Development of MALDI-TOF mass spectrometry based methods for the identification and molecular characterization of proteins, phosphoproteins and DNA adducts

Von der Naturwissenschaftlichen Fakultät

der Gottfried Wilhelm Leibniz Universität Hannover

zur Erlangung des Grades

Doktor der Naturwissenschaften

Dr. rer. nat.

genehmigte Dissertation von

Dottore in Scienze biologiche

Ignazio Garaguso

Geboren am 07 Dec. 1970 in Erice (TP), Italien

2011

Referent :   Prof. Dr. Juergen Borlak

                 Lehrstuhl für Pharmako- und Toxikogenomik

                 Medizinische Hochschule Hannover

                 Abteilung für Molekulare Medizin und Medizinische Biotechnologie

                 Fraunhofer Institut für Toxicologie und Experimentelle Medizin (ITEM), Hannover

Korreferent: Prof. Dr. rer. nat. Andreas Pich

                 Institut für Toxikologie

                 Medizinische Hochschule Hannover

Tag der Promotion: 16 März 2011

Die vorliegende Arbeit wurde in der Abteilung für Molekulare Medizin und Medizinische Biotechnologie am Fraunhofer Institut für Toxikologie und Experimentelle Medizin in Hannover angefertigt.

Erklärung:

Hiermit versichere ich, dass ich die vorliegende Arbeit selbständig und nur mit Hilfe der angegeben Hilfsmittel und Quellen verfasst habe. Die entnommenen Stellen aus benutzen Werken wurden wörtlich oder inhaltlich als solche gekennzeichnet.

Diese Dissertation wurde bisher weder als Masterarbeit, Diplomarbeit noch als sonstige Prüfungsarbeit verwendet.

Hannover, im Dezember 2010

….A Stephanie & Marta

Table of Contents

Zusammenfassung

Die MALDI-TOF Massenspektrometrie (MS) ist eine etablierte Technologie, die sich für die Analyse zahlreicher Biomoleküle wie etwa Nukleinsäuren, Proteinen und deren posttranlsationalen Modifizierungen bewährt hat.

In der vorliegenden Arbeit wurde eine neue Methode mittels MALDI-TOF Massenspektrometrie entwickelt, um bei Addukten an Nukleinsäuren einerseits die Chemikalien zu analysieren, die zu einer Interaktion mit den Nukleinsäuren führen und zum anderen auch die daraus resultierenden DNA-Addukte nachzuweisen.

Für die Entwicklung dieser Methode, war es vor allem erforderlich eine neue Probenpräparation für die massenspektrometrische Analyse zu etablieren. Diese neue „matrix layer" Probenpräparation hat sich als sehr geeignet für die Analyse der o.g. Biomoleküle erwiesen, da sowohl ein zuverlässiger Nachweis mit hoher Nachweisempfindlichkeit, sowie eine umfassende Charakterisierung der molekularen Beschaffenheit dieser Biomoleküle möglich ist.

Hervorzuheben ist die vielseitige Verwendbarkeit der MALDI-TOF MS Technologie in Kombination mit der neu entwickelten Methode zur Probenpräparation, welche eine umfassende und detaillierte Analyse von Proteinen auch mit hohem Probendurchsatz ermöglicht, wie es für die meisten Proteomstudien erforderlich ist (Kapitel I). Darüber hinaus wurden vereinfachte und robuste Verfahren entwickelt, die eine Anreicherung und Aufreinigung der zu analysierenden Biomoleküle mittels MALDI-TOF MS ermöglichen, wodurch die Identifizierung und molekulare Charakterisierung von DNA-Addukten erfolgen kann, als auch die Ermittlung von Phosphorylierungen an Proteinen (Kapitel II & III).

Für den Nachweis und die Identifizierung von phosphorylierten Proteinen wurde zudem eine neue Methode entwickelt, die auf der Kombination von zweidimensionaler Gelelektrophorese, einer Immundetektion und der Verwendung von MALDI-TOF MS beruht (Kapitel III).

Mit der vorliegenden Arbeit konnte gezeigt werden, dass die Entwicklung einer optimierten Methode zur Probenpräparation eine analytische Technologie wie die Massenspektrometrie effizient nutzbar macht, um mit hoher Nachweisempfindlichkeit verschiedenartige Krebs-relevante Biomoleküle nachzuweisen sowie deren molekulare Struktur zu charakterisieren. Die Anwendung dieser neuen MALDI MS basierten Methode für den Nachweis, die Identifizierung und Charakterisierung von so unterschiedlichen Biomolekülen wie Proteinen, phosphorylierten Proteinen und Addukten von Nukleinsäuren stellt daher einen vielversprechenden Beitrag in der Krebsforschung dar.

Schlagwörter: MALDI-TOF, Massenspektrometrie; Protein-Analyse; Proteomik, Protein-Phosphorylierung; DNA-Addukte.

x

Abstract

MALDI-TOF is versatile mass spectrometric technology which can be applied to the study of a variety of molecules such as nucleic acids, proteins and their post translational modifications.

In this thesis novel MALDI-TOF mass spectrometry (MS) based methods were developed to improve detection of chemicals interacting with nucleic acid and to identify and characterize DNA adducts, proteins, and protein posttranslational modifications. This necessitated the development of a new "matrix layer" sample preparation that showed be extremely useful for sensitive detection and molecular characterization of these biomolecules. The versatility of MALDI-TOF MS technology in combination with the new developed methods permitted in-depth analyses of proteins and automated data acquisition allowing high throughput required in proteomic studies (Chapter I). Furthermore, simplified and robust procedures for analyte enrichment/purification were implemented with MALDI tandem MS for the identification and molecular characterization of DNA adducts and determination of protein phosphorylation sites (Chapter II & III). Moreover a method was developed for the reliable detection and identification of phosphorylated proteins based on a combination of Two-dimensional gel electrophoresis, immunodetection and MALDI tandem MS (Chapter III).

Noteworthy, with this doctoral thesis it was demonstrated that a single mass spectrometric technology enables sensitive detection and reliable identification through molecular characterization of these cancer-related biomolecules. The new developed MALDI MS based methods provide the capability to detect, identify and qualify protein, protein phosphorylation and DNA-adducts in a high sensitive manner on a systematic scale, and therefore, can be applied to biomedical and cancer research.

Keywords: MALDI-TOF; mass spectrometry, protein analysis; Proteomics, protein phosphorylation; DNA adducts.

List of original publications

Publication I:

<u>Ignazio Garaguso</u> and Jürgen Borlak : Matrix Layer Sample Preparation: an Improved MALDI-MS Peptide Analysis Method for Proteomic studies. *PROTEOMICS*, 2008, 8,(13), 2583–2595.

Publication II:

<u>Ignazio Garaguso</u>, Roman Halter, Jacek Krzeminski, Shantu Amin and Jürgen Borlak: Method for the Rapid Detection and Molecular Characterization of DNA Alkylating Agents by MALDI-TOF Mass Spectrometry. *Anal. Chem.*, 2010, 82, (20), 8573–8582.

Publication III:

<u>Ignazio Garaguso</u> and Jürgen Borlak : Rapid screening assay to search for phosphorylated proteins in tissue extracts. (*Submitted*)

The original publications have been reprinted with permissions of the copyright holders.

In this thesis the original publications are referred to their numbers from this list (Publication I - III).

This thesis includes also unpublished data.

List of acronyms and abbreviations

2-DE	Two-dimensional gel electrophoresis
B[a]P	Benzo[a]pyrene
B[a]PDE	7,8 dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene,
B[c]Ch	Benzo[c]chrysene
B[c]ChDE	9,10-dihydroxy-11,12-epoxy-9,10,11,12-tetrahydrobenzo[c]chrysene
CAD	Collisional activated dissociation
CHCA	α-Cyano-4-hydroxycinnamic acid
CID	Collision-induced dissociation
CNLS	Constant neutral loss scan
CYP450	Cytochrome P450
Da	Dalton
DD	Dried-Droplet
DE	Dihydrodiol epoxide
DFF	DNA-adducts fragment fingerprint
dG	2'-deoxyguanosine
dGp	2'-deoxyguanosine-3'-monophosphate
DHB	2,5-Dihydroxybenzoic acid
DIE	delayed ion extraction
DMF	DNA-adduct mass fingerprint
Er:YAG	Yttrium-aluminum-garnet-crystals doped with erbium
ESI	Electrospray
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum
HPLC	High performance liquid chromatography
Hz	Hertz
IARC	International Agency for Research on Cancer
ICR	Ion cyclotron resonance
IEF	Isoelectric focusing
IMAC	Immobilised Metal Affinity Chromatography
IR	Infrared
IRMPD	Infrared multiphoton dissociation
ISD	In-source decay
keV	kilo electron Volts
LC	Liquid chromatography
LDI	Laser desorption/ionization
LID	Laser-induced dissociation
LMR	Low mass range
m/z	Mass to charge
MAPK	Mitogen activated protein kinase
ML	Matrix layer
MOAC	Metal Oxide Affinity Chromatography

MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MW	Molecular weight
Nd:YAG	Yttrium-aluminum-garnet-crystals doped with neodymium
ns	Nanoseconds
PAH	polycyclic aromatic hydrocarbon
PFF	peptide fragment fingerprint
PhIP	2-amino-1-methyl-6-phenyl-imidazo[4,5-ß]pyridine
pI	Isoelectric point
PIS	Precursor ion scan
PMF	Peptide mass fingerprint
PSD	Post-source decay
PTM	Posttranslational modifications
S/N	Signal-to-noise ratio
SDS-PAGE	Sodium dodecyl sulfate-Polyacrylamide gel electrophoresis
THAP	2,4,6-Trihydroxyacetophenone
TIS	Timed ion selector
TL	Thin Layer
TOF	Time-Of-Flight
TQ	Triple quadupole
UV	ultraviolet
λ	Wavelength

Preface

Cancer is one of the leading causes of morbidity and mortality accounted for arount 13% of all death worldwide [1]. In Europe, more then 3.4 million new cases are diagnosed and 1.8 million men and women died of cancer in 2008 [2] , Prognosis for lung and liver cancer patients is poor with 5-year survival rates being less than 10% and 5%, respectively [3]. The exceptionally high mortality of most neoplasms is partialy due to our inability to diagnose the diseases at an early stage. In fact, 50% of patients already have distant metastases at diagnosis [4]. Curative surgery is efficacious only for those patients who are diagnosed sufficiently early in the disease process. If lung cancer, for example, is localized at the time of diagnosis and treated promptly by surgery, the 5-year survival rate increases to 52% [5]. Therefore, early detection is vital for the effective treatment of neoplasms, providing evidence to substantiate the need of analytical methods for cancer biomarkers discovery. These biomarkers can be used in population-wide screening programs and diagnostic purposes in the hope to reduce mortality and morbidity in cancer [6]. The knowledge about the causative factors and their role in carcinogenesis is crucial to understand the evolution of the disease. This knowledge will drive the development analytical method to search, select and isolate cancer-related biomolecules which can be used as biomarker for monitoring and screening purposes. However, the etiology of most neoplasms is not always evident. For instance, in lung cancer tobacco smoking is the primary etiological factor [7], but development of neoplasms is the result of a combination of several causes [8]. Numerous environmental and industrial produced toxicants and life stile factors involving mutagens and carcinogens exposure are known to be causative elements of cancer. Specifically, molecules such as polycyclic aromatic hydrocarbons (PAH), nitrosamine, halogenated aromatic hydrocarbons, polychlorinated biphenyls (PCB), dioxin and aromatic amines are ubiquitous environmental pollutants and in part components of cigarette smoke. Indeed, increased risk of lung, skin and bladder cancer has been associated with exposure to these carcinogens [9-13]. However, most carcinogens require metabolic activation to exert their carcinogenic effects. Moreover, the balance between metabolic activation and detoxification differs among individuals and affects cancer risk [14]. Metabolic activation of carcinogens leads to intermediates that might react with nucleic acid and proteins leading to DNA, RNA, and protein adduct formation and consequently to mutation. The absolute central role of carcinogens and their DNA adducts in chemical-induced cancer is illustrated by the scheme in Figure 1 [7,8]. If DNA adducts escape cellular repair mechanisms and persist, they may lead to miscoding, resulting in permanent mutations. Cells with damaged DNA may be removed by apoptosis. If a permanent mutation occurs in a critical region, it can lead to activation of oncogenes or deactivation of tumor suppressor genes. Consequently, aberrant cell formation with loss of normal

growth control can lead to cancer development, as shown in the central track of the scheme in Figure 1 [15,16]. Therefore, carcinogens are the key connection between environmental/industrial produced toxicants, life stile factors and cancer. The upper track of the scheme in Figure 1 depicts that carcinogens such as PAH, PCB and tobacco-specific nitrosamines can bind directly to certain receptors like Aryl hydrocarbon receptor (AhR) [17] leading to activation of cellular regulatory factors such as AKT and protein kinase A (PKA) [7,17,18]. This can result ultimately in decreased apoptosis and increased angiogenesis as well as cell transformation. These changes may enhance the effects of carcinogens and their DNA adducts. Furthermore, co-factors such as co-carcinogens, tumor promoters and irradiation may amplify the activity of carcinogens through a variety of mechanisms [19,20], as illustrated in the lower track of the scheme in Figure 1.
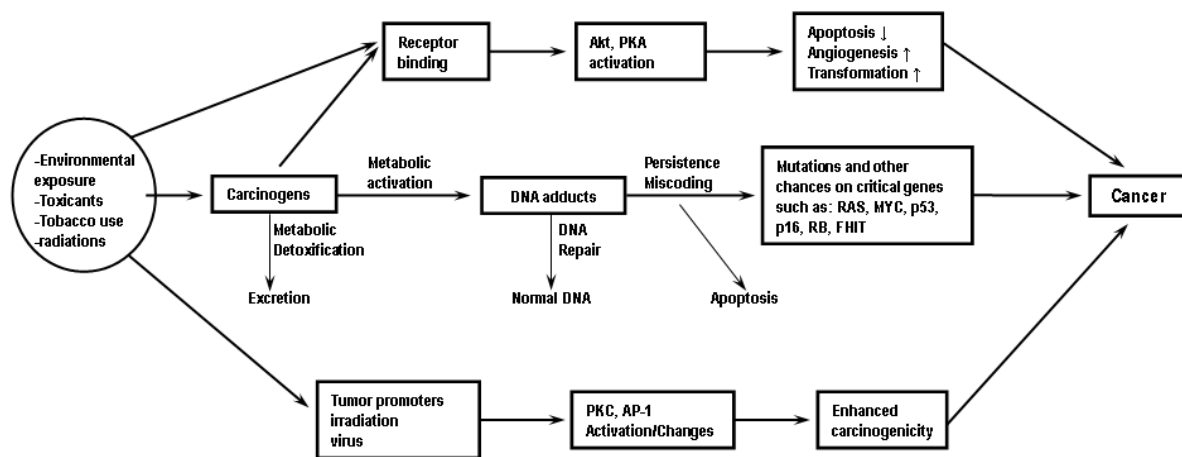
Figure 1. Scheme linking exposure to carcinogens and cancer. Adapted from ref. [7,8].

The development of cancer is generally considered a multistep process driven by carcinogen-induced genetic and epigenetic damage and activation of cellular regulatory factors of the network of signaling pathways in susceptible cells, which as result gain a selective growth advantage. Simultaneous and persistent multiple events of this type lead to aberrant cells with loss of normal growth control and ultimately to cancer [15,16]. While carcinogenesis if normally caused by genomic mutations the subsequent translational changes in the protein products indicate both the molecular mechanisms and potential markers of neoplasia [21]. Notably, the interactions and regulation of proteins function in intricate networks are indicative of biological complexity downstream from the alterations within the genes of the neoplastic cells. There are evidences which demonstrated relationships between environmental and chemical carcinogens and DNA adducts formation which leads to changes in the regulation of protein posttranslational modifications (PTM). For instance, the phosphorylation of the H2A histone variant H2AX has been directly related with the exposure to PAH [22-24]. Studies on cultured cell lines treated with PAH demonstrated direct and indirect changes of phosphorylation status of insulin growth factor

signaling pathways [25], epidermal growth factor (EGF)-receptor signaling [26], protein kinase C [27,28] as well as mitogen activated protein kinase (MAPK) such as extra cellular receptor kinase (ERK) 1/2 and p38 [19,29]. Indeed, reversible protein phosphorylation is a ubiquitous cellular mechanism for the control of protein functions and signal transduction networks that regulates diverse biological processes. The phosphorylation/deposphorylation can change the protein's enzyme activity, the cellular location, increase protein–protein interactions and target proteins for degradation, all of which induce many essential cell processes such as signal transduction, cell differentiation, proliferation, metabolic maintenance, cell division, and apoptosis [30]. Notably, receptor tyrosine kinase signal transduction pathways are the major regulators of cell proliferation and are frequently found to be mutated and activated during tumor development [31]. Moreover, the protein kinase activities, regulated through phosphorylation in turn, plays an important role in cancer and it has been well described to be involved in development of malignancies by generating inappropriate signals [32-34].

Taken collectively, understanding carcinogenesis, tumor progression, and metastasis requires not only the knowledge of the critical triggering events at the DNA level but also a careful analysis of effectors molecules such as proteins and their PTM, which act as regulatory components of the network of signaling pathways that drive neoplasia [14,34]. The availability of methods for investigation at the levels of DNA, protein and PTM and their regulation on a global scale has tremendous potential to enhance the knowledge of the cellular events that occur in response to causing agents. The study of cancer-related biomolecules will permit to understand the regulatory roles, to interpreting signaling pathways and other cellular processes that contribute to cancer development and metastasis. Therefore, the search for biomarkers for an early detection of cancer needs reliable methodology for detection of cancer-related biomolecules such as DNA-adducts, proteins and phosphoproteins. To gain insides in these processes it is important to have the use of experimental methodology for the detection and characterization of these cancer-related biomolecules. Although for this purpose a number of analytical methods are currently used, there are based on different and in some cases unrelated technologies. It will be helpful to have a unique technological platform for the analysis of all these biomolecules.

Aims of this doctoral thesis

The above examples give an overview on the variety of implications related to the DNA adducts formation, regulation of protein and protein posttranslational modifications can have in cell biology and cancer development. Moreover, underline the importance to develop analytical methodologies for the study of these biomolecules.

MALDI is a versatile mass spectrometry technology which permits high specificity and sensitivity, wide applicability, high throughput and is easy to use. The aim of this thesis was to develop more efficient methods based on MALDI-TOF MS for the study of proteins, DNA adducts and protein posttranslational modifications. These methods could be applied to solve biological questions regarding the roles played by these molecules as well as their regulation in normal cell and cancer cell metabolism.

A major focus was to develop novel MALDI matrix sample preparations to be applied for the identification and molecular characterization of proteins, PAH-DNA adducts and protein phosphorylation in a robust manner. Moreover, rapid and affordable techniques capable of efficient enrichment for these molecules from complex mixture should be developed in order to be implemented with MALDI MS. These methods should enable: a) simplified and rapid enrichment procedures; b) higher quality MS spectra, and c) acquisition of more MS and MS/MS measurements per sample.

Therefore, the specific aims of the research projects for this doctoral thesis were:

1. Development of novel MALDI MS sample preparation methods for improved peptide ionization and fragmentation, which provide higher quality MS information relevant for increased protein sequence coverage. These methods should be compatible to unambiguous peptide and protein identification and characterization as well as high throughput analyses (I).

2. Development of simple but efficient procedures for the extraction and enrichment of DNA adduct molecules. Moreover, new MALDI MS based methods for sensitive and reliable detection, identification and molecular characterization of adducted nucleotides and the alkylating molecules (II) should be developed.

3. Development of methods to detect phosphoproteins from tissue protein extracts and to efficiently enrich phosphopeptides minimizing sample treatment. Moreover, develop new MALDI-TOF MS sample preparation and procedures for improved ionization and fragmentation of phosphopeptides in order to characterize the phosphorylation sites (III).

The following chapters intend to familiarize the reader with the merits and pitfalls of the most prominent approaches used to study proteins, DNA adducts and protein phosphorylation in order to explain how each of these methods can be used to gain insight into the complex world of these biomolecules. Contributions of the author with this doctoral thesis to the relatively new fields are presented and summarized in the final paragraphs of each chapter.

Introduction
Mass spectrometry for biological sciences

Mass spectrometry (MS) is one of the most comprehensive analytical technique currently at the disposal of scientists[35]. High sensitivity and specificity are their peculiarities which led the widely use of this analytical technique in several areas of physic, chemistry, geology, nuclear science, material science, archaeology, petroleum industry, forensic science and environmental science. Indeed, MS provides high molecular specificity, analyte molecular ions and structural diagnostic fragment ions can be analyzed. MS permits high sensitivity detection, molecules can be detect at the zeptomolar level[36]. MS is versatile and it is applicable to all elements and can be used to determine the structure of most class of unknown compounds. The chemical mature of the molecules that could be analyzed is virtually unlimited: volatile or non-volatile, polar or non-polar, solid, liquid or gaseous. In addition, when used in combination with high resolution separation techniques, MS is qualified to analyze very complex samples[36]. Therefore, MS is applied to molecular mass determination, structure elucidation, mixture analysis and quantification at trace level. However, the application of MS to biological fields remained negated for long time, mostly because of the lack of suitable ionization technique for molecules of biological origin. The major challenge was to generate ions from non polar compounds of large molecular mass without molecular fragmentation. Since the introduction of the gentler modes of ionization also the analysis of large biomolecules becomes feasible. In the past, the exorbitant costs and the need of skilled operators restricted MS to the domain of elite of laboratories[37]. Currently, a number of different types of MS instrumentation for several specific applications are commercially available of reasonably price and user friendly. As result, MS has become an essential component of contemporary chemical and biochemical research laboratory. This chapter intend familiarize the reader with the basic concepts of biological mass spectrometry and in particular of matrix assisted laser desorption/ionization (MALDI), the MS technology on which are based the method developed in this doctoral thesis.

Mass spectrometry

Mass spectrometry is an analytical technique to determine the molecular mass of free ions. Mass spectrometers are extensively used in science. MS is used for determining the masses and for elucidating the chemical structures as well as the elemental composition of molecules. In order to be measured, individual molecules are converted to ions so that they can be moved and manipulated by external electric and magnetic fields for the analysis. Structural information can be generated by fragmenting the sample inside the instrument and analyzing the products generated. This is achieved using certain types of device, usually those with multiple analyzers which are known as tandem mass spectrometers. A mass spectrometer can be divided into three fundamental parts, namely the ionization

source, the analyzer, and the detector (Figure 2). First, the sample has to be introduced into the ionization source of the instrument. In the ionization source a beam of gaseous ions from the analyte is formed, then, in the mass analyzer, the ions are separates according to their mass-to-charge ratios (m/z).

Ion source	Mass analyzer	Detector

Ion formation	Ion separation	Ion detection
Electron Impact (EI)	Magnetic sector	Faraday Cup
Chemical Ionization (CI)	Electric sector	Secondary Electron multiplier (SEM)
Fast Atom Bombardment (FAB)	Quadrupole	Micro channel Plate (MCP)
Electrospray Ionization (ESI)	Ion trap	
Matrix Assisted Laser Desorption Ionization (MALDI)	Ion cyclotron	Delay detector
	Time of Flight	scintillation
	Orbitrap	

Figure 2. Schematic representation of a mass spectrometer components.

The separated ions are detected and this relative ion current (signal) sent to a data system, where the m/z ratios are stored together with their relative signal abundance to produce a mass spectrum. In modern mass spectrometers the entire operations are under complete data system control.

The sample can be inserted directly into the ionisation source. Alternatively, the sample can undergo previous chromatographic (e. g. liquid chromatography, LC; gas chromatography, GC) or electrophoretic (capillary electrophoresis, CE; gel electrophoresis: SDS-PAGE, 2-DE) separations into series of analyte components. The components are introduced directly (on-line) or sequentially (off-line) into the ionization source for individual analysis.

There are several different ionization techniques, adding or removing an electron to/from the molecules is the generally way to ionize the analytes. The analyte can be impacted by a beam of electrons with sufficient energy to ionize the molecule (electron impact, EI). The analyte can be exposed to a gentler proton transfer process that preserves and promotes the appearance of the molecular ion (chemical

ionization, CI), or by using a molecule that contains electron-capturing moieties (negative chemical ionization, NCI). The analyte can be bombarded with atoms or ions (fast atom bombardment, FAB), irradiated by photons of a high-intensity laser pulse (Laser desorption/ionization, LDI) or using an UV adsorbing molecule as matrix to mediate the process of LDI (matrix assisted desorption/Ionization, MALDI). A further ion formation process is denoted Electro spray ionization (ESI) which is included in the general term "atmospheric pressure ionization" (API) of related techniques capable of creating ions at atmospheric pressure rather than in a vacuum[35-37].

Separation and detection of analytes is subsequently exploited by the combination of several types of analyzers with suitable detectors. The main function of the mass analyzer is to separate, or resolve, the ions formed in the ionization source of the mass spectrometer according to their *m/z*. For instance, the process of ion separation can be performed a) by combining a magnetic and an electric field e.g.: sector field instruments; b) inside the radio frequency (RF) field of a quadrupole; c) by a magnetically ion trap e.g.: Ion Cyclotron Resonance Cell, ICR; d) after a flight time in an ion flight tube e.g.: Time-Of-Flight instruments; TOF; e) by an electrostatically ion trap, e.g. orbitrap. These mass analyzers have different features, including the *m/z* range that can be covered, the mass accuracy, and the achievable resolution[35-37].

The detector monitors either the charge induced or the current produced when an ion passes by or hits a surface. Because the number of ions leaving the mass analyzer at a particular instant is quite small, considerable amplification is necessary to get a signal. Therefore, the type of detector is supplied to suit the type of analyzer. Several types of electron multiplier, Faraday cups and micro-channel plate detectors are commonly used in modern commercial instruments.

Finally, the signal is then transmitted to the data system, where it is recorded in the form of mass spectra. The *m/z* values of the ions are plotted against their intensities to show the number of components in the sample, the molecular mass of each component, and the relative abundance of the various components in the sample.

Tandem (MS/MS) mass spectrometers are instruments that have more than one analyzer and so can be used for structural and sequencing studies. Two, three and four analyzers have all been incorporated into commercially available tandem instruments, and the analyzers do not necessarily have to be of the same type, in which case the instrument is a hybrid one.

Notably, with the development of soft ionization techniques such as matrix assisted laser desorption/ionization (MALDI) [38,39] and electrospray ionization (ESI) [40,41] the analysis of large biomolecules such as proteins became feasible. These ionization techniques are described as "soft" because they allow large molecules to be transferred to the gas-phase and ionized without significant

fragmentation. The following sections will cover the type of mass spectrometers and applications that are relevant for this thesis.

ESI MS

Although the phenomena of electrospray (ESI, abbreviations) have been known for hundreds of years, only in the late 1980's John Fenn and co-workers demonstrated the basic experimental principles and applied methodologies of the ESI technique for large biomolecules [40,41]. Using ESI they obtained "soft" ionization of non-volatile and thermally labile compounds such as proteins. ESI MS is now a basic tool used in probably every biological chemistry laboratory in the world. To perform ESI (Figure 3) the analyte is introduced to the source in solution either from a syringe pump or as the eluent flow from a chromatograph. The analyte solution flow passes through a metal electrospray needle (the emitter) that is subjected to a high potential difference. Typically, about 2-5 kV are applied with respect to the counter electrode, i. e. the spectrometer sampling cone. The high potential difference forces the spraying of charged droplets carrying a surface charge of the same polarity of the emitter. Liquid protrudes from the emitter forming a "Taylor cone". The droplets are repelled from the emitter towards the counter electrode on the source sampling cone. While droplets traverse the space between the emitter the cone solvent evaporates. As the solvent evaporation occurs, the droplet shrinks until it reaches the point that the surface tension can no longer sustain the charge (the Rayleigh limit) at which point a "Coulombic explosion" occurs and the droplet is ripped apart. This produces smaller droplets that can repeat the process as well as naked charged analyte molecules.

There are two accepted theories for the mechanisms in which the ions are freed into the gas phase in ESI. The charged residue model (CRM) [42] suggests that droplets undergoes a sequence of Rayleigh instabilities (Coulomb fission) together with solvent evaporation leading progeny droplets that contain on average one analyte ion. The gas-phase ions are then formed after the remaining solvent molecules evaporate, leaving the analyte with the charges that the droplet carried. The ion evaporation model (IEM) [43] suggests that the ions are emitted into the gas phase as the droplet reaches a certain radius and the field strength at the surface of the droplet becomes large enough to assist the field desorption of solvated ions. While there is no definite scientific proof, a large body of indirect evidence suggests that small ions are liberated into the gas phase through the ion evaporation mechanism, while larger ions form by charged residue mechanism [44].

ESI is a very soft ionization method as little residual energy is retained by the analyte, resulting in reduced (usually none) fragmentation upon ionization. For structural elucidation studies, tandem mass spectrometry for analyte molecules fragmentation is then required.
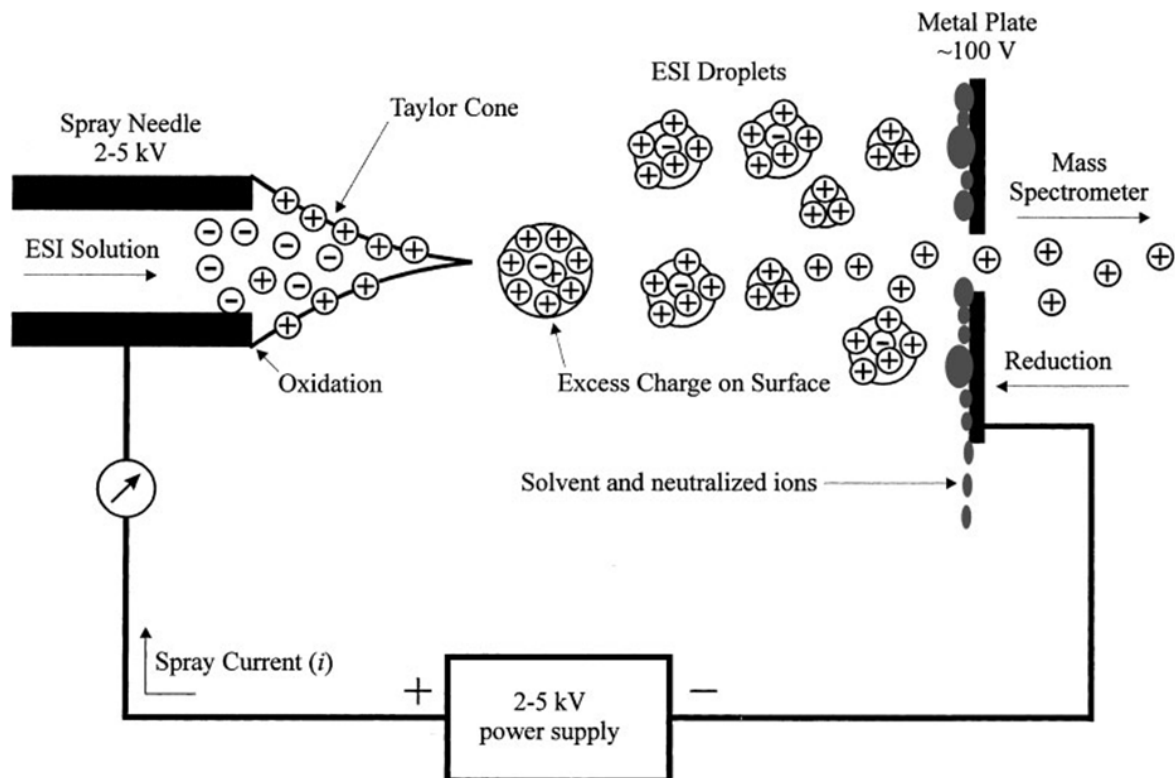
Figure 3. Schematic of the electrospray ionization process. The sample is pumped through the open capillary which a high voltage is applied. A Taylor cone is formed when charges start to move towards the counter electrode (mass spectrometer). When the Rayleigh limit has reached droplets detach from the Taylor cone. When droplets move towards the mass spectrometer, ions are freed into the gas phase either entirely by evaporation of the solvent or by effects combined from evaporation of the solvent and coulomb fission. From ref. [44]**.**

MALDI MS

Since the 1970's, laser devices have been employed in mass spectrometry to achieve a direct desorption/ionization (Laser desorption/ionization, LDI) of intact molecules. With LDI MS a layer of sample placed on a metal surface is irradiated by a laser pulse to generate analyte ions. However, reduced sensitivity, extensive fragmentation and limited detection to molecules below 1000 Da hindered the application of LDI MS for the analysis of biomolecules. In 1985, Michael Karas and Franz Hillenkamp [45] inspected the correlation process between UV laser irradiation and organic molecules, successfully demonstrated the use of a matrix (consisting of a small organic molecule) to circumvent the limitation in LDI. A low concentration of the analyte was mixed with this matrix onto a metal plate sample probe (the target) and subjected to a pulsed laser beam. Therefore, embedding the sample into suitable matrix material, the laser light was strongly absorbed and thus not only higher intensities of the analyte ions were obtained but also sample fragmentation was reduced to a minimum. Thus, a substantial burst

of ions was produced with each laser pulse. This was the foundation of matrix-assisted laser desorption/ionization (MALDI). Later, Koichi Tanaka [39] using a mixture of cobalt nanoparticles and glycerol, what he called the "ultra fine metal plus liquid matrix method of laser desorption ionization", was able to demonstrate the application to a whole range of biological macromolecules. This led him to receive a part share with John Fenn on the 2002 Noble prize for chemistry [46]. However, the MALDI technique was demonstrated (and the name coined) in 1985 by Michael Karas, Doris Bachmann, and Franz Hillenkamp [45] but ionization of proteins by MALDI was not reported until 1988 [38], immediately after Tanaka's results were reported.

To perform MALDI, the analyte is mixed with a suitable matrix compound absorbing at the used laser wavelength on a metal sample plate (the target). Therefore, a co-crystallization process of both matrix and analyte material after evaporation of the solvent take place. The target plate have defined positions (spots) on which the analyte/matrix are deposited. Then, in the high vacuum area inside the mass spectrometer's ion source the surface of the analyte/matrix sample is then exposed to an intensive pulse of short waved laser irradiation. In MALDI a two-step ionization mechanism is the most accepted model of ionization. Thus, primary ions are formed after which ion-molecule reactions give rise to secondary ions. It is essential for the matrix to be of 1000 to 10 000 times molar excess, thus leading the analyte molecules being completely isolated from each other. The incorporation of the sample molecules into the lattice structure of the matrix is supposed to be precondition of the functioning of the laser desorption/ionization process. Moreover, there are two different theories for the mechanisms in which the matrix contributes to ion formation. The cluster model (it began as the "lucky survivors" model) supports the suggestion where the matrix is mainly a medium in the desorption/ablation reaction [47,48]. According to this model, the analytes are incorporated into the matrix crystals forming clusters (Figure 4). Their charge state is determined by the pH of the solution in which they have been deposited onto the MALDI target. During the laser irradiation the clusters are desorbed and the charged analyte is freed into the gas phase through sublimation of neutral matrix. In addition, some intra-cluster charge transfer and neutralization takes place.
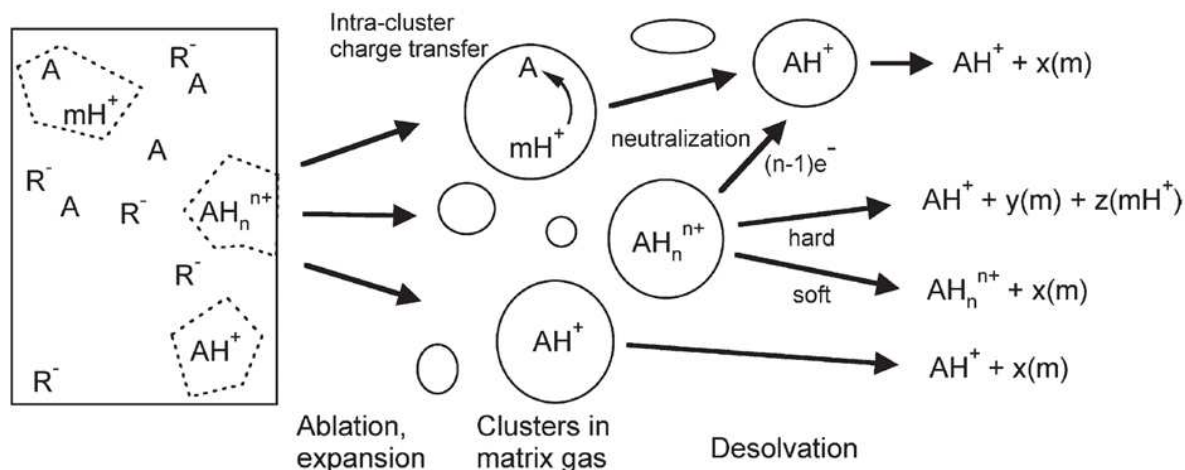
Figure 4. The "cluster model" of MALDI ionization.

Preformed ions, separated in the preparation solution, are contained in clusters ablated from the initial solid material. Some clusters contain a net excess of positive charge, others net negative (not shown). If analyte is already charged, here by protonation, cluster evaporation may free the ion. In other clusters charge may need to migrate from its initial location, e.g. on matrix, to the more favorable location on analyte (secondary reaction). For multiply charged analytes, hard and soft desolvation processes may lead to different free ions. Neutralization by electrons or counterions takes place to some degree, but is not complete. A = analyte, m = matrix, R^- = generic counter ion. From ref. [47].

In the pooling model [47,49-51], described in Figure 5, a more active role is proposed for the matrix in ionizing the analyte. This model explains the pooling of the energy in the matrix molecules. However, excitation migration and pooling have long been known in solid-state aromatics [52], but have been studied and demonstrated only in one MALDI matrix, the 2,5-dihydroxybenzoic acid (DHB) [53]. Here the laser pulse excites two matrix molecules that concentrate energy more to one of the two molecules. This combined excitation is enough to transfer the excitation stage of the matrix molecules to a higher level and ultimately produces matrix ions. Analyte ions are formed then from the primary matrix ions either by A) a proton, B) an electron or C) a cation transfer, producing singly charged analyte ions.

A) Proton transfer: $mH^+ + A \leftrightarrow m + AH^+$ and $(m-H)^- + A \leftrightarrow m + (A-H)^-$

B) Electron transfer: $m^+ + A \leftrightarrow m + A^+$ and $m^- + A \leftrightarrow m + A$

C) Cation transfer: e.g. $mNa^+ + A \leftrightarrow m + ANa^+$
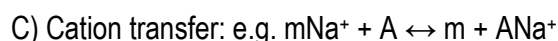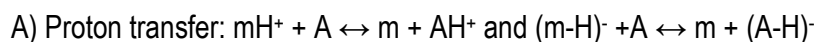
m= matrix, A=analyte
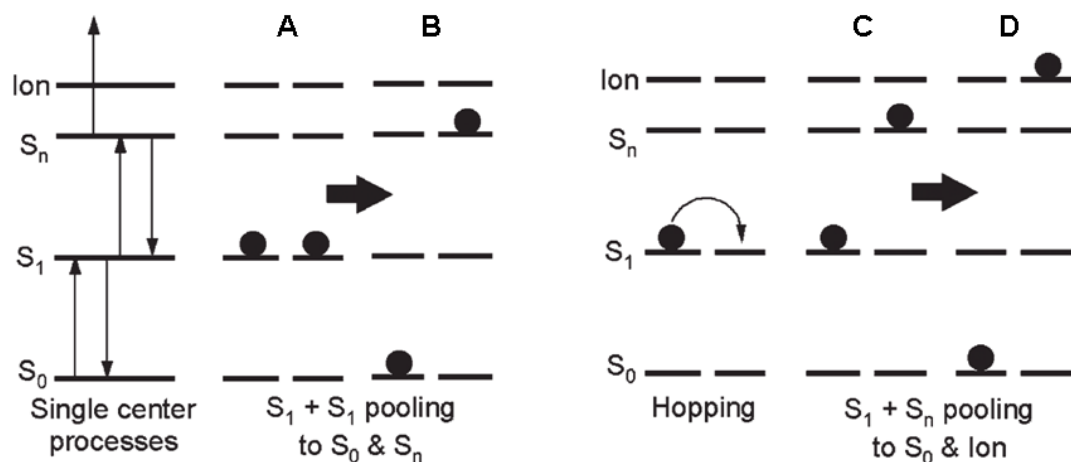
Figure 5. Unimolecular and biomolecular matrix processes included in the MALDI ionization model. Pooling reactions of matrix excited states are key steps in energy concentration and ionization. S1+ S1 pooling (A and B) takes place when the laser excites matrix molecules by one photon energy and two neighbouring molecules can transfer the energy entirely to another molecule, exciting the third molecule from S0 to Sn.
S1 + Sn pooling (C and D) takes place when a molecule from S1-S1 pooling (C) receives one photon energy from another molecule, thus, forming a matrix ion (D). After this primary ionization the secondary ionization of the analyte takes place through photon, electron or cation transfer (see text). S0 = electronic ground state; S1 = the first excited state; Sn = higher exited state, twice the energy of the S1. Adapted from ref. [47].

These ionization reactions take place in the desorbed matrix-analyte cloud just above the surface of the sample/matrix preparation onto the sample plate (target) (Figure 6). The ions are then extracted into the mass spectrometer for analysis. An electrode, which is mounted some millimeter apart opposite to the sample position, is used to generate an electrostatic field in the range of some kV/cm. Depending on the polarity, positively or negatively charged ions are accelerated from the sample surface towards the analyzer.
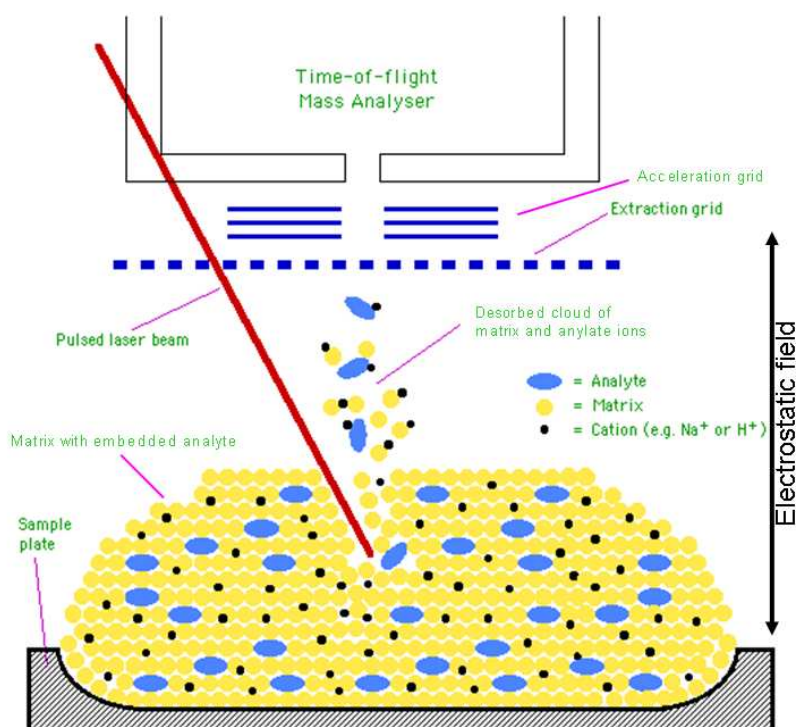
Figure 6. Schematic representation of the desorbed matrix-analyte cloud formation in the source of a MALDI-TOF mass spectrometer.

Laser devices employed in MALDI mass spectrometers are pulsed laser systems in the ultraviolet (UV) and infrared (IR) spectral ranges. Nowadays various laser types and wavelengths in the UV range are available. Nitrogen lasers are widely used in MALDI MS. In these devices, the laser medium is gaseous nitrogen (N_2), which is excited by means of an electrical discharge between two electrodes and produces a laser line. The most intensive laser line at wavelength of $\lambda= 337$ nm with pulse durations of 1-5 ns and operated at repetition rates of 10-50 Hz are used in MALDI MS. Also Excimer lasers which typically use a combination of a noble gas such as argon (Ar), krypton (Kr), or xenon (Xe) and a reactive gas such as fluorine (F) or chlorine (Cl) are used. Under electrical stimulation, a pseudo-molecule called excimer (from the words: excited and dimmer) is created which can only exist in an energized state and can give rise to laser light in the ultraviolet range. Excimer laser used in MALDI MS operate at wavelengths of $\lambda=193$ nm (ArF), at $\lambda=248$ nm (KrF) and at $\lambda=308$ nm (XeCl), using pulse durations of 2-20 ns and repetition rates of 1-20 Hz. The Nd:YAG lasers are another type of laser devices used in MALDI MS. Nd:YAG are solid-state lasers whose laser medium is a YAG crystal (yttrium aluminum garnet:$Y_3 Al_5 O_{12}$) doped with neodymium ions. The strongest excited laser line lies at a wavelength of $\lambda= 1064$ nm. This laser frequency can be doubled, tripled or quadrupled by non-linear optical processes. Thus, in addition to the fundamental wavelength at $\lambda=1064$ nm, wavelengths at $\lambda= 532$ nm, $\lambda=355$ nm and $\lambda=266$ nm arise. For MALDI mass spectrometers Nd:YAG lasers with pulse durations of about 5-15 ns and repetition rates up to 1000 Hz are used.

Concerning the MALDI MS in the IR range, Er:YAG (Yttrium-aluminum-garnet-crystals doped with erbium) solid-state laser are exploited at a wavelength of λ= 2.94 µm with pulse duration of 90 ns and repetition rates of 2-10 Htz.

The nitrogen lasers with a λ=337 nm and the Nd:YAG lasers at λ=355 nm are the most frequently implemented in modern MALDI TOF mass spectrometers. The great advantages of the Nd:YAG compared to N_2 lasers are the high repetition rate and the low energy fluctuations between individual laser pulses as well as the long life time.

In the mass spectrometer the laser beam is focused by a suitable ion optics onto the sample inside the ion source to a diameter of <= 150 µm. The laser power can be adjusted with an attenuator, i.e., a metal-coated mirror assembly. The target itself is mounted on a table movable in x and y direction allowing the systematic selection of sample positions (spots) for the measurements. Additionally, the sample can be observed by video camera to select a specific area of interest on the spot.

MALDI sources are normally coupled with TOF analyzers. Here the mass determination in the high vacuum area is performed by a very precise measurement of the period of time after acceleration process of the ions in the source and impact on the detector. An electro-static field accelerates ions formed during a short laser pulse inside the source to a kinetic energy of some keV. After leaving the source the ions pass a field-free drift region in which they are separated due to their m/z ratio. This takes place because at fixed kinetic energy, ions with different m/z values are accelerated in the ion source to different velocities. That is, knowing the acceleration voltage and the length of the drift region, the m/z ratio can be determined by measuring the flight time. Figure 7 illustrate the principle a TOF mass spectrometer measurement in linear mode. Ions formed by the laser pulse being of the same charge, but different m/z values are accelerated to different velocities. Large ions with high m/z values strike the detector at a later moment than small ions. Therefore, accelerating by means of a certain voltage U, the kinetic energy E_k of the ions is defined as:

$$E_k = \tfrac{1}{2} mv^2 = z \cdot e \cdot U \qquad\qquad (1)$$

However, the velocity v from the ion flight time t through the field-free region L of the flight tube results:

$$v = L/t$$

Therefore, replacing v in (1) by the formula given above:

$$\tfrac{1}{2} \cdot m \cdot (L/t)^2 = z \cdot e \cdot U \qquad\qquad (2)$$

Then arranging to m/z delivers:

$$m/z = (2eU/L^2) \cdot t^2 \qquad\qquad (3)$$

m = ion mass; v= velocity of the ion after acceleration; z= charge number; e= elementary charge
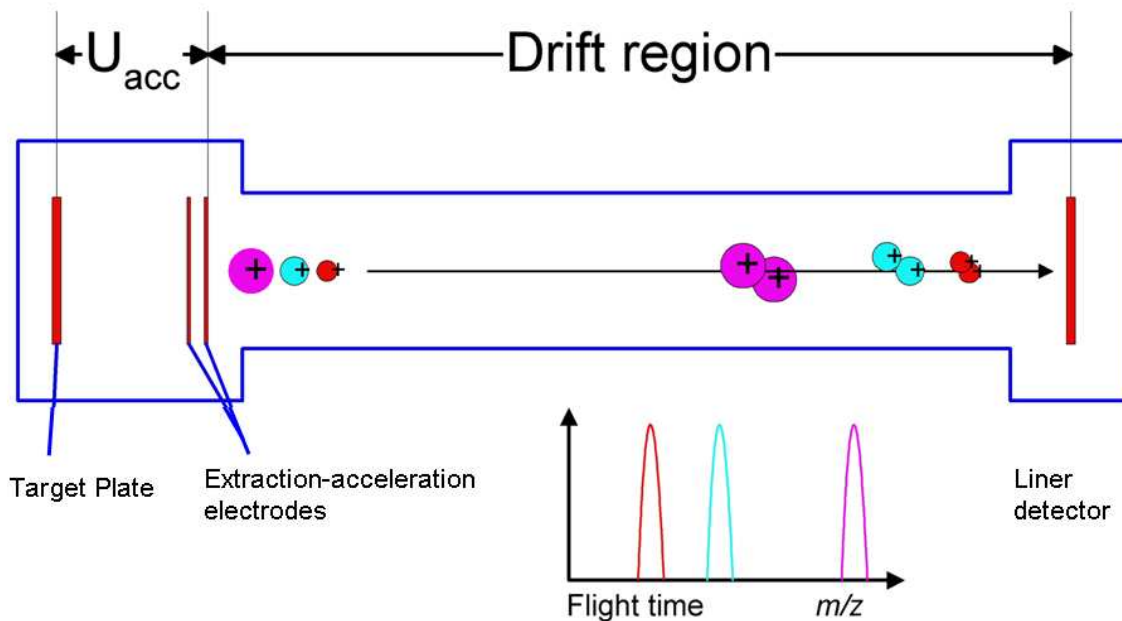
Figure 7. MALDI-TOF measurement in liner mode. Adapted from ref. [54].

Therefore, in a TOF instrument the relation of molecule mass and charge number (m/z) is proportional to the flight time. Thus, the related mass can be determined from the measured flight time. Typically, in a TOF instrument a drift region is 1-4m long and the flight times span from few to some hundreds microseconds. The calibration of the instrument is performed with molecules of known masses used as references.

Notably, not all the ions are desorbed and ionized at the same time and at the same place, thus, differences occur related to energy, location and time. Also repulsive electrical forces cause an uneven initial energy distribution of the ions. Therefore, ions of the same mass do not have the same kinetic energy after passing the acceleration field. Thus, ions leave the source with a certain energy spread, resulting in an impact on the detector at slightly different times. This affects in broadening the peak width and reducing the resolving power. Using a linear TOF instrument this relative energy spread can be compensated by increasing the acceleration voltages, however, with the disadvantage of shorter flight times which require a more precise flight time measurements. Moreover, with increasing size of the molecules the number of isotopic peaks is enhanced, i.e. more then one peak is detected for the molecular ion depending on the relative abundances of the isotope atoms. For large molecules such as peptides and proteins the number and relative intensities of isotopes also affect to greater extent the peak width in the spectrum. The peak width is an important parameter which indicates the resolving power achievable by the instrument. The resolving power ΔM defined the peak separation which allows distinguishing two ion species close to each other. There are two main ways to define the minimum peak separation ΔM in mass spectrometry: the valley and the peak width definition. The valley system

defines ΔM as the closest spacing of two peaks of equal intensity with the valley (lowest value of signal) between them less than a specified fraction of the peak height (10% or 50%). This is generally used for ion sector and triple quadrupoles instruments. In the peak width definition, the value of ΔM is the width of the peak measured at a specified fraction of the peak height (5%, 10% or 50%). Generally, for TOF analyzer the resolution is calculated using the full width at half maximum (FWHM), i.e. of the peak width at 50% of the peak height. The resolution R of a mass analyzer is defined as the ratio of the mass m and the resolving power ΔM.

The resolution R is calculated as follow:

$$R = m/\Delta M \tag{4}$$

High resolution MALDI TOF mass spectrometer are implemented with delayed ion extraction (DIE) and reflectron devices (Figure 8) to improve the mass resolution. With the DIE the electrical field between the sample surface and the related electrode is switched on delayed to the laser pulse [55]. Thus, ions with a higher initial velocity cover a greater distance from the sample surface and are exposed after switching on the electrical field to a less electrical energy than ions with a lower initial velocity. By using suitable delay time and field strength the influence of the energy distribution may be compensated at the position of the detector. Residual kinetic energy differences can be compensated by applying a reflectron. [56]. A reflectron is an electronic ion mirror which uses an oppositely polarized static electric field to reverse the trajectory of the ions entering it [56]. This field is interfaced directly to the field-free drift region. Indeed, ions with the same *m/z* but a higher kinetic energy penetrate deeper into the electric field of the reflectron. After reversion of trajectory, high kinetic energy ions reach the slower ones with the same *m/z* at a certain point of the drift region. Therefore, positioning the detector at appropriate distance (the focus of the reflectron), ions with the same *m/z* and different kinetic energies are detected simultaneously leading to improve the signals resolution. Applying a reflector, the influence of the energy distribution upon the flight time can be compensated at the location of the reflector detector.
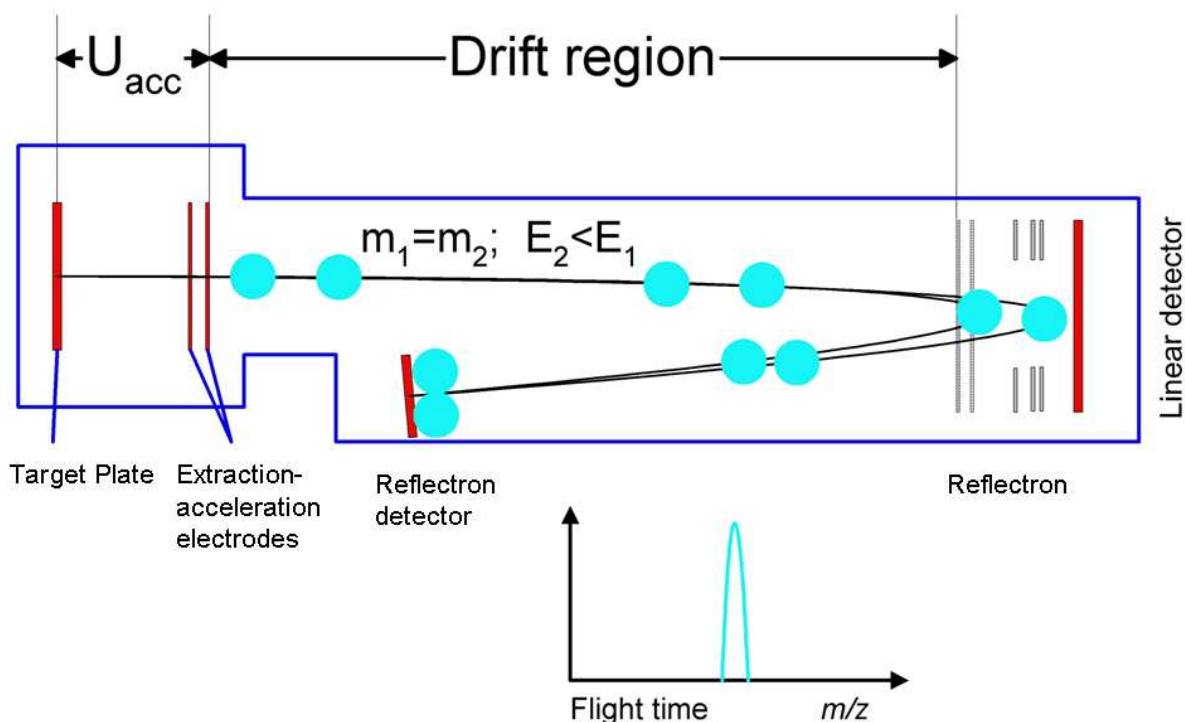
Figure 8. MALDI-TOF measurements in reflectron mode. Adapted ref. [54].

Resolutions of more than 20,000 (FWHM) within mass range up to 500 kDa and sensitivity of attomole (for peptides) can be obtained using commercial high resolution MALDI TOF instruments.

Tandem mass spectrometry

In mass spectrometric "soft" ionization processes the analyte posses a very little residual internal energy which results in the production of intact molecular ions and less or any analyte fragments. For structural elucidation purposes, subsequent analyte decomposition should be performed in order to obtain structurally diagnostic fragments. Tandem mass spectrometry also known as MS/MS employs two stage of mass analysis in order to selectively detect and generate the fragmentation of a particular analyte ions. Indeed, the targeted analyte is ionized and its specific ions are separated and selected in the first stage. The selected analyte primary ions are then decomposed to induce fragmentation, and secondary ions (fragments), characteristic of the targeted analyte, are detected and analyzed in the second stage. Therefore, analyte molecular ion and fragments can then be merged together to generate a tandem mass spectrum to obtain structural information regarding the intact molecule. Moreover, tandem mass spectrometry also enables specific compounds to be detected in complex mixtures on account of their specific and characteristic fragmentation patterns. There are two main categories of instruments that allow MS/MS experiments. The first category of instruments uses a sequence of mass spectrometer in space, while the second category uses spectrometers with ion storage capability to exploit a sequence of events in time. The first category is made up of instruments in which two or more mass spectrometers

elements are assembled series to each other. In tandem mass spectrometry in space, the separation elements are physically separated and distinct, although there is a physical connection between the elements to maintain high vacuum. These elements can be sectors, transmission quadrupole, or time-of-flight or hybrid combination of analyzers. Three quadrupole elements, two magnetic analyzer instruments or hybrids containing one magnetic and one quadrupole, one analyzing quadrupole and one time-of-flight are representative cases. The second category of MS/MS instruments comprises analyzer capable of storing ions, generally ion traps such as ICR or quadrupole ion trap. With tandem mass spectrometry in time, the selected ion can excited and caused to fragment during a selected period of time, and the fragment ions can be observed in a mass spectrum. By doing tandem mass spectrometry in time, the separation is accomplished with ions trapped in the same place, with multiple separation steps taking place over time. This process can be repeated to observe fragments of fragments over several generations. Trapping instruments can perform multiple steps of analysis, which is referred to as MS^n (which the n is the number of steps).

The ion activation step in the gas-phase is crucial in tandem mass spectrometry and ultimately defines what types of products result. Several ion activation techniques have been developed. These techniques result in different types of fragmentation and thus different information about the structure and composition of the molecule analyzed. Early experiments mainly studied metastable ion decompositions. In deed, molecular ions formed in the ion source having sufficient internal energy to fragment spontaneously in a field-free region between the source and the analyzer are classified as metastable ions. These fragments were initially identified as unknown broadened peaks in the earliest mass spectra and later explained as unimolecular decay products formed in the field-free regions of a mass spectrometer [57]. However, the ability to characterize precursor ions from metastable dissociations alone is fairly limited. Therefore, methods that cause ion activation increase internal energy and thus the number of precursors that dissociate providing structurally informative fragments were developed. Collisions between the precursor ion and a neutral target gas are accompanied by an increase in internal energy, which induces decomposition with improved probability of fragmentation as compared with metastable unimolecular dissociations. In collision-induced dissociation (CID), the selected molecular ions are usually accelerated by electrical potential to a higher kinetic energy and then allowed to collide with neutral gas molecules (helium, nitrogen or argon). In the collision some of the kinetic energy is converted into internal energy which results in bond breakage and the fragmentation of the molecular ion into smaller fragments. The acronyms collisional activated dissociation (CAD) and CID are both in use to describe this process. The electrical potential and the gas pressure can be used to get additional control on the fragmentation process, i.e. increase the degree of dissociation of higher mass ions and particularly stable ions. Therefore, the CID processes can be separated into one of two

categories based primarily on the translational energy of the precursor ion: low-energy collisions and high energy collisions. Indeed, low-energy collisions occur in the 1–100 eV range of collision energy, common in quadrupole and ion trap instruments, whereas, high energy collisions occur in the keV range, used in sector and TOF/TOF instruments.

Several relatively new techniques have been developed to complement conventional CID leading to extend the molecules and the kind of fragment obtained. These techniques includes: electron capture dissociation (ECD), blackbody infrared radiative dissociation (BIRD), electron transfer dissociation (ETD), electron-detachment dissociation (EDD), infrared multiphoton dissociation (IRMPD) and surface-induced dissociation (SID). Notably, fragmentation techniques can be implemented in a mass spectrometer using specific sequences or "routines" in order to selectively detect specific ions. In the following paragraphs a briefly description of the techniques relevant for this thesis are presented

MS/MS Routines

Using the so called routines, selective detection of ions that yield a given fragment or lose a given neutral is also possible with the appropriate tandem mass spectrometer. Tandem mass spectrometry enables a variety of experimental routines (scan mode or sequences) such as selected reaction monitoring (SRM), multiple reaction monitoring (MRM), precursor ion scan (PIS) and constant neutral loss scan (CNLS)

Selected reaction monitoring/multiple reaction monitoring (SRM)/(MRM)

The SRM exploits the unique capabilities of triple quadrupole (TQ) MS for quantitative analysis. In SRM, the first mass analyzer specifically selects predefined *m/z* values corresponding to the analyte ions (e. g. a peptide ion). A second mass analyzer then stabilizes the analyte ions while they collide with a gas, causing them to fragment (by CID). A third mass analyzer then sorts the fragment ions produced and select a specific fragment ion (of the peptide). Therefore, a complete ion spectrum is not acquired but only one or two chosen precursor/fragment ion pairs are monitored. The monitoring of more then one reaction (precursor/fragment ion pairs) is termed MRM. Several such transitions are monitored over time, yielding a set of chromatographic traces with the retention time and signal intensity for a specific transition as coordinates. These experiments are used to increase specificity of detection of known molecules [58].

Precursor ion scan (PIS)

The PIS provides the identity of al precursor ion that fragment to a preselected product ion. With this scan mode the first and second mass analyzers scan across the spectrum as partitioned by a user defined *m/z* value. The precursor ion spectrum is obtained by adjusting the second mass analyzer to transit a certain specified product ion, whereas the first mass analyzer is scanning over a certain mass range to transit only those precursor ions that generate the specified chosen product in collision cell. This experiment is used to detect specific motifs within unknown molecules [59,60].

Constant neutral loss scan (CNLS)

With the CNLS both mass analyzers are scanned simultaneously. Indeed, the first mass analyzer scans all the masses, whereas, the second mass analyzer also scans, but at a specified mass offset from the first mass analyzer [61]. This offset corresponds to a neutral loss that is commonly observed for the class of compounds under analysis. In a constant neutral loss scan, all precursors that undergo the loss of a specified common neutral are monitored. Similar to the precursor-ion scan, this technique is also useful in the selective identification of closely related class of compounds in a mixture.

MALDI tandem MS

MALDI-MS/MS with PSD

Metastable are molecular ions formed in the ion source having sufficient internal energy to fragment spontaneously in a field-free region between the source and the analyzer. The metastable ions dissociation process to produce fragment ions is defined post-source decay (PSD). The MALDI-PSD-MS/MS introduced by Kaufmann *et al.* is a technique that has been widely used for several years [55,62,63] for the characterization of proteins and peptide with high sensitivity [64]. In a MALDI TOF instrument, if an ion fragments within the source, the precursor and product ions acquire the same kinetic energy and thus can be differentiated by their velocities and flight times, which are related to their respective masses (*m/z*). When a precursor ion enters the flight tube of a TOF analyzer, PSD of ions (i. e. metastable ion dissociation) yields decomposition products reaching a linear detector with the same velocity as the original precursor ion. Hence, in TOF MS, ions that decompose during their flight through the drift tube are usually detected at the precursor's m/z. However, these products can be distinguished by the application of a reflectron [62,63]. All ions enter the reversing/retarding field of the reflectron, but the lighter ones penetrate less than the heavier ones. Therefore, fragment ions exit the reflectron earlier than their precursors and have characteristic flight times based on their masses [62]. A MALDI-PSD-

MS/MS instrument includes a precursor ion selection step for the ions undergoing PSD, performed with a so called timed ion selector (TIS) (Figure 9). Technically, the TIS is an "ion gate" which is composed of a series of wires with alternating voltages (e. g. ±1000 V) which, when switched on, prevents the transmission of any ions. When the gate is switched off, ions may be transmitted into the mass analyzer. Therefore, the ion gate is tuned to select an ion of a specific m/z by switching off the gate at the estimated time of arrival of that ion. With the TIS, the precursor is selected and detected with the instrument operated in the linear mode. Subsequently, the fragment ions are detected by switching the instrument to the reflectron mode. Reducing the reflectron voltage whilst maintaining a constant source voltage means that precursor ions become too energetic to be reflected, but fragment ions can be reflected to the detector. If the reflectron voltage is reduced to half of the voltage than when the precursor ion was focused, an ion of half the mass of the precursor will be detected. Thus, by scanning the reflectron voltage, a full range of ions from low to high mass can be observed. This process is automated and the instrument is able to scan numerous reflectron voltages and acquire an in-focus section of each spectrum from the mass of the selected precursor ion down to the smallest ion species observable.

The above described is the classical method of acquiring PSD, which requires 10–60 min per spectrum. An addition drawback of this technique is that in some instrumental combination, i.e. with curved-field reflectron [65], demonstrated reduced fragmentation efficiency.

MALDI-CID-MS/MS

The first improvement to overcome the described disadvantages of PSD relied on increasing the fragmentation yield. PSD products in MALDI-TOF instruments can be increased by either performing the experiments under low-vacuum conditions [62] and therefore causing collisions with residual gas to occur in the flight tube. Alternatively, a collision cell is placed immediately after the ion source in a TOF instrument [66]. Some tandem TOF instruments have a high-energy collision cell placed between two TOF mass analyzers. High-energy CID generally employs helium as the target gas, but collision yield may be increased by using a heavier gas, such as Ar or Xe [67,68]. Regarding the sample preparation, in PSD-CID the high efficacy of "hot" MALDI matrices, in particular of CHCA, are exploited to enhance metastable fragmentation of MALDI generated ions. MALDI-PSD has been applied to synthetic polymers, oligonucleotides, and to protein analysis. In particular, the application of MALDI PSD with CID-MS/MS to protein and peptides produce fragment ion spectra with backbone, internal and the side chain ions [66,68–71] .

MALDI-LIFT-TOF/TOF

New developments in MALDI-TOF/TOF instruments are characterized by a co-linear arrangement of two TOF mass analyzers, each equipped with an ion source (source 1 and source 2) that allow acceleration and focusing of the ions. In a particular solution (which is the instrument used for the experimental work of this doctoral thesis) termed MALDI LIFT-TOF/TOF (Figure 9), a combination of low initial accelerating voltage (in source 1), used to promote PSD fragmentation, with a post-acceleration (in source 2) allow to mass analyze precursor and fragment in one spectrum [72].
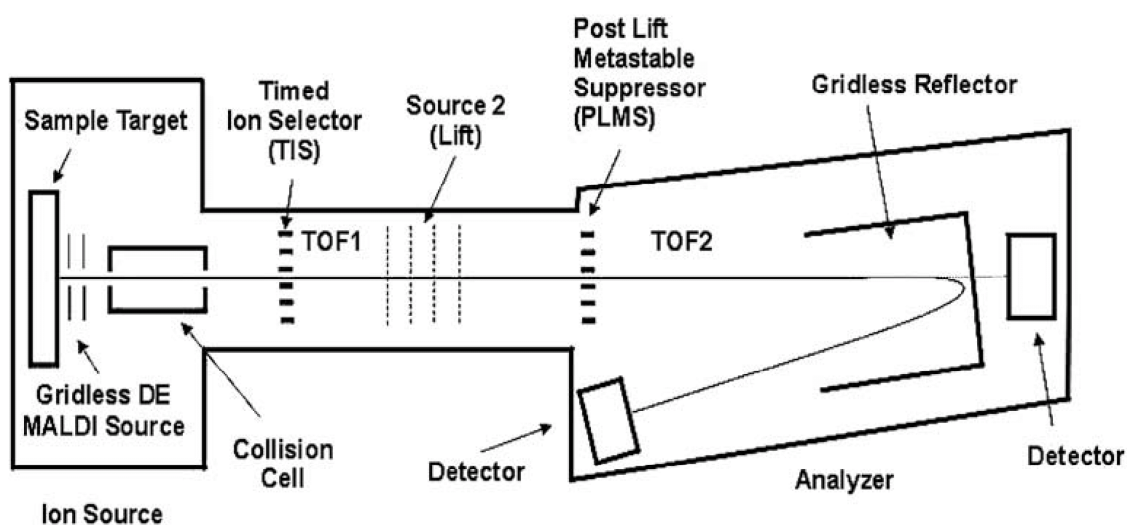
Figure 9. Schematic diagram of the LIFT-TOF/TOF mass spectrometer. TOF1 ranges from the MALDI ion source to the LIFT cell, TOF2 from the second accelerator stage in the LIFT cell to the reflectron. From ref.[72].

The basic idea in this instrumental solution relied on the subsequent acceleration, lift, of a selected precursor ion together with its fragments in the source 2. With this instrument the laser-induced dissociation (LID) is the mainly process to obtain PSD. Indeed, after the initial laser induced ion desorption process and expansion of the ion-matrix in high-vacuum MALDI systems there is still a substantial amount of internal energy left, which can lead to metastable fragmentation of the analyte ions. Therefore, LID fragmentation does not require collision gas but occurs after the acceleration stage in the ion source before the reflectron [73]. Moreover, acquisition conditions must be adapted to promote LID and generate high fragment ion yields. This is done by increasing laser fluence over the threshold, which also provides a larger number of precursor ions per shot. In addition, a low acceleration voltage in the source 1 leads to long flight time promoting fragmentation. As for traditional PSD MALDI, a TIS is used to select precursor and fragment ions by switching off the gate at the appropriate time. When precursor and fragment ions enter in the source 2, called "LIFT", they are post-accelerated, "lifted", of some keV. Here, precursor and fragment ions are subjected to a process similar to the delayed ion

extraction (DIE) in a regular MALDI ion source. Therefore, the precursor ions are fully accelerated, i. e. the sum of the kinetic energy resulted from the source 1 and LIFT acceleration, whereas, the fragment posses only a fraction of energy imparted in the source 2 proportional to their masses. An additional ion gate is situated between the LIFT device and reflector, the "post lift metastable suppressor" (PLMS), which deflects the remaining intact precursor ions and prevents further fragment ion formation after post-acceleration. Therefore, after reversion by the reflectron, the ions are time-focused onto the detector to generate a LID-MS/MS spectrum. Practically, a MS/MS spectrum is acquired in two steps: first the precursor then all fragment ions. With this instrument the entire procedure to acquire a tandem mass spectrum takes only few minutes. This system allows high resolution, mass accuracy and increased sensitivity. In case of peptides analysis MALDI-MS/MS spectra are rich in backbone ions a-, b-, y- and i-, internal ions and the side chain d- and w-ions, very useful for identification and characterization purposes [72].

MALDI-CID-LIFT-MS/MS

The MALDI LIFT-TOF/TOF instrument contains a collision cell for high-energy CID (Figure 9), with argon as collision gas introduced under computer control, to increase the source pressure to 6×10^{-6} mbar. This is useful in combination with LIFT-MS/MS to increase the fragmentation of particularly stable molecules. For peptide fragmentation, the CID-MS/MS modus generate high energy w-type ions via side-chain fragmentation which can allows to distinguish isobaric amino acid residues [70].

Reflector ISD mode for fragmenting intact proteins

The in-source decay (ISD) is a very fast fragmentation process occurring in the source region [74] during the MALDI process before the ions are accelerated out of the DIE [69]. This prevents the dissipation of the excitation energy across the peptide that would result in the preferred cleavage of the N-C_α bonds on the peptide backbone. In reflector mode (re-ISD), monoisotopic sequence fragment ions can be detected, which are predominantly c-ions and to a lower extent also y- and a-ions, well suited for sequencing the N-terminal region of intact proteins.

Chapter I
Development of matrix layer: an improved MALDI MS sample preparation method for proteomic study

Proteomics

The last century saw massive progress in knowledge of processes underlying the genetic basis of our existence. These progresses ranging from the definition of structure and sequence of DNA to the identification of specific disease-associated genetic abnormalities, and their clinical exploitation as the basis of trials of gene therapy. In particular, large-scale DNA sequencing has transformed biomedical research in a short span of time. With the sequence of human genome, it is now clear that a "global approach" is required to address a comprehensive understanding of complex biological processes. The life of a cell is a dynamic process in which it is constantly reacting to its environment. If a disease-inducing element is introduced, it may change how much and when a gene product is made, the type and extent of post-translational modifications that occur, and how these events are related together and with other genes. Because the study of this dynamic has the potential to reveal new targets for drug intervention in disease processes, emphasis is now being placed on understanding how and when genome-encoded events occur, e.g., at protein translation level. Moreover, understanding what relationship of non-genome-encoded events have to particular physiological states, e.g., posttranslational modifications of proteins, interactions between proteins, nucleic acids, lipids, carbohydrates, and combinations thereof. Proteomics is a field that promises to bridge the gap between genome sequence and cellular behaviour. It aims to study the dynamic protein products of the genome and their interactions, rather than focusing on the simple relatively static DNA of a cell.

The word "proteome" was first coined by Valerie Wasinger in 1995 [75] and Marc Wilkins in 1996 [76] do to the analogy with the term genome, while working on characterization of gene products from *Mycoplasma genitalium*. Indeed, the proteome is the protein content of biological system, cell, organ or organism. Later, the term "proteomics" was introduced to make an analogy with genomics, the study of genomes [77]. Proteomics is the large-scale study of proteins as well as their structures and functions by biochemical methods. From a technological standpoint, the essence of proteomics is protein identification and characterization. In this sense proteomics already dates back to the late 1970s when researchers started to build databases of proteins using the then newly developed technique of two-dimensional gel electrophoresis (2-DE) [78]. The 2-DE technique is capable of resolving total protein

extracts from cells into about 10,000 individual protein spots. This resulted in extensive cataloguing of spots from two-dimensional gels to create databases of all expressed proteins. However, even when such gels could be run reproducibly between laboratories, determining the identity of the proteins was difficult because of a lack of sensitive and rapid analytical methods for protein characterization (such as the polymerase chain reaction and the automated sequencer for DNA analysis). Although Edman degradation is a powerful tool for protein sequencing [79], is mainly deficient in meeting the throughput required by proteomics. In the 1990s, biological mass spectrometry emerged as a powerful analytical method that removed most of the limitations of protein analysis [80]. The most significant breakthrough in proteomics has been the mass spectrometric identification of gel-separated proteins, which extends analysis far beyond the mere display of proteins. Mass spectrometry has essentially replaced the classical technique of Edman degradation even in traditional protein chemistry, because it is much more sensitive, can deal with protein mixtures and offers much higher throughput. This development, coupled with the availability of the entire human coding sequence in public databases, marks the beginning of a new era. Today, the term proteomics covers much of the functional analysis of gene products or "functional genomics", including large-scale identification or localization as well as posttranslational modification studies of proteins and interaction studies. Proteomics in the contest of a "global approach" is used to paint a "global" picture to allow cell biologists building a complex map of cell function by discovering how changes in one signalling pathway affect other pathways, or how proteins within one signalling pathway interact with each other. The "global picture" also will allows medical researchers to look at the multiplicity of factors involved in diseases, very few of which are caused by a single gene. As for genomics, to put forward proteomics continuously needed technical improvements for increase sensitivity and specific analytical method for large-scale and highly automated analyses. Independently of the specific aims, the main steps of a proteomic project are: proteins preparation, protein separation, protein identification and characterization.

Protein preparation methods

One of the most crucial steps in proteomics is obtaining and handling the protein sample. Out of the entire complement of the genome of about 23,000 genes[81], a given cell line may express about 10,000 genes and an even higher number is expressed in tissues. Furthermore, the dynamic range of abundance of proteins in biological samples can be as high as 10^6 [82]. There are a number of different procedures that can be undertaken during protein sample preparation for proteomic analysis, which can influence the quality of the results and number of protein identifications. After protein extraction, further sample treatment is often required to reduce the sample complexity prior to MS analysis. Depending on the sample origin and the biological questions, these might include subcellular fractionation, protein

fractionation, protein digestion, peptide fractionation and desalting. Moreover, additional purification/enrichment steps are required when specific classes of proteins are under investigation e. g. posttranslational modified proteins or low abundant proteins. Indeed, additional sample handling steps to selectively enrich phosphorylated proteins or peptides are discussed more detailed in chapter III. The most common technique used to separate proteins is two-dimensional gel electrophoresis (2-DE), where proteins are firstly separated according to their isoelectric point (pI) and secondly by their apparent molecular weight (MW) [83,84]. With this approach a representation of the sample is visualized as a distribution of spots. The resulting protein spots are subjected to enzymatic digestion prior to protein identification by MS [80,85]. Alternatively, protein sample separations in only one dimension (SDS-PAGE) can be applied, i.e. separating the proteins solely according to their MW [86,87]. The advantages of SDS-PAGE as a preparation/separation method are that virtually all proteins are soluble in SDS, the range of relative molecular mass from 10,000 to 300,000 is readily covered, and extremely acidic and basic proteins are easily visualized. In this case, the bands containing proteins are enzymatic digested to peptides, and analyzed by tandem MS. To reduce complexity, additional liquid chromatography (LC) separations can be applied to the peptide samples prior to tandem MS analysis. Moreover, multidimensional LC or Isoelectric focusing (IEF) separations are frequently exploited to fractionate the peptide samples. Final peptide separations are mostly performed using reverse phase (RP) chromatography. Indeed, increasing concentration of the organic solvent result in sequentially elution of peptides from the separation column with respect to their increasing hydrophobicity.

Two-dimensional gel electrophoresis (2-DE)

The 2-DE is a powerful and widely used method for the analysis of complex protein mixtures extracted from cells, tissues, or other biological samples. This technique separates proteins according to two independent properties in two discrete steps. The first-dimension step, the IEF, separates proteins according to their isoelectric points (pI). The second-dimension step, the SDS-PAGE separates proteins according to their MW [86,87]. Each spot on the resulting 2-DE gel potentially corresponds to a single protein species in the sample. Thus, thousands of different proteins can be separated, and information such as the protein pI, the MW and the amount of each protein can be obtained.

In the original 2-DE technique, introduced by O'Farrell in 1975 [78], the first-dimension separation was performed in carrier-ampholyte-containing polyacrylamide gels cast in narrow tubes. However, its application has become significant as a result of a number of developments. Indeed, the introduction of immobilized pH gradients and Immobiline™ reagents [88] brought superior resolution and reproducibility to first-dimension IEF. Based on this concept, Görg and colleagues [89,90] developed the currently employed 2-DE technique, where carrier-ampholyte-generated pH gradients have been replaced with

immobilized pH gradients, and tube gels replaced with gels supported by a plastic backing: the precast IPG strips (Immobiline DryStrip gels) commercially available (from Biorad, GE Healthcare and other). Moreover, the 2-D difference in gel electrophoresis (DIGE) [91,92] which offers a method for controlling system variations and allowing biological changes in protein expression to be identified with statistical confidence. Also development in automation of steps after 2-DE, such as gel image analysis, spot picking, spot digestion, and sample preparation for mass spectrometry, have allowed a significant increase in the throughput of protein analysis and identification. New mass spectrometry techniques developed and data about entire genomes of a number of organisms now available allow rapid identification and characterization of very small quantities of peptides and proteins separated by 2-DE. The World Wide Web provides simple, direct access to spot-pattern databases for the comparison of electrophoresis results and genome sequence databases for assignment of sequence information.

A large and growing application of 2-DE is within the field of proteomics [76,93]. Indeed, the analysis involves the systematic separation, identification, and quantization of many proteins simultaneously from a single sample. Notably, 2-DE is used in this field due to its unparalleled ability to separate thousands of proteins simultaneously. The technique is also unique in its ability to detect post- and co-translational modifications, which cannot be predicted from the genome sequence. In addition, once 2-DE separated, proteins can also be blot onto membrane and detected using antibody leading to obtain additional specific information.

Mass spectrometry for Protein analysis

Mass spectrometry is applied in protein analysis using two different approaches: "top-down" and the "bottom-up"[94]. The top-down approach consisted in analyzing intact proteins. Indeed, the mass spectrometer is used to determine the proteins mass, whereas, protein sequence and posttranslational modification information are obtained due to protein fragmentation by tandem mass spectrometry [95-97], this approach has been applied using both ESI or MALDI instruments to characterize proteins. Recently, in two studies detailed characterization of the amino acid composition leading to 100 % sequence coverage of 15.2 kDa [98] and 13.6 kDa [99] proteins was performed, showing the power of this approach. Although this technique is beginning to become more widely used, the most common method for protein and proteome analysis is by using the bottom-up approach.

Bottom-up proteomics is a common approach to identify proteins, characterize their amino acid sequences and post-translational modifications by proteolytic digestion of proteins prior to mass spectrometry analysis [100,101]. Indeed, crude total protein extracts are directly digested using endopeptidases specific enzymes (proteases), such as trypsin, or chemical reagents which are also used in combination [102]. The proteolytic digestion is followed by one or more dimensional separation of

the resulting peptides mixture by liquid chromatography coupled to tandem mass spectrometry, a technique also known as shotgun proteomics [103]. Indeed, with a tandem mass spectrometer, the single peptides are selected and fragmented to generate tandem mass spectra. By comparing the tandem mass spectra of the proteolytic peptides with those predicted from a sequence database, peptides can be identified and multiple peptide identifications assembled into protein identification. Here, ESI tandem MS is typically employed on either hybrid quadrupole-TOF (Q-TOF) or various types of ion-trap mass spectrometers.

Alternatively, the proteins may first be separated by gel electrophoresis (SDS-PAGE or 2-DE) resulting in one or a few proteins in each proteolytic digest. Generally, trypsin is the protease utilized because of its high cleavage specificity and stablity under a wide variety of conditions. The mass spectrum of the peptide mixture resulting from a proteolytic digest of a protein produced a characteristic peptides profile providing a fingerprint of great specificity. So specific, that it is often possible to identify the protein from this information alone. This characteristic peptide profile is termed: "peptide mass fingerprint" (PMF) of the protein [80,104-109]. The peptide signals in the PMF are then selected and tandem mass spectrometry is used to obtain peptide fragments spectra. Peptides fragment along the amino acid backbone give sequence information and produced a characteristic peptides fragment profile termed: "peptide fragment fingerprint" (PFF) [110-112].

The proteolytic digests can be analyzed either by ESI or MALDI tandem MS with different types of mass spectrometers. MALDI-TOF/TOF tandem mass spectrometry is often the preferred technology because it allows high sample throughput, and several hundred samples can be analyzed in a single experiment. High confident identification and characterization of proteins is performed by correlating the PMF and PFF information with sequence databases, which has become the *de facto* the method for protein identification nowadays. Several algorithms and software have been developed to match generated PMF and PFF spectra against "in-silico-derived" spectra from protein and amino acid sequences in proteomics and genomics databases [113-115].

Peptide fragmentation

Regarding the peptide fragmentation, peptides yield a wide array of product ions depending on the quantity of vibrational energy they posses and the time-window allowed for dissociation. The ion types formed and the abundance pattern observed are influenced by the peptide sequence, the ionization technique, the charge state, the collisional energy (if any), the method of activation as well as the type of the analyzer. There are three different types of bonds that can fragment along the amino acid backbone: the NH-CH, CH-CO, and CO-NH bonds. One of the theories proposes peptide fragmentation through the "mobile proton" [116,117]. In this model any of the amide bonds in the peptide backbone can be protonated by a mobile proton making it more susceptible to fragmentation. The energy for this cleavage

comes from activation method used. Each bond breakage gives rise to two species, one neutral and the other one charged, and only the charged species is monitored by the mass spectrometer. The charge can stay on either of the two fragments depending on the chemistry and relative proton affinity of the two species. Hence, there are six possible fragment ions for each amino acid residue and these are labeled according to the accepted nomenclature in the diagram of Figure 10. The nomenclature for peptide fragments was first proposed by Roepstorff and Fohlmann [118] and later modified by Johnson *et al.* [119]. Indeed, the a, b, and c" ions having the charge retained on the N-terminal fragment, and the x, y", and z ions having the charge retained on the C-terminal fragment. The numbering indicates which peptide bond is cleaved counting from the N- and C-terminus, respectively. A subscript indicates the number of residues in the fragment. The number of hydrogens transferred to or lost from the fragment is indicated with apostrophes to the right and the left of the letter, respectively.

The cleavage along the backbone at the CO-NH bonds give rise to the b and/or the y" ions. Thus, the mass difference between two adjacent b ions, or y"; ions, is indicative of a particular amino acid residue.
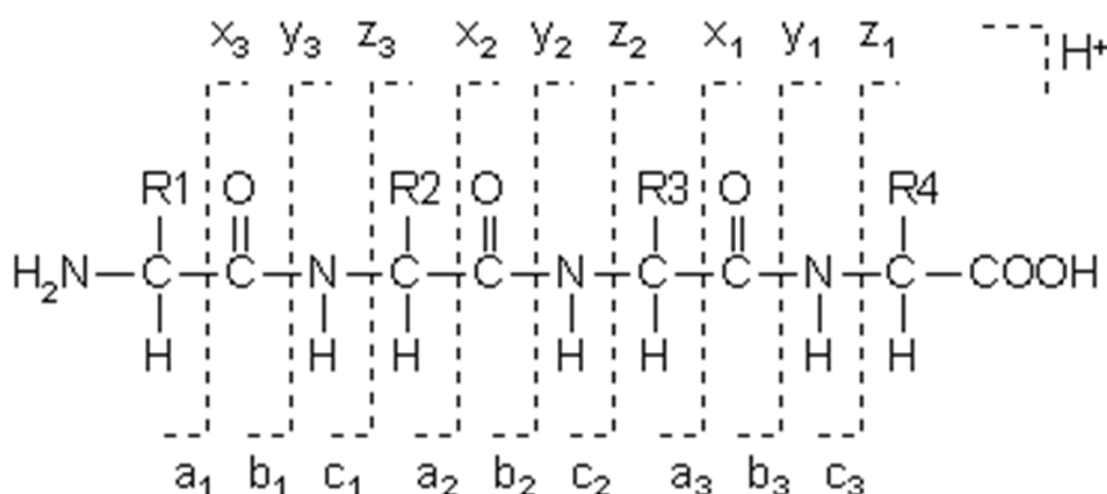
Figure 10. Diagram of the fragments formed through bond cleavages along the backbone of protonated linear peptides.

The very low mass (<150 *m/z*) region of CID spectra often contains valuable information in the form of ions generated from the individual amino acids present in the peptide, the immonium ions. Typically, high-energy CID mass spectra are richer in these ions. The immonium (HN=CH-R; labeled in a spectrum with the single letter code of the particular amino acid), are ions with a mass of 27 *m/z* lower than that of the amino acid residue (amino acid mass -CO+H). Each amino acid residue leads to a diagnostic immonium ion, with the exception of the two pairs leucine (L) and iso-leucine (I), and lysine (K) and glutamine (Q), which produce immonium ions with the same *m/z* ratio, i.e. 86 *m/z* for I and L,

101 *m/z* for K and Q. The immonium ions are useful for detecting and confirming many of the amino acid residues in a peptide, although no information regarding the position of these amino acid residues in the peptide sequence can be ascertained.

Internal fragment are ions which contain neither the N-terminus nor the C-terminus of the peptide, and their mass usually corresponds to a structure with a y-type cleavage at one end and a b-type cleavage at the other. Internal fragment ions are often labeled as **y**$_i$ **b**$_j$, in keeping with the sequence ion nomenclature and numbering. Low-energy CID spectra frequently contain abundant internal fragment ions [120], which are the result of multiple collisions that increase the probability of cleavage of two amide bonds.

Trypsin is a useful protease in mass spectrometry study as it generates peptides with a C-terminal arginine or lysine with masses in the preferred mass range for effective fragmentation [121]. This C-terminal positioning of basic residue has consequences for fragment ion formation. According to the "mobile proton" model, dissociation upon excitation is initiated by a proton that weakens an amide bond in the peptide backbone [116,117,120,122,123]. The proton affinity/gas phase basicity of the two conjugate fragments will then dictate which fragment will inherit the amide-breaking proton, leading to the formation of b- or y-ions, respectively [117]. In MALDI-MS/MS of singly charged tryptic peptides, fragmentation results in complex spectra containing not only b- and y-ions but also some a- and immonium ions, internal fragments and ions resulting from neutral loss of ammonia or water [68,124,125]. Thus, all these fragment ions are used in typical database search strategies for protein and peptides identification.

MALDI MS Sample Preparation

The matrix

The matrix plays an essential role in the MALDI ionization process [47-51]. Therefore, the choice of matrix for the sample preparation is crucial for success in MALDI MS experiments. In this regard the MALDI matrix must meet a number of generally requirements: (1) be able to embed and isolate analytes molecules (e.g., by co-crystallization), (2) be able to absorb at the laser wavelength, (3) be able to cause desorption and promote analyte ionization upon laser irradiation, (4) be vacuum stable. Initially, Karas and Hillenkamp reported that nicotinic acid was a suitable matrix for the desorption and ionization of intact protein molecules using 266-nm laser irradiation [38]. Since then, a large number of compounds have been screened as potential matrices, and a number of new matrix materials of substantial utility have been reported.
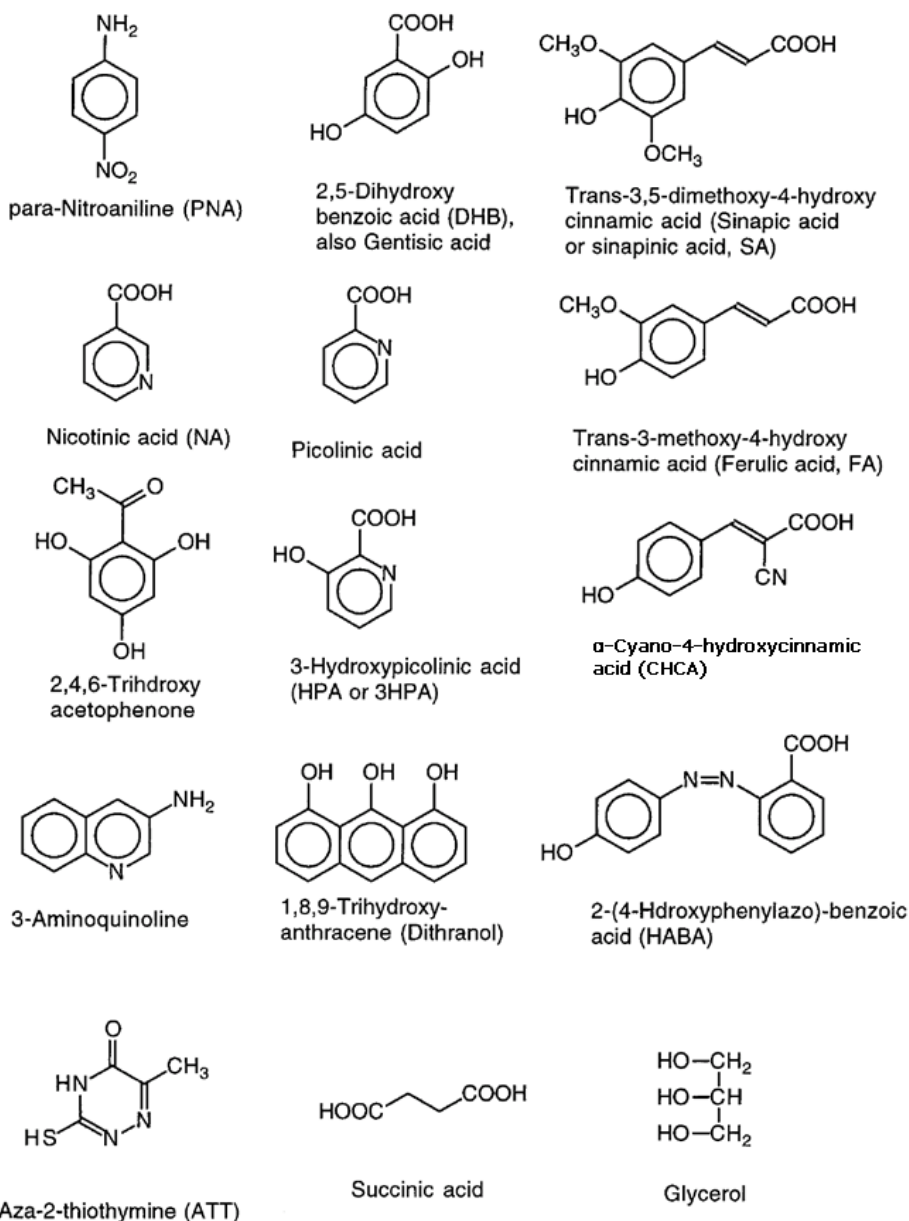
Figure 11. Structures, chemical names, trivial names, and abbreviations of frequently used MALDI matrices. Adapted from ref. [47].

Figure 11 gives the structures of frequently used MALDI matrices, their chemical names, trivial names, and abbreviations. Derivatives of benzoic acid, cinnamic acid, and related aromatic compounds were recognized early on as good MALDI matrices for proteins [126]. Matrices such as 2,5-dihydroxybenzoic acid (DHB) and several cinnamic acid derivatives, including ferulic, caffeic, and sinapinic acids (SA), have made MALDI analysis possible at longer wavelengths including the following: 308, 337, and 355 nm, accessible with excimer, nitrogen, and Nd: YAG lasers, respectively [126-129]. The α-cyano-4-hydroxycinnamic acid (CHCA) has been shown to be an effective matrix for MALDI analysis of both peptides and glycopeptides [130]. Juhasz *et al.* introduced 2-(4-hydroxyphenyazo)-benzoic acid (HABA) as

a matrix for peptides, proteins, and glycoproteins up to 250 kDa [131]. For oligonucleotides, 3-hydroxypicolinic acid (3-HPA) [132,133], glycerol [134], and succinic acid [135] are used as matrices in the UV and in the IR, respectively. However, many common UV matrices work well in the infrared too [136] and not all matrices are carboxylic acids. For example, 2,4,6-trihydroxyacetophenone (THAP) [137] has been shown to be useful for studying oligonucleotides, [138], [139]. [140]

In one of the earliest MALDI experiments Tanaka *et al.* used very fine cobalt particles mixed with liquid glycerol [39] to ionize large biomolecules. After a long lapse, liquid matrices and liquid/solid two-phase matrices (also called surface-assisted laser desorption/ionization, SALDI, to differentiate it from the MALDI using solid matrices) were proposed in attempt to circumvent some of the limitations associated with solid crystalline matrices [141-143]. For instance, some solid matrices are characterized by the production of non-homogeneous analyte/matrix crystals. Indeed, signal of the analyte are obtained only from restricted areas of the sample preparation resulting in the formation of the "sweet spot" (also known as "hot spot") phenomenon. Liquid matrices can continuously refresh their surfaces, therefore, partially circumvent the "sweet spot" phenomenon, chemical background in the low mass range as well as the co-crystallization and solubility requirements known for solid matrices. Moreover, if an absorbing solid material, e.g. a MALDI matrix, is added to a non-absorbing liquid, the advantages of liquid matrices are retained and the limitations of solid matrix reduced [144-147] [148-151] [142,152,153].

Compounds that are not easily protonated can be cationized instead, often by adding a small quantity of alkali cations, copper or silver salts to the sample preparation. Indeed, transition metals such as Cu or Ag are expected to bind to the phenylic π-system with electrons from the metal d-orbitals. [154,155].

The choice of matrix is also important for the control of fragmentation. A small number of matrices have been studied with regard to their propensity to induce fragmentation and, thereby, are classified as "hot" or "cold." Karas *et al.* [156,157] have found that, for protonated glycoproteins, post-source decay decreases in the order SA > DHB > HPA. This order agrees with Spengler et al. [62], who also described SA as a "hotter" matrix than DHB. Moreover, in MALDI mass spectrometry of oligodeoxynucleotides, DHB has also been found to induce more fragmentation than 3-HPA [158]. One proposed explanation for the "hot" and "cold" nature of matrices is simply the temperature at which they sublime, hence, desorbs the analyte [159]. Therefore, the selection of the matrix can be used to control degree the analyte fragmentation permitting additional control over the MALDI MS experiment.

To reduce unwanted salt adduct formation and to increase the homogeneity of the analyte/matrix preparation a large variety of matrix additives are used. Most of these matrix "dopant" are salt of mono- and bi- basic acids which function by improving co-crystallization and by sequestering excess salts. Difficulties with MALDI analyses can stem from many sources, not only ionization problems. Incorporation of analyte into a solid matrix and formation of suitable matrix crystallites are among the

most frequent difficulties with MALDI analyses. However, knowledge about ion formation pathways can now contribute to rational matrix selection. A fundamental consideration is the type of analyte ion expected or desired. Based on prior knowledge of the acidity or basicity of the analyte, secondary protonation or deprotonation reactions can be enhanced by selecting the appropriate matrix. This may be the most important and yet easiest type of optimization that can be carried out for MALDI MS. If cationized analytes are expected, the matrix should be chosen so as not to compete with analyte for the selected cations in the condensed or in the gas phase. Therefore, matrices that are good complexing agents, such as those with ortho-hydroxy carbonyl units, should be avoided. Divalent cations are preferable in general, due to higher electrostatic binding energies.

However, appropriate matrices for a particular analytical problem are often found only after screening of a large number of candidate compounds [133,140].

MALDI MS sample preparation methods

An optimal sample preparation method should provide a broad mass range for detection of the analyte ions, high sensitivity and mass accuracy as well as reproducibility of sample/matrix crystal preparation to obtain reliable mass spectra [160]. Moreover, the feasibility of automated spectra acquisition and data analysis are mandatory for the high throughput required in large proteomics studies

However, factors such as the choice of the matrix, the solvent composition, the pH or temperature and the preparation method can influence the rates of analyte/matrix co-crystallization and, thus, the quality and sensitivity of the MALDI sample preparations [130,161]. In addition, contaminants like buffers, salts and detergents brought in by the sample are further important factors influencing the quality of MALDI-MS analysis [162].

In the following paragraphs the available matrix sample preparations for MALDI MS are described in details and the pitfalls are discussed, together with the improvement brought in this filed by this doctoral thesis work.

Dried-Droplet (DD)

The dried-droplet is the original simple sample preparation procedure introduced in 1988 by Hillenkamp and Karas [38], which has mostly remained intact. The matrix compound dissolved in aqueous solution is mixed with analyte solution, deposited onto the sample plate metal support (the target) and dried. The target with the resulting solid deposit spot of analyte/matrix crystals on it is introduced into the mass spectrometer for analysis. Generally, the DHB, sinapinic acid (SA), and CHCA are the matrices used in DD sample preparation. The DD is a working horse method which tolerates the presence of salts and buffers very well. It is usually a good choice for samples containing more than one protein or peptide

component. However, the DD preparations are characterized by the formation of large crystals with a three-dimensional structure of the spot on the target surface limiting the resolution and spot-to-spot reproducibility of the spectra. In addition this method produces non-homogeneous distribution of analyte/matrix crystals with formation of "sweet spots". Indeed, only in restricted area of the preparation, the "sweet spots", analyte signals of acceptable signal-to-noise (S/N) and spectra of reliable quality can be obtained. These are the major drawbacks of the method impeding the implementation for automated spectra acquisition in high throughput analysis.

Fast-Evaporation, Thin Layer (TL)

The fast-evaporation method [163] was introduced with the main goal of improving the resolution and mass accuracy of MALDI measurements. In this method, matrix and sample are handled separately. An essential requirement is that the matrix compound should be insoluble in aqueous solution. The matrix, (CHCA, or THAP) is applied to the sample target in highly volatile solvent, e.g., acetone, to obtain very fast evaporation. The analyte prepared in aqueous solution is applied on top of the matrix and allowed to dry. Once the preparation has dried the target is introduced into the mass spectrometer for analysis. The preparation delivers stable thin films of matrix on the surface of the target, for this reason became known also as "thin layer"(TL) [160]. This method allows high resolution and enhanced spot-to-spot reproducibility, therefore, is suited for automated MALDI MS spectra acquisition. Notably, TL sample preparation often results in considerable oxidation of methionine and tryptophan side chains. A limitation of this preparation is that the matrix is rapidly ablated by the laser, limiting the amount of laser shots/spot allowed. To overcome this limitation a method with an additional "on-spot" re-crystallization step was introduced. Unfortunately, re-crystallization reduced drastically the sensitivity. However, the main drawback of this method is the extreme intolerance towards impurities brought in by the analyte preparation (such as those from in-gel digests) which seriously hamper the analyte/matrix crystallization process, and therefore the performance. Consequently, time consuming analyte pre-purification step are required.

Sandwich

The sandwich method [164] is derived from the fast evaporation method and the two-layer method. A sample droplet is applied on top of a fast-evaporated matrix bed as in the fast-evaporation method, followed by the deposition of a second layer of matrix in a traditional (non-volatile) solvent. The sample is basically sandwiched between the two matrix layers. This method was first used for the analysis of single mammalian cell lysates. The sandwich technique is compatible with HCCA and SA matrices and often yields significantly better peptide mapping spectra than those obtained with the DD method. With the sandwich method reduced oxidation of methionine and tryptophan side chains has reported

compared to TL method. However, reduced sensitivity and spot-to-spot reproducibility are the major limitations of this method.

Two-Layer

The two-layer method [165,166] involves the use of fast solvent evaporation to form the first layer of small crystals, followed by deposition of a mixture of matrix and analyte solution on top of the crystal layer. The difference between the sandwich and the two-layer method is in the composition and the inclusion of the analyte in the second-layer solution. The addition of matrix to the second step is believed to provide improved results, particularly for proteins and of peptide mixtures containing. Despite an increased tolerance towards impurities; drawbacks of this method are the very limited sensitivity and resolution achieved.

Noteworthy, most of sample preparations described above suffer from common disadvantages such as low shot-to-shot and spot-to-spot reproducibility, short sample life-time as well as high difference in sensitivity and tolerance toward impurities. In addition, the majority of these methods have not been successfully applied to automated spectra acquisition.

In attempt to overcome some the limitations cited above, MALDI sample supports with pre-structured surface were developed (AnchorChip™, Bruker Daltonics; and similar hydrophobic coated target samples supports) [167,168]. These supports provide a pre-structured surface consisting of small hydrophilic islands (sample spot anchors), with different diameters (200, 400, 600 or 800 μm), surrounded by the strongly solvent-repellent surface of a Teflon-like material. The sample anchors hold the sample droplets, in place and direct sample deposition onto them during solvent evaporation. The CHCA matrix using the TL preparation method in combination with pre-structured sample supports was successfully applied to automated spectra acquisition for high throughput analysis. Notably, only previous time consuming and difficult to automate micro pre-purification steps [169].

Therefore, a robust and highly sensitive sample preparation which allows reliable and automated MALDI-TOF MS and MS/MS spectra acquisition was developed.

Matrix layer sample preparation method for proteomics studies (Publication I)

DHB matrix Layer (ML) sample preparation

The new matrix layer (ML) sample preparation method is based on DHB and capitalizes its advantageous features as matrix [170], which in combination with pre-structured MALDI sample

supports[168] allows to automated MALDI-TOF-MS and tandem MS spectra acquisition. This method is simple, requires little preparation time and results in an improved MS data quality. The DHB ML sample preparation method was optimized for the acid used and the total amount of organic solvent applied. Thus, the DHB matrix was prepared at a concentration of 5g/l and dissolved in 0.1 % TFA (trifluoroacetic acid) aqueous solution containing 30% acetonitrile. Best reproducibility, high S/N and resolution, homogeneity of crystals as well as reduced ion suppression are obtained using this formulation.

The DHB ML method involves that the matrix solution is deposited onto the MALDI target and allowed to dry. Subsequently, acidified analyte solution (generally, prepared in TFA) is deposited onto the dried matrix layer. Finally the preparation is re-crystallized "on-target" using a solution containing ethanol in 0.1 % TFA (80/20, v/v).

DHB ML performance for in-solution and in-gel protein digests

The new DHB ML sample preparation enables high quality MALDI-TOF-MS and tandem MS automated spectra acquisition, due to an improved analyte/matrix distribution and high crystals surface homogeneity with small crystals covering the majority of the anchor spot surface. Indeed, an increased shot-to-shot and spot-to-spot reproducibility was evidence. Therefore, MALDI-TOF-MS spectra of equal quality (S/N and resolution) could be acquired from different positions of the sample spot surface. The laser fluence required to obtain ions was constant once the threshold was reached. Notably, crystals appearance and ion signals were unchanged after several laser shots at the same position. Indeed, in-solution and in-gel protein tryptic digest samples prepared with the DHB ML using automated spectra acquisition were compared to DHB DD and CHCA TL methods. With the DHB ML increased number of peptides were detected, improved proteins sequence coverage and reduced mass error were obtained. Notably, analysis of variance (ANOVA) confirmed that the improvements observed from DHB ML are statistically relevant. Moreover, the new DHB ML method demonstrated an improved tolerance toward sample impurities (e.g., salts and buffers), eliminating the need for pre-purification of in-gel tryptic digest of proteins. Thus, the sample preparation is less complex and avoids the loss of informative peptides. Consequently, the DHB ML sample preparation method is well suited for a fully automated MALDI-MS and MS/MS spectra acquisition, leading to reliable and rapid analysis.

Sensitivity

The sensitivity and the dynamic range of detection are two determinant parameters to evaluate the reliability and robustness of a sample preparation for MALDI MS. The DHB ML sample preparation is very sensitive either in case of analysis of purified peptides and in complex mixture of peptides

originated from protein digests. Thus, a limit of detection (LOD) of 1 and 100 attomole was determined for single peptides and protein digests, respectively. Moreover, picomole amounts of peptide mixtures were readily analyzed without performance degradation. Thus, the dynamic range of the DHB ML preparation method is about 10^6.

Phosphorylated protein analysis

The physicochemical characteristics of the negatively charged phosphate groups influence the behavior of phosphopeptides during sample preparation and during the MALDI desorption/ionization process. Therefore, phosphorylated peptides are often detected with low efficiency or not at all in MALDI MS spectra. Notably, the DHB ML method prepared with the addition of phosphoric acid as matrix additive demonstrate an improved ionization of phosphopeptides in MS spectra. Moreover, the method allowed high quality MS/MS spectra for characterization of phosphopetides leading to determine the site of phosphorylation.

Application to proteomics studies

The improved performance and the robustness of the new developed DHB ML preparation method was definitively demonstrated by the application to the proteomic map of mouse lung research project. This project involved 2-DE separation followed by MALDI TOF/TOF MS identification and characterization of protein from mouse lung extracts. Here, the DHB ML allowed a protein identification success rate of 88%. Noteworthy, based on the number identified protein and the peptide detected, the protein sequence coverage, the total number of MS/MS performed and the mascot score obtained, DHB ML demonstrate an improved performance by 44% when compared with CHCA TL method.

Taken collectively, the DHB ML sample preparation demonstrated to be a valuable and robust method for MALDI-TOF MS and MS/MS analysis in proteomics studies.

Appendix Chapter I

Garaguso I, & Borlak J. Matrix Layer Sample Preparation: an Improved MALDI-MS Peptide Analysis Method for Proteomic studies. PROTEOMICS 2008, Vol. 8, Issue 13, 2583–2595.

The following is the pre-peer reviewed version of the article:

Matrix Layer Sample Preparation: an Improved MALDI-MS Peptide Analysis Method for Proteomic studies PROTEOMICS Volume 8, Issue 13, pages 2583–2595, No. 13 July 2008.

Copyright: © 2007 WILWY-VCH Verlag GmBH & Co. KGaA, Weiheim

Web: http://onlinelibrary.wiley.com/doi/10.1002/pmic.200701147/abstract

Rapid communication

Matrix Layer Sample Preparation: an Improved MALDI-MS Peptide Analysis Method for Proteomic studies

*Ignazio Garaguso[1], Jürgen Borlak [1,2] ***

[1]Department of Drug Research and Medical Biotechnology, Fraunhofer Institute of Toxicology and Experimental Medicine, Hannover, Germany

[2]Center for Pharmacology and Toxicology, Hannover Medical School, Hannover, Germany

*Corresponding author:

Prof. Dr. J. Borlak, Department of Drug Research and Medical Biotechnology Fraunhofer Institute of Toxicology and Experimental Medicine, Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany.

Phone: +49-511-5350-559, Fax +49-511-5350-573, E-mail: borlak@item.fraunhofer.de.

List of abbreviations:

peptide mass mapping: PMM; 2,5-dihydroxybenzoic acid: DHB; matrix layer : ML; dried droplet: DD; thin-layer: TL.

ABSTRACT:

The analytical performance of MALDI-MS is highly influenced by sample preparation and the choice of matrix. Here we present an improved MALDI-MS sample preparation method for peptide mass mapping and peptide analysis, based on the use of the 2,5-dihydroxybenzoic acid matrix and pre-structured sample supports, termed: Matrix Layer. This sample preparation is easy to use and results in a rapid automated MALDI-MS and MS/MS with high quality spectra acquisition. The between-spot variation was investigated using standard peptides and statistical treatment of data confirmed the improvement gained with the matrix layer method. Furthermore, the sample preparation method proved to be highly sensitive, in the lower-attomole range for peptides, and we improved the performance of MALDI-MS/MS for characterization of phosphopeptides as well. The method is versatile for the routine analysis of in-gel tryptic digests thereby allowing for an improved protein sequence coverage. Furthermore, reliable protein identification can be achieved without the need of desalting sample preparation. We demonstrate the performance and the robustness of our method using commercially available reference proteins and automated MS and MS/MS analyses of in-gel digests from lung tissue lysate proteins separated by 2-DE.

MALDI-TOF-MS is extensively applied to protein research. Since the pioneering work of Michael Karas and Franz Hillenkamp [1], an improved understanding of the MALDI ionization mechanisms and the development of different matrix substances has extended the use of MALDI-MS beyond proteomic studies. Initially, in MALDI-MS the analytes were embedded in a UV-absorbing matrix, notably tryptophan then nicotinic acid [2]. Thereafter, new matrices as well as new laser and mass analyzer technologies were developed [3, 4] allowing the analysis of (mostly) large bio-molecules in the low fmol range [5, 6]. More recently, MALDI-MS is applied to an analysis of small molecules, small peptides with masses below 1000 amu. as well as lipids and carbohydrates [7, 8]. Additional applications of MALDI MS include an imaging of blotted samples, gels or tissues [9, 10].

In proteomic studies, the combination of high resolution protein separation by two-dimensional gel electrophoresis (2DE) and MALDI-TOF-MS analysis has been proven to be in valuable for an identification of proteins and the characterization of post-translational modifications (PTM) [11-14].

In this regard, the choice of the matrix and the sample preparation method are of critical importance for a successful peptide mass mapping (PMM) analysis, for their influence on the ionization behavior, the formation of adducts, the stability (or fragmentation) of the analytes and, other practical implications in MALDI-MS analysis.

Specifically, derivatives of benzoic acid, e.g., 2,5-dihydroxybenzoic acid (DHB) and of cinnamic acid, e.g., α-cyano-4-hydroxycinnamic acid (CHCA) are commonly used matrices in UV-MALDI. Notably, the DHB matrix is suited for the analysis of low molecular weight compounds, proteins and glycoproteins [15], whereas the CHCA matrix is frequently used for the analysis of peptides and small proteins [5].

In general an optimal sample preparation method should provide a broad mass range for detection of the analyte ions, high sensitivity and mass accuracy as well as reproducibility of sample/matrix crystal preparation to obtain mass spectra. Several methods for sample preparation have been described, the most commonly used are: the dried droplet method (DD) [1], fast solvent evaporation, known as thin-layer (TL) [16], sandwich preparations [17] and two-layer [18]. However, factors such as solvent composition, pH or temperature can influence the rates of matrix–analyte co-crystallization and thus the quality of the MALDI sample preparations [5, 19]. In addition, contaminants like buffers, salts and detergents brought in by the sample are further important factors influencing the quality of MALDI-MS analysis [20].

Here we present an improved DHB sample preparation method, termed: Matrix Layer (ML). The preparation capitalizes advantageous features of DHB as matrix and in combination with pre-structured MALDI sample supports [21, 22] allows to automated MALDI-TOF-MS and tandem MS spectra acquisition. This preparation is easy to use, requires little preparation time and results in an improved MS data quality. Furthermore, we demonstrate that our sample preparation enables high quality MALDI-

TOF-MS and tandem MS automated spectra acquisition, due to an improved analyte/matrix distribution and high crystals surface homogeneity. Indeed, we evidence an increased shot-to-shot and spot-to-spot reproducibility leading to reliable and rapid automated MS analysis. In addition, the DHB ML preparation method demonstrates a sensitivity in the lower-attomole range. Our method shows and improved tolerance toward impurities (e.g., salts and buffers), eliminating the need for pre-purification of in-gel tryptic digests samples. Furthermore, we find that addition of *o*-phosphoric acid to DHB ML preparation to improve the performance of MALDI-TOF-MS/MS analysis as well as the characterization of phosphopeptides. Finally, an increased signal-to-noise ratio and resolution of peaks, a high protein sequence coverage and the high quality MS/MS obtained in the analysis of in-gel digests from lung tissue lysate proteins separated by 2-DE, demonstrate the usefulness of this method for proteomic studies.

The new DHB Matrix Layer (DHB ML) sample preparation method was optimized for the acid used and the total amount of organic solvent applied. In our experience the best reproducibility, signal-to-noise ratio, homogeneity of crystals and reduced ion suppression are obtained when DHB is prepared in 0.1 % TFA acidified solution containing 30% acetonitrile. We used DHB at a concentration of 5g/l, in order to achieve crystallization solely on the anchor surface but keeping an optimum of analyte-to-matrix ratio. Similar to the conventional CHCA TL preparation, for the DHB ML preparation method the matrix solution is deposited onto the MALDI target and allowed to dry. Subsequently, an acidified analyte solution is deposited onto the dried matrix layer. Finally the preparation is re-crystallized "on-target" using a solution containing ethanol/ 0.1 % TFA (80/20, v/v).

In the inserts of Figure 1 photographs of sample preparations of BSA digests using the new DHB ML method and conventional DHB DD and CHCA TL methods (see Supporting Information, Materials and Methods) are depicted. The widely used CHCA TL preparation resulted in the formation of small crystals evenly distributed over the whole anchor surface [Figure 1 (b), insert]. The DHB DD preparation resulted in randomly distributed aggregates of large crystal needles pointing inward to the anchor center [Figure 1 (a), insert]. On the contrary, the new DHB ML showed better crystal distribution with small crystals covering the majority of the anchor surface [Figure 1 (c), insert].

In the case of CHCA TL, analyte ions were obtained from the entire preparation surface. Nonetheless, ion signal intensity varied depending on the position on which the spectrum was acquired. With the conventional DHB DD preparation, spectra of more variable quality and a larger spot-to-spot variation of ions were yielded. The automated spectra acquisition was hampered because of the search for "sweet spots" from which signals of acceptable signal-to-noise ratio are obtained. Only few positions on the sample spot preparation, mainly on large crystals, resulted in good quality spectra. In addition, the laser fluence required to obtain spectra of reasonable quality varied strongly from the different positions on

the surface of the DHB DD preparation, which rendered an automated spectra acquisition impossible. This in agreement well with observations reported elsewhere [6, 30]. In contrast, the new DHB ML preparation resulted in good quality spectra, with favorable signal-to-noise ratio, resolution and intense ion signals. Notably, the crystal appearance and ion signals were unchanged after several shots at the same position. Therefore, MALDI-TOF-MS spectra of equal quality (signal-to-noise ratio and resolution) could be acquired from different positions of the sample surface. The laser fluence required to obtain ions was constant once the threshold was reached. Consequently, the new DHB ML sample preparation method is well suited for a fully automated MALDI-MS spectra acquisition analysis.

To determine the significance of our observation, a standard peptides mixture composed of nine peptides with a mass range of 700 to 3500 m/z was used. Evaluations are based on 27 spots of the same peptide mixture prepared according to the three different matrix preparations described above. For each spot data derived from 400 laser shots were acquired using the same automatic acquisition parameters. The relative standard deviation (RSD) for the mass precision and the average of signal-to-noise intensities were calculated. These results are presented in Table 1.

For all observed peptides the mass precision was much better with the DHB ML preparation methode as compared to DHB DD. The RSDs of the mass precision were in the range 2,5-5,6% (for the mass range 700-2500 *m/z*), except for somatostatin 28 were the RSD was. 22,1% . but with the DHB DD method 34,6%.

Furthermore, with the DBH ML method the signal-to-noise ratio was improved as well (S/N up to 2300) especially in the low mass region (<1300 *m/z*) when compared to CHCA TL (S/N < 1700) method.

To prove superiority of the DHB ML method an analysis of variance (ANOVA) was computed. We used the relative mass error of the single peptide expressed in ppm in the ANOVA test as target factor to assess the differences measured between the methods [Figure 2 (a)] (For further details on the statistics applied see supporting Information, Materials and methods). To address which of the components have the highest significance, a post-hoc Fisher last significance difference (LSD) multiple comparison was performed. The $p < 0,05$ for the direct comparison of DHB ML with the other two method, confirm that the differences are significant (Supporting Information, Table S1). In addition, in Figure 2 [panel (b)] the mean value of measured relative mass error versus the single peptides for all method employed is plotted. The accuracy for the DHB ML method is significantly better (< 30 ppm) when compared with the other methods.

The statistical analysis and the results are presented in table 1. We therefore confirm the DHB ML preparation to be significantly better even in an automated spectra acquisition mode.

To document the advantageous features of the new DHB ML method, four tryptic digests of standard proteins were studied. Notably, BSA, ovalbumin and α-and β-casein were used to compare the DHB ML

with the conventional CHCA TL and DHB DD preparations. All experiments were performed at least three times. Each protein was reduced, alkylated, and thereafter digested with trypsin overnight. Two hundred fmol of the protein digests were prepared using the preparation methods above described.

In Figure 1, the MALDI-TOF-MS spectra of BSA tryptic digests obtained by automated acquisition of the three preparations are given. The new DHB ML preparation [Figure 1 (c)] provided an improved signal-to-noise ratio as compared with the DHB DD preparation [Figure 1 (a)]. In addition, a reduced chemical noise, reduced matrix signal and less ions suppression in the low mass region was observed as compared to CHCA TL preparation [Figure 1 (b)]. Noteworthy, the spectrum of the DHB ML preparation yielded the highest number of peptide signals, therefore leading to enhanced protein sequence coverage. In fact, a total of 52 BSA tryptic peptides were detected with the DHB ML preparation method allowing for a sequence coverage of 76%. In contrast, 33 BSA peptide detected with CHCA TL method allowed a sequence coverage of 57%, while the 45 peptides detected with the DHB DD method resulted in a sequence coverage of 68%. Notably, for all the analyzed protein digests, more peptides were detected with DHB ML than with CHCA TL and/or DHB DD preparations, resulting in higher sequence coverage (data not shown).

The differences between DHB ML and DHB DD preparations in the peak resolution, signal-to-noise ratio and the number of peptide observed are probably caused by different analyte incorporation, thus resulting in different crystal structures of the two preparations. Most likely, the formation of smaller crystals obtained with the DHB ML preparation, is beneficial for an homogeneous incorporation of the sample over the whole preparation surface. Thus, a sufficient analyte-to-matrix ratio and an improved isolation between analyte molecules were obtained, which is required for an efficient desorption and consequently ionization process [31-33]. The mass data obtained differ not only quantitatives but also qualitatives in terms of different peptides observed, probably caused by the ionization process related to the kind of matrices. Notably, DHB leads to the formation of molecular ions with low internal energy, which remains intact during mass analysis. In contrast, CHCA leads to significant decomposition of peptide ions during the analysis [34, 35]. Here, we confirm previous observations demonstrating that the representation of complex analytes in MALDI MS spectra is dependent on the matrix, the solvent composition, and the properties of the analytes it self [31, 5].

MALDI is known to tolerate small amount of impurities and salt contaminants more readily than other MS ionization methods. However, the signal of peptide ions can still be suppressed by the presence of impurities and salts, such as those from in-gel digested samples. The use of micro-columns to desalt and concentrate the analyte prior to MALDI-MS analysis is a widely applied method to get rid of impurities from the sample [30, 36].

In order to study the tolerance towards impurities of our newly developed DHB ML preparation method, we applied it to the analysis of different peptide mixtures obtained from in-gel digest of proteins. The same four standard proteins cited above were loaded on 1-D gel; the stained protein-bands were manually excised and subjected to in-gel digestion with trypsin. An aliquot of the digestion supernatant from each protein-band was prepared using the DHB ML method; same amounts were desalted and concentrated using ZipTip®-C$_{18}$ reversed-phase micro-column and prepared according to DHB DD and CHCA TL preparation methods. The mass spectra obtained from 0.8 µl of supernatant from α-casein in-gel digest, which corresponds to approximately 250 fmol are shown in Figure 3. It should be noticed that the crystals from DHB DD preparation after micro-column purification were particularly large. As previously reported, an automated MS spectra acquisition was not possible. However, from some positions of the DHB DD preparation, peptide signals with a good signal-to-noise ratio were manually acquired [Figure 3 (a)]. The spectrum of CHCA TL preparation is dominated by matrix cluster signals and shows a considerable reduced number of peptide ions [Figure 3 (b)]. In contrast, DHB ML preparation resulted in a more homogeneous crystal surface and the acquired spectrum resulted in signals of comparable signal-to-noise ratio [Figure 3 (c)]. Only in the central region of the preparation, a reduced crystal density was observed. However, the automated spectra acquisition was never hampered and it was always possible to acquire data of sufficient quality. Probably, salts and contaminants are concentrated to the central region of the preparation. This observation is supported by MS that spectra obtained from this region with significant more chemical noise. Moreover, a similar 'desalting' effect has already been described for DHB preparations [4]. This 'desalting' effect seems to be less efficient for DHB DD preparation with the conditions applied herein.

Notably, spectra of DHB DD and CHCA TL preparations were of less quality when prepared without the additional desalting and concentration procedure (data not shown). Nevertheless, an increased signal-to-noise ratio for the peptide ions was observed in the spectra after sample cleanup, but a lower number of informative peptide ions were detected. Examination of the spectra from all analyzed protein-bands digest showed that the loss of peptides included signals corresponding to small hydrophilic and large hydrophobic peptides. Probably, they are not retained on the micro-column or cannot be eluted with the solution used. Noteworthy, a database search with peak list obtained from the spectrum of α-casein, prepared according to DHB ML method, identified the two variants α-casein-S1 and α-casein-S2 with 12 and 15 peptides matched, respectively. With the conventional DHB DD preparation in combination with micro-column purification, only 8 detected peptides matched with α-casein-S1 and 12 in the case of α-casein-S2. Poorly results were obtained when the CHCA TL preparation method coupled with micro-column purification was used, i.e. 5 peptides were detected for α-casein-S1 and 5 for α-casein-S2.

Taken collectively, the DHB ML method is particularly useful for proteomic studies of complex protein samples with the presence of impurities from a gel-based separation step prior to mass spectrometric analysis. The preparation is less complex and no purification is required, which avoids the loss of informative peptides. Furthermore, the preparation results in a comparable signal-to-noise ratio. Finally, the data acquisition can be automated.

The sensitivity of our DHB ML sample preparation method was studied using the peptide Glu[1]-Fibrinopeptide B (GluFib). In Figure 4 (a) and (b) spectra obtained from a dilution series from 10 amol to 1 amol of GluFib are depicted. A 0.8 µl volume of each dilution was prepared with the DHB ML preparation method and compared with the CHCA TL method. With 10 amol of the GluFib applied on the target, spectra acquisition was straightforward. Laser fluence just above the threshold produced spectra with reliable quality for every single shot. Loading an amount of 1 amol, with the DHB ML preparation method, the peptide signal was readily detectable. Comparable results were obtained with CHCA TL preparation method but peak intensity and signal-to-noise ratio of both preparations was different. Attempts to detect 0,5 amol of GluFib prepared with the DHB ML method were successful in roughly half of all measurements, whereas, the detection of the same amount of peptide failed with the CHCA TL method. The DHB DD preparation method was also included in the study but proved to be less sensitive (data not shown).

In-solution tryptic digest of reduced and carbamidomethylated BSA was used for evaluating the sensitivity of our method for a peptide mixture. The loading on sample support of 100 fmol of the protein digest, resulted in intense signal of peptides with high signal-to-noise ratio. A total of 42 BSA peptides were detected and the sequence coverage was 63.3%. Similar sequence coverage was obtained for 10 fmol of digest, where 39 BSA peptides were detected allowing for a sequence coverage of 57%. A slight decrease in the signal-to-noise ratio and loss of some peptide signals was observed between 10 and 5 fmol. At 1 fmol of BSA digest 22 peptides were detected with a sequence coverage of 36.1% . Below 1 fmol, e.g. at the 100 amol level, the spectra deteriorated dramatically, several peptide signals were lost and the signal-to-noise ratio decreased significantly [Figure 4 (c)]. However, a database search with peak list obtained from this spectrum still identified BSA with a score of 69 and 10 peptides matched allowing to a sequence coverage of 17,3%. In general, the sensitivity of the DHB ML preparation was slightly higher than the CHCA TL preparation. Attempts to detect 100 amol with DHB DD and CHCA TL methods failed [Figure 4 c (1) and (2)].

Here, we established the lower limit of the routine working range of the DHB ML method to be in the lower-attomole range. Moreover, picomole amounts of peptide mixtures were readily analyzed without performance degradation. Thus, the dynamic range of the DHB ML preparation method is at least 10^6.

Encouraged from our findings with the DHB ML method, we tested its performance when applied to MALDI tandem mass spectrometry. Two singly protonated peptides from the spectrum of α-casein in-gel digest showed in Figure 3,were selected as precursor ions for MALDI MS/MS analysis. Figure 5 (a) displays the MS/MS spectrum of the single protonated precursor ion at *m/z* 1098.60. The complete y- and b-ion series was readily obtained, a database search using the fragmentation data identified the peptide [204]AMKPWIQPK[212] from α-casein-S2 with an ion score of 53 .The fragmentation pattern of the single protonated precursor ion at *m/z* 1267.70, is shown in Figure 5 (b). An almost complete y- and b-ion series corresponding to the α-casein-S1 peptide [106]YLGYLEQLLR[115] was identified with an ion score of 40. This confirms our initial identification of α-casein-S1 and α-casein-S2 in the PMM analysis of the in-gel digests and broadens the application of the DHB ML preparation method for MALDI tandem MS analysis as well.

Furthermore, we were interested to apply our method for the detection and characterization of phosphorylated peptides. Notably, phosphorylated peptides are often detected with low efficiency or not at all, due to the physicochemical characteristics of the negatively charged phosphate groups, which influence the behavior of phosphopeptides during sample preparation and during the MALDI desorption/ionization process. Recent reports suggest the use of *o*-phosphoric acid as matrix additive to enhance the signal of phosphorylated peptides in complex mixture [19, 37]. Thus, we tested the performance of DHB ML preparation method by addition of phosphoric acid to the matrix and the re-crystallization solutions. Best results were obtained at the concentration of 0.1% *o*-phosphoric acid which has no negative influence on the crystallization and on the hydrophobic coating of sample support. The bovine milk α-casein is phosphorylated stoichiometrically at serine residues. With the DHB ML preparation using TFA, only the peptide ion *m/z* 1951.95 was readily detectable [see figure 3 (c) ion label S1p). By use of the DHB ML preparation with addition of 0.1 % of *o*-phosphoric acid, the ions at *m/z* 1660.79, *m/z* 1927.69 and *m/z* 1951.95 were readily detected in the spectra of 400 fmol of in-solution digestion of α-casein [Figure 6 (a)]. Thus these ions were selected and tandem MS/MS spectra were acquired. Figure 6 (b) and (c) display the MS/MS spectra of ions *m/z* 1660.79 and *m/z* 1951.95. The MS/MS spectra of these peptides were dominated by an [MH - 98]+ ion and [MH - 80]+ ion, which correspond to gas phase β-elimination of a H_3PO_4 and loss of HPO_3 groups, respectively. Note, the facile loss of H_3PO_4 and HPO_3 did not prevent the formation of the amino acid sequence ions. Thus, the nearly complete series of y- ions was easily detected, corresponding to the α-casein peptide [121]VPQLEIVPN(pS)AEER[134] at *m/z* 1660.79 which was identified with an ion score of 94. The ion at *m/z* 1951.95 was identified with an ion score of 48 as the peptide [119]YKVPQLEIVPN(pS)AEER[134]. Furthermore, the MS/MS fragmentation pattern of both peptides reveled the position of the phosphate moiety on serine 130. Figure 6 (d) displays the MALDI-MS/MS spectrum of the ion *m/z* 1927.69. In this

case two abundant ion signals, at $[MH - 98]^+$ and $[MH - 196]^+$, corresponding to neutral loss of two H_3PO_4 were detected downstream of the parent ion signal. Thus, two phosphorylated amino acids in the peptide are present. In fact, a database search using the fragmentation data assigned the ion *m/z* 1927.69 to the phosphopeptide [58]DIG(pS)E(pS)TDQAMEDIK[73] with an ion score of 42. Notably, the y- and b- ion series allowed clearly to identify the two phosphorylated residues as serine 61 and 63, excluding the threonine in position 64. Similar results were obtained for phosphorylated peptides of the other analyzed proteins (data not shown). These results confirm the compatibility of our DHB ML sample preparation method with the addition of *o*-phosphoric acid for the analysis of phosphorylated peptides to characterize the phoshporylation sites.

To demonstrate the performance and the robustness of our new DHB ML preparation method in a proteomic research project, we analyzed in-gel digests from 2-DE gel separated proteins of mouse lung lysate. Arbitrary, a set of 66 protein spots were excised from different parts of 200 µg proteins loaded 2-DE mini-gel (pH gradient 3-11 non linear), followed by in-gel digestion with trypsin and the resulting spectra were analyzed. The 66 excised protein spots (Figure 7) were classified in three groups according to their intensity: weak (W: 26spots), medium (M: 22 spots), and strong (S: 18 spots) as denoted in Table S2 (Supporting information). An amount of 1 µl of the digestion supernatant from each spot was prepared according to DHB ML method and loaded onto the pre-structured MALDI sample support. MALDI-MS spectra acquisition and tandem MS precursor ions selection and acquisition was done in automated mode (see Supporting Information, Material and Methods). Figure 8 depicts representative peptide mass maps of the identified protein in the spot 44 using the two different sample preparation methods. In both cases, the protein could be identified as Selenium-binding protein 1. Notably, the CHCA TL method gave sequence coverage of 24% with 10 peptide matched and a Mascot score of 94 [Figure 8 (a)]. The sequence coverage with the DHB ML method was 70% with 32 peptide matched and a Mascot score of 233 [Figure 8 (b)]. In addition, only with the DHB ML method three ions could be automatically selected for MS/MS spectra acquisition with high ion scores, thus confirming the robustness of the DHB ML method for MALDI tandem MS analysis [Figure 8 (c), (d), (e)].

The result of the comparison showed that the DHB ML preparation method was in general superior, when compared to the CHA TL method in terms of improved number of peptides identified, sequence coverage and Mascot score (Figure 9). The DHB ML method gave an average sequence coverage of 42%, while the classical CHCA TL method average sequence coverage of 32, respectively (Figure 9). Furthermore, a total of 71 MS/MS spectra were successfully acquired with DHB ML method, compared to only 31 with the CHCA TL method. Overall, an improved performance by 44% was determined, furthermore, some peptides were detected uniquely with one or the other preparations.

Taken collectively, with the DHB ML preparation method 58 of 66 mouse protein spots analyzed were identified by database search as compared to 56 with the CHCA TL method. By iterative database searching of the data obtained with the DHB ML method a mixture of proteins could be analyzed for the protein spots: 13, 17 and 58. Using the CHCA TL method only the spot 58 a mixture could be analyzed. For the spot 31, assigned as medium, the identification was confirmed by MS/MS. Medium- and strong-intensity spots were identified with a success rate of 100%, while spots with weak intensity were identified with a success rate of 69%. Note, this corresponds to an overall success rate of 88%. Combined with reasonable high amino acid sequence coverage and the high sensitivity demonstrated, DHB ML sample preparation method is a viable and robust alternative to standard matrix preparation protocols.

The DHB ML method presented here has several advantages in terms of facile preparation, feasibility for automated data acquisition. With this method a decrease in chemical noise and a high tolerance toward impurities avoided the loss of informative peptides leading to higher proteins sequence coverage as compared with conventional sample preparation methods. With this matrix sample preparation, the signal-to-noise ratio and resolution for the ion signals over the entire preparation surface is improved, most likely due to the formation of smaller crystals which are homogeneously distributed over the whole spot surface. The high mass accuracy, the increased number of peptides ions detected, the higher sequence coverage and the enhanced tandem MS performance proven the reliability and robustness of the DHB ML method.

When prepared using *o*-phosphoric acid as additive, our DHB ML preparation method can efficiently be applied for an identification of phosphorylated proteins and the characterization of the phosphorylation site.

Acknowledgements

Supporting Information Available

Table S1: Fisher last significance difference.

Table S2: List of Mouse Liver Proteins identified on the 2DE gel.

Material and methods

Table 1

Spot to Spot Variation			
PEPTIDES	**CHCA TL**	**DHB DD**	**DHB ML**
Bradykinin (1-7)			
m/z average(Da)	757,439	757,345	757,350
m/z STDEV (%)	1,8%	4,2%	2,5%
S/N average	77,7	1157,5	1342,9
Angiotensin II			
m/z average(Da)	1046,586	1046,505	1046,503
m/z STDEV (%)	2,4%	5,1%	2,5%
S/N average	1691,6	2316,8	2346,3
Angiotensin I			
m/z average(Da)	1296,736	1296,667	1296,658
m/z STDEV (%)	2,8%	6,4%	2,6%
S/N average	544,2	1018,7	1222,7
Substance P			
m/z average(Da)	1347,789	1347,720	1347,713
m/z STDEV (%)	2,8%	6,8%	2,8%
S/N average	2071,4	770,2	1345,8
Bombesin			
m/z average(Da)	1619,880	1619,815	1619,811
m/z STDEV (%)	3,2%	8,2%	3,2%
S/N average	2264,9	685,8	1256,8
Rinein Substrate			
m/z average(Da)	1758,990	1758,940	1758,925
m/z STDEV (%)	3,5%	9,2%	3,3%
S/N average	1035,7	775,3	758,5
ACTH Clip (1-17)			
m/z average(Da)	2093,153	2093,092	2093,082
m/z STDEV (%)	4,2%	10,3%	4,1%
S/N average	932,3	877,9	762,7
ACTH Clip (18-39)			
m/z average(Da)	2465,279	2465,200	2465,192
m/z STDEV (%)	4,7%	12,0%	5,6%
S/N average	1650,3	1170,8	642,4
Somatostatin (28)			
m/z average(Da)	3147,586	3147,576	3147,479
m/z STDEV (%)	6,5%	34,6%	22,1%
S/N average	182,5	80,3	169,8

Table 1

Spot to spot variation using DHB ML and DHB DD CHCA TL methods for the obtained molecular masses [M+H]+ in spectra of 500 laser shots, acquired using the same automatic acquisition parameters, at 27 different sample spots with nine standard peptides.

Figures

Figure 1
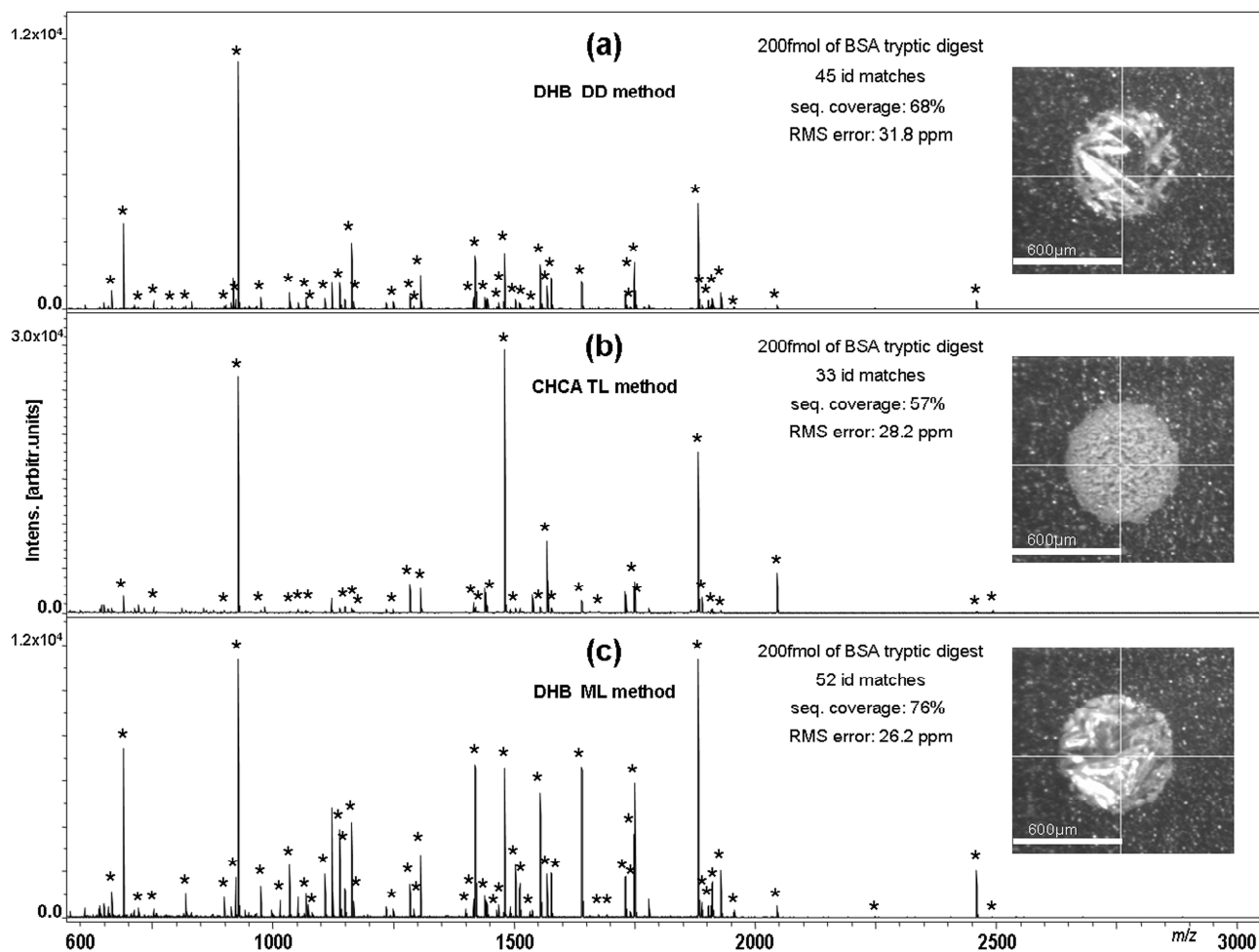
Figure 1. DHB ML preparation method.

MALDI-TOF mass spectra obtained from a BSA tryptic digest. A 200 fmol amount of the digest was prepared with (a) CHCA TL, (b) DHB DD, and (c) DHB ML.

Stars indicate signals matching tryptic peptides of BSA.

With DHB DD method 45 peaks matched with BSA allow for sequence coverage of 68%. With CHCA method 33 peaks matched and a sequence coverage of 57%: Using the DHB ML method the peaks matched were 53 allowing for a sequence coverage of 76%.

The inserts display images of the relative preparation methods.

Figure 2
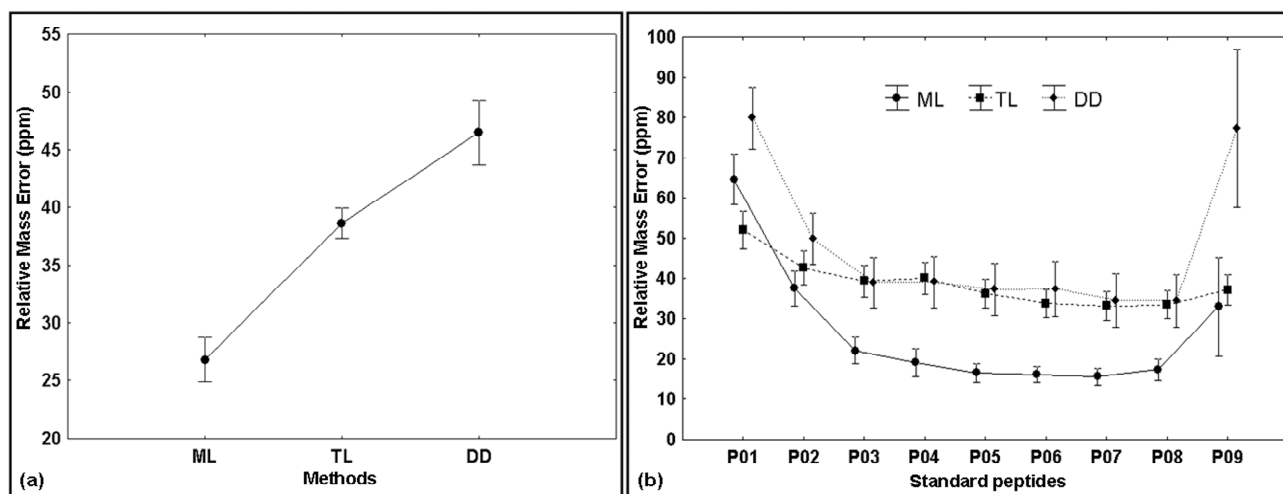
Figure 2. Statistical analysis of spot-to-spot variation.

(a) The relative mass error of the single peptide expressed in ppm was used as target factor to compare methods. We use the software STATISTICA to perform the ANOVA analysis. F(0,05;2;691) = 28,14. F_{critic} (0,05;2;691) = 3,01. Critic values of F when Alpha= 0,05, degree of freedom between= 2, degree of freedom within= 691.

(b) The DHB ML, the CHCA TL and DHB DD were compared on the basis of the mean of relative mass error. The accuracy of the DHB ML method is significantly improved when compared to the CHCA TL and DHB DD (Relative mass error <30pmm).

Vertical bars indicate standard error. For detailed peptide masses see Supporting Information, Materials and Methods.

Figure 3
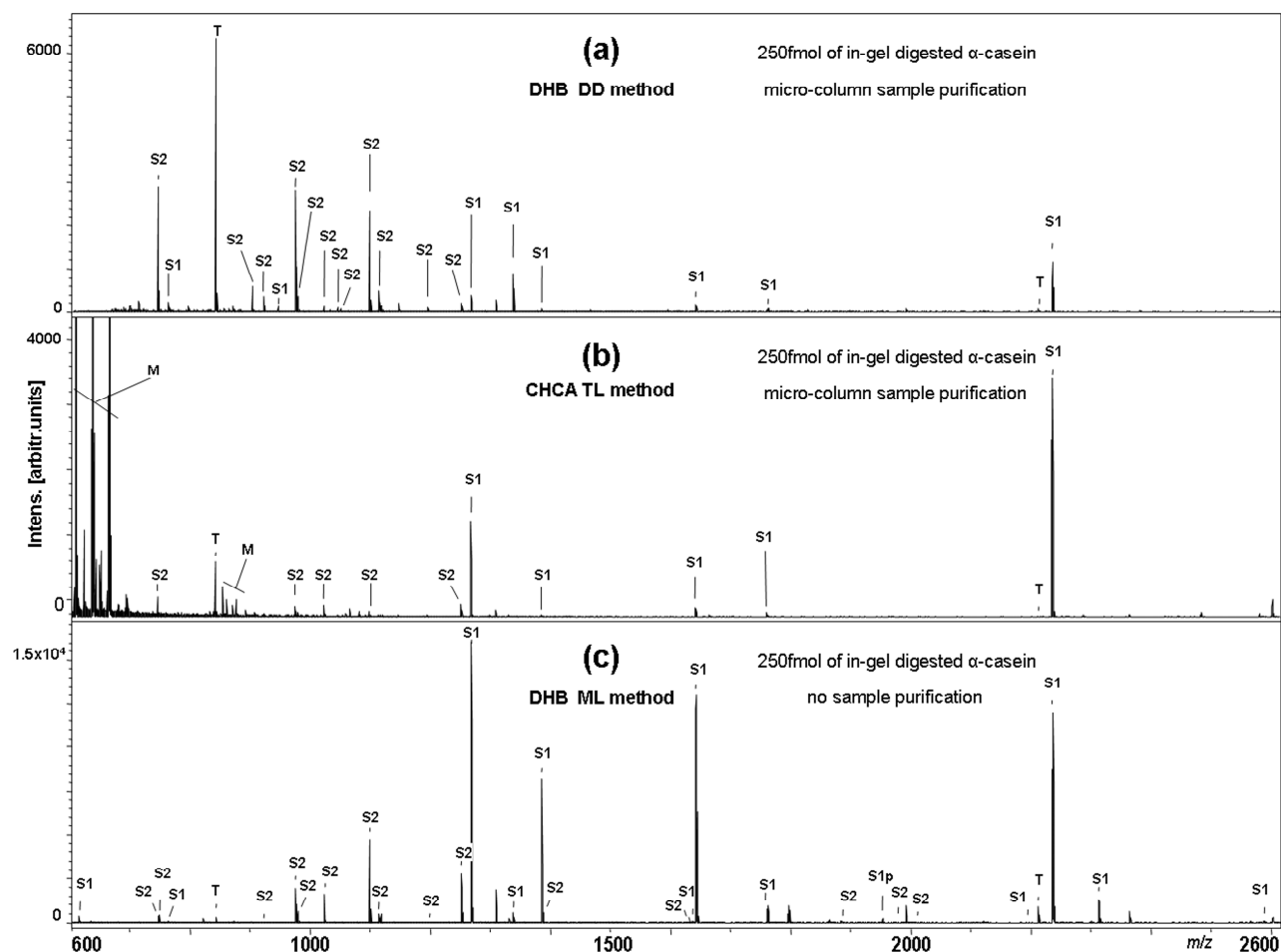
Figure 3. MALDI-TOF mass spectra obtained from a α-casein in-gel tryptic digest.

A 250 fmol amount of the digest was prepared with (a) the DHB DD and (b) the CHCA TL method and included a desalting and concentration step using micro-columns (see text), (c) the DHB ML preparation method without sample purification.

The tryptic peptide ions of α-casein-S1 and α-casein-S2 are indicated with S1 and S2 respectively.

Commonly observed matrix and trypsin auto-digest signals are indicated by M and T, respectively.

The ion peak S1p in panel (c) *m/z* 1951.95, indicate the only poshorylated peptide ion detectable with DHB ML preparation using TFA.

Figure 4
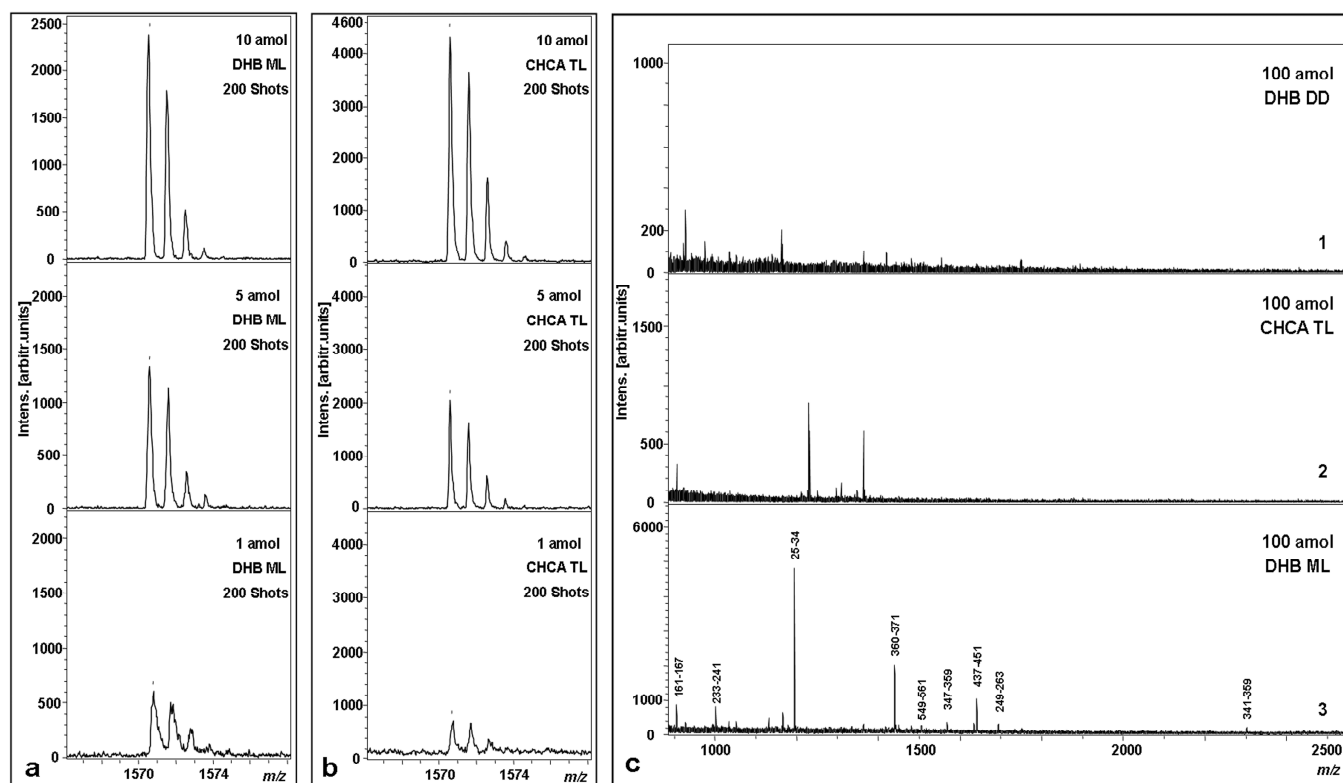
Figure 4. Sensitivity of the DHB ML sample preparation method.

MALDI-TOF mass spectra of a dilution series of the peptide Glu[1]Fibrinopeptide B (monoisotopic m/z 1570.59) using DHB ML (a) and CHCA TL (b). A 0.8 µl volume of sample, in 25% ACN, 0.1% TFA, was used for each sample.

(c) MALDI-TOF mass spectra of BSA tryptic digest. A 100 amol amount of the digest was prepared with (1) the DHB DD, (2) CHCA TL and (3) DHB ML method.

Figure 5
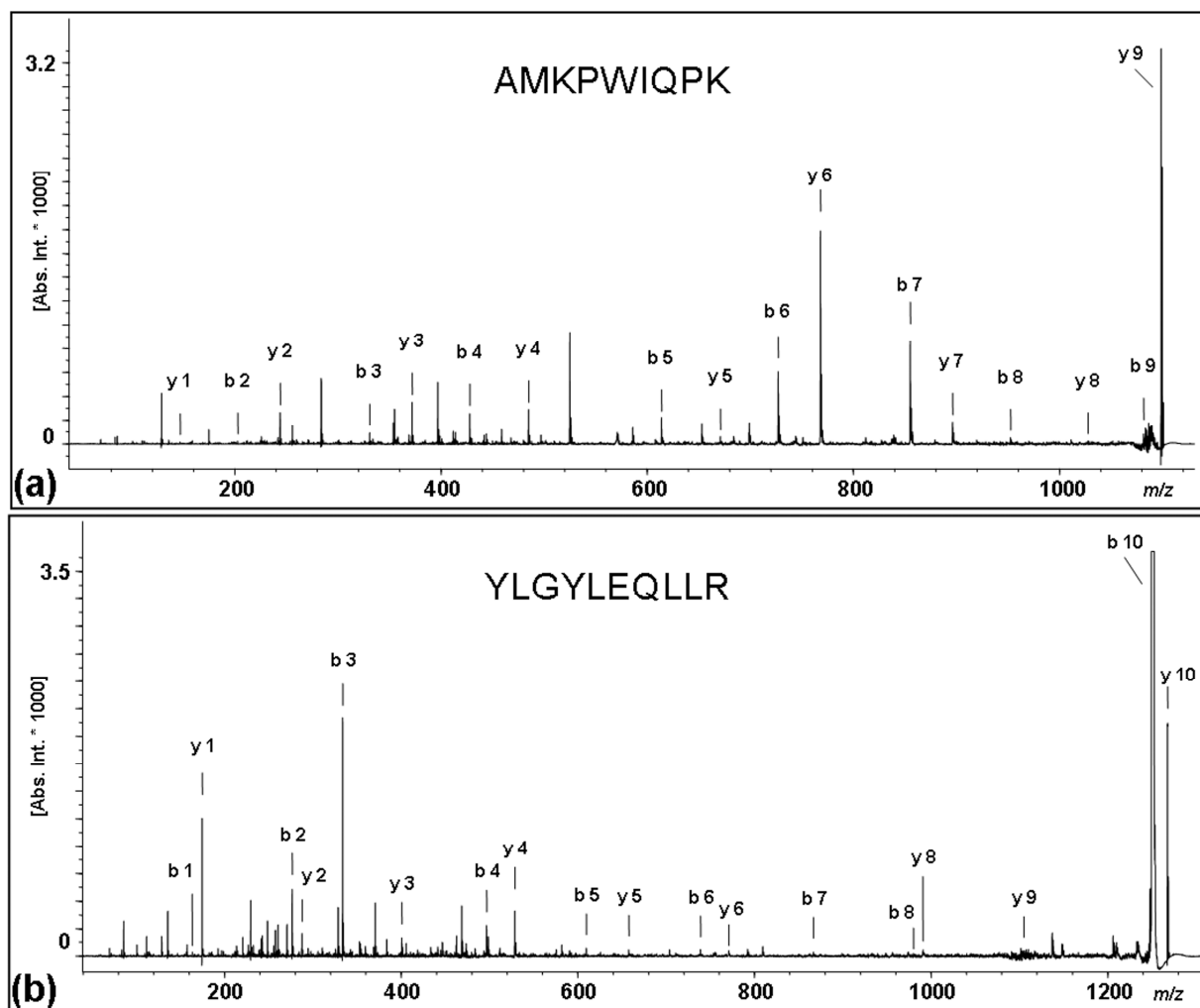
Figure 5. Peptide sequencing by MALDI-MS/MS with the DHB ML sample preparation. MALDI-TOF/TOF mass spectra of singly protonated peptide ions from α-casein. (a) The single protonated precursor ion at *m/z* 1098.60 and the y- and b-ion series could be assigned to the peptide [204]AMKPWIQPK[212] of α-casein-S2 with an ion score of 53.

(b) The fragmentation pattern of the single protonated precursor ion at *m/z* 1267.70 revealed the complete y- and b-ion series of the α-casein-S1 peptide [106]YLGYLEQLLR[115] and was identified with an ion score of 40.

The y-H2O- and b-H2O-ions are not labeled for simplicity.

Figure 6
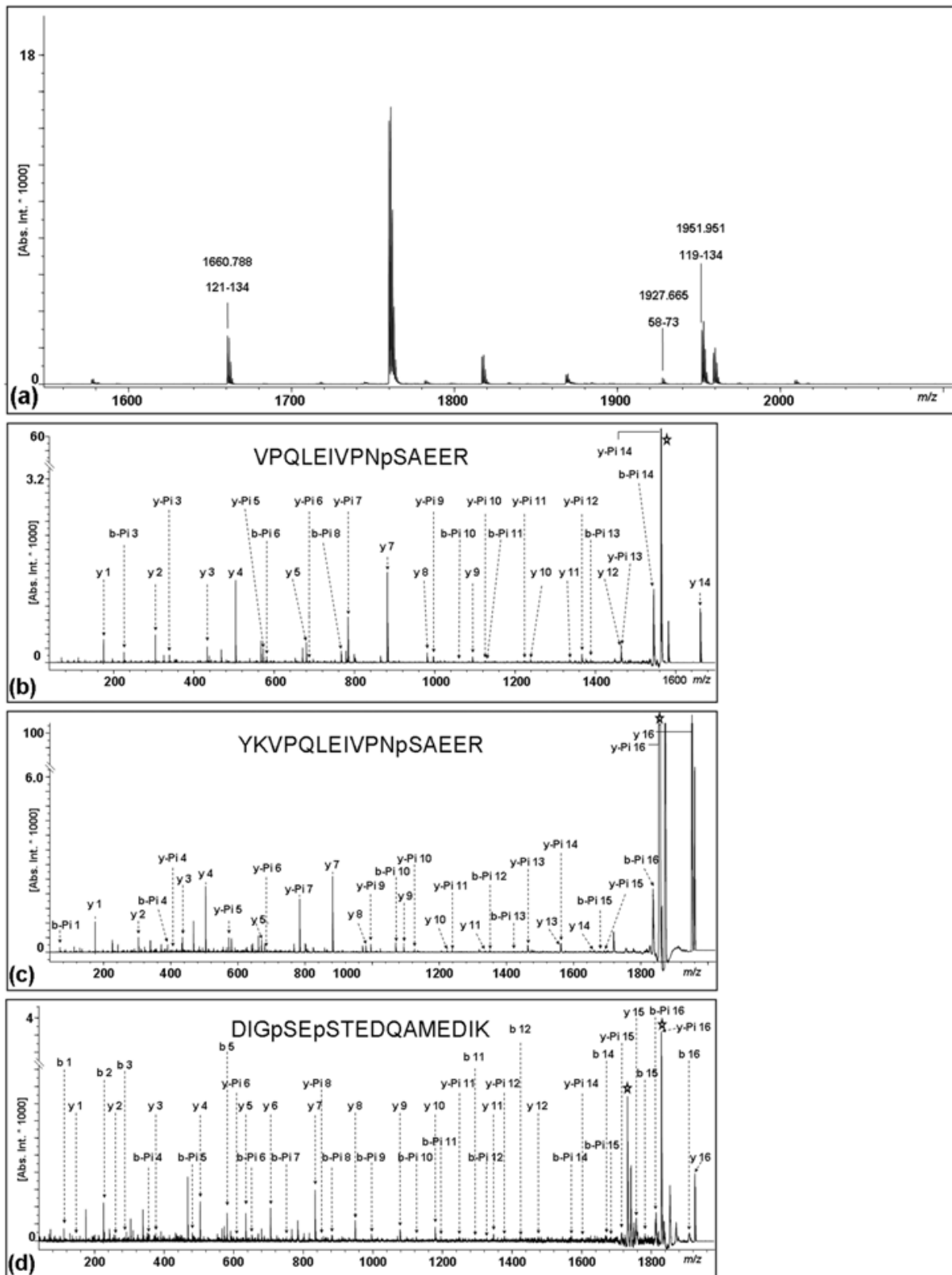
Figure 6. Enhanced desorption/ionization efficiency of phosphorylated molecular ions by addition of *o*-phosphoric acid (PA) in the DHB ML preparation solutions(matrix and re-crystallization solution). Application of o-phosphoric acid (PA) as matrix additive enabled recovery of both singly and multiply phosphorylated species in a tryptic digest of α-casein.

(a) Detailed view of the region *m/z* 1600–2000 of tryptic MALDI-TOF peptide mass map of 400 fmol α-casein tryptic digest prepared with the DHB ML method using 0.1% PA. Only the phosphorylated molecular ions are indicated.

 Note, in the spectrum of α-casein in-gel digest DHB ML prepared with TFA only the ion *m/z* 1951.95 was detectable [see Figure 2 (c), molecular ion *m/z* 1951.95 labeled as S1p].

Phosphopeptides sequencing by MALDI-MS/MS.

(b) and (c) MALDI-TOF-MS/MS of singly protonated phosphopeptide ions from α-casein-S1. The monophosphorylated peptides displayed a near complete series of y-ions confirming the sites of modification. Multiple ion signals from mainly y-H2O but also b- and b-H2O have not been labeled for simplicity. The phosphopeptide precursor ion signal [M+H]$^+$ *m/z* 1660.79 and *m/z* 1951.95 [labeled as y-14 in (b) and as y-16 in (c), respectively], were accompanied by intense -98 Da satellite ion signals due to β-elimination of phosphoric acid from phosphoserine (S130) (✮).

(d) MALDI-TOF-MS/MS of double protonated phosphopeptide ion from α-casein-S1. The phosphopeptide precursor ion signal [M+H]$^+$ *m/z* 1927.69 (labeled as y-16), was accompanied by two intense -98 Da ion signals due to successive β-elimination of phosphoric acid from the two phosphoserine (✮). The doubly phosphorylated peptide displayed a complete series of y-ions leading to assign the phosphorylation to the two serine in the sequence (S61 and S63). The, y-H2O- and b-H2O-ions are not labeled for simplicity.

Figure 7
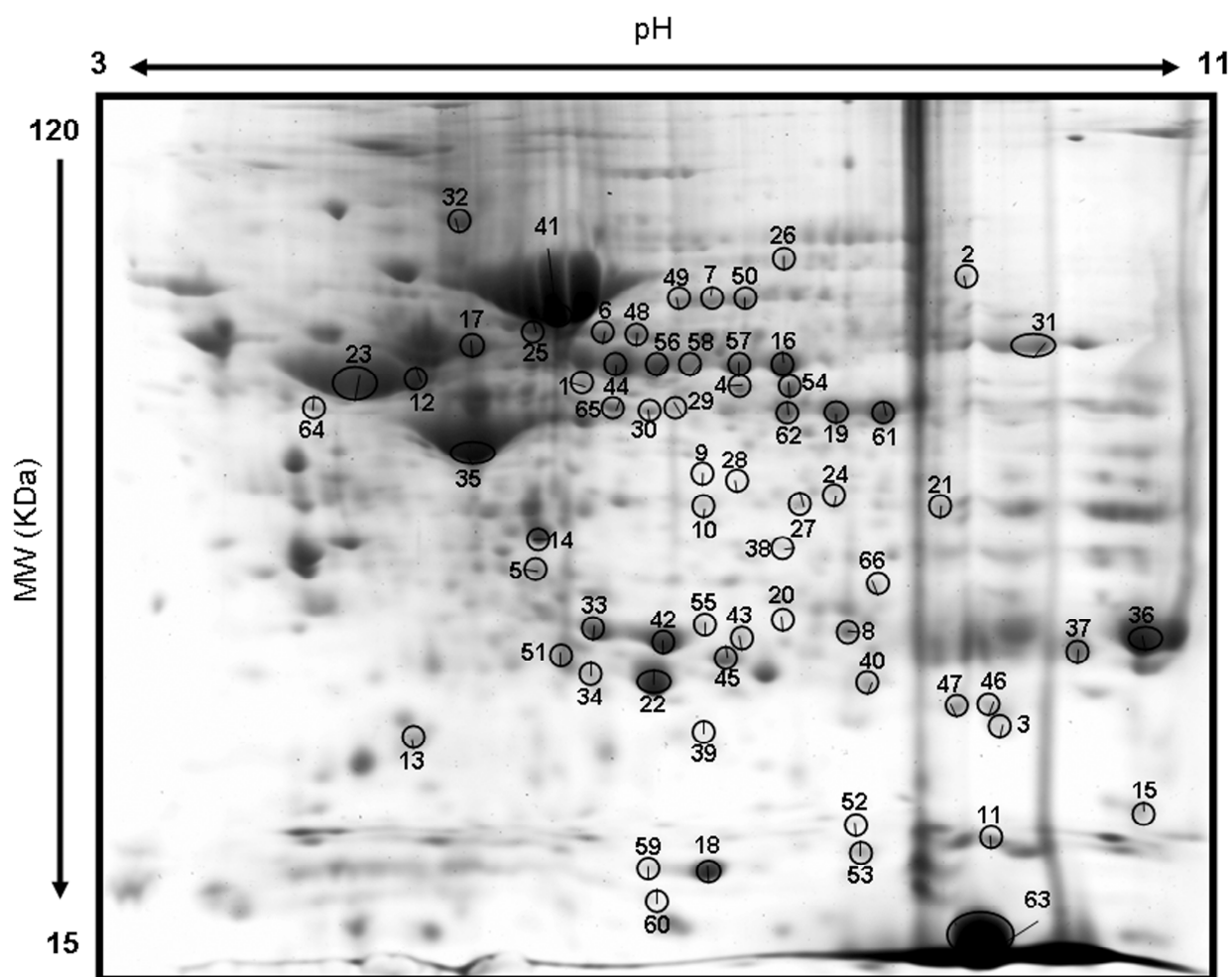
Figure 7. 2D-gel separation of mouse lung proteins visualized with colloidal Coomassie Brilliant Blue G-250.

The separation was performed with pH 3-11 IPG strips, 70 X 50 mm second dimension SDS-PAGE gel 12% total acrylamide. A total 66 protein spots with molecular masses between 15and 100 kDa were prepared using the DHB ML method and loaded onto pre-structured sample support. A total of 58 protein spot could be identified with PMM. For 3 protein spots of the 58, a mixture of two proteins each was found. A total of 71 tandem MS spectra were from the automated system selected and performed to confirm the identifications.

Figure 8
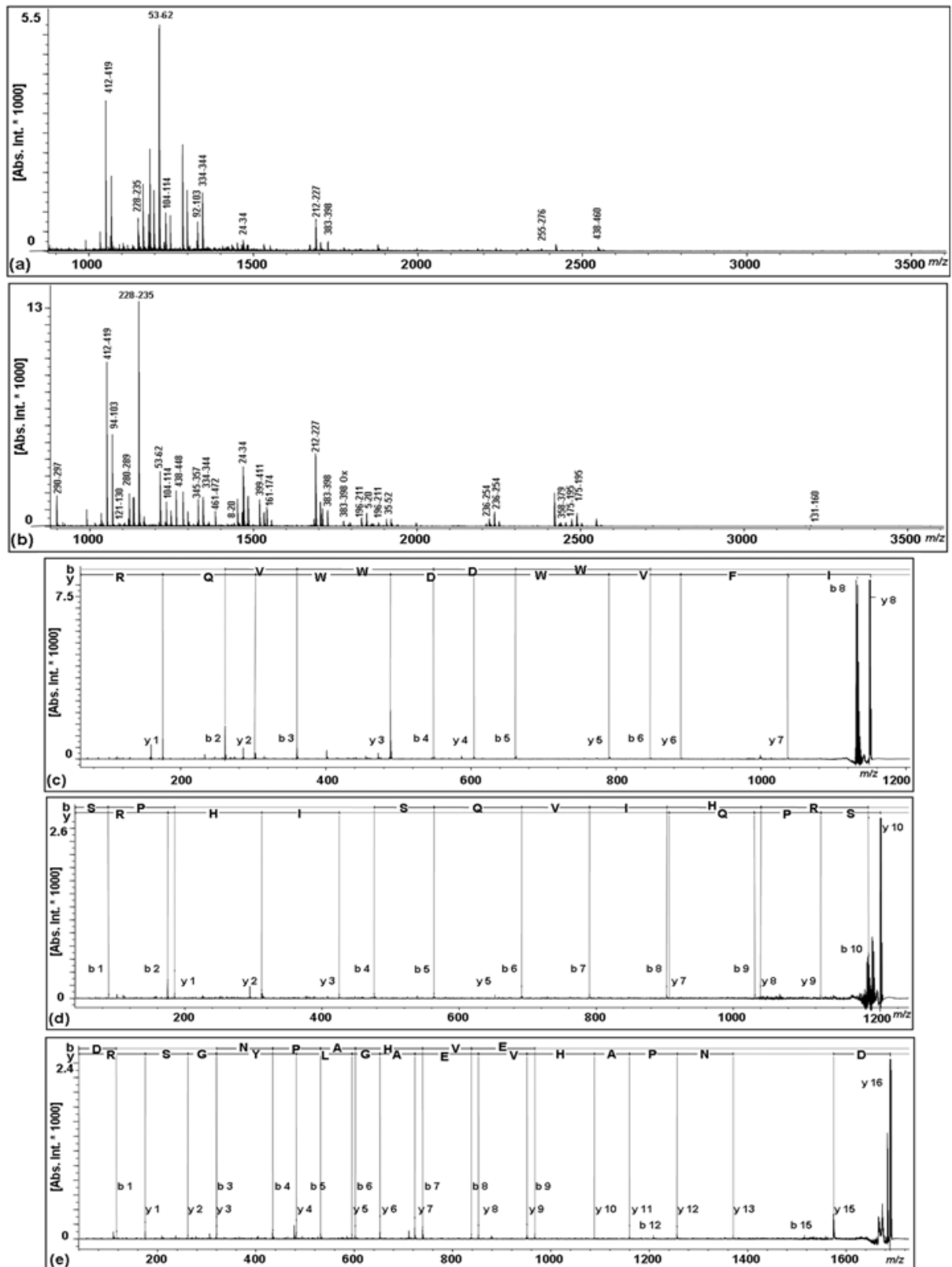
Figure 8. MALDI-TOF-MS peptide mass maps and MS/MS of in-gel digest of spot 44 separated by 2D gel electrophoresis (Figure 6) illustrating that the DHB ML sample preparation method on pre-structured sample support improve sequence coverage and the MS/MS performance.

(a) MALDI-TOF-MS peptide mass map obtained with traditional CHCA TL preparation method. A total of 10 peptides could be assigned to Selenium-binding protein 1, which corresponds to an protein sequence coverage of 24% and a MASCOT score of 94.

(b) MALDI-TOF-MS Peptide mass map using DHB ML preparation method. A protein sequence coverage of 70% could be obtained, which corresponded to 32 identified peptides with a MASCOT score of 233.

Peptide sequencing by MALDI-TOF-MS/MS of Selenium-binding protein 1, spot 44.

(c) MALDI-TOF-MS/MS of single protonated ions *m/z* 1149.60, y-ion series allowed to assign to the peptide [228]IFVWDWQR[235] of Selenium-binding protein 1 with an ion score of 33.

(d) MALDI-TOF-MS/MS of single protonated ions *m/z* 1212.64, the peptide [53]SPQYSQVIHR[62] of the same protein, was identified with an ion score of 28.

(e) MALDI-TOF-MS/MS of single protonated ions *m/z* 1689.83. The fragmentation pattern revealed the complete y- and b-ion series of the peptide [212]DGFNPAHVEAGLYGSR[227] of Selenium-binding protein 1 and was identified with an ion score of 98.

The y-H2O- and b-H2O-ions are not labeled for simplicity.

Figure 9

Figure 9. A Comparison of the DHB ML and the CHCA TL methods using in-gel digested proteins from lung tissue lysate. In general, the DHB ML method improved the sequence coverage (a), the number of identified peptides (b), and Mascot score (c). The DHB ML and the CHCA TL protocols yielded an average of protein sequence coverage of 42% and 32%, respectively.

* = second protein identified for the spot.

References

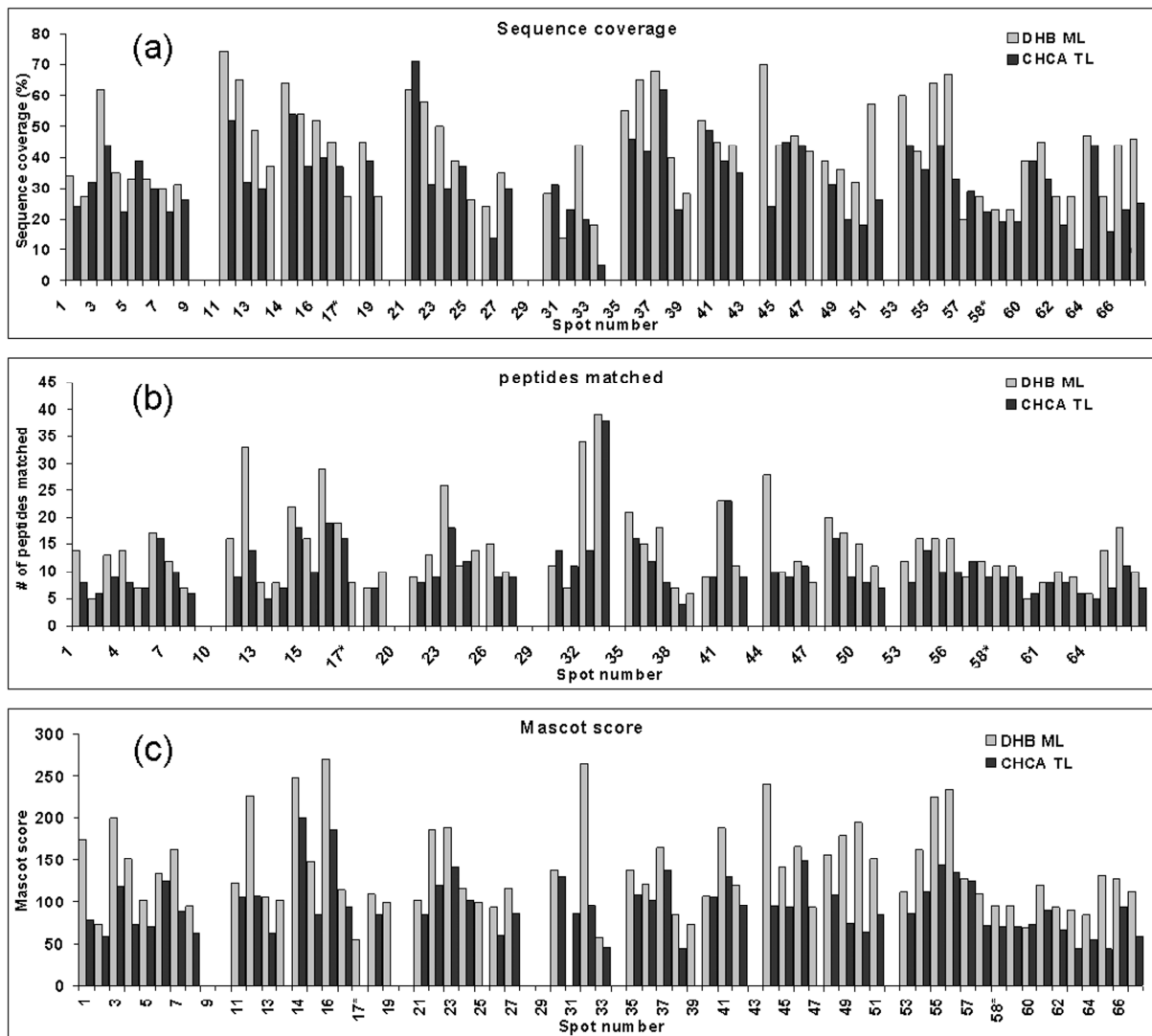[1]   Karas, M., Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 1988, 60, 2299–2301.

[2]   Karas, M., Ingendoh, A., Bahr, U., Hillenkamp, F., Ultraviolet-laser desorption/ionization mass spectrometry of femtomolar amounts of large proteins. *Biomed. Environ. Mass Spectrom.* 1989, 18, 841–843.

[3]   Beavis, R. C., Chaudhary, T., Chait, B. T., α-cyano-4-hydroxycinnamic Acid as a Matrix for Matrix-assisted Laser Desorption Mass Spectrometry. *Org. Mass Spectrom.* 1992, 27, 156-158.

[4]   Strupat, K.; Karas, M.; Hillenkamp, F., 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption—ionization mass spectrometry. *Int. J. Mass Spectrom. Ion Processes* 1991, 111, 89.

[5]   Cohen, S. L., Chait, B. T., Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal. Chem.* 1996, 68, 31–37.

[6]   Önnerfjord, P., Ekstrom, S., Bergquist, J., Nilsson, J., et al., Homogeneous sample preparation for automated high throughput analysis with matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 1999, 13, 315–322.

[7]   Andalo, C., Bocchini, P., Pozzi, R., Galletti, G. C., Accurate mass measurement of synthetic analogues of prazosine by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* 2001, 15, 665-669.

[8]   Cohen, L. H., Gusev, A. I., Small molecule analysis by MALDI mass spectrometry. *Anal. Bioanal. Chem.* 2002, 373, 571–586.

[9]   Nicola, A. J., Gusev, A. I., Proctor, A., Jackson, E. K., Hercules D. M., Application of the fast-evaporation sample preparation method for improving quantification of angiotensin II by matrix-assisted laser desorption/ionization. *Rapid Commun. Mass Spectrom.* 1995, 9, 1164–1171.

[10]   Stoeckli, M., Farmer, T. B., Caprioli, R. M., Automated mass spectrometry imaging with a matrix-assisted laser desorption ionization time-of-flight instrument. *J. Am. Soc. Mass Spectrom.* 1999, 10, 67–71.

[11]   Anderson, N. L., Anderson, N. G., The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 2002, 1, 845–867.

[12]   Lisacek, F. C., Traini, M. D., Sexton, D., Harry, J. L., Wilkins, M. R., Strategy for protein isoform identification from expressed sequence tags and its application to peptide mass fingerprinting. *Proteomics* 2001, 1, 186–193.

[13]   Claverol, S., Burlet-Schiltz, O., Girbal-Neuhauser, E., Gairin, J. E., Monsarrat, B., Mapping and structural dissection of human 20 S proteasome using proteomic approaches. *Mol. Cell. Proteomics* 2002, 1, 567–578.

[14]   Shevchenko, A., Jensen, O. N., Podtelejnikov, A. V., Sagliocco, F., et al., Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two dimensional gels. *Proc. Natl. Acad. Sci. U.S.A.* 1996, 93, 14440–14445.

[15]   Juhasz, P., Costello, C.E., Biemann, K., Matrix-assisted laser desorption ionization mass spectrometry with 2-(hydroxyphenyl-azo)benzoic acid matrix. *J Am. Soc. Mass Spectrom.* 1993, 4, 399–409.

[16]   Vorm, O., Roepstorff, P., Mann, M., Improved Resolution and Very High Sensitivity in MALDI TOF of Matrix Surfaces Made by Fast Evaporation. *Anal. Chem.* 1994, 66, 3281-3287.

[17]   Kussmann, M., Nordhoff, E., Rahbek-Nielsen, H., Haebel, S., et al., Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. *J. Mass Spectrom.* 1997, 32, 593-601.

[18]   Dai, Y. Q., Whittal, R. M., Li, L., Confocal Fluorescence Microscopic Imaging for Investigating the Analyte Distribution in MALDI Matrices. *Anal. Chem.* 1996, 68, 2494-2500.

[19]   Kjellstrom S., Jensen O. N., Phosphoric acid as a matrix additive for MALDI MS analysis of phosphopeptides and phosphoproteins. *Anal. Chem.* 2004, 76, 5109-17.

[20]   Garden R. W., Sweedler J. V., Heterogeneity within MALDI samples as revealed by mass spectrometric imaging. *Anal. Chem.* 2000, 72, 30-36.

[21]   Schuerenberg, M., Luebbert, C., Eickhoff, H., Kalkum, M., et al., Prestructured MALDI-MS sample supports. *Anal. Chem.* 2000, 72, 3436-3442.

[22]   Johnson, T., Berquist, J., Ekman, R., Nordhoff, E., et al., A CE-MALDI interface based on the use of prestructured sample supports. *Anal. Chem.* 2001, 73 1670-1675.

[23]   Stone, K. L., LoPresti, M. B., Crawford, J. M., DeAngelis, R. Williams, K. R. A Practical Guide to Protein and Peptide Purification for Microsequencing. (Matsudaira, P., ed.) Academic Press N Y 1989, pp. 33-47.

[24]   Doherty, N. S., Littman, B. H., Reilly, K., Swindell, A. C., Buss, J. M.,et al., Analysis of changes in acute-phase plasma proteins in an acute inflammatory response and in rheumatoid arthritis using two-dimensional gel electrophoresis. *Electrophoresis* 1998, 19, 355-63.

[25]   Ferrari, G., Garaguso, I., Adu-Bobie, J., Doro, F., et al., Outer membrane vesicles from group B Neisseria meningitidis delta gna33 mutant: proteomic and immunological comparison with detergent-derived outer membrane vesicles. *Proteomics* 2006, 6, 1856–1866.

[26]   Jensen, O. N., Wilm, M., Shevchenko A., Mann, A., Peptide sequencing of 2-DE gel-isolated proteins by nanoelectrospray tandem mass spectrometry. *Methods Mol.* Biol. 1999, 112, 571-588.

[27]   Wilm, M., A. Shevchenko, T. Houthaeve, S. Breit, L. et al., Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* 1996, 379, 466–469.

[28]   Nordhoff, E., Schürenberg, M., Thiele, G., Lübberta, C., et al., Sample preparation protocols for MALDI-MS of peptides and oligonucleotides using prestructured sample supports *Int. J of Mass Spectrom.* 2003, 226, 163–180.

[29]   Perkins, D. N., Pappin, D. J., Creasy, D. M., Cottrell J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, 20, 3551–3567.

[30]   Gobom, J., Nordhoff, E., Mirgorodskaya, E., Ekman, R., Roepstorff, P., Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* 1999, 34, 105-116.

[31]   Gusev, A. I., Wilkinson, W. R., Proctor, A., Hercules, D. M., Improvement of signal reproducibility and matrix/comatrix effects in MALDI analysis *Anal. Chem.* 1995, 67, 1034–1041.

[32]   Horneffer, V., Dreisewerd, K., Ludemann, H. C., Hillenkamp, F., Lage, M., Strupat, K., Is the incorporation of analytes into matrix crystals a prerequisite for matrix-assisted laser desorption/ionization mass spectrometry? A study of five positional isomers of dihydroxybenzoic acid. *Int. J. Mass Spectrom.* 1999, 187, 859-870.

[33]   Kruger, R., Pfenninger, A., Fournier, I., Gluckmann, M., Karas, M., Analyte incorporation and ionization in matrix-assisted laser desorption/ionization visualized by pH indicator molecular probes. *Anal. Chem.* 2001, 73, 5812-5821.

[34]   Cordero, M. M., Cornish, T. J., Lys, I. A., Cotter, R. J., Sequencing peptides without scanning the reflectron: post-source decay with a curved-field reflectron time-of-flight mass spectrometer. *Rapid Commun. Mass Spectrom*. 1995, 9, 1356-1361.

[35]   Karas, M., Bahr, U., Strupat, K., Hillenkamp, F., et al., Matrix Dependence of Metastable Fragmentation of Glycoproteins in MALDI TOF Mass Spectrometry *Anal. Chem.* 1995, 67, 675-679.

[36]   Bagshaw, R.D., Callahan, J.W., Mahuran, D.J., Desalting of in-gel-digested protein sample with mini-C18 columns for matrix-assisted laser desorption ionization time of flight peptide mass fingerprinting. *Anal Biochem.* 2000, 284, 432–435.

[37]   Stensballe, A., Andersen, S., Jensen, O. N., Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* 2001, 1, 207-222.

Chapter II
Development of a new method for detection and molecular characterization of DNA alkylating agents by MALDI-TOF mass spectrometry

PAH-DNA adducts

All organisms are exogenously exposed to varying amounts of a broad range of foreign substances, generally defined xenobiotica. Although most of the xenobiotica are innocuous compounds of natural source, some chemicals demonstrated to have carcinogenic or mutagenic properties in experimental systems.

The modern research on chemical carcinogenesis began with the isolation of benzo[a]pyrene (B[a]P) from coal tar in 1930 by Cook and colleagues, and the subsequent demonstration that B[a]P initiates tumors when repeatedly painted on mouse skin [171]. Cook and colleagues presented the first proof that polycyclic aromatic hydrocarbons (PAH) are capable to induce cancer, a milestone of the modern study on chemical carcinogenesis.

The PAH are a group of highly lypophilic organic compounds containing two or more condensed benzene rings representing a wide range of molecular size and structural complexity (Figure 12). PAH are primarily formed by the incomplete combustion of organic materials, although they can also be generated following the preparation of food that has been charred or grilled [172,173]. Common sources of PAH carcinogenic to humans include automobile exhaust emissions, tobacco smoke, and coal tar [173-175]. The Monographs Program on the Evaluation of Carcinogenic Risks to Humans of the International Agency for Research on Cancer (IARC) publishes authoritative carcinogenic risk assessments [9]. The IARC Monographs identify chemicals, mixtures, occupational exposures, physical and biological agents, lifestyle and environmental factors that can increase the risk of human cancer. Nearly all of these risks were first identified through observational epidemiology, and then verified by supplementary studies in animals and other experimental systems. Therefore, tested chemicals have been classified into 5 groups based on their potential to cause cancer in humans. To date, these evaluations have identified a total of 107 agents, mixtures and exposures that are classified in group 1: "carcinogenic to humans".
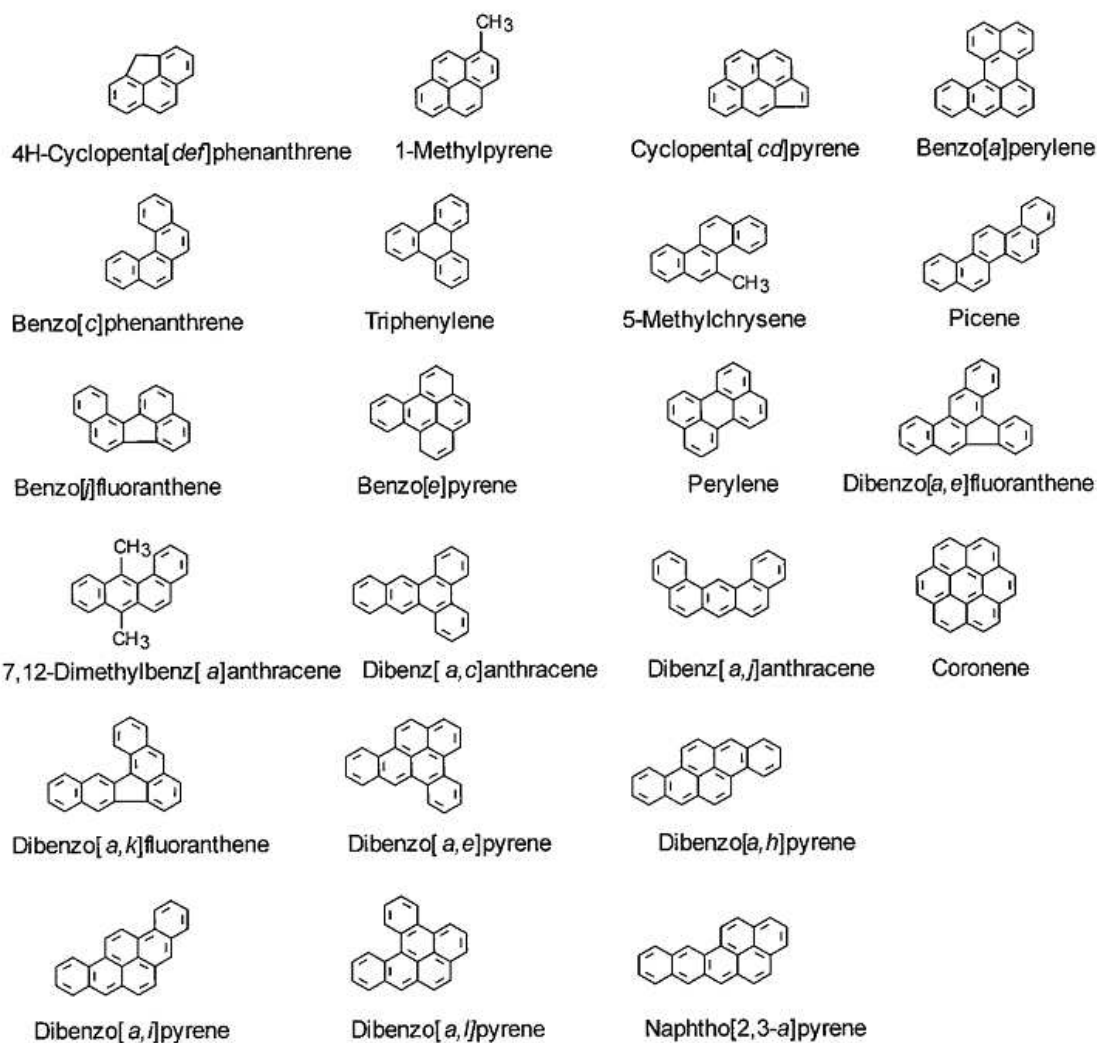
Figure 12. Chemical structures of some polycyclic aromatic hydrocarbons (PAHs). From ref. [176].

The B[a]P is included in this group 1 and most PAH are included in the first three group of risk classification. Numerous animal studies involving dermal, intraperitoneal, and intratracheal routes of exposure as well as exposure through the diet have confirmed the potent carcinogenic properties of PAH [13,177]. The impact of PAH on human health has been a great topic of concern due to their universal presence in environmental pollution and certain occupational settings [12,174,178,179]. Evidence linking PAH to the induction of carcinogenesis in humans points to an increased risk of lung, skin, and bladder cancer following environmental and occupational exposure [10].

While PAH are generally not carcinogenic *per se*, metabolic activation leads to intermediates that might react with nucleic acid and protein leading to DNA, RNA and proteins adducts. Most of the DNA adducts formed in the genome are normally removed by the DNA repair mechanisms. However, unrepaired DNA adducts can potentially lead to genetic mutation when the genome is replicated. Presently, there are three principal metabolic pathways for PAHs that have been proposed and established with

experimental evidence. A pathway proposed more than 30 years ago suggests the activation of PAH by the formation of radical cations by removal of one electron from the π electron system of the molecule through one-electron oxidation catalyzed by P450 peroxidase [180]. Radical cations are electrophilic capable of interacting with nucleophilic centers in cellular macromolecules [181,182]. DNA adducts formed through PAH radical cation are unstable leading to spontaneous depurination which would result in formation of apurinic sites as the major type of DNA damage [183,184]. In the late 1980s a pathway of PAH activation that involves the formation of *o*-quinones catalyzed by dihydrodiol dehydrogenases has been postulated [185,186]. Evidence has revealed that PAH-*o*-quinones are highly reactive Michael acceptors which can form both stable and depurinating DNA adducts [187]. Notably, the theory of bay region dihydrodiol epoxide (DE), well established through studies over the past half century, is widely accepted by scientists in this field as the dominant mechanism of PAH metabolic activation [188-190]. This mechanism involves three enzyme-mediated reactions as (Figure 13).
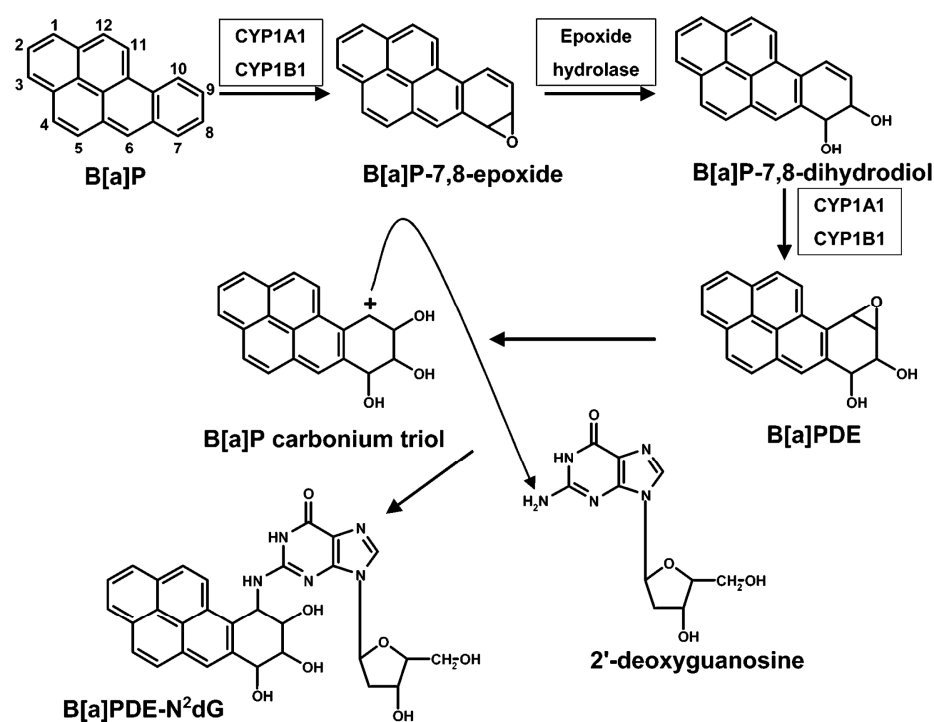
Figure 13. Metabolic activation of B[a]P and the formation of B[a]PDE-dG. From ref.[191].

Firstly, the oxidation of a double bond on the PAH catalyzed by P450 enzymes (CYP), i. e. CYP1A1 and CYP1B1, to form unstable arene oxides. Secondly, the hydrolysis of the arene oxides by microsomal epoxide hydrolase (EH) to trans dihydrodiols. Finally, a second CYP catalyzed oxidation at the double bond adjacent to the diol function to generate a vicinal DE. This pathway can lead to sterically hindered

"bay" or "fjord" region DE. The bay and/or fjord region DE are electrophiles capable of binding to DNA while some of the DE stereoisomers of PAHs are found to be ultimate carcinogens [192,193]. Two pairs of enantiomeric DE may result from the reaction, although the ultimate carcinogen-DNA adducts that are observed appear to indicate high stereoselectivity.

These DE metabolites may covalently bind via a cis or trans addition of the exocyclic amino groups of purines and pyrimidines. However, the ability of adducts to bind to DNA and persist to form mutations may depend on the conformation of the adduct, the DNA sequence and the efficiency of repair enzymes in identify the lesion [194]. Notably, some DNA adducts derived from B[a]P can be repaired by a nucleotide excision repair pathway [195,196].

Since the first report of the quantization of a DNA adduct in 1969 [197], an extensive amount of analytical work has been done to establish the relevance of DNA adducts as important biomarkers for cancer development. While some compounds such as aminobiphenyl have been closely associated with the onset of cigarette smoking-induced cancer [11,198], the complexity of environmental exposure mechanisms complicates the establishment of a direct causal link of compound such as PAH. Notably, B[a]P has been shown to cause nonrandom mutational hotspots in the p53 tumor suppressor gene and ras proto-oncogene resulting in predominantly G:C to T:A transversions [199,200].

PAH-DNA adducts detection and characterization Methods

A number of methods have been developed for the detection of PAH-derived DNA adducts, which are dictated by the inherent chemical and physical properties of these molecules. These methods can be distinguished in indirect and direct methods. With indirect method the detection of PAH is made on the compounds released from the DNA following acid hydrolysis [201]. The direct method allow the detection of PAH bond to DNA or nucleotides. Therefore, the direct methods were preferred because in most cases they permitted the determination of the part of DNA involved in the adduction event. These methods include high performance liquid chromatography (HPLC) separation technique coupled to fluorescence detection, or NMR, circular dichroism, UV/visible fluorescence spectroscopy, synchronous fluorescence spectroscopy and gas chromatography [202-207]. The production of antibodies recognizing PAH-DNA adducts has led to the development of also immunoaffinity and immunological based methods [203,208]. The most common method used for the sensitive detection of DNA adducts is the ^{32}P-postlabeling assay.Nonetheless, methods based on mass spectrometry were recently developed. The main advantage of MS based methods is that it not only provides information on the molecular mass of the DNA adducts, but can also generate structurally significant information for its characterization. In the

following paragraphs recent advances of methods for the detection and structural characterization are discussed in detail together with the advancement added by the work of this doctoral thesis.

^{32}P-postlabeling assay

^{32}P-Postlabeling analysis has been widely used for the detection of a variety of DNA adducts induced by endogenous and exogenous mutagens or carcinogens [209-213]. The principal stages of the ^{32}Ppostlabeling assay are the digestion of DNA to nucleoside 3'-monophosphates with micrococcus nuclease and calf spleen phosphodiesterase II. This is followed by an enrichment step to enhance the sensitivity, which involved extraction of the aromatic/hydrophobic adducts into butanol as a means of separating them from the un-adducted normal nucleotides. The extracted adducted nucleotides are 5'-labeled with ^{32}P-orthophosphate from [γ- ^{32}P]-ATP by the T4 polynucleotide kinase catalyzed reaction. Finally, chromatographic and/or electrophoretic separations of the labeled species are performed followed by their detection and quantitation. Indeed, the [^{32}P]-5'-labeled nucleotides adducts are commonly separated by multidimensional thin-layer chromatography (TLC) on polyethyleneimine (PEI)-cellulose plates using several different buffer conditions [210,211]. Alternatively, the use of non-denaturing polyacrylamide gel electrophoresis (PAGE) has been adapted [214]. The adducts were detected using autoradiography or by storage phosphor imaging coupled with scanning densitometry to quantify the radioactivity on the chromatograms [215,216]. This method is widely used due to the small amounts of the DNA required and its high sensitivity, i.e. one adduct per 10^6-10^8 non-adducted nucleotides (5µg DNA) can be detected [210,211]. However, the method require the handling of hazardous radioactive phosphorus, and is laborious with significant inter- and intra-laboratory variations [213]. The major drawbacks of the ^{32}P-postlabeling assay are that the chemical nature and molecular structure of the adducing agent can not be determined, and co-chromatography studies with DNA adducts reference materials resulted in erroneous information as recently reported [217].

ESI tandem MS based methods

Recently, methods for the detection of PAH-DNA adducts based on the use of LC coupled with ESI-MS were developed. The general strategy (Figure 14) used in these approaches is based on the enzymatic digestion of the adducted DNA to single nucleotides/nucleosides. A subsequent solid phase extraction (SPE) as pre-purification step is used prior to HPLC separations using several detection systems (e. g. fluorescence, UV). Finally, the adducts are characterized by ESI-MS using a number of different tandem MS fragmentation techniques [191]. Moreover, direct reaction products of purified bioactivated form of PAH with DNA or single deoxy-nucleotides were used as standard reference compounds.
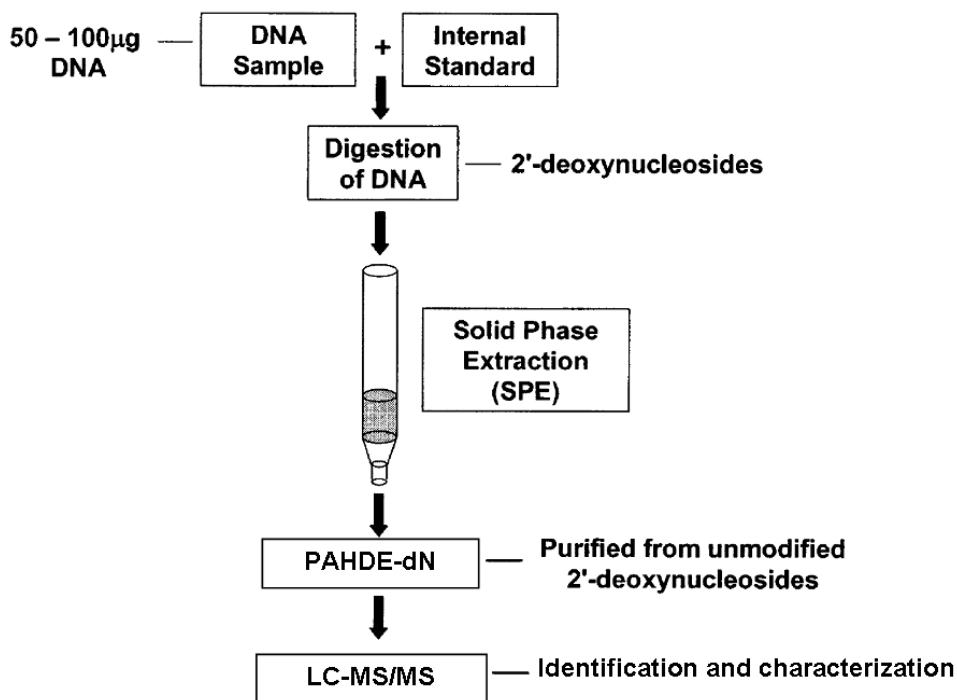
Figure 14. General procedure used for detection of PAHDE-deoxynucleotides adducts in DNA Samples using LC-MS/MS. dN= generic 2'-deoxynucleoside. Adapted from ref. [191]

LC-MS/MS methods using ESI ionization and SRM have been reported for the detection, characterization and quantification of several PAH-2'-deoxyguanosine (dG) adducts of the dietary mutagen, 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine (PhIP) [218] and for B[a]P [191]. The adducts were initially characterized in calf thymus DNA modified *in-vitro* with bioactivated form of PAH. Moreover, the adducts were detected in the colon and liver of animals treated with PAH. After, enzymatic digestion of the DNA samples to 2'-deoxynucleosides, and SPE purification to remove unmodified 2'-deoxynucleosides the analysis was followed by LC-ESI-MS/MS with triple quadrupole (TQ) operated in the SRM scan mode. In this mode, the precursorions (the protonated adducts molecular ion $[M + H]^+$) are selectively transmitted by the first quadrupole mass analyzer (Q1) and are subjected to CID in the second quadrupole (Q2). These collision conditions generally result in the loss of deoxyribose $[M+H-116]^+$ to form the protonated base adduct ion $[BH_2]^+$, which are selectively transmitted through the third quadrupole (Q3). In these repots limits of detection about 3 adducts per 10^8 2'-deoxynucleosides were determined by measurements based on isotope dilution with internal standards. However, the full product ion scan mode, used in these LC-ESI-MS/MS analysis, can not be employed for the characterization of unknown DNA adducts at trace level, because of the slow scan rate and the low sensitivity in such scan mode. Alternatively, the analysis rely on the characteristic retention time, but searches based on SRM, do not permit *de novo* identification of DNA adducts. In addition, the

needs of stable isotope internal standards further limit the wide applicability of the methods. Notably, despite the implementation with multi-stage scan events (MS[n]), which can eliminate isobaric interferences, with these approaches only the adducts on deoxyguanosine were detected [191,218]. Moreover, Wang *et. al.* reported the use of nanoflow liquid chromatography coupled to ESI-hybrid quadrupole-TOF MS using CID fragmentation for the detection of DNA adducts from lung and urine of mice expose of asphalt fume and *in-vitro* [219-221]. Notably, Gaskell *et. al.* reported along with the main base adducts the identification and characterization of phosphodiester adducts as minor product formed following reaction of B[a]PDE with 2'-deoxynucleotides *in-vitro* [222]. In addition, applications of capillary zone electrophoresis-ESI-MS were also described, leading to detect B[a]PDE of 2'-deoxyguanosine together with adducts formed with 2'-deoxyadenosine and 2'-deoxycytidine [223,224]. A particular approach reported by Feng and colleagues implemented the LC-MS/MS with diode array and fluorescence detection . In this way they were able to separate the four stereoisomeric anti-BPDE-dG and two interfering anti-BPDE tetrols by taking advantage of their distinct fluorescence quenching [225,226].
It should be noted that relevant improvements were reported with these approaches, such as the detection of many different adducts and the feasibility of a molecular characterization of some of them. However, the major drawbacks of these approaches are the need of large quantities of starting material and time consuming pre-purification steps as well as HPLC separation procedures required.

MALDI MS

There few reports on the application of MALDI MS for the analysis of DNA adducts, mostly limited to targeted analysis. Some of these reports should be considered as feasibility studies of the application of MALDI MS for DNA adducts analysis [227]. Stemmler *et. al.* reported the use of MALDI-FTICR MS to test several different matrices for the structural characterization of synthesized PAH-deoxynuxcleotides adduct [228]. Chiarelli and colleagues reported the use of MALDI for the analysis of arylamine adducts [229] and for the structural differentiation of diastereoisomers of PAH-deoxyadenosine adducts [230]. In several investigations the analysis was performed on oligonucleotide rather then single nucleotide adducts [231-234], probably due to the lack of reliable matrices and sample preparation methods for an analysis in the low mass range (<1000 Da). Recently, the detection of B[a]PDE and the dietary mutagen PhIP - deoxyguanosine adducts by MALDI with CID fragmentation has been also reported [235,236].
However, these repots were unable to demonstrate applicability to wide range of compound. Moreover, the analysis relied on the detection of the molecular ion and just few characteristic fragments. In addition, large amounts of sample were needed, and in most cases the matrix interfered with a reliable detection do to matrix cluster formation.

Therefore, efforts were invested to develop a new robust and sensitive method for the detection and characterization of PAH-DNA by MALDI-TOF MS.

A MALDI-TOF MS based method for the rapid detection and molecular characterization of DNA Alkylating Agents (Publication II)

This study aimed of the development of a rapid and sensitive method for the detection, identification and molecular characterization of PAH-DNA adducts by MALDI-TOF-MS

The strategy relied in the synthesis of standards molecules of metabolic activated forms (dihydrodiol epoxide, DE) of PAH-DNA adducts to be used as model analytes. The synthesized PAHDE-DNA adducts were used to develop a wide applicable analytical method. Moreover, their mass spectra were exploited to generate a repository of references MS and tandem MS spectra to be used for the identification and characterization of unknown adducts.

Initially, single nucleotides and nucleosides were reacted with several PAHDE. Therefore, adducts were synthesized by direct reaction of the four 2'-deoxynucleosides and the four 2'-deoxynucleosides-3'-monophosphates with *anti*-7,8 Dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene (B[a]PDE), *anti*-11,12-dihydroxi-13,14-epoxy-9,10,11,12.tetrahydrobenzo[a,l]pyrene (B[al]PDE), *anti*-11,12,dihydroxy-13,14-epoxy-11,12,13,14-tetrahydrobenzo[g]chrysene (B[g]ChDE), *anti*-3,4-dihydroxy-1,2-Epoxy-1,2,3,4-tetrahydrobenzo[c]phenanthrene (B[c]PhDE) and with *anti*-9,10-dihydroxy-11,12-epoxy-9,10,11,12-tetrahydrobenzo[c]chrysene (B[c]ChDE) (the general abbreviation of PAHDE for these PAH activated form will be used herein).

The subsequent step was to develop a reliable MALDI matrix sample preparation for analysis of such molecules. Several matrix and sample preparations were compared on basis of production of ions, signal reproducibility, degree of analyte fragmentation and sensitivity. The DHB ML (publication I) "doped" with ammonium phosphate demonstrated to be the method of choice for an analysis of PAHDE and PAHDE-DNA adducts by MALDI-TOF MS.

However, to obtain high quality MALDI-MS spectra sample purity is of great importance. Therefore, several enrichment methods were tested using commercial available solid phase extraction (SPE) cartridges devices with different sorbent bed materials. These devices demonstrated large differences in performance depending on the resin used as well as needed large amount and volume (milliliters) of sample. In some cases, adsorption of the analytes on the device components was evidenced. The N-vinylpyrrolidone/divinylbenzene (hydrophilic-lypophilic balance, HLB) copolymer sorbent demonstrated to be valuable enabling increased recovery and enhanced loading capacity as compared to other resins.

In addition, we implemented the HLB sorbent and developed self-assembled disposable micro-columns (HLB-µ-SPE, see experimental section in publication II) to meet the sample amount and the volume (micro liters) used for MALDI-MS sample preparation. The micro-columns were prepared in pipette tip, thus, micro to nano grams amount in micro liters volumes of sample could be handled.

An additional issue is that the PAHDE-DNA adducts have masses under 1000 Da, thus, in the low mass range (LMR). Unfortunately, for MALDI-TOF MS in the LMR there were no available instrument settings for spectra acquisition and, most importantly, methods for spectra calibration.

Development of a calibration method

For applications in MALDI-TOF MS the availability of methods for calibrate the spectra is mandatory for accurate measurements. Commercially available calibration mixtures are designed for MALDI spectra calibration in the mid-mass range (from 700 to 4000 Da). For this reason, some investigators reported the use of purified mixture of analyte-related compound as calibrators [237,238]. Unfortunately, not always purified molecules are available and these should be prepared new each application. To provide a reliable wide application a peptide based calibration mixture was developed. The calibration mixture will be used for the external calibration of spectra in the LMR. The calibration mixture is composed of 8 peptides with masses between 147 Da and 913 Da. Notably, the mixture was optimized to in order to be compatible of combination with a commercial calibration mixture for the mid-mass range (Peptide calibration standard II, Part-No. #222570, Bruker Daltonics, Bremen, Germany). Indeed, the Peptide calibration standard II covers a range between 750-3400 Da, whereas, the new developed calibration mixture covers a range of 140- 910 Da. Therefore, the combined calibration mixture allows the calibration of MALDI MS spectra form the low to the mid-mass range (from 100 to 4000 Da).

Figure 15 shows a typical MALDI-TOF mass spectrum of the peptides calibration mixture in the mass range 10-1200 *m/z*, prepared using the DHB ML sample preparation. The spectrum shows a notable absence of matrix cluster ions and few weak cationized ions which allows to a simple interpretation and optimal calibration.
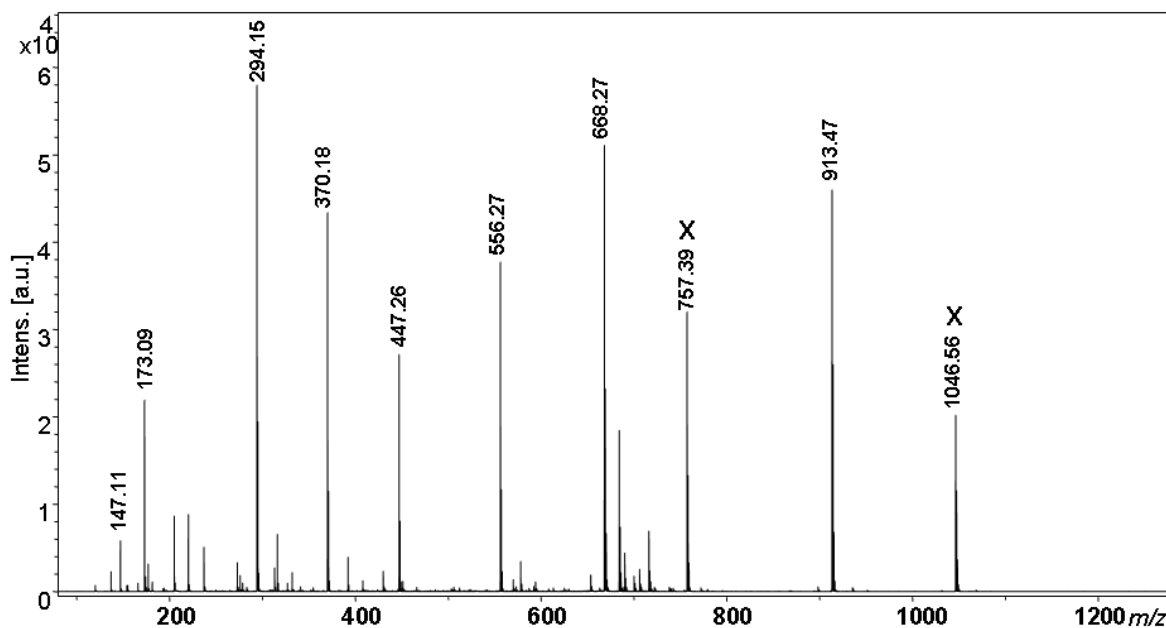
Figure 15. Typical positive ion MALDI-TOF mass spectrum of the developed calibration mixture acquired on an Ultraflex II mass spectrometer (Bruker Daltonics). X= peptide from the commercial peptide mixture which fallow in this mass range.

Reference spectra

The reaction mixtures of the four 2'-deoxynucleosides and the four 2'-deoxynucleosides-3'-monophosphates were subjected to HLB-μ-SPE sample cleanup (see experimental section). Then, aliquots of the purified analytes were loaded on MALDI target with the DHB ML sample preparation protocol described above. Using the calibration mixture developed for external calibration, a mass accuracy below 20 ppm over the whole mass range (10–1200 m/z) was obtained. Notably the accuracy was <10 ppm for the molecular ions ([M+H]$^+$) at a resolution of >5000 (FWHM). Based on dilution series of a 32 μM stock solution of purified B[a]PDE-2'-deoxyguanosine-3'-monophosphate (dGp) a limit of detection (LOD) <100 fmol loaded on target (65 pg, S/N> 3) was determined.

The MALDI-TOF MS spectra of PAHDE-DNA adducts together with the molecular ion ([M+H]$^+$) provide typical mass-signature specific for the alkylating molecule and the nucleotide involved in the alkylation reaction. Such fingerprints can thus be defined as DNA-adduct mass fingerprint (DMF). Based on the DMF, the [M+H]$^+$ and specific ions were selected for MALDI-TOF-CID-MS/MS measurements to obtain definitive structural diagnostic fragment ions, i.e DNA-adducts fragment fingerprint (DFF). Therefore, the DMF and DFF of synthesized PAH-DNA adducts were acquired to be use as reference MS and tandem MS spectra.

Determination of PAHDE-DNA adducts in calf thymus DNA

The robustness and reliability of the developed methodology was demonstrated by the application for the analysis of adducts formed with calf thymus DNA upon reaction with B[a]PDE and B[c]ChDE. Based on the full MALDI-TOF mass spectrum, a simple comparison with the reference DMF allowed specification of the molecular ion and characteristic fragment ions of the B[a]PDE-DNA adducts. Readily molecular ions and characteristic fragment for deoxynucleotides and deoxynucleosides at the same time were detected. The MALDI-TOF and –CID-MS/MS spectra yielded characteristic fragmentation patterns that allow identification of deoxyguanosine, deoxyadenosine and deoxycytidine DNA adducts. The spectra are simple to interpret and there is a notable absence of cationized sample and matrix ions. Notably, additional adducts as dinucleotide could be detected, identified and characterized by MALDI tandem MS.

Overall the method is sensitive and nanogramms of hydrolyzed DNA are sufficient for the identification and characterization of adducts.

When compared with the ^{32}P-postlabeling assay a distinct advantage of the method is the unambiguous identification of DNA adducts and the possibility to identify the chemical nature of the alkylating agent. Moreover, the developed method allowed the simultaneous and unambiguous detection and identification of deoxynucleotide and deoxynucleosides and dinucleotides PAHDE adducts at the same time, leading to a more comprehensive analysis. In addition, the analysis time is drastically reduced and a dedicated radioactivity laboratory is not required.

Appendix Chapter II

Garaguso I, Halter R, Krzeminski J, Amin S, Borlak J. Method for the Rapid Detection and Molecular Characterization of DNA Alkylating Agents by MALDI-TOF Mass Spectrometry. Anal. Chem., 2010, 82 (20), pp 8573–8582, Publication Date (Web): September 24, 2010.

The following is the pre-peer reviewed version of the article:

Method for the Rapid Detection and Molecular Characterization of DNA Alkylating Agents by MALDI-TOF Mass Spectrometry.

Web: http://pubs.acs.org/doi/abs/10.1021/ac101568h

Author contributions:
IG: conceived the study, developed the methods and carried out the experimental work: sythesized the PAH-nucleo-tides/-sides adducts, tested the matrices, developed the MALDI matrix sample preparation method, developed the enrichment/purification method, developed the calibration method and the MALDI instrumental setting up in the low mass range, performed the MALDI-TOF analysis and interpreted the results. Drafted and wrote the manuscript.
HR: performed the ^{32}P-Postlabeling assay.
KJ and AS: synthesized the PAHDE molecules.
JB: participated in the coordination and reviewed the manuscript.

Method for the Rapid Detection and Molecular

Characterization of DNA Alkylating Agents

by MALDI-TOF Mass Spectrometry

*Ignazio Garaguso[1,2], Roman Halter[1], Jacek Krzeminski[3], Shantu Amin[3] and Jürgen Borlak [1,2] **

[1]Department of Drug Research and Medical Biotechnology, Fraunhofer Institute of Toxicology and

Experimental Medicine, Hannover, Germany

[2]Center for Pharmacology and Toxicology, Hannover Medical School, Hannover, Germany

[3]Department of Pharmacology, Penn State College of Medicine, Hershey, PA 17033, USA

***Corresponding author:**

 Prof. Dr. J. Borlak, Department of Drug Research and Medical Biotechnology Fraunhofer Institute of

Toxicology and Experimental Medicine, Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany.

Phone: +49-511-5350-559, Fax +49-511-5350-573, E-mail: borlak@item.fraunhofer.de.

List of abbreviations:

PAH: polycyclic aromatic hydrocarbons; B[a]PDE: (±)-*anti*-benzo[a]pyrene-7,8-diol-9,10-epoxide;

B[c]ChDE: (±)-*anti*-benzo[c]chrysene-9,10-diol-11,12-epoxide () 2,5-dihydroxybenzoic acid: DHB; matrix

layer: ML.

Keywords:

Polycyclic Aromatic Hydrocarbons, DNA Adducts, Carcinogens, MALDI-TOF/TOF-MS; 2,5-

dihydroxybenzoic acid; Nucleosides, Nucleotides, ^{32}P-postlabeling, benzo[a]pyrene-7,8-diol-9,10-

epoxide, benzo[c]chrysene-9,10-diol-11,12-epoxide.

ABSTRACT

Metabolic activation of polycyclic aromatic hydrocarbons (PAH) may cause DNA adduct formation. While these are commonly detected by the ^{32}P-postlabeling assay, this method is not informative on the chemical nature of the alkylating agent. Here we report a simple and reliable method that employs MALDI-TOF-MS with 2,5-dihydroxybenzoic acid (DHB) matrix layer (ML) sample preparations for the detection and structural characterization of PAH-DNA adducts. The method involves the enzymatic digestion of DNA to 2'-deoxynucleotides followed by solid phase extraction to remove salt and other contaminants prior to MALDI-MS analysis. By collision induced dissociation (CID) structurally relevant fragments are obtained to permit characterization of the alkylating molecules and the adducted nucleotide. Next to guanosine, adenosine and cytidine adducts formed from reactions with (±)-*anti*-benzo[a]pyrene-7,8-diol-9,10-epoxide (B[a]PDE) are identified at a sensitivity of <100 fmol and a mass accuracy of <10ppm. Studies with (±)-*anti*-benzo[c]chrysene-9,10-diol-11,12-epoxide (B[c]ChDE) further document the versatility and usefulness of the method. When compared with the ^{32}P-postlabeling assay MALDI-MS only indentified deoxycytidine as well nucleoside and dinucleotides adducts. Therefore, this sensitive method enables molecular specification and characterization of adducted nucleotides and of the alkylating agent, and thus, provides comprehensive information that is beyond the ^{32}P-postlabeling assay.

Introduction

Polycyclic aromatic hydrocarbons (PAH) are ubiquitous environmental pollutants and are primarily formed by incomplete combustion of organic matter, such as automobile exhaust emissions, tobacco smoke, coal tar, but are also generated by cooking of food that has been charred or grilled. While PAH are not carcinogenic *per se*, metabolic activation leads to intermediates that might react with nucleic acid and proteins leading to DNA, RNA and protein adducts. Estimates suggest up to 50 000 DNA damage events per day. While DNA repair enzymes remove DNA adducts efficiently, non-repaired lesion may cause mutations to initiate malignant transformations[1].Thus, exposure to PAH has been linked to lung, skin, and bladder cancer[2-4]. Furthermore, in molecular epidemiological studies DNA adducts are used to evidence exposure to hazardous chemicals and to estimate risks in particular work places or environments[5;6]. Notably, PAH are activated by three major mechanisms involving either one-electron oxidation to form intermediate radical cations, a reactive *o*-quinone and the oxidation to dihydrodiol epoxide (DE) intermediates[7]. The DE formation requires an initial metabolic activation of PAH, which is efficiently catalyzed by the combined action of cytochrome P450 1A1 and 1B1 resulting in the oxidation of a double bond to yield an epoxide. The epoxide is then converted to trans dihydrodiol by microsomal epoxide hydrolase, which undergoes biotransformation by cytochrome P450 monooxygenase resulting in an oxidation at the adjacent double bond to generate the ultimate reactive DE species[8]. Such metabolic activation may lead to sterically hindered bay or fjord PAHDE, which are electrophiles, that covalently react by cis or trans addition with the exocyclic amino group of the DNA bases. Studies on PAH metabolism, DNA binding, mutagenicity and cell transformation assay demonstrated that PAH such as benzo[a]pyrene (B[a]P), chrysene, Benzo[c]chrysene (B[c]Ch), benzo[c]phenantrene (B[c]Ph), 5-methylchrysene, dibenzo[a,l]pyrene are activated by this pathway[7;9] and react with amino groups of the purine bases as preferential alkylating sites[10]. Typical targets in DNA are the N^2 exocyclic amino group of guanine, the N^6 of adenine and to a lesser extend the N^4 exocyclic amino group of cytosine[11].The molecular structure of PAH is an important determinant of their biological activities[12], as the metabolically activated epoxides located either in a bay or fiord region display different carcinogenic properties. For instance the tumorigenic activity of fjord region PAHDE is remarkably higher then of the structurally related bay region PAHDE[7;13]. While guanine residues are the principal site of reaction for PAHDE derived from planar hydrocarbons such as B[a]P[14], adenine residues are targeted as well or more effectively than guanine residues when the reactive metabolite is derived from a nonplanar hydrocarbon such as B[c]Ph[15].

Much research has been invested for an identification of PAH-DNA adducts and included HPLC separation techniques coupled with ultraviolet, or NMR, circular dichroism (CD), UV/visible, fluorescence spectroscopy and immunoassays[16-19]. However, the most common method used for the

sensitive detection of DNA adducts is the ^{32}P-postlabeling assay. This method involves the reduction of genomic DNA to single deoxynucleosides-3'-monophosphate by enzymatic hydrolysis. Then, the enzymatic transfer of radiolabeled phosphate groups from [^{32}P] ATP to nucleotides, followed by the multidimensional thin layer chromatography (TLC) allows the separation and detection of modified nucleotides. This method is widely used, in part, due to the small amounts of the DNA required and its high sensitivity, i.e. one adduct per 10^6-10^8 non-adducted nucleotides (5μg DNA)[20;21] can be detected, but the method is laborious with significant inter- and intra-laboratory variations[22]. Unfortunately, the chemical nature and molecular structure of the alkylating agent can not be analyzed by this method and co-chromatographic studies with DNA adduct reference materials resulted in erroneous information as recently reported[23]. There is need for a method of high sensitivity that permits chemical identification and characterization of the DNA adducts at the same time. In this regard, mass spectrometry (MS) based methods were developed for the study of PAH-DNA adducts, and are based on liquid chromatography (LC), capillary LC or capillary electrophoresis (CE) coupled to Electrospray ionization (ESI) tandem mass spectrometry (MS/MS)[24-26]. For example, LC-ESI/MS/MS has been used for the structural characterization of B[a]P-derived DNA adducts formed upon the reaction of B[a]PDE with naked DNA, or cell lines and tissue extract of mice that have been exposed to alkylating agents[27-29]. Nonetheless, these methodologies have some limitations, e.g. large amounts of sample requirement, extensive liquid–liquid extraction, time consuming HPLC or chromatographic purification, as well as expensive instrumentation. Furthermore, the full product ion scan mode, used in LC-ESI-triple quadrupole-MS/MS analysis, can not be employed for the characterization of unknown DNA adducts at trace level, because of the slow scan rate and the low sensitivity in such scan mode. Alternatively, an analysis rely on the characteristic retention time, but searches based on selected reaction monitoring (SRM), do not permit *de novo* identification of DNA adducts[30]. Taking all this into account, the availability of a sensitive, accurate and reliable methodology for the molecular characterization would be of great interest. To this end, MALDI-TOF-MS, as introduced by Karas and Hillenkamp in 1987[31], has rapidly become a valuable technique for the analysis of a wide range of molecules with sensitivities in the attomole range[32;33].

We therefore report the development of a robust and simple method for the detection and characterization of PAHDE-DNA adducts based on MALDI-MS. We employ micro solid phase extraction (μ-SPE) to remove efficiently salt and other contaminants of the PAHDE-DNA reaction products. Then MALDI-MS spectra are recorded to obtain DNA-adduct mass fingerprint (DMF). Subsequent, MS/MS-CID of selected DMF ions produce diagnostic fragments that permit an identification and molecular characterization of PAHDE and the nucleotide involved in the alkylation, i.e. DNA-adducts fragment fingerprint (DFF). The usefulness of the method is afforded by characterization of adducts formed *in*

vitro with calf thymus DNA upon reaction with several different PAHDE. Finally, the MALDI-MS method is also compared with the ^{32}P-postlabeling assay in regards to sensitivity and specificity.

Experimental section

Safety considerations

Caution: *B[a]PDE and B[c]ChDE are potential mutagenic and carcinogenic agents and must be handled with care. Protective clothing must be worn. Appropriate safety procedures should be followed when working with this compound and for discard waste materials.*

Chemicals and reagents. Acetonitrile, methanol and Water LC-MS grade were from JT Baker (Griesheim, Germany). T4 Polynucleotide kinase was from USB Corporation (Cleveland, OH). Phosphodiesterase II from bovine spleen (SPDE), micrococcal nuclease from *Staphylococcus aureus* (MN) and all other chemicals were of high-purity grade purchased from Sigma Chemical Co. (St. Louis, MO).

Synthesis of PAH-diol-epoxides. The (±)-*anti*-7,8-dihydroxy-9,10-epoxy-7,8,9,10-tetrahydrobenzo[a]pyrene (B[a]PDE) was synthesized according to the protocol given by Yagi *et al.*[34] The (±)-*anti*-9,10-dihydroxy-11,12-epoxy-9,10,11,12-tetrahydrobenzo[c]chrysene (B[c]ChDE) was synthesized as described by Amin *et al.*[9]

Reactions of 2'-deoxynucleosides and 2'-deoxynucleosides-3'-monophosphate with PAHDE. Adducts were synthesized by direct reaction of the four 2'-deoxynucleosides and the four 2'-deoxynucleosides-3'-monophosphates (500 µg, 1 mg/ml dissolved in water) with B[a]PDE and with B[c]ChDE (125 µg; 1 mg/ml dissolved in tetrahydrofuran) for 18 h at 37°C in the dark. Standard solutions were generated by purification of reaction products as described below. The concentrations were determined by UV spectroscopy performed on a LAMBDA 40 spectrophotometer (Perkin Elmer ,Waltham, MA) using known extinction coefficients[35;36]

Synthesis of PAHDE modified calf thymus DNA and enzymatic hydrolysis. The calf thymus DNA adducts of individual PAHDE were synthesized as described by Amin *et al.*[18]. Details are given in the SI.

Micro-scale solid phase extraction (µ-SPE) for sample Purification. Purification of B[a]PDE-nucleosides, -nucleotides adduct and DNA hydrolysates was accomplished by hydrophobic interaction liquid chromatography using N-vinylpyrrolidone/divinylbenzene (hydrophilic-lypophilic balance, HLB) sorbent on self-assembled micro solid phase extraction (HLB-µ-SPE) disposable columns (for details see SI). Typically, 4 pmol of B[a]PDE-nucleosides, -nucleotides or 12.5 ng of DNA hydrolysates were the amount loaded onto the MALDI target. Note, representative results of at least three independent experiments are presented herein.

DNA-adduct determination by the ^{32}P-postlabeling assay. The ^{32}P-postlabeling was carried out as described by Gupta *et al.* [37] with modification given by Halter *et al.*[38]. Details are given in the SI.

Mass spectroscopic analysis. MALDI-MS experiments were performed on an Ultraflex II MALDI-TOF/TOF mass spectrometer equipped with a SmartBeam™ laser and a LIFT-MS/MS facility. Typically, 600 spectra over a 10–1200 *m/z* mass range, acquired at 100Hz, were summed and externally calibrated using a standard mixture composed 9 peptides. (instrumentation and software from Bruker Daltonics, Bremen, Germany). For details see SI.

DHB matrix layer, MALDI matrix preparation protocol. DHB matrix layer (ML) sample preparation on AnchorChip™ sample support (Bruker Daltonics) was performed as described by Garaguso & Borlak[32], with slight modifications. The DHB was prepared at a concentration of 5 g/l in 0.1 % TFA solution containing 30 % acetonitrile and 1 g/l of ammonium phosphate monobasic. The matrix solution, 0.5 µl, was deposited onto the MALDI target and allowed to dry. Subsequently, the analyte solution was deposited onto the dried matrix layer formed.

Results

MALDI-MS analysis of the B[a]PDE-2'-deoxynucleosides reaction products.

The reaction mixtures of the four 2'-deoxynucleosides and the four 2'-deoxynucleosides-3'-monophosphates were subjected to HLB-µ-SPE sample cleanup (see experimental section). Then, aliquots of the purified analytes were loaded on MALDI target with the DHB ML sample preparation protocol described above. The calculated and measured masses are given in Table S-1, and the ion assignments are consistent with the elemental compositions of the fragment ions depicted in Scheme 1. A mass accuracy below 20 ppm over the whole mass range (10–1200 *m/z*) was obtained and was <10 ppm for the molecular ions ([M+H]$^+$) at a resolution of >5000. Based on dilution series of a 32 µM stock solution of purified B[a]PDE-2'-deoxyguanosine-3'-monophosphate (dGp) a limit of detection (LOD) <100 fmol loaded on target (65 pg, S/N> 3) was determined.

The purified reaction products of B[a]PDE-2'-deoxyguanosine (dG) and dGp yielded abundant protonated molecular ions ([M+H]$^+$) at *m/z* 570 and *m/z* 650, respectively (Figure 1A, B, and S-1A, S-2A). Only weak cationized molecular ions [M+Na]$^+$ and [M+K]$^+$ were present. The aglycon fragment ion ([BH$_2$]$^+$) at *m/z* 454 is the result of the cleavage of the glycosidic bond with transfer of one hydrogen from the sugar to the base (see scheme 1). The facile loss of the neutral deoxyribose, is a characteristic mass-signature[26;29;39;40]. The next intense ion signal at *m/z* 303 corresponds to the B[a]PDE triol ([PAH]$^+$) while the ions at *m/z* 285 and *m/z* 257 correspond to a subsequent neutral losses of H$_2$O ([PAH-H$_2$O]$^+$) and CO ([PAH-H$_2$O-CO]$^+$) from the [PAH]$^+$ ion, all of which are specific of the alkylating molecule. Furthermore, specific alkylated nucleotide ions are present, i.e. the 2'-deoxyguanosine

([dG+H]$^+$) ion at *m/z* 268 and guanine ([G+H]$^+$) ion at *m/z* 152 (Figure 1A and S-1A), whereas the ion at *m/z* 348 ([dGp+H]$^+$) is specific for the 2'-deoxyguanosine-3'-monophosphate (Figure 1B and S-2A). While B[a]PDE-2'-deoxyadenosine (dA) and -2'-deoxyadenosine-3'-monophosphate (dAp) abundant protonated molecular ions at *m/z* 554 and *m/z* 634 are observed (Figure S-1B and S-2B); whereas the alkylated nucleotide resulted in intense signals: at *m/z* 252 that corresponds to [dA+H]$^+$, and the ion at *m/z* 332 that corresponds to [dAp+H]$^+$. Moreover, in both spectra specific ions for the alkylating molecule are detected together with the adenosine [A+H]$^+$ ion at *m/z* 136. The protonated molecular ions [M+H]$^+$ of B[a]PDE-2'-deoxycytidine (dC) and 2'-deoxycytidine-3'-monophosphate (dCp) reactions yielded weaker signals (Figure S-1C, S-2C) possible due to an instability of B[a]PDE-2'-deoxycytidine adducts that undergo extensive hydrolysis to B[a]P tetrols (*m/z* 320). In addition, the aglycon [BH$_2$]$^+$ is detected as weak sodiated ([BH$_2$+Na]$^+$) ion at *m/z* 436 in both spectra, while an increased abundance of the B[a]P tetrols ions and its extensive fragmentation to [PAH]$^+$, [PAH-H$_2$O]$^+$ and [PAH-H$_2$O-CO]$^+$ is observed in both spectra. Presumably, differences in electrophilicity and steric constraints of the dC and dCp compared to other deoxynucleosides may account for a low reaction efficiency and increased instability as suggested by ohers[11;26;41]. However, no adducts were detected in the reaction mixture with 2'-deoxythymidine (dT) and 2'-deoxythymidine-3'monophosphate (dTp).

MS/MS-CID analysis for characterization of B[a]PDE-2'-deoxynucleosides reaction products. The MS/MS-CID measurements involved the molecular ions [M+H]$^+$ at *m/z* 570 and *m/z* 650 for B[a]PDE-2'deoxyguanosine adducts (Figure 1C, D and S-3A, S-4A), at *m/z* 554 and *m/z* 634 for B[a]PDE-2'deoxyadenosine adducts, and at *m/z* 530 and *m/z* 610 for B[a]PDE-2'deoxycytidine adducts (Figure S-3 and S-4). Product ions corresponding to the characteristic aglycon fragment ([BH$_2$]$^+$) at *m/z* 454 for dG and dGp, at *m/z* 438 for dA and dAp, at *m/z* 414 for dC and dCp were readily detected. Here, the [dN+H]$^+$, [dNp+H]$^+$ and [N+H]$^+$ ions of the respective B[a]PDE reactions are observed that allowed the characterization of the nucleotides involved in the alkylation reaction. The ion at *m/z* 268 ([dG+H]$^+$) and at *m/z* 152 ([G+H]$^+$) characterize the dG (Figure 1C). Likewise the ion at *m/z* 348 ([dGp+H]$^+$) and at *m/z* 152 ([G+H]$^+$) characterize the dGp (Figure 1D). With 2'-deoxyadenosine, the fragments ions at *m/z* 252 ([dA+H]$^+$) and *m/z* 136 ([A+H]$^+$) characterize the dA (Figure S-3B), whereas the ions at *m/z* 332 ([dAp+H]$^+$) and *m/z* 136 ([A+H]$^+$) characterize the dAp (Figure S-4B). Notably, from 2'deoxycytidine [M+H]$^+$ ions MS/MS-CID spectra of the same quality were obtained see Figure S-3C with the characteristic ions at *m/z* 112 ([C+H]$^+$) and *m/z* 228 ([dC+H]$^+$). Furthermore, the ions at *m/z* 112 ([C+H]$^+$) and *m/z* 308 ([dC+H]$^+$) characterize the dCp (Figure S-4C). Importantly, all spectra yielded the ions [PAH]$^+$ at *m/z* 303, [PAH-H$_2$O]$^+$ at *m/z* 285 and [PAH-H$_2$O-CO]$^+$ at *m/z* 257 that allowed the characterization of the B[a]PDE alkylating molecule.

Aglycon fragmentation. The MS/MS-CID of the $[BH_2]^+$ fragment (Figure 1E) resulted firstly in loss of H_2O ($[BH_2-H_2O]^+$), a strong PAH triol fragment ion ($[PAH]^+$) together with the production of protonated base ion ($[G+H]^+$). Subsequent isolation and MS/MS-CID fragmentation of the $[BH_2-H_2O]^+$ ion at m/z 436 (Figure 1F) show two additional losses of H_2O $[BH_2-2H_2O]^+$ and $[BH_2-3H_2O]^+$, the $[PAH]^+$ ion and its fragments to yield an abundant protonated base ion $[G+H]^+$ which dominated the spectrum.

PAH Triol Fragments. The MS/MS-CID of the $[PAH]^+$ at m/z 303 (Figure 1G) resulted in two dominant fragments at m/z 285 ($[PAH-H_2O]^+$) and m/z 257 ($[PAH-H_2O-CO]^+$). An additional loss of water yielded the $[PAH-2H_2O-CO]^+$ ion at m/z 239.

Determination of B[a]PDE adducts in Calf thymus DNA hydrolysates by MALDI-MS. B[a]PDE was reacted with calf thymus DNA and hydrolyzed as described in the experimental section. One aliquot corresponding to 1.5 µg of DNA was subjected to HLB-µ-SPE purification and about 12.5 ng of adducted DNA were loaded onto the MALDI sample support using the DHB ML protocol. The ion at m/z 570 corresponds to the $[M+H]^+$ of B[a]PDE-dG, the ion at m/z 650 corresponds to the $[M+H]^+$ of B[a]PDE-dGp, while the $[M+H]^+$ ion at m/z 554 characterizes B[a]PDE-dA (Figure 2A). Next to the $[M+H]^+$, the aglycon ($[BH_2]^+$) at m/z 454 and at m/z 438, the protonated deoxynucleoside ions ($[dN+H]^+$) at m/z 268 and at m/z 252, the deoxynucleoside-monophosphate ions ($[dNp+H]^+$) at m/z 332 and at m/z 348, the protonated nucleobase fragment ions ($[N+H]^+$) at m/z 136 and at m/z 152 of the B[a]PDE adducts are detected. The signals of the positively charged B[a]PDE triol fragment at m/z 303 and its fragments are also present. Subsequently, MS/MS-CID experiments in the positive mode are employed to confirm the identification and structural characterization of the DNA adducts (Figure 2B, C). It should be noted that the weak ion for the B[a]PDE-dGp could be fragmented in one of three MS/MS-CID experiments (data not shown). Of further interest are some ions detected at m/z 819, m/z 843 and m/z 859 labeled as α, β and γ, respectively (Figure 2A). Their selection and subsequent MS/MS-CID fragmentation disclose the presence of incomplete hydrolysis products. Thus, the MS/MS-CID of the ion at m/z 843 (Figure 2D) yield a characteristic mass-signature allowing an identification of the alkylated dinucleotide B[a]PDE-dAp-dC. The ions at m/z 112 ($[C+H]^+$) and at m/z 210 ($[dC-H_2O+H]^+$) indicate the presence of dC, while ions at m/z 136 ($[A+H]^+$) and m/z 332 ($[dAp+H]^+$) indicate the presence of dAp. The presence of B[a]PDE is confirmed at m/z 303 (PAH^+) and its characteristic fragment ions. The ion at m/z 541 is the base peak in the spectrum and correspond to the loss of B[a]PDE (-302) from the $[M+H]^+$, with a subsequent loss of cytosine at m/z 430. The ion at m/z 634 (B[a]PDE-dAp) reveal the identity of the alkylated nucleotide, which is confirmed by the presence of the $[BH_2]^+$ ion at m/z 438 and by the ion at m/z 732 resulting from the loss of a cytosine (-111) of the $[M+H]^+$. Moreover, the absence of a B[a]PDE-dC aglycon ion $[BH_2]^+$ at m/z 414 and a B[a]PDE-dCp ion at m/z 610 defines the site of adduction of the dAp. Likewise, MS/MS-CID spectra are obtained from the ion at m/z 819 leading to

unambiguous identification and characterization of the dinucleotide B[a]PDE-dCp-dC (Figure 2E). An intense ion at m/z 112 ([C+H]$^+$) and at m/z 308 ([dCp+H]$^+$) define the presence of deoxycytidine. The ions at m/z 610 (B[a]PDE-dCp) and the aglycon ion at m/z 414 characterize the alkylation event, while the ion at m/z 708 represents the loss of a cytosine (-111) from the [M+H]$^+$. The ion at m/z 517 corresponds to the loss of B[a]PDE (-302) from the [M+H]$^+$ and a subsequent loss of cytosine at m/z 406, are further informative ions in this spectrum. The MS/MS-CID of the weak ion at m/z 859 produced poor spectra.

Determination of B[a]PDE adducts in calf thymus DNA by the ^{32}P-postlabeling assay. Ten µg of B[a]PDE-DNA hydrolysates that were also used for MALDI-MS studies were enriched by butanol extraction. Eventually an equivalent of 2 ng of ^{32}P-Postlabeled DNA was applied to multidimensional TLC (Figure 2A inset). Here, one intense spot corresponding to B[a]PDE-deoxyguanosine and a faint spot tentatively assigned to the B[a]PDE-deoxyadenosine adduct, were detected, giving rise to a relative adduct level (RAL) of ~25 adducts per 10^6 deoxynucleotides and ~2 adducts per 10^6 deoxynucleotides for B[a]PDE-deoxyguanosine and B[a]PDE-deoxyadenosine, respectively.

Determination of B[c]ChDE adducts in Calf thymus DNA hydrolysates by MALDI-MS

To further demonstrate the reliability of the MALDI-MS based methodology, DNA adducts derived from reaction with (±)-*anti*-benzo[c]chrysene-9,10-diol-11,12-epoxide (B[c]ChDE), a metabolic activation product of benzo[c]chrysene, were studied. In Scheme 2 the major predicted fragment and the fragmentation pathways of B[c]ChDE-DNA adduct are depicted. Thus, one aliquot of hydrolysate that corresponds to 1.5 µg of DNA was subjected to HLB-µ-SPE and a volume corresponding to 12.5 ng of modified DNA was loaded onto the MALDI sample support. The ion at m/z 596 corresponds to the [M+H]$^+$ of B[c]ChDE-dG, the ion at m/z 580 is the [M+H]$^+$ of B[c]ChDE-dA and the ion at m/z 556 corresponds to the [M+H]$^+$ of B[c]ChDE-dC, all of which were readily detected (Figure 3A). In addition, MALDI-TOF analysis yielded signals of the 2'-deoxynucleoside-3'-monophosphate adducts: the B[c]ChDE-dGp at m/z 676, the B[c]ChDE-dAp at m/z 660 and the B[c]ChDE-dCp at m/z 636. Next to the [M+H]$^+$, unique mass-signature ions for the B[c]ChDE and the alkylated nucleoside were also detected. Here, the characteristic [BH$_2$]$^+$ ions at m/z 480 for B[c]ChDE-dG, at m/z 464 for B[c]ChDE-dA and at m/z 440 for B[c]ChDE-dC are noted. The protonated deoxynucleoside ions ([dN+H]$^+$), at m/z 268 and at m/z 252, the deoxynucleoside-monophosphate fragment ([dNp+H]$^+$) at m/z 332 and 348 and the protonated nucleobase fragment ([N+H]$^+$) at m/z 112, 136 and 152 are also detected. For the B[c]ChDE triol ([PAH]$^+$) at m/z 329 mainly loss of H$_2$O, the [PAH-H$_2$O]$^+$ ion at m/z 311 are observed. Subsequently, MS/MS-CID experiments in the positive mode were carried out to obtain fragmentation spectra of B[c]ChDE-dG, B[c]ChDE-dA, B[c]ChDE-dC (Figure 3B-D), B[c]ChDE-dGp and of B[c]ChDE-dAp (Figure 3E and F). Typically, the B[c]ChDE triol [PAH]$^+$ yielded a weak ion at m/z 329, an intense fragment ion

[PAH-H$_2$O]$^+$ at m/z 311 as well as [PAH-2H$_2$O]$^+$ at m/z 293. Moreover, consecutive loss of CO and water from the ion at m/z 311 produce [PAH-H$_2$O-CO]$^+$ at m/z 283 and [PAH-2H$_2$O-CO]$^+$ at m/z 265. (Figure 3B inset and Scheme 2). This characteristic fragmentation pattern could be further evidenced in MS/MS-CID esperiments and together with the [M+H]$^+$ and the aglycon [BH$_2$]$^+$ all the structural characteristic specific ions are present in each spectrum in agreement with the elemental compositions of the fragment ions depicted in Scheme 2. An identification and molecular characterization of the DNA adducts could thus be definitively confirmed. Moreover, the MS/MS-CID spectrum of B[c]ChDE-dC [M+H]$^+$ ion at m/z 556 yield, next to its specific fragments, additional ions at m/z 136 and at m/z 480 that correspond with high probability to the [N+H]$^+$ of dA and the aglycon ion [BH2$_2$]$^+$ of B[c]ChDE-dG, respectively. We propose that the isobaric fragment at m/z 556.1478 derived from a dinucleotide (see below) could not be efficiently discriminated by the ion-selector in the fragmentation process. We additionally observed ions at m/z 845, m/z 869, m/z 884 and m/z 893, which are labeled α, β, γ and δ in Figure 3A, respectively. The molecular ion at m/z 893 allowed an identification of the B[c]ChDE-dAp-dA dinucleotide (Figure 4A). The ion at m/z 660 (B[c]ChDE-dAp) reveal the identity of the alkylated nucleotide, which is confirmed by the presence of the aglycon [BH$_2$]$^+$ at m/z 464 and by the ion at m/z 758 derived from the loss of adenine (-135) of the [M+H]$^+$. The CID spectrum yielded the specific ions for the deoxyadenosine: the ions at m/z 136 ([A+H]$^+$), at m/z 252 ([dA+H]$^+$) and m/z 332 ([dAp+H]$^+$), whereas the ion at m/z 565 corresponded to the loss of B[c]ChDE (-328) from the [M+H]$^+$. The selection and subsequent MS/MS-CID fragmentation of the [M+H]$^+$ ion at m/z 869 led to an identification of the dinucleotide B[c]ChDE-dAp-dC (Figure 4C). Thus, ions at m/z 112 ([C+H]$^+$) and at m/z 210 ([dC-H$_2$O+H]$^+$) identify the dC, while ions at m/z 136 ([A+H]$^+$) and 332 ([dAp+H]$^+$) identify the dAp as the nucleotide components. The specific ion at m/z 660 (B[c]ChDE-dAp) reveal the identify of the alkylated nucleotide which is confirmed by the [BH$_2$]$^+$ ion at m/z 464 and by the ion at m/z 758 which resulted from the loss of cytosine (-111) of the [M+H]$^+$. Moreover, the ion at m/z 541 corresponded to the loss of B[c]ChDE (-328) from the [M+H]$^+$ with subsequent loss of cytosine (-111) at m/z 430. The MS/MS-CID spectrum of the ion at m/z 845 identified the dinucleotide B[c]ChDE-dCp-dC (Figure 4D).The ion at m/z 636 (B[c]ChDE-dCp) revealed the identity of the alkylated nucleotide which is confirmed by the presence of the [BH$_2$]$^+$ at m/z 440 and by the ion at m/z 734 derived from loss of cytosine (-111) of the [M+H]$^+$. Further specific ion for the nucleotide components includes the ions at m/z 112 ([C+H]$^+$) and a weak [dCp+H]$^+$ ion at m/z 308. Regarding the ion at m/z 884 (Figure 4B) the incomplete fragmentation pattern of the CID spectrum leads us to tentatively assigne fragment ions. Thus, the ion at m/z 660, the ion at m/z 464 (as its aglycon [BH$_2$]$^+$) together with ions at m/z 136 ([A+H]$^+$) and m/z 332 ([dAp+H]$^+$) define the presence of B[c]ChDE-dAp, while the ion at m/z 758 could be due to the loss of thymine (-126) from the [M+H]$^+$ (if B[c]ChDE-dAp-dT). Noteworthy, the ion at m/z 556 is isobaric to the B[c]ChDE-

dC (at *m/z* 556.2078) but could correspond to the loss of B[c]ChDE (-328) from the [M+H]$^+$. However, selection and fragmentation lead to the spectrum depicted in Figure 3D.

Determination of B[c]ChDE adducts in calf thymus DNA by the ^{32}P-postlabeling assay. Ten μg of B[c]ChDE-DNA hydrolysates that was also used for MALDI-MS analysis were enriched by butanol extraction followed by ^{32}P-postlabeling with 2 ng of adducted DNA hydrolysate being applied to multidimensional TLC (see inset in Figure 3A). Here, two radioactive intense spots corresponding to B[c]PhDE-deoxyguanosine and B[c]PhDE-deoxyadenosine adducts and a third very faint spot tentatively assigned as B[c]PhDE-deoxycytidine adduct were detected. The RAL was determined as ~59 adduct per 10^6 deoxynucleotides for B[c]PhDE-deoxyguanosine, ~18 adduct per 10^6 deoxynucleotides for B[c]PhDE-deoxyadenosine and ~3 adduct per 10^7 deoxynucleotides for B[c]PhDE-deoxycytidine.

Discussion

This study aimed for the development of a rapid and sensitive method for the detection, identification and characterization of PAHDE-DNA adducts by MALDI-TOF-MS and the following experimental conditions need to be considered. Firstly, to facilitate the acquisition of MALDI-MS and MS/MS-CID mass spectra, the matrix sample preparation should produce consistent, reproducible and long-lasting analyte signals. This is achieved by the homogeneous distribution of the matrix-analyte crystals on the sample support. Secondly, the matrix should produce a limited number of identifiable matrix peaks, and should not be affected by the presence of salts, buffers, and other common sample components, which could results in matrix clusters formation and ion suppression. Thirdly, the matrix influences the extent of desorption/ionization-induced fragmentation of analytes [42]. Notable, the 3-hydroxypicolinc acid (HPA) matrix proved useful for the analysis of oligonucleotides[43], while α-cyano-4-hydroxycinnamic acid (CHCA) and 2,5-dihydroxybenzoic acid (DHB) worked well for the analysis of proteins and peptides[32;44]. For an analysis of PAHDE-DNA adducts these matrices exhibited different characteristics with respect to appearance of the crystals, the production of matrix ions, signal reproducibility, and the degree of analyte fragmentation (data not shown). We found the DHB prepared by the matrix layer (ML) method[32] to be the matrix of choice for an analysis of PAHDE and PAHDE-DNA adducts. An improved homogeneity of the matrix crystals, enhanced sensitivity and reduced cationized matrix and analyte molecular ions by the addition of ammonium phosphate as matrix "dopant" was observed. Furthermore, the DHB ML sample preparation delivered consistent signals for hundreds of laser shots at one location to facilitate the acquisition of MALDI-MS and MS/MS-CID spectra. Likewise, the MALDI signals are long-lived, and a single sample can easily be used for extensive MS/MS-CID experiments. In addition, the purity of the sample is of great importance with DNA hydrolysates containing huge amounts of

unreacted nucleotides, buffers and enzymes which could hinder the matrix crystallization, the reproducibility of the spectra and provoke ion suppression effects to impair the MALDI-MS analysis. Samples were therefore subjected to solid phase extraction for purification/enrichment that either consisted of silica-based C18, polymer-based styrene/divinylbenzene (SDB) or N-vinylpyrrolidone/divinylbenzene (hydrophilic-lypophilic balance, HLB) bed materials. While many reports described the use of one of these resins alone or in combination for the purification/enrichment of PAHDE-DNA adducts[25;26;29], the HLB copolymer sorbent proofed valuable enabling increased recovery and enhanced loading capacity as compared to other resins. Therefore, we implemented the HLB sorbent and developed self-assembled disposable micro columns (HLB-μ-SPE, see experimental section) to meet the sample amount and the volume used for MALDI-MS sample preparation.

As shown in Table S-1 the reaction products of B[a]PDE with the four 2'-deoxynucleosides and the four 2'-deoxynucleosides-3'-monophosphates purified by HLB-μ-SPE and were measured with a mass accuracy of <10 ppm at a resolution of >5000, and a sensitivity of >100 fmol loaded on target.

The molecular ions ([M+H]$^+$) corresponded to the monoalkylated B[a]PDE adducts of dG, [M+H]$^+$ at *m/z* 570; dGp, [M+H]$^+$ at *m/z* 650; dA, [M+H]$^+$ at *m/z* 554; dAp, [M+H]$^+$ at *m/z* 634; dC, [M+H]$^+$ at *m/z* 530; dCp, [M+H]$^+$ at *m/z* 610. In the reaction mixture with 2'-deoxythymidine (dT) and 2'-deoxythymidine-3'monophosphate (dTp), no adducts were detected presumably due to low alkylation efficiency. This is consistent with previous reports using similar condition but other ionization techniques[26]. Noteworthy, cationazed ([M+Na]$^+$ or [M+K]$^+$) molecular ions were weak or absent, which is a particular advantage for the interpretation of spectra. The MALDI-TOF spectra of PAHDE-DNA adducts together with the molecular ion ([M+H]$^+$) provide typical mass-signature specific for the alkylating molecule and the nucleotide involved in the alkylation reaction. Indeed, next to the [M+H]$^+$, the aglycon fragment ([BH$_2$]$^+$), the protonated 2'-deoxynucleoside ([dN+H]$^+$) or the 2'-deoxynucleoside-3'-monophosphate fragments ([dNp+H]$^+$), the protonated nucleobase ([N+H]$^+$) together with signals of the positively charged B[a]PDE triol fragment ([PAH]$^+$) are the observed mass-signature (Figure 1A, B and S-1, S-2). Such fingerprints can thus be defined as DNA-adduct mass fingerprint (DMF). Based on the DMF, the [M+H]$^+$ and specific ions were selected for MS/MS-CID measurements to obtain definitive structural diagnostic fragment ions (Figure 1 C, D and S-3, S-4), i.e DNA-adducts fragment fingerprint (DFF). In Scheme 1 the major fragment and the fragmentation pathways are depicted. Several different investigations reported the detection of the aglycon ion [BH$_2$]$^+$, as major ion product for the characterization of DNA adducts[28;29;39;40;45;46].The new developed method allowed further fragmentation of this and other DFF ions component on the same target spot. Thus, MS/MS-CID fragmentation of the [BH$_2$]$^+$ and the [PAH]$^+$ ions were achieved leading to additional diagnostic fragment ions which enabled improved characterization. Indeed, fragmentation of the protonated adducted base ([BH$_2$]$^+$) produced [G+H]$^+$ and

involves hydrogen transfer from PAH triol. This, would explain the positive charged [PAH]$^+$ fragment. Moreover, the fragmentation of the [PAH]$^+$ produced the fragments [PAH-H$_2$O]$^+$, [PAH-H$_2$O-CO]$^+$ and [PAH-2H$_2$O-CO]$^+$ which are specific and characteristic of the PAH triol moiety. All together these information allow the identity of the PAH and the number and type of ring substituent to be defined. In fact, these diagnostic fragments were used to structurally differentiate diastereoisomers of PAH and PAH-DNA adducts[42;47;48]. Therefore, both the DMF and DFF permitted an identification and characterization of B[a]PDE adducts formed with the four 2'-deoxynucleosides and the four 2'-deoxynucleosides 3'-monophosphates (see Figure 1 and S-1-S-4).

Consequently, we applied this methodology for analysis of adducts formed with calf thymus DNA upon reaction with B[a]PDE. Based on the full MALDI-TOF mass spectrum, a simple comparison with the reference DMF allowed specification of the molecular ion and characteristic fragment ions of the B[a]PDE-DNA adducts. Readily molecular ions and characteristic fragment for deoxynucleotides and deoxynucleosides were detected. The B[a]PDE-dG at m/z 570, B[a]PDE-dGp at m/z 650 and B[a]PDE-dA at m/z 554, B[a]PDE-dCp at m/z 610 DNA adducts were the main adducted molecules detected with calf thymus DNA. Notably, the latter adduct could be identified as a dinucleotide component. Some of the commercially available SPDEs, used in the DNA hydrolysis, contain 3'-phospatase activity[49]. This may explain why the B[a]PDE-DNA adduct hydrolysates contained both the alkylated nucleotides and nucleosides. We exclude the possibility that these PAHDE-2'-deoxynucleosides are fragmentation products from the equivalent 3'-monophosphate as we never observed such a fragmentation in MS or MS/MS-CID reference spectra. Further adducts were identified as dinucleotide resulting from incomplete DNA hydrolysis. i.e. the ions corresponding to the B[a]PDE-dCp-dC and the B[a]PDE-dAp-dC dinucleotides. Moreover, the MALDI-MS/MS-CID analysis allowed the specification and characterization of the alkylation event as the B[a]PDE-dCp and B[a]PDE-dAp adducts, respectively. In one report and next to B[a]PDE-dGp ions at higher m/z were observed. However, the investigators were unable to determinate the identity and to characterize further adducts, despite the elaborate SPE enrichment prior to CE-coupled ESI-MS[25]. Incomplete DNA hydrolysis products are probably formed because the alkylated nucleotides are not efficiently recognized from MN and/or SPDE due to sterically unfavorable interactions of their modified structure. Nonetheless, the efficiency of MN/SPDE enzymatic hydrolysis could be improved by a preceding DNA hydrolysis step with DNA-ase I or Benzonase® endonucleases which release 5'-phosphate oligonucleotides. Importantly, the digestion protocol used in this study is widely used for ^{32}P-postlabeling assay. Noteworthy, cationazed ([M+Na]$^+$ or [M+K]$^+$) molecular ions were weak or absent. Thus sufficient cleanup of analytes was achieved. In previous report using standard solutions of adducts the cationazed form ions were the base peaks in the spectra rendering an

interpretation challenging. In contrast the here reported HLB-µ-SPE purification/enrichment protocol is simple and robust.

We also compare findings obtained by MALDI-MS with the ^{32}P-postlabeling assay. Here, two radioactive spot were detected corresponding to B[a]PDE-deoxyguanosine and the B[a]PDE-deoxyadenosine adducts. Importantly, dinucleotides which are also substrate of the T4 polynucleotide kinase were not detected. Hence, a comparison of the two methodologies clearly demonstrates that only MALDI-MS provided molecular specification to characterize the nucleotide and the alkylating agent involved. Additionally, the B[a]PDE-dCp adduct was identified by MALDI-MS. Thus, MALDI-MS provided comprehensive information such as the detection and identification of deoxynucleotides and deoxynicleosides and dinucleotides adducts. That was not detected by the ^{32}P-postlabeling assay.

The reliability of our MALDI-MS based methodology is further demonstrated by the detection, identification and characterization of the adducts derived from reaction of calf thymus DNA with (±)-*anti*-benzo[c]chrysene-9,10-diol-11,12-epoxide (B[c]ChDE). The B[c]ChDE, is an intruding PAH since is characterized by the presence of a bay and a fjord region within the same molecule (see Scheme 2). The location of the epoxide in the sterically-hindered fjord region determines its high carcinogenic activity [9]. With the ^{32}P-postlabeling assay only three radioactive spot were detected corresponding to B[c]ChDE-deoxyguanosine, -deoxyadenosine and deoxycytidine adducts. Applying the newly developed method we were able to readily identify a total of six B[c]ChDE-DNA adduct as 2'-nucleosides and 2'-nucleosides-3'monophosphate: the B[c]ChDE-dG at *m/z* 596, the B[c]ChDE-dGp at *m/z* 676, the B[c]ChDE-dA at *m/z* 580, the B[c]ChDE-dAp at *m/z* 660, the B[c]ChDE-dC at *m/z* 556 and the B[c]ChDE-dCp at *m/z* 636. In addition, three dinucleotides adducts were also detected: the B[c]ChDE-dAp-dA, the B[c]ChDE-dAp-dC and the B[c]ChDE-dCp-dC. MS/MS-CID fragmentation confirmed the identificaty and allowed the molecular characterization of the nucleotides and nucleosides with B[c]ChDE as the alkylating molecule. Again, a comparison of the two methodologies confirm that only MALDI-MS led to a comprehensive analysis and provided molecular specification for an identification and characterization of the adducts formed. In regard to the sensitivity, B[a]PDE-deoxycytidine was detected by the 32P-postlabeling with a RAL of ~3 adduct per 10^7 deoxynucleotides. Consequently, the MALDI-TOF-MS method has at list a similar sensitivity.

Conclusions

A rapid MALDI-TOF-MS based method has been developed, and applied for the sensitive detection, identification and characterization of prototypic PAH-DNA adducts derived from B[a]PDE and B[c]ChDE that were reacted with calf thymus DNA. We evidence that MALDI-TOF and MS/MS-CID spectra yielded characteristic fragmentation patterns that allow identification of deoxyguanosine, deoxyadenosine and deoxycytidine DNA adducts. The spectra are simple to interpret and there is a notable absence of

cationized sample and matrix ions. Overall the method is sensitive, and nanogramms of hydrolyzed DNA are sufficient for the identification and characterization of the adducts. When compared with the ^{32}P-postlabeling assay a distinct advantage of the method is the unambiguous identification of DNA adducts and the possibility to identify the chemical nature of the alkylating agent. Moreover, our method allowed the simultaneous and unambiguous detection and identification of deoxynucleotide and deoxynucleosides PAHDE adducts at the same time. In addition, the analysis time is drastically reduced and a dedicated radioactivity laboratory is not required. In the long term, a database of DNA-adducts mass fingerprints and fragmentation pattern of reactions products of several different PAH with nucleotides and nucleoside will be developed for an automated detection of DNA-adducts.
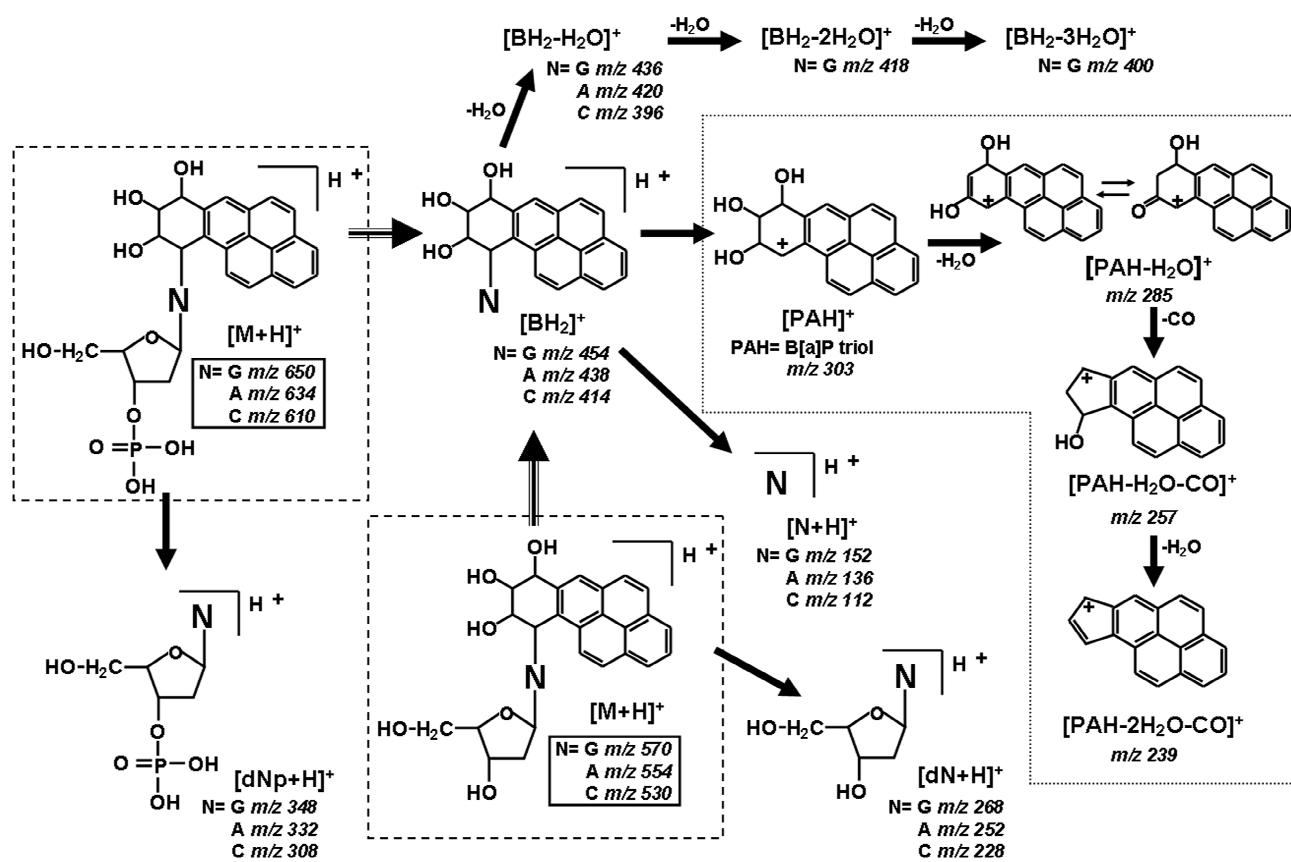
Acknowledgements

Supporting Information Available

Experimental procedures, spectral data and images. This material is available free of charge via the Internet at htpp://pubs.acs.org.

Scheme 1

Scheme 1. Proposed MALDI-MS/MS-CID fragmentation pathway of B[a]PDE-2'-deoxynucleosides and B[a]PDE-2'-deoxynucleosides 3'-phosphate. G: guanine, A: adenine; C: cytosine

Scheme 2
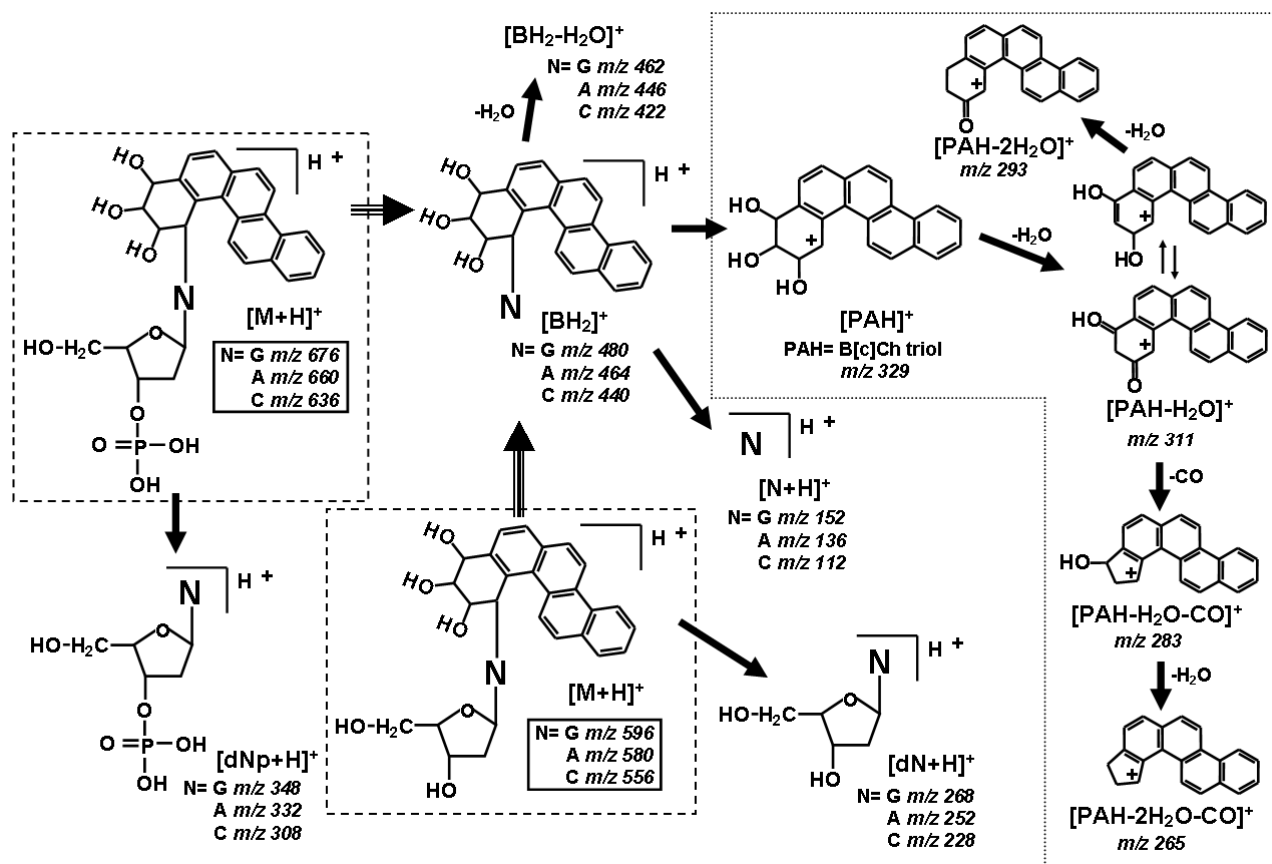
Scheme 2. Proposed MALDI-MS/MS-CID fragmentation pathway of B[c]ChDE-2'-deoxynucleosides and B[c]ChDE-2'-deoxynucleosides 3'-monophosphate. G: guanine, A: adenine; C: cytosine .

Figure 1
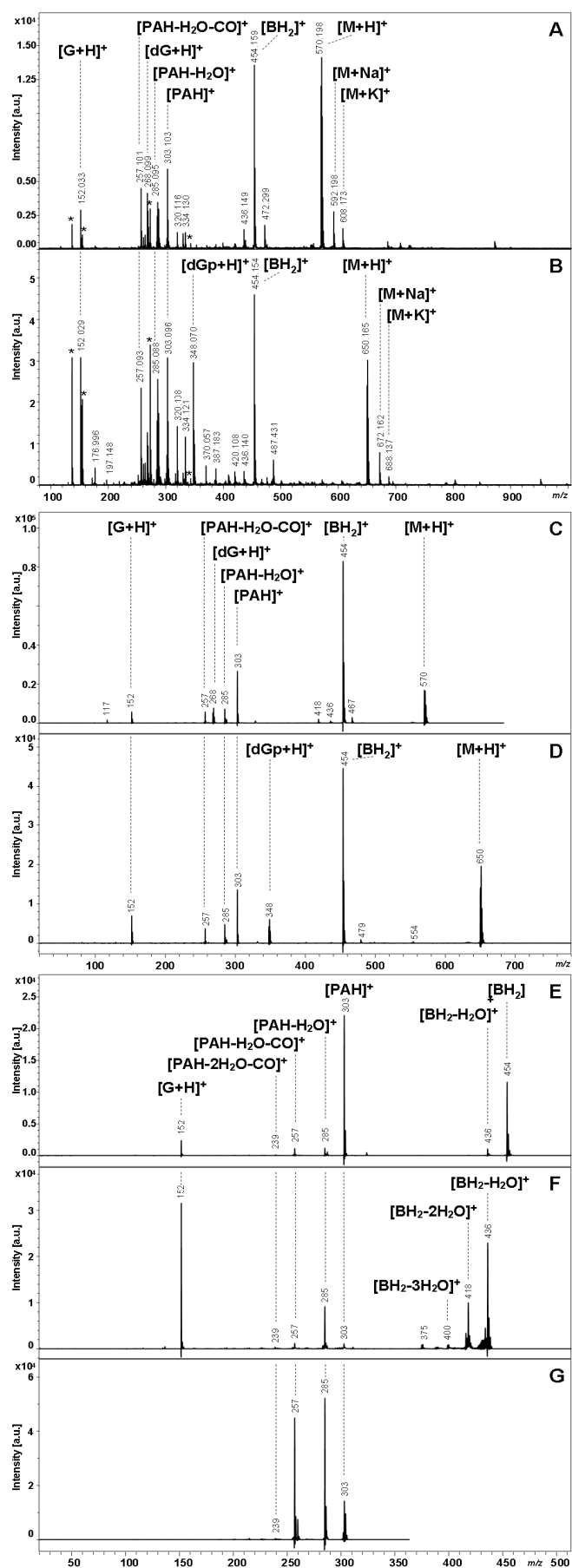
Figure 1. MALDI-TOF-MS and MS/MS-CID reference spectra of B[a]PDE adducts.

Typical positive ion MALDI-TOF mass spectra of B[a]PDE-dG (**A**) and B[a]PDE-dGp (**B**). [*= matrix ions]. Positive ion MALDI-MS/MS-CID mass spectra of the molecular ions [M+H]+: at m/z 570 of the B[a]PDE-dG in **A** (**C**), at m/z 650 of the B[a]PDE-dGp (**D**) in **B**. Positive ion MALDI-MS/MS-CID reference spectra of specific relevant fragment ions: aglycon [BH$_2$]+ at m/z 454 (**E**); [BH$_2$-H$_2$O]+ at m/z 436 (**F**) and B[a]PDE triol [PAH]+ at m/z 303 (**G**) of the B[a]PDE-dG in **A**. The peak annotations are detailed in scheme 1.

Figure 2

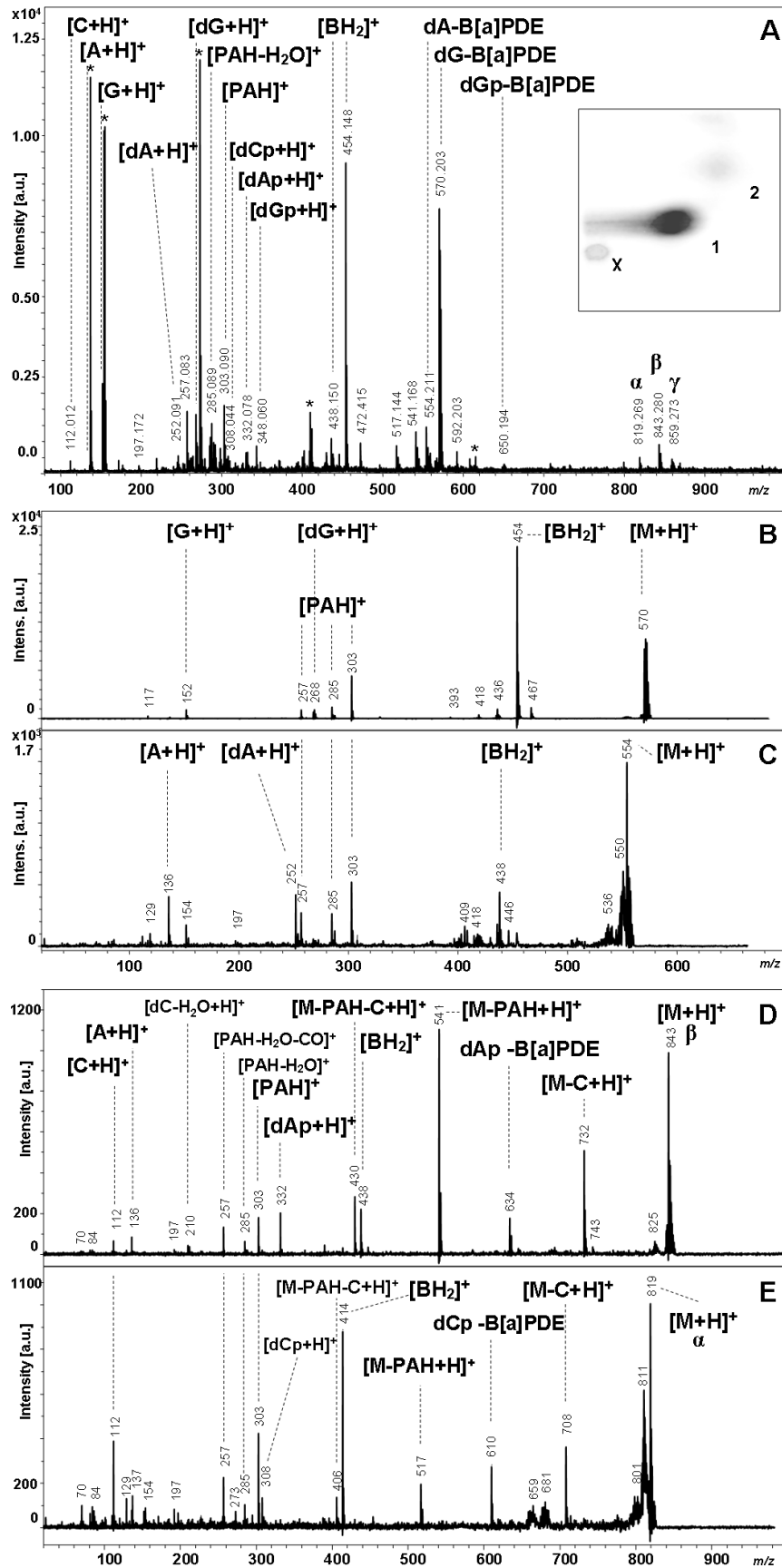**Figure 2. Analysis of B[a]PDE-DNA adducts by MALDI-MS and the ^{32}P-postlabeling assay.** Typical positive ion MALDI-TOF mass spectrum of adducts formed upon reaction of B[a]PDE with calf thymus DNA (**A**). α, β and γ are incomplete hydrolysis products (see below). [*****= matrix ions]. Typical storage phosphor image obtained following two dimensional TLC; 1 indicates B[a]PDE-dG adduct; 2 indicates the tentatively assigned B[a]PDE-dA adduct; x denotes the origin (**A**, inset).

Positive ion MALDI-MS/MS-CID mass spectra of the [M+H]+ ions of the DNA adducts present in DNA hydrolysate. Product ion spectra of the B[a]PDE-dG adduct at *m/z* 570 (**B**) and the B[a]PDE-dA adduct at *m/z* 554 (**C**) of the hydrolysates in **A**. The peak annotations are detailed in Scheme 1.

Positive ion MALDI-MS/MS-CID spectra of incomplete hydrolysis products detected in the B[a]PDE-DNA hydrolysates labeled β and α in **A**. Product ion spectra obtained from the molecular ions [M+H]+: at *m/z* 843, identified as B[a]PDE-dAp-dC (**D**); at *m/z* 819, identified as B[a]PDE-dCp-dC (**E**).
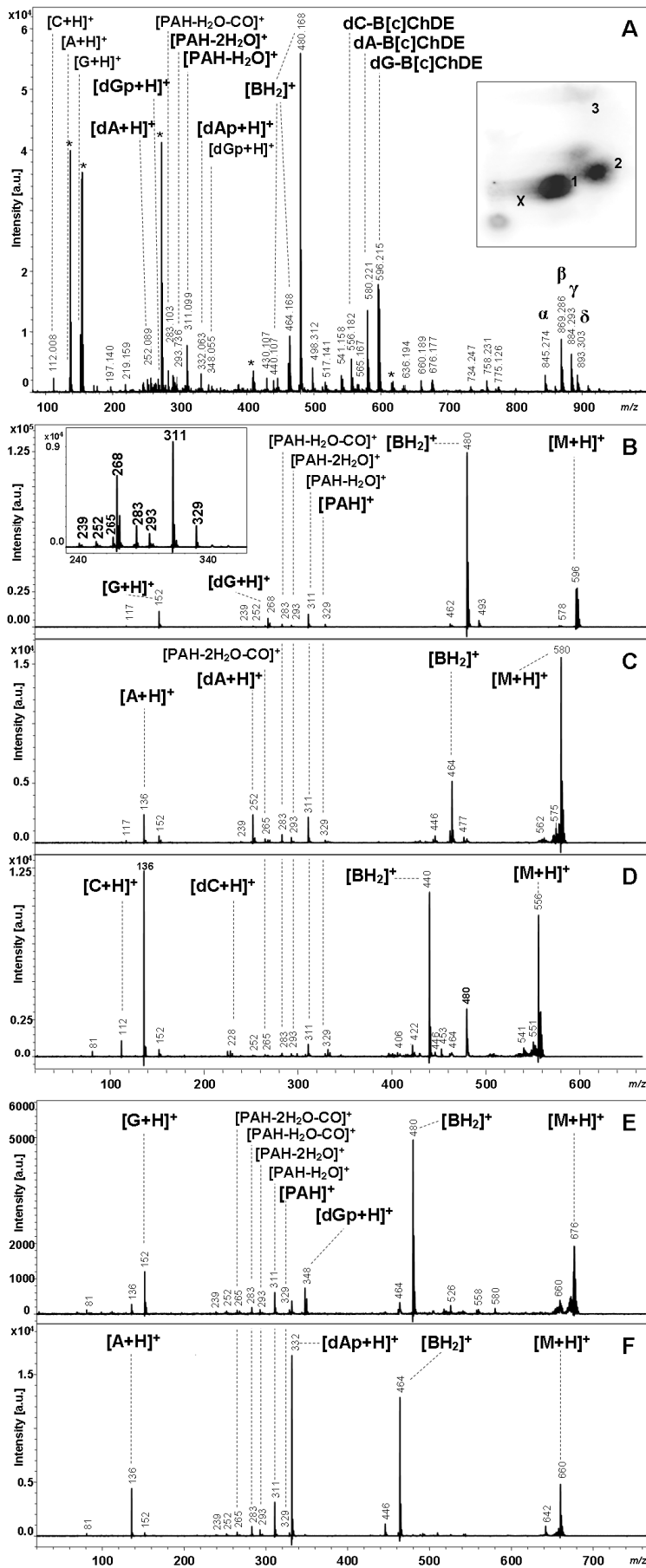
Figure 3

Figure 3. Analysis of B[c]ChDE-DNA adducts by MALDI-MS and the ^{32}P-postlabeling assay.

Typical positive ion MALDI-TOF mass spectrum of adducts formed upon reaction of B[c]ChDE with calf thymus DNA (**A**). α , β, γ and δ are incomplete hydrolysis products. [*= matrix ions]. Typical storage phosphor image obtained following two dimensional TLC; 1 indicates B[c]ChDE-dG adduct; 2 indicates B[c]ChDE-dA adduct; 3 indicates the tentatively assigned B[a]PDE-dC adduct; x denotes the origin (**A**, inset). Positive ion MALDI-MS/MS-CID mass spectra of the [M+H]+ ions of DNA adducts present in DNA hydrolysate. Product ion spectra of the B[c]ChDE-dG adduct at *m/z* 596 (**B**), the B[c]ChDE-dA adduct at *m/z* 580 (**C**), the B[c]ChDE-dC adduct at *m/z* 556 (**D**), the B[c]ChDE-dGp at *m/z* 676 (**E**) and the B[c]ChDE-dAp adduct at *m/z* 660 (**F**) of the hydrolysates in **A**. Expanded region of **B** for details of the B[c]ChDE triol ion [PAH]+ characteristic mass-signature (**B**, inset). The peak annotations are detailed in Scheme 2.

Figure 4

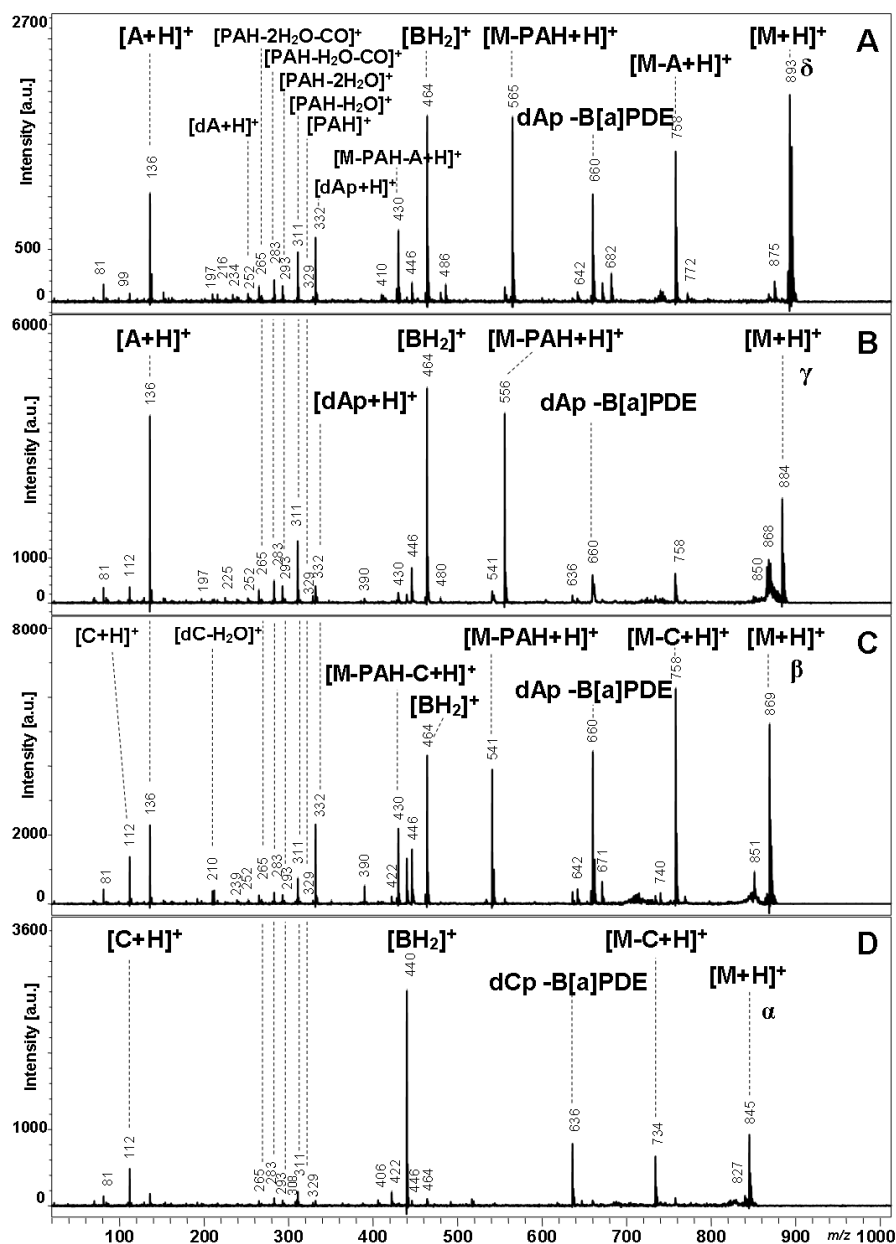**Figure 4. MALDI-MS/MS-CID spectra of incomplete hydrolysis products detected in of B[c]ChDE-DNA hydrolysates.** Positive ion MALDI-MS/MS-CID mass spectra of the dinucleotides [M+H]+ present the DNA hydrolysates labeled δ, γ, β and α in Figure 3A. The [M+H]+ ions: at m/z 893, identified as B[c]ChDE-dAp-dA (**A**); at m/z 884 (**B**); at m/z 869, identified as B[c]ChDE-dAp-dC (**C**); at m/z 845, identified as B[c]ChDE-dCp-dC (**D**)

References

1. Dipple, A. *Carcinogenesis* **1995,** *16*, 437-41.

2. Schoket, B.; Poirier, M. C.; Mayer, G.; Torok, G.; Kolozsi-Ringelhann, A.; Bognar, G.; Bigbee, W. L.; Vincze, I. *Mutat Res* **1999,** *445*, 193-203.

3. Bostrom, C. E.; Gerde, P.; Hanberg, A.; Jernstrom, B.; Johansson, C.; Kyrklund, T.; Rannug, A.; Tornqvist, M.; Victorin, K.; Westerholm, R. *Environ Health Perspect.* **2002,** *110 Suppl 3*, 451-88.

4. Phillips, D. H. *Mutat Res* **1999,** *443*, 139-47.

5. Dipple, A.; Khan, Q. A.; Page, J. E.; Ponten, I.; Szeliga, J. *Int J Oncol* **1999,** *14*, 103-11.

6. Peltonen, K.; Dipple, A. *J Occup.Environ Med* **1995,** *37*, 52-58.

7. Xue, W.; Warshawsky, D. *Toxicol Appl Pharmacol* **2005,** *206*, 73-93.

8. Shimada, T.; Fujii-Kuriyama, Y. *Cancer Sci* **2004,** *95*, 1-6.

9. Amin, S.; Lin, J. M.; Krzeminski, J.; Boyiri, T.; Desai, D.; el Bayoumy, K. *Chem Res Toxicol* **2003,** *16*, 227-31.

10. Dipple, A. *IARC Sci Publ.* **1994,** 107-29.

11. Cheng, S. C.; Hilton, B. D.; Roman, J. M.; Dipple, A. *Chem Res Toxicol* **1989,** *2*, 334-40.

12. Dipple, A.; Peltonen, K.; Cheng, S. C.; Ross, H.; Bigger, C. A. *Adv.Exp Med Biol* **1994,** *354*, 101-12.

13. Misra, B.; Amin, S.; Hecht, S. S. *Chem Res Toxicol* **1992,** *5*, 242-47.

14. Koreeda, M.; Moore, P. D.; Wislocki, P. G.; Levin, W.; Yagi, H.; Jerina, D. M. *Science* **1978,** *199*, 778-81.

15. Dipple, A.; Pigott, M. A.; Agarwal, S. K.; Yagi, H.; Sayer, J. M.; Jerina, D. M. *Nature* **1987,** *327*, 535-36.

16. Cadet, J.; Weinfeld, M. *Anal Chem* **1993,** *65*, 675A-82A.

17. Divi, R. L.; Beland, F. A.; Fu, P. P.; Von Tungeln, L. S.; Schoket, B.; Camara, J. E.; Ghei, M.; Rothman, N.; Sinha, R.; Poirier, M. C. *Carcinogenesis* **2002,** *23*, 2043-49.

18. Amin, S.; Misra, B.; Desai, D.; Huie, K.; Hecht, S. S. *Carcinogenesis* **1989,** *10*, 1971-74.

19. Lin, J. M.; Desai, D.; Chung, L.; Hecht, S. S.; Amin, S. *Chem Res Toxicol* **1999,** *12*, 341-46.

20. Gupta, R. C. *Cancer Res* **1985,** *45*, 5656-62.

21. Reddy, M. V.; Gupta, R. C.; Randerath, E.; Randerath, K. *Carcinogenesis* **1984,** *5*, 231-43.

22. Randerath, K.; Randerath, E. *Drug Metab Rev* **1994,** *26*, 67-85.

23. Arif, J. M.; Dresler, C.; Clapper, M. L.; Gairola, C. G.; Srinivasan, C.; Lubet, R. A.; Gupta, R. C. *Chem Res Toxicol* **2006,** *19*, 295-99.

24. Singh, R.; Farmer, P. B. *Carcinogenesis* **2006,** *27*, 178-96.

25. Barry, J. P.; Norwood, C.; Vouros, P. *Anal Chem* **1996,** *68*, 1432-38.

26. Willems, A. V.; Deforce, D. L.; Van den Eeckhout, E. G.; Lambert, W. E.; Van Peteghem, C. H.; De Leenheer, A. P.; Van Bocxlaer, J. F. *Electrophoresis* **2002,** *23*, 4092-103.

27. Beland, F. A.; Churchwell, M. I.; Von Tungeln, L. S.; Chen, S.; Fu, P. P.; Culp, S. J.; Schoket, B.; Gyorffy, E.; Minarovits, J.; Poirier, M. C.; Bowman, E. D.; Weston, A.; Doerge, D. R. *Chem Res Toxicol* **2005,** *18*, 1306-15.

28. Singh, R.; Gaskell, M.; Le Pla, R. C.; Kaur, B.; Azim-Araghi, A.; Roach, J.; Koukouves, G.; Souliotis, V. L.; Kyrtopoulos, S. A.; Farmer, P. B. *Chem Res Toxicol* **2006,** *19*, 868-78.

29. Wang, J. J.; Marshall, W. D.; Frazer, D. G.; Law, B.; Lewis, D. M. *Anal Biochem* **2003,** *322*, 79-88.

30. Paehler, A.; Richoz, J.; Soglia, J.; Vouros, P.; Turesky, R. *J Chem Res Toxicol* **2002,** *15*, 551-61.

31. Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal Chem* **1991,** *63*, 1193A-203A.

32. Garaguso, I.; Borlak, J. *Proteomics* **2008,** *8*, 2583-95.

33. Cohen, L. H.; Gusev, A. I. *Anal Bioanal.Chem* **2002,** *373*, 571-86.

34. Yagi, H.; Thakker, D. R.; Hernandez, O.; Koreeda, M.; Jerina, D. M. *J Am Chem Soc* **1977,** *99*, 1604-11.

35. MacLeod, M. C.; Selkirk, J. K. *Carcinogenesis* **1982,** *3*, 287-92.

36. Luch, A.; Platt, K. L.; Seidel, A. *Carcinogenesis* **1998,** *19*, 639-48.

37. Gupta, R. C.; Reddy, M. V.; Randerath, K. *Carcinogenesis* **1982,** *3*, 1081-92.

38. Halter, R.; Hansen, T.; Seidel, A.; Zieman, C.; Borlak, J. *Journal of Occupational and Environmental Hygiene* **2007,** *4*, 44-64.

39. Ricicki, E. M.; Soglia, J. R.; Teitel, C.; Kane, R.; Kadlubar, F.; Vouros, P. *Chem Res Toxicol* **2005,** *18*, 692-99.

40. Charles A.Barnes; Norman H.L.Chiu *International Journal of Mass Spectrometry* **2009,** *279*, 170-75.

41. Straub, K. M.; Meehan, T.; Burlingame, A. L.; Calvin, M. *Proc Natl Acad Sci U S A* **1977,** *74*, 5285-89.

42. Stemmler, E. A.; Hettich, R. L.; Hurst, G. B.; Buchanan, M. V. *Rapid Commun.Mass Spectrom* **1993,** *7*, 828-36.

43. Wu, K. J.; Steding, A.; Becker, C. H. *Rapid Commun.Mass Spectrom* **1993,** *7*, 142-46.

44. Beavis, R. C.; Chait, B. T. *Rapid Commun.Mass Spectrom* **1989,** *3*, 432-35.

45. Annan, R. S.; Giese, R. W.; Vouros, P. *Anal Biochem* **1990,** *191*, 86-95.

46. Lay, J. O., Jr.; Chiarelli, M. P.; Bryant, M. S.; Nelson, R. W. *Environ Health Perspect.* **1993,** *99*, 191-93.

47. Chiarelli, M. P.; Chang, H. F.; Olsen, K. W.; Barbacci, D.; Huffer, D. M.; Cho, B. P. *Chem Res Toxicol* **2003,** *16*, 1236-41.

48. Huffer, D. M.; Chang, H. F.; Cho, B. P.; Zhang, L. K.; Chiarelli, M. P. *J Am Soc Mass Spectrom* **2001,** *12*, 376-80.

49. Naomi Suzuki; Padmaja M.Prabhu; Shinya Shibutani *Optimization in Drug Discovery*, Zhengyin Yan and Gary W.Caldwell, Ed.; Humana Press: 2004; Chapter 17.

Chapter III
Development of methods for phosphoproteins identification and phosphorylation site determination

Reversible protein phosphorylation

The term "protein post-translational modification" (PTM) describes protein processing by cleavage or covalent attachment of chemical groups to a protein. These modifications increase tremendously the diversity and heterogeneity of gene products and the way in which they are regulated. Over 300 types of modifications are known [239]. Protein reversible phosphorylation is one of the most studied cellular mechanisms to regulate protein functions. The process was discovered by the pioneering studies undertaken by the Nobel Prizwinner Krebs and Fisher to elucidate the complex hormonal regulation of skeletal muscle glycogen phosphorylase. They could determine the mechanism by which the inactive phosphorylase b was converted to phosphorylase a as the result of an enzyme-catalyzed phosphorylation-dephosphorylation reaction [240,241].

Reversible protein phosphorylation is a ubiquitous mechanism for the control of signal transduction networks that regulate diverse biological processes including response to extracellular stimuli, DNA damage and cell growth and division. Changes in protein phosphorylation affect the structure and activity of proteins regulating nearly all aspects of cell life including metabolic processes, DNA replication, gene expression, and the cell cycle. The phosphorylation statuses of proteins in a cell is controlled and modulated in a highly dynamic way by two different classes of enzymes: protein kinases which catalyze the transfer of phosphoryl groups from a high-energy compound to a nucleophilic acceptor on an amino acid side-chain of proteins, and protein phosphatases which catalyze water-driven hydrolysis of phosphoester bonds (Figure 16). The importance of these two classes of enzymes is also supported by their high number of genes in the genome, constituting approximately 2% of the proteins encoded by the human genome [242,243].
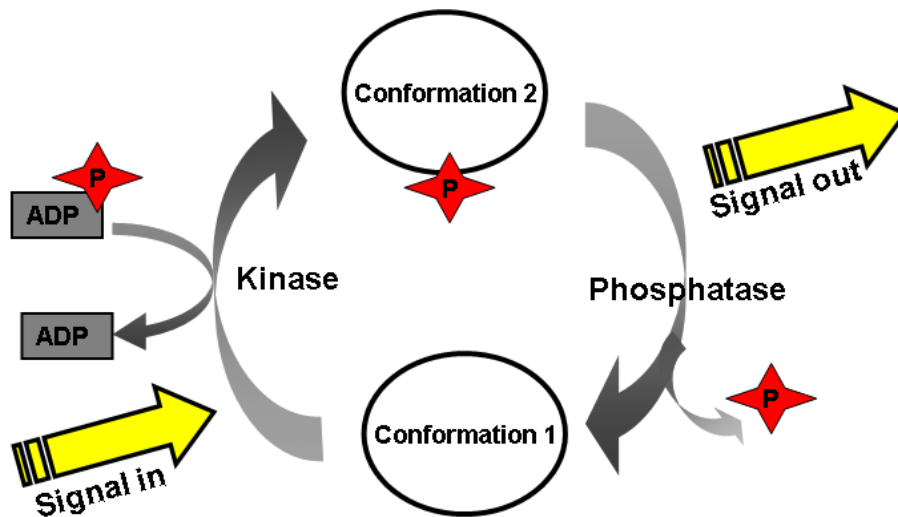
Figure 16. Protein phosphorylation is triggered by a stimulus inside or outside the cell. The protein kinase enzyme transfers the phospho-group from adenosine triphosphate (ADP+P =ATP) to the protein. Attaching the phospho-group to the protein will change its conformation and thus changes its activity. A protein phosphatase de-phosphorylates the protein and reversing its conformation.

It was estimated that the human genome codes for more than 500 protein kinases, the third most populous gene family, and 100 phosphatases [244]. The activities of protein kinases and phosphatases are subject to control both by extracellular stimuli and intracellular mechanisms and the molecular basis of this control is complex and varied.

Four types of phosphorylation are known which affect different amino acids. Indeed, the O-phosphates (O-phosphomonoesters) are formed by phosphorylation of hydroxyamino acids such as serine (S), threonine (T) and tyrosine (Y), residues. The N-, S- and acyl-phosphorylation are far less widespread and occur mostly on histidine and lysine (N-phosphates, phosphoamidates), cysteine (S-phosphothioesters) and aspartic and glutamic acid residues (acyl-phosphates, phosphate anhydrides). The predominant class of proteins phosphorylation in mammalian cells are O-phosphates which occur on Y, T and S residues with an estimated relative abundances of 1.8%, 11.8%, and 86.4%, respectively [245]. Moreover, there is a higher gain in signals involving tyrosine phosphorylation because it is less abundant and more tightly regulated [246].

Protein (de)phosphorylation can change the protein's enzyme activity, the cellular location, increase protein–protein interactions and target proteins for degradation, all of which induce cascade amplification variances of many essential cell processes such as signal transduction, cell differentiation, proliferation, metabolic maintenance, cell division, and apoptosis. Whereupon, regulation changes of the tightly controlled balance between phosphorylation and de-phosphorylation may lead to serious pathological conditions. Abnormalities in this complex (de)phosphorylation process have been shown to

be related to many diseases and cancer [21,32,247]. Therefore, identifying phosphorylation sites related to diseases might help to understand the origin, progression and hopefully the termination of these diseases. In addition, the protein kinases activity, regulated through phosphorylation in turn, plays a important role in cancer and it has been well described to be involved in development of malignancies [32-34]. The analysis of the entire cellular phosphoprotein panel, the so-called phosphoproteome, the understanding of regulatory roles and where and when protein phosphorylation takes place has become of pivotal importance in tissue injury and in particular in cancer studies. Knowledge of these molecular details will provide a framework for the development of novel therapies for the early detection and treatment of cancer. Because of this, there is considerable interest on developing methodologies to identify phosphorylated proteins and characterize its phosphorylation state in the contest of tissues/organ systems.

Protein phosphorylation analysis

Despite a growing knowledge of many phosphorylation consensus sequences, it is possible to predict whether a site could be substrate for a kinase but we cannot predict when a site will be phosphorylated. Thus, the experimental determination of sites of phosphorylation is an important task. The human proteome contains more than 100,000 phosphorylation sites [248,249]. It is estimated that one out of every three proteins is phosphorylated at some point in its life cycle [250]. Notably, phosphorylation leads to heterogeneity. Most phosphoproteins undergo phosphorylation on more than one residue, which means that all the molecules of a given protein are not identically phosphorylated. Another characteristic of phosphorylation is the ratio of phosphorylated to non-phosphorylated protein present in the cell [251]. Some protein residues may be always quantitatively phosphorylated, while others may only be transiently phosphorylated. The abundance of phosphorylated forms of a protein can be as low as 1-2% of the total amount of that protein [252]. Phosphorylation is highly dynamic and intensely regulated, and phosphorylation cycles may take place on a very short timescale. Therefore, the complexity of phosphorylation in terms of regulation, dynamics, abundance and stoichiometry render the analysis of phosphoproteins and the resulting phosphopeptides difficult. Various strategies for protein phosphorylation detection have been developed throughout the years. The combination of more different methods could lead to very articulated strategies. D' Ambrosio and colleagues [253] recently presented a survey of the different strategies for phosphoprotein enrichment which are summarized by the diagram in Figure 17. After sample pre-treatment to freeze the phosphorylation state of the proteins, detection of the phosphoprotein usually is the first step towards an identification. Consequently, sample purification/enrichment has become an essential step for the characterization of protein phosphorylation events, as it offers a reduction in complexity of the sample. Eventually, the phosphorylation sites are identified and characterized by Edman degradation or MS-based methods.
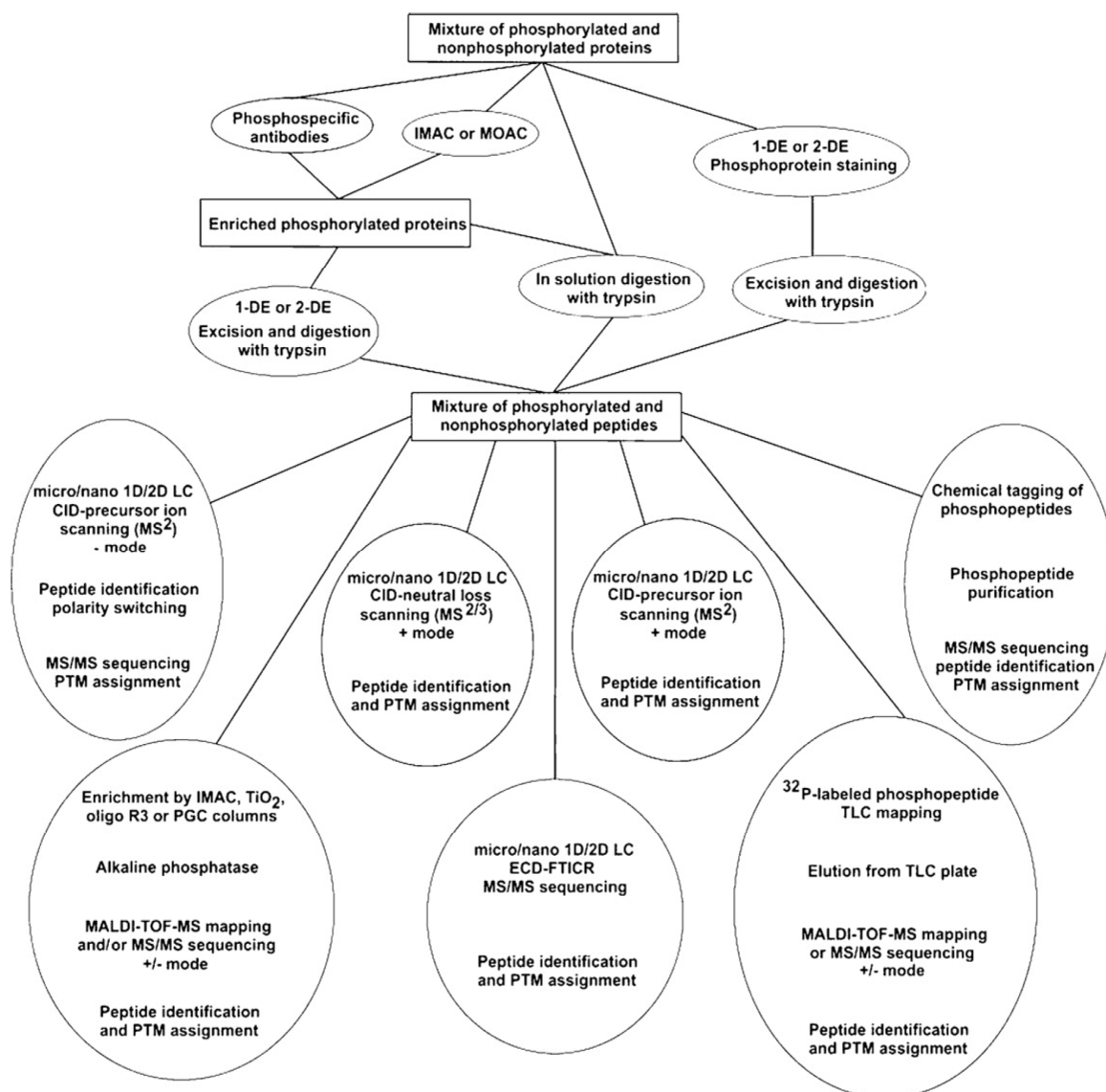
Figure 17. Different techniques for the enrichment and analysis of phosphorlylated proteins based on MS procedures. From ref. [253].

Some of the most commonly used purification, identification and characterization methods are discussed below. At the end of each section the contributions which this doctoral thesis has brought in the field are described.

Phosphorylated protein detection and identification

Phosphoprotein sample treatment

Tissue/Cell lysis frees all the phosphatases and the proteases from normal cellular regulation, thus affecting quickly and substantially the phosphoproteins stoichiometry and number. The phosphatases

dephosphorylate and proteases digest low copy number signaling proteins. Therefore, an essential part of sample preparation for phosphoprotein analysis from intact tissue/cells is to freeze the phosphorylation state of protein in the sample. This can be achieved by inhibiting the enzymatic activity of phosphatases, kinases and proteases after cell lysis. By far the most common means of inhibiting phosphatases in phosphoproteomic studies is through the use of commercially available inhibitors. While vanadium oxides such as pervanadate and orthovanadate are used to inhibit all protein tyrosine phosphatases (PTPs). The three most common serine/threonine phosphatases family, protein phosphatase 1 (PP1), PP2A and PP2B are efficiently inhibited by calyculin A (for PP1 and PP2A) and deltamethrin (PP2B) [254]. Therefore, inadvertent activities can be inhibited by adding commercial available "cocktail"of phosphatase, kinase and protease inhibitors which could lead to increase up to 40% the number of phosphoproteins identified [255]. Moreover, microwaves [256] and rapid heating [257] of samples have been also used for inactivating kinases and phosphatases.

Radioactive Labeling of Phosphoproteins

The radioactive labeling of phosphorylated proteins with ^{32}P has been for a long time the method of choice to study protein phosphorylation due to its sensitivity and selectivity. Usually, for *in-vivo* labeling the ^{32}P *ortho*-phosphate can be added to a cell culture, the endogenous kinases of the cells will incorporate radioactive phosphate groups in the phosphoproteins. Alternatively, *in vitro* labeling experiments are performed by applying radioactive γ-ATP using with specific kinase(s) that incorporate the radioactive phosphate to its specific amino acid and/or substrates. These procedures are followed by the extraction and electrophoretic separation (SDS-PAGE) or (2-DE) of proteins. The detection of phosphorylation is then performed using scintillation counting, by autoradiography or PhosphorImager systems [258]. The ^{32}P-labeling used not only to detect phosphorylation but also to identify the peptide and ultimately the phosphorylated amino acid. This can be achieved by Edman sequencing and/or by mass spectrometry [259,260]. Even though ^{32}P-labeling is still the most sensitive approach for the detection of phosphorylation it has several drawbacks. Firstly, ^{32}P is a radioactive compound must be handled in restricted dedicated laboratory under controlled and stringent safety rules. Secondly, the incorporation of radioactive phosphorus may considerably alter the in vivo state of the cells as shown in comparative studies [261]. Noteworthy, the radioactive labeling can not be applied to *in-vivo* studies of phosphoproteins in tissues.

Direct Staining of Phosphoprotein

Fluorescent-based dye systems, which selectively stain phosphorylated proteins and peptides independently of the kind of phosphorylated residues has been recently introduced [262-264]. These dyes rely on the recognition principle of inorganic phosphate receptors developed by chemists [265,266]. The

structure of the commercially available Pro-Q stain (Molecular Probes) is a proprietary patents but it is likely to be similar to the published structures. These dyes allow detection of phosphorylated proteins and the subsequent visualization of the total protein content (i.e. by coomassie staining) in the same gel. In addition, these stains are compatible with enzymatic in-gel digestion and MS [267]. However, the sensitivity of staining depends on the number of phosphorylated residues present in the protein. About 16 ng for pepsin (1 phosphorylated residue) and 2 ng for alpha-casein (8 phosphorylated residues) were the detection limits [268]. Although it is quite sensitive, it is not sufficient for comprehensive analysis of the phosphoproteome. Additional disadvantages are the failure to differentiate between the three phosphorylated residues and the lower sensitivity when compared to ^{32}P-labeling and antibody based approaches. Moreover, it should be noted that O-sulfonation can also contribute to the staining, albeit apparently to a lesser extent [263].

Detection of Phosphoproteins Employing Phosphatases

Phosphatase treatment on phosphorylated proteins results in hydrolysis of phosphate groups, which changes the net charge and molecular weight of the proteins. Consequently, phosphatase treatment changes the migration behavior of phosphorylated proteins during electrophoretic separation (such as SDS-PAGE or 2-DE). These changes have been exploited to discriminate phosphorylated from unphosphorylated proteins. The starting sample is divided in two aliquots, one of which is treated with phosphatase. Subsequently, both aliquots are 2-DE separated and the proteins properly stained. Phosphorylated proteins are consequently detected by comparison of the 2-DE protein patterns, exploiting the shift to more basic position in the gel due to loss of phosphate groups. Thus, detected phosphoproteins are identified by MS. The specific enzymatic activity of λ-phosphatase (λ-PPase), combined with the high resolution power of 2-DE and MS has led to the identification of some novel phosphoproteins in cultured rat fibroblasts [269]. Some drawbacks of this technique are that the recognition by comparison of phosphorylated/dephosphorylated forms in complex 2-DE patterns is very difficult and could lead to misinterpretations, due to the high level of complexity generated. The method can not distinguish where and which kind of residues are the sites of phosphorylation in the protein. In fact, phosphatase treatment is used as general proof of phosphorylation [267]. In addition, phosphatases are inhibited by orthovanadate and other component used to maintain the integrity of phosphorylation used in sample preparation. Finally, large amount of sample are required (one analysis before and one after phosphatase treatment).

Edman Sequencing

The development of this technique by Pehr Edman in which the amino acid sequence of a protein could be elucidated led to a breakthrough in biotechnology. Edman sequencing relies on the sequential

degradation of the amino acids at the N-terminus of a polypeptide chain, which is coupled to a solid phase. Under mildly alkaline conditions the Phenylisothiocyanate (PTC) is reacted with N-terminal amino group of the coupled peptide. Then, under acidic conditions, the phenylthiocarbamoyl derivative formed is cleaved as a thiazolinone derivative, selectively extracted into an organic solvent and treated with acid to form the more stable phenylthiohydantoin (PTH) amino acid derivative. The resulting amino acid derivatives are analysed using HPLC and the procedure repeated again. Thus, stepwise sequenced by comparing the elution profile of each the sample with standard amino acid derivatives each rum. If a phosphopeptide sample is marked with ^{32}P, phosphorylation site pinpointing can be achieved by concomitant measurement of ^{32}P activity. The major drawbacks are that no more than the first 10-15 residues of the peptides can be sequenced and the proteins have to be purified in advance. Most importantly, each single sequence need several hours, therefore, this technique can not be implemented in high throughput analysis.

Antibodies and specific domain capture molecules

Antibodies directed against phosphor-serine, -threonine, and -tyrosine have been used to generally enrich for proteins phosphorylated at the respective residues in immunoprecipitation experiments. It should be noted that the ability of global antibodies directed against phosphorylated residues is not always related to the similar efficiency in precipitate phosphorylated proteins [270]. This might be due to the better accessibility of the phosphate group located on tyrosine or because of the lower complexity of the structural nature of the threonine and serine antigens compared to tyrosine. In addition, antibodies remaining in the sample can complicate downstream analysis [271,272].

Following phosphorylation site identification, specific antibodies can be exploited for a more focused approach. Indeed, antibodies directed against regions surrounding a phosphorylated residue, are usually quite selective and can be effectively be used to enrich for a phosphoprotein of interest. The major drawbacks of this approach are that the exact residue has to be known and that the generation of these antibodies is expensive and time consuming.

In addition to antibodies there are several different protein specific domains which are capable of binding phosphorylated proteins and peptides. While SH2 (Src Homology 2) and PTB (phosphotyrosine binding) domains recognize predominantly tyrosine phosphorylation, 14-3-3 proteins as well as WW domains bind to proteins phosphorylated on serine and threonine, and FHA (forkhead associated) domains show a preference towards threonine phosphorylation [273-275]. In principle all of these domains or the respective proteins can be used for the enrichment of different sorts of phosphorylated peptides. Moreover, immobilized 14-3-3 proteins have been used for the purification of serine/threonine phosphorylated proteins [276,277]. As described, each of these domains/proteins only recognizes a subset of the total pool of phosphorylated proteins and experimental setups relying on these domains are

therefore not as widely applicable as methods relying on more general affinity mechanisms. However, a combination of a more general affinity chromatography followed by different more restrictive approaches could however, further facilitate the analysis of the phosphoproteome.

Prediction programs and Phosphoprotein databases

As the last years have seen an exponential increase in the number of identified phosphorylation sites, several groups have expended enormous efforts to curate and compile these data into on-line resources. Much of the high quality data, including some high-content data, are curated in UniProt [278], which is probably the best way to make the data available to the wider biological community. In addition, more specific compendiums are available and each provides specific types of tools (e.g., predictors, network/pathway viewers) and/or information (e.g., analytical context in which peptides were identified, quantitative profiles of phosphorylation dynamics after agonist stimulation). There are several phosphorylation site prediction programs available, which rely on different algorithms to elucidate the probability of phosphorylation on specific amino acids for any given protein. Some of these are implemented to search for motifs within proteins that are likely to be phosphorylated by specific Ser/Thr- or Tyr-kinases protein kinases or bind to specific domains. The NetPhos 2.0 server (http://www.cbs.dtu.dk/services/NetPhos/) [279] and the Scansite 2.0 (http://scansite.mit.edu) [280] belong to this predictors. However, because of the enormous complexity of the cellular proteome, prediction programs could lead to false positive results and their output has to be handled with care.

Also, several databases have been constructed, which contain data on experimentally verified phosphorylation sites. The PhosphoSitePlus® (PSP) (http://www.phosphosite.org/) [281], the Phosida [282] and Phospho.ELM 9.0 (http://phospho.elm.eu.org/) [283] contain information from literature experimental and manually cured data not only on phosphorylation but also on other protein PTMs. As all these databases are growing they are becoming more and more useful for the researcher and they will probably play an important role in the future for serving as encyclopaedias of phosphorylation sites and for the development of more sophisticated prediction programs.

A 2-DE and western blotting based approach to detect and identify phosphoproteins in tissue extracts (Publication III)

In recent years antibodies that selectively recognize phosphorylated protein residues have become available [270,284-286], making the use of antibodies a versatile and comprehensive strategy for the analysis of the protein phosphorylation. Unfortunately, some of these antibodies efficiently recognize phosphorylated proteins in western-blot applications but were ineffective to precipitate phosphoproteins [270].

Despite considerable effort to develop new separation methods for proteomics studies, 2-DE, introduced more than 25 years ago, remains one of the most efficient technique for separating the thousands of proteins expressed in a eukaryotic cell [287] including protein variants produced by the co- or post-translational processing [288,289]. Furthermore, in combination with MALDI-MS is commonly and efficiently used for large scale protein analysis and for proteome and phosphoproteome studies.

Here, a rapid and simple strategy to identify phosphorylated proteins in the contest of a tissue was developed. For this purpose, the lung phosphoproteome profiles of C57BL6 mouse strain was analyzed using several well characterized specific monoclonal antibodies (MAb) which recognize phosphoproteins. For the detection of serine phosphorylated proteins a set of 5 different anti-phoshoserine (pSset) and the MPM-2 MAbs were used. Moreover, a set of 4 different monoclonal anti-phoshotyrosine (pYset) and the 4G10 MAbs were used for the detection of tyrosine phosphorylated proteins. Notably, to screening for and to identify phosphorylated lung proteins a combination of five different techniques and technologies were employed (Figure 18) : (1) 2-DE, (2) 2-DE-western blotting with antibodies anti-phosphorylated proteins; (3) computer assisted image analysis and processing; (4) MALDI-TOF Tandem mass spectrometry (MS) with automated spectra acquisition and analysis. Therefore, 250 µg of tissue lysate proteins were separated on 2-DE and stained with Colloidal Coomassie G-250 and the gel image was acquired (2-DE-cCBB). In parallel, 40 µg of total protein extract were separated by 2-DE and subsequently transferred onto PVDF membrane for incubation with antibodies directed against phosphorylated proteins and developed (pAb). With a marker pen, landmarks points (Crosses in Figure) were set around the membranes and the image of phosphorylated proteins acquired (WB-pAb). Subsequently, the total proteins on the same membrane were revealed by cCCB-post staining and the image recorded (WB-CBB). Using the added landmark points, these two images were superimposed and combined by the image analysis software to create a virtual image showing the phosphoproteins and the total proteins together (WB-virtual). Several protein spots from this image, which are in common with the cCBB-2-DE image, were selected and as additional landmarks used to superimpose the WB-CCB image to the 2-DE-cCBB image to decipher phosphorylated proteins on the gel. The highlighted protein spots were then excised from the gel using spot cutter, subjected to in-gel digestion using trypsin, and the proteins identified using MALDI-TOF-MS/MS. Moreover, to increase the sequence coverage and for a more comprehensive analysis, the trypsin digested protein spots were analyzed using two sample preparations: the CHCA thin layer (TL) and DHB matrix layer (ML), while MS/MS data were used to reliable confirm identified proteins.
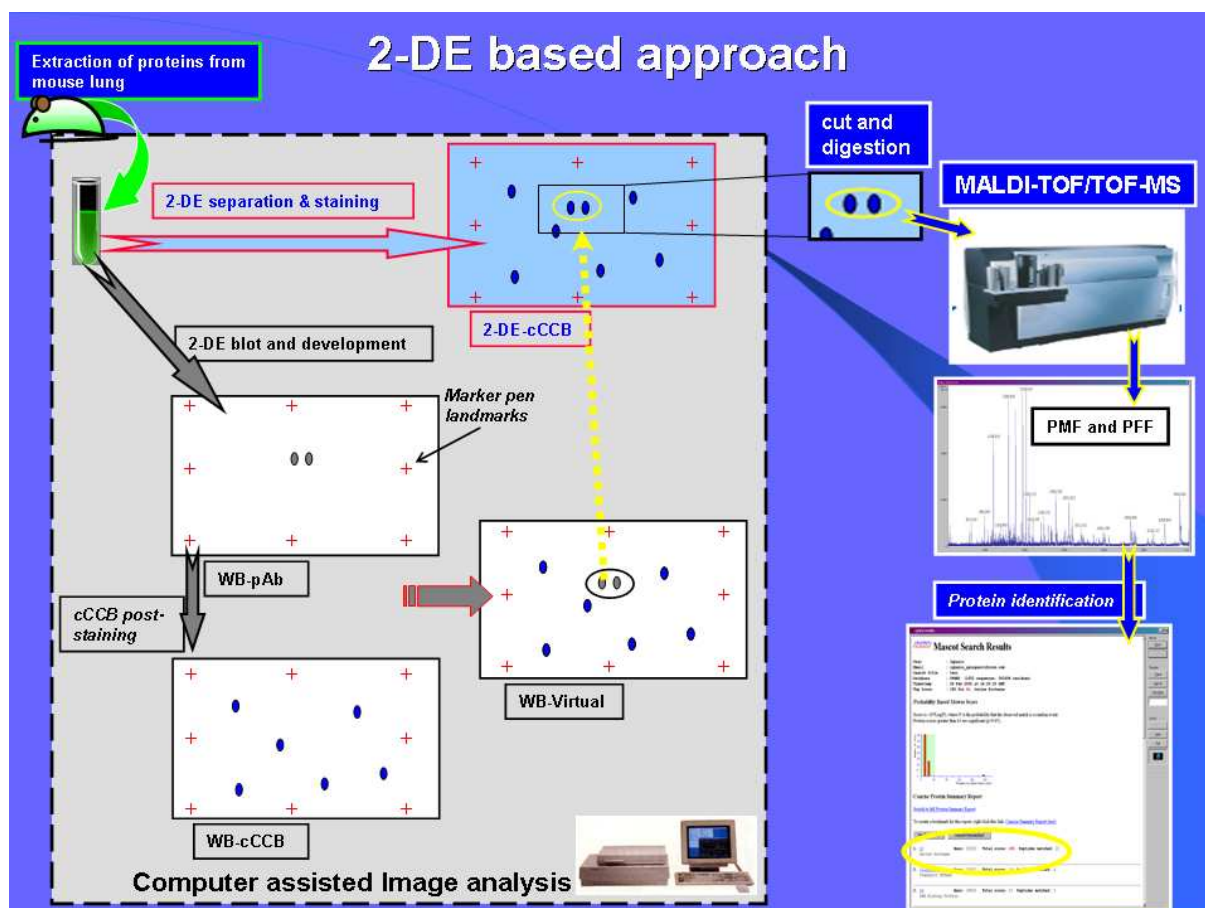
Figure 18. Schematic representation of the combined 2-DE-westr blotting approach.

With this approach 160 unique proteins were identified. The identified putative phosphoproteins were additionally characterized using bioinformatics which allowed the assignment of number and position of potential phosphorylation sites, the search for kinases and phosphorylation-dependent binding motifs, and the classification of the phosphoproteins with respect to their Gene Ontology (GO) annotations. In addition, database searches were used to provide the exact position of the known phosphorylated sites on the based on the available literature. Thus, out of 160 proteins analyzed 106 (66.2%) are already validated phosphoproteins based on the available literature, 26 proteins (16.2%) were predicted in mouse but validated as phosphoproteins either in human or in rat and 30 (18.8%) are only predicted to be phosphorylated and to enharbour a kinase docking domain. The MAbs used demonstrated different specificity in the recognition of phosphorylated proteins. Thus, by the combination of several anti phopsphoprotein MAb with different specificity a more comprehensive characterization of phosphoproteins is achieved. Moreover, broaden cellular distribution and molecular function retrieved from the functional annotation for the identified proteins, provides further evidence of the versatile character of the developed assay.

Protein Phosphorylation site determination

Phosphopeptide Enrichment

Phosphopeptide enrichment is still a necessary step for most phosphoproteins due to suppression effects from the large amount of non-phosphopeptides. In a large scale phosphorylation studies, the general strategy relied on the use of direct phosphopeptide enrichment methods after trypsin digestion of the proteins. In this case, intermediate phosphoprotein enrichments are avoided. However, most strategies applied to phosphoprotein enrichment can apply to phosphopeptide enrichment as well. Much more effort has been put into phosphopeptide enrichment method improvement and innovation, some of the most popular methods will be covered herein.

Chemical derivatization methods

Two main approaches have been presented that make use of chemical derivatization strategies for the enrichment of phosphorylated proteins and peptides. One method consists in β-elimination and Michael addition chemistry where a nucleophile which is suitable for selective enrichment is used to replace the phosphate group of phosphorylated amino acid. However, this method can only be applied to p-Ser/p-Thr residues, because p-Tyr does not undergo β-elimination. [290-293]. In addition, unwanted side reactions which can lead replacement of O-glycosides, sulfonated residues as well as hydroxyl groups on serine and threonine are the limitations of this approach [294].

The alternative is a strategy involving temporary carbodiimide coupling of the phosphate residues of p-Ser/p-Thre/p-Tyr to a solid phase, followed by washing steps to remove nonphosphorylated species [295,296]. This strategy has the distinct advantage of including phosphotyrosine in the analysis and that no side reactions are reported so far. Unfortunately, it involves several modification steps which increase the likelihood of sample loss.

Immobilized metal affinity chromatography (IMAC)

Immobilised Metal Affinity Chromatography (IMAC) is the most widely method used for phosphopeptide enrichment. This technique was initially developed by Porath *et al.* [297] for purification of His-tagged proteins. The method is based on electrostatic interactions of a negatively charged amino acid residue to a positively charged metal that is immobilized on a metal chelator matrix such as iminodiacetic acid (IDA), nitriloacetic acid (NTA) or Tris-(carboxymethyl)-ethylendiamine (TED) chelators. Since a phosphate group has a stronger net negative charge than any other amino acid residue, phosphorylated proteins/peptides are better retained on the matrix than their nonphosphorylated counterparts [294]. Iron and gallium are the used metals for the enrichment of phosphorylated species [298,299]. Despite the wide

application, the major disadvantage of IMAC based methods is a strong non-specific interaction with other negative groups such as side chain carboxylic acids from aspartate and glutamate residues [300,301]. To circumvent the unspecific binding, peptides can be methylesterified using HCl-saturated, dried methanol to convert glutamic and aspartic acid residues into their non-charged methyl esters [302]. Unfortunately, the harsh conditions used (pH 0-1) leading to side reactions such as deamination of glutamines and asparagines [303]. Moreover, incomplete esterification and site reaction render methylesterification especially unsuitable for complex protein mixtures in large scale phosphoproteome-wide analysis.

Metal Oxide Affinity Chromatography (MOAC) with titania and zirconia

Recently, it was shown that metal oxides can also be used for the enrichment of phosphorylated peptides. In these studies affinity chromatography based on titania (titanium dioxide) [304,305] or zirconia (zirconium oxide) [306] was used. Online coupling of a titania pre-column and an anion exchange or reversed phase column in an HPLC setup has been shown to be useful in the selective analysis of phosphorylated peptides derived from proteolytic digests. Identification of the phosphopeptides was achieved by monitoring the UV [307] trace or by using a mass spectrometer [308]. Similarly, Fe3O4/TiO2 core shell particles were used to specifically isolate and detect phosphopeptides [308]. Unspecific binding, which is also reported when using titania, can be reduced by methylesterification [308,309] and the use of appropriate incubation buffers [132]. However, the use of a special incubation buffer is preferable because of the side reactions occurring during methylesterification.

Phosphopeptide identification

Molecular Mass Shift and Phosphatase Treatment

With the molecular mass shift method, spectra of the protein tryptic digest were acquired, usually with MALDI-TOF-MS, and then the peptides mass detected values were compared with theoretically expected values. A molecular mass shift by single or multiples of +80 Da (HPO_3 = 80 Da) indicates a phosphopeptide [310]. Unfortunately, this method has many drawbacks. First: can be applied only to single protein analysis. Second: it is difficult to obtain MS signal of all the peptides of a protein the tryptic digest. Third: detection of phosphopeptides is worse because of the usual low stoichiometry and weak ionization of phosphopeptides as well as signal suppression from non-phosphopeptides.

The use phosphatases combined with MALDI-TOF-MS has been an improvement of this method. Indeed, by comparing spectra of protein tryptic digests with and without phosphatase treatment, a characterisitic shift of one multiple of -80 Da in molecular mass indicates phosphopeptides due to the

cleavage of HPO_3 from the phosphopeptides by phosphatases [311]. Although this approach can be more widely applied, it suffers of the drawbacks related to the decreased MS ionization and low stoichiometry of phosphopeptides.

Phosphopeptide fragmentation by tandem mass spectrometry

As discussed earlier (chapter I), peptide fragmentation occurs through charge directed dissociation, where the amide nitrogen becomes protonated and thus weakens the amide bond making the fragmentation of the peptide possible. In addition, depending on the instrument type used there is a variety of different fragmentation signatures typical for phosphopeptides, which can be used for the detection of phosphorylation and/or for the determination of phosphorylation sites. Every fragmentation technique leads to different fragmentation patterns, some of them leaving the phosphorylation intact, while others evoke the detachment of the phosphate group during the fragmentation process (see introduction/tandem MS).

However, the bond between the peptide and the phospho-group is labile which makes it favorable for fragmentation. Therefore, phospho-group detaches from the peptide easily generating an intense neutral loss ion of 98 Da or 80 Da lower than the measured peptide. These neutral losses represents the losses of H_3PO_4 (98 Da) and HPO_3 (80Da). The loss of 98 Da is observed from the pSer and pThr where as the loss of 80 Da is mostly observed from pTyr (and to some extent from pSer and pThr) [312]. Different mechanisms for neutral loss from pSer and pThr have been proposed [312,313] differing in the model for removal of the hydrogen by the departing phosphate group. The charge-remote loss of H_3PO_4 (98 Da), know as β-elimination, suggests that the hydrogen originates from the α-carbon [312], whereas the charge-directed loss theory proposes that the hydrogen is the mobile proton [314]. Notably, a neutral loss from pTyr produces the loss of HPO_3 (80 Da) [312]. The H_3PO_4 cannot be lost since the bond between the carbon in aromatic ring and the oxygen in the phospho-group is stronger than the corresponding bond in pSer and pThr due to stabilization by the aromatic ring. The bond in the phosphogroup between oxygen and phosphate is weaker so the loss of HPO_3 occurs instead of H_3PO_4 [315]. These characteristic fragmentations were exploited using a number of tandem MS methods for the detection and characterization of phosphorylated peptides [60,316-319]. In particular, Steen *et al.* [320] developed a MS/MS precursor ion scan method for detection of Tyr phosphorylated peptides where they used immonium ion of pTyr. In this method the MS instrument is set to scan 216.043 *m/z* (immonium ion of tyrosine, 136 Da and phosphogroup, 80 Da = 216 Da) which has been reported to be a characteristic fragment ion for pTyr containing peptides.

Identification and characterization of phosphopeptides by MALDI-TOF/TOF MS (Publication III)

Most phosphoproteome studies are based on optimized in-vitro cell culture studies. In order to identify *de novo* phosphoproteins in tissue extracts a new method based on MALDI-TOF MS/MS was developed. Indeed, using DHB ML as sample preparation in combination with MOAC-TiO$_2$ enrichment the peptide phosporylation sites were identified and characterized.

For a molecular characterization of the phosphorylation sites, fifty micrograms of lung extracted proteins were separated by SDS-PAGE and the gel cut in 15 regions. The phosphpeptides of the digested proteins from each region were enriched using a specific developed disposable micro-column and protocol based on MOAC-TiO$_2$ (μ-MOAC-TiO2). Eventually, the enriched samples were analyzed by MALDI-TOF-MS and each peptide signal present in the MS spectrum was subjected to MS/MS fragmentation. It should be noted that the DHB ML delivered an improved homogeneity of the sample preparation when the matrix is "doped" with diammonium hydrogen phosphate. Notably, the yielded spectra contained the characteristic neutral loss of ions at -98 and -80 from the molecular ion as result of β-elimination of H$_3$PO$_4$. Moreover, the high quality of spectra enabled detection of copious signal from the peptide back bone (y- and b- series) leading to an unambiguously characterization of the phosphorylation site(s) and in some cases even close to the complete peptide sequence. In addition, peptide ions of masses of > 2000 Da were detected to yield highly informative and structurally relevant fragment ions. Therefore, with this approach 17 proteins of and 19 phosphorylated peptides were identified and characterized by MALDI-TOF-MS/MS.

While TiO$_2$ have been widely described for the efficient enrichment of phosphorylated peptides [321] none of the presented works demonstrated the implementation with MALDI-TOF tandem MS for the specific characterization of phosphorylation sites. The ML sample preparation in combination with μ-MOAC-TiO$_2$ enrichment allows the successful fragmentation and characterization of novel phosphorylation sites. Indeed, the fragmentation pattern of the peptide at 1526.60 *m/z* (Figure 19) was identified as the peptide R.KAPESQEDEEER.A of the advanced glycosylation end product-specific receptor protein (RAGE_MOUSE). The fragmentation pattern allows not only an unambiguous identification of the peptide through the detection of the full y- and b- ion series but the presence of specific fragment such as characteristic neutral loss and internal fragment ions lead to determine the Ser 377 as the site of phosphorylation. Noteworthy, this protein is described as phosphorylated on threonine 271 but not as phosphorylated on serine (377) in mouse lung.
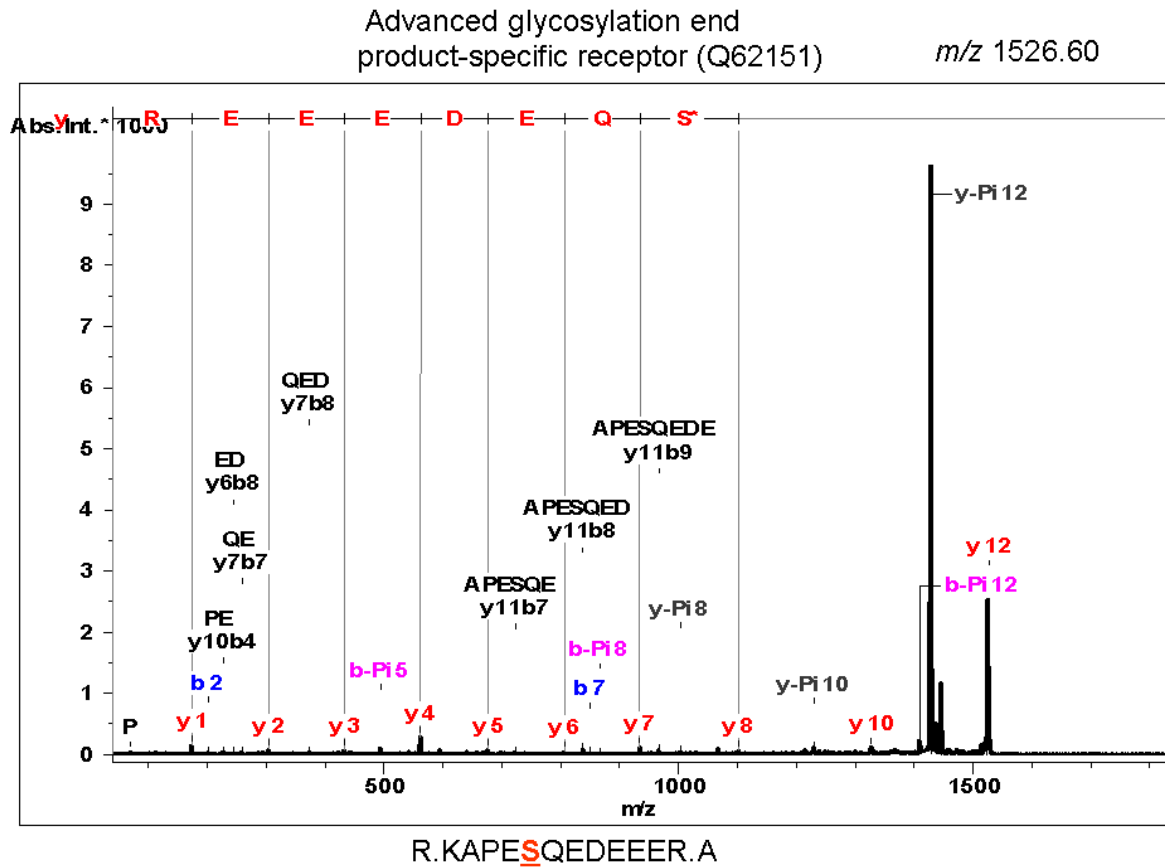
Figure 19. MALDI-TOF MS/MS of phosphopeptides.

The efficient fragmentation of phosphorylated peptide by MALDI-TOF-MS/MS lead to characterization of Ser 377 as a novel phosphorylation site of RAGE_MOUSE protein.

Appendix Chapter III

Garaguso I, & Borlak J. A Rapid screening assay to search for phosphorylated proteins in tissue extracts. Submitted, 2010.

A rapid screening assay to search for phosphorylated proteins in tissue extracts

*Ignazio Garaguso[1,2], Jürgen Borlak [1,2] **

[1]Department of Drug Research and Medical Biotechnology, Fraunhofer Institute of Toxicology and Experimental Medicine, Hannover, Germany

[2]Center for Pharmacology and Toxicology, Hannover Medical School, Hannover, Germany

***Corresponding author:**

Prof. Dr. J. Borlak, Department of Drug Research and Medical Biotechnology Fraunhofer Institute of Toxicology and Experimental Medicine, Nikolai-Fuchs-Str. 1, 30625 Hannover, Germany

Phone: +49-511-5350-559, Fax +49-511-5350-573, E-mail: borlak@item.fraunhofer.de

Keywords:

Lung proteome; MALDI-TOF-MS/MS; phosphoproteins; phosphopetides; TiO_2.

ABSTRACT

Reversible protein phosphorylation is an essential mechanism in the regulation of diverse biological process, but is frequently altered in disease. Furthermore, in order to identify *de novo* phosphoproteins in tissue extracts new methods are in need as most phosphoproteome studies are based on optimized *in-vitro* cell culture studies. Here, we describe a rapid and reliable method for the detection of phosphoproteins in tissue extract based on an immunoblotting strategy that recognize phosphorylated proteins in 2-DE and subsequent identification by MALDI-TOF-MS/MS. This method was applied to lung tissue of mice and resulted in an identification of 160 unique phosphoproteins. Notably, after TiO_2 enrichment of pulmonary protein extracts 17 additional phosphoproteins and 20 phosphorylation sites could be characterized. By use of this method, we report an identification of a phosphorylation site of an advanced glycosylation end product-specific receptor. So far this protein was unknown to be phosphorylated in lung tissue of mice. Overall the developed methodology allowed efficient and rapid screening of phosphorylated proteins and can be employed as a general experimental strategy for an identification of phosphoproteins of tissue extracts.

Introduction

Reversible protein phosphorylation is a major cellular mechanism in the regulation of protein function but such post-translational modifications are accomplished by the activities of protein kinases and reversed by phosphatases in a highly dynamic manner. It was estimated that the mammalian genome codes for more than 500 protein kinases and the proteome contains more than 100 000 phosphorylation sites [1]. The predominant class of protein phosphorylation in eukaryotic cells are O-phosphates and modifications occur on serine (S), threonine (T) and tyrosine (Y) residues, with an estimated relative abundances of 86.4% 11.8% and 1.8%, respectively [2]. Notably, phosphorylation may change activity of proteins to cause translocation, protein–protein interactions or to target proteins for degradation, thereby modulating cellular processes such as signal transduction, cell differentiation, proliferation, metabolic maintenance, cell division, as well as programmed cell death [3]. It is well established that perturbations of the tightly controlled balance between phosphorylation and de-phosphorylation is associated with numerous pathological conditions and thus are the subject of targeted therapies, most notably in the treatment of cancers where hyperactivity of protein kinases is frequently observed [4-7]. To this effect it is highly desirable to study an entire phosphoproteome, the low abundance of phosphoproteins and the stoichiometry of protein phosphorylation, however, hinder routine phosphoproteome investigations. There is therefore considerable interest in the development of facile methodologies to identify phosphorylated proteins and to asses the phosphorylation state in tissues and organs, but such studies remain demanding and challenging. Indeed, studies with tissues and particularly those of low organ weight, such as lung are complicated by the limited availability of biological material. Consequently, only a small number of investigations report the pulmonary proteome and are primarily based on studies with cell lines [8-10] or nasal/bronchoalveolar lavage [11-13]. Importantly, in biomedical research mouse models of disease are frequently employed in the study of molecular mechanisms, yet the mapping of components, regulatory events and substrates in signaling pathways, remain a challenge. Furthermore, research on phosphoproteins benefitted from the availability of antibodies that selectively recognize phosphorylated amino acid residues [14-17], thus enabling a more broad search of phosphoproteins [18] but some may proof to be ineffective in the precipitation of phosphoproteins [14]. In this regard, a variety of experimental strategy for the enrichment and detection of phosphorylated proteins have been developed. Nonetheless, none of these approach can be regarded as fit for all-proposes with the mapping and characterization of phosphoproteins requiring a combination of different methods and experimental strategies [19]. Specifically, Metal oxide affinity chromatography (MOAC) with titanium dioxide (TiO_2) has been proposed for the selective enrichment of phosphopeptides prior to MS [20;21]. This technique is based on the selective interaction of phosphopeptides with porous TiO_2 microspheres (titanspheres) via bidentate binding at the TiO_2 surface and is employed in combination with. MALDI-MS

for the detection of phosphopetides, Nonetheless, the characterization of de novo identified phosphoproteins is generally persued by HPLC coupled to ESI-MS/MS.

Finally, despite considerable advancement in the separation of complex protein mixtures, two-dimensional electrophoresis (2-DE), introduced more than 25 years ago, remains an efficient technique for the study of thousands of proteins expressed in eukaryotic cells [22] including protein variants produced by pre-and post-translational processing [23;24]. Indeed, protocols have been developed to combine 2-DE with MALDI-MS for proteome and phosphoproteome mapping studies.

Here we describe a rapid and simple assay to identify phosphorylated proteins from tissues and report the lung phosphoproteome of the C57BL6 mouse strain. The procedure involves separation of protein extracts by 2-DE and Western immunoblotting (WB). Subsequently, proteins are identified by MALDI-TOF-MS/MS. Furthermore, lung proteins extracts are separated by SDS-PAGE and after in-gel digestion phosphopeptides are enriched with TiO_2-MOAC micro columns, whereas, the phosphorylation sites are characterized by MALDI-TOF-MS/MS. The identified phosphoprotein candidates are further characterized by bioinformatics. This allowes the assignment of number and position of potential phosphorylation sites, the search for kinases and phosphorylation-dependent binding motifs, and the classification of the phosphoproteins with respect to their Gene Ontology (GO) annotations. Finally, in depth database searches provide information on the exact position of known phosphorylated sites. Overall, we demonstrate that 2-DE Western immunoblotting combined with MALDI-TOF-MS can be efficiently used to search for novel phosphorylated proteins. As an example we report the characterization of a phosphorylation site of the advanced glycosylation end product-specific receptor, which so far was not reported as a phosphorylated protein in mouse lung tissue.

Material and methods

Animal care

The C57BL/6 mice were maintained under specific pathogen free conditions. Mice were anesthetized with CO2 and sacrificed. After surgical removal the lung was washed with PBS containing protease and phosphatases inhibitors until free of blood. A total of n= 6 animals were studied. Further details are given in the SI

Materials

The α-cyano-4-hydroxycinnamic acid (CHCA), 2,5-dihydroxybenzoic acid (DHB) and MALDI pre-structured sample support (AnchorChip™ 384/600) were from Bruker Daltonics (Bremen, Germany). Sequencing grade modified trypsin was from Promega Co. (Madison, WI). Protease inhibitor kit III and Phosphatase inhibitor kit II were purchased from Merck (Darmstadt, Germany). The 4G10 and MPM-2 monoclonal antibodies were from Millipore (Billerica, MA), all other antibodies were from ALEXIS (Enzo,

Lörrach, Germany). The IPG pH 3-11 strips and buffer were from GE health Care (München, Germany).
All other chemical and reagent were of the highest grade commercially available.

Proteins extraction

The whole lung tissues were washed two times with wash buffer (20mM Tris-HCl, pH 7.4) containing
phosphatase inhibitors and protease inhibitor. The tissue was disrupted with an ultrasound sonicator in
lysis buffer (0,5%SDS, 100mM DTT in 20 mMTris-HCl, pH 7.4) containing phosphatase inhibitors
(diluted 1:100) and protease inhibitors (diluted 1:100). Subsequently, the proteins were precipitated with
TCA/Acetone. After centrifugation at 20,000xg for 20 min at 4°C, the pellet was re-suspended in buffer
reswelling buffer (8 M Urea, 2 M Thiourea, 2% CHAPS, 2% ASB 14, 65 mM DTT, 5% IPG Buffer pH 3-
11). The protein concentration was determined by the Bradford assay.

High-resolution 2-DE

Two hundred fifty micrograms of total tissue protein extracts were separated with 2-DE and gels were
stained by colloidal Coomassie Brilliant Blue G-250 (cCBB) method as described previously [25].

SDS-PAGE

The proteins were loaded on polyacrylamide gels [stacking: 4%T, 2.6%C (Bis); resolving: 12%T, 2.6%C
(Bis), 2% SDS] and run in a Mini-PROTEAN III, (Bio-Rad, Hercules, CA) according to the manufacturer
recommendations and stained by cCBB.

In-gel digestion of proteins

Protein-bands were cut out from SDS-PAGE gels using a scalpel. Protein-spots from 2-DE gel were cut
using EXQuest™ spot cutter (Bio-Rad, Hercules, CA). In-gel digestion was performed according to
standard protocols [26] with minor modifications. The Protein-bands and protein-spots were rehydrated in
30 µl and 16 µl of a digestion buffer containing 50 mM NH_4HCO_3 and 10 ng/µl of trypsin, respectively.
The digestion was allowed to proceed at 37°C for 4 hours then blocked with a 0.5% TFA solution.

TiO_2 enrichment of phosphopetides

The purification/enrichment of phosphorylated peptides was accomplished by metal oxide affinity
chromatography (MOAC) with TiO_2 beads on self-assembled disposable micro column (µ-MOAC-TiO_2).
Details are given in SI.

MALDI matrix sample preparation protocols

DHB matrix layer (ML) and CHCA thin-layer (TL) sample preparations on AnchorChip™ were carried
out as described by Garaguso & Borlak [27]. Details are given in SI.

Mass spectrometric analysis

MALDI-TOF-MS experiments were performed on an Ultraflex II MALDI-TOF/TOF mass spectrometer
equipped with a SmartBeam™ laser and a LIFT-MS/MS facility. Typically, 400 spectra over a 600–4000

m/z mass range, acquired at 100Hz, were summed and externally calibrated using a standard mixture composed 7 peptides (instrumentation and software from Bruker Daltonics). For details see SI.

Proteins Identification

Automated MALDI-MS spectra acquisition, tandem MS/MS spectra selection and acquisition as well as protein identification were carried out with the Proteinscape software (Bruker daltomics), using SwissProt protein database and the Mascot (v. 2.0, Matrix sciences; UK) software [28]. Sequence database searches were performed using the following search parameters for MS and MS/MS data: all entries; mass tolerance: 100 ppm for parent ion; 0.9 Da for fragments, one missed cleavage site; cysteine residues modified with acrylamide, whereas methionine oxidation as partial modification.

Western blot analysis

After run, 2-DE analytical gel proteins were electro-blotted onto PVDF membrane. The blotted membranes were first incubated with each set of anti-phosphoproteins specific monoclonal antibody (MAb), and then with horseradish peroxidase-goat anti-mouse conjugate polyclonal antibodies. Eventually, the protein-antibody complexes were visualized with OPTI-4-CN kit (Biorad). Details are given in SI.

Image acquisition and analysis

The images were acquired using an Expression 10000 XL scanner (Epson, Germany) and images were analyzed with PDQuest v 8.0.1 (Biorad). Additional details are given in SI.

Prediction and classification of phosphoproteins and phosphorylation sites

The NetPhos 2.0 server (http://www.cbs.dtu.dk/services/NetPhos/) [29] was used to predict the number and position of potential protein phosphorylation sites. Furthermore database searches within PhosphoSitePlus® (PSP) (http://www.phosphosite.org/) [30], Phospho.ELM 9.0 (http://phospho.elm.eu.org/) [31] and UniProt (http://www.uniprot.org/) [32]were carried out to gain information about the exact position of known phosphorylated sites. Scansite 2.0 (http://scansite.mit.edu) [33] was used to predict potential Ser/Thr- or Tyr-kinases that could be responsible for the phosphorylation. Here, motive searches of specific amino acid sequence of phosphorylated protein helped to identify kinases possibly responsible for the phosphorylation. Finally, gene ontology (GO) lists were downloaded using the protein information resource (PIR) (http://pir.georgetown.edu/) [34]. Each protein was classified with respect to its cellular component and molecular function using GO annotation.

A combined 2-DE-Western blotting approach

To screening for and to identify phosphorylated lung proteins a combination of five different techniques and technologies were employed. Notably: (1) 2-DE, (2) 2-DE-western blotting with antibodies anti-phosphorylated proteins; (3) computer assisted image analysis and processing; (4) MALDI-TOF

Tandem mass spectrometry (MS) with automated spectra acquisition and analysis. These technologies have been put together as depicted in Figure 1. A detailed description of the procedure is given in Supporting information (SI).

Results and Discussion

Representative 2-DE maps of tissue lysates together with the immunoblots are shown in SI Figure S1. On average 450 ± 30 protein spots per gel were visualized by colloidal CCB staining (cCCB) at a pI range of 3-11. The identified proteins and relative information are reported in Table S1 in SI

2-DE phosphoprotein analysis of lung lysates

Successful research into protein phosphorylation requires both a selective separation and enrichment procedure, and a reliable method for the detection, identification and characterization of phosphoproteins. When phosphorylated proteins are detected by antibodies, it is important to note that each antibody (Ab) has its own unique properties regarding sensitivity and specificity. In addition, factors such as background from contaminating proteins and nonspecific reactivity of the antibodies toward unrelated proteins could influence the performance as well. In the present study, the employed monoclonal antibodies (MAb) were previously reported for their usefulness in the enrichment of phosphorylated proteins that was followed by MS identification, therefore, confirming the specificity and selectivity [14]of the MAb. Table S2 of the SI informs on the properties and the combination of antibodies used for the search of pulmonary phosphoproteins. Notably, the various MAb were produced with slightly different antigens, that might influence the specificity and binding behavior of the Ab. Next to immunodetection of phosphoproteins their identification by MS and subsequent submission of identified sequences to software predictor and database search of already known literature information confirm the reliability of the identified phopshorylated proteins. Consequently, detection of phoshpoproteins with specific Ab and their identification by high performance MALDI-TOF-MS/MS is a simple and fast method. Moreover, to increase the sequence coverage and for a more comprehensive analysis, the trypsin digested protein spots are analyzed using two sample preparations: the CHCA thin layer (TL) and DHB matrix layer (ML), while MS/MS data are used to reliable confirm identified proteins. Once identified, bioinformatics is used to predict the potential protein phosphorylation sites and database searches provide information about the exact position of known phosphorylated sites. Next to the immunoblotting of phosphoproteins of lung tissue, extracted proteins are also separated by SDS-PAGE, and after in-gel digestion phosphopeptides are enriched with self-made TiO_2 micro column. Here, the phosphorylation sites are identified and characterized by MALDI-TOF-MS/MS.

Detection and identification of phosphorylated proteins with anti phoshoserine antibodies

For the detection of serine phosphorylated proteins a set of 5 different anti-phoshoserine (pSset) and the MPM-2 MAbs are used(SI, Table S2). After development on the PVDF membranes (Figure S1) with the pSset of monoclonal antibodies the corresponding spots are excised from the 2-DE-cCBB, in-gel digested with trypsin to yield 149 unique proteins (SI, Table S1). In Figure 2 an example of the specificity of the pSset of MAbs is given. Three distinct spots detected in the WB with the pSset MAbs also have two corresponding spots in the 2-DE-cCBB. Importantly, with of the CHCA TL sample matrix preparation only the heterogeneous nuclear ribonucleoproteins A2/B1 (O88569) is identified and confirmed by MS/MS. However, when the digests of the same spots are analyzed by the DHB ML sample preparation Glyceraldehyde-3-phosphate dehydrogenase (P16858) is additionally identified and confirmed by MS/MS. Therefore, a careful comparison of the spot in the developed immunoblot in the 2-DE-cCBB stained gel and the visual inspection of region in which the spots appeared confirm co-migration of both proteins as adjacent spots. Notably, the MS and MS/MS spectra are acquired in an automated procedure Consequently, the use of two different MALDI sample preparation methods in combination enabled more reliable identification of phosphoproteins. Furthermore, the pSset of MAbs identified a large number of kinases such as: the SRC kinase (P05480), inhibitor of nuclear factor kappa-B kinase subunit beta (O88351), integrin-linked protein kinase (O55222), serine-protein kinase ATM (Q62388) mitogen-activated protein kinase 3 (Q63844). With the MPM-2 MAb 43 unique proteins are identified (SI, Table S1) from the positive spots on the WB membrane which have an equivalent in the 2-DE-cCBB gel. Despite the reduced number of proteins detected with the MPM-2 MAb 6 unique proteins are identified which are not detected by the pSmix Abs, possible due to slight differences in specificity. One of these proteins is the heat shock protein HSP 90-alpha (HS90A). This chaperone is essential for proper protein folding, stabilization, and trafficking of an expanding list of proteins, but appears to be critical for a variety of signal transduction pathways through interaction with a wide range of transcription factors and protein kinases. Moreover, inhibition of HS90A activity is being pursued for the treatment of various kinds of cancer and particularly lung cancer [35;36]. Moreover, the use of the software Scansite predicted HS90A as a substrate for Fgr Kinase, Src Kinase, Lck Kinase, Akt Kinase, Casein Kinase 2 and Erk D-domain. In addition, database search as with Phospho.ELM and PhosphoSitePlus confirmed the phosphorylation of serine, threonine and tyrosine in mouse and human based on other experimental data.

Detection and identification of phosphorylated proteins with anti phoshotyrosine antibodies

For the detection of tyrosine phosphorylated proteins a set of 4 different monoclonal anti phoshotyrosine (pYset) and the 4G10 MAbs are used (SI, Table S2). With the pYmix MAbs (SI, Figure S1) the positive spots on the membrane are excised and digested from the 2-DE-cCBB to enable an identification of 54 unique proteins. The 4G10 is a well characterized MAb and recognizes phospho-tyrosine residues in

proteins as well as in peptides and does not cross-react with phospho-serine or phospho-threonine peptides or proteins. A comparison of the Western blots show that pYmix detected the greater number of protein spots, whereas 4G10 exhibited a similar reactivity but with fewer protein spots. With the 4G10 MAb (SI, Figure S1) 40 phosphoproteins are identified of which 6 are uniquely recognized by this MAb.

Bioinformatics analysis of phosphoproteins

Functional annotation of identified phosphoproteins was carried out by categorizing the proteins into different groups based on GO terms. Figures 3 a and b depict a pie-chart distribution of the identified proteins cataloged according to the cellular component or molecular function (SI, Table S3). Out of the 160 total phosphoproteins identified in total their distribution was assigned to 21.2% in the cytoplasm, 17.9% in the nucleus, 11.9% in the mitochondra, 16.6% in the cytoskeleton, 15.2% in the plasma membrane and 7.3% as extracellular. Such a distribution provides further evidence of the versatile character of the developed assay in the search for novel phosphoproteins. The molecular functions of the 160 annotated proteins are consistent with the roles that phosphoproteins play in biological processes including protein binding (22.2%), nucleotide binding (15.7%), catalytic activity (9.8%), enzyme regulation activity (8.5%), enzyme binding (5.9%), and signal transduction activity (6.5%). Since phosphorylation is known to extensively control biological functions and activities, it is not surprising that the functions of most of the identified phosphoproteins are related to the functions of lung.

The PhosphoSitePlus®, Phospho.ELM 9.0 and the UniProt database search were used to provide the exact position of the known phosphorylation sites based on the available literature. Furthermore, the putative phosphorylation sites of each protein with different software packages were analyzed. With Netphos 2.0, an output score of 0.5 is used as cutoff to ensure that the site is a *bona fide* phosphorylation site. Notably, for all of the 160 identified proteins several phosphorylation sites are predicted. Each phosphoprotein is further searched within Scansite 2.0, with high and medium stringency, to predict the kinase and phosphorylation-dependent binding motifs. Potential binding sites for the Erk D-domain, PDK1 Binding, DNA PK, ATM kinase and casein kinase are most commonly predicted; casein kinase 1 and 2, PKC epsilon and zeta sites are also frequently found. Thus, out of 160 proteins analyzed 106 (66.2%) are already validated phosphoproteins based on the available literature, 26 proteins (16.2%) were predicted in mouse but validated as phosphoproteins either in human or in rat and 30 (18.8%) are only predicted to be phosphorylated and to enharbour a kinase docking domain (SI, Table S1). Therefore a total of 130 (81.25%) were validated phosphoproteins and described

De novo identification of proteins and characterization of phosphoryation site by MALDI-TOF-MS/MS

Several reports describe the use of titania in metal oxide affinity chromatography (MOAC-TiO$_2$) for the successful enrichment of phosphorylated peptides. We find most of these protocols not useful for our

sample preparation due to the elevated concentration of DHB in the elution solution and the increased backpressure obtained [20]. Therefore, we developed a disposable micro-column (μ-MOAC-TiO$_2$) in which inert glass microfiber is used as frit and the composition of the elution solution is modified as reported in SI. To further characterize the position of the phosphorylation site, trypsin digests of positive spots are enriched by μ-MOAC-TiO$_2$ and subsequently applied to the pre-structured sample support for MS analysis. Then MALDI-TOF-MS spectra were searched for the presence of phosphorylated peptide signals. None of the digested proteins gave acceptable signals of phosphorylated peptides (data not shown). Considering the low stoichiometry of phosphoporylation the total amount of peptides in the 2-DE spots digests is simple insufficient for an identification of phosphopetides. Therefore, tissue extracted proteins separated by SDS-PAGE were enriched as follow: fifty micrograms of lung extracted proteins are loaded on SDS-PAGE and the lane is cut into 15 regions. The digested proteins from each region are enriched using μ-MOAC-TiO$_2$ and analyzed by MALDI-TOF-MS. Each peptide signal present in the MS spectrum is subjected to MS/MS fragmentation. With this approach 17 additional proteins of and 19 phosphorylated peptides were identified and characterized by MALDI-TOF-MS/MS. Here, 14 peptides contain a single pSer residue, 3 peptides contain two and 1 peptide contian 3 pSer residues. Moreover, one protein with a single pSer residue also contains pThre residue as reported in Table 1. The yielded spectra contained the characteristic neutral loss of ions at -98 and -80 from the molecular ion as result of β-elimination of H$_3$PO$_4$. The high quality of spectra enabled detection of copious signal from the peptide back bone (y- and b- series) leading to an unambiguously characterization of the phosphorylation site(s) and in some cases even close to the complete peptide sequence. In addition, peptide ions of masses of > 2000 Da were detected to yield highly informative and structurally relevant fragment ions as shown in Figure 4A for the peptide R.TPEELDD**S**DFETEDFDVR.S of alpha-1 catenin (P26231). In Figure 4 B the MALDI-TOF-MS/MS positive spectrum of peptide *m/z* 1526.5210 corresponds to the peptide R.KAPE**S**QEDEEER.A of the advanced glycosylation end product-specific receptor protein (RAGE_MOUSE). The fragmentation pattern allows not only an unambiguous identification of the peptide through the main fragmentation signal but the identification of the Ser 377 as the definitive phosphorylation site. Noteworthy, this protein is described as phosphorylated on threonine 271 but not as phosphorylated on serine (377) in mouse lung. While TiO$_2$ have been widely described for the efficient enrichment of phosphorylated peptide peptides none of the presented works demonstrated the feasibility of such enrichment for MALDI-MS for the specific characterization of phosphorylation sites. Indeed, the copious neutral loss in MS/MS spectra shadows the signal of the backbone peptide. With the CHCA TL sample preparation a reduced number of phosphrylated peptides, copious matrix clusters, and increased fragmentation renders the spectra impossible for data analysis (data not shown). However, the DHB ML sample preparation delivered an improved homogeneity and

increased backbone peptide fragments signal in MS/MS experiments when the matrix was doped with diammonium hydrogen phosphate. Therefore, the ML sample preparation in combination with μ-MOAC-TiO$_2$ enrichment allows the successful fragmentation and characterization of novel phosphorylation sites.

Overall, two different approaches were used in the search of phosphoproteins. The first approach is based on immunoblots with anti-phosphorylated proteins MAbs, while the second approach is based on enrichment with TiO$_2$. The immunoblot procedure does not always define all the possible residues subjected to posttranslational modifications of the investigated protein. Indeed, as depicted in Figure 5, the phosphoserine specific MAbs pSset recognize Aldehyde dehydrogenase (P47738) and Selenium-binding protein 1 (P17563), but the 4G10, MPM-2 and pYset MAbs detectes only one of these proteins. Notably, these two proteins are phosphorylated at different sites as reported elsewhere. Thus, by use of several anti phopsphoprotein Ab with different specificity a more comprehensive characterization of phosphoproteins is achieved. Furthermore, the identified proteins enharbour a kinases binding motif as determined by bioinformatics. Here, 66 % of all the detected proteins have been reported to be phosphorylated at such sites in other organisms (Human and/or rat) and with MALDI-TOF-MS/MS the phosphorylation site can be identified, as evidenced in the present study.

In regards to the sensitivity of the developed assay this is defined by the staining system of 2-DE. In other words, for a protein to be characterized, it must be visible on the gel. However, visibility is a necessary but not always a sufficient criterion. In several cases, MS failed to unambiguously assign a name to a poorly visible spot in the 2-DE-cCBB gel. Therefore, the staining system truly limits the sensitivity of the assay.

Conclusions

In conclusion an antibody based rapid screening method for detection and de novo identification of phosphoprotein is reported and this procedure is applied to lung tissue for an identification of pulmonary phosphoproteins. The facile use of MALDI-TOF-MS/MS for the characterization of phosphopetides is therefore demonstrated. The assay can readily be applied to any tissue for the search of phosphoprotein and to obtain information about on/off states with regard to protein phosphorylation. In the same way it is possible to study the effects of drug treatment with kinase inhibitors on the phosphoproteome. The proposed assay provides the capability to qualify protein phosphorylation status on a systematic scale and, therefore, can be employed in biomedical research.

Acknowledgements

Supporting Information Available

Experimental procedures and images. This material is available free of charge via the Internet at htpp://pubs.acs.org.

Table 1

Protein Entry Name	Accession	Sequence	Start end	Meas. mass	Ion Score	Miss	MW
Angiotensin-converting enzyme	P09470	R.GPQFG**S**EVELR.H	1300-1310	1298.57765	63	0	151.888
Calnexin precursor	P35564	K.AEEDEILNR**S**PR.N	573-584	1508.67402	39	1	67.733
Advanced glycosylation end product-specific receptor	Q62151	R.KAPE**S**QEDEEER.A	373-384	1526.60058	60	1	43.068
Nascent polypeptide-associated complex alpha	P70670	R.**S**V**T**DPAM**A**PRTAK.N	588-600	1600.58786	20	1	221.603
Matrix-remodeling-associated protein 7	Q9CZH7	R.VAEPEE**S**EAEEPAAEGR.Q	73-89	1879.75921	82	0	19.516
Tensin	gi|226437589	R.**S**QSFPDVEPQLPQAPTR.G	790-806	1976.91134	56	0	203.317
Arginase-1	Q61176	-.**M**SSKPK**S**LEIIGAPFSK.G	1-17	2075.89248	35	0	34.999
60S acidic ribosomal protein	P99027	K.KEE**S**EE**S**DDD**M**GFGLFD.-	99-115	2125.68687	40	1	11.651
60S acidic ribosomal protein P1	P47955	K.KEE**S**EE**S**EDD**M**GFGLFD.-	98-114	2139.70251	40	1	11.610
Myosin phosphatase Rho-interacting protein	P97434	R.AEEQLPPLLSPP**S**PSTPHSR.R	280-299	2220.06959	27	0	117.357
Alpha-1 catenin	P26231	R.TPEELDD**S**DFETEDFDVR.S	634-651	2238.85983	127	0	100.896
Membrane-associated progesterone receptor	O55022	K.EGEEPTVY**S**DDEEPKDETAR.K	173-192	2375.93983	112	1	21.692
60S acidic ribosomal protein P0	P14869	K.AEAKEE**S**EE**S**DED**M**GFGLFD.-	298 - 317	2410.81924	45	1	34.408
Septin-2	P42208	K.IYHLPDAE**S**DEDEDFKEQTR.L	210-229	2517.04527	115	0	41.783
EH domain-containing protein	Q8BH64	R.GPDEAIEDGEEG**S**EDDAEWVVTK.D	426-448	2557.01375	112	1	50.135
Elongation factor 1-delta	P57776	K.GATPAEDDEDKDIDLFG**S**DEEEEDKEAAR.L	145-173	3276.32225	144	2	31.916
Serum deprivation-response protein	Q63918	R.GNNSAVG**S**NADLTIEEDEEEEPVALQQAQQVR.Y	356-387	3520.57109	135	0	46.806
		R.RGNNSAVG**S**NADLTIEEDEEEEPVALQQAQQVR.Y	355-387	3676.67219	105	1	
		K.SSPFKV**S**PLSFGR.K	287-299	1488.72465	53	1	

Table 1. Identified phosphopeptides by MALDI tandem MS using DHB ML sample preparation.

Phosphopeptides identified by MALDI-TOF-MS/MS from 50 µg of protein extracted from mouse lung and separated by SDS-PAGE.

S= phosphorylated serine determined by MALDI TOF-MS/MS. M= oxidized methionine
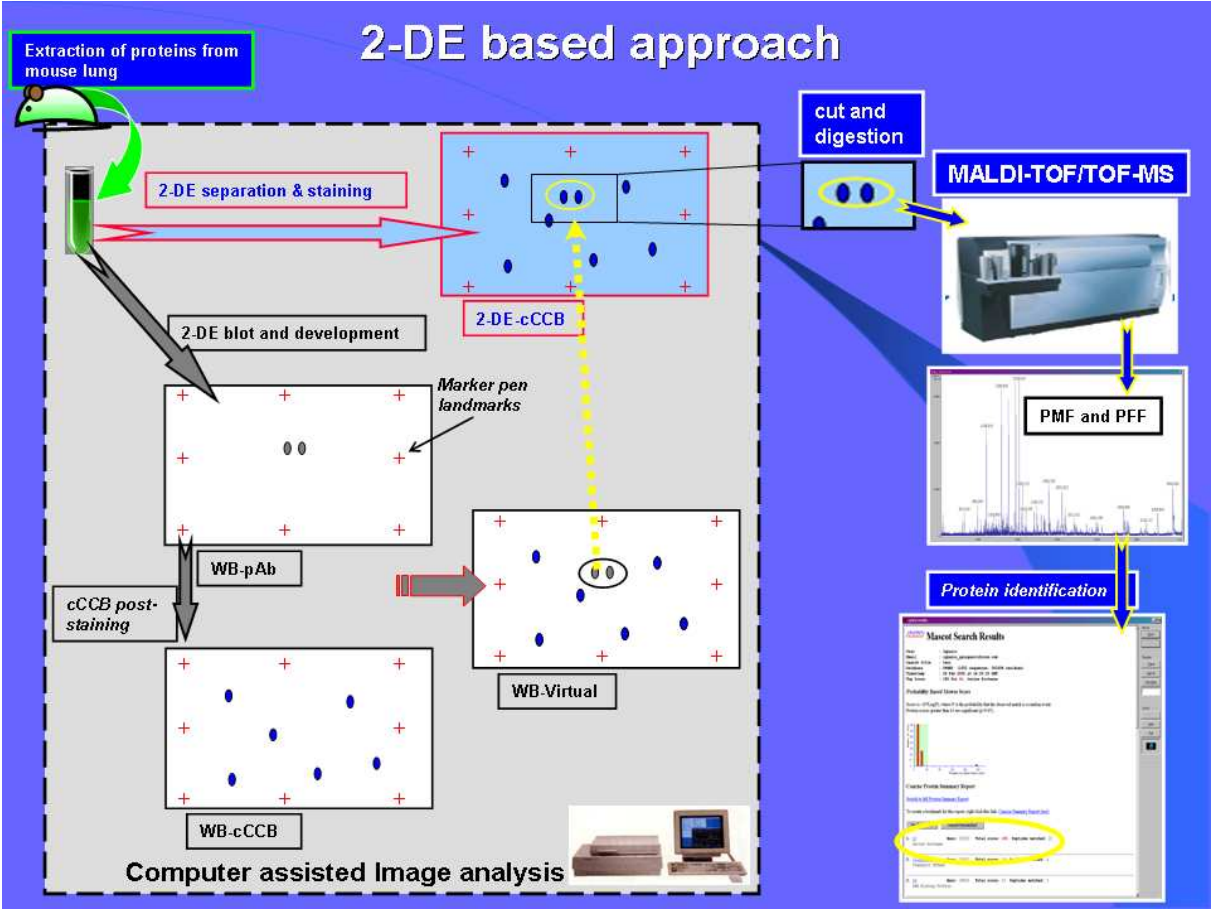
Figures

Figure 1

Figure 1. Description of the combined 2-DE-WB approach

Two hundred microgramms of tissue lysate proteins were separated on 2-DE and stained with Colloidal Coomassie G-250 and acquired as an image (2-DE-cCBB). In parallel 40 µg of total protein extract are separated by 2-DE and subsequently transferred onto PVDF membrane for incubation with antibodies directed against phosphorylated proteins and developed (pAb). With a marker pen, landmarks points (Crosses in Figure) are set around the membrane and the image of phosphorylated proteins is acquired (WB-pAb). Subsequently, the total proteins on the same membrane are revealed by cCCB-post staining and the image is recorded (WB-CBB). Using the marker added landmarks, these two images are superimposed and combined by the image analysis software to create a virtual image showing the phosphoproteins and the total proteins together (WB-virtual). Several protein spots from this image, which are in common with the cCBB-2-DE image, are selected as additional landmarks used to superimpose the WB-CCB image to the 2-DE-cCBB image to decipher phosphorylated proteins on the gel. The highlighted protein spots are then excised from the gel using spot cutter, subjected to in-gel digestion using trypsin, and the proteins identified using MALDI-TOF MS.

Figure 2
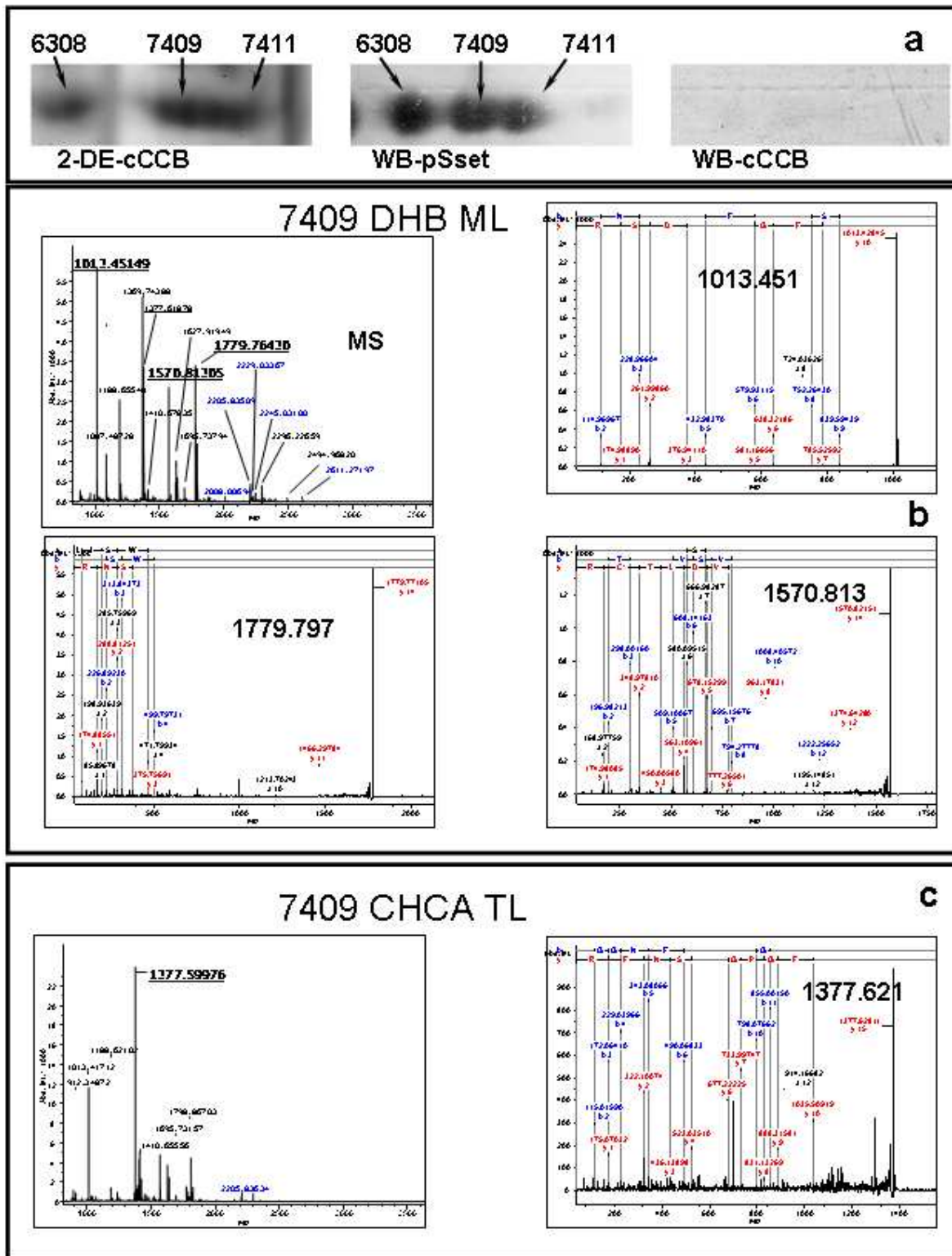
Figure 2. Specificity of the pSset MAbs and use of two MALDI sample preparation.

The use two different MALDI sample preparation methods in combination enabled more reliable identification With the CHCA TL method only the heterogeneous nuclear ribonucleoproteins A2/B1 (O88569) was identified and confirmed by MS/MS. However, when the digests of the same spots are analyzed by the DHB ML sample preparation Glyceraldehyde-3-phosphate dehydrogenase (P16858) was additionally identified and confirmed by MS/MS.

Figure 3
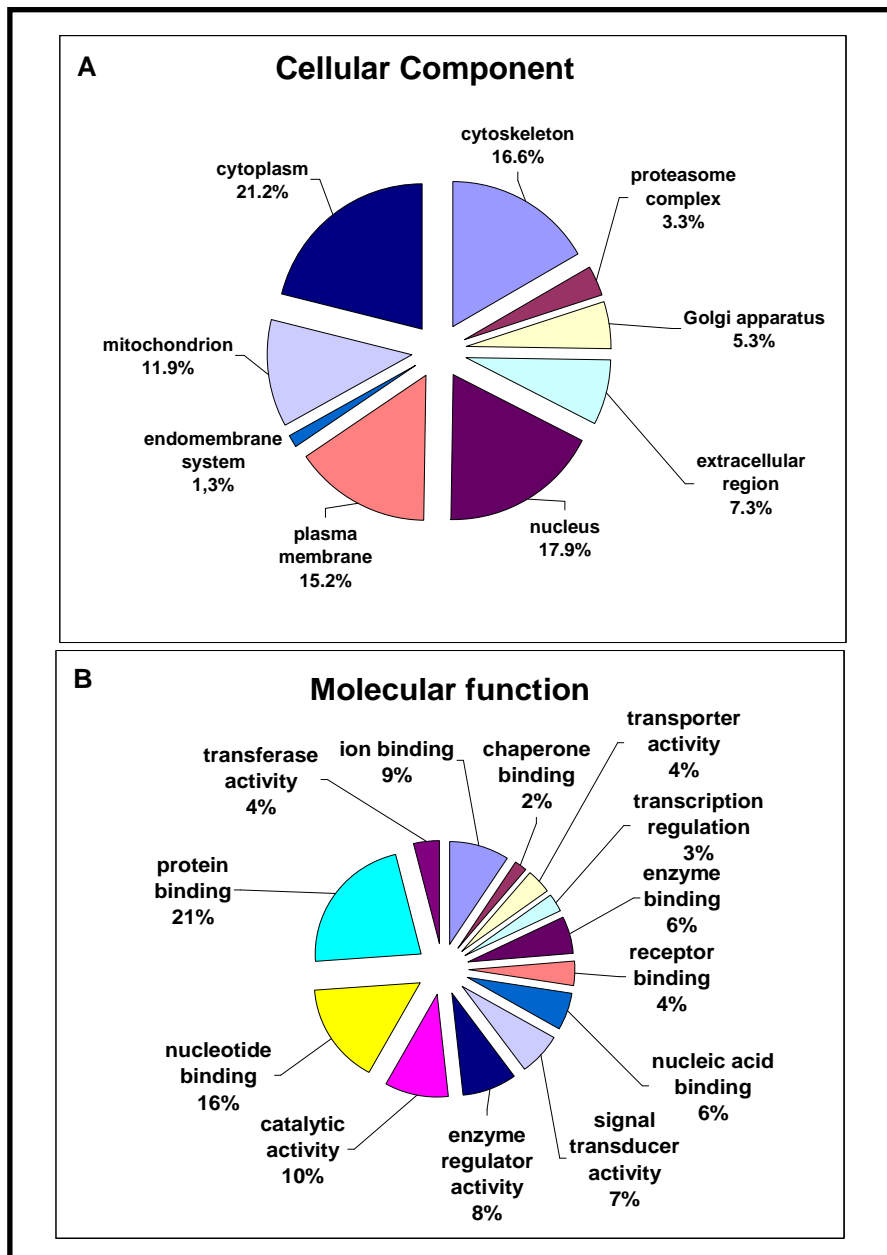
Figure 3.

Distribution of identified proteins according to the (a) cellular component or the (b) molecular function in which they are involved. Assignments were made on the basis of information provided by GO lists downloaded from PIR (http://pir.georgetown.edu/)
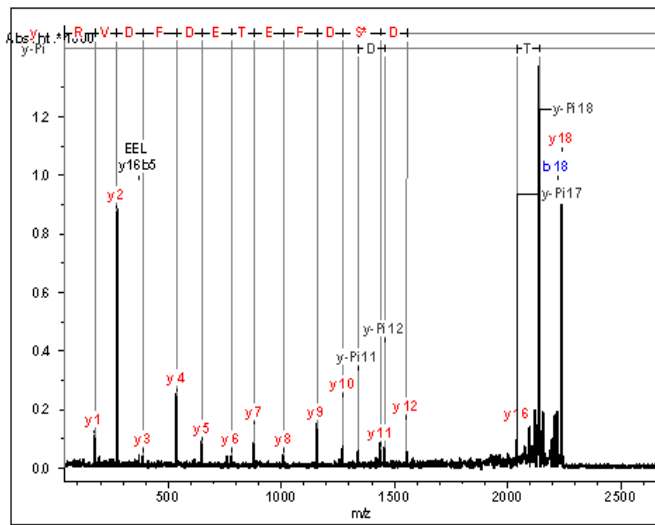
Figure 4

A

m/z m/z 2238.8598

Alpha-1 catenin (P26231)

Ion score: 127

R.TPEELDD**S**DFETEDFDVR.S
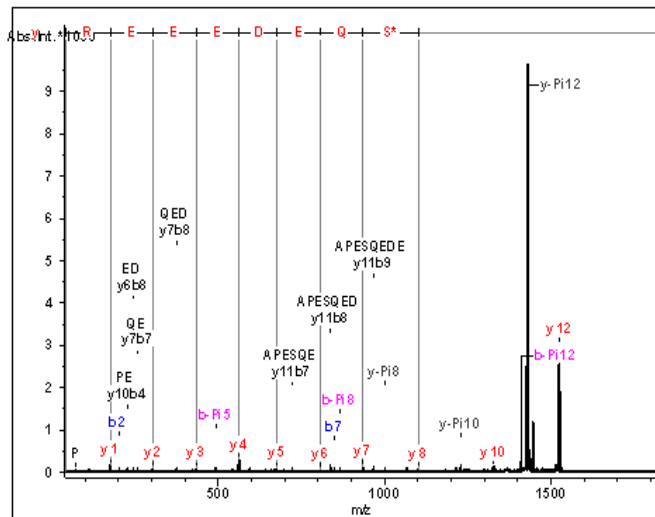
B

m/z 1526.6006

Advanced glycosylation end product-specific receptor (Q62151)

Ion score: 60

R.KAPE**S**QEDEEER.A

Figure 4. MALDI-TOF MS/MS of phosphopeptides.

The fragmentation of phosphorylated peptides by MALDI-TOF-MS/MS leads to characterization of the phosphorylation site.

Figure 5
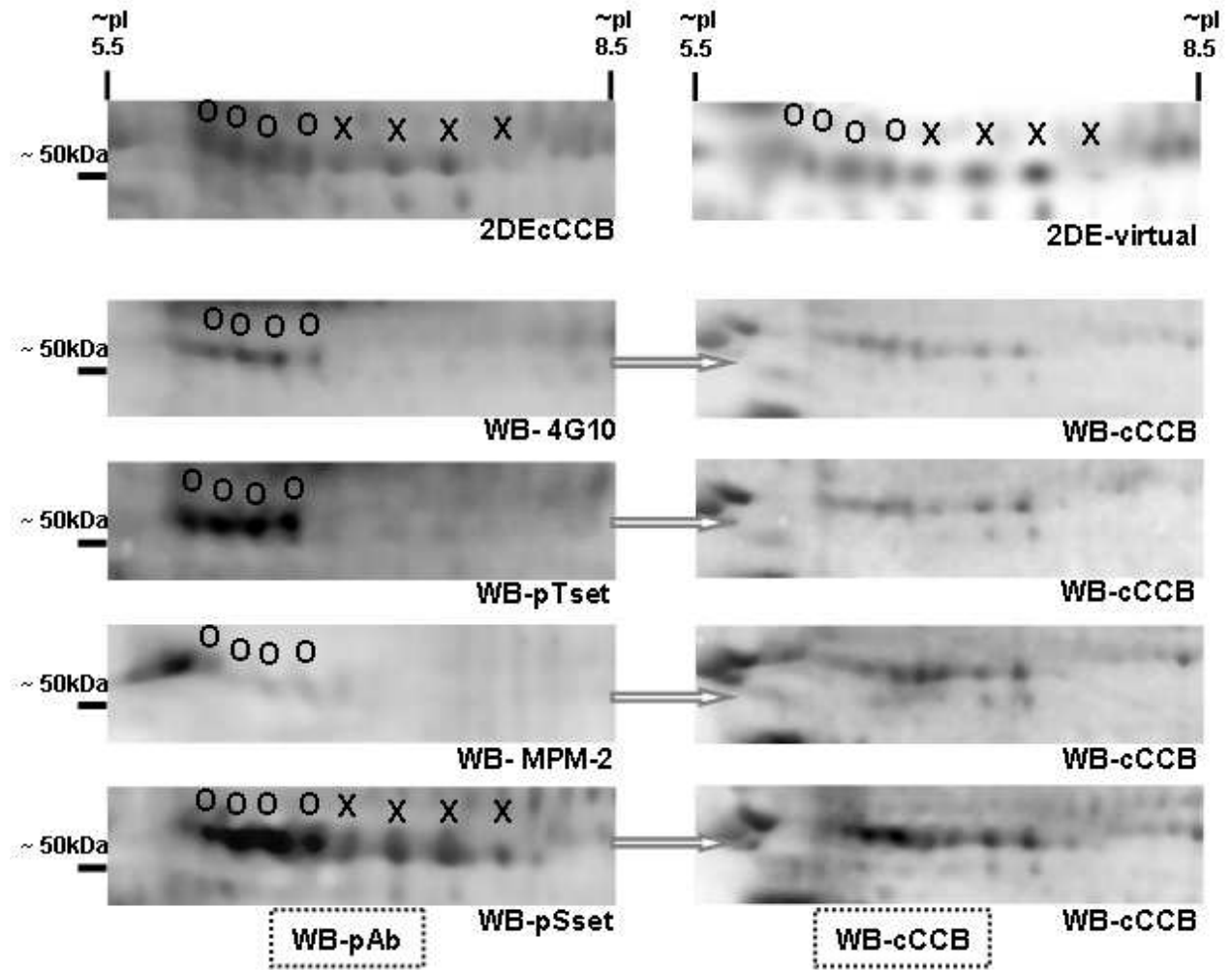
Figure 5. Specificity of the anti-phosphoproteins antibodies

The phosphoserine specific antibodies pSset recognize Aldehyde dehydrogenase (P47738) and Selenium-binding protein 1 (P17563), whereas the MPM-2, sYset and 4G10 only the first.

O= spots identified as Selenium-binding protein 1 (P17563); X= spots identified as Aldehyde dehydrogenase (P47738). On the right side are shown the same membrane after cCCB post staining. For detail on the superimposition see SI.

References

1. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. *Science* **2002,** *298*, 1912-34.

2. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. *Cell* **2006,** *127*, 635-48.

3. Zhang, X.; Jin, Q. K.; Carr, S. A.; Annan, R. S. *Rapid Commun.Mass Spectrom* **2002,** *16*, 2325-32.

4. Blume-Jensen, P.; Hunter, T. *Nature* **2001,** *411*, 355-65.

5. Zheng, X.; Wang, J.; Haerry, T. E.; Wu, A. Y.; Martin, J.; O'Connor, M. B.; Lee, C. H.; Lee, T. *Cell* **2003,** *112*, 303-15.

6. Gazdar, A. F.; Shigematsu, H.; Herz, J.; Minna, J. D. *Trends Mol Med* **2004,** *10*, 481-86.

7. Lim, Y. P. *Clin Cancer Res* **2005,** *11*, 3163-69.

8. Leung, K. Y.; Wait, R.; Welson, S. Y.; Yan, J. X.; Abraham, D. J.; Black, C. M.; Pearson, J. D.; Dunn, M. J. *Proteomics* **2001,** *1*, 787-94.

9. Predic, J.; Soskic, V.; Bradley, D.; Godovac-Zimmermann, J. *Biochemistry* **2002,** *41*, 1070-78.

10. Linthicum, D. S. *Tissue Cell* **2001,** *33*, 514-23.

11. Ghafouri, B.; Stahlbom, B.; Tagesson, C.; Lindahl, M. *Proteomics* **2002,** *2*, 112-20.

12. Noel-Georis, I.; Bernard, A.; Falmagne, P.; Wattiez, R. *Dis Markers* **2001,** *17*, 271-84.

13. Stripp, B. R.; Reynolds, S. D.; Boe, I. M.; Lund, J.; Power, J. H.; Coppens, J. T.; Wong, V.; Reynolds, P. R.; Plopper, C. G. *Am J Respir Cell Mol Biol* **2002,** *27*, 170-78.

14. Gronborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara, O.; Mann, M.; Jensen, O. N.; Pandey, A. *Mol Cell Proteomics.* **2002,** *1*, 517-27.

15. Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J. *Nat.Biotechnol.* **2005,** *23*, 94-101.

16. Davis, F. M.; Tsao, T. Y.; Fowler, S. K.; Rao, P. N. *Proc Natl Acad Sci U S A* **1983,** *80*, 2926-30.

17. Cohen, B.; Yoakim, M.; Piwnica-Worms, H.; Roberts, T. M.; Schaffhausen, B. S. *Proc Natl Acad Sci U S A* **1990,** *87*, 4458-62.

18. Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. *Proc Natl Acad Sci U S A* **1999,** *96*, 6591-96.

19. Delom, F.; Chevet, E. *Proteome Sci* **2006,** *4*, 15.

20. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. *Mol Cell Proteomics* **2005,** *4*, 873-86.

21. Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. *Anal Chem* **2004,** *76*, 3935-43.

22. Kristensen, D. B.; Kawada, N.; Imamura, K.; Miyamoto, Y.; Tateno, C.; Seki, S.; Kuroki, T.; Yoshizato, K. *Hepatology* **2000,** *32*, 268-77.

23. Yamagata, A.; Kristensen, D. B.; Takeda, Y.; Miyamoto, Y.; Okada, K.; Inamatsu, M.; Yoshizato, K. *Proteomics* **2002,** *2*, 1267-76.

24. Guy, G. R.; Philip, R.; Tan, Y. H. *Electrophoresis* **1994,** *15*, 417-40.

25. Ferrari, G.; Garaguso, I.; Adu-Bobie, J.; Doro, F.; Taddei, A. R.; Biolchi, A.; Brunelli, B.; Giuliani, M. M.; Pizza, M.; Norais, N.; Grandi, G. *Proteomics.* **2006,** *6*, 1856-66.

26. Jensen, O. N.; Wilm, M.; Shevchenko, A.; Mann, M. *Methods Mol Biol* **1999,** *112*, 513-30.

27. Garaguso, I.; Borlak, J. *Proteomics* **2008,** *8*, 2583-95.

28. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999,** *20*, 3551-67.

29. Blom, N.; Gammeltoft, S.; Brunak, S. *J Mol Biol* **1999,** *294*, 1351-62.

30. Hornbeck, P. V.; Chabra, I.; Kornhauser, J. M.; Skrzypek, E.; Zhang, B. *Proteomics* **2004,** *4*, 1551-61.

31. Diella, F.; Cameron, S.; Gemund, C.; Linding, R.; Via, A.; Kuster, B.; Sicheritz-Ponten, T.; Blom, N.; Gibson, T. J. *BMC Bioinformatics* **2004,** *5*, 79.

32. Jain, E.; Bairoch, A.; Duvaud, S.; Phan, I.; Redaschi, N.; Suzek, B. E.; Martin, M. J.; McGarvey, P.; Gasteiger, E. *BMC Bioinformatics* **2009,** *10*, 136.

33. Obenauer, J. C.; Cantley, L. C.; Yaffe, M. B. *Nucleic Acids Res* **2003,** *31*, 3635-41.

34. Wu, C. H.; Huang, H.; Arminski, L.; Castro-Alvear, J.; Chen, Y.; Hu, Z. Z.; Ledley, R. S.; Lewis, K. C.; Mewes, H. W.; Orcutt, B. C.; Suzek, B. E.; Tsugita, A.; Vinayaka, C. R.; Yeh, L. S.; Zhang, J.; Barker, W. C. *Nucleic Acids Res* **2002,** *30*, 35-37.

35. Shimamura, T.; Shapiro, G. I. *J Thorac.Oncol* **2008,** *3*, S152-S159.

36. Pratt, W. B. *Proc Soc Exp Biol Med* **1998,** *217*, 420-34.

Concluding Remarks

Mass spectrometry (MS) is increasingly becoming indispensable analytical technique in biological sciences because of its sensitivity and selectivity.

In this thesis novel MALDI-TOF MS based methods were developed to improve detection, identification and characterization of DNA adducts, proteins and protein phosphoryation. The development of a new "matrix layer" (ML) sample preparation and robust procedures for analyte enrichment/purification proofed to be extremely useful for reliable and sensitive detection and molecular characterization of these biomolecules.

The new developed MALDI-TOF DHB ML sample preparation method (Publication I) demonstrated improved peptide ionization and fragmentation, which provided higher quality MS information relevant for increased protein sequence coverage. Indeed, the new method permitted unambiguous peptides and proteins identification as well as characterization through MS/MS fragmentation. Moreover, this method enable automated MALDI MS and MS/MS data acquisition allowing high throughput required in proteomic studies.

Simple and efficient procedures to enrich/purify DNA adducts molecules were developed. In addition, a new calibration mixture and MALDI-TOF instrument settings were developed and optimized to operate in the low mass range. These improvements in combination with ML preparation method allowed sensitive detection, identification and characterization of PAH-DNA adducts (Publication II). Consequently, the MALDI-TOF MS and CID-MS/MS spectra permitted unambiguous identification of deoxynucleotide and deoxynucleosides adducts as well as the chemical nature determination of the alkylating agent at the same time.

The development of a combined approach based on 2-DE, western blotting and MALDI-TOF tandem MS (Publication III) allowed efficient detection and identification of phosphoproteins from mouse lung tissue. Moreover, the μ-MOAC-TiO$_2$ demonstrated to be a cost effective and robust alternative for the enrichment of phosphopeptides out of complex protein mixtures. Indeed, the use DHB ML sample preparation in combination with μ-MOAC-TiO$_2$ enrichment allowed the successful fragmentation and characterization of phosphopeptides leading to the determination of novel phosphorylation sites.

Taken collectively, this doctoral thesis demonstrated that a single mass spectrometric technology platform enables sensitive detection and reliable identification through molecular characterization of cancer-related biomolecules on a systematic scale. Therefore, these new developed MALDI-TOF based methods can be applied to biomedical and cancer research for the development of discovery strategies to search, select and isolate biomarker.

Acknowledgments

References

Reference List

1. WHO . Fact sheet N°297, February 2011. 2010. 2010.
Ref Type: Online Source

2. Ferlay, J.; Shin, H. R.; Bray, F.; Forman, D.; Mathers, C.; Parkin, D. M. GLOBOCAN 2008, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Internet]. 2010. Lyon, France: International Agency for Research on Cancer. 2010.
Ref Type: Online Source

3. Jemal, A.; Bray, F.; Center, M. M.; Ferlay, J.; Ward, E.; Forman, D. Global cancer statistics. *CA Cancer J. Clin.* **2011,** *61* (2), 69-90.

4. Conrad, D. H.; Goyette, J.; Thomas, P. S. Proteomics as a method for early detection of cancer: a review of proteomics, exhaled breath condensate, and lung cancer screening. *J Gen. Intern. Med* **2008,** *23 Suppl 1*, 78-84.

5. Reed, M. F.; Molloy, M.; Dalton, E. L.; Howington, J. A. Survival after resection for lung cancer is the outcome that matters. *Am J Surg.* **2004,** *188* (5), 598-602.

6. Shaw, A. T.; Kirsch, D. G.; Jacks, T. Future of early detection of lung cancer: the role of mouse models. *Clin Cancer Res* **2005,** *11* (13 Pt 2), 4999s-5003s.

7. Hecht, S. S. Tobacco carcinogens, their biomarkers and tobacco-induced cancer. *Nat Rev Cancer* **2003,** *3* (10), 733-744.

8. Wogan, G. N.; Hecht, S. S.; Felton, J. S.; Conney, A. H.; Loeb, L. A. Environmental and chemical carcinogenesis. *Semin. Cancer Biol* **2004,** *14* (6), 473-486.

9. IARC . Agents Classified by the IARC Monographs, Volumes 1–100. 2010. 10.
Ref Type: Online Source

10. Boffetta, P.; Jourenkova, N.; Gustavsson, P. Cancer risk from occupational and environmental exposure to polycyclic aromatic hydrocarbons. *Cancer Causes Control* **1997,** *8* (3), 444-472.

11. Zayas, B.; Stillwell, S. W.; Wishnok, J. S.; Trudel, L. J.; Skipper, P.; Yu, M. C.; Tannenbaum, S. R.; Wogan, G. N. Detection and quantification of 4-ABP adducts in DNA from bladder cancer patients. *Carcinogenesis* **2007,** *28* (2), 342-349.

12. Grimmer, G.; Stober, W.; Jacob, J.; Mohr, U.; Schoene, K.; Brune, H.; Misfeld, J. Inventory and biological impact of polycyclic carcinogens in the environment. *Exp Pathol* **1983,** *24* (1), 3-13.

13. Grimmer, G.; Brune, H.; Deutsch-Wenzel, R.; Naujack, K. W.; Misfeld, J.; Timm, J. On the contribution of polycyclic aromatic hydrocarbons to the carcinogenic impact of automobile exhaust condensate evaluated by local application onto mouse skin. *Cancer Lett.* **1983,** *21* (1), 105-113.

14. Schwartz, A. G.; Prysak, G. M.; Bock, C. H.; Cote, M. L. The molecular epidemiology of lung cancer. *Carcinogenesis* **2007,** *28* (3), 507-518.

15. Smith, L. E.; Denissenko, M. F.; Bennett, W. P.; Li, H.; Amin, S.; Tang, M.; Pfeifer, G. P. Targeting of lung cancer mutational hotspots by polycyclic aromatic hydrocarbons. *J Natl Cancer Inst* **2000,** *92* (10), 803-811.

16. Denissenko, M. F.; Pao, A.; Tang, M.; Pfeifer, G. P. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* **1996,** *274* (5286), 430-432.

17. Shimada, T.; Inoue, K.; Suzuki, Y.; Kawai, T.; Azuma, E.; Nakajima, T.; Shindo, M.; Kurose, K.; Sugie, A.; Yamagishi, Y.; Fujii-Kuriyama, Y.; Hashimoto, M. Arylhydrocarbon receptor-dependent induction of liver and lung cytochromes P450 1A1, 1A2, and 1B1 by polycyclic aromatic hydrocarbons and polychlorinated biphenyls in genetically engineered C57BL/6J mice. *Carcinogenesis* **2002,** *23* (7), 1199-1207.

18. Shimada, T.; Fujii-Kuriyama, Y. Metabolic activation of polycyclic aromatic hydrocarbons to carcinogens by cytochromes P450 1A1 and 1B1. *Cancer Sci* **2004,** *95* (1), 1-6.

19. Mukherjee, J. J.; Gupta, S. K.; Kumar, S. Inhibition of benzopyrene-diol-epoxide (BPDE)-induced bax and caspase-9 by cadmium: role of mitogen activated protein kinase. *Mutat Res* **2009,** *661* (1-2), 41-46.

20. Katiyar, S. K.; Matsui, M. S.; Mukhtar, H. Ultraviolet-B exposure of human skin induces cytochromes P450 1A1 and 1B1. *J Invest Dermatol.* **2000,** *114* (2), 328-333.

21. Chong, P. K.; Lee, H.; Kong, J. W.; Loh, M. C.; Wong, C. H.; Lim, Y. P. Phosphoproteomics, oncogenic signaling and cancer research. *Proteomics* **2008,** *8* (21), 4370-4382.

22. Mattsson, A.; Jernstrom, B.; Cotgreave, I. A.; Bajak, E. H2AX phosphorylation in A549 cells induced by the bulky and stable DNA adducts of benzo[a]pyrene and dibenzo[a,l]pyrene diol epoxides. *Chem Biol Interact.* **2009,** *177* (1), 40-47.

23. Yan, C.; Lu, J.; Zhang, G.; Gan, T.; Zeng, Q.; Shao, Z.; Duerksen-Hughes, P. J.; Yang, J. Benzo[a]pyrene induces complex H2AX phosphorylation patterns by multiple kinases including ATM, ATR, and DNA-PK. *Toxicol In Vitro* **2010**.

24. Albino, A. P.; Huang, X.; Jorgensen, E.; Yang, J.; Gietl, D.; Traganos, F.; Darzynkiewicz, Z. Induction of H2AX phosphorylation in pulmonary cells by tobacco smoke: a new assay for carcinogens. *Cell Cycle* **2004,** *3* (8), 1062-1068.

25. Tannheimer, S. L.; Ethier, S. P.; Caldwell, K. K.; Burchiel, S. W. Benzo[a]pyrene- and TCDD-induced alterations in tyrosine phosphorylation and insulin-like growth factor signaling pathways in the MCF-10A human mammary epithelial cell line. *Carcinogenesis* **1998,** *19* (7), 1291-1297.

26. Rodriguez-Fragoso, L.; Melendez, K.; Hudson, L. G.; Lauer, F. T.; Burchiel, S. W. EGF-receptor phosphorylation and downstream signaling are activated by benzo[a]pyrene 3,6-quinone and benzo[a]pyrene 1,6-quinone in human mammary epithelial cells. *Toxicol Appl Pharmacol* **2009,** *235* (3), 321-328.

27. Ou, X.; Weber, T. J.; Chapkin, R. S.; Ramos, K. S. Interference with protein kinase C-related signal transduction in vascular smooth muscle cells by benzo[a]pyrene. *Arch Biochem Biophys.* **1995,** *318* (1), 122-130.

28. Ou, X.; Ramos, K. S. Benzo[a]pyrene inhibits protein kinase C activity in subcultured rat aortic smooth muscle cells. *Chem Biol Interact.* **1994,** *93* (1), 29-40.

29. Henklova, P.; Vrzal, R.; Ulrichova, J.; Dvorak, Z. Role of mitogen-activated protein kinases in aryl hydrocarbon receptor signaling. *Chem Biol Interact.* **2008,** *172* (2), 93-104.

30. Zhang, X.; Jin, Q. K.; Carr, S. A.; Annan, R. S. N-Terminal peptide labeling strategy for incorporation of isotopic tags: a method for the determination of site-specific absolute phosphorylation stoichiometry. *Rapid Commun. Mass Spectrom* **2002,** *16* (24), 2325-2332.

31. Ullrich, A.; Schlessinger, J. Signal transduction by receptors with tyrosine kinase activity. *Cell* **1990,** *61* (2), 203-212.

32. Blume-Jensen, P.; Hunter, T. Oncogenic kinase signalling. *Nature* **2001,** *411* (6835), 355-365.

33. Zheng, X.; Wang, J.; Haerry, T. E.; Wu, A. Y.; Martin, J.; O'Connor, M. B.; Lee, C. H.; Lee, T. TGF-beta signaling activates steroid hormone receptor expression during neuronal remodeling in the Drosophila brain. *Cell* **2003,** *112* (3), 303-315.

34. Gazdar, A. F.; Shigematsu, H.; Herz, J.; Minna, J. D. Mutations and addiction to EGFR: the Achilles 'heal' of lung cancers? *Trends Mol Med* **2004,** *10* (10), 481-486.

35. Dass, C. *Fundamentals of Contemporary Mass Spectrometry;* Wiley-Interscience, John Wiley & Sons, Inc.: 2007.

36. Dass, C. *Principles and Practice of Biological Mass Spectrometry;* Wiley-Interscience, John Wiley & Sons, Inc.: 2000.

37. de Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications, 3rd Edition;* 3rd ed.; Wiley-Interscience, John Wiley & Sons, Inc.: 2007.

38. Karas, M.; Hillenkamp, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem* **1988,** *60* (20), 2299-2301.

39. Tanaka, K.; Waki.H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry* **1988,** *Volume 2* (Issue 8), 151-153.

40. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246* (4926), 64-71.

41. Whitehouse, C. M.; Dreyer, R. N.; Yamashita, M.; Fenn, J. B. Electrospray interface for liquid chromatographs and mass spectrometers. *Anal Chem* **1985,** *57* (3), 675-679.

42. Dole, M.; Mack, L. L.; Hines, R. L.; Fergusson, L. D.; Alice, M. B. Molecular beams of macroions. *J. Chem. Phys.* **1968,** *49*, 2240-2250.

43. Iribarne, J. V.; Thomson, B. A. On the evaporation of small ions from charged droplets. *J. Chem. Phys.* **1976,** *64*, 2287-2294.

44. Cech, N. B.; Enke, C. G. Practical implications of some recent studies in electrospray ionization fundamentals. *Mass Spectrom Rev* **2001,** *20* (6), 362-387.

45. Karas, M.; Bachmann, D.; Hillenkamp, F. Influence of the Wavelength in High-Irradiance Ultraviolet Laser Desorption Mass Spectrometry of Organic Molecules. *Anal. Chem.* **1985,** *57* (14), 2935-2939.

46. Tanaka, K. The origin of macromolecule ionization by laser irradiation (Nobel lecture). *Angew. Chem Int Ed Engl.* **2003,** *42* (33), 3860-3870.

47. Knochenmuss, R. Ion formation mechanisms in UV-MALDI. *Analyst* **2006,** *131* (9), 966-986.

48. Karas, M.; Gluckmann, M.; Schafer, J. Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom* **2000,** *35* (1), 1-12.

49. Knochenmuss, R. A quantitative model of ultraviolet matrix-assisted laser desorption/ionization. *J Mass Spectrom* **2002,** *37* (8), 867-877.

50. Knochenmuss, R. A quantitative model of ultraviolet matrix-assisted laser desorption/ionization including analyte ion generation. *Anal Chem* **2003,** *75* (10), 2199-2207.

51. Knochenmuss, R.; Zenobi, R. MALDI ionization: the role of in-plume processes. *Chem Rev* **2003,** *103* (2), 441-452.

52. Birks, J. B. *Organic molecular photophysics;* J. Wiley, Wiley-Interscience: London , 1973; Vol. 1.

53. Setz, P. D.; Knochenmuss, R. Exciton mobility and trapping in a MALDI matrix. *J Phys. Chem A* **2005,** *109* (18), 4030-4037.

54. Bruker Daltonics Internet Communication 2010.

55. Kaufmann, R.; Chaurand, P.; Kirsch, D.; Spengler, B. Post-source decay and delayed extraction in matrix-assisted laser desorption/ionization-reflectron time-of-flight mass spectrometry. Are there trade-offs? *Rapid Commun. Mass Spectrom* **1996,** *10* (10), 1199-1208.

56. Manyrin, B. A. Time-of-flight mass spectrometry (concepts, achievements, and prospects). *International Journal of Mass Spectrometry* **2001,** *206* (3), 251-266.

57. Hipple, J. A.; Condon, E. U. Detection of metastable ions with the mass spectrometer. *Phys. Rev.* **1945,** *68*, 54.

58. Lange, V.; Picotti, P.; Domon, B.; Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst. Biol* **2008,** *4*, 222.

59. Hopfgartner, G.; Varesio, E.; Tschappat, 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.

60. Wu, H. Y.; Liao, P. C. Analysis of protein phosphorylation using mass spectrometry. *Chang Gung. Med J* **2008,** *31* (3), 217-227.

61. Louris, J. N.; Wright, L. G.; Cooks, R. G.; Schoen, A. E. New scan modes accessed with a hybrid mass spectrometer. *Anal. Chem.* **2010,** *57* (14), 2918-2924.

62. Spengler, B.; Kirsch, D.; Kaufmann, R. Fundamental aspects of postsource decay in matrix_assisted laser desorption mass spectrometry. 1. Residual gas effects. *J. Phys. Chem.* **1992,** *96* (24), 9678-9684.

63. Kaufmann, R.; Spengler, B.; Lutzenkirchen, F. Mass spectrometric sequencing of linear peptides by product-ion analysis in a reflectron time-of-flight mass spectrometer using matrix-assisted laser desorption ionization. *Rapid Commun. Mass Spectrom* **1993,** *7* (10), 902-910.

References

153

64. Sickmann, A.; Marcus, K.; Schafer, H.; Butt-Dorje, E.; Lehr, S.; Herkner, A.; Suer, S.; Bahr, I.; Meyer, H. E. Identification of post-translationally modified proteins in proteome studies. *Electrophoresis* **2001,** *22* (9), 1669-1676.

65. Cornish, T. J.; Cotter, R. J. A curved-field reflectron for improved energy focusing of product ions in time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom* **1993,** *7* (11), 1037-1040.

66. Spengler, B. Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. *J Mass Spectrom* **1997,** *32* (10), 1019-1036.

67. Bienvenut, W. V.; Deon, C.; Pasquarello, C.; Campbell, J. M.; Sanchez, J. C.; Vestal, M. L.; Hochstrasser, D. F. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* **2002,** *2* (7), 868-876.

68. Medzihradszky, K. F.; Campbell, J. M.; Baldwin, M. A.; Falick, A. M.; Juhasz, P.; Vestal, M. L.; Burlingame, A. L. The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF/TOF tandem mass spectrometer. *Anal Chem* **2000,** *72* (3), 552-558.

69. Reiber, D. C.; Grover, T. A.; Brown, R. S. Identifying proteins using matrix-assisted laser desorption/ionization in-source fragmentation data combined with database searching. *Anal Chem* **1998,** *70* (4), 673-683.

70. Papayannopoulos, I. A. The interpretation of collision-induced dissociation tandem mass spectra of peptides. *Mass Spectrometry Reviews* **1995,** *14* (1), 49-73.

71. Reiber, D. C.; Brown, R. S.; Weinberger, S.; Kenny, J.; Bailey, J. Unknown peptide sequencing using matrix-assisted laser desorption/ionization and in-source decay. *Anal Chem* **1998,** *70* (6), 1214-1222.

72. Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal Bioanal. Chem* **2003,** *376* (7), 952-965.

73. Macht, M.; Asperger, A.; Deininger, S. O. Comparison of laser-induced dissociation and high-energy collision-induced dissociation using matrix-assisted laser desorption/ionization tandem time-of-flight (MALDI-TOF/TOF) for peptide and protein identification. *Rapid Commun. Mass Spectrom* **2004,** *18* (18), 2093-2105.

74. Brown, R. S.; Lennon, J. J. Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. *Anal Chem* **1995,** *67* (21), 3990-3999.

75. Wasinger, V. C.; Cordwell, S. J.; Cerpa-Poljak, A.; Yan, J. X.; Gooley, A. A.; Wilkins, M. R.; Duncan, M. W.; Harris, R.; Williams, K. L.; Humphery-Smith, I. Progress with gene-product mapping of the Mollicutes: Mycoplasma genitalium. *Electrophoresis* **1995,** *16* (7), 1090-1094.

76. Wilkins, M. R.; Pasquali, C.; Appel, R. D.; Ou, K.; Golaz, O.; Sanchez, J. C.; Yan, J. X.; Gooley, A. A.; Hughes, G.; Humphery-Smith, I.; Williams, K. L.; Hochstrasser, D. F. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y.)* **1996,** *14* (1), 61-65.

77. James, P. Protein identification in the post-genome era: the rapid rise of proteomics. *Q. Rev Biophys.* **1997,** *30* (4), 279-331.

78. O'Farrell, P. H. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* **1975,** *250* (10), 4007-4021.

79. Edman, P. A method for the determination of amino acid sequence in peptides. *Arch Biochem* **1949,** *22* (3), 475.

80. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proc Natl Acad Sci U S A* **1993,** *90* (11), 5011-5015.

81. Lander, E. S.; Linton, L. M.; Birren, B.; Nusbaum, C.; Zody, M. C.; Baldwin, J.; Devon, K.; Dewar, K.; Doyle, M.; FitzHugh, W.; Funke, R.; Gage, D.; Harris, K.; Heaford, A.; Howland, J.; Kann, L.; Lehoczky, J.; Levine, R.; McEwan, P.; McKernan, K.; Meldrim, J.; Mesirov, J. P.; Miranda, C.; Morris, W.; Naylor, J.; Raymond, C.; Rosetti, M.; Santos, R.; Sheridan, A.; Sougnez, C.; Stange-Thomann, N.; Stojanovic, N.; Subramanian, A.; Wyman, D.; Rogers, J.; Sulston, J.; Ainscough, R.; Beck, S.; Bentley, D.; Burton, J.; Clee, C.; Carter, N.; Coulson, A.; Deadman, R.; Deloukas, P.; Dunham, A.; Dunham, I.; Durbin, R.; French, L.; Grafham, D.; Gregory, S.; Hubbard, T.; Humphray, S.; Hunt, A.; Jones, M.; Lloyd, C.; McMurray, A.; Matthews, L.; Mercer, S.; Milne, S.; Mullikin, J. C.; Mungall, A.; Plumb, R.; Ross, M.; Shownkeen, R.; Sims, S.; Waterston, R. H.; Wilson, R. K.; Hillier, L. W.; McPherson, J. D.; Marra, M. A.; Mardis, E. R.; Fulton, L. A.; Chinwalla, A. T.; Pepin, K. H.; Gish, W. R.; Chissoe, S. L.; Wendl, M. C.; Delehaunty, K. D.; Miner, T. L.; Delehaunty, A.; Kramer, J. B.; Cook, L. L.; Fulton, R. S.; Johnson, D. L.; Minx, P. J.; Clifton, S. W.; Hawkins, T.; Branscomb, E.; Predki, P.; Richardson, P.; Wenning, S.; Slezak, T.; Doggett, N.; Cheng, J. F.; Olsen, A.; Lucas, S.; Elkin, C.; Uberbacher, E.; Frazier, M.; Gibbs, R. A.; Muzny, D. M.; Scherer, S. E.; Bouck, J. B.;

Sodergren, E. J.; Worley, K. C.; Rives, C. M.; Gorrell, J. H.; Metzker, M. L.; Naylor, S. L.; Kucherlapati, R. S.; Nelson, D. L.; Weinstock, G. M.; Sakaki, Y.; Fujiyama, A.; Hattori, M.; Yada, T.; Toyoda, A.; Itoh, T.; Kawagoe, C.; Watanabe, H.; Totoki, Y.; Taylor, T.; Weissenbach, J.; Heilig, R.; Saurin, W.; Artiguenave, F.; Brottier, P.; Bruls, T.; Pelletier, E.; Robert, C.; Wincker, P.; Smith, D. R.; Doucette-Stamm, L.; Rubenfield, M.; Weinstock, K.; Lee, H. M.; Dubois, J.; Rosenthal, A.; Platzer, M.; Nyakatura, G.; Taudien, S.; Rump, A.; Yang, H.; Yu, J.; Wang, J.; Huang, G.; Gu, J.; Hood, L.; Rowen, L.; Madan, A.; Qin, S.; Davis, R. W.; Federspiel, N. A.; Abola, A. P.; Proctor, M. J.; Myers, R. M.; Schmutz, J.; Dickson, M.; Grimwood, J.; Cox, D. R.; Olson, M. V.; Kaul, R.; Raymond, C.; Shimizu, N.; Kawasaki, K.; Minoshima, S.; Evans, G. A.; Athanasiou, M.; Schultz, R.; Roe, B. A.; Chen, F.; Pan, H.; Ramser, J.; Lehrach, H.; Reinhardt, R.; McCombie, W. R.; de la Bastide, M.; Dedhia, N.; Blocker, H.; Hornischer, K.; Nordsiek, G.; Agarwala, R.; Aravind, L.; Bailey, J. A.; Bateman, A.; Batzoglou, S.; Birney, E.; Bork, P.; Brown, D. G.; Burge, C. B.; Cerutti, L.; Chen, H. C.; Church, D.; Clamp, M.; Copley, R. R.; Doerks, T.; Eddy, S. R.; Eichler, E. E.; Furey, T. S.; Galagan, J.; Gilbert, J. G.; Harmon, C.; Hayashizaki, Y.; Haussler, D.; Hermjakob, H.; Hokamp, K.; Jang, W.; Johnson, L. S.; Jones, T. A.; Kasif, S.; Kaspryzk, A.; Kennedy, S.; Kent, W. J.; Kitts, P.; Koonin, E. V.; Korf, I.; Kulp, D.; Lancet, D.; Lowe, T. M.; McLysaght, A.; Mikkelsen, T.; Moran, J. V.; Mulder, N.; Pollara, V. J.; Ponting, C. P.; Schuler, G.; Schultz, J.; Slater, G.; Smit, A. F.; Stupka, E.; Szustakowski, J.; Thierry-Mieg, D.; Thierry-Mieg, J.; Wagner, L.; Wallis, J.; Wheeler, R.; Williams, A.; Wolf, Y. I.; Wolfe, K. H.; Yang, S. P.; Yeh, R. F.; Collins, F.; Guyer, M. S.; Peterson, J.; Felsenfeld, A.; Wetterstrand, K. A.; Patrinos, A.; Morgan, M. J.; de, J. P.; Catanese, J. J.; Osoegawa, K.; Shizuya, H.; Choi, S.; Chen, Y. J. Initial sequencing and analysis of the human genome. *Nature* **2001,** *409* (6822), 860-921.

82. Pandey, A.; Mann, M. Proteomics to study genes and genomes. *Nature* **2000,** *405* (6788), 837-846.

83. O'Farrell, P. Z.; Goodman, H. M.; O'Farrell, P. H. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* **1977,** *12* (4), 1133-1141.

84. Klose, J. Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik.* **1975,** *26* (3), 231-243.

85. Wahlander, A.; Arrigoni, G.; Snel, M.; Hellman, U.; James, P. Parallel post-source decay for increasing protein identification confidence levels from 2-D gels. *Proteomics* **2008,** *8* (9), 1771-1779.

86. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970,** *227* (5259), 680-685.

87. Shapiro, A. L.; Vinuela, E.; Maizel, J. V., Jr. Molecular weight estimation of polypeptide chains by electrophoresis in SDS-polyacrylamide gels. *Biochem Biophys. Res Commun.* **1967,** *28* (5), 815-820.

88. Bjellqvist, B.; Ek, K.; Righetti, P. G.; Gianazza, E.; Gorg, A.; Westermeier, R.; Postel, W. Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys. Methods* **1982,** *6* (4), 317-339.

89. Gorg, A.; Obermaier, C.; Boguth, G.; Harder, A.; Scheibe, B.; Wildgruber, R.; Weiss, W. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* **2000,** *21* (6), 1037-1053.

90. Gorg, A.; Weiss, W.; Dunn, M. J. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **2004,** *4* (12), 3665-3685.

91. Unlu, M.; Morgan, M. E.; Minden, J. S. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **1997,** *18* (11), 2071-2077.

92. Viswanathan, S.; Unlu, M.; Minden, J. S. Two-dimensional difference gel electrophoresis. *Nat Protoc.* **2006,** *1* (3), 1351-1358.

93. Pennington, S. R.; Wilkins, M. R.; Hochstrasser, D. F.; Dunn, M. J. Proteome analysis: from protein characterization to biological function. *Trends Cell Biol* **1997,** *7* (4), 168-173.

94. Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. Top Down versus Bottom Up Protein Characterization by Tandem High-Resolution Mass Spectrometry. *J. Am. Chem. Soc.* **1999,** *121* (4), 806-812.

95. Kelleher, N. L. Top-down proteomics. *Anal Chem* **2004,** *76* (11), 197A-203A.

96. Parks, B. A.; Jiang, L.; Thomas, P. M.; Wenger, C. D.; Roth, M. J.; Boyne, M. T.; Burke, P. V.; Kwast, K. E.; Kelleher, N. L. Top-down proteomics on a chromatographic time scale using linear ion trap fourier transform hybrid mass spectrometers. *Anal Chem* **2007,** *79* (21), 7984-7991.

97. Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. Top-down mass spectrometry of a 29-kDa protein for characterization of any posttranslational modification to within one residue. *Proc Natl Acad Sci U S A* **2002,** *99* (4), 1774-1779.

98. Han, J.; Borchers, C. H. Top-down analysis of recombinant histone H3 and its methylated analogs by ESI/FT-ICR mass spectrometry. *Proteomics* **2010,** *10* (20), 3621-3630.

99. Resemann, A.; Wunderlich, D.; Rothbauer, U.; Warscheid, B.; Leonhardt, H.; Fuchser, J.; Kuhlmann, K.; Suckau, D. Top-down de Novo protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption ionization-time-of-flight/time-of-flight mass spectrometry. *Anal Chem* **2010,** *82* (8), 3283-3292.

100. Aebersold, R.; Mann, M. Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.

101. Chait, B. T. Chemistry. Mass spectrometry: bottom-up or top-down? *Science* **2006,** *314* (5796), 65-66.

102. Swaney, D. L.; Wenger, C. D.; Coon, J. J. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *J Proteome Res* **2010,** *9* (3), 1323-1329.

103. Wolters, D. A.; Washburn, M. P.; Yates, J. R., III An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* **2001,** *73* (23), 5683-5690.

104. James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. Protein identification by mass profile fingerprinting. *Biochem Biophys. Res Commun.* **1993,** *195* (1), 58-64.

105. James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. Protein identification in DNA databases by peptide mass fingerprinting. *Protein Sci* **1994,** *3* (8), 1347-1350.

106. Levander, F.; Rognvaldsson, T.; Samuelsson, J.; James, P. Automated methods for improved protein identification by peptide mass fingerprinting. *Proteomics* **2004,** *4* (9), 2594-2601.

107. Samuelsson, J.; Dalevi, D.; Levander, F.; Rognvaldsson, T. Modular, scriptable and automated analysis tools for high-throughput peptide mass fingerprinting. *Bioinformatics* **2004,** *20* (18), 3628-3635.

108. Yates, J. R., III; Speicher, S.; Griffin, P. R.; Hunkapiller, T. Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* **1993,** *214* (2), 397-408.

109. Hunt, D. F.; Yates, J. R., III; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proc Natl Acad Sci U S A* **1986,** *83* (17), 6233-6237.

110. Yates, J. R., III Database searching using mass spectrometry data. *Electrophoresis* **1998,** *19* (6), 893-900.

111. Yates, J. R., III; Eng, J. K.; McCormack, A. L.; Schieltz, D. Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database. *Anal Chem* **1995,** *67* (8), 1426-1436.

112. Sadygov, R. G.; Cociorva, D.; Yates, J. R., III Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book. *Nat Methods* **2004,** *1* (3), 195-202.

113. Craig, R.; Beavis, R. C. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun. Mass Spectrom* **2003,** *17* (20), 2310-2316.

114. Geer, L. Y.; Markey, S. P.; Kowalak, J. A.; Wagner, L.; Xu, M.; Maynard, D. M.; Yang, X.; Shi, W.; Bryant, S. H. Open mass spectrometry search algorithm. *J Proteome Res* **2004,** *3* (5), 958-964.

115. Perkins, D. N.; Pappin, D. J.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999,** *20* (18), 3551-3567.

116. McCormack, A. L.; Somogyi, A.; Dongre, A. R.; Wysocki, V. H. Fragmentation of protonated peptides: surface-induced dissociation in conjunction with a quantum mechanical approach. *Anal Chem* **1993,** *65* (20), 2859-2872.

117. Wysocki, V. H.; Tsaprailis, G.; Smith, L. L.; Breci, L. A. Mobile and localized protons: a framework for understanding peptide dissociation. *J Mass Spectrom* **2000,** *35* (12), 1399-1406.

118. Roepstorff, P.; Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* **1984,** *11* (11), 601.

119. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal Chem* **1987,** *59* (21), 2621-2625.

120. Bean, M. F.; Carr, S. A.; Thorne, G. C.; Reilly, M. H.; Gaskell, S. J. Tandem mass spectrometry of peptides using hybrid and four-sector instruments: a comparative study. *Anal Chem* **1991,** *63* (14), 1473-1481.

121. Olsen, J. V.; Ong, S. E.; Mann, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* **2004,** *3* (6), 608-614.

122. Paizs, B.; Suhai, S. Fragmentation pathways of protonated peptides. *Mass Spectrom Rev* **2005,** *24* (4), 508-548.

123. Harrison, A. G.; Yalcin, T. Proton mobility in protonated amino acids and peptides. *International Journal of Mass Spectrometry and Ion Processes* **1997,** *165*, 339-347.

124. Wattenberg, A.; Organ, A. J.; Schneider, K.; Tyldesley, R.; Bordoli, R.; Bateman, R. H. Sequence dependent fragmentation of peptides generated by MALDI quadrupole time-of-flight (MALDI Q-TOF) mass spectrometry and its implications for protein identification. *J Am Soc Mass Spectrom* **2002,** *13* (7), 772-783.

125. Khatun, J.; Ramkissoon, K.; Giddings, M. C. Fragmentation characteristics of collision-induced dissociation in MALDI TOF/TOF mass spectrometry. *Anal Chem* **2007,** *79* (8), 3032-3040.

126. Beavis, R. C.; Chait, B. T. Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid Commun. Mass Spectrom* **1989,** *3* (12), 432-435.

127. Strupat, K.; Karas, M.; illenkamp, F. 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption—ionization mass spectrometry. *International Journal of Mass Spectrometry and Ion Processes* **1991,** *111*, 89-102.

128. Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal Chem* **1991,** *63* (24), 1193A-1203A.

129. Beavis, R. C.; Chait, B. T. Matrix-assisted laser-desorption mass spectrometry using 355 nm radiation. *Rapid Commun. Mass Spectrom* **1989,** *3* (12), 436-439.

130. Cohen, S. L.; Chait, B. T. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal Chem* **1996,** *68* (1), 31-37.

131. Juhasz, P.; Costello, C. E.; Biemann, K. Matrix-Assisted Laser Desorption Ionization Mass Spectrometry with 2-(4-Hydroxyphenylazo)benzoic Acid Matrix. *J Am Soc Mass Spectrom* **1993,** *4* (5), 399-409.

132. Wu, K. J.; Shaler, T. A.; Becker, C. H. Time-of-flight mass spectrometry of underivatized single-stranded DNA oligomers by matrix-assisted laser desorption. *Anal Chem* **1994,** *66* (10), 1637-1645.

133. Wu, K. J.; Steding, A.; Becker, C. H. Matrix-assisted laser desorption time-of-flight mass spectrometry of oligonucleotides using 3-hydroxypicolinic acid as an ultraviolet-sensitive matrix. *Rapid Commun. Mass Spectrom* **1993,** *7* (2), 142-146.

134. Berkenkamp, S.; Kirpekar, F.; Hillenkamp, F. Infrared MALDI mass spectrometry of large nucleic acids. *Science* **1998,** *281* (5374), 260-262.

135. Nordhoff, E.; Cramer, R.; Karas, M.; Hillenkamp, F.; Kirpekar, F.; Kristiansen, K.; Roepstorff, P. Ion stability of nucleic acids in infrared matrix-assisted laser desorption/ionization mass spectrometry. *Nucleic Acids Res* **1993,** *21* (15), 3347-3357.

136. Nordhoff, E.; Ingendoh, A.; Cramer, R.; Overberg, A.; Stahl, B.; Karas, M.; Hillenkamp, F.; Crain, P. F. Matrix-assisted laser desorption/ionization mass spectrometry of nucleic acids with wavelengths in the ultraviolet and infrared. *Rapid Commun. Mass Spectrom* **1992,** *6* (12), 771-776.

137. Pieles, U.; Zurcher, W.; Schar, M.; Moser, H. E. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for the mass and sequence analysis of natural and modified oligonucleotides. *Nucleic Acids Res* **1993,** *21* (14), 3191-3196.

138. Metzger, J. O.; Tuszynski, W.; Woisch, R. New type of matrix for matrix_assisted laser desorption mass spectrometry of polysaccharides and proteins. *Fresenius J Anal Chem* **1994,** *349* (6), 473-474.

139. Dai, Y.; Whittal, R. M.; Bridges, C. A.; Isogai, Y.; Hindsgaul, O.; Li, L. Matrix-assisted laser desorption ionization mass spectrometry for the analysis of monosulfated oligosaccharides. *Carbohydr. Res* **1997,** *304* (1), 1-9.

140. Fitzgerald, M. C.; Parr, G. R.; Smith, L. M. Basic matrices for the matrix-assisted laser desorption/ionization mass spectrometry of proteins and oligonucleotides. *Anal Chem* **1993,** *65* (22), 3204-3211.

References

141.   Sunner, J.; Dratz, E.; Chen, Y. C. Graphite surface-assisted laser desorption/ionization time-of-flight mass spectrometry of peptides and proteins from liquid solutions. *Anal Chem* **1995,** *67* (23), 4335-4342.

142.   Kraft, P.; Alimpiev, S.; Dratz, E.; Sunner, J. Infrared, surface-assisted laser desorption ionization mass spectrometry on frozen aqueous solutions of proteins and peptides using suspensions of organic solids. *J Am Soc Mass Spectrom* **1998,** *9* (9), 912-924.

143.   Dale, M. J.; Knochenmuss, R.; Zenobi, R. Two-phase Matrix-assisted Laser Desorption/Ionization: Matrix Selection and Sample Pretreatment for Complex Anionic Analytes. *Rapid Commun. Mass Spectrom.* **1997,** *11* (1), 136-142.

144.   Cornett, D. S.; Duncan, M. A.; Amster, I. J. Matrix-assisted laser desorption at visible wavelengths using a 2-component matrix. *Organic Mass Spectrometry* **1992,** *27* (7), 831-832.

145.   Cornett, D. S.; Duncan, M. A.; Amster, I. J. Liquid mixtures for matrix-assisted laser desorption. *Anal Chem* **1993,** *65* (19), 2608-2613.

146.   Zöllner, P.; Schmid, E. R.; Allmaier, G. K4(Fe(CN)6/glycerol-A new liquid matrix system for matrix-assisted laser desorption/ionization mass spectrometry of hydrophobic compounds. *Rapid Commun. Mass Spectrom* **1996,** *10* (10), 1278-1282.

147.   Sze, E. T.; Chan, T. W.; Wang, G. Formulation of matrix solutions for use in matrix-assisted laser desorption/ionization of biomolecules. *J Am Soc Mass Spectrom* **1998,** *9* (2), 166-174.

148.   Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *International Journal of Mass Spectrometry* **1987,** *78* (24), 53-68.

149.   Zhao, S.; Somayajula, K. V.; Sharkey, A. G.; Hercules, D. M.; Hillenkamp, F.; Karas, M.; Ingendoh, A. Novel method for matrix-assisted laser mass spectrometry of proteins. *Anal Chem* **1991,** *63* (5), 450-453.

150.   Chan, T. W. D.; Thomas, I.; Colburn, A. W.; Derrick, P. J. Initial velocities of positive and negative protein molecule-ions produced in matrix-assisted ultraviolet laser desorption using a liquid matrix. *Chemical Physics Letters* **1994,** *222* (6), 579-585.

151.   Williams, J. B.; Gusev, A. I.; Hercules, D. M. Use of Liquid Matrices for Matrix-Assisted Laser Desorption Ionization of Polyglycols and Poly(dimethylsiloxanes). *Macromolecules* **1996,** *29* (25), 8144-8150.

152.   Berkenkamp, S.; Karas, M.; Hillenkamp, F. Ice as a matrix for IR-matrix-assisted laser desorption/ionization: mass spectra from a protein single crystal. *Proc Natl Acad Sci U S A* **1996,** *93* (14), 7003-7007.

153.   Caldwell, K. L.; Murray, K. K. Mid-infrared matrix assisted laser desorption ionization with a water/glycerol matrix. *Applied Surface Science* **1998,** *127-129*, 242-247.

154.   Deery, M. J.; Jennings, K. R.; Jasieczek, C. B.; Haddleton, D. M.; Jackson, A. T.; Yates, H. T.; Scrivens J.H. A Study of Cation Attachment to Polystyrene by Means of Matrix-assisted Laser Desorption/Ionization and Electrospray Ionization-Mass Spectrometry. *Rapid Commun. Mass Spectrom* **1997,** *11* (1), 57-62.

155.   Mowat, I. A.; Donovan, R. J.; Maier, R. R. J. Enhanced Cationization of Polymers Using Delayed Ion Extraction with Matrix-assisted Laser Desorption/Ionization. *Rapid Commun. Mass Spectrom* **1997,** *11* (1), 89-98.

156.   Gross, J.; Leisner, A.; Hillenkamp, F.; Hahner, S.; Karas, M.; Schafer, J.; Lutzenkirchen, F.; Nordhoff, E. Investigations of the metastable decay of DNA under ultraviolet matrix-assisted laser desorption/ionization conditions with post-source-decay analysis and hydrogen/deuterium exchange. *J Am Soc Mass Spectrom* **1998,** *9* (9), 866-878.

157.   Karas, M.; Bahr, U.; Strupat, K.; Hillenkamp, F.; Tsarbopoulos, A.; Pramanik, B. N. Matrix dependence of metastable fragmentation of glycoproteins in MALDI TOF mass_spectrometry. *Anal. Chem.* **1995,** *67* (3), 675-679.

158.   Zhu, L.; Parr, G. R.; Fitzgerald, M. C.; Nelson, C. M.; Smith, L. M. Oligodoxynucleotide fragmentation in MALDI/TOFmass spectrometry using 355-nm radiation. *J. Am. Chem. Soc.* **1995,** *117* (22), 6048-6056.

159.   Karas, M.; Bahr, U.; Stah-Zeng, J. R. Steps toward a more refined picture of the matrix function in UV MALDI. In *Large Ions: Their Vaporization, Detection and Structural Analysis*, 1st ed.; Baer, T., Ng, C.-Y., Powis, I., Eds.; John Wiley & Sons: London, 1996; p 27.

160.   Kussmann, M.; Roepstorff, P. Sample preparation techniques for peptides and proteins analyzed by MALDI-MS. *Methods Mol Biol* **2000,** *146*, 405-424.

161. Kjellstrom, S.; Jensen, O. N. Phosphoric acid as a matrix additive for MALDI MS analysis of phosphopeptides and phosphoproteins. *Anal Chem* **2004,** *76* (17), 5109-5117.

162. Garden, R. W.; Sweedler, J. V. Heterogeneity within MALDI samples as revealed by mass spectrometric imaging. *Anal Chem* **2000,** *72* (1), 30-36.

163. Vorm, O.; Roepstorff, P.; Mann, M. Improved Resolution and Very High Sensitivity in MALDI TOF of Matrix Surfaces Made by Fast Evaporation. *Anal. Chem.* **1994,** *66* (19), 3281-3287.

164. Li, L.; Golding, R. E.; Whittal, R. M. Analysis of Single Mammalian Cell Lysates by Mass Spectrometry. *J. Am. Chem. Soc.* **1996,** *118*, 11662-11663.

165. Dai, Y.; Whittal, R. M.; Li, L. Two-layer sample preparation: a method for MALDI-MS analysis of complex peptide and protein mixtures. *Anal Chem* **1999,** *71* (5), 1087-1091.

166. Zhang, N.; Doucette, A.; Li, L. Two-layer sample preparation method for MALDI mass spectrometric analysis of protein and peptide samples containing sodium dodecyl sulfate. *Anal Chem* **2001,** *73* (13), 2968-2975.

167. Miliotis, T.; Marko-Varga, G.; Nilsson, J.; Laurell, T. Development of silicon microstructures and thin-film MALDI target plates for automated proteomics sample identifications. *J Neurosci Methods* **2001,** *109* (1), 41-46.

168. Schuerenberg, M.; Luebbert, C.; Eickhoff, H.; Kalkum, M.; Lehrach, H.; Nordhoff, E. Prestructured MALDI-MS sample supports. *Anal Chem* **2000,** *72* (15), 3436-3442.

169. Jahn, O.; Hesse, D.; Reinelt, M.; Kratzin, H. D. Technical innovations for the automated identification of gel-separated proteins by MALDI-TOF mass spectrometry. *Anal Bioanal. Chem* **2006,** *386* (1), 92-103.

170. Kampmeier, J.; Dreisewerd, K.; Schürenberg, M.; Strupat, K. Investigations of 2,5-DHB and succinic acid as matrices for IR and UV MALDI. Part: I UV and IR laser ablation in the MALDI process. *International Journal of Mass Spectrometry and Ion Processes* **1997,** *169-170*, 31-41.

171. Cook, J. W.; Hewett, C. L.; Hieger, I. The isolation of a cancer-producing hydrocarbon from coal tar. Parts I, II, and III. *J. Chem. Soc.* **1933,** 395-405.

172. Phillips, D. H. Polycyclic aromatic hydrocarbons in the diet. *Mutat Res* **1999,** *443* (1-2), 139-147.

173. Schoket, B. DNA damage in humans exposed to environmental and dietary polycyclic aromatic hydrocarbons. *Mutat Res* **1999,** *424* (1-2), 143-153.

174. Bostrom, C. E.; Gerde, P.; Hanberg, A.; Jernstrom, B.; Johansson, C.; Kyrklund, T.; Rannug, A.; Tornqvist, M.; Victorin, K.; Westerholm, R. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ Health Perspect.* **2002,** *110 Suppl 3*, 451-488.

175. Santella, R. M.; Grinberg-Funes, R. A.; Young, T. L.; Dickey, C.; Singh, V. N.; Wang, L. W.; Perera, F. P. Cigarette smoking related polycyclic aromatic hydrocarbon-DNA adducts in peripheral mononuclear cells. *Carcinogenesis* **1992,** *13* (11), 2041-2045.

176. Blaha, L.; Kapplova, P.; Vondracek, J.; Upham, B.; Machala, M. Inhibition of gap-junctional intercellular communication by environmentally occurring polycyclic aromatic hydrocarbons. *Toxicol Sci* **2002,** *65* (1), 43-51.

177. Deutsch-Wenzel, R. P.; Brune, H.; Grimmer, G.; Dettbarn, G.; Misfeld, J. Experimental studies in rat lungs on the carcinogenicity and dose-response relationships of eight frequently occurring environmental polycyclic aromatic hydrocarbons. *J Natl Cancer Inst* **1983,** *71* (3), 539-544.

178. Perera, F. P.; Hemminki, K.; Gryzbowska, E.; Motykiewicz, G.; Michalska, J.; Santella, R. M.; Young, T. L.; Dickey, C.; Brandt-Rauf, P.; De, V., I; . Molecular and genetic damage in humans from environmental pollution in Poland. *Nature* **1992,** *360* (6401), 256-258.

179. Motykiewicz, G.; Michalska, J.; Pendzich, J.; Perera, F. P.; Chorazy, M. A cytogenetic study of men environmentally and occupationally exposed to airborne pollutants. *Mutat Res* **1992,** *280* (4), 253-259.

180. Rogan, E. G.; Mailander, P.; Cavalieri, E. Metabolic activation of aromatic hydrocarbons in purified rat liver nuclei: induction of enzyme activities and binding to DNA with and without monooxygenase-catalyzed formation of active oxygen. *Proc Natl Acad Sci U S A* **1976,** *73* (2), 457-461.

181. Cavalieri, E.; Roth, R. Reaction of methylbenzanthracenes and pyridine by one-electron oxidation. A model for metabolic activation and binding of carcinogenic aromatic hydrocarbons. *J Org Chem* **1976,** *41* (16), 2679-2684.

182. Cavalieri, E.; Rogan, E. Role of radical cations in aromatic hydrocarbon carcinogenesis. *Environ Health Perspect.* **1985,** *64*, 69-84.

183. Cavalieri, E. L.; Rogan, E. G. Central role of radical cations in metabolic activation of polycyclic aromatic hydrocarbons. *Xenobiotica* **1995,** *25* (7), 677-688.

184. Li, K. M.; Todorovic, R.; Rogan, E. G.; Cavalieri, E. L.; Ariese, F.; Suh, M.; Jankowiak, R.; Small, G. J. Identification and quantitation of dibenzo[a,l]pyrene--DNA adducts formed by rat liver microsomes in vitro: preponderance of depurinating adducts. *Biochemistry* **1995,** *34* (25), 8043-8049.

185. Penning, T. M.; Ohnishi, S. T.; Ohnishi, T.; Harvey, R. G. Generation of reactive oxygen species during the enzymatic oxidation of polycyclic aromatic hydrocarbon trans-dihydrodiols catalyzed by dihydrodiol dehydrogenase. *Chem Res Toxicol* **1996,** *9* (1), 84-92.

186. Penning, T. M.; Burczynski, M. E.; Hung, C. F.; McCoull, K. D.; Palackal, N. T.; Tsuruda, L. S. Dihydrodiol dehydrogenases and polycyclic aromatic hydrocarbon activation: generation of reactive and redox active o-quinones. *Chem Res Toxicol* **1999,** *12* (1), 1-18.

187. McCoull, K. D.; Rindgen, D.; Blair, I. A.; Penning, T. M. Synthesis and characterization of polycyclic aromatic hydrocarbon o-quinone depurinating N7-guanine adducts. *Chem Res Toxicol* **1999,** *12* (3), 237-246.

188. Harvey, R. G. Polyciclic Aromatic Hydrocarbons. In *Chemistry and Carcinogenicity*, Cambridge Univ.Press, Ed.; Cambridge, 1991; pp 11-87.

189. Gelboin, H. V. Benzo[alpha]pyrene metabolism, activation and carcinogenesis: role and regulation of mixed-function oxidases and related enzymes. *Physiol Rev* **1980,** *60* (4), 1107-1166.

190. Baird, W. M.; Hooven, L. A.; Mahadevan, B. Carcinogenic polycyclic aromatic hydrocarbon-DNA adducts and mechanism of action. *Environ Mol Mutagen.* **2005,** *45* (2-3), 106-114.

191. Singh, R.; Gaskell, M.; Le Pla, R. C.; Kaur, B.; Azim-Araghi, A.; Roach, J.; Koukouves, G.; Souliotis, V. L.; Kyrtopoulos, S. A.; Farmer, P. B. Detection and quantitation of benzo[a]pyrene-derived DNA adducts in mouse liver by liquid chromatography-tandem mass spectrometry: comparison with 32P-postlabeling. *Chem Res Toxicol* **2006,** *19* (6), 868-878.

192. Xue, W.; Warshawsky, D. Metabolic activation of polycyclic and heterocyclic aromatic hydrocarbons and DNA damage: a review. *Toxicol Appl Pharmacol* **2005,** *206* (1), 73-93.

193. Whitlock, J. P., Jr. Induction of cytochrome P4501A1. *Annu Rev Pharmacol Toxicol* **1999,** *39*, 103-125.

194. Wu, M.; Yan, S.; Patel, D. J.; Geacintov, N. E.; Broyde, S. Relating repair susceptibility of carcinogen-damaged DNA with structural distortion and thermodynamic stability. *Nucleic Acids Res* **2002,** *30* (15), 3422-3432.

195. Braithwaite, E.; Wu, X.; Wang, Z. Repair of DNA lesions induced by polycyclic aromatic hydrocarbons in human cell-free extracts: involvement of two excision repair mechanisms in vitro. *Carcinogenesis* **1998,** *19* (7), 1239-1246.

196. Friedberg, E. C. How nucleotide excision repair protects against cancer. *Nat Rev Cancer* **2001,** *1* (1), 22-33.

197. Bordin, F.; Musajo, L.; Bevilacqua, R. Fluorescent and non fluorescent C4-cycloadducts in the photoreaction at 365 nm between psoralen-3H and DNA. *Z. Naturforsch. B* **1969,** *24* (6), 691-693.

198. Denissenko, M. F.; Pao, A.; Pfeifer, G. P.; Tang, M. Slow repair of bulky DNA adducts along the nontranscribed strand of the human p53 gene may explain the strand bias of transversion mutations in cancers. *Oncogene* **1998,** *16* (10), 1241-1247.

199. Ross, J. A.; Nesnow, S. Polycyclic aromatic hydrocarbons: correlations between DNA adducts and ras oncogene mutations. *Mutat Res* **1999,** *424* (1-2), 155-166.

200. Denissenko, M. F.; Pao, A.; Tang, M.; Pfeifer, G. P. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in P53. *Science* **1996,** *274* (5286), 430-432.

201. Boysen, G.; Hecht, S. S. Analysis of DNA and protein adducts of benzo[a]pyrene in human tissues using structure-specific methods. *Mutat Res* **2003,** *543* (1), 17-30.

202. Lin, J. M.; Desai, D.; Chung, L.; Hecht, S. S.; Amin, S. Synthesis of anti-7,8-dihydroxy-9,10-epoxy-7,8,9, 10-tetrahydro-11-methylbenzo[a]pyrene and its reaction with DNA. *Chem Res Toxicol* **1999,** *12* (4), 341-346.

203. Divi, R. L.; Beland, F. A.; Fu, P. P.; Von Tungeln, L. S.; Schoket, B.; Camara, J. E.; Ghei, M.; Rothman, N.; Sinha, R.; Poirier, M. C. Highly sensitive chemiluminescence immunoassay for benzo[a]pyrene-DNA adducts: validation by comparison with other methods, and use in human biomonitoring. *Carcinogenesis* **2002,** *23* (12), 2043-2049.

204. Cadet, J.; Weinfeld, M. Detecting DNA damage. *Anal Chem* **1993,** *65* (15), 675A-682A.

205. Straub, K. M.; Meehan, T.; Burlingame, A. L.; Calvin, M. Identification of the major adducts formed by reaction of benzo(a)pyrene diol epoxide with DNA in vitro. *Proc Natl Acad Sci U S A* **1977,** *74* (12), 5285-5289.

206. Pavanello, S.; Favretto, D.; Brugnone, F.; Mastrangelo, G.; Dal Pra, G.; Clonfero, E. HPLC/fluorescence determination of anti-BPDE-DNA adducts in mononuclear white blood cells from PAH-exposed humans. *Carcinogenesis* **1999,** *20* (3), 431-435.

207. Amin, S.; Misra, B.; Desai, D.; Huie, K.; Hecht, S. S. Chromatographic conditions for separation of 32P-labeled phosphates of major polynuclear aromatic hydrocarbon--deoxyribonucleoside adducts. *Carcinogenesis* **1989,** *10* (10), 1971-1974.

208. van Gijssel, H. E.; Divi, R. L.; Olivero, O. A.; Roth, M. J.; Wang, G. Q.; Dawsey, S. M.; Albert, P. S.; Qiao, Y. L.; Taylor, P. R.; Dong, Z. W.; Schrager, J. A.; Kleiner, D. E.; Poirier, M. C. Semiquantitation of polycyclic aromatic hydrocarbon-DNA adducts in human esophagus by immunohistochemistry and the automated cellular imaging system. *Cancer Epidemiol Biomarkers Prev* **2002,** *11* (12), 1622-1629.

209. Halter, R.; Hansen, T.; Seidel, A.; Zieman, C.; Borlak, J. Importance of DNA-Adduct Formation and Gene Expression Profiling of Disease Candidate Genes in Rats Exposed to Bitumen Fumes. *Journal of Occupational and Environmental Hygiene* **2007,** *4* (1 Suppl. 1), 44-64.

210. Reddy, M. V.; Gupta, R. C.; Randerath, E.; Randerath, K. 32P-postlabeling test for covalent DNA binding of chemicals in vivo: application to a variety of aromatic carcinogens and methylating agents. *Carcinogenesis* **1984,** *5* (2), 231-243.

211. Gupta, R. C. Enhanced sensitivity of 32P-postlabeling analysis of aromatic carcinogen:DNA adducts. *Cancer Res* **1985,** *45* (11 Pt 2), 5656-5662.

212. Gupta, R. C.; Reddy, M. V.; Randerath, K. 32P-postlabeling analysis of non-radioactive aromatic carcinogen--DNA adducts. *Carcinogenesis* **1982,** *3* (9), 1081-1092.

213. Randerath, K.; Randerath, E. 32P-postlabeling methods for DNA adduct detection: overview and critical evaluation. *Drug Metab Rev* **1994,** *26* (1-2), 67-85.

214. Terashima, I.; Suzuki, N.; Shibutani, S. 32P-Postlabeling/polyacrylamide gel electrophoresis analysis: application to the detection of DNA adducts. *Chem Res Toxicol* **2002,** *15* (3), 305-311.

215. Reichert, W. L.; Stein, J. E.; French, B.; Goodwin, P.; Varanasi, U. Storage phosphor imaging technique for detection and quantitation of DNA adducts measured by the 32P-postlabeling assay. *Carcinogenesis* **1992,** *13* (8), 1475-1479.

216. Vulimiri, S. V.; Smith, C. V.; Randerath, E.; Randerath, K. 32P-postlabeling of bile components: bulky adduct-like behavior in polyethyleneimine-cellulose thin layer chromatography. *Carcinogenesis* **1994,** *15* (9), 2061-2064.

217. Arif, J. M.; Dresler, C.; Clapper, M. L.; Gairola, C. G.; Srinivasan, C.; Lubet, R. A.; Gupta, R. C. Lung DNA adducts detected in human smokers are unrelated to typical polyaromatic carcinogens. *Chem Res Toxicol* **2006,** *19* (2), 295-299.

218. Goodenough, A. K.; Schut, H. A.; Turesky, R. J. Novel LC-ESI/MS/MS(n) method for the characterization and quantification of 2'-deoxyguanosine adducts of the dietary carcinogen 2-amino-1-methyl-6-phenylimidazo[4,5-b]pyridine by 2-D linear quadrupole ion trap mass spectrometry. *Chem Res Toxicol* **2007,** *20* (2), 263-276.

219. Wang, J. J.; Marshall, W. D.; Frazer, D. G.; Law, B.; Lewis, D. M. Characterization of DNA adducts from lung tissue of asphalt fume-exposed mice by nanoflow liquid chromatography quadrupole time-of-flight mass spectrometry. *Anal Biochem* **2003,** *322* (1), 79-88.

220. Wang, J. J.; Frazer, D. G.; Stone, S.; Goldsmith, T.; Law, B.; Moseley, A.; Simpson, J.; Afshari, A.; Lewis, D. M. Urinary benzo[a]pyrene and its metabolites as molecular biomarkers of asphalt fume exposure characterized by microflow LC coupled to hybrid quadrupole time-of-flight mass spectrometry. *Anal Chem* **2003,** *75* (21), 5953-5960.

221. Wang, J. J.; Frazer, D. G.; Law, B.; Lewis, D. M. Identification and quantification of urinary benzo[a]pyrene and its metabolites from asphalt fume exposed mice by microflow LC coupled to hybrid quadrupole time-of-flight mass spectrometry. *Analyst* **2003,** *128* (7), 864-870.

222. Gaskell, M.; Kaur, B.; Farmer, P. B.; Singh, R. Detection of phosphodiester adducts formed by the reaction of benzo[a]pyrene diol epoxide with 2'-deoxynucleotides using collision-induced dissociation electrospray ionization tandem mass spectrometry. *Nucleic Acids Res* **2007,** *35* (15), 5014-5027.

223. Barry, J. P.; Norwood, C.; Vouros, P. Detection and identification of benzo[a]pyrene diol epoxide adducts to DNA utilizing capillary electrophoresis-electrospray mass spectrometry. *Anal Chem* **1996,** *68* (8), 1432-1438.

224. Willems, A. V.; Deforce, D. L.; Van den Eeckhout, E. G.; Lambert, W. E.; Van Peteghem, C. H.; De Leenheer, A. P.; Van Bocxlaer, J. F. Analysis of benzo[a]pyrene diol epoxide-DNA adducts by capillary zone electrophoresis- electrospray ionization-mass spectrometry in conjunction with sample stacking. *Electrophoresis* **2002,** *23* (24), 4092-4103.

225. Feng, F.; Wang, H. Simultaneous analysis of four stereoisomers of anti-benzo[a]pyrene diol epoxide-deoxyguanosine adducts in short oligodeoxynucleotides using reversed-phase high-performance liquid chromatography. *J Chromatogr. A* **2007,** *1162* (2), 141-148.

226. Feng, F.; Yin, J.; Song, M.; Wang, H. Preparation, identification and analysis of stereoisomeric anti-benzo[a]pyrene diol epoxide-deoxyguanosine adducts using phenyl liquid chromatography with diode array, fluorescence and tandem mass spectrometry detection. *J Chromatogr. A* **2008,** *1183* (1-2), 119-128.

227. Lan, Z. H.; Wang, P.; Giese, R. W. Matrix-assisted laser desorption/ionization mass spectrometry of deoxynucleotides labeled with an IMI dye. *Rapid Commun. Mass Spectrom* **1999,** *13* (14), 1454-1457.

228. Stemmler, E. A.; Buchanan, M. V.; Hurst, G. B.; Hettich, R. L. Structural characterization of polycyclic aromatic hydrocarbon dihydrodiol epoxide DNA adducts using matrix-assisted laser desorption/ionization Fourier transform mass spectrometry. *Anal Chem* **1994,** *66* (8), 1274-1285.

229. Chiarelli, M. P.; Wu, H. P.; Antunes, A. M.; Branco, P. S. Product ion studies of some novel arylamine adducts of deoxyguanosine by matrix-assisted laser desorption/ionization and post-source decay. *Rapid Commun. Mass Spectrom* **1999,** *13* (20), 2004-2010.

230. Chiarelli, M. P.; Chang, H. F.; Olsen, K. W.; Barbacci, D.; Huffer, D. M.; Cho, B. P. Structural differentiation of diastereomeric benzo[ghi]fluoranthene adducts of deoxyadenosine by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and postsource decay. *Chem Res Toxicol* **2003,** *16* (10), 1236-1241.

231. Brown, K.; Harvey, C. A.; Turteltaub, K. W.; Shields, S. J. Structural characterization of carcinogen-modified oligodeoxynucleotide adducts using matrix-assisted laser desorption/ionization mass spectrometry. *J Mass Spectrom* **2003,** *38* (1), 68-79.

232. Stemmler, E. A.; Hettich, R. L.; Hurst, G. B.; Buchanan, M. V. Matrix-assisted laser desorption/ionization Fourier-transform mass spectrometry of oligodeoxyribonucleotides. *Rapid Commun. Mass Spectrom* **1993,** *7* (9), 828-836.

233. Tretyakova, N.; Matter, B.; Ogdie, A.; Wishnok, J. S.; Tannenbaum, S. R. Locating nucleobase lesions within DNA sequences by MALDI-TOF mass spectral analysis of exonuclease ladders. *Chem Res Toxicol* **2001,** *14* (8), 1058-1070.

234. Pieles, U.; Zurcher, W.; Schar, M.; Moser, H. E. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry: a powerful tool for the mass and sequence analysis of natural and modified oligonucleotides. *Nucleic Acids Res* **1993,** *21* (14), 3191-3196.

235. Barnes, C. A.; Norman H.L.Chiu, N. H. L. Accurate characterization of carcinogenic DNA adducts using MALDI tandem time-of-flight mass spectrometry. *International Journal of Mass Spectrometry* **2009,** *279* (2-3), 170-175.

236. Qiao, Z.; Udeochu, U.; Jimerson, T.; Fletcher, M.; Bakare, O.; Hosten, C. M. Detection of benzopyrene-deoxyguanosine adducts by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom* **2006,** *20* (3), 487-492.

237. Fuchs, B.; Nimptsch, A.; Suss, R.; Schiller, J. Analysis of brain lipids by directly coupled matrix-assisted laser desorption ionization time-of-flight mass spectrometry and high-performance thin-layer chromatography. *J AOAC Int* **2008,** *91* (5), 1227-1236.

238. Fuchs, B.; Schiller, J.; Suss, R.; Zscharnack, M.; Bader, A.; Muller, P.; Schurenberg, M.; Becker, M.; Suckau, D. Analysis of stem cell lipids by offline HPTLC-MALDI-TOF MS. *Anal Bioanal. Chem* **2008,** *392* (5), 849-860.

239. Jensen, O. N. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. *Curr. Opin. Chem Biol* **2004,** *8* (1), 33-41.

240. FISCHER, E. H.; KREBS, E. G. Conversion of phosphorylase b to phosphorylase a in muscle extracts. *J Biol Chem* **1955,** *216* (1), 121-132.

241. KREBS, E. G.; FISCHER, E. H. Phosphorylase activity of skeletal muscle extracts. *J Biol Chem* **1955,** *216* (1), 113-120.

242. Robinson, D. R.; Wu, Y. M.; Lin, S. F. The protein tyrosine kinase family of the human genome. *Oncogene* **2000,** *19* (49), 5548-5557.

243. Smith, C. I.; Islam, T. C.; Mattsson, P. T.; Mohamed, A. J.; Nore, B. F.; Vihinen, M. The Tec family of cytoplasmic tyrosine kinases: mammalian Btk, Bmx, Itk, Tec, Txk and homologs in other species. *Bioessays* **2001,** *23* (5), 436-446.

244. Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di, F., V; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C.; Yao, A.; Ye, J.; Zhan, M.; Zhang, W.; Zhang, H.; Zhao, Q.; Zheng, L.; Zhong, F.; Zhong, W.; Zhu, S.; Zhao, S.; Gilbert, D.; Baumhueter, S.; Spier, G.; Carter, C.; Cravchik, A.; Woodage, T.; Ali, F.; An, H.; Awe, A.; Baldwin, D.; Baden, H.; Barnstead, M.; Barrow, I.; Beeson, K.; Busam, D.; Carver, A.; Center, A.; Cheng, M. L.; Curry, L.; Danaher, S.; Davenport, L.; Desilets, R.; Dietz, S.; Dodson, K.; Doup, L.; Ferriera, S.; Garg, N.; Gluecksmann, A.; Hart, B.; Haynes, J.; Haynes, C.; Heiner, C.; Hladun, S.; Hostin, D.; Houck, J.; Howland, T.; Ibegwam, C.; Johnson, J.; Kalush, F.; Kline, L.; Koduru, S.; Love, A.; Mann, F.; May, D.; McCawley, S.; McIntosh, T.; McMullen, I.; Moy, M.; Moy, L.; Murphy, B.; Nelson, K.; Pfannkoch, C.; Pratts, E.; Puri, V.; Qureshi, H.; Reardon, M.; Rodriguez, R.; Rogers, Y. H.; Romblad, D.; Ruhfel, B.; Scott, R.; Sitter, C.; Smallwood, M.; Stewart, E.; Strong, R.; Suh, E.; Thomas, R.; Tint, N. N.; Tse, S.; Vech, C.; Wang, G.; Wetter, J.; Williams, S.; Williams, M.; Windsor, S.; Winn-Deen, E.; Wolfe, K.; Zaveri, J.; Zaveri, K.; Abril, J. F.; Guigo, R.; Campbell, M. J.; Sjolander, K. V.; Karlak, B.; Kejariwal, A.; Mi, H.; Lazareva, B.; Hatton, T.; Narechania, A.; Diemer, K.; Muruganujan, A.; Guo, N.; Sato, S.; Bafna, V.; Istrail, S.; Lippert, R.; Schwartz, R.; Walenz, B.; Yooseph, S.; Allen, D.; Basu, A.; Baxendale, J.; Blick, L.; Caminha, M.; Carnes-Stine, J.; Caulk, P.; Chiang, Y. H.; Coyne, M.; Dahlke, C.; Mays, A.; Dombroski, M.; Donnelly, M.; Ely, D.; Esparham, S.; Fosler, C.; Gire, H.; Glanowski, S.; Glasser, K.; Glodek, A.; Gorokhov, M.; Graham, K.; Gropman, B.; Harris, M.; Heil, J.; Henderson, S.; Hoover, J.; Jennings, D.; Jordan, C.; Jordan, J.; Kasha, J.; Kagan, L.; Kraft, C.; Levitsky, A.; Lewis, M.; Liu, X.; Lopez, J.; Ma, D.; Majoros, W.; McDaniel, J.; Murphy, S.; Newman, M.; Nguyen, T.; Nguyen, N.; Nodell, M. The sequence of the human genome. *Science* **2001,** *291* (5507), 1304-1351.

245. Olsen, J. V.; Blagoev, B.; Gnad, F.; Macek, B.; Kumar, C.; Mortensen, P.; Mann, M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* **2006,** *127* (3), 635-648.

246. Hunter, T. Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. *Cell* **1995,** *80* (2), 225-236.

247. Lim, Y. P. Mining the tumor phosphoproteome for cancer markers. *Clin Cancer Res* **2005,** *11* (9), 3163-3169.

248. Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002,** *298* (5600), 1912-1934.

249. Zhang, H.; Zha, X.; Tan, Y.; Hornbeck, P. V.; Mastrangelo, A. J.; Alessi, D. R.; Polakiewicz, R. D.; Comb, M. J. Phosphoprotein analysis using antibodies broadly reactive against phosphorylated motifs. *J Biol Chem* **2002,** *277* (42), 39379-39387.

250. Sickmann, A.; Meyer, H. E. Phosphoamino acid analysis. *Proteomics* **2001,** *1* (2), 200-206.

251. Schlessinger, J. Cellular signaling by receptor tyrosine kinases. *Harvey Lect.* **1993,** *89,* 105-123.

252. Kalume, D. E.; Molina, H.; Pandey, A. Tackling the phosphoproteome: tools and strategies. *Curr. Opin. Chem Biol* **2003,** *7* (1), 64-69.

253. D'Ambrosio, C.; Salzano, A. M.; Arena, S.; Renzone, G.; Scaloni, A. Analytical methodologies for the detection and structural characterization of phosphorylated proteins. *J Chromatogr. B Analyt. Technol. Biomed Life Sci* **2007,** *849* (1-2), 163-180.

254. Rogers, L. D.; Foster, L. J. Phosphoproteomics--finally fulfilling the promise? *Mol Biosyst.* **2009,** *5* (10), 1122-1129.

255. Thingholm, T. E.; Larsen, M. R.; Ingrell, C. R.; Kassem, M.; Jensen, O. N. TiO(2)-based phosphoproteomic analysis of the plasma membrane and the effects of phosphatase inhibitor treatment. *J Proteome Res* **2008,** *7* (8), 3304-3313.

256. O'Callaghan, J. P.; Sriram, K. Focused microwave irradiation of the brain preserves in vivo protein phosphorylation: comparison with other methods of sacrifice and analysis of multiple phosphoproteins. *J Neurosci Methods* **2004,** *135* (1-2), 159-168.

257. Scholz, B.; Skold, K.; Kultima, K.; Fernandez, C.; Waldemarson, S.; Savitski, M. M.; Svensson, M.; Boren, M.; Stella, R.; Andren, P. E.; Zubarev, R.; James, P. Impact of temperature dependent sampling procedures in proteomics and peptidomics - A characterization of the liver and pancreas post mortem degradome. *Mol Cell Proteomics* **2010**.

258. Boucherie, H.; Massoni, A.; Monribot-Espagne, C. Radiolabeling for two-dimensional gel analysis. *Methods Mol Biol* **2008,** *424*, 125-135.

259. Campbell, D.; Morrice, N. Identification of protein phosphorylation sites by a combination of mass spectrometry and solid phase Edman sequencing. *J Biomol. Tech.* **2002,** *13* (3), 119-130.

260. Wettenhall, R. E.; Aebersold, R. H.; Hood, L. E. Solid-phase sequencing of 32P-labeled phosphopeptides at picomole and subpicomole levels. *Methods Enzymol.* **1991,** *201*, 186-199.

261. Bond, J. A.; Webley, K.; Wyllie, F. S.; Jones, C. J.; Craig, A.; Hupp, T.; Wynford-Thomas, D. p53-Dependent growth arrest and altered p53-immunoreactivity following metabolic labelling with 32P ortho-phosphate in human fibroblasts. *Oncogene* **1999,** *18* (25), 3788-3792.

262. Kinoshita, E.; Kinoshita-Kikuta, E.; Takiyama, K.; Koike, T. Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Mol Cell Proteomics* **2006,** *5* (4), 749-757.

263. Martin, K.; Steinberg, T. H.; Goodman, T.; Schulenberg, B.; Kilgore, J. A.; Gee, K. R.; Beechem, J. M.; Patton, W. F. Strategies and solid-phase formats for the analysis of protein and peptide phosphorylation employing a novel fluorescent phosphorylation sensor dye. *Comb. Chem High Throughput. Screen.* **2003,** *6* (4), 331-339.

264. Schulenberg, B.; Aggeler, R.; Beechem, J. M.; Capaldi, R. A.; Patton, W. F. Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. *J Biol Chem* **2003,** *278* (29), 27251-27255.

265. Han, M. S.; Kim, D. H. Naked-eye detection of phosphate ions in water at physiological pH: a remarkably selective and easy-to-assemble colorimetric phosphate-sensing probe. *Angew. Chem Int Ed Engl.* **2002,** *41* (20), 3809-3811.

266. Han, M. S.; Kim, D. H. Visual detection of AMP and real-time monitoring of cyclic nucleotide phosphodiesterase (PDE) activity in neutral aqueous solution. Chemosensor-coupled assay of PDE and PDE inhibitors. *Bioorg. Med Chem Lett.* **2003,** *13* (6), 1079-1082.

267. Cecconi, D.; Zamo, A.; Bianchi, E.; Parisi, A.; Barbi, S.; Milli, A.; Rinalducci, S.; Rosenwald, A.; Hartmann, E.; Zolla, L.; Chilosi, M. Signal transduction pathways of mantle cell lymphoma: a phosphoproteome-based study. *Proteomics* **2008,** *8* (21), 4495-4506.

268. Liu, J.; Cai, Y.; Wang, J.; Zhou, Q.; Yang, B.; Lu, Z.; Jiao, L.; Zhang, D.; Sui, S.; Jiang, Y.; Ying, W.; Qian, X. Phosphoproteome profile of human liver Chang's cell based on 2-DE with fluorescence staining and MALDI-TOF/TOF-MS. *Electrophoresis* **2007,** *28* (23), 4348-4358.

269. Yamagata, A.; Kristensen, D. B.; Takeda, Y.; Miyamoto, Y.; Okada, K.; Inamatsu, M.; Yoshizato, K. Mapping of phosphorylated proteins on two-dimensional polyacrylamide gels using protein phosphatase. *Proteomics* **2002,** *2* (9), 1267-1276.

270. Gronborg, M.; Kristiansen, T. Z.; Stensballe, A.; Andersen, J. S.; Ohara, O.; Mann, M.; Jensen, O. N.; Pandey, A. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol Cell Proteomics.* **2002,** *1* (7), 517-527.

271. Bergstrom, L. S.; Hagner-McWhirter, S.; Elfineh, L.; Molin, M.; Jorsback, A.; Ohman, J.; Pettersson, U. Detection of tyrosine phosphorylated proteins by combination of immunoaffinity enrichment, two-dimensional difference gel electrophoresis and fluorescent Western blotting. *Biochem Biophys. Res Commun.* **2010,** *401* (4), 581-585.

272. Bergstrom, L. S.; Molin, M.; Savitski, M. M.; Emilsson, L.; Astrom, J.; Hedberg, L.; Adams, C.; Nielsen, M. L.; Engstrom, A.; Elfineh, L.; Andersson, E.; Zubarev, R. A.; Pettersson, U. Immunoaffinity enrichments followed by mass spectrometric detection for studying global protein tyrosine phosphorylation. *J Proteome Res* **2008,** *7* (7), 2897-2910.

273. Roque, A. C.; Lowe, C. R. Lessons from nature: On the molecular recognition elements of the phosphoprotein binding-domains. *Biotechnol Bioeng.* **2005,** *91* (5), 546-555.

274. Roque, A. C.; Gupta, G.; Lowe, C. R. Design, synthesis, and screening of biomimetic ligands for affinity chromatography. *Methods Mol Biol* **2005,** *310*, 43-62.

275. Roque, A. C.; Lowe, C. R. Advances and applications of de novo designed affinity ligands in proteomics. *Biotechnol Adv.* **2006,** *24* (1), 17-26.

276. Pozuelo, R. M.; Geraghty, K. M.; Wong, B. H.; Wood, N. T.; Campbell, D. G.; Morrice, N.; Mackintosh, C. 14-3-3-affinity purification of over 200 human phosphoproteins reveals new links to regulation of cellular metabolism, proliferation and trafficking. *Biochem J* **2004,** *379* (Pt 2), 395-408.

277. Dubois, F.; Vandermoere, F.; Gernez, A.; Murphy, J.; Toth, R.; Chen, S.; Geraghty, K. M.; Morrice, N. A.; Mackintosh, C. Differential 14-3-3 affinity capture reveals new downstream targets of phosphatidylinositol 3-kinase signaling. *Mol Cell Proteomics* **2009,** *8* (11), 2487-2499.

278. The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res* **2010,** *38* (Database issue), D142-D148.

279. Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol* **1999,** *294* (5), 1351-1362.

280. Obenauer, J. C.; Cantley, L. C.; Yaffe, M. B. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res* **2003,** *31* (13), 3635-3641.

281. Hornbeck, P. V.; Chabra, I.; Kornhauser, J. M.; Skrzypek, E.; Zhang, B. PhosphoSite: A bioinformatics resource dedicated to physiological protein phosphorylation. *Proteomics* **2004,** *4* (6), 1551-1561.

282. Gnad, F.; Gunawardena, J.; Mann, M. PHOSIDA 2011: the posttranslational modification database. *Nucleic Acids Res* **2010**.

283. Diella, F.; Cameron, S.; Gemund, C.; Linding, R.; Via, A.; Kuster, B.; Sicheritz-Ponten, T.; Blom, N.; Gibson, T. J. Phospho.ELM: a database of experimentally verified phosphorylation sites in eukaryotic proteins. *BMC Bioinformatics* **2004,** *5*, 79.

284. Rush, J.; Moritz, A.; Lee, K. A.; Guo, A.; Goss, V. L.; Spek, E. J.; Zhang, H.; Zha, X. M.; Polakiewicz, R. D.; Comb, M. J. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat. Biotechnol.* **2005,** *23* (1), 94-101.

285. Davis, F. M.; Tsao, T. Y.; Fowler, S. K.; Rao, P. N. Monoclonal antibodies to mitotic cells. *Proc Natl Acad Sci U S A* **1983,** *80* (10), 2926-2930.

286. Cohen, B.; Yoakim, M.; Piwnica-Worms, H.; Roberts, T. M.; Schaffhausen, B. S. Tyrosine phosphorylation is a signal for the trafficking of pp85, an 85-kDa phosphorylated polypeptide associated with phosphatidylinositol kinase activity. *Proc Natl Acad Sci U S A* **1990,** *87* (12), 4458-4462.

287. Kristensen, D. B.; Kawada, N.; Imamura, K.; Miyamoto, Y.; Tateno, C.; Seki, S.; Kuroki, T.; Yoshizato, K. Proteome analysis of rat hepatic stellate cells. *Hepatology* **2000,** *32* (2), 268-277.

288. Yamagata, A.; Kristensen, D. B.; Takeda, Y.; Miyamoto, Y.; Okada, K.; Inamatsu, M.; Yoshizato, K. Mapping of phosphorylated proteins on two-dimensional polyacrylamide gels using protein phosphatase. *Proteomics* **2002,** *2* (9), 1267-1276.

289. Guy, G. R.; Philip, R.; Tan, Y. H. Analysis of cellular phosphoproteins by two-dimensional gel electrophoresis: applications for cell signaling in normal and cancer cells. *Electrophoresis* **1994,** *15* (3-4), 417-440.

290. Vosseller, K.; Hansen, K. C.; Chalkley, R. J.; Trinidad, J. C.; Wells, L.; Hart, G. W.; Burlingame, A. L. Quantitative analysis of both protein expression and serine / threonine post-translational modifications through stable isotope labeling with dithiothreitol. *Proteomics* **2005,** *5* (2), 388-398.

291. Poot, A. J.; Ruijter, E.; Nuijens, T.; Dirksen, E. H.; Heck, A. J.; Slijper, M.; Rijkers, D. T.; Liskamp, R. M. Selective enrichment of Ser-/Thr-phosphorylated peptides in the presence of Ser-/Thr-glycosylated peptides. *Proteomics* **2006,** *6* (24), 6394-6399.

292.   van der Veken, P.; Dirksen, E. H.; Ruijter, E.; Elgersma, R. C.; Heck, A. J.; Rijkers, D. T.; Slijper, M.; Liskamp, R. M. Development of a novel chemical probe for the selective enrichment of phosphorylated serine- and threonine-containing peptides. *Chembiochem* **2005,** *6* (12), 2271-2280.

293.   Oda, Y.; Nagasu, T.; Chait, B. T. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat. Biotechnol.* **2001,** *19* (4), 379-382.

294.   McLachlin, D. T.; Chait, B. T. Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Anal Chem* **2003,** *75* (24), 6826-6836.

295.   Zhou, H.; Watts, J. D.; Aebersold, R. A systematic approach to the analysis of protein phosphorylation. *Nat Biotechnol* **2001,** *19* (4), 375-378.

296.   Tao, W. A.; Wollscheid, B.; O'Brien, R.; Eng, J. K.; Li, X. J.; Bodenmiller, B.; Watts, J. D.; Hood, L.; Aebersold, R. Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry. *Nat Methods* **2005,** *2* (8), 591-598.

297.   Porath, J.; Carlsson, J.; Olsson, I.; Belfrage, G. Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **1975,** *258* (5536), 598-599.

298.   Liu, H. L.; Ho, Y.; Hsu, C. M. Molecular simulations to determine the chelating mechanisms of various metal ions to the His-tag motif: a preliminary study. *J Biomol. Struct. Dyn.* **2003,** *21* (1), 31-41.

299.   Nuhse, T. S.; Stensballe, A.; Jensen, O. N.; Peck, S. C. Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol Cell Proteomics* **2003,** *2* (11), 1234-1243.

300.   Zachariou, M.; Hearn, M. T. Protein selectivity in immobilized metal affinity chromatography based on the surface accessibility of aspartic and glutamic acid residues. *J Protein Chem* **1995,** *14* (6), 419-430.

301.   Ueda, E. K.; Gout, P. W.; Morganti, L. Current and prospective applications of metal ion-protein binding. *J Chromatogr. A* **2003,** *988* (1), 1-23.

302.   Conrads, T. P.; Issaq, H. J.; Veenstra, T. D. New tools for quantitative phosphoproteome analysis. *Biochem Biophys. Res Commun.* **2002,** *290* (3), 885-890.

303.   He, T.; Alving, K.; Feild, B.; Norton, J.; Joseloff, E. G.; Patterson, S. D.; Domon, B. Quantitation of phosphopeptides using affinity chromatography and stable isotope labeling. *J Am Soc Mass Spectrom* **2004,** *15* (3), 363-373.

304.   Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* **2005,** *4* (7), 873-886.

305.   Pinkse, M. W.; Uitto, P. M.; Hilhorst, M. J.; Ooms, B.; Heck, A. J. Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal Chem* **2004,** *76* (14), 3935-3943.

306.   Kweon, H. K.; Hakansson, K. Selective zirconium dioxide-based enrichment of phosphorylated peptides for mass spectrometric analysis. *Anal Chem* **2006,** *78* (6), 1743-1749.

307.   Sano, A.; Nakamura, H. Titania as a chemo-affinity support for the column-switching HPLC analysis of phosphopeptides: application to the characterization of phosphorylation sites in proteins by combination with protease digestion and electrospray ionization mass spectrometry. *Anal Sci* **2004,** *20* (5), 861-864.

308.   Chen, C. T.; Chen, Y. C. Fe3O4/TiO2 core/shell nanoparticles as affinity probes for the analysis of phosphopeptides using TiO2 surface-assisted laser desorption/ionization mass spectrometry. *Anal Chem* **2005,** *77* (18), 5912-5919.

309.   Sano, A.; Nakamura, H. Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides. *Anal Sci* **2004,** *20* (3), 565-566.

310.   Mann, M.; Jensen, O. N. Proteomic analysis of post-translational modifications. *Nat Biotechnol* **2003,** *21* (3), 255-261.

311.   Larsen, M. R.; Sorensen, G. L.; Fey, S. J.; Larsen, P. M.; Roepstorff, P. Phospho-proteomics: evaluation of the use of enzymatic de-phosphorylation and differential mass spectrometric peptide mass mapping for site specific phosphorylation assignment in proteins separated by gel electrophoresis. *Proteomics* **2001,** *1* (2), 223-238.

312. Tholey, A.; Reed, J.; Lehmann, W. D. Electrospray tandem mass spectrometric studies of phosphopeptides and phosphopeptide analogues. *J Mass Spectrom* **1999,** *34* (2), 117-123.

313. Reid, G. E.; Simpson, R. J.; O'Hair, R. A. Leaving group and gas phase neighboring group effects in the side chain losses from protonated serine and its derivatives. *J Am Soc Mass Spectrom* **2000,** *11* (12), 1047-1060.

314. Thingholm, T. E.; Jorgensen, T. J.; Jensen, O. N.; Larsen, M. R. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat Protoc.* **2006,** *1* (4), 1929-1935.

315. Edelson-Averbukh, M.; Shevchenko, A.; Pipkorn, R.; Lehmann, W. D. Gas-phase intramolecular phosphate shift in phosphotyrosine-containing peptide monoanions. *Anal Chem* **2009,** *81* (11), 4369-4381.

316. Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc Natl Acad Sci U S A* **2007,** *104* (7), 2193-2198.

317. DeGnore, J. P.; Qin, J. Fragmentation of phosphopeptides in an ion trap mass spectrometer. *J Am Soc Mass Spectrom* **1998,** *9* (11), 1175-1188.

318. Edelson-Averbukh, M.; Pipkorn, R.; Lehmann, W. D. Phosphate group-driven fragmentation of multiply charged phosphopeptide anions. Improved recognition of peptides phosphorylated at serine, threonine, or tyrosine by negative ion electrospray tandem mass spectrometry. *Anal Chem* **2006,** *78* (4), 1249-1256.

319. Edelson-Averbukh, M.; Pipkorn, R.; Lehmann, W. D. Analysis of protein phosphorylation in the regions of consecutive serine/threonine residues by negative ion electrospray collision-induced dissociation. Approach to pinpointing of phosphorylation sites. *Anal Chem* **2007,** *79* (9), 3476-3486.

320. Steen, H.; Kuster, B.; Fernandez, M.; Pandey, A.; Mann, M. Detection of tyrosine phosphorylated peptides by precursor ion scanning quadrupole TOF mass spectrometry in positive ion mode. *Anal Chem* **2001,** *73* (7), 1440-1448.

321. Larsen, M. R.; Thingholm, T. E.; Jensen, O. N.; Roepstorff, P.; Jorgensen, T. J. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* **2005,** *4* (7), 873-886.

Curriculum vitae

Ignazio Garaguso

Born: 07 Dec.1970 at Erice (Italy)

Nationality: Italian

E-mail: igaragu@tin.it

Professional experience

Since Oct. 2005: Fraunhofer Institute of Toxicology and Experimental Medicine (ITEM), Hannover (D)
Department of Molecular Medicine and Medical Biotechnology
Research fellow/Scientist

1999 - 2005: Chiron Vaccines s.r.l., (now Novartis Vaccines & Diagnostics), Siena, (I)
Research centre, Biochemistry and Molecular Biology department- Proteomics and mass spectrometry facility
Research fellow

1997 - 1999: University of Siena (SMFN), Siena (I).
Faculty of Mathematics, Physical and natural sciences- Molecular Biology Department
Scientific Consultant

1993 - 1996: "G&G rappresentanze", Reggio Calabria (I).
Intermediation agency (owner of the agency) - Pharmaceutical products
Freelancer sales representative

1990 - 1992: Collardin GmbH, Herborn-Schönbach (D)
Department of environmental safety
Chemical technician

Fellowships

2006-2008: Research fellow as visiting scientist from "Fraunhofer Gesellschaft", Munich (D)

Education

2005: Università degli studi di Siena, Siena (I)
Degree in Biological Sciences with a specialisation in Cellular and Molecular Biology.
Dissertation: Evaluation of sera elicited against recombinant membrane proteins of *Chlamydia Pneumoniae* trough the use of Proteomic methods.
Obtained the title of Dottore in Scienze Biologiche.
Score: 105/110

1989: Istituto Tecnico Industriale Statale "A. Panella", Reggio Calabria (I)
High School Degree in Industrial Engineering with specialization in industrial experimental chemistry
Technical Thesis: Method used for the analysis of Wastewater in a Molecular sieves production site
Obtained the title of Perito chimico.
Score: 48/60

Languages

English: fluent written and spoken
German: fluent written and spoken
French: basic level
Spanish: good written and spoken

Technical competences

Molecular biology: cloning, (bacterial and mammalian systems), reporter gene systems, PCR.
Cell biology: cell lines(LLC-MK2, HepG2, A549) and bacterial (*E. Coli*, *N. Meningitidis*, *group A, B Streptococcus*, *C. Pneumoniae*) cultures.
Biochemistry: Wide experience in recombinant proteins expression: cell-based systems, in-vitro transcription/translation. Protein purification: GST, His tag, Immunoprecipitation. Chromatography: ion-exchange, reversed phase, gel filtration. Well experienced in Solid phase extraction method development. Large experience with HPLC and NanoHPLC systems using multidimensional separations techniques for advanced analyses of complex mixtures (MS coupled). Protein quantification methods: Lowry, Bradford and Kjeldahl. Automatic sequencing: "Edman degradation". DNA and Protein separation: Agarose gel electrophoresis, IEF, SDS-PAGE, 2DE, electroblotting. Technical skill in Gas Chromatography atomic absorption spectroscopy, ELISA and UV/VIS spectroscopy.
Mass Spectrometry: Profound experience with MALDI-ToF-MS, ESI-Q-ToF-MS, ESI-IonTrap-MS.

Informatics

Well experienced in the use of: Microsoft Office suite (all applications), Acrobat, Reference Manager, Photoshop and several internet browser.
Expert use of software for: image analysis (e.g.: Quantity One, PDQuest, Melanie, DeCyder); sequence and chemical structure analysis (e.g.: PdbViewer, BLAST, ChemSketch); array data analysis (GeneSpring, Array track) "Web-available" database and bioinformatics tools (e.g.: ExPasy, NCBI, PIR, ProteinProspector) and automatic analytical systems management (e.g.: HyStar, Proteinscape, Chromeleon, ProteinLynx).

List of Publications

Journal articles

-<u>Garaguso I</u> and Borlak J.
Rapid screening assay to search for phosphorylated proteins in tissue extracts.
Submitted.

-<u>Garaguso I</u> and Borlak J.
Method for the rapid detection and molecular characterization of DNA alkylating agents by MALDI-TOF Mass spectrometry.
Anal. Chem., **2010**, 82 (20), pp 8573–8582.

-<u>Garaguso I</u> and Borlak J.
Matrix Layer sample preparation: an improved MALDI-MS peptide analysis method for proteomic studies.
Proteomics. **2008**, 8 (13):2583-95.

-Bijon Chatterji, Maria Stella Ritorto, <u>Ignazio Garaguso</u>, -Eckard Belau, Giuseppe Gazzana, Jürgen Borlak
Onkoproteomics von Lungentumoren.
Biospektrum. **2006**, 12 (7): 732-35

-Germano Ferrari, <u>Ignazio Garaguso</u>, Jeannette Adu-Bobie, Francesco Doro, Anna Rita Taddei, Alessia Biolchi1, Brunella Brunelli, Marzia Monica Giuliani, Mariagrazia Pizza, Nathalie Norais, and Guido Grandi.
Outer membrane vesicles from group B *Neisseria meningitidis* Δgna33 mutant: proteomic and immunological comparison with detergent-derived outer membrane vesicles.
Proteomics **2006**, 6 (6):1856-66.

-Manuel J. Rodríguez-Ortega, Nathalie Norais, Giuliano Bensi, Sabrina Liberatori, Sabrina Capo, Marirosa Mora, Maria Scarselli, Francesco Doro, Germano Ferrari, <u>Ignazio Garaguso</u>, Tiziana Maggi, Anita Neumann, Alessia Covre, John L. Telford, and Guido Grandi. **Characterization and identification of vaccine candidate proteins through analysis of the group A Streptococcus surface proteome.**.
Nature Biotech. **2006**, 24 (2):191-7.

-Oretta Finco, Alessandra Bonci, Mauro Agnusdei, Maria Scarselli, Roberto Petracca, Nathalie Norais, Germano Ferrari, <u>Ignazio Garaguso</u>, Manuela Donati, Vittorio Sambri, Roberto Cevenini, Giulio Ratti, and Guido Grandi.
Identification of new potential vaccine candidates against *Chlamydia pneumoniae* by multiple screenings. Vaccine **2005**, 23 (9): 1178-88.

-Montigiani S, Falugi F, Scarselli M, Finco O, Petracca R, Galli G, Mariani M, Manetti R, Agnusdei M, Cevenini R, Donati M, Nogarotto R, Norais N, <u>Garaguso I</u>, Nuti S, Saletti G, Rosa D, Ratti G, Grandi G. **Genomic approach for analysis of surface proteins in *Chlamydia pneumoniæ***.
Infect Immun. **2002**, 70 (1): 368-79.

-Nathalie Norais, Renzo Nogarotto, Emilia Tiziana Iacobini, <u>Ignazio Garaguso</u>, Renata Grifantini, Giuliano Galli and Guido Grandi. **Combined automated PCR cloning, in vitro transcription/translation and two-dimensional electrophoresis for bacterial proteome analysis**. Proteomics **2001**, 1 (11): 1378-1389.

Books

-Norais N., <u>Garaguso I.</u>, Ferrari G. and Grandi G.
In Vitro Transcription and Translation Coupled to Two-Dimensional Electrophoresis for Bacterial Proteome Analysis.
Methods in molecular biology, II edition, (Grandi, G., ed.) Human Press, Totowa, N J, **2007**, 375 pp. 183-209.

-Norais N., Nogarotto R., Iacobini ET., Garaguso I., Grifantini R., Galli G., and Grandi G.

Protein identification from 2-D gels using *in vitro* transcription translation products.
Handbook of Proteomic Methods. (Conn P. M., ed.) Human Press Totowa, N J., **2003**, pp.17-28.

Oral presentations

Title: **Detection and characterization of polycyclic aromatic hydrocarbon-DNA adducts by MALDI TOF MS**.
August **2009**- 18th International Mass Spectrometry Conference (Bremen, Germany).

Title: **MALDI mass spectrometry for the analysis of DNA-Polycyclic Aromatic Hydrocarbons adducts**.
September **2008**- Gesprächskreis BITUMEN 17thMeeting (Bochum, Germany).

Title: **Experimental approaches to investigate the posttranslational modifications of proteins**.
February **2008**-Das Massenspektrometrie-Forum Hannover (Hannover, Germany).

Title: **Posttranslational modification and cancer: Methodological considerations**.
March **2007**- Mass Spectrometry Solutions for Life Science Research, Bruker Daltonik Seminar tour (Hannover, Germany).

Posters

-Garaguso I., Borlak J.
Improved MALDI MS peptide mass mapping using Matrix Layer sample preparation method.
Proteomic Forum 2009. April **2009** (Berlin, Germany).

-Garaguso I., Borlak J.
Protein phosphorylation characterization by MALDI-TOF mass spectrometry with the improved matrix layer sample preparation.
HUPO 7th World congress. August **2008** (Amsterdam, The Netherlands).

-F. Doro, M.J. Rodríguez-Ortega, G. Ferrari, I. Garaguso, A.R. Taddei, J.L. Telford, N. Norais, G. Grandi
Proteomic Analysis of membrane vesicles produced by group B *streptococcus* grown in Presence of a non lethal concentration of penicillin.
7th Siena Meeting. September **2006** (Siena, Italy)

-Garaguso I., Rodrìguez-Ortega M.J., Liberatori S., Doro F., Ferrari G., Capo S. , Bensi G., Telford J., Norais N. Grandi G.
Proteomic analysis of vesicle-like structure released by Streptococci grown in the presence of Penicillin.
Chiron R&D Days and expo. April **2005** (Emeryville, USA).

-Garaguso I., Ferrari G., Nogarotto R., Norais,N., Grandi,G.
Neisseria meningitidis outer membrane sub-proteome.
Chiron R&D Days and expo. April **2004** (Emeryville, USA).