biomolecules[239]. Because of its better accuracy, the trajectory method requires demanding computational resources, especially for very large systems. In this case, the EHSS or DHSS are often a good compromise between required calculation time and accuracy.

The schematic representation of how each calculation method takes into account the interaction with the gas and the analyte to calculate its CCS value is presented in Figure 2.29.

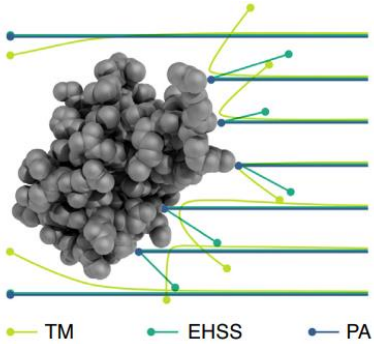


Figure 2.29 – Graphical representation of the different CCS calculation methods. PA "only" discriminates hits and misses and estimates the CCS value from the fraction of hits. The EHSS considers (multiple) scattering after initial

collision, and TM calculates the gas molecules' trajectories, whereas EHSS and PA only address direct contact between gas molecules and analyte. Figure from Gabelica *et al.* [240]

2.4.2.6. Structural information provided by IM-MS

As explained above, a detailed structural interpretation of the CCS value is generally based on X-ray or NMR-derived structures as inputs for the prediction of theoretical CCS values. However, despite the availability of these structures, IM-MS can be used as a structural tool for the study of dynamic systems where ion structures are assessed in the absence of solvation, counter-ions or any other chemical species. In this context, the measurement of ionic mobilities in the gas phase and subsequent determination of the associated CCS allows to assess the conformational landscape that studied species can access in the gas phase. One major advantage of IM-MS over classical structural characterization methods is the ability to evaluate dynamic changes in the analyte structure, which has previously been difficult to perform for most biological samples due to the averaged structures provided by NMR or X-ray crystallography for example.

However, most recent IM-MS studies still require comparison with data provided by more "traditional" structural techniques or the assistance of computational chemistry for a detailed structural evaluation of the CCS information. Compared to NMR or X-ray crystallography, IMS does not provide detailed atomic-level structural information but this lack of detailed structural information is counterbalanced by the ability to study conformational dynamics of a large range of molecular systems. Indeed, the great strength of IM-MS compared to many other biophysical techniques is its ability to interrogate dynamic heterogeneity of systems. In this regard, the determination of changes in mobility and thus conformation and CCS using IM-MS can be used to address changes in properties such as conformational dynamics[241–245], folding and unfolding intermediates[241,246,247], ligand-induced conformational changes[248,249], aggregation intermediates[250,251] and quaternary structures[252,253].

Importantly, it should be noted that IM-MS-derived structural features (i.e. gas-phase structural features) should be interpreted keeping in mind that desolvation can cause conformational changes within the structure (due to intramolecular charges solvation). The degree of change will depend on both the experimental conditions and the analyte itself.

2.4.2.6.1. Peptides

IM-MS can be used in combination with MD simulations to determine the secondary structures of peptides in the gas phase and major factors involved in the formation these secondary structures (including α-helices or β-sheets) have been studied[254,255]. It was also suggested that studying the folding process of polypeptides originating from the cell membrane in vacuo allows to mimic the low dielectric constant of such environment[256].

The incorporation of metal cations into peptidic structures and subsequent impact on the structure has also been studied by IM-MS, especially by comparing cationized (e.g. using Na^+, K^+, NH_4^+) versus protonated peptide ions[257]. Similar studies have studied conformational changes induced by the introduction of metal cations on protein-ligand complexes[258,259]. Finally, IM-MS has also helped to improve the understanding of peptide dissociation in the gas phase[260].

2.4.2.6.2. Proteins

The major difficulty in IM-MS of large biomolecules such as proteins and proteins complexes is the ability to transfer intact species from native conditions to the gas phase without inducing major structural changes due to the desolvation process[62]. Naturally, the structural characterization of

such species in the gas phase assumes that the analyzed structure resembles the native, functional species in solution. The conformation of proteins in solution is controlled by the presence of water molecules as they orient the hydrophobic regions of the protein towards the core, inducing the formation of a hydrophobic core to minimize contact with the aqueous environment[261]. Naturally, concerns about the structural integrity and subsequent major conformational changes after the removal of water molecules upon the desolvation arose[239,262].

In this context, several studies including both experimental and theoretical approaches tried to unravel the link between solution and gas-phase conformation[263]. For example, several IM-MS studies focused on the exact role of hydration in the control of stability and protein conformation in the gas phase using folded and unfolded protein ions[264–267]. In these studies, water molecules were introduced into the ion mobility cell to react with protein species and the influence of the addition of water molecules on the mobility in the gas phase was studied[266]. Based on this approach, the addition of water molecules to unfolded gas-phase cytochrome c ions suggested the refolding towards more compact, native-like structures. The study of smaller biological peptides such as gramicidin S highlighted the competition between charge solvation by water and internal 'self-solvation' by hydrogen bonds and their influence on gas-phase structures[267]. Finally, the IM-MS study of protonated basic amino acids (arginine and lysine) highlighted the effect of hydration for gas-phase ions in the competition between salt bridges and charge solvation, allowing the formation of otherwise disfavored salt bridges[264]. Such experimental approaches have contributed to a better understanding of the ESI mechanisms and especially its ability to generate gas-phase ions where solution structural features are preserved or disrupted[265]. The main result of these various approaches is that upon transfer to the gas phase, the potential structural changes of a particular ion are essentially dictated by the relative roles of electrostatic and hydrophobic interactions within the native structure.

Although the link between gas-phase and solution or solid-phase conformations is difficult to address, CCS measurements based on IM-MS of biologically relevant species, such as proteins and protein assemblies, are of great interest since they enable the investigation of conformation dynamics and folding/unfolding equilibria that are otherwise not easily accessible by solid- or solution-phase strategies. As a result, IM-MS plays an important role in unravelling structural features when the appropriate precautions related to above-mentioned concerns are addressed.

Indeed, a partial or total unfolding has a direct impact on the ion mobility. In the past, this conformational change has already been assessed by the observation of higher charge states after the ESI process due to a higher number of ionizable sites exposed[268]. ESI of proteins is known to generate several charge states. It is generally believed that a lower charge state is more representative of the native form because compact structures tends to expose a reduced number of protonation sites. Consequently, IM-MS analyses are generally performed on ions of reduced charge states as they are believed to be more representative of the biologically relevant, native folded species. However, extreme care should be taken when dealing with such assumptions since the contribution of hydrophobic, ionic and hydrogen bond interactions governing the 3D structure can be impacted by the desolvation process in a very specific way. As a result, a reliable relationship between compactness and low charge state cannot be readily derived[268].

Conformational changes induced by ligand binding[248] or post-translational modifications can also easily assessed by IM-MS, without requiring the precise evaluation of CCS values (prior detailed structural determination). Finally, the exploration of protein dynamics due to conformational changes induced by temperature was also possible[269].

The comparison between IM-MS derived CCS for lower charge states and data generated by NMR or X-ray crystallography generally shows a good agreement even though the discrepancy between solid-phase crystal structures and IM-MS structures tends to increase with the size of the protein[239]. This effect could be explained by the additional space occupied by the solvent molecules inside the crystal structure[270], coherent with a partial collapse of the gas-phase structure (more compact conformation) due to the desolvation process.

2.4.2.6.3. Nucleic acids

The IM-MS analysis of these large biomolecules was also performed since they often retain their solution-phase structure in the gas phase for (up to) milliseconds, provided that low charge states are studied[271]. In contrast, at higher charge states, these structures elongate as Watson–Crick pairs are progressively disrupted. Various secondary structures were successfully characterized by IM-MS in combination to MD simulations including hairpins, pseudoknots and cruciforms[272] as shown in Figure 2.30:

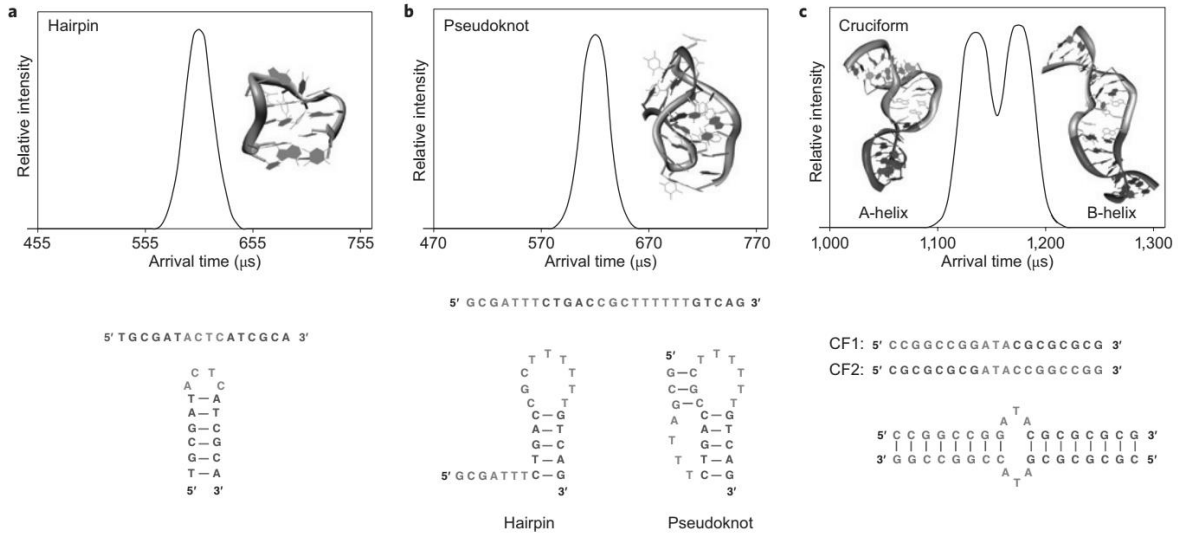


Figure 2.30 - Various nucleic acid secondary structures characterized by IM-MS from Lanucara *et al.* [207]

However, similar to proteins, the desolvation of such analytes can also lead to structural changes. This effect was shown on cruciforms that were known to adopt a β-helix structure in solution while IM-MS measurements revealed the existence of both a β-helix and an α-helix when desolvated[272].

Also, the study of G-quadruplex formation is of particular importance in understanding the mechanism of cellular ageing. The solution characterization of such structures revealed the formation of stacked planar rings composed of four guanosine residues stabilized by atypical Hoogsteen hydrogen bonds where cations are interlocked between the planes[273]. The IM-MS characterization of these structures showed the retention of conformation upon desolvation[274]. Furthermore, IM-MS allowed the characterization of the intermediates in the assembly of these G-quadruplexes. Using this strategy, it could be demonstrated that the final kinetically stable tetramers form an equilibrium state of DNA monomers, dimers and trimers, upon addition of ammonium cations[275].

2.4.2.7. Comparison of CZE with IMS for the assessment of gas phase and solution structures

As demonstrated in the above sections, CZE and IMS are both mobility-based techniques and the only noticeable difference lies in the medium in which the migration takes place; solution for CZE and gas phase for IMS. As a result, the respective mobilities in solution and in the gas-phase are governed by similar separation parameters, as demonstrated in the equations below:

$$\text{IMS}: \ K = \frac{3}{16N}\left(\frac{2\pi}{\mu\, k_B T}\right)^{1/2}\frac{q}{\Omega}$$

$$\text{CZE}: \ \mu_e = \frac{q}{6\pi.\eta.R_h} = \frac{1}{6\pi.\eta}\,\frac{q}{R_h}$$

By comparing these 2 equations, it is clear that both separation techniques are governed by the charge of the analyte (q) and a structural parameter, which is the collision cross section **Ω** for IMS and the hydrodynamic radius R_h for CZE. It should be mentioned that the charge of the analyte "q" is always an integer value in IMS since it corresponds to the charge state produced by the electrospray ion source (m/z) while "q" in CZE refers to the average charge state of the analyte in solution (mainly dictated by the pH of the solution and the pK_a values of the ionizable functional groups) if the interconversion rates between the different charge states is sufficient.

IMS is a powerful structural characterization technique in the gas phase. Fitted with the appropriate ion source, the ability to generate gas-phase ions of biologically relevant molecules from near-physiological conditions (i.e. relatively high ionic strength and buffered pH) enables the structural assessment of these species. This assessment is of particular interest if the derived structural features can be linked to the biological activity in solution, as characteristics of the solution conformation may sometimes be preserved in the gas-phase[276,277].

Consequently, the retention of intact structural features of the native form in solution is a serious concern about structural information deduced from IMS measurements. In this regard, a direct comparison of the results obtained in the gas phase by IMS with data obtained for the same system in solution by CZE could improve our understanding of the link between gas-phase and solution structures and the role played by solvent molecules, counter-ions and other chemical entities on structural features in solution. In particular, the ability of CZE to perform separations in "native" or "near-native" conditions could be compared to IMS results obtained by spraying the same species of interest from the same solutions.

Table 2.4 summarizes various characteristics of each technique.

Table 2.4 - Comparative table of respective characteristics of IM-MS and CZE-MS

Parameter	IM-MS	CZE-MS
Migration medium	Gas phase	Solution (either aqueous or (hydro)organic solvent)
Analysis duration	Milliseconds range	Minutes range
Sample consumption per analysis	µL range	nL range
Accessible charge states	Limited to the charge states generated by the ESI process	Tunable according to the pH of the BGE if the molecule possesses ionizable groups
Structural information	CCS	R_h
Structural elucidation	Structural elucidation based on MS (/MS) CID or ETD)	Structural elucidation based on MS and MS^n
Computational chemistry effort	Straightforward since performed in vacuo	Difficulty to properly describe the solvation and counterions effects
Hyphenation	Commercially available instrument	Coupling between both techniques using a suitable CE-MS interface
Resolution	Low (50-100 CCS/ΔCCS)	(very) High (> 10^5 theoretical plates)
Separable and detectable species	Kinetically trapped species detectable	Species at chemical equilibrium

2.5. References

(1) REUSS; F., F. Sur Un Nouvel Effet de l'electricite Galvanique. *Mem. Soc. Imp. Natur. Moscou* **1809**, *2*, 327–337.

(2) TISELIUS; A. The Moving Boundary Method of Studying the Electrophoresis of Proteins. *Nov. Acta Regiae Soc. Sci. Uppsala, Ser. 4* **1930**, *7* (4), 1–107.

(3) Tiselius, A. A New Apparatus for Electrophoretic Analysis of Colloidal Mixtures. *Trans. Faraday Soc.* **1937**, *33* (0), 524–531. https://doi.org/10.1039/TF9373300524.

(4) Egli, M.; Flavell, A.; Allen, S.; Fisher, J.; Haq, S. I.; Engels, J. W.; Grasby, J. A.; Pyle, A. M.; Luisi, B.; Laughton, C. *Nucleic Acids in Chemistry and Biology*; Royal Society of Chemistry, 2006.

(5) Southern, E. M. Detection of Specific Sequences among DNA Fragments Separated by Gel Electrophoresis. *J. Mol. Biol.* **1975**, *98* (3), 503–517.

(6) Chrambach, A.; Reisfeld, R. A.; Wyckoff, M.; Zaccari, J. A Procedure for Rapid and Sensitive Staining of Protein Fractionated by Polyacrylamide Gel Electrophoresis. *Anal. Biochem.* **1967**, *20* (1), 150–154.

(7) Neville, D. M. Molecular Weight Determination of Protein-Dodecyl Sulfate Complexes by Gel Electrophoresis in a Discontinuous Buffer System. *J. Biol. Chem.* **1971**, *246* (20), 6328–6334.

(8) Hjertén, S. Free Zone Electrophoresis. *Chromatogr. Rev.* **1967**, *9* (2), 122–219. https://doi.org/10.1016/0009-5907(67)80003-6.

(9) Zhu, Z.; Lu, J. J.; Liu, S. Protein Separation by Capillary Gel Electrophoresis: A Review. *Anal. Chim. Acta* **2012**, *709*, 21–31. https://doi.org/10.1016/J.ACA.2011.10.022.

(10) Hunt, G.; Nashabeh, W. Capillary Electrophoresis Sodium Dodecyl Sulfate Nongel Sieving Analysis of a Therapeutic Recombinant Monoclonal Antibody: A Biotechnology Perspective. *Anal. Chem.* **1999**, *71* (13), 2390–2397. https://doi.org/10.1021/ac981209m.

(11) Krull, I. S.; Liu, X.; Dai, J.; Gendreau, C.; Li, G. HPCE Methods for the Identification and Quantitation of Antibodies, Their Conjugates and Complexes. *J. Pharm. Biomed. Anal.* **1997**, *16* (3), 377–393. https://doi.org/10.1016/S0731-7085(97)00071-X.

(12) Lee, H. G.; Chang, S.; Fritsche, E. Rational Approach to Quantitative Sodium Dodecyl Sulfate Capillary Gel Electrophoresis of Monoclonal Antibodies. *J. Chromatogr. A* **2002**, *947* (1), 143–149. https://doi.org/10.1016/S0021-9673(01)01590-4.

(13) Tous, G. I.; Wei, Z.; Feng, J.; Bilbulian, S.; Bowen, S.; Smith, J.; Strouse, R.; McGeehan, P.; Casas-Finet, J.; Schenerman, M. A. Characterization of a Novel Modification to Monoclonal Antibodies: Thioether Cross-Link of Heavy and Light Chains. *Anal. Chem.* **2005**, *77* (9), 2675–2682. https://doi.org/10.1021/ac0500582.

(14) Lu, G.; Crihfield, C. L.; Gattu, S.; Veltri, L. M.; Holland, L. A. Capillary Electrophoresis Separations of Glycans. *Chem. Rev.* **2018**, *118* (17), 7867–7885. https://doi.org/10.1021/acs.chemrev.7b00669.

(15) Mittermayr, S.; Guttman, A. Influence of Molecular Configuration and Conformation on the Electromigration of Oligosaccharides in Narrow Bore Capillaries. *Electrophoresis* **2012**, *33* (6), 1000–1007.

(16) Butler, J. M. *Forensic DNA Typing: Biology, Technology, and Genetics of STR Markers*; Elsevier, 2005.

(17) Butler, J. M. *Fundamentals of Forensic DNA Typing*; Academic press, 2009.

(18) Kremser, L.; Blaas, D.; Kenndler, E. Capillary Electrophoresis of Biological Particles: Viruses, Bacteria, and Eukaryotic Cells. *Electrophoresis* **2004**, *25* (14), 2282–2291.

(19) Lancioni, C.; Keunchkarian, S.; Castells, C. B.; Gagliardi, L. G. Determination of Thermodynamic Binding Constants by Affinity Capillary Electrophoresis. *Talanta* **2019**, *192*, 448–454.

(20) Opekar, F.; Coufal, P.; Štulík, K. Rapid Capillary Zone Electrophoresis along Short Separation Pathways and Its Use in Some Hyphenated Systems: A Critical Review. *Chem. Rev.* **2009**, *109* (9), 4487–4499.

(21) Tůma, P. Rapid and Sensitive Determination of Branched-Chain Amino Acids in Human Plasma by Capillary Electrophoresis with Contactless Conductivity Detection for Physiological Studies. In *Clinical Applications of Capillary Electrophoresis*; Springer, 2019; pp 15–24.

(22) Štěpánová, S.; Kašička, V. Recent Developments and Applications of Capillary and Microchip Electrophoresis in Proteomics and Peptidomics (2015–Mid 2018). *J. Sep. Sci.* **2019**, *42* (1), 398–414.

(23) Fanali, S.; Chankvetadze, B. Some Thoughts about Enantioseparations in Capillary Electrophoresis. *Electrophoresis* **2019**.

(24) He, Y.; Lacher, N. A.; Hou, W.; Wang, Q.; Isele, C.; Starkey, J.; Ruesch, M. Analysis of Identity, Charge Variants, and Disulfide Isomers of Monoclonal Antibodies with Capillary Zone Electrophoresis in an Uncoated Capillary Column. **2010**, *82* (8), 3222–3230. https://doi.org/10.1021/ac9028856.

(25) He, Y.; Isele, C.; Hou, W.; Ruesch, M. Rapid Analysis of Charge Variants of Monoclonal Antibodies with Capillary Zone Electrophoresis in Dynamically Coated Fused-silica Capillary. *J. Sep. Sci.* **2011**, *34* (5), 548–555.

(26) Rochu, D.; Masson, P. Multiple Advantages of Capillary Zone Electrophoresis for Exploring Protein Conformational Stability. *Electrophoresis* **2002**, *23* (2), 189–202.

(27) Righetti, P. G.; Verzola, B. Folding/Unfolding/Refolding of Proteins: Present Methodologies in Comparison with Capillary Zone Electrophoresis. *Electrophoresis* **2001**, *22* (12), 2359–2374.

(28) Lukacs, K. D.; Jorgenson, J. W. Capillary Zone Electrophoresis: Effect of Physical Parameters on Separation Efficiency and Quantitation. *J. High Resolut. Chromatogr.* **1985**, *8* (8), 407–411.
(29) Jorgenson, J. W.; Lukacs, K. D. Zone Electrophoresis in Open-Tubular Glass Capillaries. *Anal. Chem.* **1981**, *53* (8), 1298–1302.
(30) Horvath, J.; Dolník, V. Polymer Wall Coatings for Capillary Electrophoresis. *Electrophoresis* **2001**, *22* (4), 644–655.
(31) Lucy, C. A.; MacDonald, A. M.; Gulcev, M. D. Non-Covalent Capillary Coatings for Protein Separations in Capillary Electrophoresis. *J. Chromatogr. A* **2008**, *1184* (1–2), 81–105. https://doi.org/10.1016/j.chroma.2007.10.114.
(32) Gstaiger, M.; Aebersold, R. Applying Mass Spectrometry-Based Proteomics to Genetics, Genomics and Network Biology. *Nat. Rev. Genet.* **2009**, *10* (9), 617.
(33) Aebersold, R.; Mann, M. Mass Spectrometry-Based Proteomics. *Nature* **2003**, *422* (6928), 198.
(34) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass Spectrometry-based Metabolomics. *Mass Spectrom. Rev.* **2007**, *26* (1), 51–78.
(35) Rossi, D. T.; Sinz, M. *Mass Spectrometry in Drug Discovery*; CRC Press, 2001.
(36) Eberlin, L. S.; Norton, I.; Orringer, D.; Dunn, I. F.; Liu, X.; Ide, J. L.; Jarmusch, A. K.; Ligon, K. L.; Jolesz, F. A.; Golby, A. J. Ambient Mass Spectrometry for the Intraoperative Molecular Diagnosis of Human Brain Tumors. *Proc. Natl. Acad. Sci.* **2013**, *110* (5), 1611–1616.
(37) Eberlin, M. N. Electrospray Ionization Mass Spectrometry: A Major Tool to Investigate Reaction Mechanisms in Both Solution and the Gas Phase. *Eur. J. Mass Spectrom.* **2007**, *13* (1), 19–28.
(38) Biemann, K.; Scoble, H. A. Characterization by Tandem Mass Spectrometry of Structural Modifications in Proteins. *Science (80-.).* **1987**, *237* (4818), 992–998.
(39) West, G. M.; Tang, L.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification-and Mass Spectrometry-Based Strategy. *Anal. Chem.* **2008**, *80* (11), 4175–4185.
(40) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; Pauw, E. De. Triplex and Quadruplex DNA Structures Studied by Electrospray Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2002**, *16* (18), 1729–1736.
(41) Yamashita, M.; Fenn, J. B. Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* **1984**, *88* (20), 4451–4459.
(42) Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. Molecular Beams of Macroions. *J. Chem. Phys.* **1968**, *49* (5), 2240–2249.
(43) Nohmi, T.; Fenn, J. B. Electrospray Mass Spectrometry of Poly (Ethylene Glycols) with Molecular Weights up to Five Million. *J. Am. Chem. Soc.* **1992**, *114* (9), 3241–3246.
(44) Hawkridge, A. M. Practical Considerations and Current Limitations in Quantitative Mass Spectrometry-Based Proteomics. *Quant. proteomics* **2014**, 1–25.
(45) Gomez, A.; Tang, K. Charge and Fission of Droplets in Electrostatic Sprays. *Phys. Fluids* **1994**, *6* (1), 404–414.
(46) 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.
(47) 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.
(48) Iribarne, J. V; Thomson, B. A. On the Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, *64* (6), 2287–2294.
(49) Thomson, B. A.; Iribarne, J. V. Field Induced Ion Evaporation from Liquid Surfaces at Atmospheric Pressure. *J. Chem. Phys.* **1979**, *71* (11), 4451–4463.
(50) Iribarne, J. V; Dziedzic, P. J.; Thomson, B. A. Atmospheric Pressure Ion Evaporation-Mass Spectrometry. *Int. J. mass Spectrom. ion Phys.* **1983**, *50* (3), 331–347.
(51) Schmelzeisen-Redeker, G.; Bütfering, L.; Röllgen, F. W. Desolvation of Ions and Molecules in Thermospray Mass Spectrometry. *Int. J. Mass Spectrom. Ion Process.* **1989**, *90* (2), 139–150.
(52) Kebarle, P.; Tang, L. From Ions in Solution to Ions in the Gas Phase-the Mechanism of Electrospray Mass Spectrometry. *Anal. Chem.* **1993**, *65* (22), 972A-986A.
(53) Wang, G.; Cole, R. B. Charged Residue versus Ion Evaporation for Formation of Alkali Metal Halide Cluster Ions in ESI. *Anal. Chim. Acta* **2000**, *406* (1), 53–65. https://doi.org/http://dx.doi.org/10.1016/S0003-2670(99)00599-1.
(54) Gamero-Castaño, M.; Mora, J. F. de la. Kinetics of Small Ion Evaporation from the Charge and Mass Distribution of Multiply Charged Clusters in Electrosprays. *J. Mass Spectrom.* **2000**, *35* (7), 790–803.
(55) Cole, R. B. Some Tenets Pertaining to Electrospray Ionization Mass Spectrometry. *J. Mass Spectrom.* **2000**, *35* (7), 763–772.
(56) Gamero-Castano, M.; de la Mora, J. F. Mechanisms of Electrospray Ionization of Singly and Multiply Charged Salt Clusters. *Anal. Chim. Acta* **2000**, *406* (1), 67–91.
(57) Hogan Jr, C. J.; Carroll, J. A.; Rohrs, H. W.; Biswas, P.; Gross, M. L. Combined Charged Residue-Field Emission Model of Macromolecular Electrospray Ionization. *Anal. Chem.* **2008**, *81* (1), 369–377.

(58) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. ACS Publications 2012.
(59) Wilm, M. S.; Mann, M. Electrospray and Taylor-Cone Theory, Dole's Beam of Macromolecules at Last? *Int. J. Mass Spectrom. Ion Process.* **1994**, *136* (2–3), 167–180.
(60) Wilm, M.; Mann, M. Analytical Properties of the Nanoelectrospray Ion Source. *Anal. Chem.* **1996**, *68* (1), 1–8.
(61) Juraschek, R.; Dülcks, T.; Karas, M. Nanoelectrospray—More than Just a Minimized-Flow Electrospray Ionization Source. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (4), 300–308. https://doi.org/http://dx.doi.org/10.1016/S1044-0305(98)00157-3.
(62) Heck, A. J. R. Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology. *Nat. Methods* **2008**, *5* (11), 927.
(63) Fenn, J. B. Ion Formation from Charged Droplets: Roles of Geometry, Energy, and Time. *J. Am. Soc. Mass Spectrom.* **1993**, *4* (7), 524–535.
(64) El-Faramawy, A.; Siu, K. W. M.; Thomson, B. A. Efficiency of Nano-Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **2005**, *16* (10), 1702–1707.
(65) Ruff, M.; Mueller, M. S.; Loos, M.; Singer, H. P. Quantitative Target and Systematic Non-Target Analysis of Polar Organic Micro-Pollutants along the River Rhine Using High-Resolution Mass-Spectrometry–Identification of Unknown Sources and Compounds. *Water Res.* **2015**, *87*, 145–154.
(66) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. ACS Publications 2014.
(67) Hoffman, E. De; Stroobant, V. Mass Spectrometry: Principles and Applications. *West Sussex John Wiley Sons, Bruxellas, Bélgica* **2007**, *1* (2), 85.
(68) Shi, S. D.-H.; Drader, J. J.; Hendrickson, C. L.; Marshall, A. G. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry in a High Homogeneity 25 Tesla Resistive Magnet. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (3), 265–268.
(69) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. Fourier Transform Ion Cyclotron Resonance Mass Spectrometry: A Primer. *Mass Spectrom. Rev.* **1998**, *17* (1), 1–35.
(70) Laskin, J.; Futrell, J. H. New Approach for Studying Slow Fragmentation Kinetics in FT-ICR: Surface-Induced Dissociation Combined with Resonant Ejection. *Int. J. Mass Spectrom.* **2015**, *378*, 160–168.
(71) Dearden, D. V; Liang, Y.; Nicoll, J. B.; Kellersberger, K. A. Study of Gas-phase Molecular Recognition Using Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR/MS). *J. mass Spectrom.* **2001**, *36* (9), 989–997.
(72) Nibbering, N. M. M. Gas-Phase Ion/Molecule Reactions as Studied by Fourier Transform Ion Cyclotron Resonance. *Acc. Chem. Res.* **1990**, *23* (9), 279–285.
(73) Makarov, A. A. Mass Spectrometer. Google Patents March 23, 1999.
(74) Makarov, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Anal. Chem.* **2000**, *72* (6), 1156–1162.
(75) Zubarev, R. A.; Makarov, A. Orbitrap Mass Spectrometry. ACS Publications 2013.
(76) Savaryn, J. P.; Toby, T. K.; Kelleher, N. L. A Researcher's Guide to Mass Spectrometry-based Proteomics. *Proteomics* **2016**, *16* (18), 2435–2443.
(77) Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S. Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *Anal. Chem.* **2006**, *78* (7), 2113–2120.
(78) Olsen, J. V; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-Energy C-Trap Dissociation for Peptide Modification Analysis. *Nat. Methods* **2007**, *4* (9), 709.
(79) McAlister, G. C.; Berggren, W. T.; Griep-Raming, J.; Horning, S.; Makarov, A.; Phanstiel, D.; Stafford, G.; Swaney, D. L.; Syka, J. E. P.; Zabrouskov, V. A Proteomics Grade Electron Transfer Dissociation-Enabled Hybrid Linear Ion Trap-Orbitrap Mass Spectrometer. *J. Proteome Res.* **2008**, *7* (8), 3127–3136.
(80) Michalski, A.; Damoc, E.; Lange, O.; Denisov, E.; Nolting, D.; Müller, M.; Viner, R.; Schwartz, J.; Remes, P.; Belford, M. Ultra High Resolution Linear Ion Trap Orbitrap Mass Spectrometer (Orbitrap Elite) Facilitates Top down LC MS/MS and Versatile Peptide Fragmentation Modes. *Mol. Cell. Proteomics* **2012**, *11* (3), O111-013698.
(81) Michalski, A.; Damoc, E.; Hauschild, J.-P.; Lange, O.; Wieghaus, A.; Makarov, A.; Nagaraj, N.; Cox, J.; Mann, M.; Horning, S. Mass Spectrometry-Based Proteomics Using Q Exactive, a High-Performance Benchtop Quadrupole Orbitrap Mass Spectrometer. *Mol. Cell. Proteomics* **2011**, *10* (9), M111-011015.
(82) Hopfgartner, G.; Varesio, E.; Tschäppät, V.; Grivet, C.; Bourgogne, E.; Leuthold, L. A. Triple Quadrupole Linear Ion Trap Mass Spectrometer for the Analysis of Small Molecules and Macromolecules. *J. Mass Spectrom.* **2004**, *39* (8), 845–855.
(83) Domon, B.; Aebersold, R. Mass Spectrometry and Protein Analysis. *Science (80-.).* **2006**, *312* (5771), 212–217.
(84) Henry, H.; Sobhi, H. R.; Scheibner, O.; Bromirski, M.; Nimkar, S. B.; Rochat, B. Comparison between a High-resolution Single-stage Orbitrap and a Triple Quadrupole Mass Spectrometer for Quantitative Analyses of Drugs. *Rapid Commun. Mass Spectrom.* **2012**, *26* (5), 499–509.
(85) Gallien, S.; Duriez, E.; Crone, C.; Kellmann, M.; Moehring, T.; Domon, B. Targeted Proteomic Quantification on Quadrupole-Orbitrap Mass Spectrometer. *Mol. Cell. proteomics* **2012**, *11* (12), 1709–1723.

(86) Xian, F.; Hendrickson, C. L.; Blakney, G. T.; Beu, S. C.; Marshall, A. G. Automated Broadband Phase Correction of Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal. Chem.* **2010**, *82* (21), 8807–8812.
(87) Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. High-Performance Mass Spectrometry: Fourier Transform Ion Cyclotron Resonance at 14.5 Tesla. *Anal. Chem.* **2008**, *80* (11), 3985–3990.
(88) Rose, R. J.; Damoc, E.; Denisov, E.; Makarov, A.; Heck, A. J. R. High-Sensitivity Orbitrap Mass Analysis of Intact Macromolecular Assemblies. *Nat. Methods* **2012**, *9* (11), 1084.
(89) Makarov, A.; Denisov, E.; Lange, O.; Horning, S. Dynamic Range of Mass Accuracy in LTQ Orbitrap Hybrid Mass Spectrometer. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (7), 977–982.
(90) Hilton, G. C.; Martinis, J. M.; Wollman, D. A.; Irwin, K. D.; Dulcie, L. L.; Gerber, D.; Gillevet, P. M.; Twerenbold, D. Impact Energy Measurement in Time-of-Flight Mass Spectrometry with Cryogenic Microcalorimeters. *Nature* **1998**, *391* (6668), 672.
(91) Cheng, Y.-F.; Dovichi, N. J. Subattomole Amino Acid Analysis by Capillary Zone Electrophoresis and Laser-Induced Fluorescence. *Science (80-.).* **1988**, *242* (4878), 562–564.
(92) Liu, J. P.; Shirota, O.; Wiesler, D.; Novotny, M. Ultrasensitive Fluorometric Detection of Carbohydrates as Derivatives in Mixtures Separated by Capillary Electrophoresis. *Proc. Natl. Acad. Sci.* **1991**, *88* (6), 2302–2306.
(93) Huang, X. C.; Quesada, M. A.; Mathies, R. A. DNA Sequencing Using Capillary Array Electrophoresis. *Anal. Chem.* **1992**, *64* (18), 2149–2154.
(94) Ganzler, K.; Greve, K. S.; Cohen, A. S.; Karger, B. L.; Guttman, A.; Cooke, N. C. High-Performance Capillary Electrophoresis of SDS-Protein Complexes Using UV-Transparent Polymer Networks. *Anal. Chem.* **1992**, *64* (22), 2665–2671.
(95) Kolch, W.; Neusüß, C.; Pelzing, M.; Mischak, H. Capillary Electrophoresis–Mass Spectrometry as a Powerful Tool in Clinical Diagnosis and Biomarker Discovery. *Mass Spectrom. Rev.* **2005**, *24* (6), 959–977.
(96) Mischak, H.; Coon, J. J.; Novak, J.; Weissinger, E. M.; Schanstra, J. P.; Dominiczak, A. F. Capillary Electrophoresis–Mass Spectrometry as a Powerful Tool in Biomarker Discovery and Clinical Diagnosis: An Update of Recent Developments. *Mass Spectrom. Rev.* **2009**, *28* (5), 703–724.
(97) Klepárník, K. Recent Advances in the Combination of Capillary Electrophoresis with Mass Spectrometry: From Element to Single-Cell Analysis. *Electrophoresis* **2013**, *34* (1), 70–85. https://doi.org/10.1002/elps.201200488.
(98) Zhang, J.; Xie, J.; Chen, X.; Hu, Z. Sensitive Determination of Ephedrine and Pseudoephedrine by Capillary Electrophoresis with Laser-Induced Fluorescence Detection. *Analyst* **2003**, *128* (4), 369–372.
(99) Lee, R.; Ptolemy, A. S.; Niewczas, L.; Britz-McKibbin, P. Integrative Metabolomics for Characterizing Unknown Low-Abundance Metabolites by Capillary Electrophoresis-Mass Spectrometry with Computer Simulations. *Anal. Chem.* **2007**, *79* (2), 403–415.
(100) Hirayama, A.; Soga, T. Amino Acid Analysis by Capillary Electrophoresis-Mass Spectrometry. In *Amino acid analysis*; Springer, 2012; pp 77–82.
(101) Simionato, A. V. C.; Moraes, E. P.; Carrilho, E.; Tavares, M. F. M.; Kenndler, E. Determination of Amino Acids by Capillary Electrophoresis-electrospray Ionization-mass Spectrometry: An Evaluation of Different Protein Hydrolysis Procedures. *Electrophoresis* **2008**, *29* (10), 2051–2058.
(102) Campa, C.; Coslovi, A.; Flamigni, A.; Rossi, M. Overview on Advances in Capillary Electrophoresis-mass Spectrometry of Carbohydrates: A Tabulated Review. *Electrophoresis* **2006**, *27* (11), 2027–2050.
(103) Haselberg, R.; de Jong, G. J.; Somsen, G. W. Capillary Electrophoresis–Mass Spectrometry for the Analysis of Intact Proteins. *J. Chromatogr. a* **2007**, *1159* (1–2), 81–109.
(104) Ohnesorge, J.; Neusüß, C.; Wätzig, H. Quantitation in Capillary Electrophoresis-mass Spectrometry. *Electrophoresis* **2005**, *26* (21), 3973–3987.
(105) Deforce, D. L. D.; Raymackers, J.; Meheus, L.; Van Wijnendaele, F.; De Leenheer, A.; Van den Eeckhout, E. G. Characterization of DNA Oligonucleotides by Coupling of Capillary Zone Electrophoresis to Electrospray Ionization Q-TOF Mass Spectrometry. *Anal. Chem.* **1998**, *70* (14), 3060–3068.
(106) Amenson-Lamar, E. A.; Sun, L.; Zhang, Z.; Bohn, P. W.; Dovichi, N. J. Detection of 1 zmol Injection of Angiotensin Using Capillary Zone Electrophoresis Coupled to a Q-Exactive HF Mass Spectrometer with an Electrokinetically Pumped Sheath-Flow Electrospray Interface. *Talanta* **2019**, *204*, 70–73. https://doi.org/10.1016/J.TALANTA.2019.05.079.
(107) Zürbig, P.; Renfrow, M. B.; Schiffer, E.; Novak, J.; Walden, M.; Wittke, S.; Just, I.; Pelzing, M.; Neusüß, C.; Theodorescu, D. Biomarker Discovery by CE-MS Enables Sequence Analysis via MS/MS with Platform-independent Separation. *Electrophoresis* **2006**, *27* (11), 2111–2125.
(108) Von Brocke, A.; Nicholson, G.; Bayer, E. Recent Advances in Capillary Electrophoresis/Electrospray-mass Spectrometry. *Electrophoresis* **2001**, *22* (7), 1251–1266.
(109) Cole, R. Electrospray Ionization Mass Spectrometry. *Fundam. Instrum. Appl.* **1997**.
(110) Banks, J. F. Recent Advances in Capillary Electrophoresis/Electrospray/Mass Spectrometry. *Electrophoresis* **1997**, *18* (12-13), 2255–2266.

(111) Cai, J.; Henion, J. Capillary Electrophoresis-Mass Spectrometry. *J. Chromatogr. A* **1995**, *703* (1–2), 667–692. https://doi.org/10.1016/0021-9673(94)01178-H.
(112) Gelpí, E. Interfaces for Coupled Liquid-phase Separation/Mass Spectrometry Techniques. An Update on Recent Developments. *J. mass Spectrom.* **2002**, *37* (3), 241–253.
(113) Moini, M. Capillary Electrophoresis Mass Spectrometry and Its Application to the Analysis of Biological Mixtures. *Anal. Bioanal. Chem.* **2002**, *373* (6), 466–480.
(114) Tomer, K. B. Separations Combined with Mass Spectrometry. *Chem. Rev.* **2001**, *101* (2), 297–328.
(115) Kannamkumarath, S. S.; Wrobel, K.; Wrobel, K.; B'Hymer, C.; Caruso, J. A. Capillary Electrophoresis–Inductively Coupled Plasma-Mass Spectrometry: An Attractive Complementary Technique for Elemental Speciation Analysis. *J. Chromatogr. A* **2002**, *975* (2), 245–266.
(116) Olesik, J. W.; Kinzer, J. A.; Olesik, S. V. Capillary Electrophoresis Inductively Coupled Plasma Spectrometry for Rapid Elemental Speciation. *Anal. Chem.* **1995**, *67* (1), 1–12.
(117) Hommerson, P.; Khan, A. M.; de Jong, G. J.; Somsen, G. W. Capillary Electrophoresis-Atmospheric Pressure Chemical Ionization-Mass Spectrometry Using an Orthogonal Interface: Set-up and System Parameters. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (7), 1311–1318.
(118) Mol, R.; de Jong, G. J.; Somsen, G. W. On-line Capillary Electrophoresis-mass Spectrometry Using Dopant-assisted Atmospheric Pressure Photoionization: Setup and System Performance. *Electrophoresis* **2005**, *26* (1), 146–154.
(119) Nilsson, S. L.; Andersson, C.; Sjöberg, P. J. R.; Bylund, D.; Petersson, P.; Jörntén-Karlsson, M.; Markides, K. E. Phosphate Buffers in Capillary Electrophoresis/Mass Spectrometry Using Atmospheric Pressure Photoionization and Electrospray Ionization. *Rapid Commun. mass Spectrom.* **2003**, *17* (20), 2267–2272.
(120) Bonvin, G.; Schappler, J.; Rudaz, S. Non-Aqueous Capillary Electrophoresis for the Analysis of Acidic Compounds Using Negative Electrospray Ionization Mass Spectrometry. *J. Chromatogr. A* **2014**, *1323*, 163–173.
(121) Klampfl, C. W. Review Coupling of Capillary Electrochromatography to Mass Spectrometry. *J. Chromatogr. A* **2004**, *1044* (1–2), 131–144.
(122) Neusüß, C.; Pelzing, M.; Macht, M. A Robust Approach for the Analysis of Peptides in the Low Femtomole Range by Capillary Electrophoresis-tandem Mass Spectrometry. *Electrophoresis* **2002**, *23* (18), 3149–3159.
(123) Schappler, J.; Veuthey, J.-L.; Rudaz, S. 18 Coupling CE and Microchip-Based Devices with Mass Spectrometry. *Sep. Sci. Technol.* **2008**, *9*, 477–521.
(124) Nashabeh, W.; Greve, K. F.; Kirby, D.; Foret, F.; Karger, B. L.; Reifsnyder, D. H.; Builder, S. E. Incorporation of Hydrophobic Selectivity in Capillary Electrophoresis: Analysis of Recombinant Insulin-like Growth Factor I Variants. *Anal. Chem.* **1994**, *66* (13), 2148–2154.
(125) Busnel, J. M.; Schoenmaker, B.; Ramautar, R.; Carrasco-Pancorbo, A.; Ratnayake, C.; Feitelson, J. S.; Chapman, J. D.; Deelder, A. M.; Mayboroda, O. A. High Capacity Capillary Electrophoresis-Electrospray Ionization Mass Spectrometry: Coupling a Porous Sheathless Interface with Transient- Isotachophoresis. *Anal. Chem.* **2010**, *82* (22), 9476–9483. https://doi.org/10.1021/ac102159d.
(126) Covey, T. R.; Thomson, B. A.; Schneider, B. B. Atmospheric Pressure Ion Sources. *Mass Spectrom. Rev.* **2009**, *28* (6), 870–897.
(127) Kostiainen, R.; Bruins, A. P. Effect of Multiple Sprayers on Dynamic Range and Flow Rate Limitations in Electrospray and Ionspray Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **1994**, *8* (7), 549–558.
(128) Smith, R. D.; Olivares, J. A.; Nguyen, N. T.; Udseth, H. R. Capillary Zone Electrophoresis-Mass Spectrometry Using an Electrospray Ionization Interface. *Anal. Chem.* **1988**, *60* (5), 436–441.
(129) Nilsson, S. L.; Bylund, D.; Jörntén-Karlsson, M.; Petersson, P.; Markides, K. E. A Chemometric Study of Active Parameters and Their Interaction Effects in a Nebulized Sheath-liquid Electrospray Interface for Capillary Electrophoresis-mass Spectrometry. *Electrophoresis* **2004**, *25* (13), 2100–2107.
(130) Varesio, E.; Cherkaoui, S.; Veuthey, J. Optimization of CE-ESI-MS Parameters for the Analysis of Ecstasy and Derivatives in Urine. *J. High Resolut. Chromatogr.* **1998**, *21* (12), 653–657.
(131) Schappler, J.; Guillarme, D.; Prat, J.; Veuthey, J.; Rudaz, S. Coupling CE with Atmospheric Pressure Photoionization MS for Pharmaceutical Basic Compounds: Optimization of Operating Parameters. *Electrophoresis* **2007**, *28* (17), 3078–3087.
(132) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. Improved Electrospray Ionization Interface for Capillary Zone Electrophoresis-Mass Spectrometry. *Anal. Chem.* **1988**, *60* (18), 1948–1952.
(133) Axén, J.; Axelsson, B.; Jörntén-Karlsson, M.; Petersson, P.; Sjöberg, P. J. R. An Investigation of Peak-broadening Effects Arising When Combining CE with MS. *Electrophoresis* **2007**, *28* (18), 3207–3213.
(134) Mokaddem, M.; Gareil, P.; Belgaied, J.; Varenne, A. A New Insight into Suction and Dilution Effects in Capillary Electrophoresis Coupled to Mass Spectrometry via an Electrospray Ionization Interface. Part I-Suction Effect. *Electrophoresis* **2008**, *29* (10), 1957–1964.
(135) Herrero, M.; Ibañez, E.; Cifuentes, A. Capillary Electrophoresis-Electrospray-Mass Spectrometry in Peptide Analysis and Peptidomics. *Electrophoresis* **2008**, *29* (10), 2148–2160. https://doi.org/10.1002/elps.200700404.
(136) Font, G.; Ruiz, M. J.; Fernández, M.; Picó, Y. Application of Capillary Electrophoresis-mass Spectrometry

for Determining Organic Food Contaminants and Residues. *Electrophoresis* **2008**, *29* (10), 2059–2078.
(137) Robledo, V. R.; Smyth, W. F. The Application of CE-MS in the Trace Analysis of Environmental Pollutants and Food Contaminants. *Electrophoresis* **2009**, *30* (10), 1647–1660.
(138) Mechref, Y.; Novotny, M. V. Glycomic Analysis by Capillary Electrophoresis–Mass Spectrometry. *Mass Spectrom. Rev.* **2009**, *28* (2), 207–222.
(139) Servais, A.; Crommen, J.; Fillet, M. Capillary Electrophoresis-mass Spectrometry, an Attractive Tool for Drug Bioanalysis and Biomarker Discovery. *Electrophoresis* **2006**, *27* (13), 2616–2629.
(140) Ramautar, R.; Somsen, G. W.; de Jong, G. J. CE-MS in Metabolomics. *Electrophoresis* **2009**, *30* (1), 276–291.
(141) Ramautar, R.; Mayboroda, O. A.; Somsen, G. W.; de Jong, G. J. CE-MS for Metabolomics: Developments and Applications in the Period 2008–2010. *Electrophoresis* **2011**, *32* (1), 52–65.
(142) Lee, E. D.; Mück, W.; Henion, J. D.; Covey, T. R. On-Line Capillary Zone Electrophoresis-Ion Spray Tandem Mass Spectrometry for the Determination of Dynorphins. *J. Chromatogr. A* **1988**, *458*, 313–321.
(143) Lee, E. D.; Mück, W.; Henion, J. D.; Covey, T. R. Liquid Junction Coupling for Capillary Zone Electrophoresis/Ion Spray Mass Spectrometry. *Biomed. Environ. Mass Spectrom.* **1989**, *18* (9), 844–850.
(144) Maxwell, E. J.; Chen, D. D. Y. Twenty Years of Interface Development for Capillary Electrophoresis–Electrospray Ionization–Mass Spectrometry. *Anal. Chim. Acta* **2008**, *627* (1), 25–33.
(145) Severs, J. C.; Harms, A. C.; Smith, R. D. A New High-performance Interface for Capillary Electrophoresis/Electrospray Ionization Mass Spectrometry. *Rapid Commun. mass Spectrom.* **1996**, *10* (10), 1175–1178.
(146) Pleasance, S.; Thibault, P.; Kelly, J. Comparison of Liquid-Junction and Coaxial Interfaces for Capillary Electrophoresis-Mass Spectrometry with Application to Compounds of Concern to the Aquaculture Industry. *J. Chromatogr. A* **1992**, *591* (1–2), 325–339.
(147) Sheppard, R. L.; Henion, J. Quantitative Capillary Electrophoresis/Ion Spray Tandem Mass Spectrometry Determination of EDTA in Human Plasma and Urine. *Anal. Chem.* **1997**, *69* (15), 2901–2907.
(148) Garcia, F.; Henion, J. D. Gel-Filled Capillary Electrophoresis/Mass Spectrometry Using a Liquid Junction Ion Spray Interface. *Anal. Chem.* **1992**, *64* (9), 985–990.
(149) Schmidt, A.; Karas, M.; Dülcks, T. Effect of Different Solution Flow Rates on Analyte Ion Signals in Nano-ESI MS, or: When Does ESI Turn into Nano-ESI? *J. Am. Soc. Mass Spectrom.* **2003**, *14* (5), 492–500.
(150) Olivares, J. A.; Nguyen, N. T.; Yonker, C. R.; Smith, R. D. On-Line Mass Spectrometric Detection for Capillary Zone Electrophoresis. *Anal. Chem.* **1987**, *59* (8), 1230–1232.
(151) Sanz-Nebot, V.; Balaguer, E.; Benavente, F.; Barbosa, J. Comparison of Sheathless and Sheath-flow Electrospray Interfaces for the Capillary Electrophoresis-electrospray Ionization-mass Spectrometry Analysis of Peptides. *Electrophoresis* **2005**, *26* (7-8), 1457–1465.
(152) Bonvin, G.; Schappler, J.; Rudaz, S. Capillary Electrophoresis–Electrospray Ionization-Mass Spectrometry Interfaces: Fundamental Concepts and Technical Developments. *J. Chromatogr. A* **2012**, *1267* (0), 17–31. https://doi.org/10.1016/j.chroma.2012.07.019.
(153) Ramsey, R. S.; McLuckey, S. A. Capillary Electrophoresis/Electrospray Ionization Ion Trap Mass Spectrometry Using a Sheathless Interface. *J. Microcolumn Sep.* **1995**, *7* (5), 461–469.
(154) Trapp, O.; Pearce, E. W.; Kimmel, J. R.; Yoon, O. K.; Zuleta, I. A.; Zare, R. N. A Soft On-column Metal Coating Procedure for Robust Sheathless Electrospray Emitters Used in Capillary Electrophoresis-mass Spectrometry. *Electrophoresis* **2005**, *26* (7-8), 1358–1365.
(155) Chen, Y.; Her, G. A Simple Method for Fabrication of Silver-coated Sheathless Electrospray Emitters. *Rapid Commun. mass Spectrom.* **2003**, *17* (5), 437–441.
(156) Zamfir, A. D.; Dinca, N.; Sisu, E.; Peter-Katalinić, J. Copper-coated Microsprayer Interface for On-line Sheathless Capillary Electrophoresis Electrospray Mass Spectrometry of Carbohydrates. *J. Sep. Sci.* **2006**, *29* (3), 414–422.
(157) Bendahl, L.; Hansen, S. H.; Olsen, J. A New Sheathless Electrospray Interface for Coupling of Capillary Electrophoresis to Ion-trap Mass Spectrometry. *Rapid Commun. mass Spectrom.* **2002**, *16* (24), 2333–2340.
(158) Chang, Y. Z.; Her, G. R. Sheathless Capillary Electrophoresis/Electrospray Mass Spectrometry Using a Carbon-Coated Fused-Silica Capillary. *Anal. Chem.* **2000**, *72* (3), 626–630.
(159) Nilsson, S.; Wetterhall, M.; Bergquist, J.; Nyholm, L.; Markides, K. E. A Simple and Robust Conductive Graphite Coating for Sheathless Electrospray Emitters Used in Capillary Electrophoresis/Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2001**, *15* (21), 1997–2000.
(160) Zhu, X.; Thiam, S.; Valle, B. C.; Warner, I. M. A Colloidal Graphite-Coated Emitter for Sheathless Capillary Electrophoresis/Nanoelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **2002**, *74* (20), 5405–5409.
(161) Bigwarfe Jr, P. M.; White, T. P.; Wood, T. D. Polyaniline-coated Nanoelectrospray Emitters: Performance Characteristics in the Negative Ion Mode. *Rapid Commun. mass Spectrom.* **2002**, *16* (24), 2266–2272.
(162) Maziarz, E. P.; Lorenz, S. A.; White, T. P.; Wood, T. D. Polyaniline: A Conductive Polymer Coating for Durable Nanospray Emitters. *J. Am. Soc. Mass Spectrom.* **2000**, *11* (7), 659–663.
(163) Fang, L.; Zhang, R.; Williams, E. R.; Zare, R. N. Online Time-of-Flight Mass Spectrometric Analysis of Peptides Separated by Capillary Electrophoresis. *Anal. Chem.* **1994**, *66* (21), 3696–3701.

(164) Cao, P.; Moini, M. A Novel Sheathless Interface for Capillary Electrophoresis/Electrospray Ionization Mass Spectrometry Using an in-Capillary Electrode. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (5), 561–564.
(165) Figeys, D.; van Oostveen, I.; Ducret, A.; Aebersold, R. Protein Identification by Capillary Zone Electrophoresis/Microelectrospray Ionization-Tandem Mass Spectrometry at the Subfemtomole Level. *Anal. Chem.* **1996**, *68* (11), 1822–1828.
(166) Figeys, D.; Ducret, A.; Yates, J. R.; Aebersold, R. Protein Identification by Solid Phase Microextraction—Capillary Zone Electrophoresis—Microelectrospray—Tandem Mass Spectrometry. *Nat. Biotechnol.* **1996**, *14* (11), 1579.
(167) Severs, J. C.; Smith, R. D. Characterization of the Microdialysis Junction Interface for Capillary Electrophoresis/Microelectrospray Ionization Mass Spectrometry. *Anal. Chem.* **1997**, *69* (11), 2154–2158.
(168) Viberg, P.; Nilsson, S.; Skog, K. Nanospray Mass Spectrometry with Indirect Conductive Graphite Coating. *Anal. Chem.* **2004**, *76* (14), 4241–4244.
(169) Bateman, K.; L. White, R.; Thibault, P. *Disposable Emitters for On-line Capillary Zone Electrophoresis/Nanoelectrospray Mass Spectrometry*; 1997; Vol. 11. https://doi.org/10.1002/(SICI)1097-0231(19970215)11:3<307::AID-RCM850>3.0.CO;2-M.
(170) Moini, M. Design and Performance of a Universal Sheathless Capillary Electrophoresis to Mass Spectrometry Interface Using a Split-Flow Technique. *Anal. Chem.* **2001**, *73* (14), 3497–3501. https://doi.org/10.1021/ac010189c.
(171) Settlage, R. E.; Russo, P. S.; Shabanowitz, J.; Hunt, D. F. A Novel μ-ESI Source for Coupling Capillary Electrophoresis and Mass Spectrometry: Sequence Determination of Tumor Peptides at the Attomole Level. *J. Microcolumn Sep.* **1998**, *10* (3), 281–285.
(172) Whitt, J. T.; Moini, M. Capillary Electrophoresis to Mass Spectrometry Interface Using a Porous Junction. *Anal. Chem.* **2003**, *75* (9), 2188–2191. https://doi.org/10.1021/ac026380j.
(173) Moini, M. Simplifying CE-MS Operation. 2. Interfacing Low-Flow Separation Techniques to Mass Spectrometry Using a Porous Tip. *Anal. Chem.* **2007**, *79* (11), 4241–4246. https://doi.org/10.1021/ac0704560.
(174) Haselberg, R.; Ratnayake, C. K.; de Jong, G. J.; Somsen, G. W. Performance of a Sheathless Porous Tip Sprayer for Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry of Intact Proteins. *J. Chromatogr. A* **2010**, *1217* (48), 7605–7611. https://doi.org/10.1016/j.chroma.2010.10.006.
(175) Bonvin, G.; Veuthey, J.-L.; Rudaz, S.; Schappler, J. Evaluation of a Sheathless Nanospray Interface Based on a Porous Tip Sprayer for CE-ESI-MS Coupling. *Electrophoresis* **2012**, *33* (4), 552–562. https://doi.org/10.1002/elps.201100461.
(176) Ramautar, R.; Busnel, J.-M.; Deelder, A. M.; Mayboroda, O. A. Enhancing the Coverage of the Urinary Metabolome by Sheathless Capillary Electrophoresis-Mass Spectrometry. *Anal. Chem.* **2012**, *84* (2), 885–892. https://doi.org/10.1021/ac202407v.
(177) Tseng, M.-C.; Chen, Y.-R.; Her, G.-R. A Beveled Tip Sheath Liquid Interface for Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry. *Electrophoresis* **2004**, *25* (13), 2084–2089. https://doi.org/10.1002/elps.200305907.
(178) Hashimoto, M.; Ishihama, Y.; Tomita, M.; Soga, T. Microelectrospray Interface with Coaxial Sheath Flow for High-resolution Capillary Electrophoresis/Mass Spectrometry Separation. *Rapid Commun. Mass Spectrom. An Int. J. Devoted to Rapid Dissem. Up-to-the-Minute Res. Mass Spectrom.* **2007**, *21* (22), 3579–3584.
(179) Tseng, M.-C.; Chen, Y.-R.; Her, G.-R. *A Low-Makeup Beveled Tip Capillary Electrophoresis /Electrospray Ionization Mass Spectrometry Interface for Micellar Electrokinetic Chromatography and Nonvolatile Buffer Capillary Electrophoresis*; 2004; Vol. 76. https://doi.org/10.1021/ac049330i.
(180) Chen, Y.-R.; Tseng, M.-C.; Chang, Y.-Z.; Her, G.-R. A Low-Flow CE/Electrospray Ionization MS Interface for Capillary Zone Electrophoresis, Large-Volume Sample Stacking, and Micellar Electrokinetic Chromatography. *Anal. Chem.* **2003**, *75* (3), 503–508. https://doi.org/10.1021/ac026098c.
(181) Liu, C. C.; Alary, J.-F.; Vollmerhaus, P.; Kadkhodayan, M. Design, Optimisation, and Evaluation of a Sheath Flow Interface for Automated Capillary Electrophoresis-Electrospray-Mass Spectrometry. *Electrophoresis* **2005**, *26* (7-8), 1366–1375. https://doi.org/10.1002/elps.200410133.
(182) Zhu, G.; Sun, L.; Yan, X.; Dovichi, N. J. Stable, Reproducible, and Automated Capillary Zone Electrophoresis–Tandem Mass Spectrometry System with an Electrokinetically Pumped Sheath–Flow Nanospray Interface. *Anal. Chim. Acta* **2014**, *810*, 94–98.
(183) Sun, L.; Zhu, G.; Zhang, Z.; Mou, S.; Dovichi, N. J. Third-Generation Electrokinetically Pumped Sheath-Flow Nanospray Interface with Improved Stability and Sensitivity for Automated Capillary Zone Electrophoresis–Mass Spectrometry Analysis of Complex Proteome Digests. *J. Proteome Res.* **2015**, *14* (5), 2312–2321.
(184) Höcker, O.; Montealegre, C.; Neusüß, C. Characterization of a Nanoflow Sheath Liquid Interface and Comparison to a Sheath Liquid and a Sheathless Porous-Tip Interface for CE-ESI-MS in Positive and Negative Ionization. *Anal. Bioanal. Chem.* **2018**, *410* (21), 5265–5275.
(185) Causon, T. J.; Maringer, L.; Buchberger, W.; Klampfl, C. W. Addition of Reagents to the Sheath Liquid: A

Novel Concept in Capillary Electrophoresis-Mass Spectrometry. *J. Chromatogr. A* **2014**, *1343*, 182–187. https://doi.org/10.1016/J.CHROMA.2014.04.002.
(186) González-Ruiz, V.; Codesido, S.; Far, J.; Rudaz, S.; Schappler, J. Evaluation of a New Low Sheath–Flow Interface for CE-MS. *Electrophoresis* **2016**, *37* (7–8), 936–946.
(187) Pabst, M.; Altmann, F. Glycan Analysis by Modern Instrumental Methods. *Proteomics* **2011**, *11* (4), 631–643.
(188) Mant, C. T.; Hodges, R. S. High Performance Liquid Chromatography of Peptides and Proteins. *Sep. Anal. Conform. CRC Press. Boca Raton, FL* **1991**, 2–4.
(189) Brown, P. R. *HPLC in Nucleic Acid Research: Methods and Applications*; CRC Press, 1984; Vol. 28.
(190) Weston, A.; Brown, P. R. *High Performance Liquid Chromatography & Capillary Electrophoresis: Principles and Practices*; Elsevier, 1997.
(191) Chankvetadze, B. *Capillary Electrophoresis in Chiral Analysis*; John Wiley & Sons, 1997.
(192) Visser, N. F. C.; Lingeman, H.; Irth, H. Sample Preparation for Peptides and Proteins in Biological Matrices Prior to Liquid Chromatography and Capillary Zone Electrophoresis. *Anal. Bioanal. Chem.* **2005**, *382* (3), 535–558.
(193) Wilkins, M. R.; Williams, K. L.; Appel, R. D.; Hochstrasser, D. F. *Proteome Research: New Frontiers in Functional Genomics*; Springer Science & Business Media, 2013.
(194) FDA, U. S. Guidance for Industry: Pharmacogenomic Data Submissions. *Rockv. FDA* **2005**.
(195) Bischoff, R.; Luider, T. M. Methodological Advances in the Discovery of Protein and Peptide Disease Markers. *J. Chromatogr. B* **2004**, *803* (1), 27–40.
(196) Pang, J. X.; Ginanni, N.; Dongre, A. R.; Hefta, S. A.; Opiteck, G. J. Biomarker Discovery in Urine by Proteomics. *J. Proteome Res.* **2002**, *1* (2), 161–169.
(197) Gao, J.; Garulacan, L.-A.; Storm, S. M.; Opiteck, G. J.; Dubaquie, Y.; Hefta, S. A.; Dambach, D. M.; Dongre, A. R. Biomarker Discovery in Biological Fluids. *Methods* **2005**, *35* (3), 291–302.
(198) Shen, Y.; Smith, R. D. Proteomics Based on High-efficiency Capillary Separations. *Electrophoresis* **2002**, *23* (18), 3106–3124.
(199) Dakna, M.; He, Z.; Yu, W. C.; Mischak, H.; Kolch, W. Technical, Bioinformatical and Statistical Aspects of LC-MS and CE-MS Based Clinical Proteomics: A Critical Assessment. *J Chromatogr. B Biomed. Sci. Appl* **2008**, *877*, 1250–1258.
(200) Kaiser, T.; Wittke, S.; Just, I.; Krebs, R.; Bartel, S.; Fliser, D.; Mischak, H.; Weissinger, E. M. Capillary Electrophoresis Coupled to Mass Spectrometer for Automated and Robust Polypeptide Determination in Body Fluids for Clinical Use. *Electrophoresis* **2004**, *25* (13), 2044–2055.
(201) Hjertén, S. High-Performance Electrophoresis : Elimination of Electroendosmosis and Solute Adsorption. *J. Chromatogr. A* **1985**, *347*, 191–198. https://doi.org/10.1016/S0021-9673(01)95485-8.
(202) Klein, J.; Papadopoulos, T.; Mischak, H.; Mullen, W. Comparison of CE-MS/MS and LC-MS/MS Sequencing Demonstrates Significant Complementarity in Natural Peptide Identification in Human Urine. *Electrophoresis* **2014**, *35* (7), 1060–1064. https://doi.org/10.1002/elps.201300327.
(203) Li, Y.; Champion, M. M.; Sun, L.; Champion, P. A. D.; Wojcik, R.; Dovichi, N. J. Capillary Zone Electrophoresis-Electrospray Ionization-Tandem Mass Spectrometry as an Alternative Proteomics Platform to Ultraperformance Liquid Chromatography-Electrospray Ionization-Tandem Mass Spectrometry for Samples of Intermediate Complexity. *Anal. Chem.* **2012**, *84* (3), 1617–1622.
(204) Sun, L.; Zhu, G.; Zhao, Y.; Yan, X.; Mou, S.; Dovichi, N. J. Ultrasensitive and Fast Bottom-up Analysis of Femtogram Amounts of Complex Proteome Digests. *Angew. Chemie Int. Ed.* **2013**, *52* (51), 13661–13664. https://doi.org/10.1002/anie.201308139.
(205) Wei, J.; Yang, L.; Harrata, A. K.; Lee, C. S. High Resolution Analysis of Protein Phosphorylation Using Capillary Isoelectric Focusing-electrospray Ionization-mass Spectrometry. *Electrophoresis* **1998**, *19* (13), 2356–2360.
(206) Mariño, K.; Bones, J.; Kattla, J. J.; Rudd, P. M. A Systematic Approach to Protein Glycosylation Analysis: A Path through the Maze. *Nat. Chem. Biol.* **2010**, *6* (10), 713.
(207) Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E. The Power of Ion Mobility-Mass Spectrometry for Structural Characterization and the Study of Conformational Dynamics. *Nat. Chem.* **2014**, *6* (4), 281.
(208) Mason, E. A.; Schamp, H. W. Mobility of Gaseous Lons in Weak Electric Fields. *Ann. Phys. (N. Y).* **1958**, *4* (3), 233–270. https://doi.org/10.1016/0003-4916(58)90049-6.
(209) Creaser, C. S.; Griffiths, J. R.; Bramwell, C. J.; Noreen, S.; Hill, C. A.; Thomas, C. L. P. Ion Mobility Spectrometry: A Review. Part 1. Structural Analysis by Mobility Measurement. *Analyst* **2004**, *129* (11), 984–994. https://doi.org/10.1039/B404531A.
(210) Giles, K.; Pringle, S.; R Worthington, K.; Little, D.; L Wildgoose, J.; H Bateman, R. *Applications of a Travelling Wave-Based Radio-Frequency Only Stacked Ring Ion Guide*; 2004; Vol. 18. https://doi.org/10.1002/rcm.1641.
(211) Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V; Ruotolo, B. T. Collision Cross Sections of Proteins and Their Complexes: A Calibration Framework and Database for Gas-Phase Structural Biology. *Anal. Chem.* **2010**, *82* (22), 9557–9565. https://doi.org/10.1021/ac1022953.
(212) Campuzano, I.; F Bush, M.; Robinson, C.; Beaumont, C.; Richardson, K.; Kim, H.; I Kim, H. *Structural*

Characterization of Drug-like Compounds by Ion Mobility Mass Spectrometry: Comparison of Theoretical and Experimentally Derived Nitrogen Collision Cross Sections; 2011; Vol. 84. https://doi.org/10.1021/ac202625t.

(213) Bush, M. F.; Campuzano, I. D. G.; Robinson, C. V. Ion Mobility Mass Spectrometry of Peptide Ions: Effects of Drift Gas and Calibration Strategies. *Anal. Chem.* **2012**, *84* (16), 7124–7130. https://doi.org/10.1021/ac3014498.

(214) Chawner, R.; Mccullough, B.; Giles, K.; Barran, P.; J Gaskell, S.; Eyers, C. *QconCAT Standard for Calibration of Ion Mobility-Mass Spectrometry Systems*; 2012; Vol. 11. https://doi.org/10.1021/pr3005327.

(215) Stoeckli, M.; Chaurand, P.; E. Hallahan, D.; M. Caprioli, R. *Imaging Mass Spectrometry: A New Technology for the Analysis of Protein Expression in Mammalian Tissues*; 2001; Vol. 7. https://doi.org/10.1038/86573.

(216) McDonnell, L. A.; Heeren, R. M. A. Imaging Mass Spectrometry. *Mass Spectrom. Rev.* **2007**, *26* (4), 606–643. https://doi.org/10.1002/mas.20124.

(217) Amstalden van Hove, E. R.; Smith, D. F.; Heeren, R. M. A. A Concise Review of Mass Spectrometry Imaging. *J. Chromatogr. A* **2010**, *1217* (25), 3946–3954. https://doi.org/10.1016/J.CHROMA.2010.01.033.

(218) Stauber, J.; MacAleese, L.; Franck, J.; Claude, E.; Snel, M.; Kaletas, B. K.; Wiel, I. M. V. D.; Wisztorski, M.; Fournier, I.; Heeren, R. M. A. On-Tissue Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (3), 338–347. https://doi.org/10.1016/J.JASMS.2009.09.016.

(219) Trim, P. J.; Henson, C. M.; Avery, J. L.; McEwen, A.; Snel, M. F.; Claude, E.; Marshall, P. S.; West, A.; Princivalle, A. P.; Clench, M. R. Matrix-Assisted Laser Desorption/Ionization-Ion Mobility Separation-Mass Spectrometry Imaging of Vinblastine in Whole Body Tissue Sections. *Anal. Chem.* **2008**, *80* (22), 8628–8634. https://doi.org/10.1021/ac8015467.

(220) Goscinny, S.; Joly, L.; De Pauw, E.; Hanot, V.; Eppe, G. Travelling-Wave Ion Mobility Time-of-Flight Mass Spectrometry as an Alternative Strategy for Screening of Multi-Class Pesticides in Fruits and Vegetables. *J. Chromatogr. A* **2015**, *1405*, 85–93. https://doi.org/10.1016/J.CHROMA.2015.05.057.

(221) Picotti, P.; Bodenmiller, B.; Mueller, L. N.; Domon, B.; Aebersold, R. Full Dynamic Range Proteome Analysis of S. Cerevisiae by Targeted Proteomics. *Cell* **2009**, *138* (4), 795–806. https://doi.org/10.1016/J.CELL.2009.05.051.

(222) Shliaha, P. V; Bond, N. J.; Gatto, L.; Lilley, K. S. Effects of Traveling Wave Ion Mobility Separation on Data Independent Acquisition in Proteomics Studies. *J. Proteome Res.* **2013**, *12* (6), 2323–2339. https://doi.org/10.1021/pr300775k.

(223) Langridge, J.; Rodriguez-Suarez, E.; Hughes, C.; Gethings, L.; Giles, K.; Wildgoose, J.; Stapels, M.; Fadgen, K.; Geromanos, S.; Vissers, H.; et al. *An Ion Mobility Assisted Data Independent LC-MS Strategy for the Analysis of Complex Biological Samples*; 2013; Vol. 9. https://doi.org/10.2174/1573411011309020006.

(224) Saba, J.; Bonneil, E.; Pomiès, C.; Eng, K.; Thibault, P. Enhanced Sensitivity in Proteomics Experiments Using FAIMS Coupled with a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer. *J. Proteome Res.* **2009**, *8* (7), 3355–3366. https://doi.org/10.1021/pr801106a.

(225) Hallen, R. W.; Shumate, C. B.; Siems, W. F.; Tsuda, T.; Hill, H. H. Preliminary Investigation of Ion Mobility Spectrometry after Capillary Electrophoretic Introduction. *J. Chromatogr. A* **1989**, *480*, 233–245. https://doi.org/10.1016/S0021-9673(01)84291-6.

(226) Jooß, K.; Meckelmann, S. W.; Klein, J.; Schmitz, O. J.; Neusüß, C. Capillary Zone Electrophoresis Coupled to Drift Tube Ion Mobility-Mass Spectrometry for the Analysis of Native and APTS-Labeled N-Glycans. *Anal. Bioanal. Chem.* **2018**, 1–10.

(227) Li, J.; Purves, R. W.; Richards, J. C. Coupling Capillary Electrophoresis and High-Field Asymmetric Waveform Ion Mobility Spectrometry Mass Spectrometry for the Analysis of Complex Lipopolysaccharides. *Anal. Chem.* **2004**, *76* (16), 4676–4683. https://doi.org/10.1021/ac049850d.

(228) Regueiro, J.; Negreira, N.; Berntssen, M. H. G. Ion-Mobility-Derived Collision Cross Section as an Additional Identification Point for Multiresidue Screening of Pesticides in Fish Feed. *Anal. Chem.* **2016**, *88* (22), 11169–11177. https://doi.org/10.1021/acs.analchem.6b03381.

(229) Scarff, C. A.; Thalassinos, K.; Hilton, G. R.; Scrivens, J. H. Travelling Wave Ion Mobility Mass Spectrometry Studies of Protein Structure: Biological Significance and Comparison with X-Ray Crystallography and Nuclear Magnetic Resonance Spectroscopy Measurements. *Rapid Commun. Mass Spectrom.* **2008**, *22* (20), 3297–3304. https://doi.org/10.1002/rcm.3737.

(230) Murray, K. *Definitions of Terms Relating to Mass Spectrometry (IUPAC Recommendations 2013)*; 2013; Vol. 35. https://doi.org/10.1515/ci-2013-0516.

(231) Wyttenbach, T.; von Helden, G.; Bowers, M. T. Gas-Phase Conformation of Biological Molecules: Bradykinin. *J. Am. Chem. Soc.* **1996**, *118* (35), 8355–8364.

(232) Baumketner, A.; Bernstein, S. L.; Wyttenbach, T.; Bitan, G.; Teplow, D. B.; Bowers, M. T.; Shea, J. Amyloid B-protein Monomer Structure: A Computational and Experimental Study. *Protein Sci.* **2006**, *15* (3), 420–428.

(233) Dear, G. J.; Munoz-Muriedas, J.; Beaumont, C.; Roberts, A.; Kirk, J.; Williams, J. P.; Campuzano, I. Sites of Metabolic Substitution: Investigating Metabolite Structures Utilising Ion Mobility and Molecular Modelling. *Rapid Commun. Mass Spectrom.* **2010**, *24* (21), 3157–3162.

(234) Wyttenbach, T.; von Helden, G.; Batka, J. J.; Carlat, D.; Bowers, M. T. Effect of the Long-Range Potential

on Ion Mobility Measurements. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (3), 275–282.
(235) Shvartsburg, A. A.; Jarrold, M. F. An Exact Hard-Spheres Scattering Model for the Mobilities of Polyatomic Ions. *Chem. Phys. Lett.* **1996**, *261* (1–2), 86–91.
(236) Larriba, C.; Hogan Jr, C. J. Ion Mobilities in Diatomic Gases: Measurement versus Prediction with Non-Specular Scattering Models. *J. Phys. Chem. A* **2013**, *117* (19), 3887–3901.
(237) Larriba, C.; Hogan Jr, C. J. Free Molecular Collision Cross Section Calculation Methods for Nanoparticles and Complex Ions with Energy Accommodation. *J. Comput. Phys.* **2013**, *251*, 344–363.
(238) Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. Mobilities of Carbon Cluster Ions: Critical Importance of the Molecular Attractive Potential. *J. Chem. Phys.* **1998**, *108* (6), 2416–2423.
(239) Jurneczko, E.; Barran, P. E. How Useful Is Ion Mobility Mass Spectrometry for Structural Biology? The Relationship between Protein Crystal Structures and Their Collision Cross Sections in the Gas Phase. *Analyst* **2011**, *136* (1), 20–28.
(240) Gabelica, V.; Marklund, E. Fundamentals of Ion Mobility Spectrometry. *Curr. Opin. Chem. Biol.* **2018**, *42*, 51–59. https://doi.org/10.1016/J.CBPA.2017.10.022.
(241) Shi, H.; Pierson, N. A.; Valentine, S. J.; Clemmer, D. E. Conformation Types of Ubiquitin [M+ 8H] 8+ Ions from Water: Methanol Solutions: Evidence for the N and A States in Aqueous Solution. *J. Phys. Chem. B* **2012**, *116* (10), 3344–3352.
(242) Bereszczak, J. Z.; Barbu, I. M.; Tan, M.; Xia, M.; Jiang, X.; van Duijn, E.; Heck, A. J. R. Structure, Stability and Dynamics of Norovirus P Domain Derived Protein Complexes Studied by Native Mass Spectrometry. *J. Struct. Biol.* **2012**, *177* (2), 273–282.
(243) Jenner, M.; Ellis, J.; Huang, W.; Lloyd Raven, E.; Roberts, G. C. K.; Oldham, N. J. Detection of a Protein Conformational Equilibrium by Electrospray Ionisation-ion Mobility-mass Spectrometry. *Angew. Chemie Int. Ed.* **2011**, *50* (36), 8291–8294.
(244) Wyttenbach, T.; Grabenauer, M.; Thalassinos, K.; Scrivens, J. H.; Bowers, M. T. The Effect of Calcium Ions and Peptide Ligands on the Relative Stabilities of the Calmodulin Dumbbell and Compact Structures. *J. Phys. Chem. B* **2009**, *114* (1), 437–447.
(245) Gidden, J.; Bushnell, J. E.; Bowers, M. T. Gas-Phase Conformations and Folding Energetics of Oligonucleotides: DTG-and DGT. *J. Am. Chem. Soc.* **2001**, *123* (23), 5610–5611.
(246) Nabuchi, Y.; Hirose, K.; Takayama, M. Ion Mobility and Collision-Induced Dissociation Analysis of Carbonic Anhydrase 2. *Anal. Chem.* **2010**, *82* (21), 8890–8896.
(247) Smith, D. P.; Giles, K.; Bateman, R. H.; Radford, S. E.; Ashcroft, A. E. Monitoring Copopulated Conformational States during Protein Folding Events Using Electrospray Ionization-Ion Mobility Spectrometry-Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (12), 2180–2190.
(248) Hopper, J. T. S.; Oldham, N. J. Collision Induced Unfolding of Protein Ions in the Gas Phase Studied by Ion Mobility-Mass Spectrometry: The Effect of Ligand Binding on Conformational Stability. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (10), 1851–1858.
(249) Hyung, S.-J.; Robinson, C. V; Ruotolo, B. T. Gas-Phase Unfolding and Disassembly Reveals Stability Differences in Ligand-Bound Multiprotein Complexes. *Chem. Biol.* **2009**, *16* (4), 382–390.
(250) Beveridge, R.; Chappuis, Q.; Macphee, C.; Barran, P. Mass Spectrometry Methods for Intrinsically Disordered Proteins. *Analyst* **2013**, *138* (1), 32–42.
(251) Smith, D. P.; Radford, S. E.; Ashcroft, A. E. Elongated Oligomers in B2-Microglobulin Amyloid Assembly Revealed by Ion Mobility Spectrometry-Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2010**, *107* (15), 6794–6798.
(252) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. Evidence for Macromolecular Protein Rings in the Absence of Bulk Water. *Science (80-.).* **2005**, *310* (5754), 1658–1661.
(253) Uetrecht, C.; Versluis, C.; Watts, N. R.; Wingfield, P. T.; Steven, A. C.; Heck, A. J. R. Stability and Shape of Hepatitis B Virus Capsids in Vacuo. *Angew. Chemie Int. Ed.* **2008**, *47* (33), 6247–6251.
(254) Jarrold, M. F. Helices and Sheets in Vacuo. *Phys. Chem. Chem. Phys.* **2007**, *9* (14), 1659–1671. https://doi.org/10.1039/B612615D.
(255) Zilch, L. W.; Kaleta, D. T.; Kohtani, M.; Krishnan, R.; Jarrold, M. F. Folding and Unfolding of Helix-Turn-Helix Motifs in the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2007**, *18* (7), 1239–1248.
(256) Florance, H. V; Stopford, A. P.; Kalapothakis, J. M.; McCullough, B. J.; Bretherick, A.; Barran, P. E. Evidence for α-Helices in the Gas Phase: A Case Study Using Melittin from Honey Bee Venom. *Analyst* **2011**, *136* (17), 3446–3452.
(257) Wu, C.; Klasmeier, J.; Hill Jr, H. H. Atmospheric Pressure Ion Mobility Spectrometry of Protonated and Sodiated Peptides. *Rapid Commun. mass Spectrom.* **1999**, *13* (12), 1138–1142.
(258) Liu, D.; Seuthe, A. B.; Ehrler, O. T.; Zhang, X.; Wyttenbach, T.; Hsu, J. F.; Bowers, M. T. Oxytocin-Receptor Binding: Why Divalent Metals Are Essential. *J. Am. Chem. Soc.* **2005**, *127* (7), 2024–2025.
(259) Chen, L.; Gao, Y. Q.; Russell, D. H. How Alkali Metal Ion Binding Alters the Conformation Preferences of Gramicidin A: A Molecular Dynamics and Ion Mobility Study. *J. Phys. Chem. A* **2011**, *116* (1), 689–696.
(260) Garcia, I. R.; Giles, K.; Bateman, R. H.; Gaskell, S. J. Studies of Peptide A-and b-Type Fragment Ions Using Stable Isotope Labeling and Integrated Ion Mobility/Tandem Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (12), 1781–1787.

(261) Raschke, T. M. Water Structure and Interactions with Protein Surfaces. *Curr. Opin. Struct. Biol.* **2006**, *16* (2), 152–159.
(262) Wolynes, P. G. Biomolecular Folding in Vacuo!!!(?). *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92* (7), 2426.
(263) Ruotolo, B. T.; Robinson, C. V. Aspects of Native Proteins Are Retained in Vacuum. *Curr. Opin. Chem. Biol.* **2006**, *10* (5), 402–408.
(264) Gao, B.; Wyttenbach, T.; Bowers, M. T. Protonated Arginine and Protonated Lysine: Hydration and Its Effect on the Stability of Salt-Bridge Structures. *J. Phys. Chem. B* **2009**, *113* (29), 9995–10000.
(265) Rodriguez-Cruz, S. E.; Klassen, J. S.; Williams, E. R. Hydration of Gas-Phase Ions Formed by Electrospray Ionization. *J. Am. Soc. Mass Spectrom.* **1999**, *10* (10), 958–968.
(266) Fye, J. L.; Woenckhaus, J.; Jarrold, M. F. Hydration of Folded and Unfolded Gas-Phase Proteins: Saturation of Cytochrome c and Apomyoglobin. *J. Am. Chem. Soc.* **1998**, *120* (6), 1327–1328.
(267) Rodriguez-Cruz, S. E.; Klassen, J. S.; Williams, E. R. Hydration of Gas-Phase Gramicidin S (M+ 2H) 2+ Ions Formed by Electrospray: The Transition from Solution to Gas-Phase Structure. *J. Am. Soc. Mass Spectrom.* **1997**, *8* (5), 565–568.
(268) Hall, Z.; Robinson, C. V. Do Charge State Signatures Guarantee Protein Conformations? *J. Am. Soc. Mass Spectrom.* **2012**, *23* (7), 1161–1168.
(269) Berezovskaya, Y.; Porrini, M.; Barran, P. E. The Effect of Salt on the Conformations of Three Model Proteins Is Revealed by Variable Temperature Ion Mobility Mass Spectrometry. *Int. J. Mass Spectrom.* **2013**, *345*, 8–18.
(270) Kantardjieff, K. A.; Rupp, B. Matthews Coefficient Probabilities: Improved Estimates for Unit Cell Contents of Proteins, DNA, and Protein–Nucleic Acid Complex Crystals. *Protein Sci.* **2003**, *12* (9), 1865–1871.
(271) Arcella, A.; Portella, G.; Ruiz, M. L.; Eritja, R.; Vilaseca, M.; Gabelica, V.; Orozco, M. Structure of Triplex DNA in the Gas Phase. *J. Am. Chem. Soc.* **2012**, *134* (15), 6596–6606.
(272) Baker, E. S.; Dupuis, N. F.; Bowers, M. T. DNA Hairpin, Pseudoknot, and Cruciform Stability in a Solvent-Free Environment. *J. Phys. Chem. B* **2009**, *113* (6), 1722–1727.
(273) Nagesh, N.; Chatterji, D. Ammonium Ion at Low Concentration Stabilizes the G-Quadruplex Formation by Telomeric Sequence. *J. Biochem. Biophys. Methods* **1995**, *30* (1), 1–8.
(274) Baker, E. S.; Bernstein, S. L.; Gabelica, V.; De Pauw, E.; Bowers, M. T. G-Quadruplexes in Telomeric Repeats Are Conserved in a Solvent-Free Environment. *Int. J. Mass Spectrom.* **2006**, *253* (3), 225–237.
(275) Rosu, F.; Gabelica, V.; Poncelet, H.; De Pauw, E. Tetramolecular G-Quadruplex Formation Pathways Studied by Electrospray Mass Spectrometry. *Nucleic Acids Res.* **2010**, *38* (15), 5217–5225.
(276) Konijnenberg, A.; Butterer, A.; Sobott, F. Native Ion Mobility-Mass Spectrometry and Related Methods in Structural Biology. *Biochim. Biophys. Acta (BBA)-Proteins Proteomics* **2013**, *1834* (6), 1239–1256.
(277) Kaddis, C. S.; Loo, J. A. Native Protein MS and Ion Mobility: Large Flying Proteins with ESI. ACS Publications 2007.

CHAPTER 3

Bacillus licheniformis peptidoglycan characterization by CZE-MS: assessment with the benchmark RP-HPLC-MS method

3.1. Context of the chapter

The peptidoglycan (PGN) layer found in the cell wall of bacteria is a characteristic and essential structure that forms a sacculus surrounding the cytoplasmic membrane as shown in Figure 3.1.

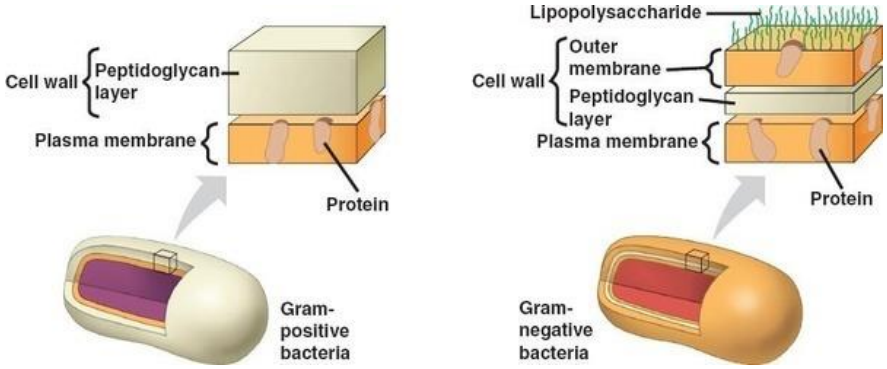


Figure 3.1 - Schematic representation of Gram-positive and Gram-negative bacteria with a focus on the structure of their cell wall from Klug, W. S.; Cummings, M. R. Concepts of Genetics; Upper Saddle River, NJ: Pearson Education, 2006

The glycan strands are composed of alternating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) residues linked by β_{1-4} bonds. Each Mur*N*Ac residue is then substituted on its *D*-lactyl moiety by a peptide stem whose primary structure varies according to the bacterial species.

In general, the peptide stem consists of the pentapeptide NH_2-*L*-Ala-*D*-*i*Glu-*m*-A_2pm (or *L*-Lys)-*D*-Ala-*D*-Ala-OH. The *m*-A_2pm residue is mostly found in Gram negative bacteria and *Bacillus* species while *L*-Lys is mostly found in Gram positive bacteria. The chemical composition of this type of PGN is presented in Figure 3.2.

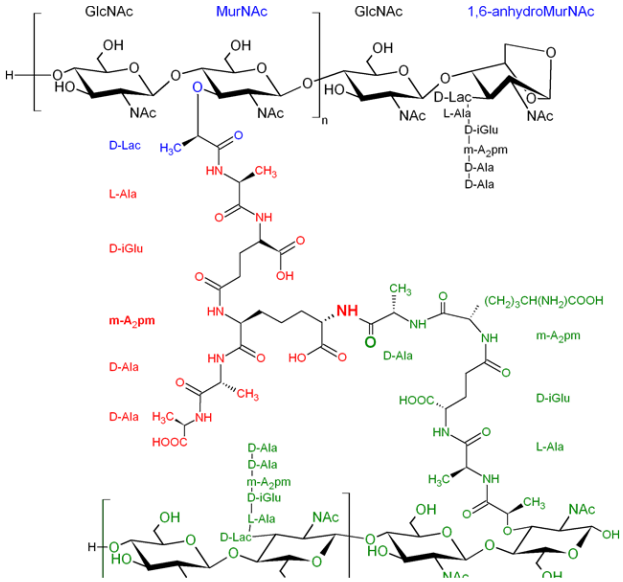


Figure 3.2 - General structure of the peptidoglycan in Gram negative and *Bacillus* bacteria from The Structure of Bacterial Cell Wall of N. Jean, C. Bougault & J.P. Simorre

PGN fragments (often referred as "muropeptides") are soluble fragments that can act as messengers in diverse cell-signaling events. In bacteria, they are used for bacterial communication

as signal molecules in spore resuscitation and germination. The induction of β-lactamase production can also be triggered by an accumulation of PGN fragments. In eukaryotes, the detection of muropeptides is perceived as a sign of bacterial invasion and initiates an immune response.

Because of its importance for bacteria and almost unique chemical composition in the living world, PGN is an ideal target for the use of antibiotics. Not surprisingly, numerous antibacterial antibiotics aim at inhibiting the peptidoglycan biosynthesis. For example, β-lactams are widely used in clinical treatment for their antibacterial effects. However, the recurrent clinical use of such antibiotics has ultimately led to the emergence of bacterial resistance, which is now commonly found for β-lactams.

The mechanism of bacterial resistance to antibiotics mainly relies on the introduction of various PGN chemical modifications. For example, the amidation of the carboxyl group of *D*-Glu present in the PGN peptide moiety contributes to the antibiotics resistance of methicillin-resistant *Staphylococcus aureus* (MRSA). Additionally, the deacetylation of *N*-acetyl glucosamine contained in the PGN glycan moiety prevents the binding of enzymes such as autolysins or PGN hydrolases and therefore protects the PGN from degradation. Finally, the O-acetylation of *N*-acetyl muramic acid is linked to the emergence of multiple resistances such as resistance to lysozyme or penicillin, endogenous autolysis or even macrophage killing.

The chemical PGN modification is also advantageous for bacteria as the introduction of an amidation on the α-carboxyl of the ε-carboxyl group of the *m*-A_2pm affects PGN recognition by the mammalian host innate immune system, facilitating bacterial infection.

In this context, the development of reliable analytical methods to identify the muropeptide composition and especially, the chemical modifications of such fragments in complex samples is of major interest.

In general, the peptidoglycan is enzymatically digested by a mutanolysin to generate the corresponding muropeptides, which are subsequently analyzed by reversed-phase (C18) high performance liquid chromatography (RP-HPLC). Though effective, the major limitations of this technique are the meticulous control over several experimental parameters, the large quantities of sample consumed and the difficulties to efficiently separate the chemically modified muropeptides from their unmodified equivalents. Also, because PGN fragments are naturally hydrophilic, the use of RP-HPLC is poorly adapted, even when an ion pairing agent is used (e.g. trifluoroacetic acid). The use of alternative more adapted chromatographic modes such as hydrophilic interaction chromatography (HILIC) could be considered, despite its lower separation efficiency compared to RP-HPLC.

Because capillary zone electrophoresis (CZE) is particularly well-adapted for the separation of hydrophilic ionized and/or polar compounds, we introduced capillary zone electrophoresis coupled to mass spectrometry (CZE-MS) as an online tool for the separation and characterization of muropeptide. Besides, the major chemical modifications of muropeptides (the deacetylation of the glycan moiety and the amidation of the peptide moiety) are both involved in a modification of the global charge of the muropeptides in solution. As a result, CZE appears to be perfectly adapted to assess these chemical modifications, which are co-eluted by RP-HPLC.

This work was initiated by the NetRBI ARC project (Action de Recherche Concertée) and the NetRBI action (Modelling of the Network Regulating *Bacillus licheniformis* BlaP β-Lactamase Induction) of Professor Bernard Joris from the Center of Protein Engineering of the University

of Liège (Centre d'Ingénierie des Protéines, CIP) in collaboration with the Mass Spectrometry Laboratory. This project mainly concerns the refinement of the metabolic flux of PGN fragments to understand the resistance mechanisms implied in the antibiotic resistance in the model bacteria *Bacillus spp.*

In our workflow, we first used the conventional RP-HPLC method to characterize the muropeptides contained in the PGN of *Bacillus licheniformis*. The identification of the muropeptides was performed by offline MALDI-MS and ESI-MS on a collection of 100 fractions collected during the chromatographic separation. Then, online CZE-ESI-MS was successfully applied on the same muropeptide mixture of *B. licheniformis*. A particular attention was taken concerning the chemical modifications of the muropeptides using CZE-ESI-MS and major chemical modifications including the amidation and the deacetylation were well characterized based on the use of tandem mass spectrometry. In particular, the development of an online CZE-ESI-MS^3 method allowed for the first time to identify the amidation site among amidation isomers containing several amidation sites.

The results are presented in the following section as a published research article in Electrophoresis. The related Supporting Information (SI) are presented directly after the article. In the present work, the sample preparation of the bacterial samples and HPLC-UV or HPLC-MS analysis were performed by Madeleine Boulanger while the CZE-ESI-MS^n experiments and optimization were performed by myself.

3.2. Results – published research article in Electrophoresis

Bacillus licheniformis peptidoglycan characterization by CZE-MS: assessment with the benchmark RP-HPLC-MS method

Madeleine Boulanger and Cédric Delvaux; Loïc Quinton; Bernard Joris; Edwin De Pauw and Johann Far

Electrophoresis; June 2019

DOI: 10.1002/elps.201900147

Electrophoresis 2019, *0*, 1–11

Madeleine Boulanger[1]*
Cédric Delvaux[2]*
Loïc Quinton[2]
Bernard Joris[1]
Edwin De Pauw[2]
Johann Far[2]

[1]Center for Protein Engineering, InBioS Research Unit, Quartier Agora, University of Liège, Liège, Belgium
[2]Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Liège, Belgium

Received March 26, 2019
Revised May 22, 2019
Accepted May 23, 2019

Research Article

Bacillus licheniformis peptidoglycan characterization by CZE–MS: Assessment with the benchmark RP-HPLC-MS method

Peptidoglycan or murein is an essential polymer found in bacterial cell wall. It is a dynamic structure that is continuously remodeled or modified during bacterial cell growth or in presence of cell wall stresses. These modifications are still poorly understood mainly due to the peptidoglycan, which is rather non-soluble, and the difficulties to separate the hydrophilic glycopeptides (muropeptides) by reversed phase liquid chromatography, generated by the enzymatic digestion using mutanolysin, an *N*-acetyl-muramidase, cleaving the $\beta_{1\rightarrow4}$ bound between *N*-acetylglucosamine and *N*-acetylmuramic acid. Here, we report the use of CZE–MS for an easy and fast screening of muropeptides generated by the action of muramidase on the *Bacillus licheniformis* cell wall. Electron transfer and CID–MS were also used to unambiguously identify and localize the presence or the absence of amidation and acetylation moieties on muropeptide variants. The reference method to analyse muropeptides by reversed phase chromatography was also tested and the advantages and disadvantages of both methods were evaluated.

Keywords:
Acetylation / Amidation / CZE-MS / HPLC-MS / Peptidoglycan
DOI 10.1002/elps.201900147

Additional supporting information may be found online in the Supporting Information section at the end of the article.

1 Introduction

The peptidoglycan (PGN) or murein is an essential extracellular polymer found in the cell wall of most bacteria. It forms a mesh-like layer surrounding the cytoplasmic membrane and consists of glycan strands cross-linked by short peptides [1]. It defines the shape of the bacterial cells and protects them against their own high internal osmotic pressure. The glycan strands are made up of alternating *N*-acetylglucosamine (Glc*N*Ac) and *N*-acetylmuramic acid (Mur*N*Ac) residues linked by β_{1-4} bonds (Fig. 1). Each Mur*N*Ac residue is substituted on its D-lactyl moiety by a peptide stem whose primary structure varies according to the bacterial species. In general, in the nascent PGN, the peptide stem consists of the pentapeptide NH_2-L-Ala-D-*i*Glu-*m*-A_2pm (or L-Lys)-D-Ala-D-Ala-OH (*m*-A_2pm, *meso*-2,6-diaminopimelic acid, *i*Glu for isoglutamyl residue). The *m*-A_2pm residue is mostly found in Gram negative bacteria while L-Lys is mostly found in Gram positive bacteria, except for *Bacillus* species (*m*-A_2pm). In the mature PGN, the last D-Ala residue of the pentapeptide stem is cleaved and the remaining D-Ala carboxyl end can either be involved in a new amide bond with the free *m*-A_2pm amino group of a pentapeptide located on an adjacent glycan chain (crosslinking reaction), or remains free (Fig. 1) [2]. In *Escherichia coli* and *Bacillus licheniformis*, the mature PGN composition follows the above classical description as presented in Fig. 1. However, variations can be found such as in *Staphylococcus aureus* where the *m*-A_2pm is replaced by a lysine and the cross-linking between the peptide stems is achieved *via* an additional pentaglycyl bridge [3].

During the bacterial life, the PGN is continuously remodelled and chemically modified. Indeed, the PGN assembly requires both biosynthetic and hydrolytic enzymes to introduce new glycan chains in PGN mesh [4]. Moreover, chemical modifications such as Mur*N*Ac *O*-acetylation and Glc*N*Ac deacetylation contribute to bacterial pathogenicity by

Correspondence: Dr. Johann Far, Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Août 11, B-4000 Liège, Belgium
E-mail: Johann.Far@uliege.be

Abbreviations: **ETD**, Electron transfer dissociation; **Glc*N*Ac**, *N*-acetylglucosamine; ***m*-A$_2$pm**, *meso*-diaminopimelic acid; **Mur*N*Ac**, *N*-acetylmuramic acid; **PGN**, peptidoglycan

*These authors equally contributed to this work.

Color online: See article online to view Fig. 1 in color.

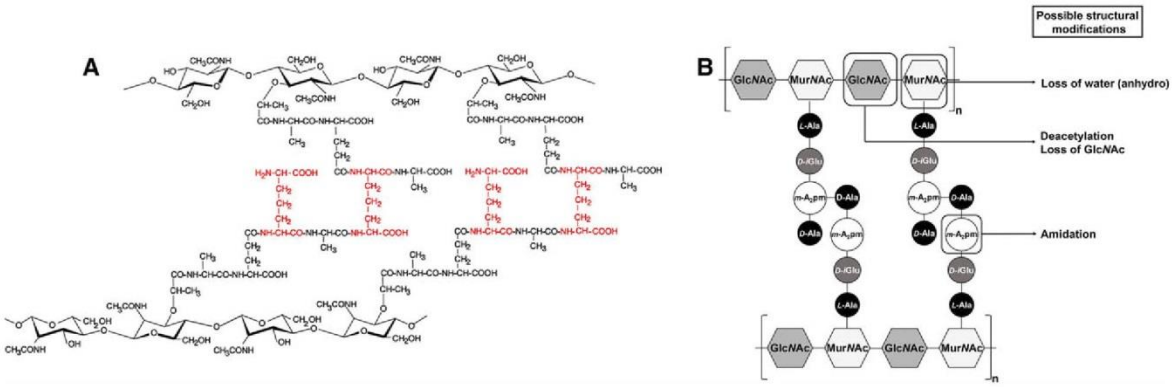


Figure 1. Peptidoglycan of *Bacillus licheniformis*. (A) Structure of the PGN (adapted from Vollmer et al. [2]); (B) PGN schematic representation with possible structural modifications. Meso-diaminopimelic acid residues are in red.

protecting the bacteria against the (e.g., human) lysozyme action [5,6]. In addition, the amidation of free α- and ε-carboxyl groups of *D*-*i*Glu and *m*-A_2pm affects PGN recognition by the mammalian host innate immune system [7,8]. The PGN biosynthesis is therefore a well-established target for antibiotics, such as penicillin (a β-lactam antibiotic), vancomycin, collicin or bacitracin [9]. In fact, sublethal concentrations of β-lactam antibiotics induce an antibiotic stress of the cell wall that causes the accumulation of PGN fragments inside the cell. In some pathogenic bacterial strains, the PGN fragments accumulation in the cytoplasm induces the production of a β-lactamase that inactivates the β-lactam antibiotics [10,11]. The characterization of PGN metabolism is then crucial for a better understanding of the bacterial cell life and to avoid the bacterial resistance to antibiotics. In this context, efficient and fast analytical methods for PGN characterization are needed with the aim of improving our knowledge concerning PGN metabolism as well as refining the model of β-lactamase induction due to PGN fragments accumulation.

As the PGN is a large insoluble macromolecule, it is easily purified after a hot SDS treatment of bacterial cells. The proteins covalently linked to mature PGN are discarded by a proteolytic treatment, whereas for Gram-positive bacteria an additional treatment with hydrofluoric acid is performed to remove the bound teichoic acids. At this stage, highly pure PGN is obtained and can be characterized after a partial digestion with a muramidase to generate soluble PGN fragments (muropeptides) bearing a Mur*N*Ac as reducing end. The reference method for these muropeptides separation is based on the off-line combination of RP-HPLC with UV detection at 200–205 nm. In this method, the terminal Mur*N*Ac is reduced into *N*-acetyl-muraminitol (Mol*N*Ac) to avoid peak duplication due to the anomeric isomerisation at the C1 position [12]. Additionally, the off-line muropeptide analysis by high resolution MALDI or ESI-MS allows the molecular mass and empiric formula determination while MS/MS leads to the structural elucidation of the PGN fragments [13,14]. Recently, the method was improved to an online UPLC-ESI-MS method, thus allowing the immediate determination of exact masses and the structure elucidation of the eluted muropeptides [15,16]. Alternatively, two electrophoretic methods were developed for PGN analysis: PAGE was used to study the muropeptide composition of various bacterial strains whereas CZE was applied for quantifying PGN precursors (UDP-Mur*N*Ac-pentapeptide and disaccharide pentapeptide) and certain muropeptides [17,18]. In the latter case, muropeptides were not directly identified but the effect of an antibiotic exposure on PGN composition was highlighted by the change of profile of muropeptides during the CZE separation.

In this study, CZE coupled to high-resolution MS (FT-ICR-MS) was applied for the first time to analyse the muropeptides generated by the digestion of pure *B. licheniformis* PGN by a muramidase. Complementary, CID and Electron transfer dissociation (ETD) fragmentations were also used to elucidate fine structural properties. Finally, the performance of CZE was compared to the RP-HPLC-UV/MS reference method and critically evaluated.

2 Materials and method

2.1 Chemicals and reagents

All eluents, background electrolytes, reagents, and enzymes were purchased from Sigma-Aldrich (Bornem, Belgium) except for LB Broth medium (MP Biomedicals, Huissen, The Netherlands). TFA and ACN were obtained from ThermoFisher Scientific (Asse, Belgium). The chromatographic/electrophoretic eluents were "HPLC-MS grade". MilliQ ultra-pure water (resistivity of 18.2 MΩ.cm) was daily produced using a system from Millipore (Molsheim, France).

2.2 Sample preparation

2.2.1 Isolation and purification of PGN

B. licheniformis 749/I was grown in LB Broth medium at 37°C and its PGN was isolated from cells in exponential phase by a

method adapted from Atrih and coworkers for *Bacillus subtilis* (see the Supporting Information) [19].

2.2.2 Muropeptides preparation

Purified PGN samples (200 μL) were digested at 37°C by the addition of 25 μL of mutanolysin from *Streptomyces globisporus* ATCC 21553 (Sigma Aldrich M9901, 5 U/μL) in 25 μL of ammonium acetate 0.2 M (pH 6.9) for 16 hours at 37°C. Insoluble material was removed by centrifugation (14 000 g, 10 min, at 4°C), and the resulting PGN hydrolysate (muropeptides) was stored at 4°C. Muropeptides were reduced by using sodium borohydride ($NaBH_4$) according to the protocol described by Dougherty et al. [20] where the reaction was stopped by an addition of formic acid (FA, final concentration of 0.50 M) instead of phosphoric acid to avoid any significant chemical noise and ion suppression during MS analyses.

2.3 Muropeptides offline analysis by RP-HPLC-MS

A Waters HPLC Alliance System (Manchester, UK) coupled with a 2998 Photodiode Array detector (180-800 nm, Waters) were used. The muropeptides separation was achieved on a RP Nucleodur C_{18} EC column (250 × 4.6 mm i.d., 3 μm, Macherey-Nagel) at 40°C. The elution buffers were as follows: buffer A: TFA/water (0.05/99.95, v/v); buffer B: TFA/ACN (0.05/99.95, v/v). After the column equilibration in buffer A, a linear gradient from 0 to 18% buffer B over a period of 100 min was started 6 min after the 100 μL sample injection (100% buffer A). The flow rate was fixed to 0.50 mL/min and the eluted compounds were detected by monitoring A^{201nm}. Each detected peak was collected at the detector outlet, freeze-dried and suspended in 100 μL of ACN/water/FA (50/49.9/0.1, v/v/v). The analysis of HPLC fractions was performed by ESI-MS using the SYNAPT G2 HDMS (Waters, Manchester, UK) fitted with an electrospray ion source or by MALDI-MS using the MALDI-TOF/TOF UltrafleXtreme (Bruker Daltonics, Bremen, Germany) (see the Supporting Information for more details). The SYNAPT G2 HDMS data were reprocessed using MassLynx v4.1 and the MALDI UltrafleXtreme data were reprocessed using DataAnalysis v4.0.

2.4 Muropeptides online analysis by CZE-MS

The CZE separation was performed on a Beckman P/ACE MDQ system (Beckman Coulter – SCIEX Separations, Brea, CA, USA) running under 32 karat v7.0 software. The bare fused-silica capillary (50 μm i.d. and 365 μm o.d.) was obtained from Polymicro (Optronis GmbH, Kehl, Germany) and cut to a total length of 90 cm. This capillary was conditioned following a standard activation procedure (see the Supporting Information). The sample was loaded hydrodynamically at 3 psi for 10 s. A solution – containing 10 μM of two peptides, NH_2-MRFA-OH and human [Glu^1]-Fibrinopeptide B (Glufib) (Sigma-Aldrich) – was co-injected as controls. The separation was performed applying 30 kV (normal polarity, cathode at the CE inlet) during 120 min with a maintained capillary temperature of 25°C using 100 mM formic acid as BGE (pH 2.38 according to Peakmaster v5.3).

The CE system was connected to the mass spectrometer via an Analis SL CE-MS coaxial sprayer (PN 10–301 347, Analis, Suarlée, Belgium). The sheath liquid consisted of water/2-propanol (50/50, v/v) with 10 mM FA and was delivered at 1.5 μL/min using a syringe pump. ESI-MS detection, CID-MS^2 and MS^3 experiments were performed on a Thermo Finnigan LTQ FT-ICR Ultra Hybrid mass spectrometer 7T (Thermo Scientific, Merelbeke, Belgium) fitted with the nanospray Flex ion source. For the hybrid CID and ETD experiments, a Bruker Amazon ETD Speed 3D ion trap (Bruker Daltonics, Bremen, Germany) was directly coupled with the coaxial sprayer on a modified online nanoElectrospray ion source (obsolete ion source design). Other MS-related parameters are detailed in the Supporting Information.

3 Results and discussion

3.1 Muropeptides analysis by offline RP-HPLC and MALDI or ESI-MS

Bacillus licheniformis is a well-known Gram positive bacterium species producing PGN that has a composition similar to *E. coli* as reported by Schleifer and coworkers [21]. Nonetheless, the produced PGN by *B. licheniformis* may be subjected to amidation on the *m*-A_2pm moieties. One the other hand, the occurrence of Glc*N*Ac deacetylation is commonly reported, especially for the *Bacillus* species [3]. For this reason, *B. licheniformis* is a good model organism to produce a naturally modified PGN (i.e., amidation and deacetylation). The separation of muropeptides generated from *B. licheniformis* PGN, mutanolysin-digested, $NaBH_4$-reduced, and purified by RP-HPLC applying a linear TFA-ACN gradient is shown in Fig. 2A (see Supporting Information Fig. 1 and Table 2 for details about elution time and estimated PGN fragment contents). Five independent replicates of PGN preparations were compared and provided repeatable elution profiles (see Supporting Information Fig. 1 for more details), resulting in 39 baseline or almost baseline resolved peaks. One hundred fractions (0.25 mL) were collected from 15 to 75 min. Each fraction was analysed by MALDI-TOF MS and ESI-TOF MS in positive ion mode. The comparison between MALDI and ESI as ionization sources showed that the spectral interpretation using ESI was facilitated due to decreased intensity of Na^+ and K^+ adducts originating from the $NaBH_4$ reduction step [22]. However, the loss of one Glc*N*Ac moiety is more frequent when ESI mode is used (see Supporting Information for more details).

The ESI- and MALDI-MS results for HPLC peaks 1–39 (chromatogram in Fig. 2A) are summarized in

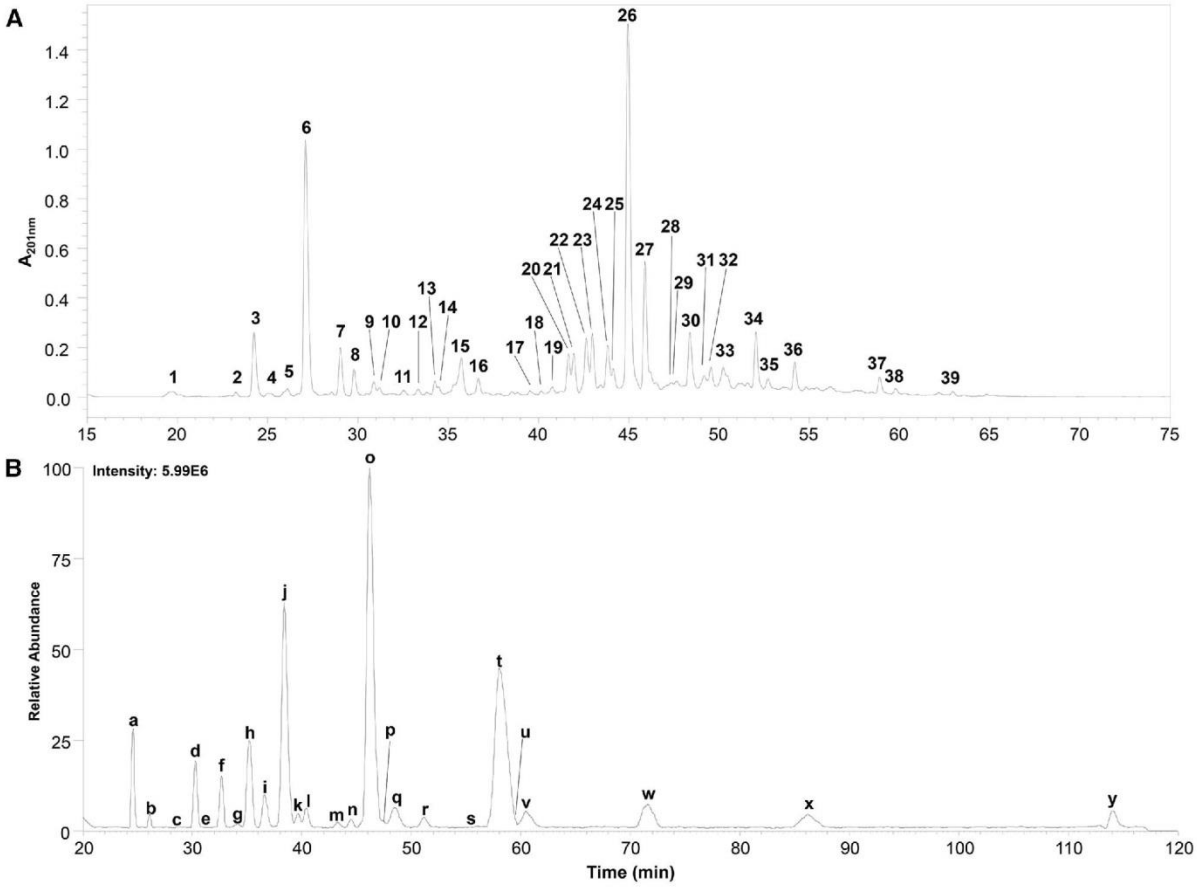


Figure 2. Muropeptide elution patterns of $NaBH_4$ reduced peptidoglycan of *Bacillus licheniformis*. (A) RP-LC-UV chromatogram; (B) CZE-FT-ICR-MS base peak electropherogram.

Table 1. Finally, a total of 62 chemically different species were detected including disaccharide-peptides (monomers or MONO:peptide, for more details see Table 1), bis-disaccharide-peptides (DI:peptide), tris-disaccharide-peptides (TRI:peptide), and one tetrakis-disaccharide-peptide (TETRA:peptide) (see Supporting Information Tables 1 and 4 for more nomenclature details and for details, respectively). As for *B. subtilis* PGN [19], multiple muropeptide modifications were identified: the amidation of the *m*-A_2pm residue ($\Delta_{m/z} = -0.984$ Da for the theoretical monoisotopic mass), the loss of the *N*-acetyl group on Glc*N*Ac ($\Delta_{m/z} = -42.010$ Da for the theoretical monoisotopic mass), the Mur*N*Ac modification to 1,6-anhydro Mur*N*Ac (loss of water and loss of H_2 due to $NaBH_4$ reduction step, $\Delta_{m/z} = -20.026$ Da for the theoretical monoisotopic mass), the loss of a Glc*N*Ac ($\Delta_{m/z} = -203.079$ Da for the theoretical monoisotopic mass) and a 77.94 Da shift attributed to a phosphate moiety bound to C6 of one Mur*N*Ac. The ESI analysis highlighted that the *m*-A_2pm amidation of the carboxylic group is present in the vast majority of analysed muropeptides and the Glc*N*Ac deacetylation occurs frequently.

3.2 Muropeptides analysis by online CZE-ESI-MS

The base peak current electropherogram obtained on the released muropeptides after enzymatic digestion with subsequent $NaBH_4$ reduction of *B. licheniformis* PGN is shown in Fig. 2B (see Supporting Information Table 3 for details). The optimized separation conditions resulted in 25 peaks (base peak electropherogram) detected in a 120 min separation window. No prominent peak was detected after a migration time higher than 120 min. Interestingly, in contrast to RP-HPLC, the $NaBH_4$ reduction step is not mandatory since the Mur*N*Ac anomerization does not lead to peak duplication in CZE-MS (data not shown). The analysis of the non-reduced muropeptide mixture showed the presence of the same muropeptides with a 2 Da mass shift. In addition, the absence of this reduction step significantly decreased the presence of salt adducts in the MS spectrum.

As could be expected from the Stoke-Einstein equation, muropeptides with a higher positive charge and/or lower molecular mass presented lower migration times (higher velocity) than less charged and/or larger (heavier) muropeptides. The exact mass obtained by FT-ICR MS allowed the

Table 1. Comparison of RP-HPLC and CZE for the detection of PGN muropeptides

CZE peak (Fig. 2B)	Experimental m/z (mass accuracy) [a]	Charge state	Muropeptide[b]	Corresponding peak in RP-HPLC (Fig. 2A)
a	524.26507 (0.21 ppm)	$[M+H]^+$	MRFA (internal standard 1)	-
b	556.76939 (−0.22 ppm)	$[M+2H]^{2+}$	MONO:tri-di+2A	7
	535.76428 (0.09 ppm)	**$[M+2H]^{2+}$**	**MONO:tri-di+2A–1Ac**	7
c	**635.80431 (0.36 ppm)**	**$[M+2H]^{2+}$**	**MONO:tri-tetra+2A–1Ac**	15
N.D[c]	**950.338 (1.13 ppm)**	**$[M+H]^+$**	**MONO:tri+1A+PO_4**	2
d	435.70080 (0.55 ppm)	$[M+2H]^{2+}$	MONO:tri+1A	6
	828.38349 (0.25 ppm)	**$[M+H]^+$**	**MONO:tri+1A–1Ac**	4, 5
	667.31475 (0.40 ppm)	$[M+H]^+$	MONO:tri–Glc*N*Ac	3
	854.89672 (0.19 ppm)	$[M+2H]^{2+}$	DI:tri-tetra+2A–2Ac	23
e	899.42075 (0.39 ppm)	$[M+H]^+$	MONO:tetra+1A–1Ac	9
	855.38782 (−0.88 ppm)	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A–2Ac**	28
	536.25568 (−1.03 ppm)	$[M+2H]^{2+}$	MONO:tri-di+1A–1Ac	8
f	546.75630 (−0.18 ppm)	$[M+2H]^{2+}$	ANHYDRO MONO:tri-di+2A	7
	556.76994 (0.77 ppm)	**$[M+2H]^{2+}$**	**MONO:tri-di+2A**	7
	909.45229 (−0.09 ppm)	$[M+H]^+$	MONO:tri:di+2A–Glc*N*Ac	7
g	**557.26178 (0.47 ppm)**	**$[M+2H]^{2+}$**	**MONO:tri:di+1A**	8
h	785.84211 (0.06 ppm)	$[M+2H]^{2+}$	Glufib (internal standard 2)	-
	855.38848 (−0.11 ppm)	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A–2Ac**	28
N.D[c]	773.344 (0.25 ppm)	$[M+2H]^{2+}$	Di:tri-tetra+1A–1Ac–Glc*N*Ac	27
i	646.79600 (−0.39 ppm)	$[M+2H]^{2+}$	ANHYDRO MONO:tri-tetra+2A	11
	656.80953 (0.26 ppm)	**$[M+2H]^{2+}$**	**MONO:tri-tetra+2A**	12
	671.81496 (0.48 ppm)	$[M+2H]^{2+}$	MONO:tetra-tetra+1A–1Ac	13
j	415.18747 (0.43 ppm)	$[M+2H]^{2+}$	MONO:tri–1Ac	7
	657.30186 (0.75 ppm)	$[M+2H]^{2+}$	MONO:tri-tetra+1A	15, 16
	774.36132 (−1.07 ppm)	$[M+2H]^{2+}$	DI:tri-tetra+2A–1Ac–Glc*N*Ac	22
	795.36698 (−0.58 ppm)	$[M+2H]^{2+}$	DI:tri-tetra+2A–Glc*N*Ac	22
	865.88827 (−0.53 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tri-tetra+2A–1Ac	20
	875.90198 (0.16 ppm)	**$[M+2H]^{2+}$**	**DI:tri-tetra+2A–1Ac**	24
	911.42011 (−0.32 ppm)	$[M+2H]^{2+}$	DI:tetra-tetra+2A–1Ac–Glc*N*Ac	33
N.D[c]	886.392 (0.58 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tetra-tetra+2A	31
	886.885 (2.26 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tetra-tetra+1A	32
k	556.76972 (0.38 ppm)	$[M+2H]^{2+}$	MONO:tri-di+2A	7
	876.39420 (0.40 ppm)	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A–1Ac**	2
l	699.29292 (−0.44 ppm)	$[M+H]^+$	MONO:di	1, 14
	657.28291 (0.61 ppm)	**$[M+H]^+$**	**MONO:di–1Ac**	1, 14
m	**557.26127 (−0.45 ppm)**	**$[M+2H]^{2+}$**	**MONO:tri-di+1A**	8
	621.29089 (0.14 ppm)	$[M+2H]^{2+}$	MONO:tri-tri+2A[d]	N.D[c]
n	**667.31474 (0.39 ppm)**	**$[M+H]^+$**	**MONO:tri+1A–1Ac**	3
o	850.36797 (0.40 ppm)	$[M+H]^+$	ANHYDRO MONO:tri+1A	6
	870.39399 (0.16 ppm)	**$[M+H]^+$**	**MONO:tri+1A**	6
	667.31500 (0.78 ppm)	$[M+H]^+$	MONO:tri+1A–Glc*N*Ac	3
	657.30091 (−0.70 ppm)	$[M+2H]^{2+}$	MONO:tri-tetra+1A	15,16
p	921.40435 (−0.43 ppm)	$[M+H]^+$	ANHYDRO MONO:tetra+1A	15
	941.43008 (−0.93 ppm)	**$[M+H]^+$**	**MONO:tetra+1A**	10
	506.73774 (0.12 ppm)	$[M+2H]^{2+}$	MONO:penta+1A	18, 19
	657.30124 (−0.20 ppm)	$[M+2H]^{2+}$	MONO:tri-tetra+1A	15, 16
q	**876.39376 (−0.10 ppm)**	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A–1Ac**	29
r	**942.41422 (−0.81 ppm)**	**$[M+H]^+$**	**MONO:tetra**	17
s	**692.81977 (−0.23 ppm)**	**$[M+2H]^{2+}$**	**MONO:tetra-tetra–1Ac**	N.D[c]
t	**896.90766 (0.60 ppm)**	**$[M+2H]^{2+}$**	**DI:tri-tetra+2A**	26
	795.36751 (0.09 ppm)	$[M+2H]^{2+}$	DI:tri-tetra+2A–Glc*N*Ac	22
	886.89410 (0.09 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tri-tetra+2A	21
	1195.54096 (−1.87 ppm)	$[M+3H]^{3+}$	TETRA:tetra-tetra-tetra-tetra+2A–3Ac	39
u	**967.94450 (0.27 ppm)**	**$[M+2H]^{2+}$**	**DI:tetra-penta+2A**	N.D[c]
	932.42428 (−1.50 ppm)	$[M+2H]^{2+}$	DI:tetra-tetra+2A	33
v	**897.39945 (0.36 ppm)**	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A**	30
	887.38506 (−0.61 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tri-tetra+1A	25

(*Continued*)

Table 1. Continued

CZE peak (Fig. 2B)	Experimental m/z (mass accuracy) [a]	Charge state	Muropeptide[b]	Corresponding peak in RP-HPLC (Fig. 2A)
w	**871.37811 (0.29 ppm)**	**$[M+H]^+$**	**MONO:tri**	7
	668.29848 (−0.01 ppm)	$[M+H]^+$	MONO:tri–Glc*N*Ac	7
	851.35125 (−0.47 ppm)	$[M+H]^+$	ANHYDRO MONO:tri	7
x	**897.39893 (-0.22 ppm)**	**$[M+2H]^{2+}$**	**DI:tri-tetra+1A**	30
	887.38453 (−1.68 ppm)	$[M+2H]^{2+}$	ANHYDRO DI:tri-tetra+1A	25
y	**699.29307 (−0.66 ppm)**	**$[M+H]^+$**	**MONO:di**	1, 14
	496.21372 (0.04 ppm)	$[M+H]^+$	MONO:di–Glc*N*Ac	1, 14
	679.26692 (0.09 ppm)	$[M+H]^+$	ANHYDRO MONO:di	1, 14
N.D[c]	838.072 (0.64 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+3A–Glc*N*Ac	34
	770.688 (1.77 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+2A–2Glc*N*Ac	34
	824.046 (−2.71 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+3A–1Ac –Glc*N*Ac	34
	891.739 (−2.64 ppm)	**$[M+3H]^{3+}$**	**TRI:tri-tetra-tetra+3A–1Ac**	34
	877.735 (−3.23 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+3A–2Ac	34
	905.739 (−6.49 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+3A	34
N.D[c]	838.384 (4.90 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+2A–Glc*N*Ac	35, 36
	824.378 (1.98 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+2A–1Ac–Glc*N*Ac	35, 36
	892.068 (−1.52 ppm)	**$[M+3H]^{3+}$**	**TRI:tri-tetra-tetra+2A–1Ac**	35
	878.067 (1.31 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+2A–2Ac	35, 36
	906.078 (5.64 ppm)	**$[M+3H]^{3+}$**	**TRI:tri-tetra-tetra+2A**	36
N.D[c]	838.705 (−3.44 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+1A–Glc*N*Ac	37
	824.706 (1.97 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+1A–1Ac –Glc*N*Ac	37
	892.396 (−1.53 ppm)	$[M+3H]^{3+}$	TRI:tri-tetra-tetra+1A–1Ac	37, 38
	878.393 (-0.96 ppm)	**$[M+3H]^{3+}$**	**TRI:tri-tetra-tetra+1A–2Ac**	37
	906.406 (5.63 ppm)	**$[M+3H]^{3+}$**	**TRI:tri-tetra-tetra+1A**	38

a) Exact masses from the LTQ-FT-ICR provided by default. If not the exact masses from the QTOF Synapt G2 is provided instead.
b) See Figure 1 for the peptidoglycan composition
c) N.D for not detected
d) Disaccharide tripeptide – tripeptide (*m*-A_2pm-*D*-Ala-*m*-A_2pm-*D*-iGlu) with 2 amidations (fraction "m")
Muropeptides printed in **bold** refer to most prominent variant in the peak. –Ac stands for deacetylation, +A stands for amidation and -Glc*N*Ac for *N*-acetylglucosamine loss. Experimental *m/z* for CZE-MS were measured by the LTQ FT-ICR Ultra Hybrid mass spectrometer (typical mass accuracy below 2 ppm) while those measured for RP-HPLC were measured on the SYNAPT G2 HDMS mass spectrometer (ESI-MS).

confident identification of all detected muropeptides thanks to the high mass accuracy (mainly <1 ppm, see Table 1 for details). During this work, the detected muropeptide modifications by CZE-ESI-MS were identical to those observed using RP-HPLC-MS, except the presence of the phosphate group. Such species should probably carry a global negative charge providing a net negative electrophoretic mobility. The use of a (dynamic) coating providing a higher electro-osmotic flow should allow the detection of these types of PGN fragments, even in acidic conditions. Unlike the RP-HPLC separation, the chemically modified muropeptides were baseline separated from their unmodified counterparts (except for the anhydromuropeptides) due to the change in charge and hydrodynamic radius of the analyte. The detailed discussion of the characterization of such modified species is provided in the next section.

A total of 49 chemically different species were detected, among which 46 were common to RP-HPLC. Indeed, the disaccharide tripeptide – tripeptide with two amidations, the deacetylated disaccharide tetrapeptide – tetrapeptide and the disaccharide tetrapeptide – disaccharide pentapeptide with two amidations were not found with the RP-HPLC reference method (see Table 1).

3.3 PGN chemical modifications characterization by CZE-ESI-MS

As chemical modifications such as amidation and deacetylation of PGN cause a change in the average charge state in solution or in the mass of the concerned muropeptide, CZE is a particularly well suited separation technique for their characterization. Indeed, the amidation of *m*-A_2pm and the deacetylation of the glycan moiety cause an increase in the muropeptide positive average charge under acidic separation conditions. The major detected structural modification was the *m*-A_2pm amidation. As seen in Table 1, the muropeptides were frequently detected during this work with a partial or full *m*-A_2pm amidation. When the muropeptide consists of at least a dimer, single or multiple amidation(s) of the *m*-A_2pm were detected. Interestingly, when the muropeptide possesses at least two *m*-A_2pm within its peptide moiety, various possible amidation sites were available and consequently, several amidation isomers could be formed. The behaviour of migration is affected by the number of carboxylic moieties from the *m*-A_2pm. Higher the rate of amidation of the *m*-A_2pm, higher is the apparent electrophoretic mobility and lower the MT. This effect is due to the conversion of a partially

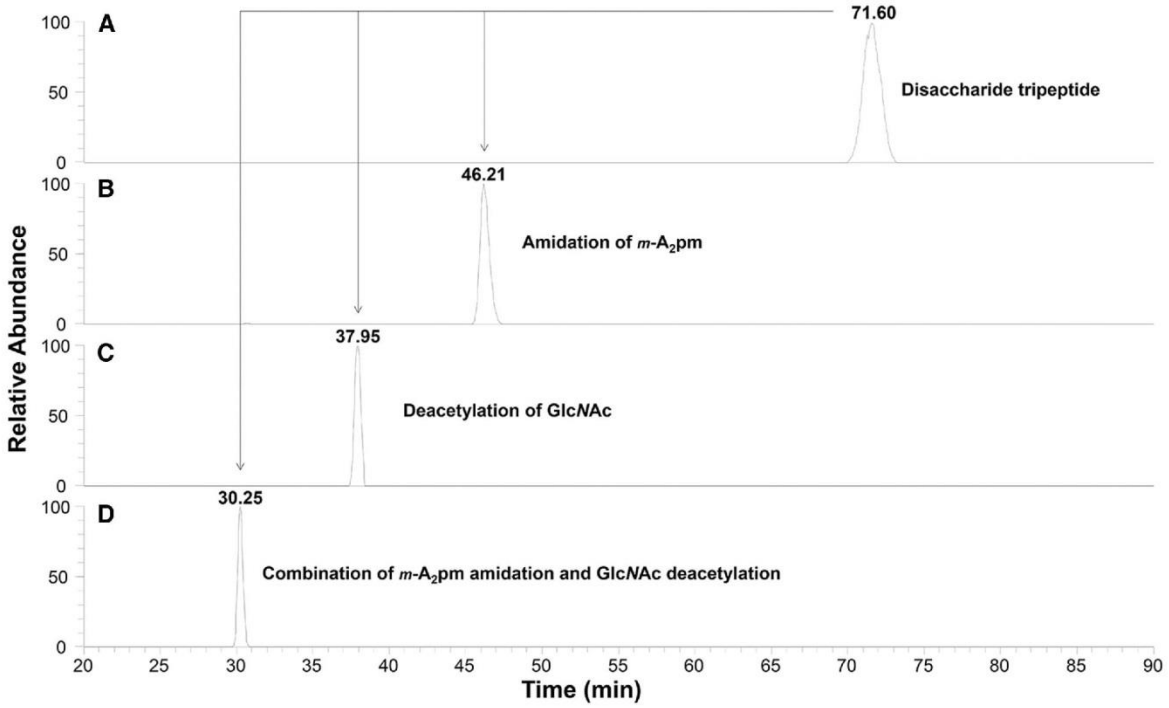


Figure 3. Comparison of the disaccharide tripeptide migration times depending on its modifications. (A) Extracted electropherogram of the disaccharide tripeptide (871.38 *m/z*); (B) Extracted electropherogram of the disaccharide tripeptide with one *m*-A_2pm amidation (870.39 *m/z*) (C) Extracted electropherogram of the disaccharide tripeptide with the loss of the Glc*N*Ac acetyl group (829.37 *m/z*); (D) Extracted electropherogram of the disaccharide tripeptide with one *m*-A_2pm amidation and the loss of the Glc*N*Ac acetyl group (828.38 *m/z*).

negatively charged carboxylic group at the pH of separation (2.38) into a non-ionizable functional group in solution after amidation. The separation and identification of such species is discussed in the next section.

The second major modification highlighted by this work was an acetyl group loss on the glycan moiety of the muropeptide. Similarly to the amidation, the deacetylation was observed in a relevant number of the detected muropeptides. When the muropeptide consists of at least a dimer, mono-deacetylated, and multi-deacetylated species were detected. Once again, higher the rate of deacetylation of the glycosidic moieties, higher is the apparent electrophoretic mobility and lower the MT. Indeed, the deacetylation led to the recovery of a basic amino group that is almost fully positively charged in the background electrolyte used during this work. In addition, experimental data showed that deacetylation could also occur during the ESI process, yielding to the detection of deacetylated species at the same migration time than the natively acetylated one (see the example in Supporting Information Figure 2 of peak "b" (from Table 1)).

Figure 3 shows the respective influences of the amidation of *m*-A_2pm, the deacetylation, and the combination of both modifications on the detected migration time, compared to the unmodified muropeptide in the case of the disaccharide tripeptide.

As seen in Fig. 3, the migration time is more affected by the deacetylation ($\Delta_{MT} = -33.65$ min) than by the *m*-A_2pm amidation ($\Delta_{MT} = -25.39$ min). Finally, as could be expected, the combination of both had the largest influence ($\Delta_{MT} = -41.35$ min). The increase of the muropeptide average charge in solution is due to either by removing a partial negative charge (amidation) or by adding a positive charge on the analyte (deacetylation), leading to a faster migration during the electrophoretic process. The pK_a value for the carboxylic group of *m*-A_2pm described in the literature is quite dispersed with values ranging from 1.85 (theoretical prediction, https://www.drugbank.ca/drugs/DB03590) to 3.75 (experimental work [23]). In this case, the ratio of the migration times of the disaccharide tripeptide bearing no or one amidation combined to the use of the Henderson-Hasselbalch equation yields to an average dissociation coefficient ($\alpha \approx 0.645$) for the carboxylic acid of the *m*-A_2pm, consistent with a computed pK_a value of about 2.1. For all stated pK_a values, a partial deprotonation of the most acidic carboxylic group of *m*-A_2pm occurs at the separation pH of 2.38.

Besides the amidation and deacetylation, some muropeptides showed the loss of a Glc*N*Ac subunit or the loss of a water molecule on the Mur*N*Ac group (anhydromuropeptides). The example of such modifications on the disaccharide tripeptide with one amidation is given in Fig. 4.

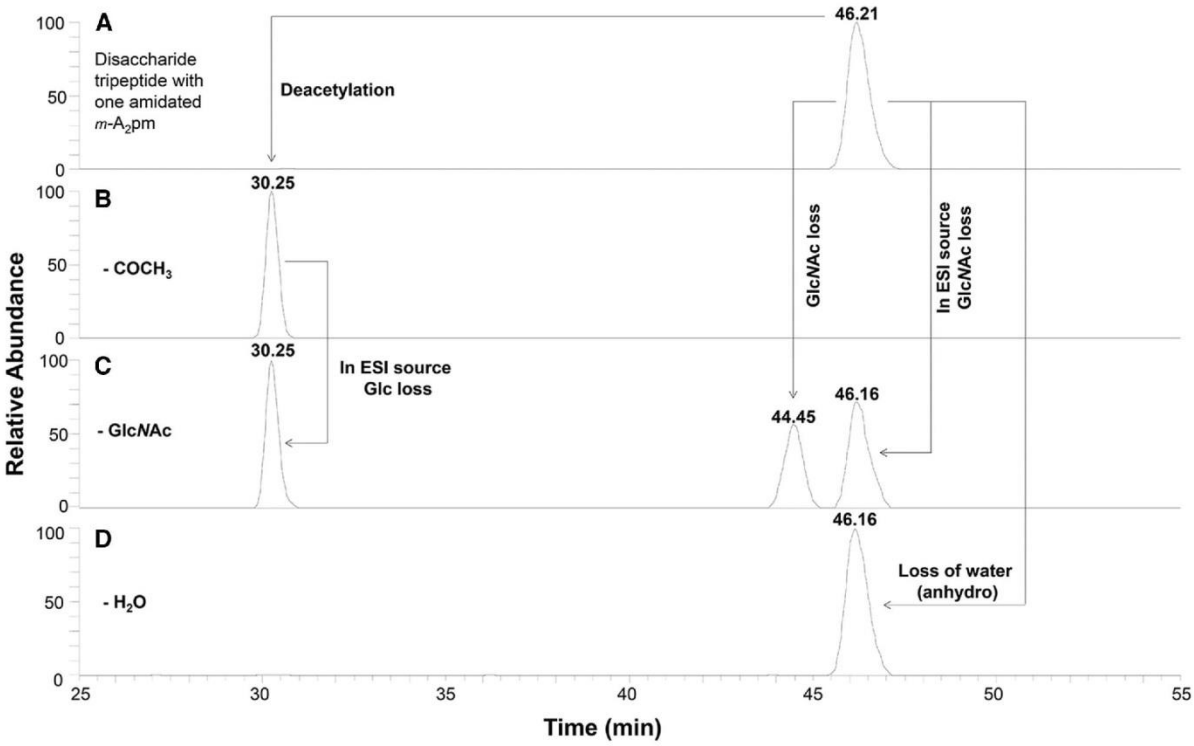


Figure 4. Influence of the loss of Glc*N*Ac and water moieties on the migration time of the disaccharide tripeptide. (A) Extracted electropherogram of the disaccharide tripeptide with one *m*-A_2pm amidation (870.39 *m/z*,); (B) Extracted electropherogram of the disaccharide tripeptide with one *m*-A_2pm amidation and one acetyl missing (828.37 *m/z*); (C) Extracted electropherogram of the disaccharide tripeptide with one *m*-A_2pm amidation and one Glc*N*Ac missing (667.31 *m/z*); (D) Extracted electropherogram of the anhydro-disaccharide tripeptide with one *m*-A_2pm amidation (850.37 *m/z*).

Figure 4 shows that the loss of a Glc*N*Ac subunit could be generated by different mechanisms. First, the presence of such species at nearly identical migration times compared to the amidated (46.21 min) and deacetylated counterparts (30.25 min) suggests that the loss of the labile Glc*N*Ac (or remaining Glc after the loss of the acetyl group) subunit could take place during the ESI process. This finding was in accordance with previously published results [13]. Secondly, the presence of the muropeptide with one amidation missing a Glc*N*Ac detected at a slightly reduced migration time compared to the corresponding amidated one (44.45 min versus 46.16 min) indicates that the loss of Glc*N*Ac could also take place in the sample prior to the electrophoretic separation. Finally, some muropeptides were also detected as anhydromuropeptides ($\Delta_{m/z} = -20.026$ Da as water and H_2 loss for the theoretical monoisotopic mass). In all cases, these modified muropeptides had an identical migration time than the unmodified species, suggesting that the water loss either did not influence the migration time of the analyte or took place during the ESI process. Nevertheless, the formation of anhydromuropeptides was a low abundant modification as it was systematically observed in the MS spectrum as small satellite peaks next to their unmodified counterparts (e.g., the peak in Fig. 4D is about 15 times less abundant than the peak in Fig. 4A).

3.4 The discrimination of the amidation site by CZE-ESI-MS^3

When a muropeptide possesses at least two *m*-A_2pm within its peptide structure, various possible amidation sites are available (see Fig. 1 panel A). Consequently, several amidation isomers can be formed, with specific amidation patterns and possibly different migration times for each isomer. This hypothesis was univocally tested on the disaccharide tripeptide dipeptide bearing one amidation. To this end, the co-eluting disaccharide tripeptide dipeptide bearing none, one or two amidations were isolated by RP-HPLC (see Supporting Information) and analysed using CZE-ESI-MS^3 with a focus on the disaccharide tripeptide dipeptide bearing one amidation. Note that the tandem MS experiments were also successfully performed by data dependent acquisition from the total PGN processed sample instead of fractions coming from chromatography. This species is represented in Fig. 5 with the two possible amidation sites (due to the presence of two *m*-A_2pm within the structure) located either on the tripeptide or on the dipeptide. Interestingly, the two amidation isomers were baseline separated by CZE (Fig. 5A) with a pronounced difference in migration times (28.25 min versus 37.18 min), which is remarkably different compared to their co-elution in RP-HPLC.

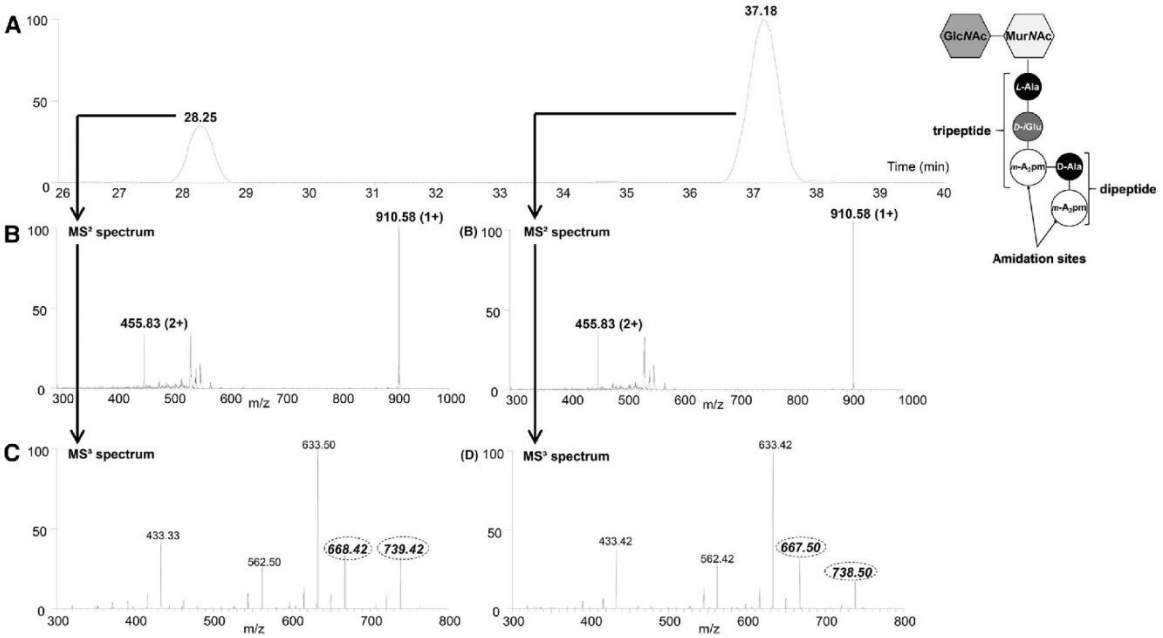


Figure 5. Amidation site identification on the disaccharide tripeptide dipeptide by MS^3 fragmentation. (A) Extracted electropherogram of the disaccharide tripeptide dipeptide with one *m*-A_2pm amidation (556.26 *m/z*, +2 charge state); (B) CID MS/MS spectra of the disaccharide tripeptide dipeptide with one *m*-A_2pm amidation show one dominant fragment ion corresponding to the loss of the labile Glc*N*Ac subunit for both isomers (910.58 *m/z*, +1 charge state; 455.83 *m/z*, +2 charge state); (C) MS^3 spectrum of the 28.25 min peak (transition 556.26 → 910.58 → fragments); (D) MS^3 spectrum of the 37.18 min peak (transition 556.26 → 910.58 → fragments)
633.50 *m/z*: *L*-Ala-*D*-iGlu-*m*-A_2pm*-*D*-Ala-*m*-A_2pm or *L*-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm*
562.50 *m/z*: *D*-iGlu-*m*-A_2pm*-*D*-Ala-*m*-A_2pm or *D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm*
433.50 *m/z*: *m*-A_2pm*-*D*-Ala-*m*-A_2pm or *m*-A_2pm-*D*-Ala-*m*-A_2pm*
668.42 *m/z*: Mur*N*Ac- *L*-Ala- *D*-iGlu- *m*-A_2pm
739.42 *m/z*: Mur*N*Ac- *L*-Ala- *D*-iGlu- *m*-A_2pm- *D*-Ala
667.50 *m/z*: Mur*N*Ac- *L*-Ala- *D*-iGlu- *m*-A_2pm*
738.50 *m/z*: Mur*N*Ac- *L*-Ala- *D*-iGlu- *m*-A_2pm*- *D*-Ala
where * represents an amidation on the *m*-A_2pm

Using online CZE-CID MS/MS, the fragmentation data obtained did not allow the unequivocal identification of the amidation site within the two isomers. Indeed, the CID MS/MS spectra were dominated by the loss of the labile Glc*N*Ac subunit in both isomers without any additional characteristic product ion (Fig. 5B). Therefore, online CZE-CID MS^3 experiments were performed to further fragment these muropeptides. The MS^3 spectra showed a clear difference with the presence of three common and two characteristic fragment ions in each case (Fig. 5C and D), allowing the discrimination of the amidation isomers. In the MS^3 spectrum corresponding to the peak detected at 28.25 min, the characteristic fragments confirmed that the amidation site is located on the "dipeptide" part of the muropeptide, therefore allowing to structurally assign this species as Glc*N*Ac-Mur*N*Ac-L-Ala-D-iGlu-*m*-A_2pm-D-Ala-*m*-A_2pm* (* denote the amidation of the *m*-A_2pm). Alternatively, in the MS^3 spectrum at 37.18 min, the characteristic fragment ions confirmed that the amidation site was located on the "tripeptide" part, leading to the structure Glc*N*Ac-Mur*N*Ac-L-Ala-D-iGlu-*m*-A_2pm*-D-Ala-*m*-A_2pm. These results clearly showed that the online combination of CZE-ESI-MS and CID MS^3 allows the unambiguous assignation of the amidation site between different amidation isomers. Moreover, the observed migration time difference between both amidation isomers could originate from the interaction between the *m*-A_2pm primary amine and the *D-i*Glu carboxylic acid. Indeed, when these two are close within the structure (amidation present on the dipeptide), their interaction increases the muropeptide partial positive charge (by hiding the partial negative charge carried by the carboxylic group) and leads to a faster migration velocity. Adversely, when the amidation is located on the tripeptide, the *D-i*Glu carboxylic acid from the tripeptide is not sufficiently close to the *m*-A_2pm primary amine located on the dipeptide and does not lead to any significant interaction, which results in a slower migration velocity.

Besides the CID-based fragmentation, ETD was also performed for structural elucidation purpose (see Supporting Information Fig. 3). To this end, an online CZE-ETD MS/MS experiment was also tested on the disaccharide tripeptide dipeptide bearing one amidation. Contrary to CID-MS/MS data, ETD-MS/MS spectra were much more informative with the presence of specific peptide moiety fragments and fragments retaining the disaccharide moiety after peptide

backbone fragmentation (see Supporting Information). Unfortunately, in the present experimental conditions, ETD spectra could only be recorded on one of the amidation isomer due to insufficient intensity of the 28.25 min peak. However, the ETD spectra recorded on the second amidation isomer (37.18 min) showed the great potential of ETD MS/MS spectra for the amidation isomers characterization in association with the CID-MS^3 results. Indeed, as CID requires up to MS^3 to provide relevant structural information (unambiguous position of the amidation position), ETD MS/MS is a convenient fragmentation alternative to access this information. Combining ETD-MS/MS and CID-MS^2 or MS^3 allowed structural elucidation of amidation and acetylation of the enzymatically digested PGN samples.

4 Concluding remarks

In conclusion, 62 chemically different muropeptides were detected in RP-HPLC ranging from monomers to a tetramer thanks to the use of the combination of separation data and MS instruments. Several by-products of the tetramer species were produced during the sample preparation that were detected by RP-HPLC with offline MS detection and during CZE-MS experiments. In CZE, 49 unique species were observed, essentially monomers and dimers but also a unique tetramer. Interestingly, the RP-HPLC allowed the identification of a trimer containing a phosphate moiety, which was not detected using CZE, most probably caused by the neutral or slightly negative charge of this muropeptide (i.e., having a negative effective mobility) despite the CZE acidic conditions. As this trimer presents multiple modifications (including various amidations and deacetylations), this leads to the large difference between the total number of observed modified species between RP-HPLC and CZE. Nevertheless, the CZE showed different migration times depending on the modifications of the muropeptide while the modified and unmodified species were often co-eluted in RP-HPLC. This suggests that LC is an adequate tool for global PGN characterization (with information about the muropeptides cross-linking level) but CZE provides more complete information with both the PGN composition and the separation of amidated and deacetylated muropeptides from their unmodified counterparts. In addition, the use of CZE-CID-MS^3 allowed the identification and localization of the amidation position in the case of isomers. Finally, our work also shows that the muramic acid reduction with $NaBH_4$ is not mandatory prior to the CZE analysis, which extensively reduces the sample preparation compared to the LC sample preparation.

As perspective, the developed CZE separation method could be of particular (biochemical) interest to better understand the PGN modifications of bacterial strains induced under stress conditions such as the exposition to antibiotics or the adaptation of bacterial PGN composition as a defence against the host immune system [24, 25]. For example, methicillin-resistant *S. aureus*, one of the most important human Gram-positive bacterial pathogen, presents an amidated *D*-Glu which leads to beta-lactam resistance [26,27]. The comparison of these PGN modifications requires the development of robust quantification methods, which can be performed by the addition of internal standards for electropherogram realignment (as performed during this work). The absolute quantification of the modified muropeptides by CZE-MS is also possible by the use of isotopically labelled standards that are not yet available, mainly due to the presence of the diaminopimelic acid amino acid into the muropeptide chains, which is still challenging to synthesize. The method duration could also be decreased by the use of a negatively charged coating (e.g., double coating strategy) to accelerate the EOF towards the mass spectrometer without impairing the separation efficiency of the muropeptides. Besides, the neutral and negatively charged analytes might become accessible within a single run, providing a more complete PGN characterization. Note that nowadays, modern Q-TOF coupled to CZE has sufficient performance in terms of sensitivity, dynamic range and mass accuracy to provide similar quality data than our study. Nevertheless, the use of CID MS^3 and optionally ETD MS/MS or ECD MS/MS (e.g., Orbitrap-based instruments and FT-ICR instruments) with exact mass capability will be mandatory to achieve the unambiguous localization of the muropeptides amidation (in the case of possible amidation isomers) in the absence of authentic standards.

M.B has a PhD fellowship from the joint Research Action (ARC NetRBI) of the French Community (Belgium). ANALIS is also acknowledged for sharing the CZE apparatus and CEMS interface with technical supports.

The authors have declared no conflict of interest.

5 References

[1] Rogers, H. J., *Bacteriol. Rev.* 1970, *34*, 194–214.

[2] Vollmer, W., Blanot, D., De Pedro, M. A., *FEMS Microbiol. Rev.* 2008, *32*, 149–167.

[3] Vollmer, W., *FEMS Microbiol. Rev.* 2008, *32*, 287-306.

[4] Vollmer, W., Joris, B., Charlier, P., Foster, S., *FEMS Microbiol. Rev.* 2008, *32*, 259–286.

[5] Bernard, E., Rolain, T., Courtin, P., Guillot, A., Langella, P., Hols, P., Chapot-Chartier, M. P., *J. Biol. Chem.* 2011, *286*, 23950–23958.

[6] Peltier, J., Courtin, P., El Meouche, I., Lemée, L., Chapot-Chartier, M. P., Pons, J. L., *J. Biol. Chem.* 2011, *286*, 29053–29062.

[7] Bernard, E., Rolain, T., Courtin, P., Hols, P., Chapot-Chartier, M. P., *J. Bacteriol.* 2011, *193*, 6323–6330.

[8] Girardin, S. E., Travassos, L. H., Hervé, M., Blanot, D., Boneca, I. G., Philpott, D. J., Sansonetti, P. J., Mengin-Lecreulx, D., *J. Biol. Chem.* 2003, *278*, 41702–41708.

[9] Bugg, T. D. H., Braddick, D., Dowson, C. G., Roper, D. I., *Trends Biotechnol.* 2011, *29*, 167–173.

[10] Amoroso, A., Boudet, J., Berzigotti, S., Duval, V., Teller, N., Mengin-Lecreulx, D., Luxen, A., Simorre, J. P., Joris, B., *PLoS Pathog.* 2012, *8*, e1002571.

[11] Johnson, J., Fisher, J., Mobashery, S., *Ann. N. Y. Acad. Sci.* 2013, *1277*, 54–75.

[12] Glauner, B., *Anal. Biochem.* 1988, *172*, 451–464.

[13] Bacher, G., Körner, R., Atrih, A., *J. Mass Spectrom.* 2001, *36*, 124–139.

[14] Xu, N., Huang, Z.-H., de Jonge, B. L. M., Gage, D. A., *Anal. Biochem.* 1997, *248*, 7–14.

[15] Kühner, D., Stahl, M., Demircioglu, D.D., Bertsche, U., *Sci. Rep.*2014, *4*, 7494.

[16] Desmarais, S. M., Tropini, C., Miguel, A., Cava, F., Monds, R. D., De Pedro, M. A., Huang, K. C., *J. Biol. Chem.* 2015, *290*, 31090–31100.

[17] Young, K. D., *J. Bacteriol.* 1996, *178*, 3962-3966.

[18] Richards, E. M., Xing, D. K., *J. Pharm. Pharmacol.* 1994, *46*, 690–696.

[19] Atrih, A., Bacher, G., Allmaier, G., Williamson, M. P., Foster, S. J., *J. Bacteriol.* 1999, *181*, 3956–3966.

[20] Dougherty, T. J., *J. Bacteriol.* 1985, *163*, 69–74.

[21] Schleifer, K. H., Kandler, O., *Bacteriol. Rev.* 1972, *36*, 407–477.

[22] Atrih, A., Zöllner, P., Allmaier, G., Foster, S. J., *J. Bacteriol.* 1996, *178*, 6173–6183.

[23] Koch, A. L., Woeste, S., *J. Bacteriol.* 1992, *174*, 4811–4819.

[24] Boneca, I. G., Dussurget, O., Cabanes, D., Nahori, M.-A., Sousa, S., Lecuit, M., Psylinakis, E., Bouriotis, V., Hugot, J.-P., Giovannini, M., *Proc. Natl. Acad. Sci. USA* 2007, *104*, 997–1002.

[25] Münch, D., Sahl, H.-G., *Biochim. Biophys. Acta, Biomembr.* 2015, *1848*, 3062–3071.

[26] Figueiredo, T. A., Ludovice, A. M., Sobral, R. G., *Microbial Drug Resistance* 2014, *20*, 238–249.

[27] Leisico, F., Vieira, D., Figueiredo, T. A., Silva, M., Cabrita, E. J., Sobral, R. G., Ludovice, A. M., Trincão, J., Romão, M. J., Lencastre, H., *Sci. Rep.* 2018, *8*, 5313.

3.3. Results – Supporting Information

Bacillus licheniformis peptidoglycan characterization by CZE-MS: assessment with the benchmark RP-HPLC-MS method

Madeleine BOULANGER[1,+] and Cédric DELVAUX[2,+], Loïc QUINTON[2], Bernard JORIS[1], Edwin DE PAUW[2] and Johann FAR[2]*

[1] Center for Protein Engineering, InBioS Research Unit, Quartier Agora, University of Liège, Allée du Six Août 13, B-4000 Liège, Belgium

[2] Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium

[+] These authors equally contributed to this work

***Corresponding author: Johann.Far@uliege.be**

SUPPORTING INFORMATION

Table S1: SIMPLIFIED AND DETAILED NOMENCLATURES FOR THE MUROPEPTIDE IDENTIFICATION

Simplified nomenclature	Detailed nomenclature
MONO:di	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu
MONO:tri	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm
MONO:tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala
MONO:penta	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*D*-Ala
MONO:tri-di	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm
MONO:tri-tri	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm-*D*-iGlu
MONO:tri-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala
MONO:tetra-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm(-*D*-Ala)-*D*-Ala-*m*-A_2pm-*D*-iGlu -Ala
DI:tri-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac
DI:tetra-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-(-*D*-Ala)-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac
DI:tetra-penta	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm(-*D*-Ala-*D*-Ala)-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac
TRI:tri-tetra-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac | *D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac
TETRA:tetra-tetra-tetra-tetra	Glc*N*Ac-Mur*N*Ac-Ala-*D*-iGlu-*m*-A_2pm(-*D*-Ala)-*D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac | *D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac | *D*-Ala-*m*-A_2pm-*D*-iGlu-Ala-Mur*N*Ac-Glc*N*Ac

EXPERIMENTAL SECTION

Peptidoglycan isolation

During the growth exponential phase, *Bacillus licheniformis* cells were pelleted (4,300 g, 10 min, 4°C) and then resuspended in 20 mL of cold water. This suspension was added dropwise to 20 mL of a boiling solution of SDS/water (8/92, *w/v*) and boiled for one extra hour in a water bath. After an ultracentrifugation step (165,000 g, 60 min, 4°C), the insoluble material was washed with cold water until free of SDS. Covalently attached sugars and proteins were removed by successive treatments with α-amylase (final concentration of 200 µg.mL^{-1}, 16-20h at RT), trypsin (200 µg.mL^{-1}, 3h at 25°C) and pronase (500 µg.mL^{-1}, 3h at 25°C). The insoluble material was recovered by ultracentrifugation (160,000 g, 60 min, 4°C), washed once in water, and suspended in 20 mL of hydrofluoric acid (48/52, *v/v*) to remove the teichoic acids. The mixture was incubated at 4°C for 48 hours under stirring. The insoluble material was recovered by ultracentrifugation (165,000 g, 60 min, 4°C) and washed successively with 8 M LiCl (twice), 0.1 M EDTA (twice) and water (twice). The insoluble material was freeze-dried and stored at 4°C.

Before muramidase digestion, the PGN was weighted with an analytical balance and resuspended in water to obtain a final concentration of 11 mg of purified PGN per mL.

Amidation isomers isolation

In order to discriminate the amidation position on the *m*-A$_2$pm, the disaccharide tripeptide - dipeptide bearing one amidation was chosen. To this end, the disaccharide tripeptide - dipeptide bearing none, one or two amidations were collected by RP-HPLC between 28 and 30 minutes as they are co-eluted in HPLC. The collected fractions were pooled together, evaporated to dryness and suspended in 100 µL of 10 mM FA for the CZE-MS analysis.

HPLC – Mass Spectrometry detection

The MALDI-MS analyses were achieved on the MALDI-TOF/TOF UltrafleXtreme (Bruker Daltonics, Bremen, Germany) using the reflectron mode and a mass range of 500-4,000 *m/z* (external TOF calibration using the "peptide calibration standard II" (Bruker)). The samples were crystallized with 20 $mg.mL^{-1}$ of 2,5-dihydroxybenzoic acid matrix in acetonitrile/water (50/50, *v/v*) on an MTP AnchorChip 384 MALDI plate from Bruker. For the ESI-MS analyses, the Synapt G2 was used in positive ionization mode (+2.5 kV) and the mass range was fixed to 50-2,000 *m/z* using the resolution mode. The TOF mass spectrometer was externally calibrated using 10 µM sodium iodide in water/isopropanol (50/50, *v/v*)

The SYNAPT G2 HDMS data were reprocessed using MassLynx v4.1 and the MALDI UltrafleXtreme data were reprocessed using DataAnalysis v4.0.

Capillary activation and regeneration

The capillary was conditioned by flushing successively water (25 psi, 5 min), MeOH/water (50/50, *v/v)* (25 psi, 5 min), water (25 psi, 5 min), 0.1 M NaOH (25 psi, 5 min), water (25 psi, 5 min), 0.2 M HCl (25 psi, 10 min), water (25 psi, 10 min) and 100 mM FA (pH 2.38, BGE) (50 psi, 20 min). Prior to each run, the capillary was rinsed using BGE (25 psi, 5 min).

LTQ-FT-ICR and AmaZon speed ETD analysers parameters

All analyses were made in the positive ionization mode (+2.5 and +2.0 kV for LTQ-FT-ICR and AmaZon speed ETD analysers, respectively).

The FT-ICR mass range was fixed to 400-2,000 *m/z* for a resolving power of 50,000 (*@m/z*: 400) and was externally calibrated using the Thermo Scientific calibration mixture as recommended by the manufacturer. Data dependent MS^n (2 or 3) were performed using the linear trap quadrupole (LTQ). These data were collected using a generic CID fragmentation method with a normalized collision

energy fixed at 35% and a mass selection window of 3 amu. Data were reprocessed using XCalibur software v2.0.

The AmaZon speed ETD was used to perform ETD experiments with a 200-2,000 mass range. The acquisition of data was performed using the CZE-MS mode and the data dependent MS/MS and MS^3 were obtained with the default settings as recommended by the manufacturer, using either CID or ETD for ion activation. Data were reprocessed using DataAnalysis 4.0.

LIQUID CHROMATOGRAPHY AND CAPILLARY ZONE ELECTROPHORESIS DETAILS

Table S2: LC elution times

Peak[a]	Elution time (min)	Estimated conc. ($\mu g.mL^{-1}$)[b]	Peak[a]	Elution time (min)	Estimated conc. ($\mu g.mL^{-1}$)	Peak[a]	Elution time (min)	Estimated conc. ($\mu g.mL^{-1}$)[b]
1	19.69	7.89	14	34.50	4.27	27	45.91	128.46
2	23.14	2.17	15	35.79	70.80	28	47.33	11.00
3	24.16	134.54	16	36.77	28.59	29	47.78	12.73
4	25.00	4.57	17	39.48	1.91	30	48.25	83.08
5	26.06	6.71	18	40.04	1.76	31	49.12	11.39
6	27.08	362.05	19	40.88	2.47	32	49.54	13.86
7	29.02	92.91	20	41.69	57.42	33	50.21	17.73
8	29.91	55.87	21	41.99	57.26	34	52.03	81.10
9	30.95	8.31	22	42.73	68.03	35	52.78	16.69
10	31.19	5.45	23	43.01	69.45	36	54.11	29.56
11	32.51	3.10	24	43.92	64.12	37	58.96	23.42
12	33.27	4.37	25	44.07	25.614	38	59.81	5.74
13	34.18	5.70	26	44.96	457.79	39	62.97	5.96

a. Peak fraction numbering according to the Figure 2A in this article
b. The content of PGN fragments were estimated from the peak area of the UV trace at 201nm proportionally to the injected PGN quantity in the column (1 mg/100 µL) assuming a constant response factor of the detector. The volume of each fraction was 0.5 mL.

Table S3: CZE migration times

Peak[a]	Migration time (min)	Peak[a]	Migration time (min)
a	24.52	n	44.45
b	26.06	o	46.21
c	28.97	p	47.02
d	30.21	q	48.49
e	30.67	r	51.15
f	32.66	s	55.71
g	34.11	t	58.04
h	35.19	u	59.87
i	36.58	v	60.48
j	38.38	w	71.60
k	39.65	x	86.22
l	40.37	y	114.02
m	43.01		

a. Identified peak as shown in the Figure 2B in this article

Figure S1: MUROPEPTIDE ELUTION PATTERNS OF FIVE REPLICATES OF $NaBH_4$ REDUCED PEPTIDOGLYCAN FRAGMENTS OF *BACILLUS LICHENIFORMIS*

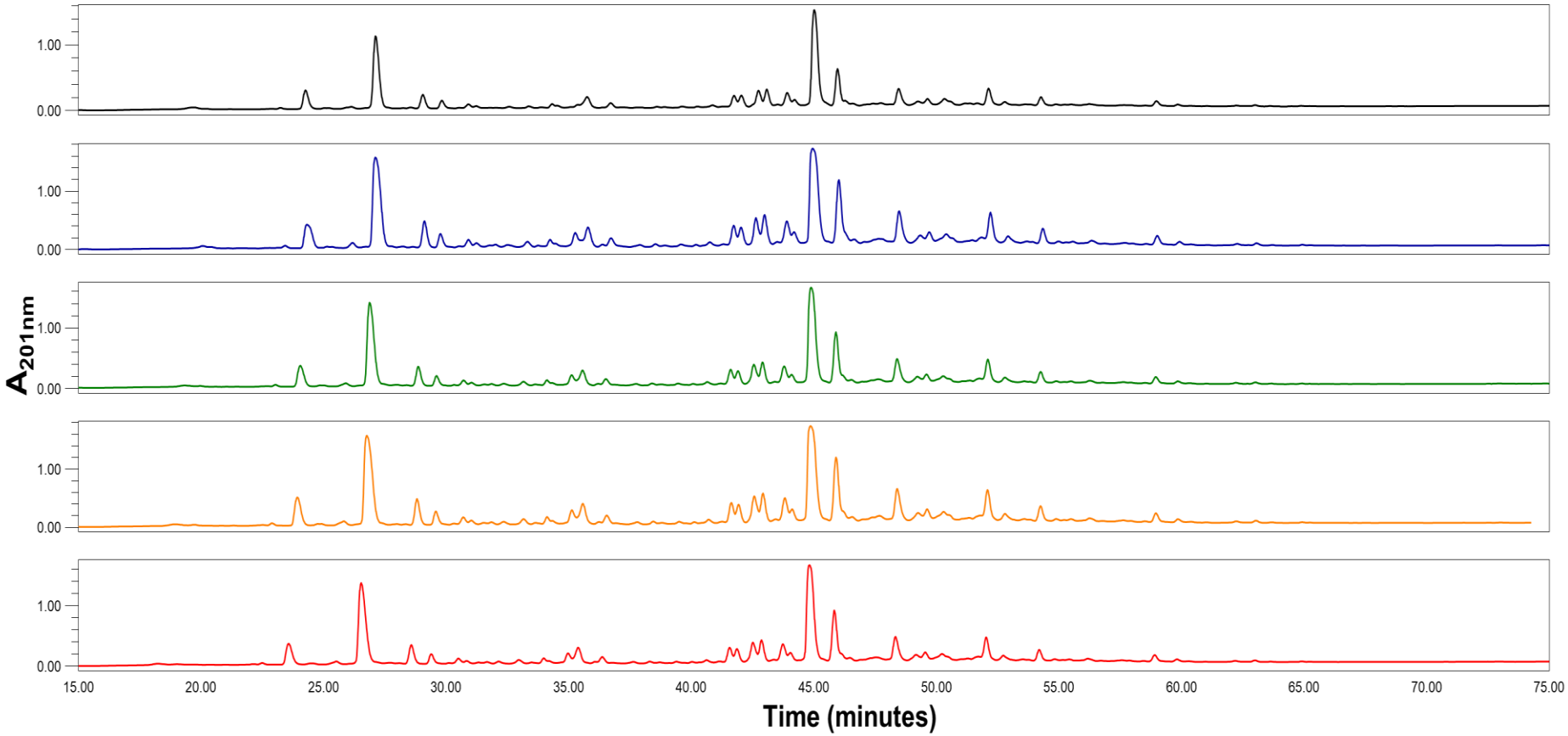


Table S4: DIFFERENCES BETWEEN MALDI AND ESI OFFLINE ANALYSES OF RP-HPLC PEAKS

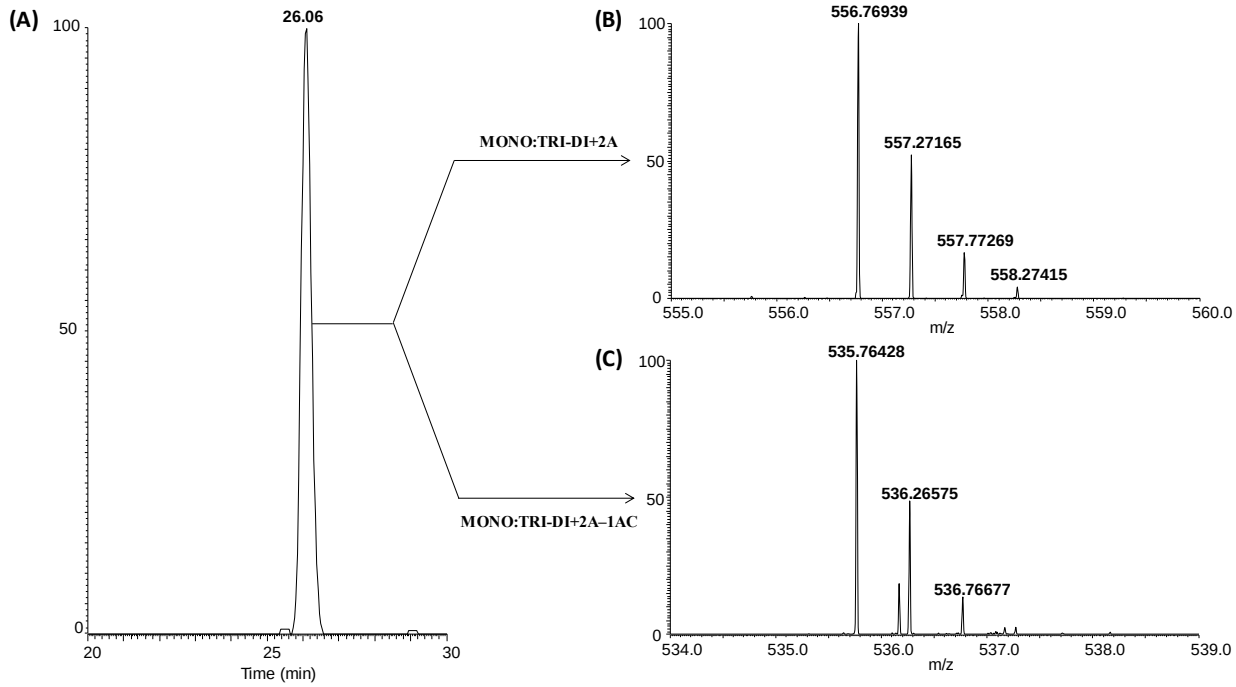
RP-HPLC peak	Muropeptides	Detected charge state in ESI	Detected charge state in MALDI
1	MONO:di–Glc*N*Ac	$[M+H]^+$	N.D.
	MONO:di–1Ac	$[M+H]^+$	$[M+H]^+$
	ANHYDRO MONO:di	$[M+H]^+$	$[M+H]^+$
	MONO:di	**$[M+H]^+$**	**$[M+Na]^+$**
2	**MONO:tri+1A+PO_4**	**$[M+H]^+$**	**$[M+Na]^+$**
	DI:tri-tetra+1A–1Ac	$[M+2H]^{2+}$	$[M+Na]^+$
3	MONO:tri–Glc*N*Ac	$[M+H]^+$	N.D.
	MONO:tri+1A–1Ac	$[M+H]^+$	$[M+Na]^+$
	MONO:tri+1A–Glc*N*Ac	**$[M+H]^+$**	**$[M+Na]^+$**
4	**MONO:tri+1A–1Ac**	**$[M+H]^+$**	**$[M+Na]^+$**
5	**MONO:tri+1A–1Ac**	**$[M+H]^+$**	**$[M+Na]^+$**
6	ANHYDRO MONO:tri+1A	$[M+H]^+$	$[M+H]^+$
	MONO:tri+1A	**$[M+H]^+$**	**$[M+Na]^+$**
7	MONO:tri–1Ac	$[M+2H]^{2+}$	$[M+Na]^+$
	MONO:tri–Glc*N*Ac	$[M+H]^+$	$[M+Na]^+$
	ANHYDRO MONO:tri	$[M+H]^+$	$[M+H]^+$
	MONO:tri	**$[M+H]^+$**	**$[M+Na]^+$**
	MONO:tri-di+2A	$[M+2H]^{2+}$	$[M+Na]^+$
	MONO:tri:di+2A–Glc*N*Ac	$[M+H]^+$	N.D.
	ANHYDRO MONO:tri-di+2A	$[M+2H]^{2+}$	N.D.
	MONO:tri-di+2A–1Ac	$[M+2H]^{2+}$	N.D.
8	MONO:tri-di+1A–1Ac	$[M+2H]^{2+}$	$[M+Na]^+$
	MONO:tri-di+1A	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
9	**MONO:tetra+1A–1Ac**	**$[M+H]^+$**	**$[M+Na]^+$**
10	**MONO:tetra+1A**	**$[M+H]^+$**	**$[M+Na]^+$**
11	**ANHYDRO MONO:tri-tetra+2A**	**$[M+2H]^{2+}$**	**N.D.**
12	**MONO:tri-tetra+2A**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
13	**MONO:tetra-tetra+1A–1Ac**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
14	MONO:di–Glc*N*Ac	$[M+H]^+$	N.D.
	MONO:di–1Ac	$[M+H]^+$	$[M+Na]^+$
	ANHYDRO MONO:di	$[M+H]^+$	$[M+H]^+$
	MONO:di	**$[M+H]^+$**	**$[M+Na]^+$**
15	**MONO:tri-tetra+2A–1Ac**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
	ANHYDRO MONO:tetra+1A	$[M+H]^+$	$[M+H]^+$
	MONO:tri-tetra+1A	$[M+2H]^{2+}$	$[M+Na]^+$
16	**MONO:tri-tetra+1A**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
17	**MONO:tetra**	**$[M+H]^+$**	**$[M+Na]^+$**
18	**MONO:penta+1A**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
19	**MONO:penta+1A**	**$[M+2H]^{2+}$**	**$[M+Na]^+$**
20	**ANHYDRO DI:tri-tetra+2A–1Ac**	**$[M+2H]^{2+}$**	**N.D.**

21	**ANHYDRO DI:tri-tetra+2A**	**$[M+2H]^{2+}$**	**N.D.**
22	DI:tri-tetra+2A–1Ac–Glc*N*Ac	$[M+2H]^{2+}$	$[M+Na]^{+}$
	DI:tri-tetra+2A–Glc*N*Ac	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
23	**DI:tri-tetra+2A–2Ac**	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
24	**DI:tri-tetra+2A–1Ac**	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
25	**ANHYDRO DI:tri-tetra+1A**	**$[M+2H]^{2+}$**	**N.D.**
26	**DI:tri-tetra+2A**	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
27	**Di:tri-tetra+1A–1Ac–Glc*N*Ac**	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
28	**DI:tri-tetra+1A–2Ac**	**$[M+2H]^{2+}$**	**N.D.**
29	**DI:tri-tetra+1A–1Ac**	**$[M+2H]^{2+}$**	**N.D.**
30	**DI:tri-tetra+1A**	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
31	**ANHYDRO DI:tetra-tetra+2A**	**$[M+2H]^{2+}$**	**N.D.**
32	**ANHYDRO DI:tetra-tetra+1A**	**$[M+2H]^{2+}$**	**N.D.**
33	DI:tetra-tetra+2A–1Ac–Glc*N*Ac	$[M+2H]^{2+}$	N.D.
	DI:tetra-tetra+2A	**$[M+2H]^{2+}$**	**$[M+Na]^{+}$**
34	TRI:tri-tetra-tetra+3A–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+2A–2Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+3A–1Ac –Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+3A–1Ac	**$[M+3H]^{3+}$**	**$[M+Na]^{+}$**
	TRI:tri-tetra-tetra+3A–2Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+3A	$[M+3H]^{3+}$	$[M+Na]^{+}$
35	**TRI:tri-tetra-tetra+2A–1Ac**	**$[M+3H]^{3+}$**	**$[M+Na]^{+}$**
	TRI:tri-tetra-tetra+2A–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+2A–1Ac–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+2A–2Ac	$[M+3H]^{3+}$	N.D.
36	**TRI:tri-tetra-tetra+2A**	**$[M+3H]^{3+}$**	**$[M+Na]^{+}$**
	TRI:tri-tetra-tetra+2A–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+2A–1Ac–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+2A–2Ac	$[M+3H]^{3+}$	$[M+Na]^{+}$
37	TRI:tri-tetra-tetra+1A–Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+1A–1Ac –Glc*N*Ac	$[M+3H]^{3+}$	N.D.
	TRI:tri-tetra-tetra+1A–2Ac	**$[M+3H]^{3+}$**	**$[M+Na]^{+}$**
	TRI:tri-tetra-tetra+1A–1Ac	$[M+3H]^{3+}$	N.D.
38	**TRI:tri-tetra-tetra+1A**	**$[M+3H]^{3+}$**	**$[M+Na]^{+}$**
	TRI:tri-tetra-tetra+1A–1Ac	$[M+3H]^{3+}$	N.D.
39	**TETRA:tetra-tetra-tetra-tetra+2A–3Ac**	**$[M+3H]^{3+}$**	**$[M+3H]^{3+}$**

N.D for not detected

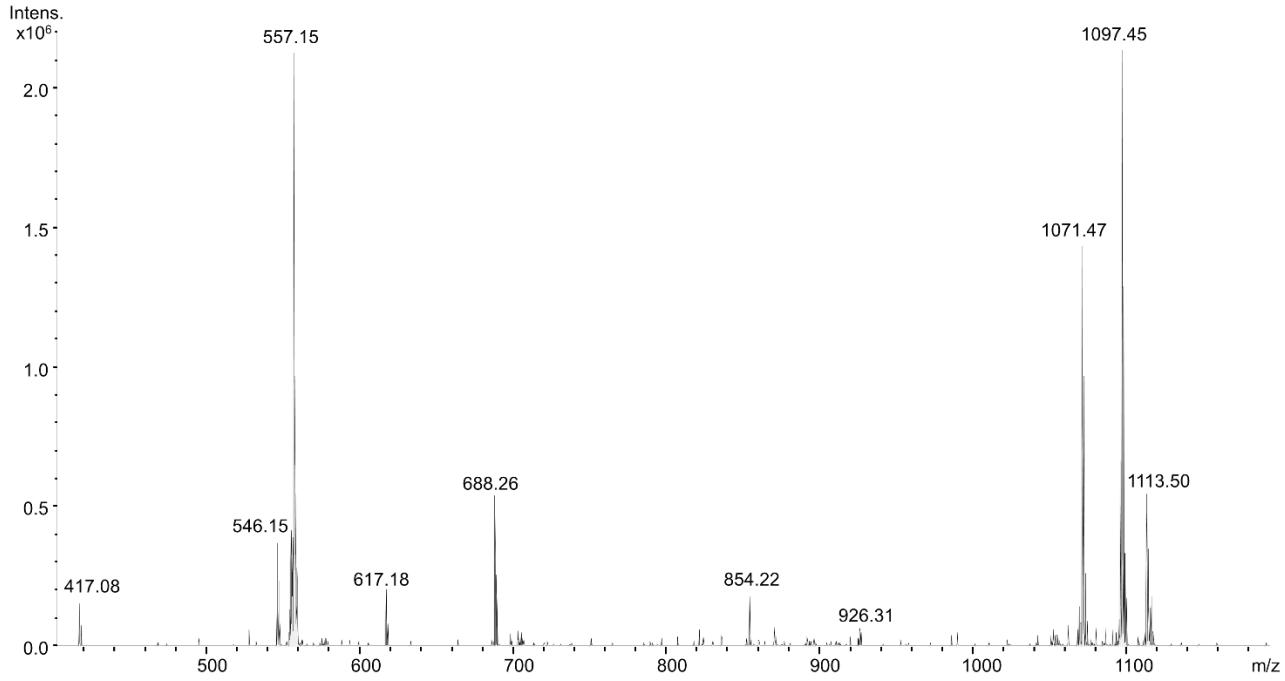
Muropeptides printed in **bold** refer to most prominent variant in the peak

Figure S2: IN-SOURCE DISSOCIATION PRODUCING THE DEACETYLATED FORM OF THE DISACCHARIDE TRIPEPTIDE - DIPEPTIDE BEARING 2 AMIDATIONS (MONO:TRI-DI+2A, SEE PEAK "B" FROM TABLE 1) OBTAINED DURING CZE-ESI-MS EXPERIMENTS



(A) Extracted electropherogram of peak "b" from Table 1; **(B)** mass spectrum at 26.06min showing the presence of disaccharide tripeptide – dipeptide bearing two amidations (*m/z:* 556.76939 as $[M+2H]^{2+}$) and **(C))** mass spectrum at 26.06min showing the presence of the in-source formation of the deacetylated disaccharide tripeptide – dipeptide bearing two amidations (*m/z:* 535.76428 as $[M+2H]^{2+}$). Loss of the acetyl group from Glc*N*Ac is 42.010 Da.

Figure S3: ETD SPECTRUM OF THE DISACCHARIDE TRIPEPTIDE - DIPEPTIDE WITH ONE AMIDATION



ESI- ETD MS^2 spectrum of the disaccharide tripeptide dipeptide with 1 amidation at 38.17 min (parent ions: 1113.50 *m/z*, +1 charge state; 557.15 *m/z*, +2 charge state). 1071.47 *m/z* is the loss of the remaining acetyl on the glycan moiety. 926.31 *m/z* is Glc*N*Ac-Mur*N*Ac-*L*-Ala-*D*-iGlu-*m*-A_2pm*-*D*-Ala. 854.22 *m/z* is Glc*N*Ac-Mur*N*Ac-*L*-Ala-*D*-iGlu-*m*-A_2pm*. 688.26 *m/z* and 617.18 *m/z* are characteristic of *L*-Ala-*D*-iGlu-*m*-A_2pm*-*D*-Ala-*m*-A_2pm. 546.15 *m/z* is *D*-iGlu-*m*-A_2pm*-*D*-Ala-*m*-A_2pm. 417.08 m/z is *m*-A_2pm*-*D*-Ala-*m*-A_2pm.

* for amidated *m*-A_2pm

3.4. Conclusion and perspectives

In this chapter, the use of Capillary Zone Electrophoresis coupled to Mass Spectrometry as a valuable alternative to the classical Reversed Phase Liquid Chromatography was demonstrated. The proposed CZE-MS provides several advantages.

First, the online coupling between CZE and MS avoids the tedious and time-consuming offline ESI-MS or MALDI-MS characterization required after the fractionation of the RP-HPLC eluent to identify the different muropeptides. Moreover, this work demonstrated that the tedious reduction step using $NaBH_4$ of the PGN fragment can be skipped for the CZE-MS characterization of these compounds.

Then, the proposed CZE-MS method baseline separates the two major types of variants of PGN muropeptides, i.e. muropeptides bearing an amidation on their *m*-A_2pm or muropeptides deacetylated on their glycan moiety. This result is in large contrast with the RP-HPLC separation where these species are most of the time co-eluted. Because these chemical variations are linked to bacterial antibiotics resistance and immune defense recognition, the proposed CZE-MS method in this chapter is of great biochemical interest.

Finally, the use of tandem mass spectrometry including Collision Induced Dissociation (CID) and Electron Transfer Dissociation (ETD) permits the fine structural assignment of muropeptides. This chapter showed that CZE can separate amidation isomers when multiple amidations sites are available (i.e. species with a specific pattern of amidation but an identical chemical composition). Again, this separation contrasts with the separation obtained by RP-HPLC, where such species are co-eluted. The use of ETD-MS^2/MS^3 or CID-MS^3 allowed the structural assignment of the position of the amidation within this kind of analyte when the intensity of the species is sufficient.

This work opens up exciting opportunities about the characterization of PGN and especially, the chemical modifications bacteria implement on their PGN in response to various external stresses, including exposure to antibiotics. This method could also help in understanding the chemical modifications that super-resistant bacteria provide to their PGN composition.

To fully benefit of the MS capabilities, the CZE-MS characterization of the PGN fragments could include an absolute quantification step based on the use of stable (heavy) isotope labelling similar to those commonly used for the absolute quantitation of peptides by LC-MS. This method has the potential to allow the exact determination of metabolic flux of PGN fragments of the bacteria under different (stress) condition. In addition, the use of the FLAG method, which mainly consists to introduce a gene within the bacterial genome to constitutively produce a marker peptide reflecting the size of the cell population, should also allow very accurate normalization of the quantified targets per cell amount and allow to refine the biological model of PGN pathways. Further development of the analytical method as well as cell engineering is required to achieve this level of information but the proposed method in this chapter is a first valuable step for such future development.

CHAPTER 4

Capillary Zone Electrophoresis Coupled to Mass Spectrometry for the Detection and Absolute Quantitation of Peptidoglycan-derived Peptides in Bacterial Cytoplasmic Extracts

4.1. Context of the chapter

As explained in Chapter 3, the peptidoglycan (PGN) is a layer found in the cell wall of bacteria that forms a structure surrounding the cytoplasmic membrane. Its primary function is to prevent lysis of the bacterium due to its own high intracellular osmotic pressure.

During the lifecycle of the bacterium, the peptidoglycan is involved in various biochemical processes such as growth and division and is therefore continuously remodeled. This constant reshape induces the release of various PGN fragments, including peptides and muropeptides. Naturally, the release of cell-wall fragments would represent a substantial loss of resources if they were not recovered and recycled. Consequently, PGN fragments can be reincorporated in the cytoplasm's cell to limit nutrient waste, as shown in Figure 4.1 (yellow part).

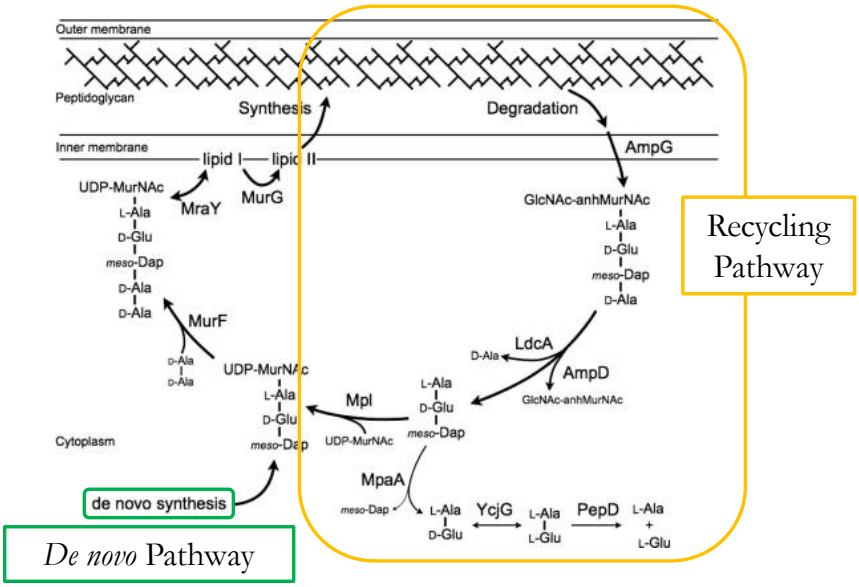


Figure 4.1 - Biosynthesis and recycling of PGN fragments in bacteria, adapted from Park, J. T., Uehara, T. *Microbiol Mol Bio Rev* 2008, 72 (2), 211–227.

Once recaptured, PGN fragments can be reallocated to *de novo* PGN biosynthesis (e.g. in *Escherichia coli*) or degraded to supply the metabolism of the cell (e.g. in *Bacillus subtilis*) according to the bacterial species.

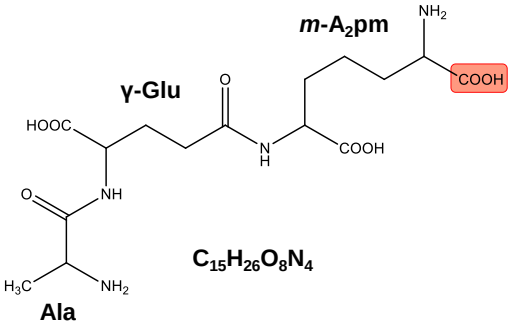
As explained in the previous chapter, bacteria introduce various chemical modifications to their PGN to acquire antibiotics resistance or reduce recognition by the mammalian host innate immune system during infection. Among these modifications, the amidation of carboxylic groups contained in the peptide moiety of PGN plays a crucial role. Indeed, the introduction of an amidation on the α-carboxyl of the ε-carboxyl group of the *m*-A_2pm was shown to affect PGN recognition by the mammalian host innate immune system. Also, in methicillin-resistant *Staphylococcus aureus* (MRSA), the amidation of the *D*-Glu contributes to the outstanding antibiotics resistance of this strain.

In this context, the development of reliable and robust analytical methods to detect and quantify cytoplasmic PGN fragments, along with their modified equivalents bearing an amidation is of great biochemical interest to better understand the PGN biosynthesis mechanism and the acquisition of antibiotic resistance.

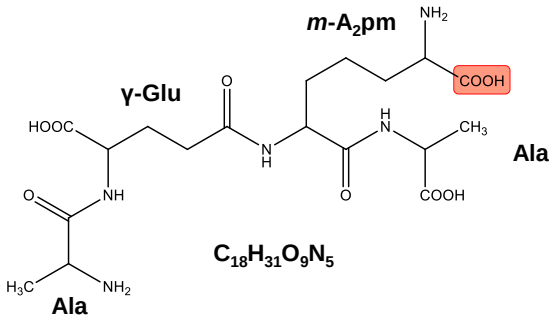
Once again, this work was realized within the collaboration between Prof. Joris from the Center of Protein Engineering of the University of Liège (Centre d'Ingénierie des Protéines, CIP) and the Mass Spectrometry Laboratory during the NetRBI ARC project (Action de Recherche Concertée).

In this chapter, we introduce the use of CZE-ESI-MS as a tool for the detection of three PGN-derived peptides with their equivalents bearing one amidation in crude bacterial cytoplasmic extracts, as shown in Figure 4.2.

(A) Tripeptide : *L*-Ala – γ-Glu – *m*-A_2pm



(B) Tetrapeptide : *L*-Ala – γ-Glu – *m*-A_2pm – *D*-Ala



(C) Pentapeptide : *L*-Ala – γ-Glu – *m*-A_2pm – *D*-Ala – *D*-Ala

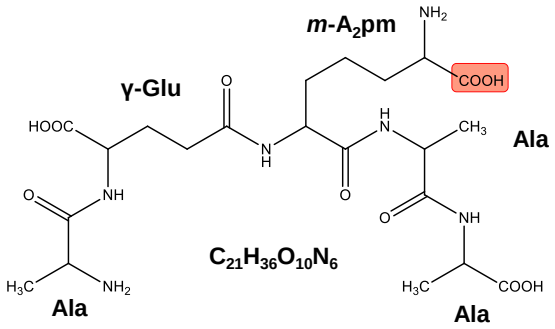


Figure 4.2 - **(A)** Structure and chemical formula of the tripeptide; **(B)** Structure and chemical formula of the tetrapeptide; **(C)** Structure and chemical formula of the pentapeptide. Boxes indicate the potential amidation site

The results showed that the CZE-ESI-MS method could successfully detect and separate all peptides of interest, including the peptides bearing an amidation. As expected, the species bearing an amidation presented a systematic increase of mobility (i.e. a lower migration time) compared to the unmodified species, in agreement with results obtained on muropeptides in Chapter 3. This effect was attributed to the neutralization of a partially negatively charged carboxylate group from the *m*-A_2pm which provides a net gain of positive charge of the species in acidic conditions such

as the BGE used during this work. The pK_a value of the meso-diaminopimelic acid contained in the peptide moiety was also estimated based on the respective migration times of the unmodified species and the equivalent bearing an amidation. Finally, we investigated the use of a ^{13}C ^{15}N labelled tripeptide standard as used for the absolute quantitation using LC-MS in proteomics. This isotopically labelled standard allowed the absolute quantitation of the endogenous tripeptide in bacterial cytoplasmic extracts of bacteria grown in the presence or absence of antibiotics (cephalosporin).

The results are presented in the following section as a submitted research article (the targeted journal is **Analytical Chemistry** from the American Chemical Society). The related Supporting Information (SI) are presented directly after the article.

4.2. Results – research article to be submitted in Analytical Chemistry

The Use of Capillary Zone Electrophoresis Coupled to Electrospray Mass Spectrometry for the Detection and Absolute Quantitation of Peptidoglycan-derived Peptides in Bacterial Cytoplasmic Extracts

Cédric DELVAUX[1], Marjorie DAUVIN[2], Madeleine BOULANGER[2], Loïc QUINTON[1], Bernard JORIS[2], Edwin DE PAUW[2] and Johann FAR[1]*

[1] Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium

[2] Center for Protein Engineering, InBioS Research Unit, Quartier Agora, University of Liège, Allée du Six Août 13, B-4000 Liège, Belgium

***Corresponding author:** c.delvaux@uliege.be

Keywords: amidation, CZE-ESI-MS, peptidoglycan, absolute quantitation, bacterial cytoplasmic extracts

Abbreviations:

***D*-Ala**: *D*-Alanine

***D*-γ-Glu**: *D*-isoglutamic acid

Glc*N*Ac: *N*-acetylglucosamine

***m*-A_2pm:** *meso*-diaminopimelic acid

Mur*N*Ac: *N*-acetylmuramic acid

PGN: peptidoglycan

BGE: background electrolyte

ESI-MS: electrospray mass spectrometry

ABSTRACT

Peptidoglycan (PGN) is an essential structure found in the bacterial cell wall. During the lifetime of the bacteria, PGN continuously undergoes degradation and biosynthesis to ensure various biochemical processes. PGN recycling process releases a large amount of fragments including peptides and muropeptides. PGN fragments also display a messenger function such as the induction of β-lactamase production in bacteria or the trigger of the immune response for mammals. PGN is also chemically modified by the introduction of an amidation on the carboxylic moiety of the diaminopimelic acid (*m*-A_2pm) to alter PGN degradation by autolysins and affect the recognition by the mammalian immune system. Currently, the detection and quantitation of PGN-derived peptides in bacterial cytoplasm is still challenging, mainly due to the difficulty to separate these very hydrophilic peptides by RP-HPLC. The assessment of the amidation in these peptides by RP-HPLC is also a major issue due to their co-elution.

Here we report the use of capillary zone electrophoresis coupled to mass spectrometry for the detection of three PGN peptides of interest and their equivalents bearing one amidation in bacterial cytoplasmic extracts. The quantitation of the tripeptide based on the ^{13}C ^{15}N isotopically labelled standard was also performed in crude cytoplasmic extracts of bacteria grown in the presence or absence of an antibiotics (cephalosporin). Despite the high complexity of the samples, the repeatability of the CZE-MS quantitation results was excellent, with standard deviations close to 1%. The global reproducibility of the method including biological handling was better than 20%.

INTRODUCTION

Peptidoglycan (PGN) is a biopolymer composed of glycans and peptides that forms a 3D network surrounding the cytoplasmic membrane of most bacteria. Peptidoglycan ensures an essential structural role for the bacterial cell wall, strong enough to maintain the high osmotic pressure of the cytoplasm and yet sufficient flexible to allow cell reproduction.

The PGN structure is composed of repeating β-(1→4)-(*N*-acetylglucosamine-β-(1→4)-*N*-acetylmuramic) disaccharide usually cross-linked by short peptides, therefore named peptidoglycan.[1] According to the bacterial strain, the *D*-lactyl group of each Mur*N*Ac residue is bonded to an oligopeptide with various possible sequences. In most cases, this peptide is composed of *L*-Ala - *D*-γ-Glu - *m*-A_2pm(or *L*-Lys) - *D*-Ala - *D*-Ala where the last *D*-Ala residue is eliminated in the mature PGN. The final 3D network is obtained by cross-linking the peptide moieties of neighboring PGN fragments to form the cell wall. Again, the nature of this cross-link is characteristic of the bacterial strain. For example, the PGN cross-link in *Escherichia coli* is achieved through a covalent bond between the *m*-A_2pm and *D*-Ala residues of adjacent muropeptides[2]. Alternatively, *Staphylococcus aureus* uses a Gly_5 interpeptide bridge to cross-link its PGN[3].

During the lifetime of the bacteria, the PGN is constantly reshaped[4] to adjust to the lifecycle of the bacteria, with a subsequent release of various PGN fragments, including peptides and muropeptides (i.e. peptides still bound to a glycan moiety). In some bacteria, these PGN fragments are reincorporated in the cytoplasm's cell to limit nutrient waste. This phenomenon applies to Gram negatives such as *E. coli* but also to Gram positives such as *B. subtilis*[5]. Once recaptured, PGN fragments can be reallocated to *de novo* PGN synthesis[6] (*E. coli*) or degraded by the cell to supply the metabolism[7,8] (*B. subtilis)*.

Alternatively, cytoplasmic PGN fragments also display a messenger function for bacteria[9]. For

example, in numerous Gram negative bacteria[10] and in the Gram positive environmental species *B. licheniformis*[11], the production of β-lactamase is induced by the presence of PGN fragments. Interestingly, some bacteria chemically alter the structure of their PGN by the introduction of an amidation on the free carboxylic moiety of their diaminopimelic acid (*m*-A_2pm)[12,13,14]. This chemical modification was shown to be linked to the regulation of autolysins activity[15] and PGN recognition by the mammalian host innate immune system[16,17]. As a result, the amidation of the *m*-A_2pm seems to preserve the PGN from degradation and to promote bacterial infection.

Because of its critical role for bacteria and almost unique composition in the living world, PGN is a target of choice for the use of antibiotics[18]. Recently, the emergence of bacterial resistance to antibiotics, including the appearance of so-called "super-bugs", has requested the introduction of new strategies to combat bacterial infections[19]. To this end, a detailed understanding of the PGN lifecycle including both its biosynthesis and recycling processes is mandatory. In this context, the development of analytical approaches to detect and quantify cytoplasmic PGN peptides is required.

Currently, the detection and quantitation of these PGN cytoplasmic peptides faces several analytical issues. First, the use of classical reversed-phase HPLC (which is the standard separation method for biochemists) fails to efficiently separate these small hydrophilic peptides, as they are eluted with the void volume. Then, a method for the efficient assessment of the amidation of these PGN peptides is also missing in the literature. Finally, the quantitation of these analytes in crude cytoplasmic extracts is still a very challenging analytical task due to the lack of separation or detection specificity of these analytes in complex matrices by the most frequently used analytical methods in biochemistry.

Recent work by Boulanger and coworkers demonstrated the potency of a CZE-ESI-MS-based method to baseline separate various PGN muropeptides after digestion by a muramidase,

including species bearing one or several amidation(s)[20]. CZE-ESI-MS was also successfully applied for the absolute quantitation of both peptides and proteins using the isotope-labelled standards method[21,22]. Currently, this approach is the golden standard for the absolute quantitation by MS[23,24] and consists in the addition of isotopically heavy-labeled peptide standards to the samples at known concentrations. The ratio of the co-eluted peak areas between the endogenous and spiked isotopic-labeled peptide yields the absolute quantity of the targeted peptide. This method provides excellent quantitation results due to the high specificity (identical physico-chemical properties except the mass) of the isotope-labelled standards to the targeted peptides.

In our study, we propose to introduce the use of CZE-ESI-MS for the analysis of cytoplasmic PGN peptides and their amidated counterparts in bacterial crude cytoplasmic extracts. The use of isotopically labelled peptide (^{13}C ^{15}N tripeptide) allows the absolute quantitation of the corresponding endogenous PGN peptide in crude bacterial cytoplasmic extracts of bacteria grown in the presence or absence of antibiotics (cephalosporin).

MATERIALS AND METHODS

Chemicals

All background electrolytes reagents (formic acid) and solvents (isopropanol) were purchased from Sigma-Aldrich (Bornem, Belgium) and were « metal trace analysis » grade. Milli-Q ultra-pure water was daily produced using system from Millipore (Molsheim, France) throughout the study.

Cytoplasmic PGN peptides

The PGN peptides targeted by this study are listed in the Table 1. When used as standards for method development, the tripeptide, tetrapeptide and pentapeptide were either synthesized following a reported synthesis protocol[25] or purified from bacterial cytoplasmic extracts. In all

cases, the structure of the standards was confirmed by tandem mass spectrometry (see Supporting Information Figures S1 to S3 for MS/MS spectra). The ^{13}C ^{15}N labelled tripeptide used for the quantitation method was prepared by the incorporation of a fully labelled ^{13}C ^{15}N *m*-A_2pm into the tripeptide structure, resulting in a +9Da shift in the molecular mass (seven ^{13}C and two ^{15}N in the *m*-A_2pm. The detailed preparation of this standard will be reported in detail in a future publication. The structure of this standard was confirmed by tandem mass spectrometry (see Supporting Information Figure S4 for MS/MS spectra).

Table 1: List of the peptides studied in this work with their respective sequences, chemical formulas and molecular weights.

Peptide	Sequence	Chemical Formula	Molecular weight
Tripeptide	*L*-Ala – *D*-γ-Glu – *m*-A_2pm	$C_{15}H_{26}O_8N_4$	390.175 g.mol^{-1}
Tripeptide*	*L*-Ala – *D*-γ-Glu – *m*-A_2pm*	$C_{15}H_{27}O_7N_5$	389.190 g.mol^{-1}
Tripeptide	*L*-Ala – *D*-γ-Glu – ***m*-A_2pm**	$\mathbf{^{13}C_7\,^{15}N_2}C_8H_{26}O_8N_2$	399.192 g.mol^{-1}
Tetrapeptide	*L*-Ala – *D*- γ-Glu – *m*-A_2pm – *D*-Ala	$C_{18}H_{31}O_9N_5$	461.212 g.mol^{-1}
Tetrapeptide*	*L*-Ala – *D*- γ-Glu – *m*-A_2pm* – *D*-Ala	$C_{18}H_{32}O_8N_6$	460.228 g.mol^{-1}
Pentapeptide	*L*-Ala – *D*- γ-Glu – *m*-A_2pm – *D*-Ala – *D*-Ala	$C_{21}H_{36}O_{10}N_6$	532.249 g.mol^{-1}
Pentapeptide*	*L*-Ala – *D*- γ-Glu – *m*-A_2pm* – *D*-Ala – *D*-Ala	$C_{21}H_{37}O_9N_7$	531.265 g.mol^{-1}

** indicates an amidation on the m-A_2pm.* ***Bold*** *indicates the ^{13}C ^{15}N labelling of m-A_2pm*

Sample preparation

Preparation of bacterial cytoplasmic extracts with and without the presence of antibiotics

Bacterial cells were cultivated in 100 ml LB medium at 37°C under 220 rpm agitation in a water incubator. Inoculation was made with a 16 h culture diluted 50 times. Cells were cultivated till mid-log phase before β-lactam addition when indicated (cephalosporin C with a final concentration of 7.5 µg.mL^{-1}), then further cultivated for 2 hours and centrifuged. Pellets were washed 3 times in pure ice-cold water. Final suspension was heated 10 minutes at 100°C then cells were disrupted by sonication. Lysate was centrifuged for 30 minutes at 15,000g then the supernatant was filtered on a 10 kDa cut off. Samples were lyophilized and weighted with an analytical balance before final suspension in formic acid (10 mM). For PGN peptides quantification experiments, number of bacteria in each sample was quantified by flow cytometry using a FACSVerse® instrument (BD Biosciences).

Cytoplasmic PGN peptides analysis by online CZE-ESI-MS

The electrophoretic separation of the cytoplasmic peptides was performed on a P/ACE MDQ (Beckman Coulter – SCIEX Separations, Brea, California, USA) running under 32karat v7.0 software. The CE apparatus was equipped with a bare fused-silica capillary of 50 µm inner diameter and 365 µm outer diameter with a total length of 90 cm. The capillary was conditioned using successively deionized water (25 psi, 5 min), MeOH/H_2O 50/50 (25 psi, 5 min), deionized water (25 psi, 5 min), 0.1 M NaOH (25psi, 5min), deionized water (25 psi, 5 min), 0.2 M HCl (25 psi, 10 min) and deionized water (25 psi, 10 min). Finally, the capillary was flushed with the 100 mM FA BGE during 20 min at 50 psi.

Before the first CZE run and between the different runs, the capillary was rinsed using BGE (25 psi, 5 min) before hydrodynamically loading the different samples (mix of PGN peptides or cytoplasmic extracts) at 2 psi for 15 seconds (estimated injected volume of 44nL or 2.5% of the capillary volume)

The CZE separation was performed applying 30 kV (normal polarity) for 60 minutes with a 1-minute voltage ramping. The capillary was maintained at 25 °C during the entire sets of runs.

The CE system was connected to the mass spectrometer through the EDA external adapter (Beckman, by passing the UV detector) to a sheath liquid microfluidic sprayer CEMS interface (PN 10-301347 ANALIS SL CE-MS Sprayer, Analis, Suarlée Belgium) operating in the microliter range. The selected sheath liquid was 50% iso-propanol in 10 mM FA delivered at 1.5 µL.min^{-1} using a 1mL Hamilton syringe and an external syringe pump to ensure stable Taylor cone for ESI and minimum chemical noise during detection.

ESI-MS detection was performed on a Thermo Scientific Q Exactive Orbitrap mass spectrometer (Thermo Scientific, Merelbeke, Belgium) equipped with the nanospray Flex ion source. The source temperature was set to 250°C. The Q Exactive mass range was fixed to 200-1,000 m/z with a 70,000 resolving power at 200 m/z and was externally calibrated using the

Thermo Scientific calibration mixture as recommended by the manufacturer. Data were reprocessed using the XCalibur software v3.1.

RESULTS AND DISCUSSION

Cytoplasmic PGN peptides analysis by online CZE-ESI-MS

First, the method development for the separation of the unmodified PGN peptides was performed on a mixture of the tripeptide, tetrapeptide and pentapeptide (see Figure S5 in the Supporting Information for detailed structures). These peptides were prepared at approximatively 10uM each in a 10mM formic acid buffer. The obtained electropherogram is shown in Figure 1:

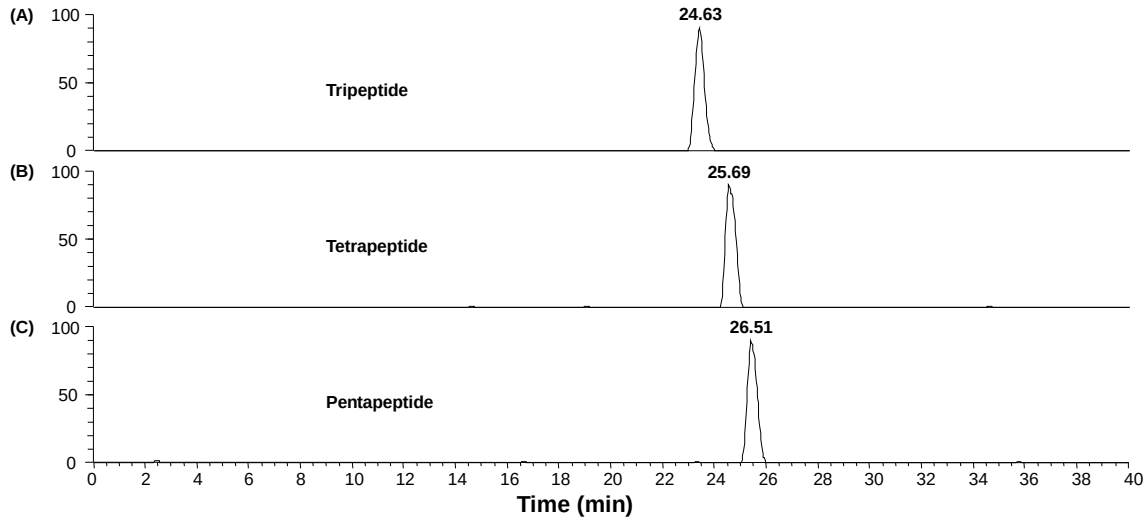


Figure 1: ***(A)*** *extracted electropherogram of the tripeptide (391.182 m/z as $[M+H]^+$);* ***(B)*** *extracted electropherogram of the tetrapeptide (462.219 m/z as $[M+H]^+$) and* ***(C)*** *extracted electropherogram of the pentapeptide (533.257 m/z as $[M+H]^+$)*

As can be seen in Figure 1, these different PGN peptides are all baseline separated in less than 30 minutes. The constant increase in the observed migration time from the tripeptide to the pentapeptide can be correlated to the increase in the respective molecular weight of these various PGN peptides at almost constant charge of the peptides due to the acidic background electrolyte used for the CE separation.

Amidation of the *m*-A2pm of PGN peptides: the example of the tripeptide

As various bacterial strains are reported to present amidation on the *m*-A2pm contained in their PGN, the effect of this chemical PGN modification was investigated on a mixture of the tripeptide and the tripeptide bearing an amidation (see Figure S6 in Supporting Information for detailed structures). The obtained electropherogram is shown in Figure 2:

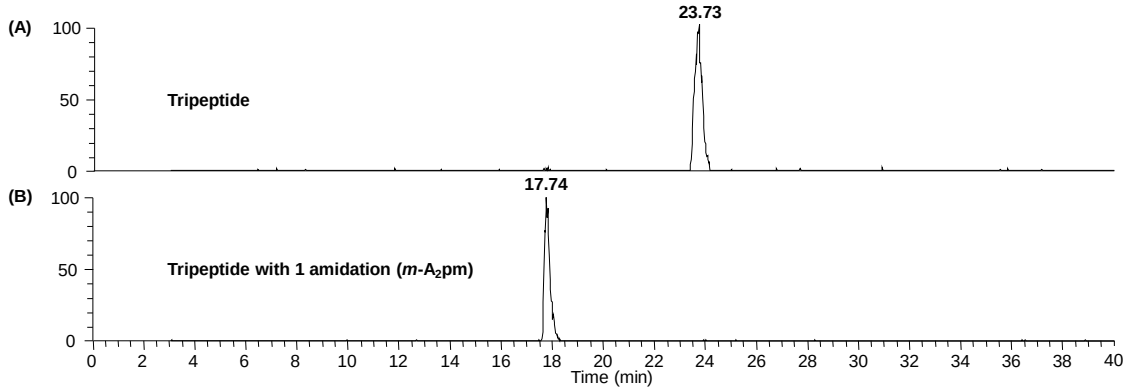


Figure 2: ***(A)*** *extracted electropherogram of the tripeptide (391.182 m/z as $[M+H]^+$); and* ***(B)*** *extracted electropherogram of the tripeptide bearing one amidation on its m-A_2pm (390.198 m/z as $[M+H]^+$)*

Figure 2 shows a substantial decrease of the migration time of the tripeptide bearing one amidation compared to the unmodified tripeptide. This result is in perfect agreement with recently published literature by Boulanger and coworkers for muropeptides[20]. This result points out an increase of the mobility of the analyte upon amidation, which can be associated to the conversion of the partially negatively charged carboxylic group of *m*-A2pm to a neutral amide group (-COO^- is converted into -CO-NH_2-). The comparison between the non-amidated and amidated species leads to an estimated pK_a value for the carboxylic acid of *m*-A2pm of 2.25; consistent with the estimated value of 2.2 determined for muropeptides[20]. The complete demonstration of the determination of the pK_a values can be found in the Supporting Information.

Detection of PGN tripeptide, tetrapeptide, pentapeptide and equivalents bearing one amidation in a crude bacterial cytoplasmic extract

The detection of three different PGN peptides (i.e. the tripeptide, tetrapeptide and pentapeptide) along with their equivalents bearing one amidation was performed on a crude bacterial cytoplasmic extract. The obtained electropherogram is shown in Figure 3:

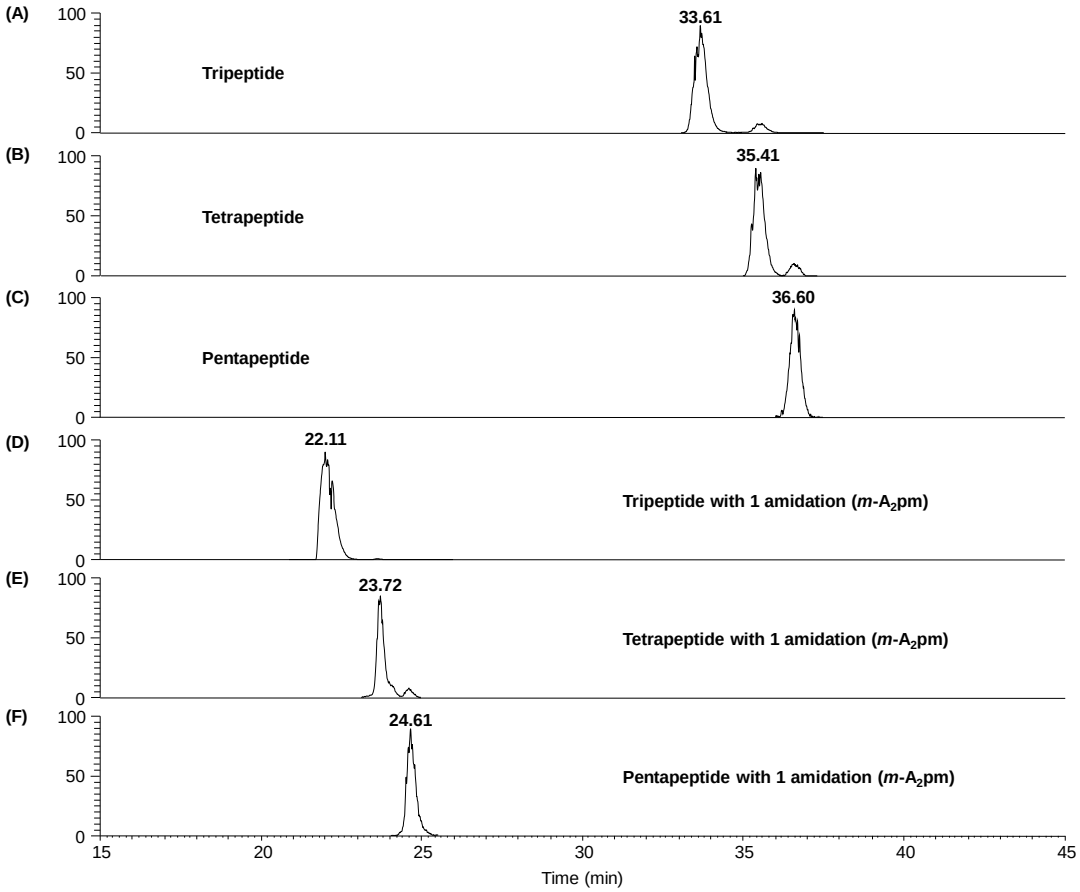


Figure 3: ***(A)*** *extracted electropherogram of the tripeptide (391.18278 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract;* ***(B)*** *extracted electropherogram of the tetrapeptide (462.219 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract;* ***(C)*** *extracted electropherogram of the pentapeptide (533.257 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract;* ***(D)*** *extracted electropherogram of the tripeptide bearing one amidation (390.198 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract;* ***(E)*** *extracted electropherogram of the tetrapeptide bearing one amidation (461.235 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract and* ***(F)*** *extracted electropherogram of the pentapeptide bearing one amidation (532.273 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract*

As can be seen, the different targeted cytoplasmic peptides (unmodified or bearing one amidation) could all be detected, and baseline separated. Consistent with the results of the tripeptide (Figure 2), the peptides bearing one amidation systematically presented a lower

migration time than their equivalents without amidation. Again, the comparison between the species with and without amidation leads to estimated pK_a values of 2.20, 2.21 and 2.21 for the tripeptide, tetrapeptide and pentapeptide respectively. This result is in excellent agreement with the calculated values on the mix of the tripeptide standards with and without amidation (Figure 2, pK_a value of 2.25) and recent literature on muropeptides[20] (pK_a value of 2.2). In the present analysis, the migration times of the different PGN peptides are shifted compared to the observed migration time for the standard peptides due to the higher complexity of the sample (crude cytoplasmic extracts). The use of a reference peptide as internal standard could correct these shifts by working with the relative migration time instead of the absolute migration time. In any case, the exact mass determination combined to the fragmentation pattern provided by the Q Exactive allows the confident identification of the targeted peptides despite the shift in migration time.

Quantitation of the cytoplasmic PGN tripeptide in bacterial cytoplasmic extracts grown in the presence or absence of antibiotics (cephalosporin C)

The absolute quantitation of the PGN tripeptide in cytoplasmic extracts of bacteria grown with or without the presence of cephalosporin (antibiotics) was performed using the ^{13}C ^{15}N labelled tripeptide. To this end, triplicates of bacterial cultures (biological triplicates) were grown with and without the presence of antibiotics (see Table 2). Each cytoplasmic extract was then analyzed by CZE-ESI-MS in triplicates (analytical triplicates). The normalization of the results between biological triplicates was performed by evaluating the number of cells in each sample using FACSVerse (a high-performance flow cytometer). The ^{13}C ^{15}N tripeptide was added at known concentration (see Table 2) to each cytoplasmic extract of these bacterial cultures prior to analysis for the absolute quantitation. An example of electropherogram is depicted in Figure 4:

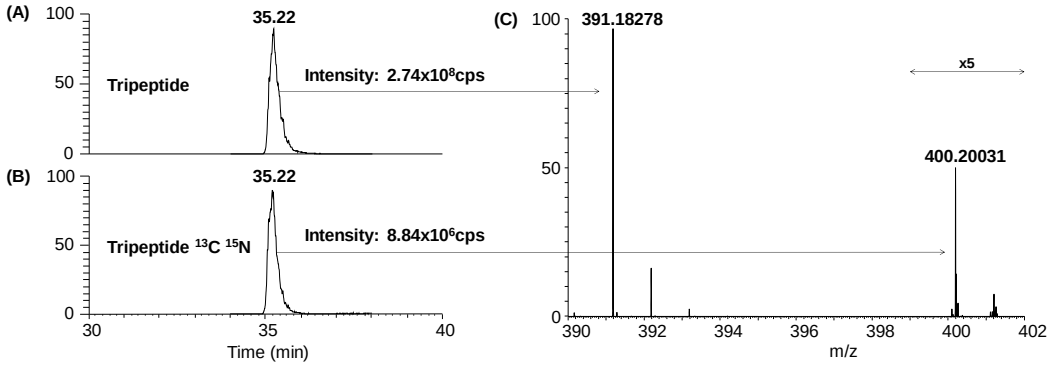


Figure 4: ***(A)*** *extracted electropherogram of the tripeptide (391.18278 m/z as $[M+H]^+$) in a bacterial cytoplasmic extract with antibiotics; and* ***(B)*** *extracted electropherogram of the ^{13}C ^{15}N labelled tripeptide (400.20031 as $[M+H]^+$) in a bacterial cytoplasmic extract with antibiotics and* ***(C)*** *associated mass spectrum with the extracted electropherogram at 35.22 min. The 399 to 402 m/z range was magnified 5x for better visualization*

As can be seen from Figure 4, the ^{13}C ^{15}N labelled and endogenous tripeptide present an identical migration time, supporting the merit of the proposed strategy for the absolute quantitation of the tripeptide in the bacterial cytoplasmic extracts. The direct ratio between the peak areas yields the concentration ratio of the species. Table 2 shows the results of the quantitation of the tripeptide in the different sample.

Table 2: List of the different samples of crude bacterial cytoplasmic extracts used in this study along with the measured number of cells (FACS), the concentration of added ^{13}C ^{15}N labeled tripeptide, the calculated concentration of cytoplasmic tripeptide and normalized concentration of tripeptide (FACS). ***Bold*** *indicates the average values with SD values for the condition in presence of antibiotics and in absence of antibiotics.*

Sample	**Sample #1**	**Sample #2**	**Sample #3**	**Sample #4**	**Sample #5**	**Sample #6**
Addition of cephalosporin (antibiotics)	No	No	No	Yes	Yes	Yes
Sample quantity (number of cells.μL^{-1})	28,226	23,937	25,892	25,860	28,202	25,734
Spiked ^{13}C ^{15}N tripeptide ($\mu mol.L^{-1}$)	18.78	20.59	19.90	22.10	17.16	19.23
Injected tripeptide ($\mu mol.L^{-1}$)	969.70 ± 9.64	1,122.99 ± 10.60	967.49 ± 19.43	1,475.32 ± 17.53	1,195.74 ± 1.95	1,132.03 ± 1.77
Normalized intracellular tripeptide ($mmol.L^{-1}$)	61.35 ± 0.61	84.30 ± 0.79	66.73 ± 1.34	101.88 ± 1.21	75.71 ± 0.12	78.55 ± 0.12
	70.79 ± 12.00 (RSD=16.9%)			**85.38 ± 14.36 (RSD=16.8%)**		

Table 2 highlights different important features of the method. First, the standard deviation (SD) determined for the analytical triplicates (same analyzed sample) is generally close to 1%. This low SD value points out the high repeatability of the CZE-ESI-MS method; even for the analysis of highly complex matrices such as crude bacterial cytoplasmic extracts. Then, the comparison between biological replicates (1 to 3 and 4 to 6) reveals that the biological repeatability is lower, with SD values around 16%. This higher value can be explained by the inherent biological heterogeneity of bacterial phenotype in stressful conditions, where different sub-populations can coexist[26,27]. The comparison between samples grown in the absence of cephalosporin (1 to 3) and in the presence of cephalosporin (4 to 6) shows an overall increase in the concentration of cytoplasmic tripeptide. This result highlights the ability of the method to perform the absolute quantitation of cytoplasmic PGN peptide concentrations in different stress conditions (for example here: the presence or absence of antibiotics during growth) as long as the isotope-labelled standards are available.

CONCLUDING REMARKS

In this work, we introduce the use of CZE-ESI-MS for the baseline separation of PGN cytoplasmic peptides ranging from the tripeptide to the pentapeptide. The effect of the amidation of the carboxylic group of *m*-A_2pm was also assessed and showed to increase the mobility of the PGN peptides bearing an amidation in acidic conditions. The comparison between unmodified species and their equivalents bearing one amidation leads to a pK_a estimation for the free carboxylic group of m-A_2pm around 2.2-2.3. This value is in excellent agreement with published results for muropeptides with a pK_a estimation close to 2.2[20]. The method was also tested on a crude bacterial cytoplasmic extract to examine the performance of the method on a complex matrix. In this case, the method detected, and baseline resolved the tripeptide, tetrapeptide and pentapeptide with their counterparts bearing an amidation. Again, the amidated species systematically presented a lower migration time, consistent with previous

observations. Finally, a quantitation method was developed based on the use of the isotope-labelled (^{13}C and ^{15}N) tripeptide. As this species shares almost an identical migration time to the endogenous cytoplasmic tripeptide, the absolute quantitation of this PGN peptide was performed based on the comparison of their respective peak areas. The analytical performances of the method were assessed on biological and analytical triplicates. The quantitation results were normalized by the accurate counting of the cell number in each sample using high-performance flow cytometry (FACS). The analytical repeatability of the CZE-ESI-MS method was excellent, with standard deviations close to 1% for each sample. In contrast, the biological variability was significantly higher with standard deviations close to 16% but was still excellent. The comparison of the measured tripeptide concentrations between conditions with and without antibiotics revealed an increased concentration of tripeptide in samples grown in the presence of antibiotics. This increase is in agreement with the results previously found in *E. coli* cells treated with β-lactam, where PGN degradation system activity is increased, generating more PGN degradation products (among others the ^{anh}MurNAc-peptide)[28]. This result also indicates the great potential of the present method to precisely quantify the bacterial cytoplasmic concentrations of PGN peptides (including species bearing an amidation) in samples grown in various stress conditions. The determination of the cytoplasmic PGN peptide levels could also be used to assess the *de novo* biosynthesis or degradation of such species by mutation of the putative involved genes. The profiling of the bacterial strain could also be considered by comparing the respective amount of each cytoplasmic PGN peptide along with its level of amidation. From a methodological point of view, the applied quantitation method for the tripeptide could also be applied for the other PGN peptides (with or without amidation), peptides or even any kind of molecules if the isotope-labelled equivalent is available.

Acknowledgments

Analis (Suarlée, Belgium) is acknowledged for providing the CE instrument and for the useful discussions. This work was supported by the Wallonia Region and the FIRST Entreprise 6592 funding and by Actions de Recherche Concertées (ARC) NetRBI 2012–2017 of the Belgian French Community. Dr Gabriel Mazzucchelli and Dominique Baiwir from the proteomics facility of the University of Liege are also acknowledged for providing access to the Q-Exactive instrument. Elodie Grifnée is acknowledged for her skills in the preparation of the samples and the sample list configuration. The GIGA Flow Cytometry Platform is acknowledged for providing access to the FACS instrument.

REFERENCES

(1) Rogers, H. J. Bacterial Growth and the Cell Envelope. *Bacteriol. Rev.* **1970**, *34* (2), 194–214.
(2) Vollmer, W.; Bertsche, U. Murein (Peptidoglycan) Structure, Architecture and Biosynthesis in Escherichia Coli. *Biochim. Biophys. Acta (BBA)-Biomembranes* **2008**, *1778* (9), 1714–1734.
(3) De Jonge, B. L.; Chang, Y.-S.; Gage, D.; Tomasz, A. Peptidoglycan Composition of a Highly Methicillin-Resistant Staphylococcus Aureus Strain. The Role of Penicillin Binding Protein 2A. *J. Biol. Chem.* **1992**, *267* (16), 11248–11254.
(4) Barreteau, H.; Kovač, A.; Boniface, A.; Sova, M.; Gobec, S.; Blanot, D. Cytoplasmic Steps of Peptidoglycan Biosynthesis. *FEMS Microbiol. Rev.* **2008**, *32* (2), 168–207.
(5) Park, J. T.; Uehara, T. How Bacteria Consume Their Own Exoskeletons (Turnover and Recycling of Cell Wall Peptidoglycan). *Microbiol. Mol. Biol. Rev.* **2008**, *72* (2), 211–227, table of contents. https://doi.org/10.1128/MMBR.00027-07.
(6) Goodell, E. W. Recycling of Murein by Escherichia Coli. *J. Bacteriol.* **1985**, *163* (1), 305–310.
(7) Mauck, J.; Chan, L.; Glaser, L. Turnover of the Cell Wall of Gram-Positive Bacteria. *J. Biol. Chem.* **1971**, *246* (6), 1820–1827.
(8) Mauck, J.; Glaser, L. Turnover of the Cell Wall of Bacillus Subtilis W-23 during Logarithmic Growth. *Biochem. Biophys. Res. Commun.* **1970**, *39* (4), 699–706. https://doi.org/10.1016/0006-291X(70)90261-5.
(9) Boudreau, M. A.; Fisher, J. F.; Mobashery, S. Messenger Functions of the Bacterial Cell Wall-Derived Muropeptides. *Biochemistry* **2012**, *51* (14), 2974–2990.
(10) Juan, C.; Torrens, G.; González-Nicolau, M.; Oliver, A. Diversity and Regulation of Intrinsic β-Lactamases from Non-Fermenting and Other Gram-Negative Opportunistic Pathogens. *FEMS Microbiol. Rev.* **2017**, *41* (6), 781–815. https://doi.org/10.1093/femsre/fux043.
(11) Amoroso, A.; Boudet, J.; Berzigotti, S.; Duval, V.; Teller, N.; Mengin-Lecreulx, D.; Luxen, A.; Simorre, J.-P.; Joris, B. A Peptidoglycan Fragment Triggers β-Lactam Resistance in Bacillus Licheniformis. *PLoS Pathog.* **2012**, *8* (3), e1002571. https://doi.org/10.1371/journal.ppat.1002571.

(12) Atrih, A.; Bacher, G.; Allmaier, G.; Williamson, M. P.; Foster, S. J. Analysis of Peptidoglycan Structure from Vegetative Cells of Bacillus Subtilis 168 and Role of PBP 5 in Peptidoglycan Maturation. *J. Bacteriol.* **1999**, *181* (13), 3956–3966.
(13) Warth, A. D.; Strominger, J. L. Structure of the Peptidoglycan from Vegetative Cell Walls of Bacillus Subtilis. *Biochemistry* **1971**, *10* (24), 4349–4358.
(14) Heinz, K.; Kandler, O. *Peptidoglycan .Types of Bacterial Cell Walls and Their Taxonomic Implications*; 1972; Vol. 36.
(15) Dajkovic, A.; Tesson, B.; Chauhan, S.; Courtin, P.; Keary, R.; Flores, P.; Marlière, C.; Filipe, S. R.; Chapot-Chartier, M.-P.; Carballido-Lopez, R. Hydrolysis of Peptidoglycan Is Modulated by Amidation of Meso-Diaminopimelic Acid and Mg2+ in Bacillus Subtilis. *Mol. Microbiol.* **2017**, *104* (6), 972–988. https://doi.org/10.1111/mmi.13673.
(16) Girardin, S. E.; Travassos, L. H.; Hervé, M.; Blanot, D.; Boneca, I. G.; Philpott, D. J.; Sansonetti, P. J.; Mengin-Lecreulx, D. Peptidoglycan Molecular Requirements Allowing Detection by Nod1 and Nod2. *J. Biol. Chem.* **2003**, *278* (43), 41702–41708.
(17) Bernard, E.; Rolain, T.; Courtin, P.; Hols, P.; Chapot-Chartier, M.-P. Identification of the Amidotransferase AsnB1 as Being Responsible for Meso-Diaminopimelic Acid Amidation in Lactobacillus Plantarum Peptidoglycan. *J. Bacteriol.* **2011**, *193* (22), 6323–6330.
(18) Bugg, T. D. H.; Braddick, D.; Dowson, C. G.; Roper, D. I. Bacterial Cell Wall Assembly: Still an Attractive Antibacterial Target. *Trends Biotechnol.* **2011**, *29* (4), 167–173.
(19) Bravo, A.; Ruiz-Cruz, S.; Alkorta, I.; Espinosa, M. When Humans Met Superbugs: Strategies to Tackle Bacterial Resistances to Antibiotics. *Biomol. Concepts* **2018**, *9* (1), 216–226.
(20) Boulanger, M.; Delvaux, C.; Quinton, L.; Joris, B.; Pauw, E. De; Far, J. Bacillus Licheniformis Peptidoglycan Characterization by CZE-MS: Assessment with the Benchmark RP-HPLC-MS Method. *Electrophoresis* **2019**.
(21) Jantos-Siwy, J.; Schiffer, E.; Brand, K.; Schumann, G.; Rossing, K.; Delles, C.; Mischak, H.; Metzger, J. Quantitative Urinary Proteome Analysis for Biomarker Evaluation in Chronic Kidney Disease. *J. Proteome Res.* **2008**, *8* (1), 268–281.
(22) Zhu, G.; Sun, L.; Linkous, T.; Kernaghan, D.; McGivney IV, J. B.; Dovichi, N. J. Absolute Quantitation of Host Cell Proteins in Recombinant Human Monoclonal Antibodies with an Automated CZE-ESI-MS/MS System. *Electrophoresis* **2014**, *35* (10), 1448–1452.
(23) Ong, S.-E.; Mann, M. Mass Spectrometry–Based Proteomics Turns Quantitative. *Nat. Chem. Biol.* **2005**, *1* (5), 252.
(24) Ong, S.-E.; Foster, L. J.; Mann, M. Mass Spectrometric-Based Approaches in Quantitative Proteomics. *Methods* **2003**, *29* (2), 124–130.
(25) Simon, J. F.; Lamborelle, N.; Zervosen, A.; Lemaire, C.; Joris, B.; Luxen, A. Development of Solid-Supported Methodology for the Preparation of Peptidoglycan Fragments Containing (2S, 6R)-Diaminopimelic Acid. *Tetrahedron Lett.* **2016**, *57* (14), 1572–1575.
(26) Golding, I.; Paulsson, J.; Zawilski, S. M.; Cox, E. C. Real-Time Kinetics of Gene Activity in Individual Bacteria. *Cell* **2005**, *123* (6), 1025–1036. https://doi.org/10.1016/J.CELL.2005.09.031.
(27) Cai, L.; Friedman, N.; Xie, X. S. Stochastic Protein Expression in Individual Cells at the Single Molecule Level. *Nature* **2006**, *440* (7082), 358–362. https://doi.org/10.1038/nature04599.
(28) Cho, H., Uehara, T. & Bernhardt, T. G. Beta-lactam antibiotics induce a lethal malfunctioning of the bacterial cell wall synthesis machinery. *Cell* **159**, 1300–1311 (2014).

4.3. Results - Supporting Information

The Use of Capillary Zone Electrophoresis Coupled to Mass Spectrometry for the Detection and Absolute Quantitation of Peptidoglycan-derived Peptides in Bacterial Cytoplasmic Extracts

Cédric DELVAUX[1], Marjorie DAUVIN[2], Madeleine BOULANGER[2], Loïc QUINTON[1], Bernard JORIS[2], Edwin DE PAUW[2] and Johann FAR[1]*

[1] Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium

[2] Center for Protein Engineering, InBioS Research Unit, Quartier Agora, University of Liège, Allée du Six Août 13, B-4000 Liège, Belgium

***Corresponding author:** c.delvaux@uliege.be

SUPPORTING INFORMATION

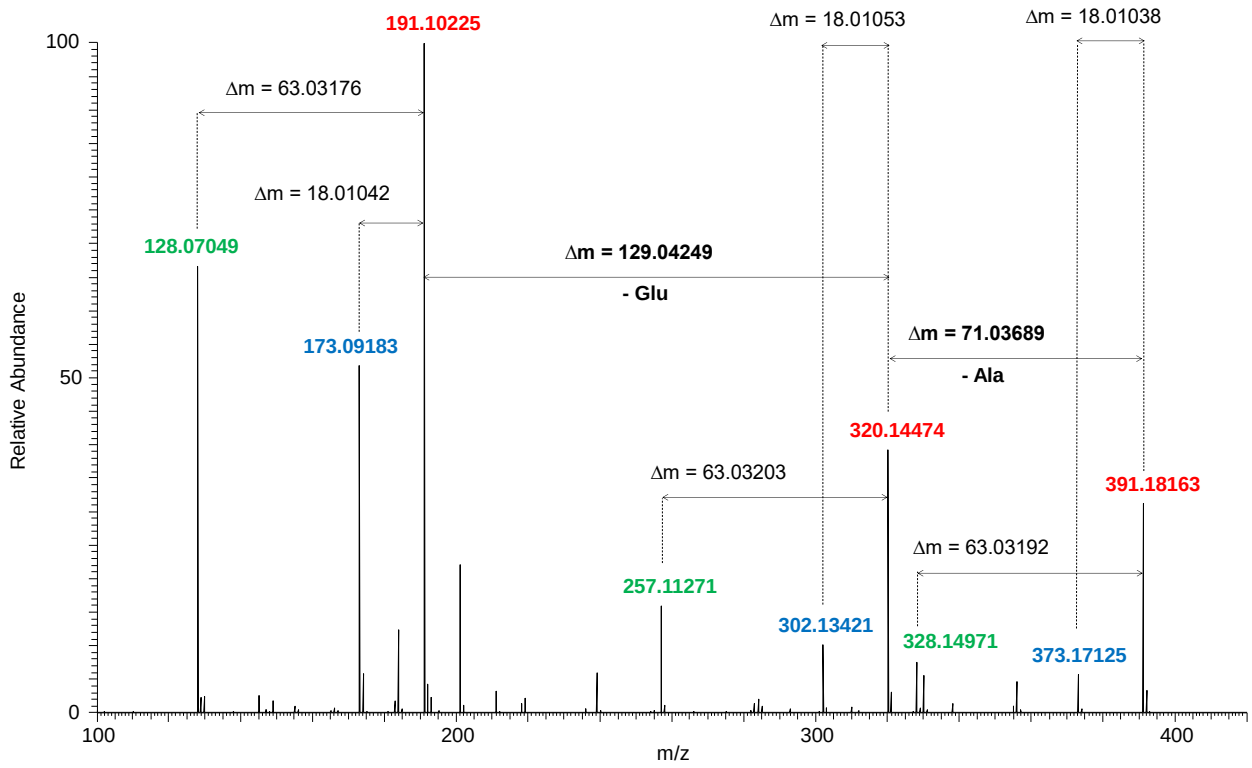


Figure S1: HCD MS/MS spectrum of the tripeptide standard. Bold red indicates the *m/z* of the characteristic ions for structural assignment, bold blue indicates loss of water and bold green indicates the combined loss of HCOOH and NH_3. The list below summarizes the different *m/z* of the fragment ions, their structures, their chemical formulas and their respective mass accuracies:

m/z	Structural assignment	Chemical formula	Mass accuracy
391.18163	Ala-γ-Glu-*m*-A_2pm	$C_{15}H_{26}O_8N_4$	-1.81ppm
373.17125	Ala-γ-Glu-*m*-A_2pm with a H_2O loss	$C_{15}H_{24}O_7N_4$	-1.42ppm
328.14971	Ala-γ-Glu-*m*-A_2pm with a "HCO_2H + NH_3" loss	$C_{14}H_{21}O_6N_3$	-1.82ppm
320.14474	γ-Glu-*m*-A_2pm	$C_{12}H_{21}O_7N_3$	-1.53ppm
302.13421	γ-Glu-*m*-A_2pm with a H_2O loss	$C_{12}H_{19}O_6N_3$	-1.49ppm
257.11271	γ-Glu-*m*-A_2pm with a "HCO_2H + NH_3" loss	$C_{11}H_{16}O_5N_2$	-1.91ppm
191.10225	*m*-A_2pm	$C_7H_{14}N_2O_4$	-1.99ppm
173.09183	*m*-A_2pm with a H_2O loss	$C_7H_{12}N_2O_3$	-1.39ppm
128.07049	*m*-A_2pm with a "HCO_2H + NH_3" loss	$C_6H_9NO_2$	-0.94ppm

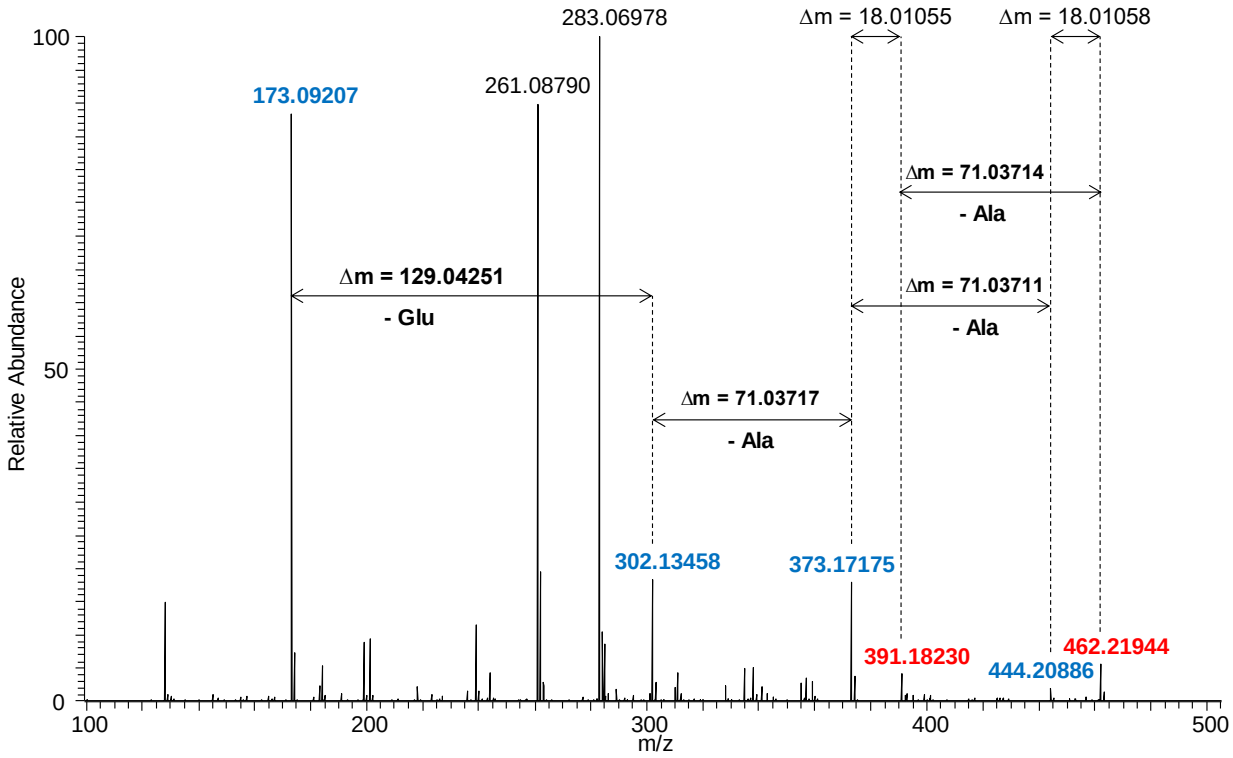


Figure S2: HCD MS/MS spectrum of the tetrapeptide standard. Bold red indicates the *m/z* of the characteristic ions for structural assignment and bold blue indicates loss of water. The list below summarizes the different *m/z* of the fragment ions, their structures, their chemical formulas and their respective mass accuracies:

m/z	Structural assignment	Chemical formula	Mass accuracy
462.21944	Ala-γ-Glu-*m*-A_2pm-Ala	$C_{18}H_{31}O_9N_5$	-0.02ppm
444.20886	Ala-γ-Glu-*m*-A_2pm-Ala with a H_2O loss	$C_{18}H_{29}O_8N_5$	-0.07ppm
391.18163	Ala-γ-Glu-*m*-A_2pm	$C_{15}H_{26}O_8N_4$	-0.07ppm
373.17125	Ala-γ-Glu-*m*-A_2pm with a H_2O loss	$C_{15}H_{24}O_7N_4$	-0.08ppm
302.13421	γ-Glu-*m*-A_2pm with a H_2O loss	$C_{12}H_{19}O_6N_3$	-0.26ppm
173.09183	*m*-A_2pm with a H_2O loss	$C_7H_{12}N_2O_3$	0.00 ppm

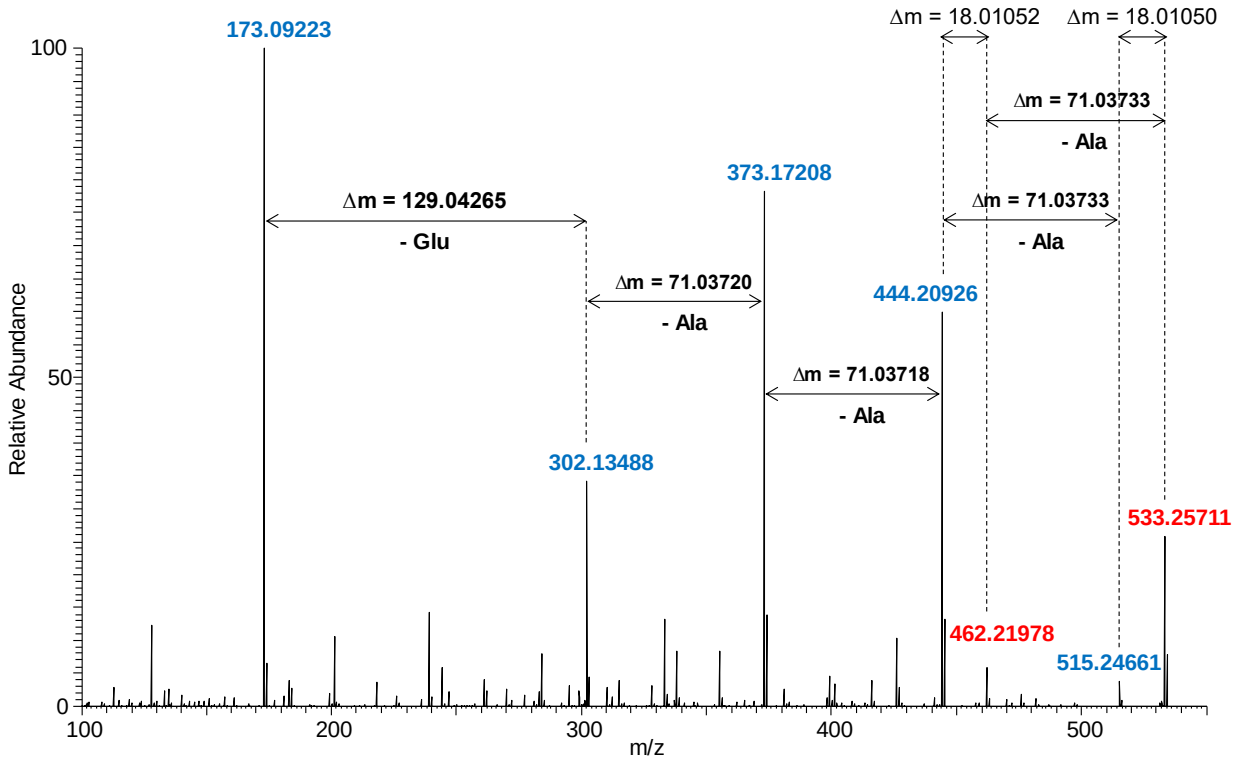


Figure S3: HCD MS/MS spectrum of the pentapeptide standard. Bold red indicates the *m/z* of the characteristic ions for structural assignment and bold blue indicates loss of water. The list below summarizes the different *m/z* of the fragment ions, their structures, their chemical formulas and their respective mass accuracies:

m/z	Structural assignment	Chemical formula	Mass accuracy
533.25711	Ala-γ-Glu-*m*-A_2pm-Ala-Ala	$C_{21}H_{36}O_{10}N_6$	+1.01ppm
515.24661	Ala-γ-Glu-*m*-A_2pm-Ala-Ala with a H_2O loss	$C_{21}H_{34}O_9N_6$	+1.18ppm
462.21978	Ala-γ-Glu-*m*-A_2pm-Ala	$C_{18}H_{31}O_9N_5$	+0.71ppm
444.20926	Ala-γ-Glu-*m*-A_2pm-Ala with a H_2O loss	$C_{18}H_{29}O_8N_5$	+0.83ppm
373.17208	Ala-γ-Glu-*m*-A_2pm with a H_2O loss	$C_{15}H_{24}O_7N_4$	+0.80ppm
302.13488	γ-Glu-*m*-A_2pm with a H_2O loss	$C_{12}H_{19}O_6N_3$	+0.73ppm
173.09223	*m*-A_2pm with a H_2O loss	$C_7H_{12}N_2O_3$	+0.92 ppm

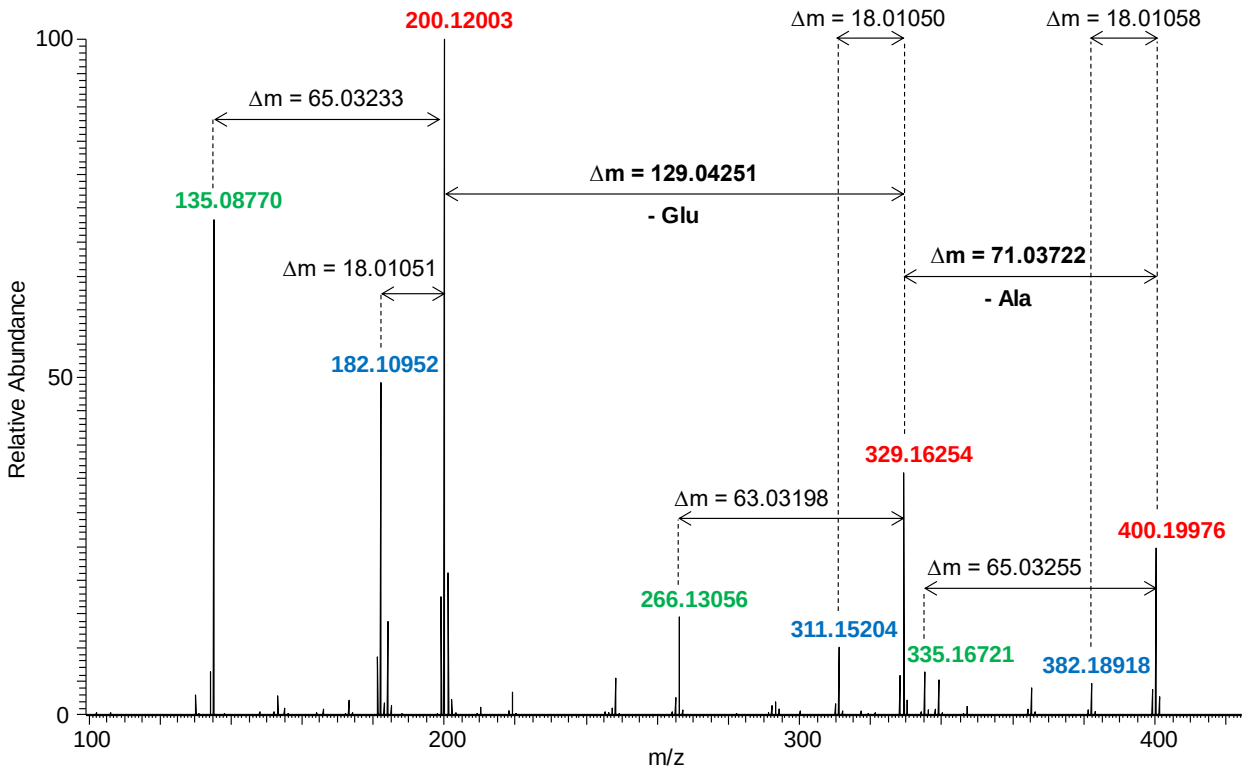
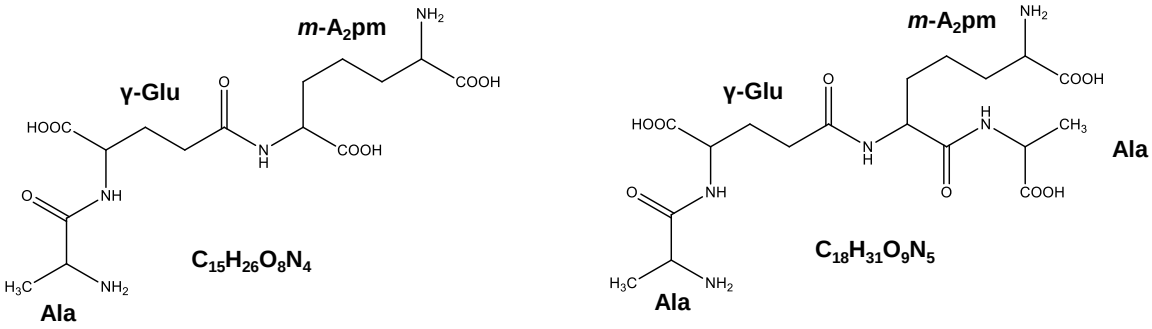


Figure S4: HCD MS/MS spectrum of the ^{13}C ^{15}N tripeptide standard. Bold red indicates the *m/z* of the characteristic ions for structural assignment, bold blue indicates loss of water and bold green indicates the combined loss of HCOOH and NH_3. The list below summarizes the different *m/z* of the fragment ions, their structures, their chemical formulas and their respective mass accuracies (**bold** *m*-A_2pm stands for ^{13}C ^{15}N *m*-A_2pm):

m/z	Structural assignment	Chemical formula	Mass accuracy
400.19976	Ala-γ-Glu-***m*-A$_2$pm**	**$^{13}C_7$ $^{15}N_2$** $C_8H_{26}O_8N_2$	-0.32ppm
382.18918	Ala-γ-Glu-***m*-A$_2$pm** with a H_2O loss	**$^{13}C_7$ $^{15}N_2$** $C_8H_{24}O_7N_2$	-0.39ppm
335.16721	Ala-γ-Glu-***m*-A$_2$pm** with a "$H^{13}CO_2H$ + $^{15}NH_3$" loss	**$^{13}C_6$ ^{15}N** $C_8H_{21}O_6N_2$	-0.81ppm
329.16254	γ-Glu-***m*-A$_2$pm**	**$^{13}C_7$ $^{15}N_2$** $C_5H_{21}O_7N$	-0.73ppm
311.15204	γ-Glu-***m*-A$_2$pm** with a H_2O loss	**$^{13}C_7$ $^{15}N_2$** $C_5H_{19}O_6N$	-0.59ppm
266.13056	γ-Glu-***m*-A$_2$pm** with a "HCO_2H + NH_3" loss	**$^{13}C_7$ $^{15}N_2$** $C_4H_{16}O_5$	-0.71ppm
200.12003	***m*-A$_2$pm**	**$^{13}C_7H_{14}{}^{15}N_2O_4$**	-0.80pm
182.10952	***m*-A$_2$pm** with a H_2O loss	**$^{13}C_7H_{12}{}^{15}N_2O_3$**	-0.55ppm
135.08770	***m*-A$_2$pm** with a "$H^{13}CO_2H$ + $^{15}NH_3$" loss	**$^{13}C_6H_9{}^{15}NO_2$**	-0.52ppm

(A) Tripeptide : *L*-Ala – γ-Glu – *m*-A_2pm **(B) Tetrapeptide : *L*-Ala – γ-Glu – *m*-A_2pm – *D*-Ala**



(C) Pentapeptide : *L*-Ala – γ-Glu – *m*-A_2pm – *D*-Ala – *D*-Ala

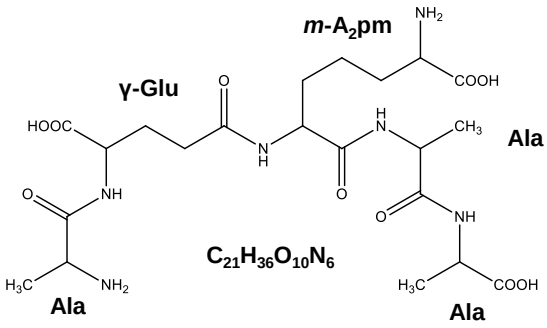


Figure S5: Structure and chemical formula of the tripeptide; **(B)** Structure and chemical formula of the tetrapeptide; **(C)** Structure and chemical formula of the pentapeptide

(A) Tripeptide : *L*-Ala – γ-Glu – *m*-A_2pm **(B) Tripeptide : *L*-Ala – γ-Glu – *m*-A_2pm***

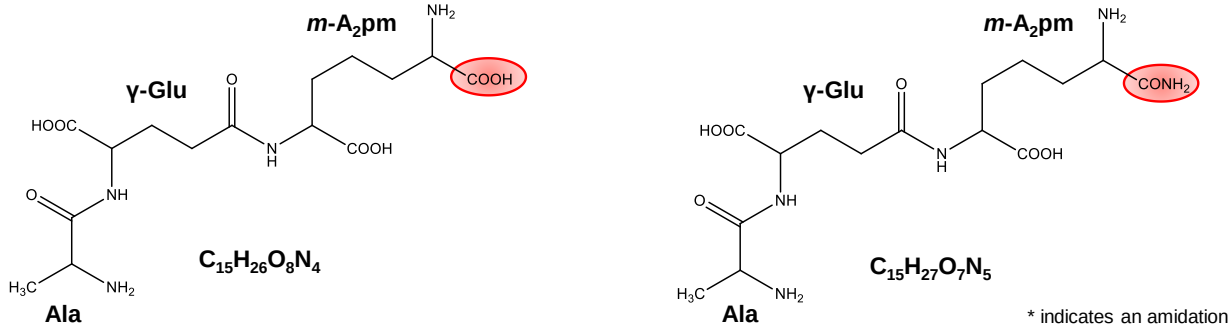


Figure S6: **(A)** Structure and chemical formula of the tripeptide and **(B)** Structure and chemical formula of the tripeptide bearing one amidation on its *m*-A_2pm

Demonstration of the pK_a estimation of the *m*-A_2pm in PGN-derived peptides:

The reference used for this demonstration is the book and eBook: "High Performance Capillary Electrophoresis: A primer" edited by Agilent Technologies, by B. Gaš, C. Sänger van de Griend, and P. Muijselaar.

- μ_{eff} : effective electrophoretic mobility (experimental electrophoretic mobility)
- α: dissociation or ionization coefficient, as ratio of ionized species content over non-ionized species content
- μ_e: electrophoretic mobility of the fully charged species in solution (i.e. maximum mobility)

$$\mu_{eff} = \alpha \times \mu_e$$

This formula is usually applied for weak mono-acids or mono-bases. Here, we only consider the free carboxylic moiety of the diaminopimelic acid to access its pK_a value. All amino-groups present in the species are considered to be fully ionized (i.e. in the form of ammonium groups $-NH_3^+$) at the pH of the background electrolyte (pH=2.38) because their respective pK_a are assumed to be greater than 7.

We also assume that the peptide bearing an amidation is never ionized on its *m*-A_2pm carboxylic moiety since amide groups cannot dissociate or accept a charge in aqueous solution. Because the species bearing the amidation presents the highest mobility (i.e. the lowest migration time), $\mu_{amidated}$ is in this case equivalent to μ_e.

The unmodified peptide (that is non-amidated species) has the lowest mobility (i.e. the largest migration time) as its carboxylic moiety is in equilibrium between carboxylate ($-COO^-$) and carboxylic form (-COOH) in solution: $\mu_{non\text{-}amidated}$ is therefore equivalent to μ_{eff}.

If we rewrite the above-mentioned equation with the introduced terms, we obtain:

$$\mu_{non-amidated} = \alpha \times \mu_{amidated}$$

We also know that:

$$\mu = \frac{l \times L}{MT \times E}$$

where "μ" is the experimental electrophoretic mobility, "l" and "L" are the capillary length and the capillary length until the detector (in the case of CZE-MS, these distances are identical), "MT" is the migration time, and "E" is the applied electric voltage on the capillary during the separation.

If we combine and rework the previous equations, we obtain:

$$\alpha = \frac{\mu_{non-amidated}}{\mu_{amidated}} = \frac{l \times L}{MT_{non-amidated} \times E} \times \frac{MT_{amidated} \times E}{l \times L} = \frac{MT_{amidated}}{MT_{non-amidated}}$$

As a result, the comparison of the migration times of the unmodified peptide and the peptide bearing an amidation reflects the ionization coefficient of the carboxylic group of the m-A_2pm.

We also know that:

$$\alpha = \frac{K_a}{[H^+] + K_a}$$

By re-writing this equation to isolate K_a, it yields:

$$K_a = \frac{\alpha \times [H^+]}{1 - \alpha}$$

$[H^+]$ can easily be found since $[H^+] = 10^{-pH}$ and $pK_a = -\log_{10}(K_a)$

The incorporation of the numerical MT values yields:

1) Tripeptide:

$$\alpha = \frac{22.11\text{min}}{33.61\text{min}} \approx 0.659$$

$$K_a = \frac{0.659 \times 10^{-2.38}}{1 - 0.659} \approx 8.22 \times 10^{-3}$$

$$\mathbf{pK_a} = -\log_{10}(8.22 \times 10^{-3}) \approx \mathbf{2.08}$$

2) Tetrapeptide:

$$\alpha = \frac{23.72\text{min}}{35.41\text{min}} \approx 0.670$$

$$\mathrm{K_a} = \frac{0.670 \times 10^{-2.38}}{1 - 0.670} \approx 8.46 \times 10^{-3}$$

$$\mathbf{pK_a} = -\log_{10}(8.46 \times 10^{-3}) \approx \mathbf{2.07}$$

3) Pentapeptide:

$$\alpha = \frac{24.61\text{min}}{36.60\text{min}} \approx 0.672$$

$$\mathrm{K_a} = \frac{0.672 \times 10^{-2.38}}{1 - 0.672} \approx 8.54 \times 10^{-3}$$

$$\mathbf{pK_a} = -\log_{10}(8.54 \times 10^{-3}) \approx \mathbf{2.07}$$

4.4. Conclusion and perspectives

In this chapter, the use of Capillary Zone Electrophoresis coupled to Mass Spectrometry as a valuable tool for the baseline separation of three PGN-derived peptides in crude bacterial cytoplasmic extracts has been demonstrated. In addition, the separation of their equivalents bearing an amidation on the meso-diaminopimelic acid (*m*-A_2pm) was also performed, as CZE provides a baseline resolution of these modified peptides.

The pK_a value of the *m*-A_2pm could be estimated based on the comparison of the migration time of the unmodified peptide and the equivalent bearing the amidation. The complete demonstration was included into the Supporting Information (see Appendix). The estimated pK_a values of the carboxylic moiety of the diaminopimelic acid is then around 2.1, which is consistent with the value found in Chapter 3 for muropeptides ($pK_a \approx 2.2$).

In complement, the quantitation of the endogenous tripeptide was performed based on the addition of ^{13}C ^{15}N tripeptide standard prior to the analysis. The results show a high repeatability of the analytical triplicates, highlighting the robustness of the method, even with highly complex samples such as crude bacterial cytoplasmic extracts.

The comparison between samples where bacteria were grown in the presence of antibiotics with samples where the bacteria were grown in the absence of antibiotics shows a slight increased concentration (+20%) of the tripeptide in the samples grown with antibiotics, although the relatively high "biological" standard deviation does not make this result statistically significantly different (RSD≈16%). This result is in agreement with reported evidence of an increase in the generation of PGN fragments in the presence of β-lactam antibiotics in *E. coli*. This result shows the great potency of the developed method to follow the intensity of multiple PGN-derived peptide levels in bacteria grown in different stress conditions. As a result, the developed methodology could provide insights in the different PGN-related mechanisms / pathways inside bacteria.

Finally, the proposed method for the quantitation could be implemented for any cytoplasmic PGN-derived peptides of interest (including the peptides bearing one amidation) as long as an isotopically labelled standard of known concentration is available.

This work provides a robust and easy method for the assessment of PGN-derived peptides in crude cytoplasmic extracts. In particular, the baseline separation of the equivalents bearing an amidation on the *m*-A_2pm, which is not accessible by classical reversed-phase HPLC, is of particular biochemical importance, since such chemical modification was linked to antibiotics resistance and PGN recognition by the mammalian host innate immune system. Finally, the absolute quantitation of these analytes based on the incorporation of the isotopically labelled standard was also demonstrated on the case of the tripeptide. This method could be expanded to the other PGN-derived peptides, as long as the isotopically labelled equivalents are available.

CHAPTER 5

Combination of Capillary Zone Electrophoresis-Mass Spectrometry, Ion Mobility-Mass Spectrometry, and Theoretical Calculations for cysteine connectivity identification in peptides bearing two intramolecular disulfide bonds

5.1. Context of the chapter

Disulfide bonds are post translational modification playing essential roles in the biological activity and stability of numerous peptides and proteins. Intra-molecular disulfide bonds are found in various natural-occurring peptides such as animal venoms. In such peptides, the appropriate cysteine connectivity provides the required conformation for the efficient binding to the molecular targets, which ensures their bioactivity. The characterization of cysteine pairing is still a challenging issue in the analysis of peptides targeting pharmaceutical or pharmacological utilizations. Thus, the use of sensitive, robust and efficient characterization techniques to assess the cysteine pairing is crucial.

The recent introduction of the gas-phase separation of disulfide isomers of peptides bearing two intra-molecular disulfide bonds (i.e. peptides with identical amino acid sequence but different cysteine connectivities) by Massonnet and coworkers using TWIM-MS paved the way for the use of mobility-based techniques for this challenging analytical issue. As demonstrated in the following Figure extracted from the article "Ion Mobility Mass Spectrometry as a tool for the structural characterization of peptides bearing intramolecular disulfide bonds" by Massonnet et al, JASMS (2016), 27 (10), 1637–1646, TWIM-MS was able to discriminate the three different disulfide isomers of the same peptide under appropriate IMS conditions:

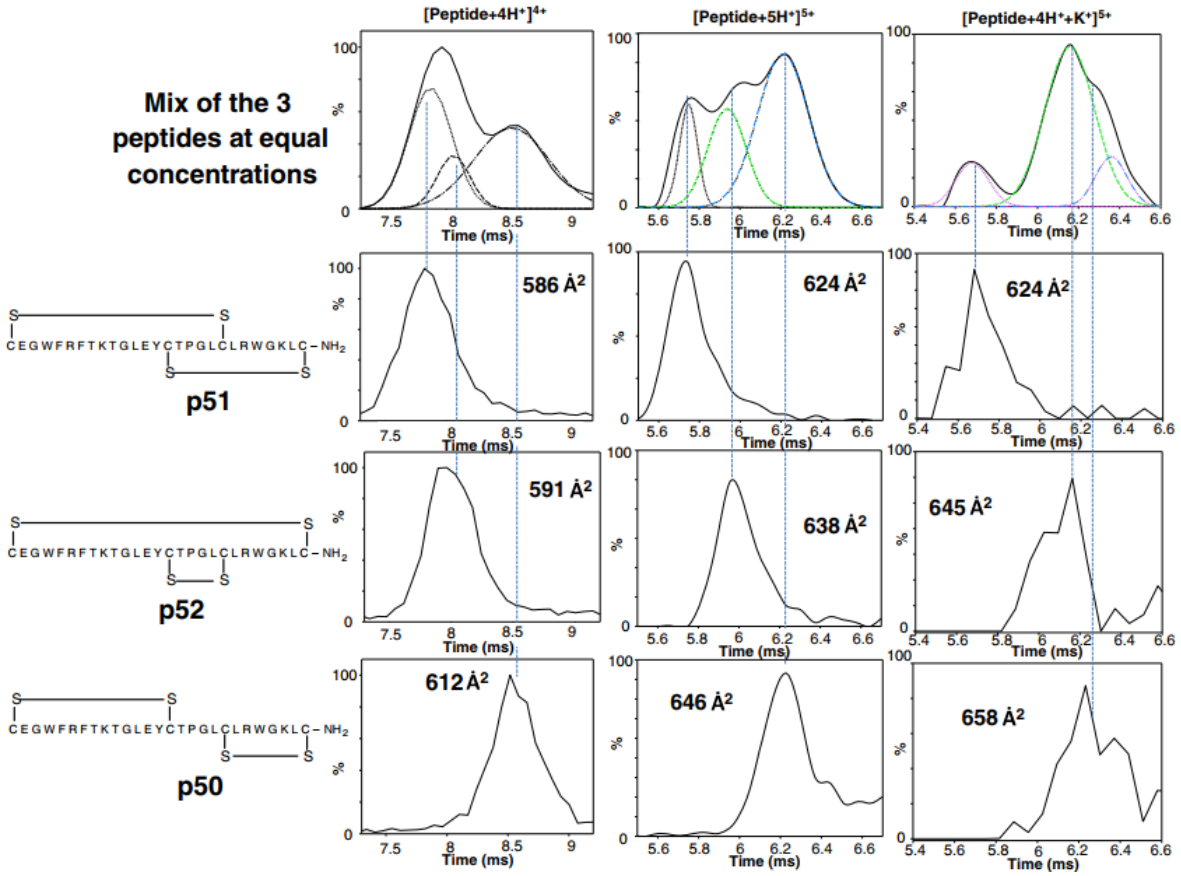
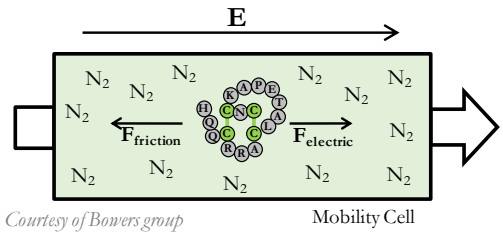


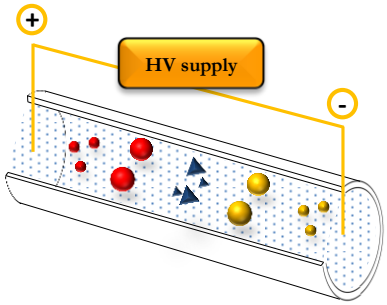
Figure 5.1 - Plots of the TWIM-MS ATDs of different peptide adducts ($[M+4H]^{4+}$, $[M+5H]^{5+}$ and $[M+4H+K]^{5+}$) for sequence-identical peptides having different cysteine connectivities from Massonnet *et al.* Ion Mobility-Mass Spectrometry as a Tool for the Structural Characterization of Peptides Bearing Intramolecular Disulfide Bond(S). J. Am. Soc. Mass Spectrom. 2016, 27 (10).

In this chapter, the separation of the disulfide isomers of three peptides of various lengths bearing two intra-molecular disulfide bonds but different cysteine connectivities have been tested using Capillary Zone Electrophoresis (CZE) and Travelling Wave Ion Mobility (TWIM) coupled to Mass Spectrometry (MS). As indicated in Figure 5.2, these two separation techniques are mobility-based techniques (TWIM in the gas phase, CZE in solution). Consequently, the relevant separation parameters (highlighted in red in Figure 5.2) are the charge of the analyte and a structural parameter which is the collision cross section (CCS) in TWIM and the hydrodynamic radius R_h in CZE.

Ion Mobility :
Separation **in the gas phase** according to both charge (**q**) and collision cross section (**Ω**)

Capillary Electrophoresis :
Separation **in solution** according to both charge (**q**) (pH dependent) and hydrodynamic radius ($\mathbf{R_h}$)





$$F_{electric} = F_{friction} \rightarrow v_{stat} = \text{mobility constant}.\ E$$

$$K = \frac{3q}{16N}.\left(\frac{2\pi}{kT}\right)^{\frac{1}{2}}.\left(\frac{m+M}{mM}\right)^{\frac{1}{2}}.\left(\frac{1}{\Omega}\right) = \text{constant}.\left(\frac{m+M}{mM}\right)^{\frac{1}{2}}.\frac{q}{\Omega}$$

$$\mu_e = \frac{q}{6\pi\eta R_h} = \frac{1}{6\pi\eta}.\frac{q}{R_h}$$

Figure 5.2 - Schematic representation of IMS and CZE along with the respective separation equations

In complement to these experimental approaches, a theoretical calculation method was also implemented to get insights into the structures of these highly constrained peptides in solution and in the gas phase. This theoretical method was based on the optimization of the structures of each disulfide isomer in vacuo and in solution to extract various physicochemical parameters. From these values, the separation efficiency of each method was assessed. Finally, the comparison of gas phase and in-solution results and in particular, the respective migration orders allowed addressing potential structural changes after desolvation by the electrospray ionization.

The results are presented in the following section as a submitted research article in **Analytical Chemistry** (submitted on July 15th, 2019). The related Supporting Information (SI) are presented directly after the article. In the present work, the IM-MS experiments were performed by Philippe Massonnet and Christopher Kune, the CZE-MS(MS) experiments were performed by myself and the theoretical calculations were performed by Christopher Kune and Johann Far.

5.2. Results – research article submitted in Analytical Chemistry

Combination of Capillary Zone Electrophoresis-Mass Spectrometry, Ion Mobility-Mass Spectrometry, and Theoretical Calculations for cysteine connectivity identification in peptides bearing two intramolecular disulfide bonds

Cédric Delvaux(1)† and Philippe Massonnet(1)(2)†, Christopher Kune(1), Jean R.N. Haler(1)(3), Gregory Upert(4), Gilles Mourier (4), Nicolas Gilles(4), Loïc Quinton(1), Edwin De Pauw(1) and Johann Far(1)

(1) Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium

(2) Maastricht Multimodal Molecular Imaging (M4I) Institute, Division of Imaging Mass Spectrometry, Maastricht, Limburg, Netherlands

(3) Department of Chemistry and Biochemistry, Florida International University, Miami, Florida 33199, USA

(4) Commissariat à l'Energie Atomique, DRF/Institut Joliot/SIMOPRO, Université Paris Sud, 91191, Gif-sur-Yvette, France

† Cédric Delvaux and Philippe Massonnet contributed equally as first authors

AUTHOR EMAIL ADDRESS: c.delvaux@uliege.be

ABSTRACT

Disulfide bonds between cysteine residues are commonly involved in the stability of numerous peptides and proteins and are crucial for providing biological activities. In such peptides, the appropriate cysteine connectivity ensures the proper conformation allowing an efficient binding

to their molecular targets. Disulfide bond connectivity characterization is still challenging and is a critical issue in the analysis of structured peptides/proteins targeting pharmaceutical or pharmacological utilizations. This study describes the development of new and fast gas-phase and in-solution electrophoretic methods coupled to mass spectrometry to characterize the cysteine connectivity of disulfide bonds. For this purpose, disulfide isomers of three peptides bearing two intra-molecular disulfide bonds but different cysteine connectivity have been investigated. Capillary Zone Electrophoresis (CZE) and Ion Mobility (IM) both coupled to Mass Spectrometry (MS) were used to perform the separation in both aqueous and gas phases, respectively. The separation efficiency of each technique has been critically evaluated and compared. Finally, theoretical calculations were performed to support and explain the experimental data based on the predicted physicochemical properties of the different peptides.

INTRODUCTION

Disulfide bonds are post-translational modifications (PTMs) playing crucial roles in the foundation of biological activities and in the stabilization of peptide or protein structures[1,2] . Intramolecular disulfide bonds are found in several naturally occurring peptides such as cyclotides[3–6], defensins[7–9] or animal venoms[10–12]. The native pairing of intramolecular disulfide bonds is in most cases mandatory to preserve the biological activity[13,14]. Indeed, misfolded variants generally lead to reduced biological activity and are degraded or recycled by enzymes to the native form[15]. For these reasons, the use of efficient characterization techniques to unambiguously access the cysteine pairing is of crucial importance. For this purpose, analytical methods involving nuclear magnetic resonance (NMR)[16,17], X-Ray crystallography[18] or enzymatic digestion followed by liquid chromatography and mass spectrometry analysis[19,20] have been developed. These strategies, while effective, present several disadvantages including large time and/or large sample consumption. In some cases, these traditional methods can even lead to an erroneous assignment of the disulfide bond connectivity, which then requires

confirmation or rectification by an alternative method. In a recent example, the NMR disulfide bonds connectivity of peptides found in marine sponges has to be rectified by an alternative approach based on total chemical synthesis followed by HPLC[21].

More recently, gas phase fragmentation techniques coupled to mass spectrometry have been introduced on peptides bearing intramolecular disulfide bond(s)[22–27] for disulfide connectivity elucidation. In this context, tandem mass spectrometry (MS/MS) such as Collision Induced Dissociation (CID)[22,25] and Electron Transfer Dissociation (ETD)[23,24,26] have been used successfully. However, these methods lead to poorly fragmented precursors when this strategy is applied on peptides with at least two intramolecular disulfide bonds. Consequently, the manual interpretation of complex fragmentation spectra is required, which hinders the systematic unambiguous disulfide bond connectivity assignment.

Alternative analytical methods are then required to overcome these various issues. Interestingly, different groups showed that ion mobility-based techniques (*e.g.* gas-phase ion mobility and liquid-phase capillary electrophoresis) could be successfully used for disulfide bond connectivity assignment in peptides and proteins[27,28]. In these studies, Capillary Zone Electrophoresis (CZE) and Capillary Gel Electrophoresis (CGE) were used for the characterization of disulfide variants of recombinant IgG2 monoclonal antibodies[28,29]. More recently, He *et al* used a set of CZE-based methods for the analysis of identity, charge variants, and disulfide isomers of IgG monoclonal antibodies (mAbs)[30]. On the other hand, traveling wave ion mobility was also used in the gas phase for the separation of disulfide-bridged peptide isomers[31,32].

Capillary Zone Electrophoresis coupled to Mass Spectrometry (CZE-MS) presents several advantages over traditional methods such as NMR, X-ray crystallography or LC-MS(/MS). First, the significantly lower injected volume needed for CE analyses (usually in the range of

tens of nanoliters) makes it particularly suited for samples available in very limited amounts such as toxins extracted from tiny venomous animals, e.g. bees, ants or wasps. Moreover, because CZE-based methods allow working with bare fused silica capillaries instead of expensive liquid chromatography (LC) columns, this analytical method offers a simple, organic solvent free, and economical alternative compared to LC.

We show here that liquid-phase CZE coupled to MS can be used as a powerful complementary analytical method to gas-phase ion-mobility mass spectrometry (IM-MS) for the characterization of disulfide isomers both in the gas phase and in aqueous solutions. The complementarity of the techniques and their respective drawbacks are discussed. A workflow including molecular modeling was also introduced to support the experimental results obtained in the gas phase and in solution.

MATERIALS AND METHODS

Peptide synthesis

All peptides used in this study are listed in table 1. Apamin was purchased from Sigma-Aldrich (Bornem, Belgium). All other peptides were synthesized following a previously reported synthesis protocol[24,33].

*Table 1: List of the peptides used in this study with their sequences, connectivity of the cysteine residues, and molecular masses. All peptides bear two intramolecular disulfide bonds. * indicates a C-terminal amidation.*

Peptide name	Sequence[a]	Cysteine connectivity[a]	Mass (Da)
ModGlo[b]	**C$_1$**EGWFRFTKTGLEY**C$_2$**TPGL**C$_3$**LRWGKL**C$_4$**	**C$_1$-C$_3$/C$_2$-C$_4$**	3161.49
ModRib[b]	**C$_1$**EGWFRFTKTGLEY**C$_2$**TPGL**C$_3$**LRWGKL**C$_4$**	**C$_1$-C$_4$/C$_2$-C$_3$**	3161.49
ModBea[b]	**C$_1$**EGWFRFTKTGLEY**C$_2$**TPGL**C$_3$**LRWGKL**C$_4$**	**C$_1$-C$_2$/C$_3$-C$_4$**	3161.49
Apamin	**C$_1$**N**C$_2$**KAPETAL**C$_3$**ARR**C$_4$**QQH*	**C$_1$-C$_3$/C$_2$-C$_4$**	2025.88
ApaRib	**C$_1$**N**C$_2$**KAPETAL**C$_3$**ARR**C$_4$**QQH*	**C$_1$-C$_4$/C$_2$-C$_3$**	2025.88
ApaBea	**C$_1$**N**C$_2$**KAPETAL**C$_3$**ARR**C$_4$**QQH*	**C$_1$-C$_2$/C$_3$-C$_4$**	2025.88
α-conotoxin	**C$_1$C$_2$**HSSW**C$_3$**KHL**C$_4$**	**C$_1$-C$_3$/C$_2$-C$_4$**	1301.46
χ-conotoxin	**C$_1$C$_2$**HSSW**C$_3$**KHL**C$_4$**	**C$_1$-C$_4$/C$_2$-C$_3$**	1301.46

[a] the cysteine residues are identified by the number in subscript as the first, second, third or fourth cysteine residue present in the sequence from the N terminus to the C terminus

[b] These peptides are synthetic peptides whose primary sequence was selected to provide optimized ionization efficiency using electrospray ion source

Reagents

All background electrolytes reagents (formic acid, acetic acid, and ammonium acetate) and solvents (isopropanol and methanol) were purchased from Sigma-Aldrich (Bornem, Belgium) and were "metal trace analysis" grade. A Milli-Q ultra-pure water system from Millipore (Millipore, Molsheim, France) was used throughout the study.

Ion Mobility - Mass Spectrometry (IM-MS)

Ion mobility experiments were performed on a SYNAPT G2 HDMS mass spectrometer (Manchester, UK) equipped with a traveling-wave ion mobility cell in positive ion mode using an electrospray (ESI) source. For all peptides, the capillary and sampling cone voltages were respectively 3.0kV and 20V. All peptides were infused at 1µmol.L^{-1} in a H_2O/Acetonitrile (50/50 v/v) solvent mixture spiked with 0.1% formic acid. Data acquisition and data processing were performed using MassLynx v4.1.

For the apamin isomers (Apamin, ApaRib and ApaBea), the wave velocity and the wave height in the ion mobility cell were, respectively, 2000 m.s^{-1} and 30 V, while the wave height and the wave velocity in the transfer cell were, respectively, 508 m/s and 0.1 V. The wave velocity and wave height in the trap were, respectively 500 m.s^{-1} and 0.5 V. The IMS cell pressure (N_2) was 1.64 mbar while the trap and the transfer pressures were 0.53 mbar of argon. Finally, the He cell pressure was 0.18 mbar.

For the conotoxins (conotoxin α and conotoxin χ), the wave velocity and the wave height in the ion mobility cell were, respectively, 1250 m.s^{-1} and 30 V, while the wave height and the wave velocity in the transfer cell were respectively 508 m.s^{-1} and 0.1 V. The wave velocity and wave height in the trap were, respectively, 508 m.s^{-1} and 0.5 V. The IMS cell pressure (N_2) was 2.00 mbar while the trap and the transfer pressures were 0.61 mbar. Finally, the He cell pressure was 0.24 mbar.

Capillary Zone Electrophoresis-Mass Spectrometry (CZE-MS)

A CESI8000 instrument (Beckman Coulter – SCIEX Separations) equipped with the External Detector Adaptor and a bare fused-silica capillary of 30µm inner diameter and 150µm outer diameter with a total length of 90cm were used. 32karats v10.1 was used to operate the CE apparatus. ESI-MS detection was performed on a LTQ FT Ultra 7T Hybrid Mass Spectrometer (Thermo Scientific Merelbeke, Belgium) instrument operating in positive ion mode (+2.5kV) with both Linear Trap Quadrupole (LTQ) and Fourier Transform – Ion Cyclotron Resonance (FT-ICR) mass analyzers. Mass ranges were set to 300-2000 m/z to reduce chemical noise during data acquisition. XCalibur v2.2 was used for data acquisition and data processing. The coupling between CZE and electrospray MS was performed using a microfluidic SL CE-MS coaxial sprayer from Analis (PN 10-301347, Analis, Suarlée, Belgium).

All details regarding the capillary activation, the BackGround Electrolyte (BGE) and Sheath Liquid (SL) compositions are detailed in the Supporting Information.

All CZE-MS runs were performed using the same procedure: the different peptides used in this study (listed in table 1) were co-injected with a reference peptide before the migration. The use of a reference peptide allows working with Relative Migration Times (RMT) instead of absolute migration times. The RMTs were calculated by dividing the migration time of the disulfide isomer by the migration time of the reference peptide as shown in the equation (1):

$$\text{RMT} = \frac{\text{absolute migration time (analyte)}}{\text{absolute migration time (reference)}} \quad (1)$$

The reference peptides used in this study were chosen for their availability and widespread use as MS calibration compounds. For this purpose, [Glu^1]-Fibrinopeptide B human or Glufib (EGVNDNEEGFFSAR, MW = 1570.57 Da) and MRFA (MRFA, MW = 523.65 Da) were selected as reference peptides. The reference peptide was selected according to the optimized pH conditions for efficient separation of the disulfide isomers.

All data were collected and treated as follows. First, each pure disulfide isomer was injected in the presence of a reference peptide and electrophoretically migrated using the optimized BGE (Supporting Information – Table S1). The extracted electropherograms were recorded following the most abundant charge state detected in the MS spectrum. The RMTs were calculated for each run and an average RMT and its standard deviation were calculated for 6 replicate measurements. In a second series of experiments, the mixture of all the disulfide isomers was injected and separated using the same method than previously described. The different isomers in the mixture were attributed by matching the RMTs of the pure disulfide isomers and the RMTs of the different isomers detected in the mixture. pH optimization of the BGE was required to obtain a baseline separation of the apamin and the conotoxin isomers.

Theoretical Calculations: Structure Optimization using Molecular Mechanics

Part 1: Structure optimization in vacuo (gas-phase optimization)

In order to gain a better understanding of the measured arrival time distributions (IM-MS), theoretical calculations were performed using Molecular Mechanics (MM). For the gas-phase calculations, the structure of each disulfide isomer of each peptide was first drawn according to both its primary sequence and specific disulfide pattern. Chem3D software (Perkin Elmer) was used for low-energy conformation search using the Mercks Molecular Force Field MMFF94[34]. The structure of the most stable conformation (*i.e.* with the lowest steric energy) of each peptide was then refined using Gaussian09d software[35] using the included Assisted Model Building with Energy Refinement (AMBER) force field[36]. The theoretical Collision Cross Sections (CCS) of these structures in N_2 were finally calculated with IMoS[37,38] using the exact Hard Sphere Scattering model (EHSS). The workflow of these calculations is detailed in the Supporting Information (Supporting Information - Figure S1).

Part 2: Structure optimization in water (solution optimization)

Similarly to the workflow used for the gas-phase modeling, the structure of each disulfide isomer used in this study was drawn according to its primary sequence and specific disulfide pattern. From this non-optimized linear structure, an assessment of the charge state distribution as a function of pH was calculated based on the primary structure using the Henderson-Hasselbalch equation. The structure of each disulfide isomer was then iteratively optimized using Gaussian09d by MM with the AMBER force field and the Polarizable Continuum Model[39] (PCM) as a model to mimic the solvent effect for the different considered charge states. The pK_a shifts of the residues in the computed isomer due to the folding were assessed after computing the new pK_a values using H++[40–42] (http://biophysics.cs.vt.edu) and the preferentially charged residues were selected before structure re-optimization using Gaussian09d. When several residues could carry the charge, all of the possible structures were computed using Gaussian09d for structure optimization and the lowest total energy candidate was selected for the iterative structure optimization process. Typically, these apamins, conotoxins, and synthetic peptides required less than 10 iterations until the pK_a values of the residues became stable. Finally, the individual gyration radii of each different structure (e.g. every considered charge state for each optimized structure of a disulfide isomer) were determined using PyMol[43] v1.7 and the rgyrate python script. Because of the charge state of the peptides in solution is not an integer value, the final average gyration radii were calculated by a gyration radii ponderation by the averaged in-solution charge states (for example, if the considered charge state is 2.5, the final gyration radius was calculated using 50% of the gyration radius of the 2+ charge state and 50% of the gyration radius of the 3+ charge state). Finally, the gyration radii were converted into the corresponding hydrodynamic radii assuming a globular shape conversion factor ($R_g = (3/5)^{0.5} \times R_h \approx 0.775 \times R_h$)[44] and the electrophoretic mobilities were estimated from the Stokes-Einstein equation (Supporting Information - Figure S2).

RESULTS AND DISCUSSION

IM-MS and CZE-MS characterization of the peptide isomers

In an effort to evaluate the efficiency for the identification of the cysteine residues connectivity by CZE-MS (mobility measurement in solution) and IM-MS (mobility measurement in the gas phase), three peptides of various lengths (see Table 1) were analyzed by both techniques. These peptides all contain four cysteine residues within their structures and present three possible disulfide connectivity patterns (except for the conotoxin peptides where only two disulfide patterns were available). The same workflow was applied to the three systems. First, the analysis of the arrival time distributions provided by IM-MS was performed to retain the most efficient separation according to the different charge states observed in the gas phase (see Figures 2 and 3 for details). The experimental results were then supported by theoretical calculations of the CCS values on the studied charge states. In detail, each disulfide isomer of the different peptides was modeled in the gas phase. Charge states were attributed to match the experimental charge state by protonating the appropriate residues according to the proton affinity values of each residue[45]. After structure optimization, the CCS values were computed, and the separation efficiency of IM-MS was assessed in each case based on computed theoretical CCS values.

Secondly, CZE-MS experiments were performed on the same peptides for the separation of the disulfide isomers in solution. The pH of the background electrolyte in CZE-MS was optimized to obtain the efficient separation of the isomers. The experimental results were supported by computational chemistry to calculate the theoretical average charge in solution and the optimized tridimensional structures were used to derive their respective theoretical gyration radii and the average hydrodynamic radii at the experimentally optimized pH. These values allowed the estimation of electrophoretic mobility values, leading to the assessment of the

theoretical separation capabilities of CZE-MS in each case. The results of this workflow are presented in Figures 1, 2 and 3.

Part 1: Synthetic Model Peptide

First, the CZE-MS method development was performed on a synthetic 27-residue peptide whose disulfide isomers have recently been characterized by IM-MS[31]. In this work, Massonnet and co-workers showed that IM-MS enables a partial separation of the different disulfide isomers using the 4+ and 5+ charge states from proton adducts. The three different disulfide patterns of this peptide were labeled *"ModBea"* (C_1-C_2/C_3-C_4), "*ModGlo"* (C_1-C_3/C_2-C_4) and *"ModRib"* (C_1-C_4/C_2-C_3) following the commonly used Globular, Ribbon and Bead nomenclature. This peptide was selected as the first candidate to evaluate the performance of CZE-MS compared to those recently published using IM-MS[31]. The results are presented in Figure 1.

(A) **Determination of RMT (separate disulfide isomers)**

Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
ModGlo	0.6974	0.0016	0.23%
ModBea	0.7065	0.0009	0.12%
ModRib	0.7288	0.0011	0.15%

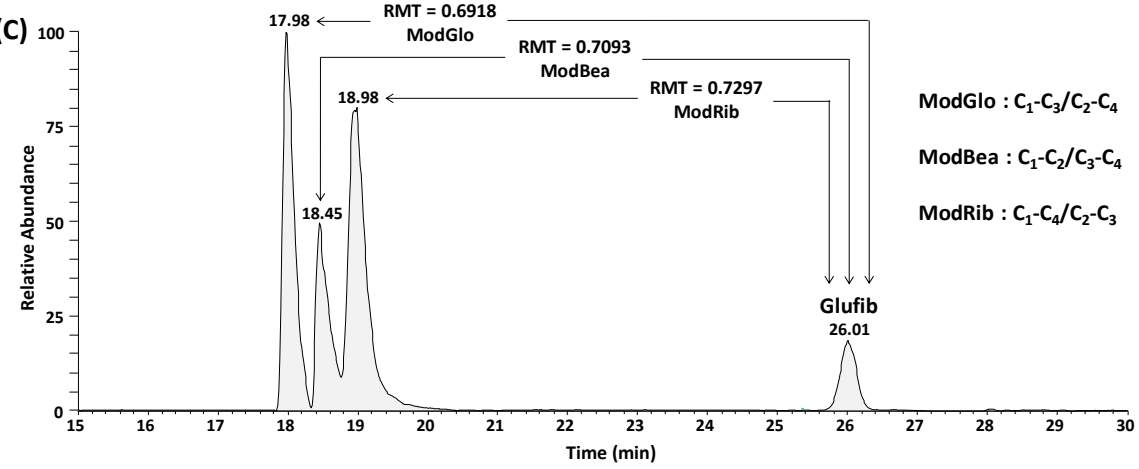
RMT = relative migration time σ = standard deviation
n = number of replicates

(B) **Determination of RMT (mix of the disulfide isomers)**

Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
ModGlo	0.6939	0.0028	0.31%
ModBea	0.7123	0.0023	0.32%
ModRib	0.7317	0.0019	0.25%

RMT = relative migration time σ = standard deviation
n = number of replicates

(C)



(D)

Peptide	Experimental ATD (ms)[a]	Charge state	Theoretical CCS (Å²)	Experimental RMT	Theoretical charge state at pH 2.26[b]	Theoretical R_h (Å)	Theoretical μ_e $m^2.(V.s)^{-1}$
ModGlo	5.70	5+	708	0.6939	+4.7	12.2	3.24×10^{-8}
ModRib	5.95	5+	722	0.7317	+4.9	13.9	2.96×10^{-8}
ModBea	6.23	5+	777	0.7123	+4.8	13.1	3.13×10^{-8}

Gas phase | Solution

Figure 1: *CZE-MS determination of the relative migration time (RMT) of each disulfide isomer of the model peptide in 10% acetic acid - 20% isopropanol BGE,* ***(A)*** *performed on 6 separate replicates and* ***(B)*** *on 6 replicates of the isomer mix* ***(C)*** *Typical CZE-MS electropherogram obtained for a run in 10% acetic acid - 20% isopropanol BGE of a mix of the 3 disulfide isomers of the model peptide co-injected with Glufib as internal standard (the calculated RMTs allow isomer identification).* ***(D)*** *Table of the model peptides with their experimentally measured IM-MS arrival time distributions[a] at the considered charge state, the theoretically predicted CCS obtained from the corresponding optimized structure, the experimentally measured CZE-MS RMT obtained at the optimized pH, the theoretically predicted average charge state and theoretically predicted average hydrodynamic radius (R_h) from the corresponding optimized structure, and the theoretically predicted electrophoretic mobility (μ_e)*

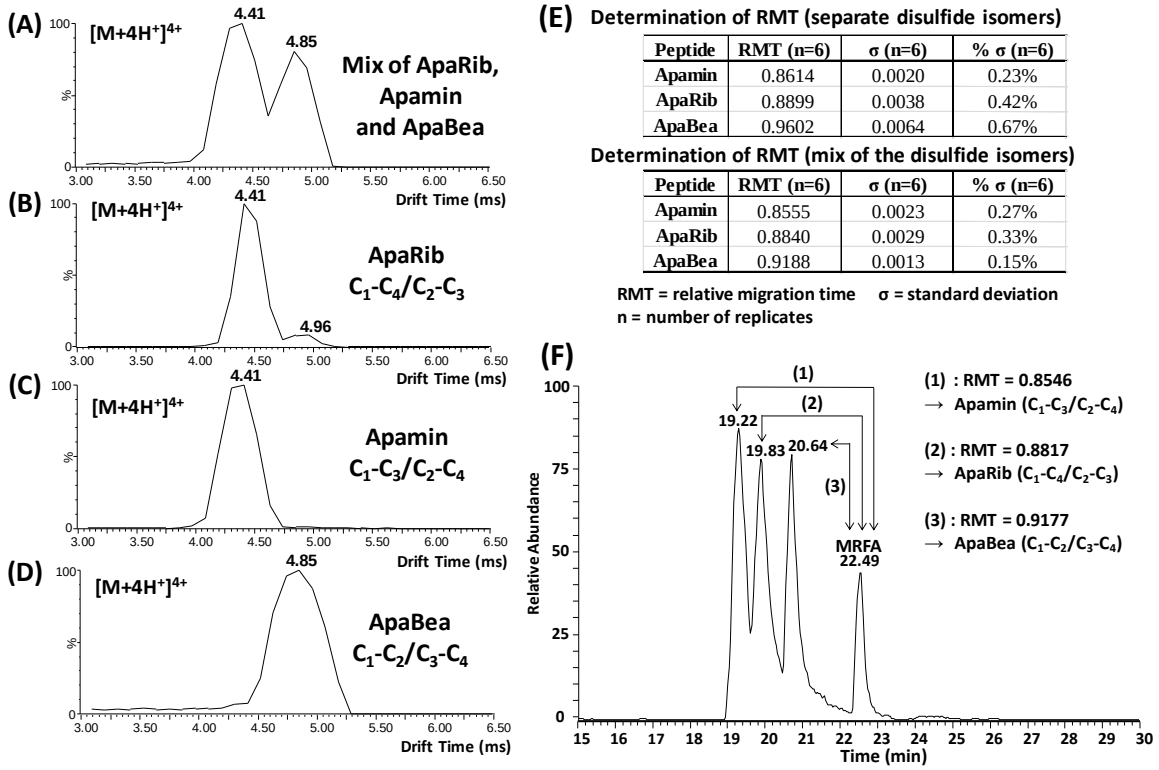
[a] measured arrival time distributions (ATDs) from Massonnet, P. et al, Ion Mobility-Mass Spectrometry as a Tool for the Structural Characterization of Peptides Bearing Intramolecular Disulfide Bond(S). J. Am. Soc. Mass Spectrom. 2016, 27 (10), 1637–1646 (Figure 6).

[b] the charge state of these species was estimated at pH 2.26, the calculated pH for a 10% acetic acid background electrolyte (BGE), therefore not considering the influence of isopropanol on the pH of the BGE

Part 2: Apamin

The first biological peptide analyzed during this study, i.e. the apamin, is a naturally occurring C-amidated 18-residue peptide found in bee venom[46]. The amino acid sequence presents four

cysteines allowing three possible folds, labeled *"ApaBea"* (C_1-C_2/C_3-C_4), the naturally occurring "*Apamin"* (C_1-C_3/C_2-C_4) and *"ApaRib"* (C_1-C_4/C_2-C_3). Figure 2 summarizes the experimental and theoretical results obtained for this peptide in the optimized conditions.



(G)

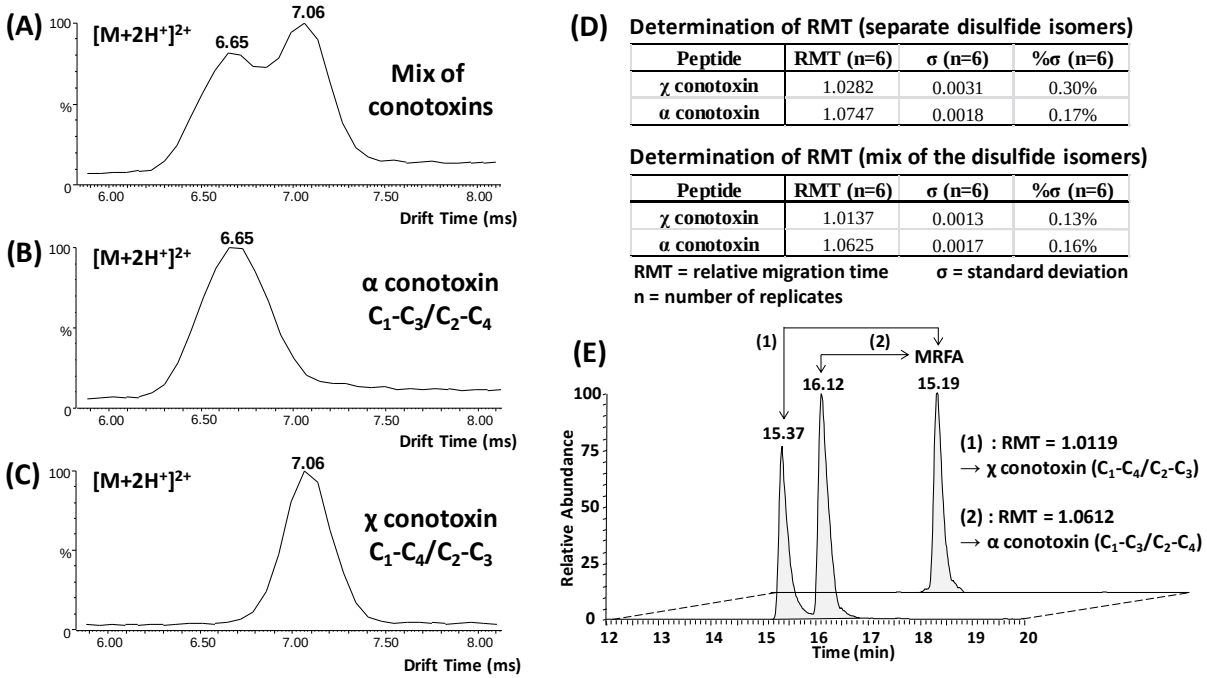
Peptide	Experimental ATD (ms)	Charge state	Theoretical CCS (Å²)	Experimental RMT	Theoretical charge state at pH 6.98	Theoretical R_h (Å)	Theoretical μ_e $m^2.(V.s)^{-1}$
ApaRib	4.41	4+	490	0.8840	+3.1	9.3	2.77×10^{-8}
Apamin	4.41	4+	504	0.8555	+2.9	9.2	2.80×10^{-8}
ApaBea	4.85	4+	619	0.9188	+2.9	9.6	2.59×10^{-8}

Gas phase | Solution

Figure 2: *IM-MS-obtained arrival time distributions (ATDs) of the 4+ charge state of* ***(A)*** *a mix of the 3 apamin peptides (Apamin, ApaBea and ApaRib);* ***(B)*** *arrival time distributions (ATDs) of the 4+ charge state of ApaRib;* ***(C)*** *arrival time distributions (ATDs) of the 4+ charge state of Apamin;* ***(D)*** *arrival time distributions (ATDs) of the 4+ charge state of ApaBea;* ***(E)*** *CZE-MS determination of the relative migration time of each disulfide isomer of apamin on 6 separate replicates (upper part) followed by 6 replicates of the isomer mix (lower part) in 100mM ammonium acetate BGE (pH 6.98) co-injected with MRFA (the calculated RMTs allow isomer identification)* ***(F)*** *Typical CZE-MS electropherogram obtained for a run in 100mM ammonium acetate BGE of a mix of the 3 disulfide isomers of Apamin co-injected with MRFA (the calculated RMTs allow isomer identification)* ***(G)*** *Table of the apamins with their experimentally measured IM-MS arrival time distributions (IMS) at the considered charge state, the theoretically predicted CCS obtained from the corresponding optimized structure, the experimentally measured CZE-MS RMT obtained at the optimized pH, the theoretically predicted average charge state and theoretically predicted average hydrodynamic radius (R_h) from the corresponding optimized structure, and the theoretically predicted electrophoretic mobility (μ_e)*

Part 3: Conotoxins

Another biological peptide used in this study is a naturally occurring 11-residue conotoxin found in the venom of marine cone snails[47]. The amino acid sequence also displays four cysteines as presented in Table 1. Because of the vicinal position of the two first cysteines within the sequence, the chemical synthesis of the disulfide isomer C_1-C_2/C_3-C_4 was not attempted, therefore leading to the analysis of only two disulfide isomers, which are respectively labelled *"α conotoxin"* (C_1-C_3/C_2-C_4) and the naturally occurring *"χ conotoxin"* (C_1-C_4/C_2-C_3), in accordance with the nomenclature of conotoxins[48]. Figure 3 summarizes the experimental and theoretical results obtained for this peptide in the optimized conditions.



(F)

Peptide	Experimental ATD (ms)	Charge state	Theoretical CCS (Å²)	Experimental RMT	Theoretical charge state at pH 6.98	Theoretical R_h (Å)	Theoretical μ_e $m^2.(V.s)^{-1}$
α conotoxin	6.65	2+	346	1.0625	+1.1	8.3	1.19×10^{-8}
χ conotoxin	7.06	2+	353	1.0137	+1.3	8.6	1.32×10^{-8}
			Gas phase			Solution	

***Figure 3**: **(A)** IM-MS-obtained arrival time distributions (ATDs) of the 2+ charge state of a mix of the conotoxins; **(B)** arrival time distributions (ATDs) of the 2+ charge state of the α-conotoxin; **(C)** arrival time distributions (ATDs) of the 2+ charge state of χ-conotoxin; **(D)** CZE-MS determination of the relative migration time of each disulfide isomer of apamin on 6 separate replicates (upper part) followed by 6 replicates of the isomer mix (lower part) in 50mM ammonium acetate BGE (pH 6.98) co-injected*

with MRFA (the calculated RMTs allow isomer identification) ***(E)*** *Typical CZE-MS electropherogram obtained for a run in 50mM ammonium acetate BGE of a mix of the 2 conotoxins co-injected with MRFA (the calculated RMTs allow isomer identification); the electropherogram of MRFA in* ***(E)*** *was extracted separately from the electropherogram of the conotoxins for better visualization.* ***(F)*** *Table of the conotoxins with their experimentally measured IM-MS arrival time distributions (IMS) at the considered charge state, the theoretically predicted CCS obtained from the corresponding optimized structure, the experimentally measured CZE-MS RMT obtained at the optimized pH, the theoretically predicted average charge state and theoretically predicted average hydrodynamic radius (R_h) from the corresponding optimized structure, and the theoretically predicted electrophoretic mobility (μ_e)*

pH influence on the efficiency of the CZE-MS separations

The pH of the background electrolyte is a critical parameter to obtain an efficient separation of the peptide isomers. Indeed, pH influences the average charge state in solution of the peptides and therefore directly affects the separation efficiency of the method. Figure 4 shows the influence of a pH change for conotoxins from an unresolved isomers mixture in acidic conditions to a baseline separation in neutral conditions.

(A) Determination of RMT (separate disulfide isomers)

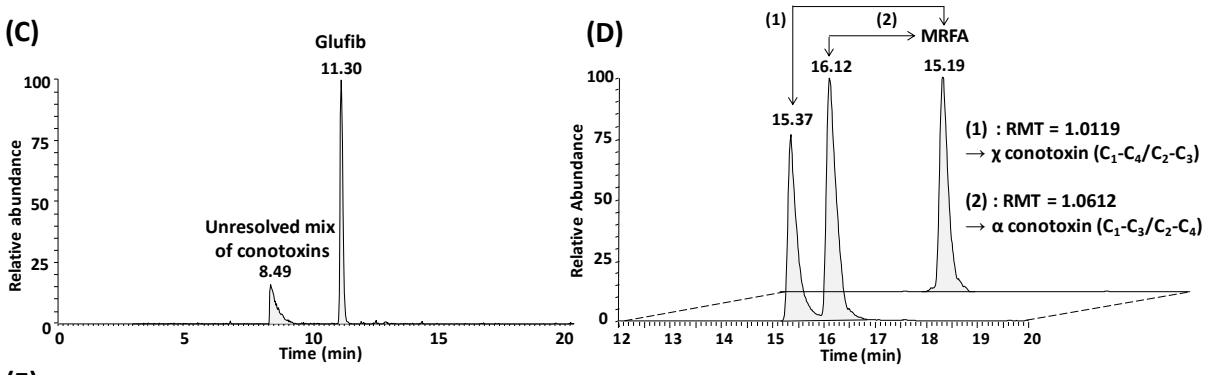
Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
χ conotoxin	0.7040	0.0009	0.13%
α conotoxin	0.7028	0.0002	0.02%

RMT= relative migration time σ = standard deviation
n = number of replicates

(B) Determination of RMT (separate disulfide isomers)

Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
χ conotoxin	1.0282	0.0031	0.30%
α conotoxin	1.0747	0.0018	0.17%

RMT = relative migration time σ = standard deviation
n = number of replicates



(E)

Peptide	Experimental RMT	Theoretical charge state at pH 2.38	Theoretical R_h (Å)	Theoretical μ_e $m^2.(V.s)^{-1}$	Experimental RMT	Theoretical charge state at pH 6.98	Theoretical R_h (Å)	Theoretical μ_e $m^2.(V.s)^{-1}$
χ conotoxin	0.7040	+3.8	8.7	3.74×10^{-8}	1.0282	+1.3	8.6	1.32×10^{-8}
α conotoxin	0.7028	+3.8	8.3	3.92×10^{-8}	1.0747	+1.1	8.3	1.19×10^{-8}

Formic acid : pH = 2.38 | NH_4CH_3COO : pH = 6.98

Figure 4: *CZE-MS determination of the relative migration time of each conotoxin on 6 separate replicates in* ***(A)*** *a 100 mM formic acid (pH 2.38) co-injected with Glufib and* ***(B)*** *in a 50 mM ammonium acetate BGE (pH 6.98) co-injected with MRFA.* ***(C)*** *and* ***(D)*** *represent typical electropherograms obtained for a mix of the two conotoxins in a 100 mM formic acid BGE (pH 2.38) co-injected with Glufib*

***(C)** and in a 50 mM ammonium acetate BGE co-injected with MRFA **(D)**. The calculated RMTs allow isomer identification. The electropherogram of MRFA in **(D)** was extracted separately from the electropherogram of the conotoxins for better visualization. **(F)** Table of the conotoxins with their the experimentally measured CZE-MS RMT obtained in both pH conditions, the theoretically predicted average charge state and theoretically predicted average hydrodynamic radius (R_h) from the corresponding optimized structure, and the theoretically predicted electrophoretic mobility (μ_e) in both pH conditions*

Experimentally, the switch from acidic (100mM formic acid (FA), pH 2.38) to neutral conditions (50mM ammonium acetate, pH 6.98) allows the disulfide isomers to acquire a slight but sufficient charge state difference in solution to obtain baseline separation. Similar results were obtained for Apamin and ApaRib, with almost no separation under acidic conditions (Supporting Information – Figure S4) but an efficient separation was obtained under neutral conditions (Figure 3F). In this example, the effect of the pH on the separation is decisive because of the relatively small size of this peptide. In this peptide, the presence of two disulfide bonds highly constrains the structures of the corresponding disulfide isomers, leading to close hydrodynamic radius values. Consequently, the average charge state in solution is the critical parameter for an efficient separation by CZE-MS. When working in acidic conditions (100mM FA, pH 2.38), the calculated charges of both isomers are equal, leading to a lack of separation.

Evaluation of the parameters affecting the separation of the ionized peptides in the gas phase and in solution with computational chemistry support

The experimental results using IM-MS showed that a partial gas phase separation of the disulfide isomers was obtained in almost all studied cases. An efficient separation could be obtained if the CCS difference of the analyzed charge states was greater than the IMS resolution (as CCS/ΔCCS). In this context, the optimization of the ion mobility parameters (mainly wave velocity and wave height for a traveling wave instrument) is of paramount importance to obtain the most efficient gas-phase separation as they directly influence the experimental ion mobility resolution. This effect is demonstrated by the comparison of the experimental conditions of the apamins and the conotoxins. Despite the larger theoretical CCS difference for Apamin and

ApaRib in comparison to the conotoxins, ApaRib and Apamin were experimentally unresolved while the conotoxins were partially resolved. The measured IMS diffusion using tetraalkylammonium salts[49] suggested that the IMS resolving power (Peak APEX value divided by peak FWHM) was greater for the conotoxins than the apamins (Supporting Information - Table S2). In this study, the critical gas phase separation parameter, the CCS value, is directly linked to the balance between the way the disulfide bonds constrain the structure of the various disulfide isomers and the Coulomb repulsion between charged residues. Most of our disulfide peptide isomers were efficiently separated by IM-MS and only the Apamin and ApaRib isomers were unresolved. From a theoretical point of view, the predicted CCS values (see Figure 1, 2, and 3) supported the experimental migration order and the previously published results[50]. These values also pointed out that the structuring effect of the disulfide bonds directly affects the CCS values in the gas phase. Indeed, the peptide having the most constrained structure (i.e. the most compact disulfide isomer) presents the smallest CCS value and vice versa. Their CCS deviations are also less affected when the charge state of the peptide increased. When the disulfide patterns have a comparable structuring effect (leading to very close CCS values), an IM-MS separation of the peptide isomers is generally inefficient with the current resolving power ($CCS/\Delta CCS \approx 40\text{-}50$) of IM-MS. As expected, the use of Collision Induced Dissociation (see Supporting Info – Figure S3) on unresolved species did not allow an unambiguous identification since MS/MS spectra were nearly identical. The assignment of the disulfide connectivity with high confidence based on the combination of arrival time distributions and fragmentation patterns is then not always achievable.

In complement to IM-MS experiments, experimental results using CZE-MS showed that an almost baseline separation in solution was achieved in all studied cases, eventually after a pH optimization of the BGE (see next section for details). In the case of CZE-MS, the critical separation parameters are the average charge state in solution (q) and the average hydrodynamic

radius (R_h) of the species, which are both influenced by the folding and consequently by the disulfide pattern. The most obvious effect of different disulfide patterns is the impact on the hydrodynamic radius. As the disulfide bonds are differently arranged in the peptide's structure, the peptide is constrained into different conformations (one conformation per disulfide isomer) leading to different hydrodynamic radii. The second effect of the disulfide patterns is the influence on the average charge in solution. As demonstrated by several groups[44,51], the pK_a values of the ionizable side chains of the residues contained in peptides and proteins are directly affected by their chemical environment. In our study, the differences of q and R_h of all the disulfide isomers were sufficiently different to allow an efficient separation in the optimized pH of the background electrolyte by CZE-MS. In addition, the theoretical calculations unraveled the relative influence of these parameters in each case. Similar to the IM-MS results, these theoretical results matched with the experimental results in all case, except for Apamin and ApaRib. This slight discrepancy between the theoretical and experimental results may be caused by several approximations and initial guesses of the theoretical method in solution such as the presence of counter-ions, the conversion factor from R_g to R_h, or the dynamics of the system. Indeed, the conversion value used to obtain the hydrodynamic radii from the gyration radii is only valid for globular systems[44]. This value should be refined for each disulfide isomer to better match their respective genuine shapes.

In general, the trend in the hydrodynamic radius is followed by the estimated electrophoretic mobility, especially for the larger systems (model peptides and apamins). This suggests that the influence of the hydrodynamic radius is the critical separation parameter for these systems by CZE-MS, with a limited impact of the average charge state. However, the influence of the charge state in solution becomes decisive for smaller systems such as the conotoxins. In this case, the hydrodynamic radius difference is too low to provide an efficient separation (because of the closely related constrained structures) and the average charge difference in solution

overcomes the influence of the hydrodynamic radius difference. The accuracy of the computed charge state in solution for such small systems is therefore of crucial importance for an efficient prediction of the peptide isomers separation by CZE-MS. This also highlights the balance in the importance of both parameters in the separation capabilities of CZE-MS.

Comparison of the separations in the gas phase and in solution

Because of the presence or absence of solvent, the structures of the different disulfide isomers could be influenced by solvation effects during the migration. In solution, the peptide is surrounded by its hydration sphere and counter-ions, while in the gas phase, the peptide's charges require intramolecular solvation by the residues of the peptide.

A comparison of the migrations in the gas phase and solution can only be relevant if the compared charge states are identical, where the only difference relies on the structural parameter (CCS for the gas phase and R_h for the solution). For this purpose, the arrival time distributions of the different charge states generated by the ESI process were measured using IM-MS and compared to the analyzed average charge states obtained during the CZE-MS experiments.

In the case of the conotoxins, the plot of the arrival time distributions as a function of the charge state (Supporting Information - Figure S5) reveals that the initially well-separated species at low charge state (1+ or 2+) asymptotically tend to the same arrival time ditribution at high charge state (4+). As the peak width was significantly larger for the 1+ charge state than for the 2+ (due to ion cloud diffusion[49]), the 2+ charge state was selected as the best separation condition in IM-MS. The CZE-MS results showed that the conotoxins could not be separated when working under acidic conditions (pH 2.38). At this pH, the theoretical calculations predicted an average charge state of +3.84 for both disulfide isomers. Using the arrival time distributions of the different charge states of the conotoxins, a virtual 3.84+ charge state could

be estimated in IM-MS for each conotoxin, mostly based on the 4+ and 3+ charge states (Supporting Information - Figure S5). The estimated IM-MS arrival time distributions for the virtual +3.84 charge states were indeed similar, leading to non-separated species in IM-MS, in perfect agreement with the results obtained by CZE-MS. In contrast, the calculated average charge states in solution at pH 6.98 revealed that the α-conotoxin presented an average charge state of +1.12 while the χ-conotoxin had an average charge state of +1.29. Again, the estimation of the virtual arrival time distributions for these species showed that the χ-conotoxin had a lower estimated arrival time ditribution than α-conotoxin, in agreement with CZE-MS results. In this system, the migration order does not seem to be altered by the transfer of the species to the gas phase, probably because of the small size of the peptide combined to the presence of two disulfide bonds that structurally constrain the folding of the peptides.

In the case of the apamins, the plot of the arrival time distributions as a function of the charge state (Supporting Information - Figure S6) reveals that for almost all charge states, ApaRib and Apamin could not be separated by IM-MS. Interestingly, although the arrival time distributions significantly decreased from 2+ to 5+ charge states, the normalized arrival time distributions difference (i.e. the difference of the arrival time distributions of Apamin and ApaBea divided by the arrival time distribution of Apamin) between the co-migrating ApaRib/Apamin and ApaBea increased from 2+ to 5+. This effect could be explained by the increasing Coulombic repulsion within the structures of these disulfide isomers from 2+ to 5+ charge states. In ApaRib and Apamin, the structures cannot unfold due to the interconnected disulfide bonds that highly constrain the structures of these disulfide isomers (similarly to the observation on the conotoxins) and prevent (or limit) a possible unfolding due to Coulombic repulsions. In contrast, the disulfide bonds of ApaBea are not interconnected, forming two separate loops at each end of the peptide chain, allowing the structure to partially unfold from the 2+ charge state to the 5+ charge state due to Coulombic repulsions. The increasing relative arrival time

ditribution difference could thus be attributed to the partial unfolding of ApaBea compared to the constrained ApaRib and Apamin.

The CZE-MS in acidic conditions (pH 2.38) highlighted similar results to IM-MS at all charge states, with a near co-migration of ApaRib and Apamin and a larger separation efficiency for ApaBea (Supporting Information - Figure S4C). However, unlike the IM-MS results, CZE-MS experiments revealed baseline separation of Apamin, ApaRib, and ApaBea at pH 6.98. The computed theoretical values of the charge state in solution (close to 3+) and R_h at this pH (Figure G) revealed that the combination of slight charge differences to hydrodynamic radius differences was in this case sufficient to provide baseline separation of the peptide isomers. It should be noted that the theoretical calculations either underestimates the solution mobility of Apamin or overestimates the solution mobility of ApaRib as the results of the calculations indicate that these species should be co-migrating instead of being fully separated. The limitation of this theoretical approach is probably linked to the several hypotheses and approximations mentioned above. Interestingly, the comparison of 3+ charge states in IM-MS and CZE-MS for these disulfide isomers also highlighted identical migration order between the solution and the gas phase, pointing out similar results for this system than for the conotoxins about the preservation of migration orders from the solution to the gas phase.

Finally, the comparison of the disulfide isomers of the model peptide was not attempted because of the presence of a relatively high proportion of isopropanol (20%) in the BGE, preventing an accurate description of the solution properties (especially in terms of charge) required for the simulations.

The comparison of the migration behavior between the gas phase and the solution suggests the structural features of highly structured systems such as the small conotoxin peptide isomers were not drastically affected. Indeed, when comparing identical charge states, there was a

perfect agreement between the migration orders in solution and in the gas phase. In the future, the present methodology will be implemented on non-constrained (i.e. linear) structures such as peptides without disulfide bonds to investigate differential migration behaviors between the liquid phase (i.e. CZE-MS) and the gas phase (i.e. IM-MS).

Finally, from an analytical point of view, these results show that IM-MS and CZE-MS provide similar migration behaviors at similar charge (average) states. Nonetheless, the charge state of the ions is imposed by the electrospray ionization or by the pH of the background electrolyte. The difference of behavior in terms of migration in the gas phase and the solution is charge-dependent and could be tuned accordingly to differentiate the migration behavior between the gas phase and the liquid phase. In the highly constrained disulfide isomers used in this study, the structural parameter (CCS in IMS and R_h in CZE) is sometimes not sufficiently different to provide efficient separation. In such cases, IM-MS separations can be optimized by investigating different charge states produced by the electrospray ionization and adapting the IMS-related separation settings (essentially the wave velocity, wave height, and gas pressure). In CZE-MS, an efficient way to optimize the separation is by adjusting the pH.

CONCLUSION

The aim of this study was to evaluate the ability and possible complementarity of CZE-MS and IM-MS to separate peptides presenting the same amino acid sequence and bearing two intramolecular disulfide bonds but different disulfide bond patterns. For this purpose, three peptides of various lengths were selected and analyzed by both analytical techniques. The results revealed that CZE-MS was able to discriminate the disulfide isomers of each peptide, either on a hydrodynamic radius difference and/or on an in-solution average charge state difference. In all cases, near baseline resolution was achieved using traditional MS-friendly background electrolytes. When using gas-phase IM-MS, results showed that the separation is

sometimes compromised when the conformations of the disulfide isomers are very similar and/or the resolving power of the ion mobility cell is insufficient. Also, the comparison of the migration behavior between the gas phase and the solution for the peptides at identical charge states indicates that the structures after the desolvation are not drastically affected. All experimental results were supported by computational chemistry. These calculations were based on structure optimizations in the gas phase and in solution to determine various physicochemical properties including collision cross sections (gas phase), average charge state (solution) and average hydrodynamic radius (solution). These parameters led to the estimation and in-depth understanding of the migration behaviors of the studied species. The proposed approach paves the way for the theoretical prediction of structures in the gas phase and the solution to access the theoretical ATDs and the relative migration times obtained from the solution.

As demonstrated in this work, IM-MS is a fast analytical technique that enables the attribution of the disulfide connectivity when the IM-MS provides a resolved separation of the different disulfide isomers. In the present work, we achieved at least partial IM-MS separations of most of the disulfide isomers with a commercial traveling wave Synapt G2 HDMS having a resolving power (CCS/ΔCCS) around 40-50. This work showed the contribution of IM-MS in the challenging attribution of disulfide patterns in peptides in a couple of minutes. On the other hand and despite a longer duration of analysis, the use of CZE-MS provides several advantages including high robustness (all standard deviations of the RMTs are below 1%), the ability to work in “native” or “native-like” conditions and the minimal sample consumption per analysis (in the nanoliter range). The versatility of CZE relies on the use of both the charge (q) and the hydrodynamic radius (R_h) as separation parameters and is particularly useful for disulfide pattern attribution. As presented in this work, both parameters can be exploited to optimize the separation conditions (e.g., pH of background electrolyte, addition of organic modifiers) to

obtain a baseline separation of all the peptides isomers investigated, showing the great potential of CZE-MS for disulfide connectivity attribution.

The comparison between the migration orders in IM-MS and CZE-MS at equivalent average charge state revealed similar migration behavior in both the liquid and the gas phases, suggesting that structural features from the solution (R_h) are not drastically affected after the desolvation and transfer into the gas phase by electrospray ionization. This result was expected due to the highly constrained structures of the studied peptides.

In the future, the use of peptides bearing a larger number of intramolecular disulfide bonds or even proteins can be considered. The use of non-constrained peptides in comparison to the studied species in this work could be performed to check the assumption concerning the preservation of structural features from the solution to the gas phase.

ACKNOWLEDGMENTS

The authors thank the FRS-FNRS for the financial support (FRIA and instrumentation), the Fonds Européen de Développement Régional (FEDER), the Walloon region and the European Commission (F.P. 7 VENOMICS project) for financial support. We also thank the Walloon region and the FIRST Entreprise 6592 project for the financial support of this work. Analis is also acknowledged for valuable CE expertise and discussions about this work.

SUPPORTING INFORMATION AVAILABLE

This information is available free of charge via the Internet at http://pubs.acs.org/

CONFLICT OF INTEREST DISCLOSURE

The authors declare no competing financial interest

REFERENCES

(1) Matsumura, M.; Signor, G.; Matthews, B. W. Substantial Increase of Protein Stability by Multiple Disulphide Bonds. *Nature* **1989**, *342* (6247), 291–293.

(2) Góngora-Benítez, M.; Tulla-Puche, J.; Albericio, F. Multifaceted Roles of Disulfide Bonds. Peptides as Therapeutics. *Chem. Rev.* **2013**, *114*, 901–926. https://doi.org/10.1021/cr400031z.

(3) Craik, D. J.; Simonsen, S.; Daly, N. L. The Cyclotides: Novel Macrocyclic Peptides as Scaffolds in Drug Design. *Curr. Opin. Drug Discov. Devel.* **2002**, *5* (2), 251–260.

(4) Craik, D. J.; Cemazar, M.; Daly, N. L. The Cyclotides and Related Macrocyclic Peptides as Scaffolds in Drug Design. *Curr. Opin. Drug Discov. Devel.* **2006**, *9* (2), 251–260.

(5) Henriques, S. T.; Craik, D. J. Cyclotides as Templates in Drug Design. *Drug Discov. Today* **2010**, *15* (1–2), 57–64. https://doi.org/10.1016/j.drudis.2009.10.007.

(6) Gunasekera, S.; Daly, N. L.; Anderson, M. A.; Craik, D. J. Chemical Synthesis and Biosynthesis of the Cyclotide Family of Circular Proteins. *IUBMB Life* **2006**, *58* (9), 515–524. https://doi.org/10.1080/15216540600889532.

(7) Wang, C. K. L.; Kaas, Q.; Chiche, L.; Craik, D. J. CyBase: A Database of Cyclic Protein Sequences and Structures, with Applications in Protein Discovery and Engineering. *Nucleic Acids Res.* **2008**, *36* (Database issue), D206-10. https://doi.org/10.1093/nar/gkm953.

(8) Lehrer, R. I.; Ganz, T. Antimicrobial Peptides in Mammalian and Insect Host Defence. *Curr. Opin. Immunol.* **1999**, *11* (1), 23–27. https://doi.org/http://dx.doi.org/10.1016/S0952-7915(99)80005-3.

(9) Ganz, T. Defensins: Antimicrobial Peptides of Innate Immunity. *Nat Rev Immunol* **2003**, *3* (9), 710–720.

(10) Terlau, H.; Olivera, B. M. Conus Venoms: A Rich Source of Novel Ion Channel-Targeted Peptides. *Physiol. Rev.* **2004**, *84* (1), 41–68. https://doi.org/10.1152/physrev.00020.2003.

(11) Lewis, R. J.; Garcia, M. L. Therapeutic Potential of Venom Peptides. *Nat Rev Drug Discov* **2003**, *2* (10), 790–802.

(12) Gehrmann, J.; Alewood, P. F.; Craik, D. J. Structure Determination of the Three Disulfide Bond Isomers of α-Conotoxin GI: A Model for the Role of Disulfide Bonds in Structural Stability. *J. Mol. Biol.* **1998**, *278* (2), 401–415. https://doi.org/http://dx.doi.org/10.1006/jmbi.1998.1701.

(13) Wu, Y.; Wu, X.; Yu, J.; Zhu, X.; Zhangsun, D.; Luo, S. Influence of Disulfide Connectivity on Structure and Bioactivity of α-Conotoxin TxIA. *Molecules* **2014**, *19* (1), 966–979. https://doi.org/10.3390/molecules19010966.

(14) Moroder, L.; Besse, D.; Musiol, H.-J.; Rudolph-Böhner, S.; Siedler, F. Oxidative Folding of Cystine-Rich Peptides vs Regioselective Cysteine Pairing Strategies. *Biopolymers* **1996**, *40* (2).

(15) Trivedi, M. V.; Laurence, J. S.; Siahaan, T. J. The Role of Thiols and Disulfides on Protein Stability. *Curr. Protein Pept. Sci.* **2009**, *10* (6). https://doi.org/10.2174/138920309789630534.

(16) Walewska, A.; Skalicky, J. J.; Davis, D. R.; Zhang, M.-M.; Lopez-Vera, E.; Watkins, M.; Han, T. S.; Yoshikami, D.; Olivera, B. M.; Bulaj, G. NMR-Based Mapping of Disulfide Bridges in Cysteine-Rich Peptides: Application to the μ-Conotoxin SxIIIA. *J. Am. Chem. Soc.* **2008**, *130* (43), 14280–14286. https://doi.org/10.1021/ja804303p.

(17) Mobli, M.; King, G. F. NMR Methods for Determining Disulfide-Bond Connectivities. *Toxicon* **2010**, *56* (6), 849–854. https://doi.org/10.1016/j.toxicon.2010.06.018.

(18) Clarke, J.; Henrick, K.; Fersht, A. R. Disulfide Mutants of Barnase I: Changes in Stability and Structure Assessed by Biophysical Methods and X-Ray Crystallography. *J. Mol. Biol.* **1995**, *253* (3), 493–504. https://doi.org/10.1006/JMBI.1995.0568.

(19) Calvete, J. J.; Schrader, M.; Raida, M.; McLane, M. A.; Romero, A.; Niewiarowski, S. The Disulphide Bond Pattern of Bitistatin, a Disintegrin Isolated from the Venom of the Viper Bitis Arietans. *FEBS Lett.* **1997**, *416* (2), 197–202. https://doi.org/http://dx.doi.org/10.1016/S0014-5793(97)01203-9.

(20) Bauer, M.; Sun, Y.; Degenhardt, C.; Kozikowski, B. Assignment of All Four Disulfide Bridges in Echistatin. *J. Protein Chem.* **1993**, *12* (6), 759–764. https://doi.org/10.1007/BF01024934.

(21) Liu, H.; Boudreau, M. A.; Zheng, J.; Whittal, R. M.; Austin, P.; Roskelley, C. D.; Roberge, M.; Andersen, R. J.; Vederas, J. C. Chemical Synthesis and Biological Activity of the Neopetrosiamides and Their Analogues: Revision of Disulfide Bond Connectivity. *J. Am. Chem. Soc.* **2010**, *132* (5). https://doi.org/10.1021/ja9102925.

(22) Durand, K. L.; Ma, X.; Plummer, C. E.; Xia, Y. Tandem Mass Spectrometry (MSn) of Peptide Disulfide Regio-Isomers via Collision-Induced Dissociation: Utility and Limits in Disulfide Bond Characterization. *Int. J. Mass Spectrom.* **2013**, *343–344*, 50–57. https://doi.org/http://dx.doi.org/10.1016/j.ijms.2013.04.005.

(23) Tan, L.; Durand, K. L.; Ma, X.; Xia, Y. Radical Cascades in Electron Transfer Dissociation (ETD) - Implications for Characterizing Peptide Disulfide Regio-Isomers. *Analyst* **2013**. https://doi.org/10.1039/C3AN01333B.

(24) Massonnet, P.; Upert, G.; Smargiasso, N.; Gilles, N.; Quinton, L.; De Pauw, E. Combined Use of Ion Mobility and Collision-Induced Dissociation to Investigate the Opening of Disulfide Bridges by Electron-Transfer Dissociation in Peptides Bearing Two Disulfide Bonds. *Anal. Chem.* **2015**, *87* (10), 5240–5246. https://doi.org/10.1021/acs.analchem.5b00245.

(25) Papayannopoulos, I. A. The Interpretation of Collision-Induced Dissociation Tandem Mass Spectra of Peptides. *Mass Spectrom. Rev.* **1995**, *14* (1), 49–73. https://doi.org/10.1002/mas.1280140104.

(26) Cole, S.; Ma, X.; Zhang, X.; Xia, Y. Electron Transfer Dissociation (ETD) of Peptides Containing Intrachain Disulfide Bonds. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (2), 310–320. https://doi.org/10.1007/s13361-011-0300-z.

(27) Massonnet, P.; Haler, J. R. N.; Upert, G.; Smargiasso, N.; Mourier, G.; Gilles, N.; Quinton, L.; De Pauw, E. Disulfide Connectivity Analysis of Peptides Bearing Two Intramolecular Disulfide Bonds Using MALDI In-Source Decay. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (10), 1995–2002. https://doi.org/10.1007/s13361-018-2022-y.

(28) Martinez, T.; Guo, A.; Allen, M. J.; Han, M.; Pace, D.; Jones, J.; Gillespie, R.; Ketchem, R. R.; Zhang, Y.; Balland, A. Disulfide Connectivity of Human Immunoglobulin G2 Structural Isoforms. **2008**, *47* (28), 7496–7508. https://doi.org/10.1021/bi800576c.

(29) Lacher, N. A.; Wang, Q.; Roberts, R. K.; Holovics, H. J.; Aykent, S.; Schlittler, M. R.; Thompson, M. R.; Demarest, C. W. Development of a Capillary Gel Electrophoresis Method for Monitoring Disulfide Isomer Heterogeneity in IgG2 Antibodies. **2010**, *31* (3), 448–458. https://doi.org/10.1002/elps.200900371.

(30) He, Y.; Lacher, N. A.; Hou, W.; Wang, Q.; Isele, C.; Starkey, J.; Ruesch, M. Analysis of Identity, Charge Variants, and Disulfide Isomers of Monoclonal Antibodies with Capillary Zone Electrophoresis in an Uncoated Capillary Column. **2010**, *82* (8), 3222–3230. https://doi.org/10.1021/ac9028856.

(31) Massonnet, P.; Haler, J. R. N.; Upert, G.; Degueldre, M.; Morsa, D.; Smargiasso, N.; Mourier, G.; Gilles, N.; Quinton, L.; De Pauw, E. Ion Mobility-Mass Spectrometry as a Tool for the Structural Characterization of Peptides Bearing Intramolecular Disulfide Bond(S). *J. Am. Soc. Mass Spectrom.* **2016**, *27* (10), 1637–1646. https://doi.org/10.1007/s13361-016-1443-8.

(32) Echterbille, J.; Quinton, L.; Gilles, N.; De Pauw, E. Ion Mobility Mass Spectrometry as a Potential Tool to Assign Disulfide Bonds Arrangements in Peptides with Multiple Disulfide Bridges. *Anal. Chem.* **2013**, *85* (9), 4405–4413. https://doi.org/10.1021/ac303686w.

(33) Upert, G.; Mourier, G.; Pastor, A.; Verdenaud, M.; Alili, D.; Servent, D.; Gilles, N. High-Throughput Production of Two Disulphide-Bridge Toxins. *Chem. Commun.* **2014**, *50* (61), 8408–8411.

(34) Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J. Comput. Chem.* **1996**, *17* (5-6), 490–519.

(35) 09, G.; A. 02, R.; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V. Gaussian, Inc. **2016**.

(36) Weiner, P. K.; Kollman, P. A. AMBER: Assisted Model Building with Energy Refinement. A General Program for Modeling Molecules and Their Interactions. *J. Comput. Chem.* **1981**, *2* (3), 287–303. https://doi.org/10.1002/jcc.540020311.

(37) Larriba, C.; Hogan Jr, C. J. Free Molecular Collision Cross Section Calculation Methods for Nanoparticles and Complex Ions with Energy Accommodation. *J. Comput. Phys.* **2013**, *251*, 344–363.

(38) Larriba, C.; Hogan Jr, C. J. Ion Mobilities in Diatomic Gases: Measurement versus Prediction with Non-Specular Scattering Models. *J. Phys. Chem. A* **2013**, *117* (19), 3887–3901.

(39) Barone, V.; Cossi, M.; Tomasi, J. Geometry Optimization of Molecular Structures in Solution by the Polarizable Continuum Model. *J. Comput. Chem.* **1998**, *19* (4), 404–417.

(40) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating p K as and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33* (suppl_2), W368–W371.

(41) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of PKs in Macromolecules. *Proteins Struct. Funct. Bioinforma.* **2006**, *63* (4), 928–938.

(42) Anandakrishnan, R.; Aguilar, B.; Onufriev, A. V. H++ 3.0: Automating p K Prediction and the Preparation of Biomolecular Structures for Atomistic Molecular Modeling and Simulations. *Nucleic Acids Res.* **2012**, *40* (W1), W537–W541.

(43) DeLano, W. L. The PyMOL Molecular Graphics System (DeLano Scientific LLC, San Carlos, CA). *PyMOL Mol. Graph. Syst. World Wide Web URL http//www. pymol. org* **2002**.

(44) Tande, B. M.; Wagner, N. J.; Mackay, M. E.; Hawker, C. J.; Jeong, M. Viscosimetric, Hydrodynamic, and Conformational Properties of Dendrimers and Dendrons. *Macromolecules* **2001**, *34* (24), 8580–8585. https://doi.org/10.1021/ma011265g.

(45) Bleiholder, C.; Suhai, S.; Paizs, B. Revising the Proton Affinity Scale of the Naturally Occurring α-Amino Acids. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (9), 1275–1281.

(46) Habermann, E. Apamin. *Pharmacol. Ther.* **1984**, *25* (2), 255–270. https://doi.org/10.1016/0163-7258(84)90046-9.

(47) Wang, Y.; Shao, X.; Li, M.; Wang, S.; Chi, C.; Wang, C. Mr1e, a Conotoxin from Conus Marmoreus with a Novel Disulfide Pattern. **2008**, *40* (5), 391–396. https://doi.org/10.1111/j.1745-7270.2008.00414.x.

(48) King, G. F.; Gentz, M. C.; Escoubas, P.; Nicholson, G. M. A Rational Nomenclature for Naming Peptide Toxins from Spiders and Other Venomous Animals. *Toxicon* **2008**, *52* (2), 264–276. https://doi.org/10.1016/j.toxicon.2008.05.020.

(49) Kune, C.; Far, J.; De Pauw, E. Accurate Drift Time Determination by Traveling Wave Ion Mobility Spectrometry: The Concept of the Diffusion Calibration. *Anal. Chem.* **2016**, *88* (23), 11639–11646. https://doi.org/10.1021/acs.analchem.6b03215.

(50) Pace, C. N.; Grimsley, G. R.; Scholtz, J. M. Protein Ionizable Groups: PK Values and Their Contribution to Protein Stability and Solubility. *J. Biol. Chem.* **2009**, *284* (20), 13285–13289. https://doi.org/10.1074/jbc.R800080200.

(51) Grimsley, G. R.; Scholtz, J. M.; Pace, C. N. A Summary of the Measured PK Values of the Ionizable Groups in Folded Proteins. *Protein Sci.* **2009**. https://doi.org/10.1002/pro.19.

5.3. Results - Supporting Information

Combination of Capillary Zone Electrophoresis-Mass Spectrometry, Ion Mobility-Mass Spectrometry, and Theoretical Calculations for cysteine connectivity identification in peptides bearing two intramolecular disulfide bonds

Cédric Delvaux(1)† and Philippe Massonnet(1)(2)†, Christopher Kune(1), Jean R.N. Haler(1)(3), Gregory Upert(4), Gilles Mourier (4), Nicolas Gilles(4), Loïc Quinton(1), Edwin De Pauw(1) and Johann Far(1)

Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium

Maastricht Multimodal Molecular Imaging (M4I) Institute, Division of Imaging Mass Spectrometry, Maastricht, Limburg, Netherlands

Department of Chemistry and Biochemistry, Florida International University, Miami, Florida 33199, USA

Commissariat à l'Energie Atomique, DRF/Institut Joliot/SIMOPRO, Université Paris Sud, 91191, Gif-sur-Yvette, France

† Cédric Delvaux and Philippe Massonnet contributed equally as first authors

AUTHOR EMAIL ADDRESS: c.delvaux@uliege.be

SUPPORTING INFORMATION

Materials and Methods

Capillary Zone Electrophoresis-Mass Spectrometry (CZE-MS)

The capillary was conditioned using successively deionized water (25 psi, 5 min), MeOH/H_2O 50/50 (25 psi, 5 min), deionized water (25 psi, 5 min), 0.1 M NaOH (25psi, 5min), deionized water (25 psi, 5 min), 0.2 M HCl (25 psi, 10 min) and deionized water (25 psi, 10 min). Finally, the capillary was flushed with the background electrolyte (BGE) as 100 mM FA during 20 min at 50 psi.

Before the first CE run and between the different runs, the capillary was rinsed using BGE (50 psi during 5 min) before loading the sample and a reference compound hydrodynamically into the capillary at 5 psi for 30 seconds. The CE separation was performed applying 30 kV (normal polarity) for 40 minutes. The capillary was maintained at 25°C while the autosampler was maintained at 15°C during the entire sets of runs.

The CE system was connected to the mass spectrometer through the EDA external adapter (Beckman, by passing the UV detector) to a sheath liquid sprayer CEMS interface (PN 10-301347 ANALIS SL CE-MS Sprayer, Analis, Suarlée Belgium) operating in the microliter range. The composition of the sheath liquid was adapted according to the nature of the BGE as shown in table S1. The sheath liquid flow rate was fixed at 0.7µL.min^{-1} to ensure stable Taylor cone for electrospray ionization and minimum chemical noise during MS detection.

The compositions of both BackGround Electrolyte (BGE) and Sheath Liquid (SL) were optimized for each peptide as summarized in Table1.

The samples were all prepared following the same procedure: 100µL of each peptide of interest (or the mix of the 3 disulfide isomers) were prepared at 10µM from the stock solutions available after the synthesis procedure in the appropriate 10-fold diluted BGE.

Table S1: List of the background electrolyte and sheath liquid compositions for the different peptides used in the study.

Peptide	Composition of the BGE	Composition of the SL
ModGlo ModRib ModBea	10% Acetic Acid in 20% IPA	10% Acetic Acid in 50% IPA
Apamin ApaRib ApaBea	100mM Ammonium Acetate pH 6.98	10% Acetic Acid in 50% IPA
α conotoxin χ conotoxin	50mM Ammonium Acetate pH 6.98	1% Acetic Acid, 25µM Ammonium Acetate pH 7 in 50% IPA

Theoretical Calculations: Molecular Mechanics

Part 1: Structure optimization in vacuo (gas phase optimization)

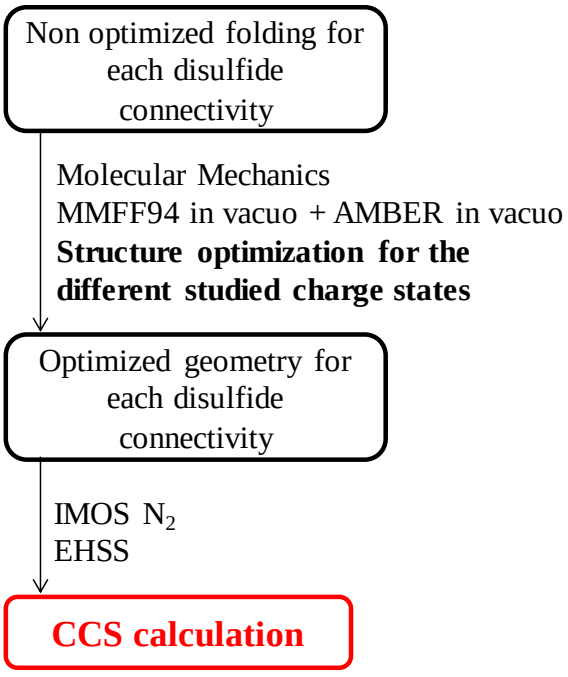
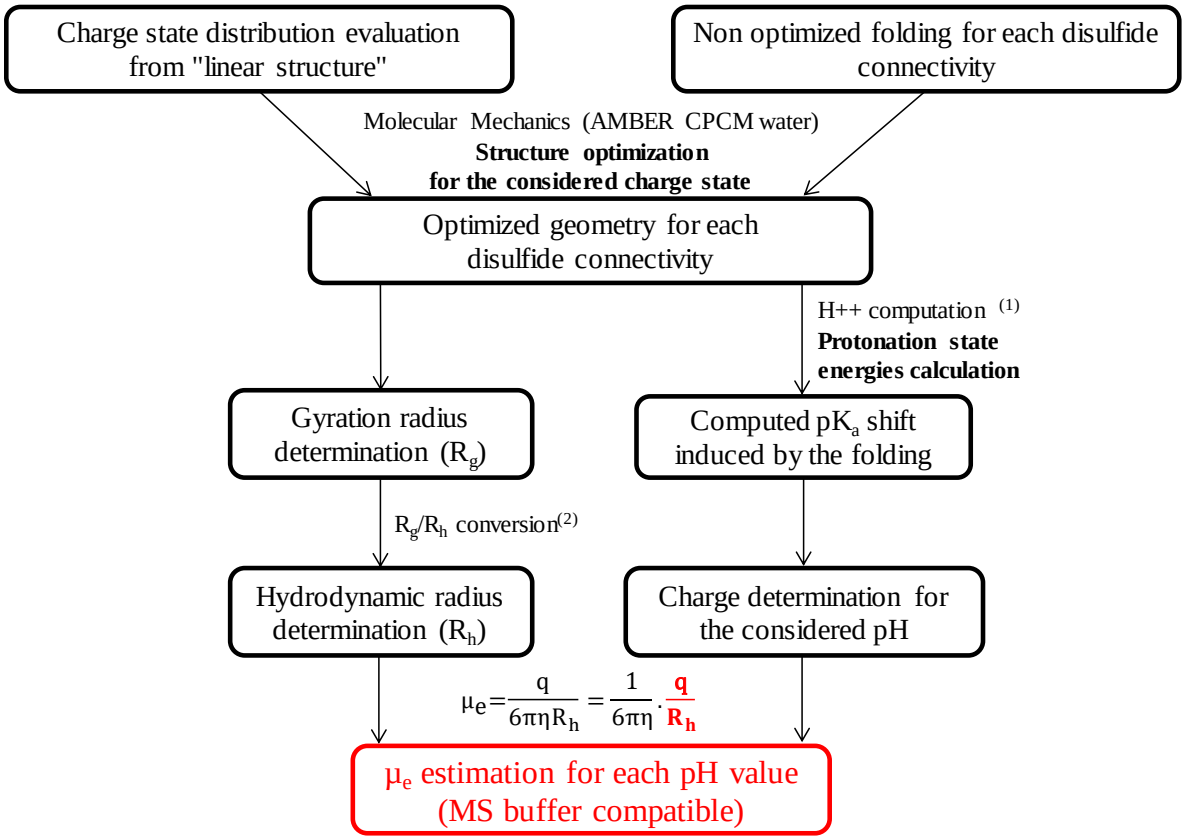


Figure S1: Workflow of the theoretical calculations performed for the gas phase structures and the determination of their respective CCS values

Part 2: Structure optimization in water (solution optimization)



(1) http://biophysics.cs.vt.edu/

(2) Tande and Wagner; Macromolecules, 2001, 34, 8580-8585

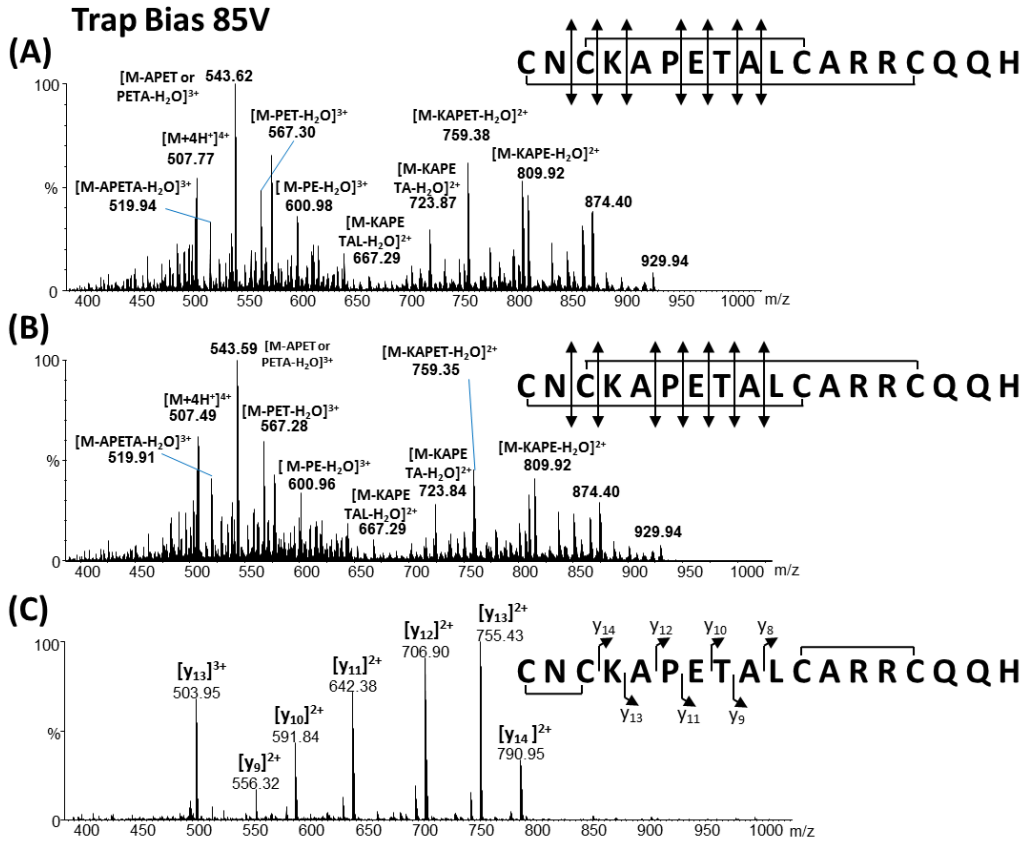
Figure S2: Workflow of the theoretical calculations performed for in-solution structures and the estimation of their respective electrophoretic mobilities

Table S2: Evolution of IMS resolution of the Synapt G2 as a function of the IMS settings used in this study. The resolution is calculated as the drift time divided by the FWHM for various tetraalkylammoniums.

Conotoxins IMS settings				
Ions	**m/z**	**Drift time**	**FWHM**	**Resolution**
Tetrapropylammonium	186.22	5.91	0.32	18.53
Tetrabutylammonium	242.28	7.89	0.41	19.29
Tetrapentylammonium	298.35	10.28	0.55	18.60
Tetrahexylammonium	354.41	12.90	0.73	17.65
Tetraheptylammonium	410.47	15.65	0.94	16.68
Tetraoctylammonium	466.54	18.50	1.14	16.19

Apamins IMS settings				
Ions	**m/z**	**Drift time**	**FWHM**	**Resolution**
Tetrapropylammonium	186.22	6.84	0.46	14.92
Tetrabutylammonium	242.28	9.30	0.66	14.06
Tetrapentylammonium	298.35	12.35	0.93	13.22
Tetrahexylammonium	354.41	15.87	1.37	11.62
Tetraheptylammonium	410.47	19.75	1.90	10.42
Tetraoctylammonium	466.54	24.04	2.49	9.66

Model Peptides IMS settings				
Ions	**m/z**	**Drift time**	**FWHM**	**Resolution**
Tetrapropylammonium	186.22	6.84	0.46	14.92
Tetrabutylammonium	242.28	9.30	0.66	14.06
Tetrapentylammonium	298.35	12.35	0.93	13.22
Tetrahexylammonium	354.41	15.87	1.37	11.62
Tetraheptylammonium	410.47	19.75	1.90	10.42
Tetraoctylammonium	466.54	24.04	2.49	9.66



<u>Figure S3:</u> (A) CID MS/MS spectrum of ApaRib (C_1-C_4/C_2-C_3); (B) CID MS/MS spectrum of Apamin (C_1-C_3/C_2-C_4); (C) CID MS/MS spectrum of ApaBea (C_1-C_2/C_3-C_4). All spectra were recorded with a trap bias of 85V.

(A) **Determination of RMT (separate disulfide isomers)**

Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
ApaRib	0.4676	0.0009	0.18%
Apamin	0.4925	0.0008	0.15%
ApaBea	0.5287	0.0031	0.59%

RMT = relative migration time σ = standard deviation
n = number of replicates

(B) **Determination of RMT (mix of the disulfide isomers)**

Peptide	RMT (n=6)	σ (n=6)	%σ (n=6)
ApaRib	0.4753	0.0015	0.32%
Apamin	0.4857	0.0021	0.43%
ApaBea	0.5212	0.0017	0.32%

RMT = relative migration time σ = standard deviation
n = number of replicates

(C)

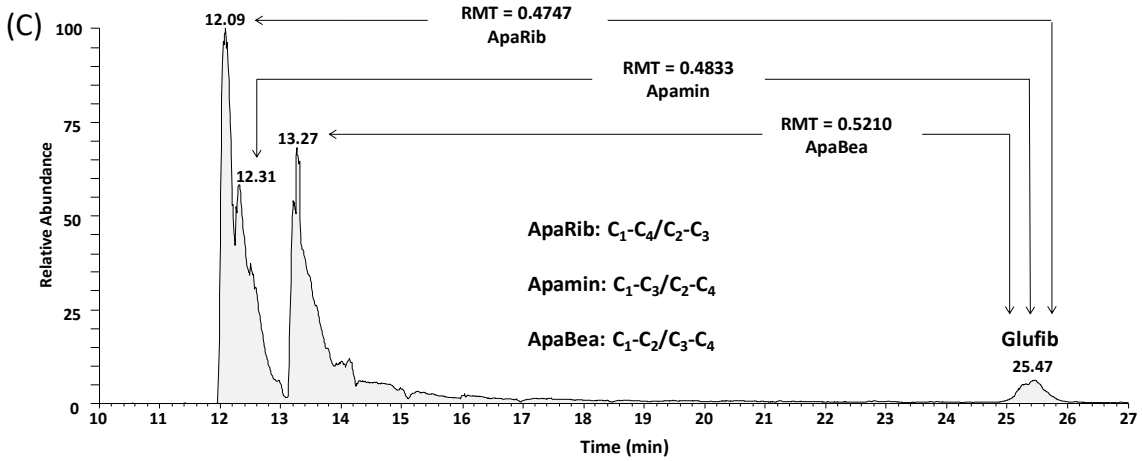


Figure S4: CZE-MS determination of the relative migration time (RMT) of each disulfide isomer of the apamin peptide in 100mM formic acid BGE, ***(A)*** *performed on 6 separate replicates and* ***(B)*** *on 6 replicates of the isomer mix* ***(C)*** *Typical CZE-MS electropherogram obtained for a run in 100mM formic acid BGE of a mix of the 3 disulfide isomers of the apamin peptide co-injected with Glufib as internal standard (the calculated RMTs allow the isomer identification).*

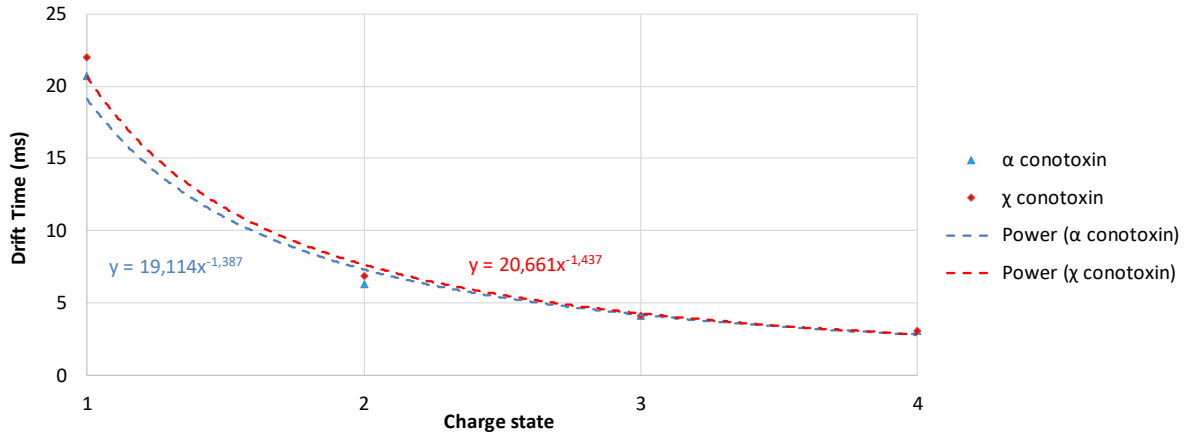
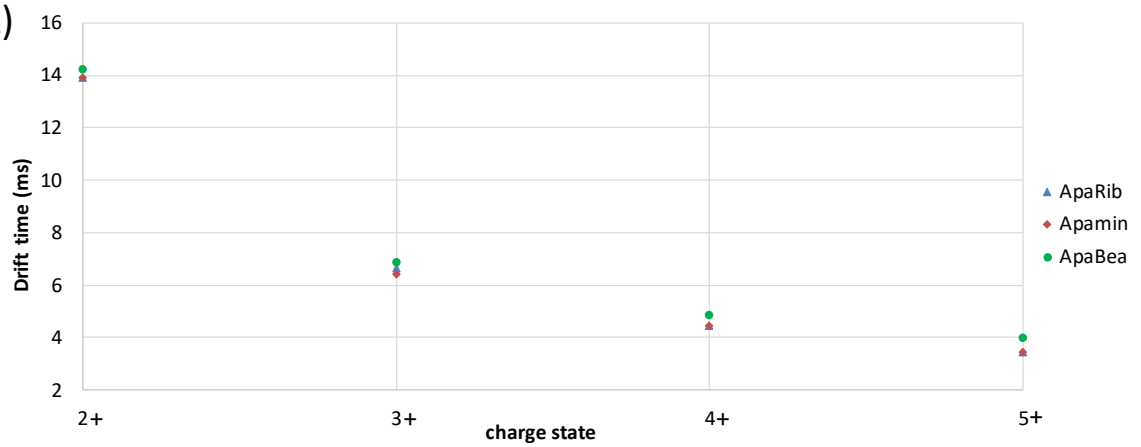


Figure S5: Evolution of the drift time of the α and χ conotoxins as a function of the considered charge state and power fit with the related equation

(A)



(B)

	2+	3+	4+	5+
drift time ApaRib (ms)	13.89	6.61	4.41	3.42
drift time Apamin (ms)	13.89	6.39	4.41	3.42
drift time ApaBea (ms)	14.22	6.84	4.85	3.97
Normalized % of drift time difference between ApaBea and Apamin/ApaRib	2.38%	7.04%	9.98%	16.08%

<u>Figure S6:</u> (A) Evolution of the drift time of ApaRib, Apamin and ApaBea as a function of the considered charge state (+1 charge state could not be detected) and (B) Table with the respective drift times of the different charge states of ApaRib, Apamin and ApaBea and the normalized percentage of drift time difference between ApaBea and ApaRib/Apamin

5.4. Conclusion and perspectives

In this chapter, the ability of mobility-based techniques including Capillary Zone Electrophoresis (CZE) and Ion Mobility (IM) to separate disulfide isomers of peptides with various lengths has been successfully demonstrated.

In ion mobility, the separation of disulfide isomers on a particular charge state relies on a CCS difference at a specific charge state. In the studied peptides, the CCS difference is generally sufficient to lead to at least a partial separation, except in the case of ApaRib and Apamin. This result highlights the great potential of this fast analytical technique for the separation of closely related species, provided that the resolution of the gas phase structures is sufficient. In the future, the use of instrument characterized by a higher IMS resolution such as Trapped Ion Mobility Spectrometry (TIMS) could be considered to increase the separation efficiency of the gas-phase disulfide isomers.

In addition, the use of collision induced dissociation tandem mass spectrometry (CID-MS^2) did not provide characteristic fragments of the unseparated disulfide isomers by IMS (see Supporting Information of this Chapter in the Appendix), probably because of the particularly constrained structures due to the presence of two disulfide bonds. In the future, the use of alternative fragmentation methods such as electron transfer dissociation (ETD) or spectroscopic activation methods could be considered.

On the other hand, the separation in CZE is more versatile as it relies on both the charge (q) and the hydrodynamic radius (R_h) as separation parameters. In this context, a pH optimization of the method is necessary in some cases (such as the apamins and the conotoxins) to achieve an efficient separation based on a charge difference. Although the separation itself is much longer than for IMS (several minutes compared to milliseconds), CZE-MS provided a baseline separation of the different disulfide isomers and allows to work in native or native-like solutions.

Because these two techniques only require minute amounts of sample, they could be of major interest to determine the disulfide pattern of peptides originating from venoms that are extremely difficult to obtain, such as those produced by tiny venomous animals.

As expected for highly constrained peptides, the comparison between the migration orders in IM-MS and CZE-MS at equivalent average charge state revealed that structural features from the solution are mostly retained after the desolvation and transfer into the gas phase by electrospray ionization. In the future, this comparison could be tested on non-constrained peptides to assess the preservation of structural features from the solution to the gas phase.

Finally, the developed calculation method (see Figures S1 and S2 in the Supporting Information), allows the estimation of several physico-chemical parameters based on the optimized structures in the gas phase (Figure S1) and solution (Figure S2).

Concerning the theoretical calculations, the estimation of the different separation parameters by the proposed method suffers from various approximations and initial guesses that should be improved to increase the confidence of the computed properties, especially for solution calculations.

For the gas phase structures, the trajectory method (TM) could be implemented instead of the EHSS method as it generally provides a more accurate description of the CCS. Of course, this

should be attempted keeping in mind that the invested computational time must provide a valuable contribution to the CCS prediction.

For the structures in solution, it should be mentioned that the structures were optimized in water, as water is one of the only surrounding solution directly implemented into the force field. Of course, water does not present the same physicochemical properties than the utilized BGE in this work (such as 100mM formic acid or 50mM ammonium acetate). As a result, the influence of the ions contained in the migration buffer is not properly taken into account.

Finally, the conversion factor between the gyration radius (R_g) provided by the calculation and the hydrodynamic radius (R_h) should also be revised as all optimized structures do not present the same "shape". The value used in this study 0.775 is only valid for globular shapes. However, other conversion factors exist when molecules deviate from globular to non-spherical or elongated structures and the ratio R_g/R_h tend to be larger than 0.775. This conversion factor should therefore be improved to better reflect the genuine shape of each optimized structure.

CHAPTER 6

Determination of the Electroosmotic Flow and Suction Effect in Capillary Electrophoresis coupled to Mass Spectrometry

6.1. Context of the chapter

As explained in the introduction (chapter 2), in capillary electrophoresis, two electro-driven phenomena take place upon the application of the electric field inside the capillary. Next to the well-known electrophoretic mobility of the analytes, a flow of bulk solution referred as the "electroosmotic flow" (EOF) also occurs. This flow is due to the displacement of the diffuse layer of ions in solution from the double layer in direct contact with the capillary wall upon the application of the electric field. Various experimental parameters such as the ionic strength, pH and composition of the BGE or the capillary surface properties influence the amplitude and the direction of the electroosmotic flow (see chapter 2 for related details). As the apparent mobility of an analyte (directly determined from its detected migration time) is the superimposition of its intrinsic electrophoretic mobility and the EOF mobility, a strict control of EOF is crucial for efficient analytical method development.

From the analytical point of view, the EOF can be beneficial or detrimental. Indeed, CE separations can be compromised if the EOF mobility is much larger and in the same direction than the intrinsic electrophoretic mobilities of the analytes of interest, providing too short analysis times to obtain sufficiently separated analytes. On the other hand, as the EOF sweeps all analytes towards the direction of the detector regardless of their charge (generally the cathodic end when working with the normal polarity mode with classical bare fused silica capillaries), it enables single-point detection for cations, neutral molecules and anions in a single run.
Unlike pressure-driven phenomena characterized by a parabolic flow profile, the characteristic flat plug-like profile of EOF preserves the CE resolution as it doesn't cause any additional diffusion (see Figure 6.1).

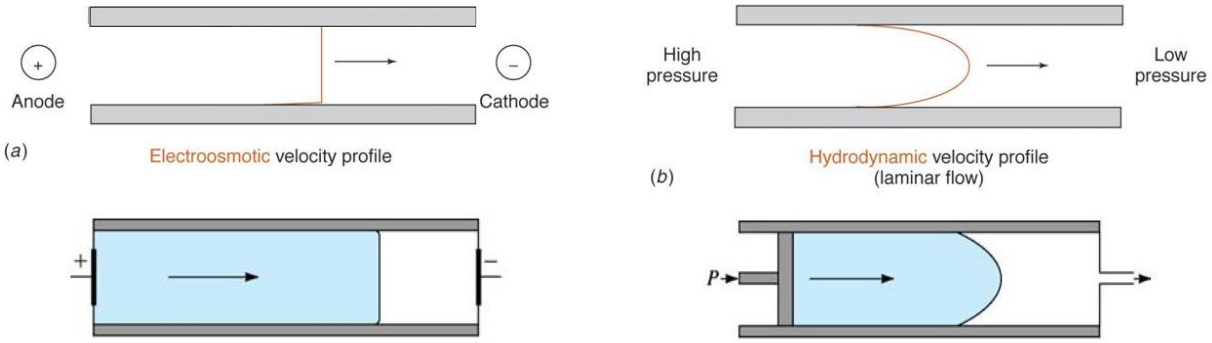


Figure 6.1 – flow profiles for liquids under (a) electroosmotic flow (capillary electrophoresis) and (b) pressure-induced flow (liquid chromatography) adapted from Dong-Sun Lee in Capillary Electrophoresis (chapter 33)

Because EOF plays such an important role in the separation mechanism, intense efforts have been devoted to its accurate control by modifying, annihilating or even reversing it. The most straightforward way to adjust EOF is to modify one or several physicochemical properties of the separation buffer such as concentration, ionic strength, pH and/or temperature. Unfortunately, because these parameters also strongly influence the electrophoretic migration of the analytes, the global outcome of such changes on the separation efficiency is difficult to predict. Another possible strategy consists in modifying the surface properties of the inner capillary wall. This strategy can be easily applied by using capillaries of different chemical composition than traditional silica-based capillaries such as Teflon or PEEK. However, the mainstream method to modify the surface properties of the capillary relies on the use of dynamic or covalent coatings of the capillary wall of bare fused silica capillaries to achieve the desired EOF characteristics. Historically, it is mainly due to the use of optic detection of the analytes (such as performed in CE-UV) which requires a transparent window to the light. Figure 6.2 shows an example where a cationic layer is non-covalently coated on the inner capillary surface, reversing the direction of EOF:

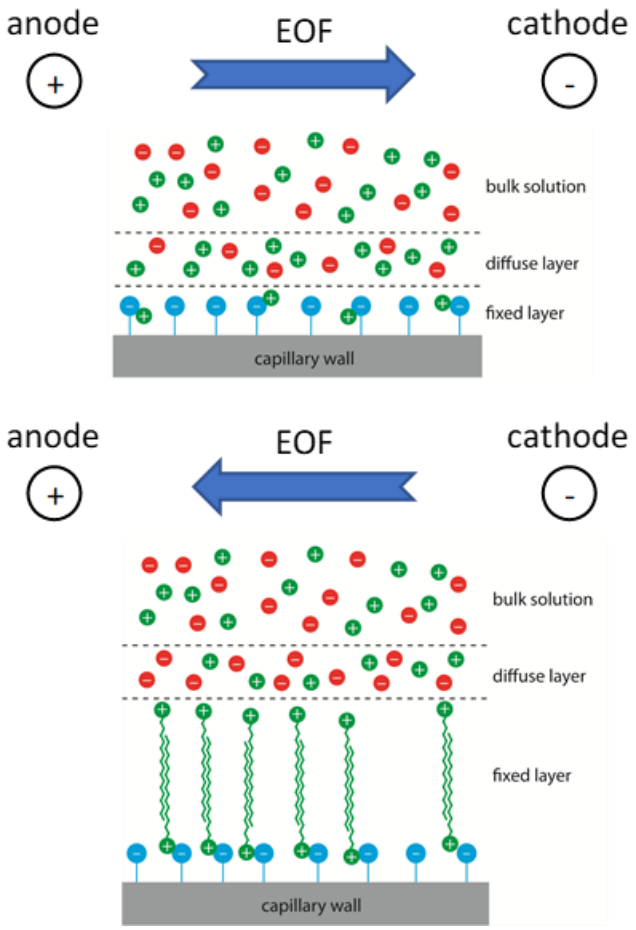


Figure 6.2 - Different EOF direction according to the surface properties of the capillary, adapted from Analytical Chemistry, D. Harvey, 2018.

In classical CE-UV, various methods have been developed to measure the EOF mobility such as the measurement of liquid mass transmitted by the EOF, the measurement of flow velocity of detectable analyte with zero electrophoretic mobility (referred as the neutral marker method, acetone is one of the most popular EOF markers), the monitoring of a current drop (a relatively large difference of conductivities is required between the solutions at the inlet and outlet of the CE capillary) and the measurement of the effective mobility of charged analyte of well characterized electrophoretic mobilities.

Due to its relative recent introduction compared to CE-UV, CE-MS still mainly relies on CE-UV methods to determine the EOF mobility. CE-UV-MS is still difficult to implement online, mostly because the required window for UV detection creates a fragile point in the capillary, which easily breaks during MS hyphenation. Consequently, the determination of the EOF in CE-MS often relies on the determination of the EOF by classical CE-UV methods using similar experimental conditions than the CE-MS method. Nonetheless, this method assumes that the CE-MS interface (or the detector in a general way) does not interfere with the EOF measurement.

Currently, the hyphenation of CE to MS is almost exclusively performed using an electrospray interface. Among the possible interface designs, the sheath liquid interface is often favored due to the increased spray stability and higher versatility towards BGE composition. However, because of the use of sheath liquid and the ESI process, a suction effect inside the CE capillary has often been reported for sheath liquid interfaces.

As a result, the apparent mobility of a particular analyte in CE-MS using a sheath liquid interface is the combination of its own intrinsic electrophoretic mobility, the EOF mobility and the suction effect due to the CE-MS hyphenation. From an analytical but also from a physical chemistry point of view, the characterization of the so-called apparent EOF (i.e. the combined effect of the effective EOF and the suction effect) is critical. First, the evaluation of the impact of the suction effect caused by the CE-MS interface is required for the efficient development of CE-MS analytical methods. Then, the determination of the respective contributions of the effective EOF and the suction effect is also crucial for determining accurate electrophoretic mobilities required to determinate the unbiased hydrodynamic radius of the species of interest.

In this chapter, we introduce the use of EOF markers for the characterization of the apparent EOF in CE-MS. Two approaches were developed based on the use of two neutral markers (acetamide and sucrose). In addition, these EOF markers were also applied in CE-MS experiments using normal and reverse polarities to determine the respective contributions of the effective EOF and the suction effect in the global apparent EOF. A system of 2 linear equations was resolved to

extract these values, assuming than the suction effect is greater than the EOF in the experimental conditions used in this study, and that the absolute value of EOF does not change between the runs using normal and reverse polarity.

Finally, an alternative method based on two fully charged analytes, i.e. tetraethylammonium and tetrapropylammonium, was also developed for the characterization of the apparent EOF in different commonly used background electrolytes in CE-MS.

The results are presented in the following section as a short communication to be submitted in Electrophoresis. The related Supporting Information (SI) are presented directly after the short communication.

6.2. Results – short communication to be submitted in Electrophoresis

Determination of the Electroosmotic Flow and Suction Effect in Capillary Electrophoresis coupled to Mass Spectrometry

Cédric Delvaux[(1)], Edwin de Pauw[(1)], and Johann Far[(1)]

(1) *Mass Spectrometry Laboratory, MolSys Research Unit, Quartier Agora, University of Liège, Allée du Six Août 11, B-4000 Liège, Belgium*

Keywords: Capillary Electrophoresis, Mass Spectrometry, Electroosmotic Flow, Suction Effect, Low Sheath-flow Interface

ABSTRACT

In this Short Communication, a method for the determination of the combination of the Electroosmotic Flow (EOF) and the suction effect (referred as the apparent EOF) in Capillary Electrophoresis coupled to Mass Spectrometry (CE-MS) using a microfluidic CE-MS interface is described. The use of neutral (acetamide and sucrose) and charged (tetraalkylammoniums) mobility markers is demonstrated and the determination of the respective contributions of the effective EOF and suction effect is performed. Finally, the impact of the sheath liquid flowrate of a microfluidic CE-MS interface on the apparent EOF was also evaluated.

INTRODUCTION

Recently, the hyphenation of Capillary Electrophoresis to Mass Spectrometry (CE-MS) has opened new exciting opportunities by combining the high resolving power and efficiency, short analysis times, and small sample consumption of CE to the outstanding characterization capabilities of MS[1]. Currently, CE-MS hyphenation is mainly implemented using an electrospray ionization (ESI) interface. In general, these interfaces are classified into sheathless, sheath liquid, and liquid junction interfaces[2]. Among these types of interfaces, the sheath liquid

configuration offers a good compromise between sensitivity and robustness as it tolerates higher concentration background electrolytes (BGE) with increased spray stability.

In classical CE separations, the electroosmotic flow (EOF) plays a critical role as its mobility (μ_{EOF}) superimposes to the intrinsic electrophoretic mobility (μ_e) of the analyte to provide its apparent mobility (μ_{app}), as described by equation (1):

$$\mu_{app} = \mu_e + \mu_{EOF} \quad (1)$$

In CE-MS, an additional effect due to the ESI process and sheath liquid supply, referred as the "suction effect", also takes place during the separation. As a result, in CE-MS, the apparent EOF mobility ($\mu_{EOF,app}$) is the combined effect of the effective EOF and the suction effect ($\mu_{suction}$). In this case, equation (1) must be expressed as:

$$\mu_{app} = \mu_e + \mu_{EOF} + \mu_{suction} = \mu_e + \mu_{EOF,app} \quad (2)$$

In general, the EOF is often considered valuable as it allows a single-point detection for all analytes regardless of their charge within a single run, as long as their electrophoretic mobility is lower than the (apparent) EOF mobility. However, it requires a strict control to avoid impairing the separation efficiency[3]. In this context, the proper characterization of the (apparent) EOF is critical, including in CE-MS. Currently, the EOF determination relies on the use of classical CE-UV methods with similar experimental conditions than the CE-MS method[4]. This method assumes that the CE-MS interface does not interfere with the EOF value due to the suction effect. A popular EOF marker in CE-UV is acetone, which is neutral in any aqueous background electrolyte and presents a fair UV absorbance.

In our work, we introduce two methods based on neutral and charged mobility markers to measure the apparent EOF mobility in CE-MS. Besides, an additional method was developed to estimate the respective contributions of the effective EOF and the suction effect. Finally, we investigated the impact of the sheath liquid flowrate on the apparent EOF for the microfluidic CE-MS interface daily used in our laboratory.

MATERIAL AND METHODS

For this purpose, a P/ACE MDQ capillary electrophoresis (Beckman Coulter – SCIEX Separations, Brea, CA, USA) operating with bare fused-silica capillaries (50 µm i.d., 365 µm o.d., 90 or 100cm in length) was either coupled to a Q Exactive or LTQ FT Hybrid mass spectrometer fitted with the nanospray Flex ion source (Thermo) via an Analis microfluidic SL CE-MS coaxial sprayer (PN 10–301 347, Analis, Suarlee, Belgium). The capillary was conditioned following a standard activation procedure described elsewhere[5]. For each run, the sample was loaded hydrodynamically at 2 psi for 15 s and the separations were performed applying 30 kV in normal polarity (except when stated otherwise) using 50 mM ammonium acetate (pH 7.00 or 6.15 measured by a pH-meter) or 10% acetic acid. The sheath liquid (SL) consisted of water/2-propanol/acetic acid (49/50/1, *v/v/v*) delivered at 1.0 µL.min^{-1} using an Hamilton syringe (250µL or 1mL depending on the desired flow rate) and a syringe pump, except when the SL flowrate was varied (between 1.0 and 3 µL.min^{-1}). The neutral markers (i.e. acetamide and sucrose), were prepared at 1mM in the 10-fold diluted background electrolyte before injection while the charged markers (i.e. tetraethyl ammonium and tetrapropylammonium) were prepared at 10uM in the same conditions. Xcalibur v2.2 or v3.1 was used to reprocess the MS data. The CE instrument was operated by 32karat v7. PeakMaster 5.3 (Echmet group[6]) was used to predict the theoretical migration times of the tetraalkylammonium compounds. All reagents were purchased from Sigma Aldrich, Belgium. MilliQ water (18.2MΩ) was daily produced using a Millipore system. The pH-meter was purchased from WTW (pH 3110 model) and was 3-point calibrated using the reference pH solutions recommended by the manufacturer (pH 4.00 – pH 7.00 – pH 10.00)

RESULTS AND DISCUSSION

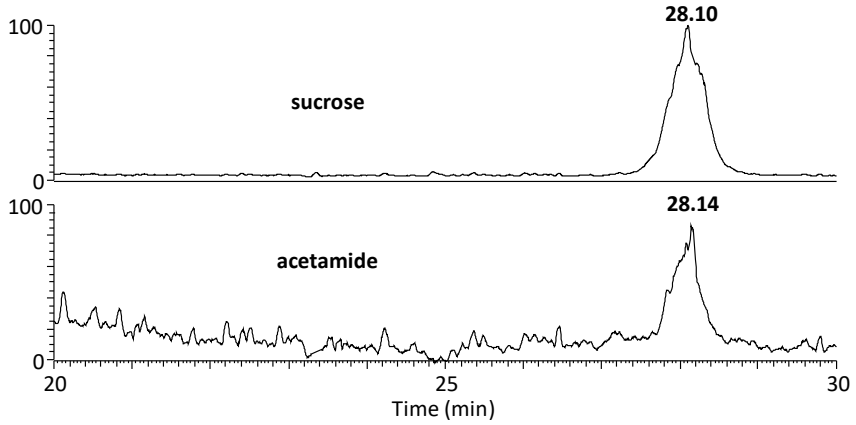
The use of acetamide and sucrose as neutral apparent EOF mobility markers

The selection of such markers was challenging due the requirement to remain a neutral species in any MS friendly aqueous background electrolyte but efficiently transferred to the gas phase and charged during the electrospray ionization (ESI). Acetamide was ionized as a proton adduct (*m/z* = 60.044) during ESI in positive mode while sucrose was either ionized as a proton adduct (m/z: 343.123) or cation adducts including sodium (m/z: 365.105), potassium (m/z: 381.079), but most interestingly ammonium (m/z: 360.150), which is commonly used in MS friendly background electrolytes. Sucrose adduct ions were also observed at 2:1 stoichiometry (*m/z* 702.266 for $[(C_{12}H_{22}O_{11})_2 + NH_4]^+$) and even higher sucrose:cation stoichiometries; fitting with the typical mass range covered by most of the MS omics application.

Using these neutral markers, the apparent EOF mobilities of a 50mM acetate ammonium background electrolyte (BGE) at pH 7.00 and pH 6.15 (measured by a pH meter) could be determined on 3 replicates, as shown in Figure 1:

(A)

NH_4Ac 50mM (pH 7.00)	Migration time (min)	$\mu_{EOF,app}$ ($m^2.(V.s)^{-1}$)
Sucrose	28.59 ± 0.37	$1.57 \times 10^{-8} \pm 2.02 \times 10^{-10}$
Acetamide	28.55 ± 0.32	$1.58 \times 10^{-8} \pm 1.78 \times 10^{-10}$



(B)

NH_4Ac 50mM (pH 6.15)	Migration time (min)	$\mu_{EOF,app}$ ($m^2.(V.s)^{-1}$)
Sucrose	33.90 ± 0.25	$1.33 \times 10^{-8} \pm 9.94 \times 10^{-11}$
Acetamide	33.84 ± 0.25	$1.33 \times 10^{-8} \pm 9.82 \times 10^{-11}$

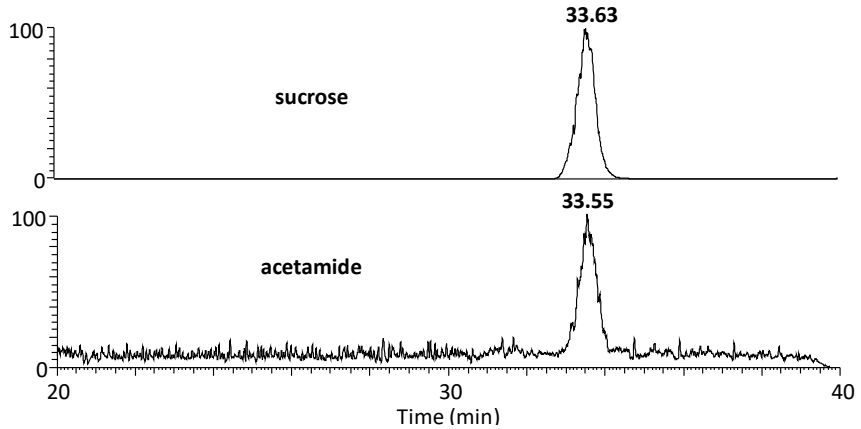


***Figure 1**: **(A)** Determination of the apparent EOF mobility on triplicates in a 50mM acetate ammonium buffer at pH 7.00 based on sucrose (detected as $[M+NH_4]^+$ at 360.15 m/z) and acetamide (detected as $[M+H]^+$ at 60.04 m/z). Typical extracted electropherograms of sucrose (middle) and acetamide (down) during separation. **(B)** Determination of the apparent EOF mobility on triplicates in a 50mM acetate ammonium buffer at pH 6.15 based on sucrose (detected as $[M+NH_4]^+$ at 360.15 m/z) and acetamide (detected as $[M+H]^+$ at 60.04 m/z). Typical extracted electropherograms of sucrose (middle) and acetamide (down) during separation*

The experimental EOF at pH 7.00 in a aqueous 50mM ammonium acetate BGE was around 1.6×10^{-8} $m^2.(V.s)^{-1}$ and 1.3×10^{-8} $m^2.(V.s)^{-1}$ for the same BGE at pH 6.15. As expected, the apparent EOF mobility decreases when the pH decreases due to the higher protonation of the silanol groups from the capillary wall, responsible for the magnitude of the effective EOF. As can be seen from Figure 1, the standard deviations of the measured apparent EOF mobilities

are close to 1%. In the present experimental conditions, acetamide suffers from a higher chemical noise than sucrose due to the use of acetate based BGEs.

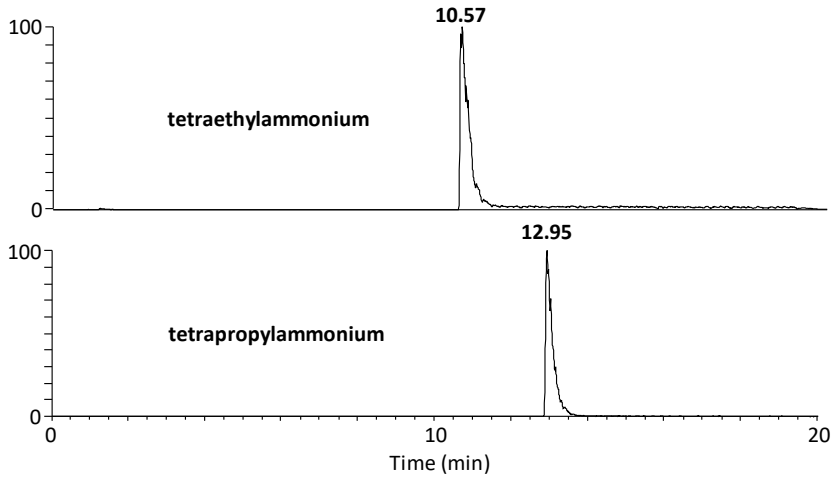
The use of tetraethylammonium and tetrapropylammonium as charged apparent EOF mobility markers

In complement to the use of acetamide and sucrose, tetraethylammonium (m/z: 130.25) and tetrapropylammonium (m/z: 186.36) can also be used for the apparent EOF mobility measurement. In this case, these species bear a constant positive charge at all pH values and their respective electrophoretic mobilities are well-defined and predictable. As a result, comparing the experimental migration times of these species to their respective computed migration times allows the determination of the apparent EOF mobility. The computation of the electrophoretic mobilities was performed using PeakMaster v5.3[6]. Based on these charged markers, the apparent EOF mobilities of the same 50mM acetate ammonium background electrolyte at pH 7.00 previously characterized by the neutral marker method could be determined on 3 replicates, as shown in Figure 2:

(A)

NH_4Ac 50mM (pH 7.00)	Migration time (min)	μ_{eff} ($m^2.(V.s)^{-1}$)
Tetraethylammonium	10.57 ± 0.01	4.26 x 10^{-8} ± 1.90 x 10^{-11}
Tetrapropylammonium	12.96 ± 0.02	3.47 x 10^{-8} ± 5.04 x 10^{-11}

(B)



(C)

NH_4Ac 50mM (pH 7.00)	Theoretical μ_e ($m^2.(V.s)^{-1}$)	$\mu_{EOF,app}$ ($m^2.(V.s)^{-1}$)
Tetraethylammonium	2.83 x 10^{-8}	1.43 x 10^{-8}
Tetrapropylammonium	1.82 x 10^{-8}	1.65 x 10^{-8}

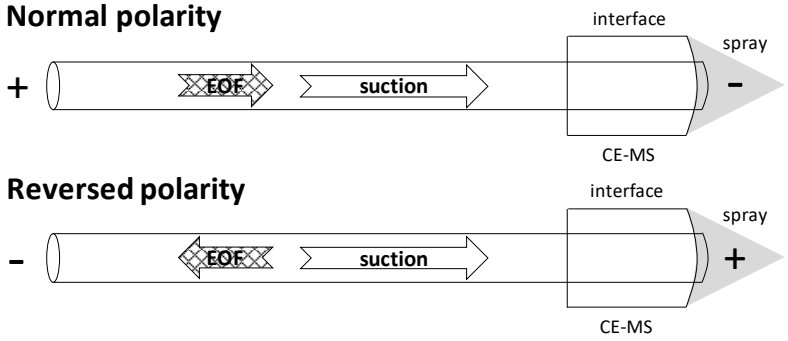
Figure 2: ***(A)*** *Determination of the effective mobility of tetraethylammonium (detected as the molecular charged species at 130.25 m/z) and tetrapropylammonium (detected as the molecular charged species at 186.36 m/z) on triplicates in an 50mM acetate ammonium buffer at pH 7.00.* ***(B)*** *Typical extracted electropherograms of tetraethylammonium (up) and tetrapropylammonium (down) during separation in an 50mM acetate ammonium buffer at pH 7.00.* ***(C)*** *Determination of the apparent EOF mobility based on the PeakMasterv5.3 computed electrophoretic mobilities subtracted to the apparent electrophoretic mobilities of the markers (from Figure. 2A) in these separation conditions.*

Interestingly, the apparent EOF mobility values determined using the charged and neutral mobility markers are closely related, especially for the values using tetrapropylammonium. The main advantage of the charged mobility markers compared to the neutral mobility markers is the reduced migration times, allowing the determination of the apparent EOF mobility 3 times faster with significantly lower standard deviations (<0.2%) than the neutral marker method.

The determination of the respective contributions of suction effect and effective EOF mobility to the apparent EOF mobility

To characterize each contribution, a method based on the neutral mobility marker acetamide was developed where its migration time was recorded in normal polarity (cathode at the CE outlet) and in reversed polarity (cathode at the CE inlet) with the same applied voltage (30kV). Because of the ESI process, the suction effect is always directed from the CE inlet to the outlet (see Figure 3A for details). Conversely, using regular and uncoated bare fused silica capillary, the EOF direction depends on the applied polarity and is either directed from the CE inlet to the outlet (normal polarity) or from the CE outlet to the inlet (reversed polarity). As a result, assuming that the EOF in normal and reversed polarities is identical, the comparison of the migration times in both polarities allows the determination of the specific contributions of the suction effect and the effective EOF (Figure 3B, see Supporting Information for detailed demonstration). Using this strategy, the EOF and suction flowrates and mobilities could be determined for a 10% acetic acid BGE. In a second series of experiments, the influence of the sheath liquid flowrate on the apparent EOF mobility was also assessed (Figure 3D to 3F).

(A)



(B)

$$\text{flowrate}_{\text{normal polarity}} = \text{flowrate}_{\text{suction}} + \text{flowrate}_{\text{EOF normal polarity}}$$

$$\text{flowrate}_{\text{reversed polarity}} = \text{flowrate}_{\text{suction}} - \text{flowrate}_{\text{EOF reversed polarity}}$$

$$\text{Hypothesis}: \ \text{flowrate}_{\text{EOF normal polarity}} \approx \text{flowrate}_{\text{EOF reversed polarity}} = \text{flowrate}_{\text{EOF}}$$

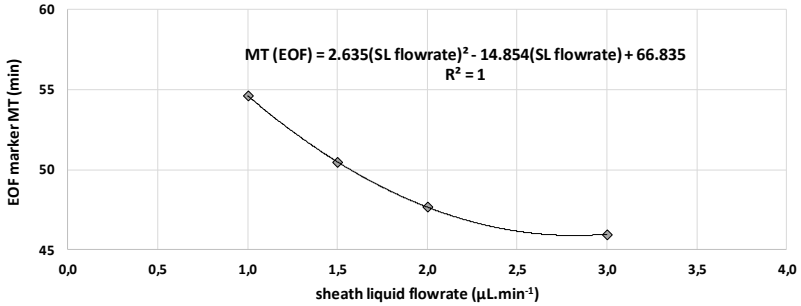
$$\text{flowrate}_{\text{normal polarity}} + \text{flowrate}_{\text{reversed polarity}} = 2\ \text{flowrate}_{\text{suction}}$$

$$\text{flowrate}_{\text{normal polarity}} - \text{flowrate}_{\text{reversed polarity}} = 2\ \text{flowrate}_{\text{EOF}}$$

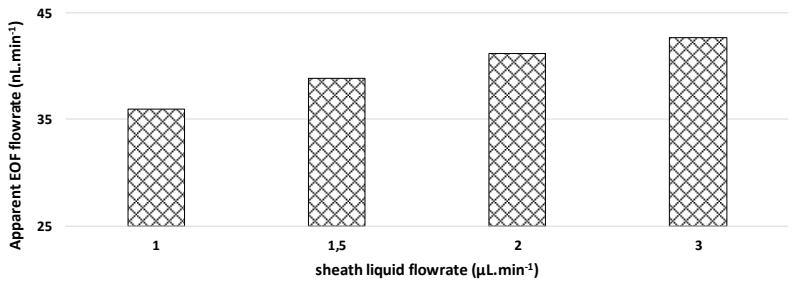
(C)

Polarity	EOF_{app} ($nL.min^{-1}$)	Suction ($nL.min^{-1}$)	EOF ($nL.min^{-1}$)	$\mu_{EOF.app}$ ($m^2.(V.s)^{-1}$)	$\mu_{suction}$ ($m^2.(V.s)^{-1}$)	μ_{EOF} ($m^2.(V.s)^{-1}$)
Normal	35.93	28.89	7.04	1.017×10^{-8}	8.18×10^{-9}	1.99×10^{-9}
Reversed	21.86					
Normal	34.77	28.68	6.10	9.88×10^{-9}	8.12×10^{-9}	1.76×10^{-9}
Reversed	22.58					

(D)



(E)



(F)

Sheath liquid flowrate ($\mu L.min^{-1}$)	EOF_{app} ($nL.min^{-1}$)	ΔEOF_{app} ($nL.min^{-1}$)	$\mu_{EOF.\ app}$ ($m^2.(V.s)^{-1}$)	$\Delta\mu_{EOF.app}$ ($m^2.(V.s)^{-1}$)
1.0	35.94	6.57	1.02×10^{-8}	1.86×10^{-9}
1.5	38.89	9.52	1.10×10^{-8}	2.70×10^{-9}
2.0	41.17	11.80	1.17×10^{-8}	3.34×10^{-9}
3.0	42.69	13.32	1.21×10^{-8}	3.77×10^{-9}

Figure 3**: **(A)** Schematic representation of the flowrates due to the EOF and the suction effect in normal and reversed polarities **(B)** Equations corresponding to the normal and reversed polarities **(C)** Determination of the apparent EOF, EOF and suction flowrates and corresponding mobilities for 2 replicates in a 10% acetic acid background electrolyte **(D)** Influence of the sheath liquid flowrate on the detected migration time of acetamide **(E)

Evolution of the apparent EOF flowrate as a function of the sheath liquid flowrate from 1.0 to 3.0 µL.min^{-1} ***(F)*** *Determination of the increment of apparent EOF as a function of the sheath liquid flowrate*

As can be seen in Figure 3C, the major contribution to the apparent EOF under these conditions is the suction effect, mostly because acetic acid-based BGEs provide very low effective EOF[7]. Interestingly, the suction flowrate obtained by the present method for a sheath liquid delivered at 1.0 µL.min^{-1} (Figure 3C) is in perfect agreement with the reported value of 28.9nL.min^{-1} by Gonzáles-Ruiz et al[4] using the classical CE-UV / CE-MS comparison method. The comparison of a sheath liquid delivered between 1.0 µL.min^{-1} and 3.0 µL.min^{-1} shows a decrease in the observed migration time of acetamide that could be perfectly fitted by a 2nd degree polynomial equation (Figure 3D). From this equation, the extrapolated migration time of acetamide at 0µL.min^{-1} (i.e. 66.835min) allows to calculate the increase of the apparent EOF mobility due to the sheath liquid flowrate. As expected, this value increases when the sheath liquid flowrate increases, as demonstrated in Figures 3E and 3F.

CONCLUSION

This work presents two efficient methods for the determination of the apparent EOF mobility (i.e. the combination of the effective EOF and the suction effect associated with sheath liquid CE-MS interfaces) based on neutral or charged mobility markers. Although both techniques allow to efficiently determine apparent EOF values, the charged mobility marker method performs 3 times faster than the neutral mobility marker with lower standard deviations. Finally, a method based on migrations in normal and reversed polarities of a neutral marker allowed to unravel the specific contributions of the effective EOF and the suction effect to the global apparent EOF was also introduced. The developed EOF and suction effect determination methods are flexible enough to be implemented in virtually any omics experiments because the proposed markers will be observed in the typical mass ranges and run duration used during CE-MS runs. Moreover, the proposed markers belong to different family of compounds and should prevent chemical (cross reactions) or mass interferences of the markers with the compounds of

interest after careful selection of the most appropriate EOF markers. These methods also correct any potential bias from the parallel CE-UV and CE-MS experiments due to the assumption that the CE-MS interface does not interfere with the EOF.

REFERENCES

(1) Schmitt-Kopplin, P.; Frommberger, M. Capillary Electrophoresis–Mass Spectrometry: 15 Years of Developments and Applications. *Electrophoresis* **2003**, *24* (22-23), 3837–3867.

(2) Bonvin, G.; Schappler, J.; Rudaz, S. Capillary Electrophoresis–Electrospray Ionization-Mass Spectrometry Interfaces: Fundamental Concepts and Technical Developments. *J. Chromatogr. A* **2012**, *1267* (0), 17–31. https://doi.org/10.1016/j.chroma.2012.07.019.

(3) Lukacs, K. D.; Jorgenson, J. W. Capillary Zone Electrophoresis: Effect of Physical Parameters on Separation Efficiency and Quantitation. *J. High Resolut. Chromatogr.* **1985**, *8* (8), 407–411.

(4) González-Ruiz, V.; Codesido, S.; Far, J.; Rudaz, S.; Schappler, J. Evaluation of a New Low Sheath–Flow Interface for CE-MS. *Electrophoresis* **2016**, *37* (7–8), 936–946.

(5) Boulanger, M.; Delvaux, C.; Quinton, L.; Joris, B.; Pauw, E. De; Far, J. Bacillus Licheniformis Peptidoglycan Characterization by CZE-MS: Assessment with the Benchmark RP-HPLC-MS Method. *Electrophoresis* **2019**.

(6) Gas, B.; Jaros, M.; Hruska, V.; Zuskova, I.; Stedry, M. PeakMaster-A Freeware Simulator of Capillary Zone Electrophoresis. *Lc Gc Eur.* **2005**, *18* (5), 282.

(7) Valkó, I. E.; Sirén, H.; Riekkola, M. Characteristics of Electroosmotic Flow in Capillary Electrophoresis in Water and in Organic Solvents without Added Ionic Species. *J. Microcolumn Sep.* **1999**, *11* (3), 199–208.

6.3. Results - Supporting Information

Demonstration of the determination of the respective influence of the effective EOF and suction effect in the global apparent EOF mobility based on a neutral marker

1) Demonstration in flowrate

If the flowrate observed in normal polarity is the combination of the effective EOF and the suction effect (both directed towards the MS), then the flowrate observed in reversed polarity is the substrction of the flowrate due to the suction effect (still directed to the MS) and the flowrate of the effective EOF (directed to the CE inlet) :

$$\text{flowrate}_{\text{normal polarity}} = \text{flowrate}_{\text{suction}} + \text{flowrate}_{\text{EOF normal polarity}}$$

$$\text{flowrate}_{\text{reversed polarity}} = \text{flowrate}_{\text{suction}} - \text{flowrate}_{\text{EOF reversed polarity}}$$

If the EOF in normal and reversed polarities are hypothesized to be equivalent:

$$\text{flowrate}_{\text{EOF normal polarity}} \approx \text{flowrate}_{\text{EOF reversed polarity}}$$

Then, the addition of the flowrates in normal and reversed polarity allows to access the contribution of the suction effect:

$$\text{flowrate}_{\text{normal polarity}} + \text{flowrate}_{\text{reversed polarity}} = 2\,\text{flowrate}_{\text{suction}}$$

$$\text{flowrate}_{\text{suction}} = \frac{\text{flowrate}_{\text{normal polarity}} + \text{flowrate}_{\text{reversed polarity}}}{2}$$

Conversely, the subtraction of the flowrates in normal and reversed polarity allows to access the contribution of the effective EOF:

$$\text{flowrate}_{\text{normal polarity}} - \text{flowrate}_{\text{reversed polarity}} = 2\,\text{flowrate}_{\text{EOF}}$$

$$\text{flowrate}_{\text{EOF}} = \frac{\text{flowrate}_{\text{normal polarity}} - \text{flowrate}_{\text{reversed polarity}}}{2}$$

2) Demonstration in apparent migration time

The flowrates observed in both polarities are linked to the capillary volume (V_{cap}) and the apparent migration time of the neutral marker (MT_{app})

$$\text{flowrate}_{\text{normal polarity}} = \frac{V_{cap}}{MT_{app}}$$

$$\text{flowrate}_{\text{reversed polarity}} = \frac{V_{cap}}{MT_{app}}$$

As expressed above, the flowrates observed in normal and reversed polarities are the addition or subtraction of the flowrate of the suction effect and the flowrate of the effective EOF. Consequently, each equation can be further developed according to the capillary volume (V_{cap}), the

migration time of the effective EOF (MT_{EOF}) and the migration time of the suction effect ($MT_{suction}$):

$$\text{flowrate}_{\text{normal polarity}} = \text{flowrate}_{\text{suction}} + \text{flowrate}_{\text{EOF normal polarity}}$$

$$= \frac{V_{cap}}{MT_{suction}} + \frac{V_{cap}}{MT_{EOF}} = \frac{V_{cap}.MT_{EOF} + V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}}$$

$$\text{flowrate}_{\text{revererse polarity}} = \text{flowrate}_{\text{suction}} - \text{flowrate}_{\text{EOF reversed polarity}}$$

$$= \frac{V_{cap}}{MT_{suction}} - \frac{V_{cap}}{MT_{EOF}} = \frac{V_{cap}.MT_{EOF} - V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}}$$

As a result, the addition of the flowrate in normal and reversed polarities allows to access the migration time of the suction effect:

$$\text{flowrate}_{\text{normal polarity}} + \text{flowrate}_{\text{reversed polarity}}$$

$$= \frac{V_{cap}.MT_{EOF} + V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}} + \frac{V_{cap}.MT_{EOF} - V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}}$$

$$= \frac{2V_{cap}.MT_{EOF}}{MT_{suction}.MT_{EOF}} = \frac{2V_{cap}}{MT_{suction}}$$

$$MT_{suction} = \frac{2V_{cap}}{\text{flowrate}_{\text{normal polarity}} + \text{flowrate}_{\text{reversed polarity}}}$$

In complement, the subtraction of the flowrate in normal and reversed polarities allows to access the migration time of the effective EOF:

$$\text{flowrate}_{\text{normal polarity}} - \text{flowrate}_{\text{reversed polarity}}$$

$$= \frac{V_{cap}.MT_{EOF} + V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}} - \frac{V_{cap}.MT_{EOF} - V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}}$$

$$= \frac{2V_{cap}.MT_{suction}}{MT_{suction}.MT_{EOF}} = \frac{2V_{cap}}{MT_{EOF}}$$

$$MT_{EOF} = \frac{2V_{cap}}{\text{flowrate}_{\text{normal polarity}} - \text{flowrate}_{\text{reversed polarity}}}$$

6.4. Conclusion and perspectives

In this chapter, the characterization of the apparent EOF mobility in CE-MS was performed based on the use of neutral and charged markers. Both methods provided similar results, although the charged marker method allows the determination of the apparent EOF 3 times faster than the neutral mobility marker with lower standard deviations.

The method developed to unravel the specific contribution of the effective EOF and suction effect revealed that the suction effect was the major contributor to the apparent EOF mobility for a 10% acetic acid because of the weak effective EOF value due to the low pH to this BGE (pH≈2.26). It is anticipated that the suction effect will have a relative lower contribution to the apparent EOF mobility for BGE characterized by higher effective EOF value (such as ammonium acetate of ammonium formate).

The impact of the sheath liquid flowrate was also investigated and revealed that an increase in the sheath liquid flowrate increases the apparent EOF, most probably because of an increase of the suction effect by the Venturi effect. Consequently, the delivered sheath liquid flowrate should be lowered to its minimum value to maintain a stable spray in order to minimize the suction effect. Lowering the sheath liquid flowrate has a beneficial effect on the detection limit but restricts the robustness to the method if non-MS friendly BGE are used.

This preliminary work opens various possible prospects for the future. First, the present methods for apparent EOF determination could be used to characterize the EOF of commonly used background electrolytes in CE-MS, including acetic acid, ammonium acetate, formic acid, ammonium formate and ammonium carbonate BGEs. This characterization could then be used to better assist the implementation of analytical procedures for efficient separations using this microfluidic sheath flow interface.

These methods could also be employed on coated capillaries (either dynamically or covalently) to characterize their apparent EOF. This prospect is of particular interest for analytical method targeting large molecules (including intact proteins) as coatings are mandatory to avoid significant adsorption of the analytes on the capillary walls for these kinds of application.

In parallel, further work to characterize the suction effect of this low sheath flow interface should be performed. For example, the use of different CE capillaries, mainly with other inner and outer diameters should be tested to characterize the suction effect associated to these parameters. It is anticipated that the use of smaller ID capillaries should decrease the suction effect. Also, the use of larger OD capillaries while maintaining the diameter of the interface needle should also decrease

the suction effect. The use of different diameters for the interface needle could also be envisaged to reduce the sheath liquid flowrate and consequently, the suction effect. Finally, the use of other background electrolytes should be considered to verify that the suction effect is not BGE dependent. The suction effect of the CE-MS interface should also depend on the viscosity of the BGE.

It should be noted that the charged markers can't be used to characterize the respective suction and effective EOF contributions because of their high intrinsic electrophoretic mobilities compared to the apparent EOF mobility. Nonetheless, the method could be adapted by using other charged markers with known mobility values matching the expected values of the apparent EOF.

Finally, the proposed method for the suction effect determination should be validated by the comparison between the same experimental conditions for this low sheath flow interface and the sheathless interface.

CHAPTER 7

Conclusions and Prospects

The main objective of this thesis was to further expand the usefulness of capillary electrophoresis coupled to mass spectrometry (CE-MS) as a valuable analytical tool for the analysis of biomolecules especially related to the omics fields. A particular focus was taken to evaluate the contribution of CE-MS in terms of the analytical performances with established analytical methods hyphenated to mass spectrometry, including High Performance Liquid Chromatography (Chapters 3 and 4) and Ion Mobility (Chapter 5). Finally, different methods were developed in Chapter 6 to evaluate and characterize the low sheath flow CE-MS interface used throughout this thesis.

In Chapter 3, the introduction of capillary zone electrophoresis coupled to mass spectrometry (CZE-MS) as a valuable alternative to RP-HPLC was demonstrated for the characterization of bacterial peptidoglycan muropeptides. Although both methods provide similar numbers of identified muropeptides, CZE-MS allowed a straightforward and in-depth characterization of the chemical modifications (mainly the amidation of the peptide moiety and the deacetylation of the glycan moiety) of PGN muropeptides. This additional characterization level is in clear contrast with the classical reversed phase HPLC method, where modified and unmodified species are most of the time co-eluted. In addition, CZE-MS allows the separation of closely related amidation isomers, which can be unambiguously identified by the use of CID or ETD tandem mass spectrometry. As a result, the developed methods in this chapter now allow biochemists to access a higher level of peptidoglycan characterization compared to the HPLC method.

Considering the efficiency of the characterization of peptidoglycan fragments that can be achieved using CZE-MS, it is anticipated that this method could become the future reference method to assess many aspects of the bacterial peptidoglycan, including the absolute quantitation and the characterization of chemical modifications induced by various growth/stress conditions. The profiling of different bacterial strains could also be considered based on a series of muropeptides that differ from strain to strain.

Further method development could be performed, in particular the reduction of the analysis time. To this end, the use of a coating to accelerate the electroosmotic flow, allowing a greater number of muropeptides to be detected (and in particular, those suspected to carry a global negative charge) or other MS-friendly background electrolytes could be tested. Finally, the development of tandem mass spectrometry, including CID and ETD fragmentations but also alternative fragmentation methods such as UVPD and IRMPD, should be further explored to extensively characterize muropeptides, with a special focus on their chemical modifications. Given the importance of these modifications to the rise of resistant bacteria, a better characterization of such modifications should bring relevant information to circumvent this public health concern.

Chapter 4 was based on the experience acquired in Chapter 3 and a new analytical method for the detection and quantitation of peptidoglycan-derived peptides was developed. These peptides are involved in the recycling pathway of peptidoglycan but the complete mechanism, including the genes, enzymes, and the biochemical pathways involved in the process, is not yet fully understood. Because these small hydrophilic peptides are mostly not retained by classical reversed phase HPLC, the development of an efficient analytical method was necessary to detect and quantify these species in bacterial cytoplasmic extracts. In this context, CZE-MS was successfully developed on three different peptides and their equivalents bearing an amidation. As for the muropeptides, the equivalents bearing an amidation were successfully separated and a total of six different species were baseline resolved by the developed method.

In addition, the absolute quantitation of one of these species (the peptide referred as "tripeptide", see Chapter 4 for related details) was developed based on the isotopically labelled ($^{13}C^{15}N$ *m*-A_2pm)

standard. The results showed that the quantitation method was reproducible, with standard deviations close to 1%, despite the sample complexity of crude bacterial cytoplasmic extracts. Also, this method demonstrates the feasibility of the absolute quantitation of all targeted peptides in samples produced in different growth/stress conditions, as long as the isotopically labelled standards are available. The normalization of the data using the flow cytometry methods allowed reaching a global reproducibility of the method better than 20%, also including the biological replicates. As a proof of concept, the quantitation of the endogenous tripeptide in bacteria grown in the presence or absence of antibiotics (cephalosporin) was shown.

In a similar way to the characterization of muropeptides, the detection and quantitation of these peptides could bring insights into the recycling pathway of PGN in bacteria. Indeed, the correlation of PGN-derived peptide levels with different growth conditions could help deciphering the impact of the use of antibiotics on the biosynthesis of the peptidoglycan. In particular, the amount of amidation or deacetylation compared to the unmodified peptides could be used as an indicator of the appearance of antibiotics resistance due to its implication into bacterial antibiotics resistance.

From the methodological point of view, in a similar way to Chapter 3, the use of capillary coating to reduce the analysis time (if sufficient separation is retained on the targeted analytes) could be considered. The use of different bacterial strains should also be tested to check the robustness of the method towards different cytoplasmic contents. In particular, it should be verified that species contained in the cytoplasmic extracts do not progressively adsorb on the capillary, altering the method separation efficiency from run to run. This remark is especially true in regards of the large difference of the wall composition of Gram positive and Gam negative bacteria and their periplasm content. Nonetheless, our successful experiments based on *Escherichia coli* (Gram negative bacillus) and *Bacillus licheniformis* (Gram positive bacillus) peptidoglycan fragment characterization provide reliable indication concerning the robustness of CZE-MS for this field.

In Chapter 5, Capillary Zone Electrophoresis coupled to Mass Spectrometry was introduced in comparison to Travelling Wave Ion Mobility-Mass Spectrometry for the characterization of cysteine connectivity of peptides bearing two intramolecular disulfide bonds. First, the analytical performance of each method, with a specific focus on the separation efficiency, was assessed. This comparison showed that baseline separation of the disulfide isomers can be achieved using CZE-MS, although preliminary pH optimization must be performed. In TWIM-MS, the separation of these closely related structures is often partial but the overall analytical procedure is shorter.

In ion mobility, the CCS difference must be greater than the ion mobility resolving power to obtain a peak separation and consequently, if the CCS values of the different disulfide isomers are too close, an efficient separation is difficult or sometimes impossible. However, in most studied cases in this chapter, the resolution of the ion mobility mass spectrometer (Synapt G2, Waters) allowed at least the partial separation of almost all studied species.

In capillary zone electrophoresis, the separation is based on both the average charge state in solution and the average hydrodynamic radius. For larger peptides, the difference in average hydrodynamic radius is the driving separation parameter. However, when the size of the peptide decreases, the difference in terms of hydrodynamic radius also decreases. In this case, the influence of a difference in the average charge state becomes a decisive separation parameter. In all studied cases, baseline resolution was achieved for the different tested disulfide isomers. In this context, the comparison of these two analytical techniques shows that CZE-MS brings a reliable alternative to IM-MS for the elucidation of cysteine connectivity in peptides bearing multiple disulfide bonds and especially, for species that cannot be resolved based on their gas phase conformations.

Moreover, the very limited sample consumption of capillary electrophoresis (around 50 nanoliters) makes this technique especially well fitted in case of extremely low availability of sample for the characterization.

In complement, theoretical calculations were also performed to unravel the influence of the different separation parameters in each analytical approach. Although this theoretical approach should certainly be further refined, the main results are in agreement with experimental results and allow to unravel the impact of each separation parameter and to predict the separation efficiency. As explained in this chapter, additional improvements of the calculation method especially for the solution are required for a more accurate and reliable prediction of the experimental results.

In addition, the comparison between results obtained in the gas phase and solution paves the way for assessing the impact of solvation during migration process. Here, the comparison between IM-MS and CZE-MS migration behaviors for identical charge states of these highly structured peptides showed that no significant structural change was observed after desolvation, as expected. The extension of this method to less structured molecular species (such as peptides without disulfide bonds) will be attempted to further explore the influence of (de)solvation on the retention of structural features.

In a more general context, the introduction of mobility-based techniques coupled to mass spectrometry represents an attractive method compared to other analytical approaches such as NMR or X ray crystallography for the elucidation of cysteine connectivity in peptides and proteins, mainly due to the low required amount of material to perform the analysis. The analysis of peptides or proteins bearing a larger number of disulfide bonds should be attempted in the future.

Finally, in Chapter 6, a preliminary investigation of the low sheath flow interface used in this thesis is presented. In particular, the characterization of the apparent electroosmotic flow (combination of the effective electroosmotic flow and the suction effect due to the sheath liquid interface) based on neutral and charged markers demonstrated a great potential for the efficient characterization of this type of interface. The presented methods are well fitted to be implementable in the omics fields due to the compatibility in terms of run duration, covered mass range, and the virtual absence of potential cross reactions (interferences) between the different EOF markers for CE-MS and the omics samples.

Besides, a method for the specific contribution of the effective EOF and suction effect to the global apparent EOF was implemented. The result of the suction effect was in perfect agreement with the reported literature on this interface based on the traditional CE-UV/CE-MS comparison using identical experimental parameters.

To conclude this thesis, the usefulness of capillary electrophoresis coupled to mass spectrometry was demonstrated in several analytically challenging areas, where either HPLC-based or IMS-based methods provide only partial information. Based on the very promising results of CE-MS presented in this work, it is anticipated that this analytical method will be increasingly implemented in analytical laboratories for various types of analyses and in particular when other analytical approaches are unable to provide the required information. However, this thesis, through very specific examples treated, has explored a small fraction of the achievable potential of CE-MS. The implementation of various modes of CE (such as capillary (micellar) electrochromatography, capillary isotachophoresis, capillary gel electrophoresis or even capillary isoelectric focusing) coupled to MS, its utilization to assess solution structural features (by the assessment of the hydrodynamic radius) or the coupling of CE to other analytical methods prior to MS analysis are

just some of several possibilities for further implementation of CE-MS in analytical chemistry. A challenging objective for the next months will be to apply CE-MS to the analysis of naturally occurring oligonucleotides and modified oligonucleotides, and their complexes.

After several decades of intense research, the development of reliable interfaces turned CE-MS into a very competing analytical platform. Compared to well-established analytical methods such as HPLC-MS, CE-MS is still not yet a mature technique. It requires further research work and, in particular, fully validated applications developed to get acceptance by the analytical community. In this context, it is anticipated to claim that CE-MS will be progressively imposed as a commonly accepted tool for diverse analytical analyses, including the analysis of biomolecules. The characterization of antibodies and nanobodies is already massively implemented using CZE-MS, providing the necessary compliance with pharmaceutical regulatory requirements.

The main limitation to the use of CE in academic or company laboratories is probably the requirement to "forget" all the troubleshooting instincts and discipline acquired from HPLC and U(H)PLC background during method transfer or the design of new CE methods from scratch. Nonetheless, the time committed to learn about the CE theory and its practical aspects could save huge amounts of time compared to the time spent for developing HPLC methods which are not well adapted for some families of compounds (e.g. analysis of very polar zwitterions analysis is straightforward using CE-MS) or poorly adapted with MS detection (e.g. methods based on size exclusion or ionic chromatography hyphenation with MS should be easier performed when switching to CE-MS).