Applications of Mass Spectrometry in Life Safety

Edited by

Crisan Popescu

Alina D. Zamfir

Nicolae Dinca

Springer

Applications of Mass Spectrometry in Life Safety

NATO Science for Peace and Security Series

This Series presents the results of scientific meetings supported under the NATO Programme: Science for Peace and Security (SPS).

The NATO SPS Programme supports meetings in the following Key Priority areas: (1) Defence Against Terrorism; (2) Countering other Threats to Security and (3) NATO, Partner and Mediterranean Dialogue Country Priorities. The types of meeting supported are generally "Advanced Study Institutes" and "Advanced Research Workshops". The NATO SPS Series collects together the results of these meetings. The meetings are co-organized by scientists from NATO countries and scientists from NATO's "Partner" or "Mediterranean Dialogue" countries. The observations and recommendations made at the meetings, as well as the contents of the volumes in the Series, reflect those of parti-cipants and contributors only; they should not necessarily be regarded as reflecting NATO views or policy.

Advanced Study Institutes (ASI) are high-level tutorial courses intended to convey the latest developments in a subject to an advanced-level audience

Advanced Research Workshops (ARW) are expert meetings where an intense but informal exchange of views at the frontiers of a subject aims at identifying directions for future action

Following a transformation of the programme in 2006 the Series has been re-named and re-organised. Recent volumes on topics not related to security, which result from meetings supported under the programme earlier, may be found in the NATO Science Series.

The Series is published by IOS Press, Amsterdam, and Springer, Dordrecht, in conjunction with the NATO Public Diplomacy Division.

Sub-Series

A.	Chemistry and Biology	Springer
B.	Physics and Biophysics	Springer
C.	Environmental Security	Springer
D.	Information and Communication Security	IOS Press
E.	Human and Societal Dynamics	IOS Press

http://www.nato.int/science
http://www.springer.com
http://www.iospress.nl

Series A: Chemistry and Biology

Applications of Mass Spectrometry in Life Safety

edited by

Crisan Popescu

DWI e.V. and Institute of Technical and Macromolecular Chemistry,
RWTH Aachen University,
Germany

Alina D. Zamfir

University "Aurel Vlaicu",
Arad, Romania

and

Nicolae Dinca

University "Aurel Vlaicu",
Arad, Romania

Springer

CONTENT

PREFACE

Mass spectrometry (MS) along with its hyphenated techniques is capable of high throughput, sensitivity, accuracy and selectivity for the analysis of structure and composition of almost any product. Like in electrophoresis, MS separates molecules based on the mass-to-charge ratio. In case of gel electrophoresis (SDS-PAGE), a well-known and efficient bioanalytical technique, proteins bear negative charges but have the same charge density, so proteins are separated according to their size. Similarly, in case of MS analysis, proteins carry the same charge, and are separated by their molecular weight. Unlike SDS-PAGE, however, modern ultra high resolution MS discerns very small mass differences and can resolve and completely identify in a single experiment species of the same nominal mass in complex biological mixtures. Consequently, MS can be used for the structural characterization, identification and sensitive detection of mixtures of biomolecules or for assessing the quality of isolated proteins (purity, integrity, or post-translational modifications, for example), carbohydrates, nucleic acids, drugs, metabolites, pollutants etc.

In the post-genome era, MS is continuously developing as one of the most reliable analytical method for elucidating the structure of molecules originating from various biological matrices. The potential of MS for high-sensitive structural analyses became unsurpassable after the introduction of electrospray (ESI) and matrix assisted laser/desorption ionization (MALDI) methods, on one hand, and the possibility to deduce in detail unknown biopolymer structures by highly accurate molecular mass measurement followed by sequencing using dissociation techniques based on multiple stage MS, on the other.

For food chemistry and environmental studies MS was uniquely shown to be capable to provide straightforward the most valuable *de novo* identification at the highest sensitivity.

Such achievements of mass spectrometry are continuously signalled worldwide for already 20 years. In view of the high speed data acquisition and the precision of data interpretation offered nowadays by the modern equipment, MS is currently by far the most performant method for monitoring the life chemical and biochemical security.

Although mass spectrometry can be well considered a mature field, researchers still unveil new applications or develop new spin-offs of the classical techniques. Some of them find surprising clinical and diagnostic applications. In this respect, MS is in its early years, but it is spectacularly growing in three areas: tissue imaging, cancer biomarker detection, and bioterrorism detection.

The aim of this book is to bring into focus the hot topics of life science for which mass spectrometry is becoming the tool of analysis, and to provide a comprehensive review of the research topics most pertinent to the advancement of mass spectrometry and its general applications. The volume brings together contributions

from eminent international researchers in the field, covering various aspects ranging from the fundamentals of MALDI and ESI to their applications to most actual omics (proteomics, lipidomics, glycomics) or to immunology. The large group of small molecules analyzed with mass spectrometry techniques is also represented by brominates and keratins, or plant extracts.

It is hoped that this volume will also provide added stimulus for further progress in the field, where major challenges still remain and great potential exists for new breakthroughs. The ability of mass spectrometry for becoming a sound diagnosis technique of neurodegenerative diseases like Alzheimer, and for being a useful tool in immunotherapy is a fertile ground for further innovation and exploitation in the future.

In preparing the book, we have relied on the contribution of the authors, without whose expert insight, motivation and commitment the publication of this volume would not have been possible. We, thus, extend our appreciation to all the authors.

We also convey our thanks to Springer for affording us the opportunity to publish this volume and to the editorial and publishing staff, in particular Wil Bruins, for the assistance, organization and efficiency in coordinating the timely preparation and production of the book.

We are also grateful to the North Atlantic Treaty Organization for their valuable support of the Advanced Research Workshop on *Mass spectrometry for Life science 2007*, Arad/Herculane, Romania, which served as the original impetus for the publication of this volume.

<div align="right">

Crisan Popescu
Alina D. Zamfir
Nicolae Dinca
Aachen, Arad, February 2008

</div>

Part I

PROTEOMICS

AND

PEPTIDOMICS

1. BLUE NATIVE PAGE AND MASS SPECTROMETRY ANALYSIS OF EPHRIN STIMULATION-DEPENDENT PROTEIN-PROTEIN INTERACTIONS IN NG108-EPHB2 CELLS

COSTEL C. DARIE[*1], VIVEKANANDA SHETTY, DANIEL S. SPELLMAN, GUOAN ZHANG, CHONGFENG XU, HELENE L. CARDASIS, STEVEN BLAIS, DAVID FENYO, AND THOMAS A. NEUBERT
Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18, New York, NY 10016, USA

Abstract. Receptor tyrosine kinases (RTK) are proteins that undergo dimerization and/or multimerization and autophosphorylation in response to ligand stimulation. Members of the RTK family are receptors for a series of growth factors that, upon stimulation, are able to start signaling events that promote cell growth and differentiation. A class of RTKs, the Eph receptors (EphRs), are found in a variety of cell types and play important roles in patterning the central and peripheral nervous systems, as well as in synapse and neural crest formation. Interaction of Eph receptors with their ephrin ligands activates signal transduction pathways that lead to cytoskeletal remodeling through formation of many stable or transient protein-protein interactions. However, these intracellular signal transduction pathways that lead to cytoskeletal remodeling are not well understood. Here, we combined Blue Native PAGE (BN-PAGE) and mass spectrometry (MS) to analyze protein-protein interactions as a result of ephrin stimulation. We analyzed both lysates and phosphotyrosine immunoprecipitate (pY99-IP) of unstimulated and ephrin-stimulated cells. Our experiments allowed us to characterize many constitutive homo- and hetero-protein complexes from the cell lysate. Furthermore, BN-PAGE and MS of the pY99-IPs from both unstimulated and stimulated cells allowed us to analyze protein-protein interactions that resulted upon ephrin stimulation. Combination of BN-PAGE and MS also has the potential for the analysis of stable and transient protein-protein interactions in other ligand-stimulated RTK-dependent signal transduction pathways.

1.   Introduction

Receptor tyrosine kinases are proteins that undergo dimerization and/or multimerization and autophosphorylation in response to ligand stimulation. Members of the RTK family are receptors for a series of growth factors like epidermal growth factor (EGF), platelet derived growth factor (PDGF), vascular endothelial growth factor (VEGF) that, upon stimulation, are able to start signaling events that promote

---

[*]Corresponding author. E-mail: darie@saturn.med.nyu.edu

cell growth and differentiation. A common feature for these RTKs is that their growth factor ligands are secreted proteins.

A different class of RTKs is represented by Eph receptors (EphRs) and is the largest sub-family of RTKs. These EphRs are found in a variety of cell types from developing and mature tissues and play important roles in patterning the central and peripheral nervous systems (Holland et al., 1996, 1998; Wilkinson, 2001), synapse formation (Dalva et al., 2000), neural crest formation (Robinson et al., 1997), as well as angiogenesis and vascular system remodeling (Yancopoulos et al., 1998; Cheng et al., 2002; Adams, 2003; Wang et al., 2006). EphRs differ from other RTKs in three regards: (1) EphRs do not promote cell growth and differentiation, but rather play important roles in attractive and repellent reactions between the cells. (2) Activation of EphRs is performed by ephrins that, unlike the other RTK soluble ligands, are anchored to cell membranes. (3) During stimulation, EphRs may be either receptors that are activated by ephrin ligands and trigger the downstream signaling or they may be ligands that activate upstream signaling in ephrin-bearing cells.

To date, at least 14 different EphRs have been identified, and based on the type of ligand with which they interact, they have been divided into two classes: EphAs and EphBs (Gale et al., 1996). EphA receptors comprise nine members: EphA1-A8 and EphA10. They interact with ephrins A which consist of five members: ephrin A1-A5. EphB receptors comprise five members: EphB1-B4 and EphB6. They interact with ephrins B which consist of three members: ephrinB1-B3 (Gale et al., 1996). Although all EphRs and ephrinBs are transmembrane proteins that contain extracellular and cytosolic domains, ephrinsA are connected to the cell surface by glycosylphosphatidylinositol anchors, and therefore do not contain a cytosolic domain.

It is commonly agreed that interaction of ephrins with Eph receptors leads to a gradual formation of ephrin-EphRs multimers that trigger cytoskeletal remodeling. This underlies cell adhesion and motility in the ligand and receptor bearing cells (Miao et al., 2000; Carter et al., 2002; Vearing and Lackmann, 2005). However, the intracellular signal transduction pathways that lead to cytoskeletal remodeling in the ephrin- or EphR-bearing cells are not well understood. Understanding of these pathways is further complicated by the bidirectional signaling that occurs between EphB receptors and their ephrinB ligands (Holland et al., 1996; Bruckner et al., 1997; Holland et al., 1998; Cowan and Henkemeyer, 2001).

Upon activation of signal transduction pathways, many transient protein-protein interactions occur, leading to the formation of different protein complexes. These complexes form during the transition of the cells from the unstimulated to stimulated state. Knowledge about the structural and functional interaction of the proteins into protein complexes as a result of ephrin stimulation will lead to a more comprehensive understanding of ephrin signaling.

BN-PAGE has long been used to analyze protein-protein interactions and protein complexes (Schagger and von Jagow, 1991; Schagger et al., 1994; Darie et al., 2005; Darie et al., 2007). This method separates protein complexes based on external charge induced by Coomassie dye and according to their molecular mass

(mass). Separation of protein complexes from the BN-PAGE lane by second dimension SDS-PAGE may reveal the subunit composition of a particular protein complex, as well as the interacting partners of a particular protein. MS (Aebersold and Mann, 2003) is another tool that allows scientists to identify not only one but most potential interacting partners of a protein involved in signal transduction pathways.

In this paper, we combined BN-PAGE and MS to analyze protein complexes from the lysates and pY99-IPs of unstimulated and ephrin-stimulated cells. Our experiments allowed us to determine the molecular mass and subunit composition of many constitutive homo- and hetero-protein complexes from cell lysates. Furthermore, BN-PAGE and MS of the pY99-IPs from both unstimulated and stimulated cells allowed us to analyze protein-protein interactions that resulted from ephrin stimulation. Combination of BN-PAGE and MS also has the potential to be applied to the analysis of stable and transient protein-protein interactions in other ligand-stimulated RTK-dependent signal transduction pathways.

2.   Experimental Design

These studies have been designed based on the hypothesis that specific protein complexes are formed to accomplish cell signaling and that deciphering the composition of these complexes will reveal significant insights about mechanisms of ephrin signaling.

2.1.   CELL SYSTEM AND ACTIVATION OF THE SIGNAL TRANSDUCTION PATHWAYS

The strategy for analyzing protein complexes from EphB2 signaling in unstimulated and ephrinB1-Fc-stimulated NG 108 cells is shown in Figure 1. Unstimulated NG 108 cells that stably express EphB2 receptors (NG108-EphB2) were stimulated with soluble, pre-clustered ephrinB1-Fc for 45 min at 37°C, according to a published procedure (Zhang et al., 2006). The control (unstimulated) cells were treated with soluble, pre-clustered IgG-Fc. Once the cells had been stimulated and the signal transduction pathways had been activated, the cells were lysed. The lysates of unstimulated and stimulated NG 108 cells were separated by BN-PAGE or were used as starting material for antibody-based purification (phosphotyrosine immuno-purification: pY99-IP). The pY99-IPs were also separated by BN-PAGE. The protein complexes that resulted were then subjected to biochemical and/or MS analysis (BN-PAGE and Western blotting [WB] or BN-PAGE and MS), followed by comparison of the data that resulted from analysis of unstimulated and stimulated cells and ultimately with the current literature.
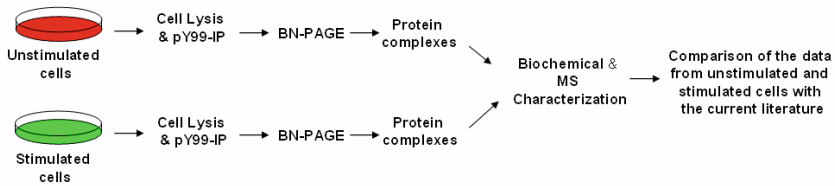
Figure 1. A workflow for the biochemical and MS characterization of ephrin signaling. The unstimulated or stimulated cells were lysed and the lysates (and/or the pY99-IPs from the lysate) were separated by BN-PAGE. The protein complexes/protein interactions were characterized biochemically and/or by MS, followed by comparison with the current literature

2.2.   BIOCHEMICAL CHARACTERIZATION OF PROTEIN COMPLEXES

BN-PAGE was used for separation of stable or transient protein complexes from cell lysates or pY99-IPs by using one-dimensional (BN-PAGE; 1D) or two-dimensional (BN-PAGE; 2D) electrophoresis. A schematic of biochemical characterization of the protein complexes is presented in Figure 2. BN-PAGE (1D) gel lanes were silver stained or electroblotted and analyzed with antibodies against proteins of interest. In addition, the BN-PAGE gel lanes were used for MS analysis. The protein complexes that were separated in the BN-PAGE (1D) were also separated under reducing and denaturing conditions in (2D) and analyzed by either silver staining (2D-SS) or WB (2D-WB) (Figure 2A).

The data that resulted from these experiments were interpreted based on differences between the lysates/pY99-IPs of unstimulated and stimulated cells that were separated by BN-PAGE and either 2D-SS, 2D-WB or MS. For example, the simplest scenario for the biochemical analysis of EphB2-associated proteins upon ephrin-Fc stimulation is schematically represented in Figures 2B and C. It is known that EphRs assemble into multimers upon ephrinB1-Fc stimulation. Therefore, the mass of the EphB2-containing protein complex will shift towards a higher mass, and will be visible as a shift in BN-PAGE (1D-WB) and/or 2D-WB (Figure 2B). By comparing the WB results from the unstimulated and stimulated cells, the mass of the EphB2-containing protein complex may be determined. Furthermore, MS analysis of the BN-PAGE gel bands that contain EphB2 from both unstimulated and stimulated cells may provide the identity of the EphB2 interactors (Figure 2B). Another scenario for the biochemical analysis of EphB2-associated proteins upon ephrin-Fc stimulation is shown in (Figure 2C). In this case, it may be determined whether two different proteins (e.g., EphB2 and FAK) are or are not part of the same protein complex before stimulation and what are their protein interactors. Furthermore, it may also be determined whether these two proteins associate into a protein complex or form two different complexes as a result of ephrin-Fc stimulation as well as their subunit composition (Figure 2C).
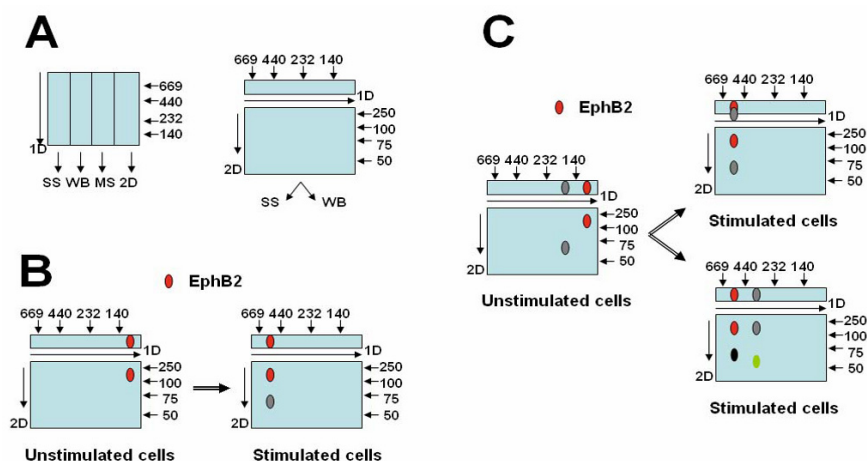
Figure 2. Biochemical characterization of ephrin signaling. (A) The starting material (lysate or pY99-IP) was separated by BN-PAGE (1D) and further used for SS, WB, MS or 2D analysis. The (2D) gel that resulted from separation of the (1D) gel lane was further used for SS or WB. (B) Comparison of the data resulting from either BN-PAGE (1D) and MS or BN-PAGE (2D) and WB of unstimulated and stimulated cells may lead to identification of the ligand stimulation-dependent protein interactors of a particular protein. (C) As in (B), except that the two different proteins investigated may interact with each other as a result of ligand stimulation, or form separate complexes with different interactors

Therefore, combination of the BN-PAGE and MS may provide information about the size of the protein complex and about the number and identities of the proteins that associate with EphB2-receptors and other signal transducing molecules. Furthermore, these experiments may allow analysis of both stable and transient protein-protein interactions as a result of ephrinB1-Fc stimulation.

2.3. MASS SPECTROMETRIC CHARACTERIZATION OF PROTEIN COMPLEXES

The general strategy for the MS characterization of the protein complexes involved in ephrin signaling is shown in Figure 3. The BN-PAGE gel lanes that contained the protein complexes were cut into 20–30 gel pieces and subjected to enzymatic digestion (tryptic digest). The resulting peptide mixture was then analyzed by reverse phase liquid chromatography and MS (LC-MS-MS/MS: NanoAquity LC and Waters Q-Tof Premier MS), and the raw data were processed and submitted to a database search using Mascot for protein identification. These experiments provided a list of proteins for each gel band that may be part of the same protein complex. The outcomes from these experiments were compared between the unstimulated and stimulated cells and then compared with the results obtained using biochemical analysis of the protein complexes. All these results were ultimately compared with the current literature.

BN-PAGE (1D)

669
440
232
140

| Cut gel into 20-30 pieces and enzymatic digestion of the gel bands | RP-HPLC and MS/MS and Identification of the proteins | Comparison with the biochemical results and current literature |
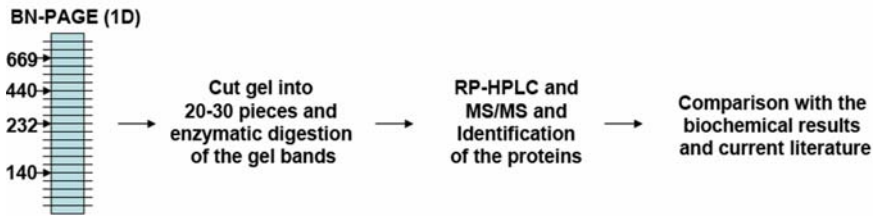
Figure 3. MS characterization of ephrin signaling. BN-PAGE (1D) gel bands were cut into pieces, digested with trypsin, and the peptides separated by LC (Nano-Aquity) and analyzed by MS (Waters Q-Tof Premier). The identified proteins were compared with the biochemical results and current literature

3.   Results

3.1. ANALYSIS BY GEL ELECTROPHORESIS OF LYSATES FROM UNSTIMULATED AND EPHRINB1-FC STIMULATED CELLS

Initially, the lysates from both unstimulated and stimulated cells were separated by SDS-PAGE and then stained with either silver (Figure 4A) or Coomassie (data not shown). No significant or visible differences in the protein pattern between the two gel lanes were observed. However, when the SDS-PAGE gel was electroblotted and incubated with pY99 antibodies, significant differences between the two lysates were observed. The lysate from the stimulated cells contained more phosphorylated proteins than the lysate from the unstimulated ones, especially in the 100–200 kDa range and 60–75 kDa range. This experiment suggested that, although not visible in either silver- or Coomassie-stained gels, differences between the unstimulated and stimulated cells may be detected. Based on these data, we also concluded that the differences between the unstimulated and stimulated cells observed by WB were as a result of ephrinB1-Fc stimulation, and that the signal transduction pathways were activated.

The lysates from both unstimulated and stimulated cells were also analyzed by BN-PAGE. BN-PAGE separates protein complexes based on the external charge induced by Coomassie Blue dye under native conditions and according to their mass. Therefore, under these experimental conditions, both stable and many transient protein interactions are preserved. In order to determine the differences between the lysates from unstimulated and stimulated cells that appeared as a result of eprinB1-Fc stimulation, we separated these lysates by BN-PAGE (1D) or (2D) (Figure 5). In the silver-stained BN-PAGE (1D) gel lanes, after visual inspection and based on the mass markers, we could observe intense bands that corresponded to protein complexes with a mass of about 750, 670, 600, 440 and 400 kDa (Figure 5A). In the silver stained (2D) gels (Figure 5B), some proteins with a mass of
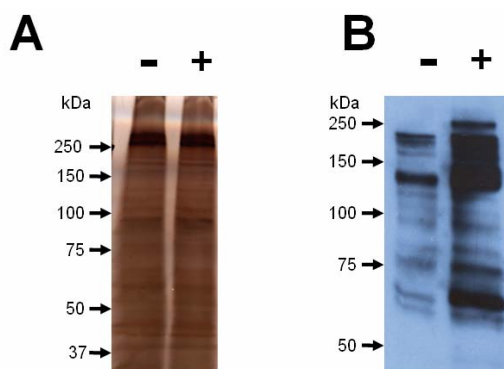
Figure 4. Analysis of lysates of unstimulated (–) and ephrin-stimulated (+) cells by SDS-PAGE. (A) The gel was silver stained (A, 1 µl loaded/lane) or electroblotted and incubated with pY99 antibodies (B, 5 µl loaded/lane). The mass markers are shown to the left of each gel

around 100–150 kDa, corresponded in (1D) gels (Figure 5A) to a mass between 150–700 kDa. This suggests that BN-PAGE separates in (1D) protein complexes that, upon reducing and denaturing are resolved into their subunits in (2D), with a mass smaller than the protein complex itself (assuming that the mass standards in the 1D [BN] gel are equivalent to the mass standards in the second [SDS-PAGE] dimension).

In order to identify differences between the unstimulated and stimulated cells as a result of ligand stimulation, we initially inspected the gel lanes that contained lysates from unstimulated or stimulated cells in either BN-PAGE (1D; Figure 5A) or 2D (Figure 5B) gels. No significant differences were observed. However, when the BN-PAGE (1D) gel lanes (data not shown) or (2D) gels (Figure 5C) were electroblotted and incubated with pY99 antibodies, significant differences between the lysates from unstimulated and stimulated cells were observed (Figure 5C). The lysate from the stimulated cells contained more phosphorylated proteins than the lysate from the unstimulated ones, in agreement with our previous data (Figure 4B) and current literature (Zhang et al., 2006), and further confirmed that the differences between the unstimulated and stimulated cells are as a result of ephrinB1-Fc stimulation. The most significant changes in the lysate from the stimulated cells in the 2D gels were observed between 100–200 kDa, 60–70 kDa, and 35–50 kDa (Figure 5C). The phosphorylated proteins from the 2D gels corresponded to a mass of approximately 100–250 kDa in the BN-PAGE (1D) under native conditions. This suggests that, under these experimental conditions (native, non-denaturing conditions), most of the tyrosine phosphorylated proteins (especially the ones with low molecular mass) observed in (2D) are part of different protein complexes in BN-PAGE (1D), with a mass ranging from 100–250 kDa (Figure 5C), again assuming the mass determinations based on standards in BN-PAGE are equivalent to
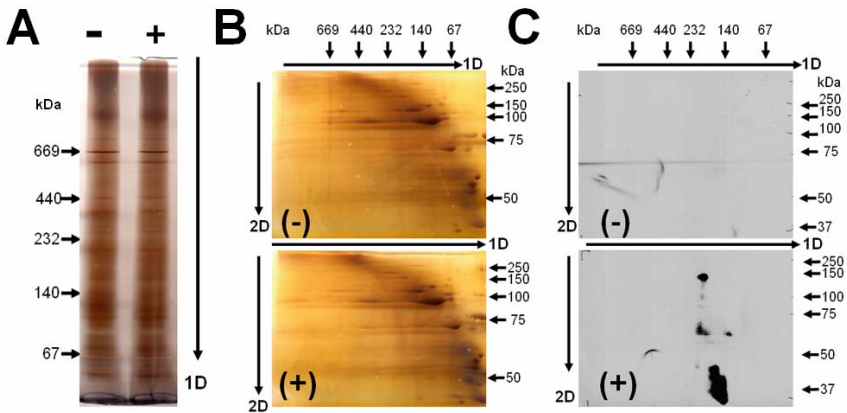
Figure 5. Analysis of the lysate of unstimulated (–) and ephrin-stimulated (+) cells by BN-PAGE. The lysates were separated in BN-PAGE (1D) (A) or BN-PAGE (1D) and SDS-PAGE (2D) (B, C) and either silver stained (A, B) or electroblotted and analyzed with pY99 antibodies. For (A), 10 µl were loaded/lane. For (B and C), 40 µl were loaded/lane in (1D). The mass markers are indicated for each gel

those in SDS-PAGE. Taken together, these data suggest that, upon stimulation, the signal transduction pathways are stimulated and the proteins involved in signaling may become phosphorylated and transiently associate in protein complexes.

3.2. ANALYSIS OF LYSATES FROM UNSTIMULATED AND BDNF-STIMULATED NEURONS BY BN-PAGE

To further confirm that BN-PAGE is suitable for analysis of protein-protein inter-actions as a result of ligand stimulation, we analyzed cell lysates of mouse cortical neurons that were stimulated by a different ligand, brain-derived neurotrophic factor (BDNF) that binds to Trk receptors. In this system, we focused on one particular protein, Hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs, mass 115 kDa). The Hrs protein is involved in endocytic processes and delivers receptors from endosomes to lysosomes where they are degraded as a result of ligand stimu-lation (Lu et al., 1998; Kobayashi et al., 2005). Therefore, we separated the cell lysates of neurons that were treated or not treated with BDNF by BN-PAGE (2D-WB) and incubated the blot membranes with HRS antibodies, followed by enhanced chemiluminescence detection. As may be seen in Figure 6, the Hrs anti-bodies recognized a band in the 2D blot at about 120 kDa, in agreement with the theoretical mass of Hrs. The Hrs antibody-reactive band corresponded in the native (1D) blot to a mass close to 200 kDa (unstimulated lysate) or 200 and 500 kDa (stimulated lysate). It is unclear if the Hrs proteins detected at about 200 kDa in (1D) are monomers or not.
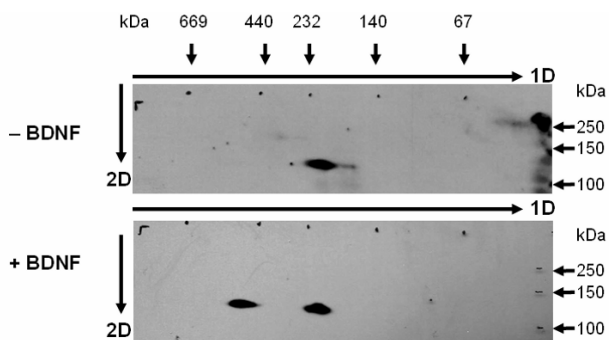
Figure 6. Analysis of lysates of unstimulated (–) and BDNF-stimulated (+) primary neuronal cell cultures by BN-PAGE. The lysates were separated in BN-PAGE (1D) and SDS-PAGE (2D), electroblotted, and visualized with Hrs antibodies. Forty μl of lysate were loaded/lane

Taken together, these experiments suggest that, upon stimulation, the Hrs protein assembles into higher mass complexes as a result of BDNF treatment. These data also suggest that, regardless of the cell system, as long as there are two different experimental conditions, they may be compared by BN-PAGE and the differences between the two states may be determined. Furthermore, the protein interactions that appear as a result of ligand stimulation (e.g., the protein(s) that interact(s) with Hrs as a result of BDNF stimulation) may be identified by MS (see below).

3.3. ANALYSIS OF LYSATES FROM UNSTIMULATED AND EPHRINB1-FC STIMULATED CELLS BY BN-PAGE AND MS

By visual inspection, in the silver-stained (Figure 5A) and Coomassie-stained (Figure 7) BN-PAGE (1D) gel lanes we could observe putative protein complexes with masses ranging from about 750 to about 400 kDa. In order to reveal the protein composition of some of these complexes, we cut out individual gel bands that corresponded to masses between 400 and 750 kDa (Figure 7), digested them with trypsin and analyzed them by LC-MS/MS. Once a subunit of a particular protein complex with a specified mass was identified, we looked for additional subunits in the same gel band and further compared these data with the current literature in terms of subunit composition, mass, protein interactions, growth factor-dependent inducibility, etc. In these experiments, we identified both homo- and hetero-complexes, most of them constitutively expressed in many mammalian cells. Among putative homocomplexes, we identified valosin-containing protein,, also named transitional endoplasmic reticulum ATPase, an 89 kDa endoplasmic reticulum membrane protein. This protein was found in bands B3 and B4, which corresponded
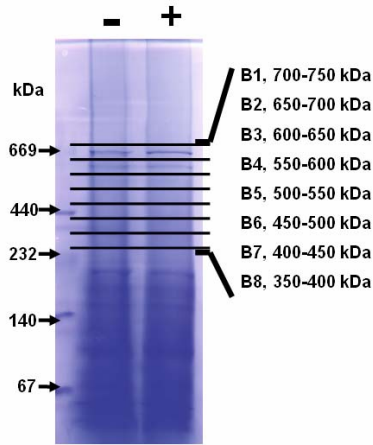
Figure 7. BN-PAGE of lysates of unstimulated (–) and ephrin-stimulated (+) NG108-EphB2 cells. The gel was Coomassie stained and the gel bands of interest (bands B1 to B8) were excised and subjected to MS analysis. The approximate masses of the bands were between 700–750 (band B1) and 350–400 kDa (band B8), as shown on the right side of the gel. The molecular mass markers are shown on the left side of the gel. Eighty microliters of lysate were loaded/lane

to masses of 550–650 kDa. A report that this 89 kDa protein is usually a homo-hexamer (DeLaBarre and Brunger, 2003), with predicted molecular mass (under native conditions) of 540 kDa, is in agreement with the mass determined experimentally by BN-PAGE.

The proteasome (prosome) was among the heterocomplexes identified by BN-PAGE and MS. This complex is involved in protein degradation and is composed of 28 alpha (mass 27 kDa) and beta (mass 29 kDa) subunits (McNaught et al., 2001), of which we identified five different subtypes of subunit alpha (alpha 1, 3, 4, 6 and 7 subunits) and five different subtypes of beta subunit (beta 1, 4, 5, 6 and 8 subunits) in a single experiment. Our experimentally determined mass of the proteasome was 650–750 kDa (detected in bands B1 and B2), in agreement with its calculated theoretical mass of 700 kDa. Examples of MSMS spectra of identified peptides that were part of either valosin-containing protein (Figure 8A) or proteasome subunits alpha 6 (Figure 8B) and beta 6 (Figure 8C) are shown in Figure 8. Taken together, these data suggest that the combination of BN-PAGE and MS is a powerful tool for determining the mass of a particular protein complex and identity of its subunits.
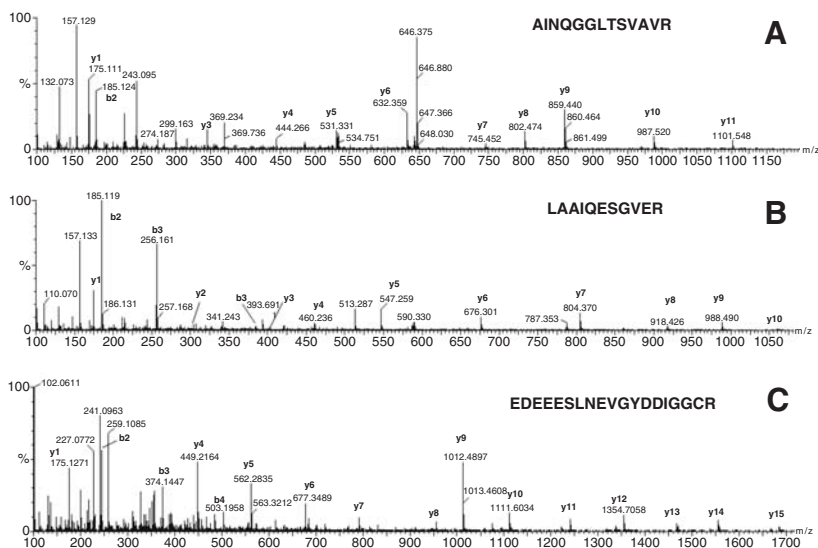
Figure 8. MS analysis of the BN-PAGE gel bands B1 (700–750 kDa) and B4 (550–600 kDa). (A) MS/MS of the *m/z* (2+) 643.33 (calculated 643.35) peak that corresponds to peptide AINQGGLTSVAVR, which is part of the proteosome alpha subunit 6 (mass 27 kDa), a 700 kDa alpha-beta heterocomplex. (B) MS/MS of the *m/z* (2+) 586.79 (calculated 586.81) peak that corresponds to peptide LAAIQESGVER, which is part of the proteosome beta subunit 6 (mass 29 kDa), a 700 kDa alpha-beta heterocomplex. (C) MS/MS of the *m/z* (2+) 1,093.49 (calculated 1,093.43) peak that corresponds to peptide EDEEESLNEVGYDDIGGCR, which is part of the valosin-containing protein (mass 89 kDa), which forms a 540 kDa homohexamer

3.4. ANALYSIS OF PY99-IPS FROM UNSTIMULATED AND EPHRINB1-FC STIMULATED CELLS BY GEL ELECTROPHORESIS AND MS

In the previous analyses of the lysates from unstimulated and stimulated cells using BN-PAGE, we could observe differences between the two states only at the WB level, with weak signals suggesting low amounts of protein. In order to identify the protein complexes that resulted upon stimulation, the cell lysates could not be used for MS analysis, due to the high complexity of the samples, large amounts of background proteins and low abundance of the proteins of interest. Therefore, we focused only on the proteins that were tyrosine phosphorylated or associated with phosphorylated proteins as a result of ephrinB1-Fc stimulation. Specifically, we performed pY99-IPs from the lysates of both unstimulated and stimulated cells. Initially we performed a small scale analysis of the pY99-IPs and separated them by either SDS-PAGE or BN-PAGE, followed by MS analysis. SDS-PAGE and MS were used to identify the differences between the proteins immunopurified by pY99 antibodies of the unstimulated and stimulated cells, while BN-PAGE and

MS were used to determine the interactions between the proteins identified by SDS-PAGE and MS that resulted upon ephrinB1-Fc stimulation.

Tables 1 and 2 show the proteins identified by SDS-PAGE and MS of pY99-IPs from unstimulated (Table 1) and stimulated (Table 2) cells. Very few proteins known to be involved in signal transduction were identified in the pY99-IPs of

TABLE 1. Proteins identified by MS analysis of the SDS-PAGE of the pY99-IP of unstimulated cells. The searching parameters were: database: NCBI_mouse_rat_FR, enzyme: trypsin, maximum missed cleavages: 1, fixed modifications: cysteine to carbamidomethyl, variable modifications: none, mass values: monoisotopic, peptide mass tolerance: 0.3 Da, fragment mass tolerance: 0.6 Da. The pkl files were recalibrated post-acquisition using in-house designed software. The NCBI accession number, protein name, Mascot score and molecular mass are shown for each identified protein

Protein Acc. #	Protein description	Score	Mw
gi|51921343	hypothetical protein LOC432987	1445	35144
gi|29244176	hypothetical protein 4732456N10	752	58230
gi|19705459	poly(A) binding protein, cytoplasmic 1	648	70656
gi|14861844	PL10 protein	507	73095
gi|13242237	heat shock protein 8	473	70827
gi|6671569	acidic ribosomal phosphoprotein P0	328	34195
gi|40068493	DEAD box polypeptide 17 isoform 1	301	72539
gi|42558248	GPI-anchored membrane protein 1	267	78121
gi|24429590	DEAH (Asp-Glu-Ala-His) box polypeptide 9	202	149489
gi|7949053	heterogeneous nuclear ribonucleoprotein A2/B1 isoform 1	198	35971
gi|31559916	heterogeneous nuclear ribonucleoprotein A3 isoform b	193	39628
gi|6754256	heat shock protein 9A	191	73483
gi|25742763	heat shock 70kD protein 5	177	72302
gi|23956214	splicing factor proline/glutamine rich	175	75394
gi|34328400	splicing factor, arginine/serine-rich 1 (ASF/SF2)	174	27728
gi|9506497	clathrin, heavy polypeptide (Hc)	168	191477
gi|6679741	PTK2 protein tyrosine kinase 2	129	119141
gi|10946928	heterogeneous nuclear ribonucleoprotein H1	110	49168
gi|6754222	heterogeneous nuclear ribonucleoprotein A/B	107	30812
gi|6754220	heterogeneous nuclear ribonucleoprotein A1 isoform a	101	34175
gi|7949051	heterogenous nuclear ribonucleoprotein U	94	87837
gi|20982845	pigpen	93	52642
gi|30794412	TAF15 RNA polymerase II, TATA box binding protein-associated	93	58566
gi|21703842	hypothetical protein LOC28088	87	55214
gi|31980689	small nuclear ribonucleoprotein polypeptide A	81	32245
gi|19527256	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	77	82448
gi|18087805	ribosomal protein S2	75	31212
gi|9789893	BRG1/brm-associated factor 53A	75	47399
gi|13385872	interleukin enhancer binding factor 2	74	43035
gi|7305247	P lysozyme structural	66	16783
gi|6671509	actin, beta, cytoplasmic	62	41710
gi|9790109	heterogeneous nuclear ribonucleoproteins methyltransferase-like 2	57	42408
gi|11024680	GERp95	57	97597
gi|13384620	heterogeneous nuclear ribonucleoprotein K	54	50944
gi|13507601	regulator of nonsense transcripts 1	51	122606
gi|31982724	MYB binding protein (P160) 1a	50	151942
gi|13242328	NS1-associated protein 1 isoform 1	50	69727
gi|34867525	PREDICTED: similar to T-complex protein 1 subunit theta	48	59550
gi|6754254	heat shock protein 1, alpha	47	84735
gi|33859724	heterogeneous nuclear ribonucleoprotein R	45	70844
gi|13592057	ribosomal protein L18	45	21645
gi|7305443	ribosomal protein L7a	42	29958
gi|51593084	SWI/SNF related, matrix associated, actin dependent regulator	38	180609
gi|6677805	ribosomal protein S4, X-linked	36	29579
gi|8567402	splicing factor, arginine/serine-rich 3 (SRp20)	36	19318

TABLE 2. Proteins identified by MS analysis of the SDS-PAGE gel of the pY99-IP of eph-rinB1-Fc stimulated cells. The searching parameters were: database: NCBI_Mouse_rat_FR, enzyme: trypsin, maximum missed cleavages: 1, fixed modifications: cysteine to carbamido-methyl, variable modifications: none, mass values: monoisotopic, peptide mass tolerance: 0.3 Da, fragment mass tolerance: 0.6 Da. The pkl files were recalibrated post-acquisition using in-house designed software. The NCBI accession number, protein name, Mascot score and molecular mass are shown for each identified protein

Protein Acc.#	Protein description	Score	Mw
gi|47777351	Eph receptor B2	2306	111311
gi|51921343	hypothetical protein LOC432987	1120	35315
gi|13242237	heat shock protein 8	1111	71055
gi|6671569	acidic ribosomal phosphoprotein P0	667	34366
gi|6754256	heat shock protein 9A	521	73768
gi|25742763	heat shock 70kD protein 5	492	72473
gi|19705459	poly(A) binding protein, cytoplasmic 1	479	70884
gi|47059093	Eph receptor B4	354	110370
gi|9506497	clathrin, heavy polypeptide (Hc)	277	193187
gi|29244176	hypothetical protein 4732456N10	233	58629
gi|40068491	DEAD box polypeptide 17 isoform 2	223	46905
gi|6671509	actin, beta, cytoplasmic	206	42052
gi|42558248	GPI-anchored membrane protein 1	169	78292
gi|6679741	PTK2 protein tyrosine kinase 2	167	119939
gi|9790069	HLA-B-associated transcript 1A	166	49460
gi|9789893	BRG1/brm-associated factor 53A	160	47913
gi|6753620	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3, X-linked	155	73455
gi|9790109	heterogeneous nuclear ribonucleoproteins methyltransferase-like 2	155	43035
gi|9506371	actin, alpha 1, skeletal muscle	136	42366
gi|14192922	actin, alpha, cardiac	136	42334
gi|6754222	heterogeneous nuclear ribonucleoprotein A/B	131	30926
gi|7949051	heterogenous nuclear ribonucleoprotein U	131	88635
gi|34867525	PREDICTED: similar to T-complex protein 1 subunit theta	129	60121
gi|31560613	chaperonin subunit 8 (theta)	118	60087
gi|31980689	small nuclear ribonucleoprotein polypeptide A	114	32245
gi|9506571	eukaryotic translation initiation factor 2, subunit 1 alpha	110	36371
gi|8393418	glyceraldehyde-3-phosphate dehydrogenase	95	36090
gi|21450625	eukaryotic translation initiation factor 4A1	92	46353
gi|6677813	ribosomal protein S8	92	24475
gi|7305443	ribosomal protein L7a	91	30129
gi|30794450	ribosomal protein L4	90	47409
gi|6755682	serine/threonine kinase receptor associated protein	90	38774
gi|7949053	heterogeneous nuclear ribonucleoprotein A2/B1 isoform 1	89	36028
gi|22122585	splicing factor, arginine/serine-rich 7	82	27589
gi|12018306	nucleolin-related protein	82	77450
gi|27687455	PREDICTED: similar to cytoplasmic beta-actin	75	42278
gi|6981248	nucleolin	74	77158
gi|6679939	glyceraldehyde-3-phosphate dehydrogenase, spermatogenic	72	48096
gi|6678469	tubulin, alpha 6	70	50562
gi|6754254	heat shock protein 1, alpha	68	85134
gi|13385998	TNF receptor-associated protein 1	68	80501
gi|46275814	neurofilament, heavy polypeptide	66	115498
gi|6679108	nucleophosmin 1	63	32711
gi|6755110	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3	62	85439
gi|8393693	laminin receptor 1	60	32917
gi|9507245	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, gamma	60	28456
gi|24429590	DEAH (Asp-Glu-Ala-His) box polypeptide 9	56	150914
gi|37674216	RIKEN cDNA G430041M01	54	32354
gi|6753324	chaperonin subunit 6a (zeta)	53	58424
gi|6756041	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta	51	27925
gi|31543689	arginine/serine-rich splicing factor 6	50	39120
gi|31543313	non-catalytic region of tyrosine kinase adaptor protein 2	50	43008
gi|8394018	protein phosphatase 2a, catalytic subunit, alpha isoform	49	36156
gi|6678053	small nuclear ribonucleoprotein B	45	23811
gi|6680047	guanine nucleotide binding protein (G protein), beta polypeptide 2 like 1	39	35511
gi|9845257	histone 1, H1c	38	21254
gi|40254593	breast cancer anti-estrogen resistance 1	37	94455

unstimulated (Table 1) compared with the stimulated (Table 2) cells. Since we analyzed pY99-IP samples, these data suggest that many of the proteins from the unstimulated cells involved in ephrinB1 signaling are not phosphorylated at tyrosine residues and not associated with a phosphotyrosine protein (Table 1). On the other hand, most of the pY99-immunopurified proteins from the stimulated cells involved in ephrinB1 signaling are either tyrosine phosphorylated, or associated with a phosphotyrosine protein, or both (Table 2). These data are in agreement with our previous experiments, in which WB with pY99 antibodies of the lysate separated by both SDS-PAGE (Figure 4B) and BN-PAGE (Figure 5C) showed a higher degree of tyrosine phosphorylation of the proteins from the stimulated cells. These data are also in agreement with a study from our laboratory, in which SILAC technology was employed (Zhang et al., 2006). Perhaps one of the most important proteins identified in the SDS-PAGE of pY99-IP of unstimulated cells is PTK2 (protein tyrosine kinase 2), a 130 kDa protein, also named focal adhesion kinase (FAK) (Table 1; see below). On the other hand, among the proteins identified in the SDS-PAGE of pY99-IP of stimulated cells and known to be involved in ephrin signaling (Zhang et al., 2006), two important proteins were identified: PTK2 (protein tyrosine kinase 2 or FAK, 130 kDa) and breast cancer anti-estrogen resistance 1, a 95 kDa protein, also named p130Cas. These two proteins are known to interact with each other via a SH3 domain (Polte and Hanks, 1995) and their identification in the pY99-IP from the stimulated cells (but only of FAK in the unstimulated ones) may suggest that FAK and p130Cas interact with each other as a result of ephrin stimulation. Since FAK was also identified in the pY99-IP of unstimulated cells, it is possible that the phosphorylated FAK interacts with p130Cas (and phosphorylates it) as a result of ephrin stimulation. Alternatively, FAK may not be phosphorylated in the pY99-IP of stimulated cells, but rather associated with phosphorylated p130Cas or with other unknown phosphorylated proteins.

In order to evaluate these hypotheses, we separated the pY99-IPs from both unstimulated and stimulated cells by BN-PAGE (1D), cut out the gel bands as shown in Figure 9 (left) and analyzed them by MS (Darie and Neubert, 2008). Not surprisingly, we identified FAK in the gel bands from unstimulated (estimated masses in the 130–170 kDa and 190–210 kDa range) and stimulated (estimated masses in the 150–170 kDa and 230–250 kDa range) cells (Figure 9, right). Since the mass of FAK is 130 kDa, this suggests that the protein identified at about 130–170 kDa is in a monomeric state, while the 190–210 kDa (from unstimulated cells) and 230–250 kDa (from stimulated cells) bands contain FAK in protein complexes. We also identified p130Cas (95 kDa) in the gel band from stimulated (estimated mass 230–300 kDa) but not unstimulated cells (Figure 9, right), in agreement with the previous SDS-PAGE and MS experiments (Tables 1 and 2). Since (1) FAK shifted from the monomeric state in unstimulated cells towards a higher mass in the stimulated cells upon stimulation; (2) p130Cas was identified in both SDS-PAGE & MS and BN-PAGE & MS as a result of ephrin stimulation; (3) it is well established that FAK and p130Cas interact with each other; and (4) FAK

(130 kDa) and p130Cas (95 kDa) were identified in the same BN-PAGE gel band with an approximate mass of 230–300 kDa, it is reasonable to conclude that FAK and p130Cas form a heterocomplex as a result of ephrin stimulation. It remains to be determined which one of these proteins (FAK, p130Cas, neither or both) are phosphorylated in the FAK-p130Cas complex. Taken together, these data suggest that SDS-PAGE, MS and BN-PAGE and MS of the pY-IPs from both unstimulated and stimulated cells are a useful system for analysis of both stable and transient protein-protein interactions.
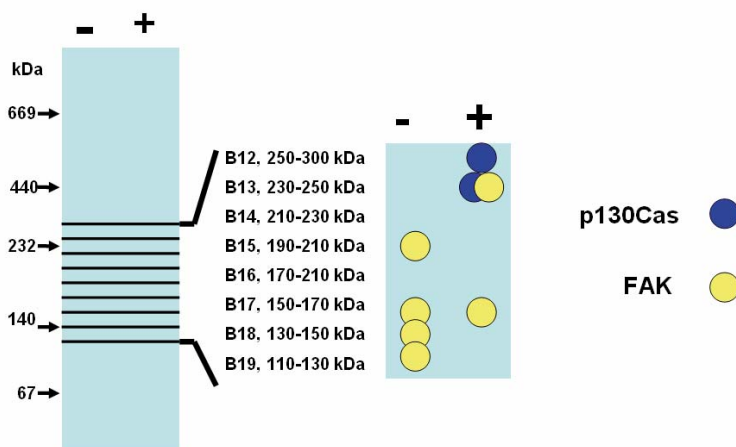
Figure 9. Representation of BN-PAGE of the pY99-IPs of unstimulated (–) and ephrin-stimulated (+) cells. Left: The gel was Coomassie stained and the gel bands of interest (bands B12 to B19) were excised and analyzed by MS. The approximate mass of the bands was between 250–300 (band B12) and 110–130 kDa thousand (band B19), as shown on the right side of the gel. The molecular mass markers are shown on the left side of the gel. Right: The approximate position in the BN-PAGE gel (mass) of the FAK and p130 Cas proteins, as determined by MS experiments. The legend of the FAK and p130 Cas proteins is also shown (right)

4.   Discussion

Upon activation of signal transduction pathways, many transient protein-protein interactions occur, leading to the formation of different protein complexes. These complexes form during the transition of the cells from the unstimulated to stimulated state. Analysis of signal transduction pathways generally starts with the identification of proteins that interact with the activated receptors. Once these proteins are initially identified, further interacting partners may be identified and the function of these proteins may be studied. A very common method for studying protein-protein interactions is immunoprecipitation by one antibody followed by WB of the immunoprecipitate with a different antibody. Unfortunately, this method may be employed only if preliminary data already exists and candidate proteins

are identified. For example, if the interaction of two proteins is already established by a different method such as two-hybrid screen, immunoprecipitation and WB experiments may be confirmative. Therefore, analysis of protein-protein interactions has been limited to immunopurifications or two hybrid screens (Ito et al., 2000; Uetz et al., 2000; Gavin et al., 2002; Ho et al., 2002) that led to construction of interaction networks (Schwikowski et al., 2000; von Mering et al., 2002).

Recent advances in mass spectrometry have allowed scientists to identify by proteomic approaches not only one but many or most potential interacting partners of a particular protein involved in a signal transduction pathway by analyzing its immunoprecipitate by SDS-PAGE and MS (Steen et al., 2002; Wang et al., 2006). Other methods such as isotope-coded affinity tag (ICAT) (Gygi et al., 1999) and stable isotope labeling by amino acids in cell culture (SILAC) (Ong et al., 2002; Blagoev et al., 2003; Ong et al., 2003) have allowed detection and quantification of the phosphorylation levels of proteins involved in signal transduction from unstimulated and stimulated cells (Zhang et al., 2006). These and similar approaches are known as functional proteomics.

Although the above mentioned approaches have advanced studies of signal transduction pathways, many questions still remain unanswered due to the technical limitations of the methodology. For example, combining SDS-PAGE (1D or 2D) with MS–although a very powerful method–is limited at the gel electrophoresis level. High or low mass proteins, very hydrophobic proteins, and proteins with high or low pI may not be identified via 2D PAGE. In addition, in experiments with unstimulated and stimulated cells, when a particular protein is identified as a participant in signal transduction pathways, no functional assignment of the protein is made unless its post-translational modifications such as phosphorylation are dramatically increased or decreased. Furthermore, even when the interacting partners of the proteins involved in signaling are identified and characterized, the functional significance of the association of these proteins into multi-subunit protein complexes via ligand stimulation is still missed. Therefore, any optimization of the current technology would be a great improvement. Analysis of protein complexes from signal transduction pathways in terms of size, composition, post-translation modifications, multimerization level, and abundance would also benefit from the introduction of new methods.

Blue Native PAGE (BN-PAGE) has long been used to analyze protein-protein interactions and protein complexes (Schagger and von Jagow, 1991; Schagger et al., 1994; Darie et al., 2005; Darie et al., 2007; Litscher et al., 2008). This method separates protein complexes based on external charge induced by Coomassie dye and according to their mass. BN-PAGE experiments may provide information about the size, number, subunit composition, stoichiometry and relative abundance of these protein complexes. Compared with previous methods used to study protein-protein interactions, there are several advantages of BN-PAGE: (1) separation of the protein complexes takes place under native conditions, so even the transient interactions between proteins may be identified, (2) separation of the protein complexes in the second dimension may reveal their subunit composition, (3) the

method may analyze the association of proteins into protein complexes as a result of ligand stimulation, independent of their levels of phosphorylation, (4) BN-PAGE may confirm the results obtained by immunopurification experiments. For example, if proteins A, B and C co-immunopurify, BN-PAGE will often distinguish between A-B-C, A-B, A-C, and B-C complexes. (5) By combining BN-PAGE with MS, both structural and functional information may be obtained (Camacho-Carvajal et al., 2004; Darie et al., 2005; Aivaliotis et al., 2006; Reifschneider et al., 2006).

In this paper, we analyzed proteins and protein complexes involved in EphB2 signal transduction pathways as a result of in ephrinB1-Fc stimulation by using SDS-PAGE, BN-PAGE, WB and MS. These experiments allowed us to conclude that the differences between the unstimulated and stimulated cells in terms of protein tyrosine phosphorylation are small and detectable only by WB experiments. These experiments also allowed us to conclude that, upon stimulation, the proteins involved in ephrin signaling associate into protein complexes that may be studied by BN-PAGE and MS. In addition, we concluded that BN-PAGE and MS analysis are suitable for separation of the protein complexes by BN-PAGE, which preserves their integrity, regardless of whether it is a homo- or a hetero-complex. Furthermore, combination of BN-PAGE and MS to study protein complexes from unstimulated and ephrin-stimulated cells also has the potential to be applied to other ligand-stimulated RTK-dependent signal transduction pathways.

The first consequences of EphB2 receptor activation are dimerization and multimerization of the receptor, and autophosphorylation of its tyrosine residues from the cytoplasmic side of the receptor (Vearing and Lackmann, 2005). The phosphorylated sites provide docking sites for a number of SH2-domain containing signaling and adaptor proteins that modulate cytoskeletal plasticity and a number of proteins that interact directly or indirectly with the phosphorylated EphRs or mediate downstream signaling have been identified (Vearing and Lackmann, 2005). Recently, a study of EphB2 signaling from unstimulated and ephrinB1-stimulated cells has been performed in our lab using SILAC technology (Zhang et al., 2006). Our group not only confirmed that some known proteins are components of the EphB2 signaling, but also identified previously unknown members that participate in the EphB2-dependent pathways (Zhang et al., 2006).

In our SDS-PAGE and MS experiments (Tables 1 and 2) we identified a number of proteins that became phosphorylated or associated with phosphorylated proteins as a result of ephrin stimulation, some of which are well known to be involved in ephrin signaling (Zhang et al., 2006). Eph receptors (B2 & B4), PTK2 protein tyrosine kinase 2/FAK, serine/threonine kinase receptor associated protein, tyrosine 3-monooxgenase/tryptophan 5-monooxgenase activation proteins (gamma & zeta polypeptides), non-catalytic region of tyrosine kinase adaptor protein 2, protein phosphatase 2a, G protein and breast cancer anti-estrogen resistance 1/p130Cas are some of the proteins identified by SDS-PAGE and MS analysis of pY99-IP of stimulated cells. Two of these proteins were of great interest to us: PTK2 protein tyrosine kinase 2 (FAK; mass 130 kDa) and breast cancer anti-estrogen resistance 1 (p130Cas; mass 95 kDa). By using BN-PAGE and MS, we were able to

conclude that these two proteins interact with each other as a result of ephrin stimulation, in agreement with the published literature (Polte and Hanks, 1995). Because we analyzed pY99-IP samples, it is likely that at least one of these proteins is phosphorylated on tyrosine. Therefore, the combination of BN-PAGE and MS is also suitable for analysis of transient protein-protein interactions.

In conclusion, we used BN-PAGE and MS to analyze protein-protein interactions from signal transduction pathways that resulted upon ephrinB1-Fc stimulation. These experiments are a starting basis for further analysis of transient protein-protein interactions in ligand-stimulated cells.

Acknowledgments

References

Adams, R. H. (2003). "Molecular control of arterial-venous blood vessel identity." *J Anat* **202**(1): 105–12.

Aebersold, R. and M. Mann (2003). "Mass spectrometry-based proteomics." *Nature* **422**(6928): 198–207.

Aivaliotis, M., M. Karas, et al. (2006). "High throughput two-dimensional blue-native electrophoresis: a tool for functional proteomics of cytoplasmatic protein complexes from Chlorobium tepidum." *Photosynth Res* **88**(2): 143–57.

Blagoev, B., I. Kratchmarova, et al. (2003). "A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling." *Nat Biotechnol* **21**(3): 315–8.

Bruckner, K., E. B. Pasquale, et al. (1997). "Tyrosine phosphorylation of transmembrane ligands for Eph receptors." *Science* **275**(5306): 1640–3.

Camacho-Carvajal, M. M., B. Wollscheid, et al. (2004). "Two-dimensional Blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach." *Mol Cell Proteomics* **3**(2): 176–82.

Carter, N., T. Nakamoto, et al. (2002). "EphrinA1-induced cytoskeletal re-organization requires FAK and p130(cas)." *Nat Cell Biol* **4**(8): 565–73.

Cheng, N., D. M. Brantley, et al. (2002). "The ephrins and Eph receptors in angiogenesis." *Cytokine Growth Factor Rev* **13**(1): 75–85.

Cowan, C. A. and M. Henkemeyer (2001). "The SH2/SH3 adaptor Grb4 transduces B-ephrin reverse signals." *Nature* **413**(6852): 174–9.

Dalva, M. B., M. A. Takasu, et al. (2000). "EphB receptors interact with NMDA receptors and regulate excitatory synapse formation." *Cell* **103**(6): 945–56.

Darie, C. C., M. L. Biniossek, et al. (2005). "Isolation and structural characterization of the Ndh complex from mesophyll and bundle sheath chloroplasts of Zea mays." *Febs J* **272**(11): 2705–16.

Darie, C. C., W. G. Janssen, et al. (2007). "Purified trout egg vitelline envelope proteins VEbeta and VEgamma polymerize into homomeric fibrils from dimers in vitro." *Biochim Biophys Acta* doi:10.1016/j.bbapap.2007.10.011

Darie, C. C. and T. A. Neubert (2008). "Analysis of protein-protein interactions in EphB2-NG108 cells as a result of ephrinB1 stimulation." *In preparation*.

DeLaBarre, B. and A. T. Brunger (2003). "Complete structure of p97/valosin-containing protein reveals communication between nucleotide domains." *Nat Struct Biol* **10**(10): 856–63.

Gale, N. W., S. J. Holland, et al. (1996). "Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis." *Neuron* **17**(1): 9–19.

Gavin, A. C., M. Bosche, et al. (2002). "Functional organization of the yeast proteome by systematic analysis of protein complexes." *Nature* **415**(6868): 141–7.

Gygi, S. P., B. Rist, et al. (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." *Nat Biotechnol* **17**(10): 994–9.

Ho, Y., A. Gruhler, et al. (2002). "Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry." *Nature* **415**(6868): 180–3.

Holland, S. J., N. W. Gale, et al. (1996). "Bidirectional signalling through the EPH-family receptor Nuk and its transmembrane ligands." *Nature* **383**(6602): 722–5.

Holland, S. J., E. Peles, et al. (1998). "Cell-contact-dependent signalling in axon growth and guidance: Eph receptor tyrosine kinases and receptor protein tyrosine phosphatase beta." *Curr Opin Neurobiol* **8**(1): 117–27.

Ito, T., K. Tashiro, et al. (2000). "Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins." *Proc Natl Acad Sci U S A* **97**(3): 1143–7.

Kobayashi, H., N. Tanaka, et al. (2005). "Hrs, a mammalian master molecule in vesicular transport and protein sorting, suppresses the degradation of ESCRT proteins signal transducing adaptor molecule 1 and 2." *J Biol Chem* **280**(11): 10468–77.

Litscher, E. S., W. G. Janssen, et al. (2008). "Purified mouse egg zona pellucida glycoproteins polymerize into homomeric fibrils under non-denaturing conditions." *J Cell Physiol* **214**(1): 153–7.

Lu, L., M. Komada, et al. (1998). "Human Hrs, a tyrosine kinase substrate in growth factor-stimulated cells: cDNA cloning and mapping of the gene to chromosome 17." *Gene* **213**(1-2): 125–32.

McNaught, K. S., C. W. Olanow, et al. (2001). "Failure of the ubiquitin-proteasome system in Parkinson's disease." *Nat Rev Neurosci* **2**(8): 589–94.

Miao, H., E. Burnett, et al. (2000). "Activation of EphA2 kinase suppresses integrin function and causes focal-adhesion-kinase dephosphorylation." *Nat Cell Biol* **2**(2): 62–9.

Ong, S. E., B. Blagoev, et al. (2002). "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics." *Mol Cell Proteomics* **1**(5): 376–86.

Ong, S. E., L. J. Foster, et al. (2003). "Mass spectrometric-based approaches in quantitative proteomics." *Methods* **29**(2): 124–30.

Polte, T. R. and S. K. Hanks (1995). "Interaction between focal adhesion kinase and Crk-associated tyrosine kinase substrate p130Cas." *Proc Natl Acad Sci U S A* **92**(23): 10678–82.

Reifschneider, N. H., S. Goto, et al. (2006). "Defining the mitochondrial proteomes from five rat organs in a physiologically significant context using 2D blue-native/SDS-PAGE." *J Proteome Res* **5**(5): 1117–32.

Robinson, V., A. Smith, et al. (1997). "Roles of Eph receptors and ephrins in neural crest pathfinding." *Cell Tissue Res* **290**(2): 265–74.

Schagger, H., W. A. Cramer, et al. (1994). "Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis." *Anal Biochem* **217**(2): 220–30.

Schagger, H. and G. von Jagow (1991). "Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form." *Anal Biochem* **199**(2): 223–31.

Schwikowski, B., P. Uetz, et al. (2000). "A network of protein-protein interactions in yeast." *Nat Biotechnol* **18**(12): 1257–61.

Steen, H., B. Kuster, et al. (2002). "Tyrosine phosphorylation mapping of the epidermal growth factor receptor signaling pathway." *J Biol Chem* **277**(2): 1031–9.

Uetz, P., L. Giot, et al. (2000). "A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae." *Nature* **403**(6770): 623–7.

Vearing, C. J. and M. Lackmann (2005). "Eph receptor signalling; dimerisation just isn't enough." *Growth Factors* **23**(1): 67–76.

von Mering, C., R. Krause, et al. (2002). "Comparative assessment of large-scale data sets of protein-protein interactions." *Nature* **417**(6887): 399–403.

Wang, Y., R. Li, et al. (2006). "Proteomic analysis reveals novel molecules involved in insulin signaling pathway." *J Proteome Res* **5**(4): 846–55.

Wilkinson, D. G. (2001). "Multiple roles of EPH receptors and ephrins in neural development." *Nat Rev Neurosci* **2**(3): 155–64.

Yancopoulos, G. D., M. Klagsbrun, et al. (1998). "Vasculogenesis, angiogenesis, and growth factors: ephrins enter the fray at the border." *Cell* **93**(5): 661–4.

Zhang, G., D. S. Spellman, et al. (2006). "Quantitative phosphotyrosine proteomics of EphB2 signaling by stable isotope labeling with amino acids in cell culture (SILAC)." *J Proteome Res* **5**(3): 581–8.

2. STRUCTURE, PROCESSING, AND POLYMERIZATION OF RAINBOW TROUT EGG VITELLINE ENVELOPE PROTEINS

COSTEL C. DARIE[1,2*], EVELINE S. LITSCHER[1], AND PAUL M. WASSARMAN[1*]

[1]*Department of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029-6574*

[2]*Present address: Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University, 540 First Avenue, New York, NY 10016*

Abstract. Mammalian and non-mammalian eggs are surrounded by a zona pellucida (ZP) and vitelline envelope (VE), respectively. The rainbow trout egg VE consists of three proteins, VEα, VEβ, and VEγ, which are related to mouse egg ZP proteins ZP1, ZP2, and ZP3. Mass spectrometry (MS) has been used extensively to identify the intramolecular disulfide linkages and the cellular site of proteolytic processing of trout VE proteins. Additionally, Blue Native-PAGE (BN-PAGE) has been used to investigate polymerization of purified trout VE proteins under non-denaturing conditions. Results of these experiments reveal that, despite ~400 million years separating the appearance of trout and mice, and the change from external to internal fertilization and development, VE and ZP proteins have a great deal in common.

1. Introduction

The plasma membrane of all vertebrate eggs is surrounded by an extracellular coat (Dumont and Brummett, 1985). Those surrounding fish and mammalian eggs are often referred to as the vitelline envelope (VE) and zona pellucida (ZP), respectively. These coats function to restrict to different degrees fertilization of eggs to sperm from the same species, to insure fertilization by only a single sperm, and to protect developing embryos either outside or within the female reproductive tract. Fish eggs are surrounded by a VE and have a funnel-shaped micropyle at the animal pole region which attracts sperm to this region and limits the number of sperm entering and attaching to the egg plasma membrane (Iwamatsu et al., 1997). Following fertilization by a single sperm, fish eggs form a fertilization cone to plug the micropyle and the VE, like the mammalian egg ZP, is hardened as part of the block to polyspermy (Shibata et al., 2000).

---

[*]Corresponding authors. E-mail: darie@saturn.med.nyu.edu; paul.wassarman@mssm.edu

Extensive molecular analyses of egg coat proteins from many mammalian and non-mammalian animal species have revealed that the proteins are highly conserved and related to each other (Litscher and Wassarman, 2007). A feature common to these proteins is the presence of a so-called ZP domain, a sequence of ~260 amino acids containing eight conserved cysteine (C) residues (Jovine et al., 2005). The ZP domain functions as a polymerization module enabling VE and ZP proteins to form the extremely long fibrils that constitute the extracellular coats (Jovine et al., 2002, 2005, 2006).

Rainbow trout egg VEs consist of at least three proteins, called VEα, VEβ, and VEγ (Brivio et al., 1991; Hyllner et al., 2001). All three proteins have an N-terminal signal sequence, a proline-glutamine- (PQ-) rich region, a ZP domain, and a C-terminal propeptide. VEα and VEβ also have a trefoil (P) domain, just upstream of the ZP domain. The proteins are synthesized by the liver under hormonal control and are transported in the bloodstream to the ovary where they assemble into a VE around eggs (Hyllner and Haux, 1992). The trout egg VE is ~50 µm thick, is composed of a wave-shaped fibrillary component embedded in an amorphous matrix, and has a very thin outer layer and a much thicker inner layer (Hagen-maier, 1973; Brivio et al., 1991; Hyllner and Haux, 1992).

Here, we present some of our recent findings that bear on the structure, processing, and polymerization of rainbow trout egg VE proteins. Application of a variety of experimental approaches, including mass spectrometry (MS), Blue Native-polyacrylamide gel-electrophoresis (BN-PAGE), and transmission electron microscopy (TEM), to purified VE proteins has revealed a great deal about the arrangement of their disulfide linkages, the cellular site of proteolytic processing, and the propensity of the proteins to polymerize into long fibrils (Darie et al., 2004, 2005, 2007).

2.   Results

2.1. RAINBOW TROUT VE PROTEINS

Eggs were obtained from rainbow trout (*Oncorhynchus mykiss*) by abdominal squeezing of females in a bloodless procedure that avoided contamination of eggs with proteins from blood. Harvested eggs (~3 mm in diameter) were frozen immediately by immersion in an ethanol and dry ice mixture. The procedure described by Brivio et al. (Brivio et al., 1991) was then used to isolate VEs from trout eggs. This involved solubilization of VE proteins in guanidine hydrochloride, dialysis against urea, and preparative SDS-PAGE.

Under non-reducing conditions on SDS-PAGE, rainbow trout VEs contain four major proteins, three of which are called VEα (M_ρ ~52 kDa), VEβ (M_ρ ~48kDa), and VEγ (M_r ~44 kDa) (Figure 1). In addition, a fourth protein (M_r ~110 kDa) is present that is referred to as HMWP (high molecular weight protein). Based on immunoblots and MS analysis, HMWP is a mixture of VEα-VEγ and VEβ-VEγ

heterodimers (Darie et al., 2004) whose N-terminal extensions, rich in proline and glutamine residues, are probably covalently crosslinked by transglutaminase. As expected, and consistent with the idea that their C residues are only involved in intramolecular disulfides, under reducing conditions on SDS-PAGE, the M_rs of these four proteins shift to ~58 (VEα), ~54 (VEβ), ~47 (VEγ), and ~120–130 (HMWP) kDa (Figure 1). VEβ, VEγ, and HMWP are stained by Coomassie blue to almost the same intensity, whereas VEα is stained less intensely, perhaps suggesting a less than stoichiometric ratio of VE proteins.
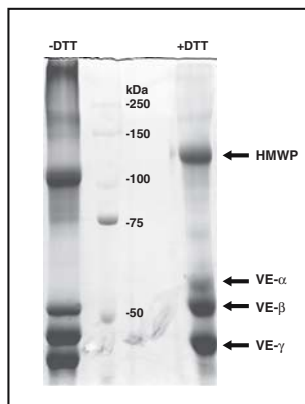
Figure 1. Coomassie-stained VE proteins after 10% SDS-PAGE under non-reducing (-DTT) and reducing (+DTT) conditions. Four main bands are resolved: VEα, β, γ, and HMWP, with a higher electrophoretic mobility under non-reducing conditions. The position of each VE protein and M_ρ standards are shown

2.2.   INTRAMOLECULAR DISULFIDES OF VE PROTEINS

VE proteins contain a large number of C residues (VEα, 18; VEβ, 18; VEγ, 12), with the first six C residues of VEα and VEβ present as disulfides (C1-4, C2-5, C3-6) in trefoil domains (Carr et al., 1994). The ZP domain contains eight C residues that are conserved in all ZP proteins and these residues are thought to be present as four intramolecular disulfides. The C residues of VE proteins were designated C1 to C8 after those conserved in the ZP domain of ZP proteins. The ZP domain of ZP1- or ZP2-like proteins (i.e., ZP1, ZP2, VEα, and VEβ) contain, in addition to the eight conserved C residues found in ZP3-like proteins (i.e., ZP3 and VEγ), two extra C residues that are conserved from fish to mammals and are designated Ca and Cb. In addition, within the ZP domain of ZP1-like proteins (VEα and VEβ) from fish, there are two C residues that have not been found in mammals; these are designated Cx and Cy. VEγ also has four additional C residues, designated C9 to C12, downstream of residue C8 of the ZP domain.

To determine the remaining disulfide linkages, MS (MALDI-TOF-MS; ESI-Q-TOF-MS) and gel-electrophoresis of chemically and enzymatically digested individual VE proteins were used (Darie et al., 2004). Each VE protein was analyzed individually by MALDI-TOF-MS in linear and reflective modes and identified by using Mascot peptide map fingerprinting (http://www.matrixscience. com). Sequence coverage was improved by using FindPept and FindMod (http://www.expasy.ch) programs.

Figure 2. Disulfide linkage determination within VEβ by ESI-Q-TOF-MS. The peaks represent the monoisotopic m/z. (A) The m/z 992.55 ion represents the double charged state of the intramolecular disulfide-linked peptide shown in Figure 2B. The m/z 855.20 ion represents the triple charged state of the intermolecularly linked peptides shown in Figure 2C. The m/z 759.78 ion represents the triple charged state of the intramolecular disulfide-linked peptide [468]IHCNTAVCLPSLGDSCEPRCY[488]. (B) CID spectrum of the 2+ charged ion m/z 992.55. The precursor ion of m/z 1,984.10 and the resulting fragments correspond to peptide [338]RLGSGGCLTKGCNEEEVAY[356] minus 2 Da, due to the intramolecular linkage. (C) CID spectrum of the 3+ charged ion m/z 855.20. The precursor ion of m/z 2,563.60 and the fragment ions correspond to the peptides [257]KVTECGTVM[265] and [236]GTNGAHCHPIGTTSVF[251] disulfide-linked

MALDI-TOF-MS was used initially to assign the disulfide bridges of VE proteins. Several criteria were imposed to assure that bridges were assigned accurately: (1) The two C residues that constitute a disulfide bridge cannot participate in other bridges. For example, if a protein contains four disulfide-linked C residues and a bridge is detected between C-3 and C-4, these residues cannot form disulfide bridges with C-1 and/or C-2. (2) Every peak that corresponds to a disulfide bridge detected by MALDI-MS under non-reducing conditions must disappear under reducing conditions. (3) At least one of the C-containing peptides participating in a disulfide bridge must be detected under reducing conditions (Note: Some peptides less than ~800 Da M_r were below the detection limits of the mass spectrometer). (4) A disulfide bridge detected in an enzymatic digest by MALDI-MS must be confirmed by detection in a digest in which a different protease was used. (5) Disulfide bridges detected by MALDI-MS must be confirmed by using a different means of generating peptides (e.g., chemical digest).
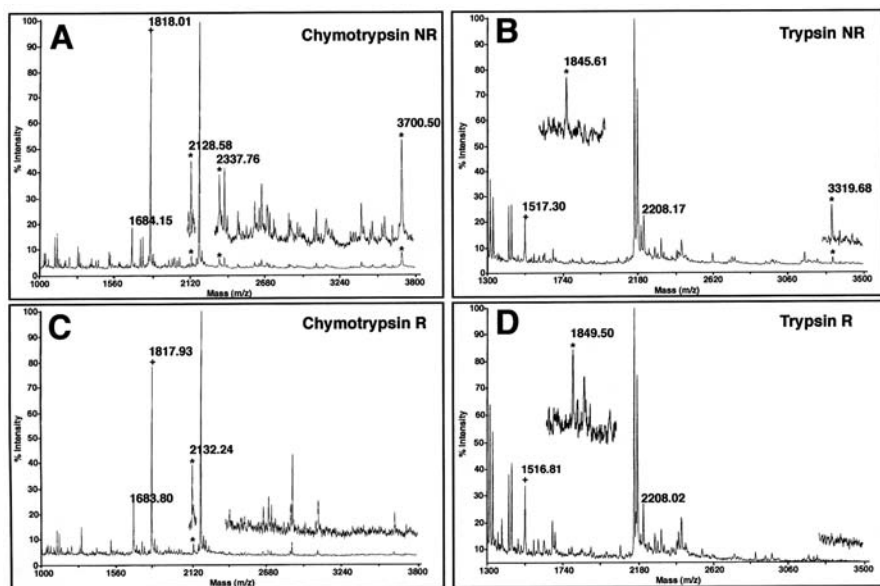
Figure 3. Assignment of the disulfide linkages of VEγ by MALDI-TOF-MS (linear mode). The protein was digested with either chymotrypsin (A, C) or trypsin (B, D) and measured under NR (A, B) and R (C, D) conditions. Disulfides are marked with (*) and the C residue-containing peptides that resulted upon reduction or as a laser-induced homolytic cleavage of the disulfides are marked with (+). The intramolecular disulfides (peaks m/z 2,128.58 and m/z 1,845.61) gained 4 Da upon reduction. The peaks of m/z 1,845.61 and m/z 1,849.50 (Figure 6B, D, inset) are from a different experiment. The chymotrypsin digest peaks of m/z 1,684.15 (Figure 6A), m/z 1,683.80 (Figure 6C) and the trypsin digest peaks of m/z 2,208.17 (Figure 6B) and m/z 2,208.02 (Figure 6D) represent the N-terminus of VEγ. The peaks represent the average m/z

Disulfide bridges of VE proteins were determined in the following manner. An intermolecular bridge represents the sum of the individual unprotonated peptides minus 2 Da (C oxidation) plus 1 proton (peaks represent the protonated form of the peptides/bridges in MALDI-MS). Intramolecular bridges gain a net mass of 2 (1 bridge) or 4 (2 bridges) Da after reduction. Examples of such MS analyses are shown for VEβ in Figure 2 and for VEγ in Figure 3. For VEγ the disulfide linkages assigned to the eight conserved C residues of its ZP domain are C1-C4, C2-C3, C5-C7, and C6-C8 (Type 1 disulfides; Figure 4). Furthermore, for VEγ residues C9 to C12 are present in disulfide linkages, C9-C11 and C10-C12. For VEβ, disulfide linkages could be assigned C1-C4, C2-C3, Cx-Cy, C5-C6, C7-Ca, and Cb-C8 (Type 2 disulfides; Figure 4). For VEα, disulfide linkages could be assigned C1-C4, C2-C3, Cx-Cy, C5-C6, C7-Ca, and Cb-C8 (Type 2 disulfides; Figure 4).

Intramolecular disulfide linkages in type-1 and type-2 ZP domains

TYPE-1 (Mouse ZP3 and Trout VEγ)
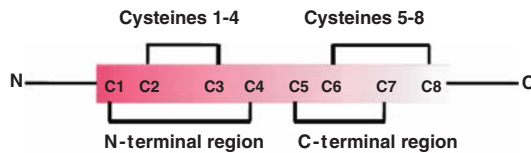
TYPE-2 (Mouse ZP1/2 and Trout VEα/β)
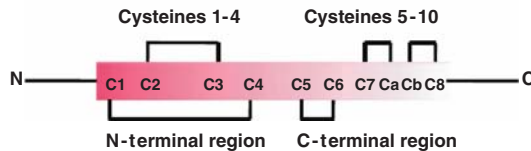
Figure 4. Molecular disulfide linkages in type-1 and type-2 ZP domains. Type-1 ZP domain is common for mouse ZP3 and fish VEγ, while type-2 ZP domain is common for mouse ZP1/ZP2 and fish VEα/VEβ. The N-terminal region has the same disulfide linkages in both type-1 and type-2 ZP domains, while the C-terminal region has different disulfide linkages in type-1 ZP domain, compared to type-2 ZP domain

2.3.   CELLULAR SITE OF PROTEOLYTIC PROCESSING OF VE PROTEINS

Where proteolytic processing of trout VE precursor proteins occurs during VE formation around eggs is of considerable interest. VE precursor proteins are secreted from hepatocytes into the bloodstream by female rainbow trout and by estrogen-induced males (Hyllner et al., 1991; Hyllner and Haux, 1992; Oppen-Berntsen et al., 1992; Fujita et al., 2002). These precursors contain a C-terminal propeptide that

must be removed by proteolytic cleavage at a consensus furin-like cleavage site (CFLCS) in order to produce mature protein that can be incorporated into VE fibrils (Sugiyama et al., 1999; Jovine et al., 2004, 2006).

MS (MALDI-TOF-MS, linear and reflective mode, and LC-MS) and gel-electrophoresis of enzymatically digested individual VE proteins were used in a peptidomics-based approach to determine the cellular site of proteolytic processing at the CFLCS of VE precursor proteins (Darie et al., 2005). Because only VE proteins that had recently arrived in the vicinity of eggs would be expected to retain a CFLCS and C-terminal propeptide, MS provides the level of sensitivity necessary to make such an assessment. MS was used to identify the C-terminal amino acid of VE proteins present on eggs and to determine whether the C-terminal propeptide, as well as an intact CFLCS, are present on at least a portion of the VE proteins surrounding eggs.

Figure 5. Analysis of peptides containing part of the C-terminal propeptide, CFLCS, and part of the ZP domain in VE proteins. (A) LC-MS(QTOF) analysis of a tryptic digest of VEβ revealed a triply charged monoisotopic peak with m/z 697.08 (calc. 697.05) (inset) which corresponds to peptide [487]CYRKRRDIPAAVQKTAR[503] (CM-C) from VEβ. The fragmentation spectrum of the precursor ion confirmed the identity of this peptide. The diagram in (A) represents the position of the peptides detected from VE proteins. (B) LC-MS(QSTAR) analysis of AspN digests of VEα revealed a quadruply charged monoisotopic peak with m/z 630.65 (calc. 630.57) (inset) that corresponds to a peptide with the sequence [525]EQMCNRQRRDL-SAQGQKKKTKG[545] (CM-C) from VEα. The CID spectrum of the parent ion produced a series of peaks that confirm the identity of the peptide

We found that the major C-terminal amino acid for VE proteins lies within the CFLCS and may be any one of the four CFLCS amino acids (Arg-X-Arg-Arg for VEα and β, and Arg-Lys-X-Arg for VEγ), consistent with previous results from our laboratory (Darie et al., 2004) and with studies of other mature ZP domain proteins (Litscher et al., 1999; Sasanami et al., 2002; Boja et al., 2003). Furthermore, MS peaks were found corresponding to sequences derived from the C-terminal propeptides of VEα, VEβ, and VEγ. Perhaps most convincing was the identification of MS peaks corresponding to a portion of mature VE proteins (i.e., upstream of the CFLCS), the CFLCS, and a portion of the C-terminal propeptide (i.e., downstream of the CFLCS) for VEα, VEβ, and VEγ (Fig 5). Collectively, these results suggest that removal of the C-terminal propeptide from VE precursor proteins takes place on the egg.

2.4. POLYMERIZATION OF PURIFIED VE PROTEINS

The fibrillar nature of the VE is consistent with results of structural studies of mouse egg ZP and other ZP domain-containing proteins, such as tectorins and uromodulin (Jovine et al., 2005). For example, the mouse egg ZP consists of
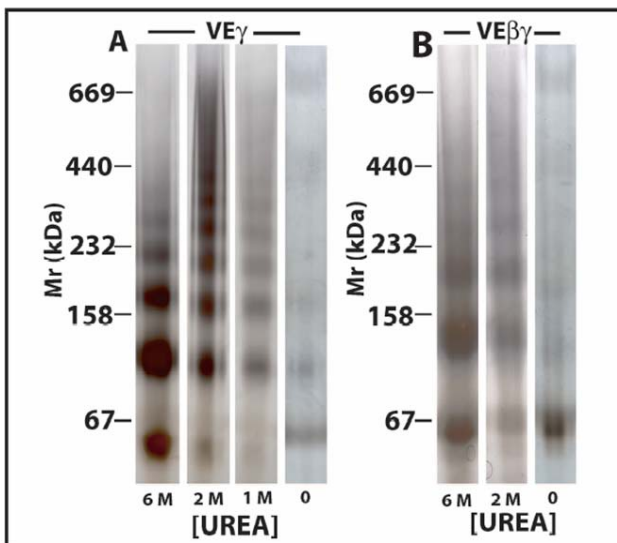
Figure 6. BN-PAGE of trout VEγ and VEβγ. Initially, samples were in 6 M urea at a concentration of ~1 mg/ml and ~2 mg/ml for VEγ and VEβγ, respectively. Samples in 2 M, 1 M, and no urea were obtained by extensive dialysis of VE proteins in 6 M urea which were then passed through a 200 μm filter. It was noted that VEγ in 1 M urea and VEβγ in 2 M and no urea produced visible aggregates. BN-PAGE lanes for VEγ with 6 M, 2 M, and 1 M urea (panel A) and for VEβγ with 6 M and 2 M urea (panel B) are portions of the same gel. BN-PAGE lanes with no urea (panels A and B) are also portions of the same gel. M_ρs are indicated and gels were stained with silver

long, interconnected fibrils exhibiting a 14–15 nm repeat (Greve and Wassarman, 1985; Wassarman and Mortillo, 1991). Recent evidence suggests that, under non-denaturing conditions (BN-PAGE), individual ZP proteins polymerize into higher order structures that are visualized by TEM as long interconnected fibrils that consist of contiguous beads (Litscher et al., 2008). Accordingly, we purified trout VEβ and VEγ to homogeneity and examined their behavior in the presence and absence of urea to determine whether individual VE proteins can also polymerize into oligomers and fibrils (Darie et al., 2007).

In our experiments, VEβ and VEγ were examined by SDS-PAGE, size-exclusion chromatography (SEC), TEM, and BN-PAGE. The latter technique separates proteins and multiprotein complexes under non-denaturing conditions on the basis of M_ρ due to external negative charge provided to proteins by Coomassie Brilliant Blue (CBB) binding to hydrophobic domains on the surface of proteins (Schagger and von Jagow, 1991; Schagger et al., 1994; Darie et al., 2005; Swamy et al., 2006). Such analyses revealed that in the presence of 6 M urea each VE protein is present primarily as monomers and as small oligomers (dimers, tetramers, etc.). However, either a reduction in urea concentration or complete removal of urea results in polymerization of VEβ and VEγ homodimers into very large oligomers (Figure 6). Mixtures of VEβ and VEγ also give rise to large oligomers, although it is unclear whether they represent homomeric or heteromeric oligomers. Under these conditions, VE proteins are visualized by TEM as aggregates of long fibrils, with each fibril composed of contiguous beads located periodically along the fibril
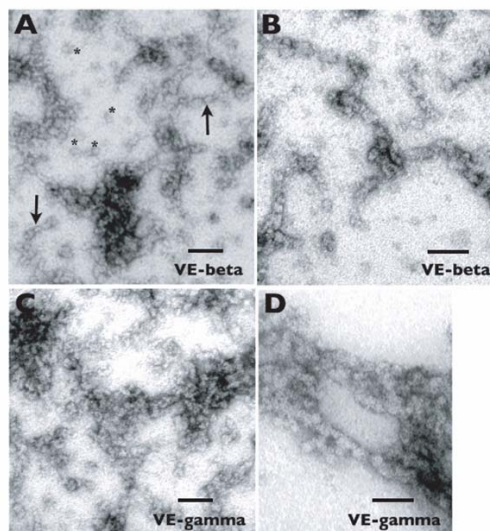
Figure 7. Transmission electron microscopy of VEβ and VEγ in buffer. Electron micrographs showing VEβ (A, ~200,000X; B, ~265,000X) and VEγ (C, ~222,000X; D, ~370,000X) fibrils. Arrows in panel A indicate contiguous beads along a fibril and asterisks indicate individual beads in the field

(Figure 7). Beads of VEβ and VEγ fibrils are 15.1 (±2.3) and 13.5 (±1.6) nm in diameter, respectively, suggesting that each bead consists of a homodimer of VE proteins. These results suggest that purified VEβ and VEγ can assemble *in vitro* into dimers and these in turn polymerize into oligomers and then fibrils of variable size depending on the experimental conditions.

3.   Discussion

Results summarized here, involving the use of MS, BN-PAGE, and TEM in particular, have provided considerable information about the structure, processing, and polymerization of rainbow trout VE proteins. They have permitted assignment of intramolecular disulfide linkages and the cellular site of proteolytic processing of VE proteins, as well as insight into the behavior of purified VE proteins *in vitro*. In addition, the results have permitted us to compare features of fish VE proteins with those of their mammalian egg counterparts, ZP proteins.

During the evolution of animal species the sites of synthesis of egg extracellular coat proteins have included the liver and ovary and, in the latter case, either growing oocytes or follicle cells, or both. Mammalian ZP proteins and amphibian VE proteins are synthesized in the ovary, whereas fish and bird VE proteins are synthesized either in the liver or ovary, or both (Litscher and Wassarman, 2007). It has been suggested that the acquisition of dual sites of synthesis or ZP-like proteins is the result of an ancient polyploidization event, followed by additional species-specific amplifications (Conner and Hughes, 2003). Whereas VE proteins of cyprinoid fish (e.g., zebrafish, goldfish, and carp) are synthesized in the ovary, VE proteins in a large number of non-cyprinoid fish (e.g., winter flounder, seabream, cod, and medaka), including rainbow trout, are synthesized in the liver and transported in blood to the ovary (Litscher and Wassarman, 2007).

Apparently, both VE and ZP proteins assemble into dimers which then polymerize into the long fibrils that constitute the extracellular coats of non-mammalian and mammalian eggs, respectively. Both types of proteins share several features, including a ZP domain, EHP and IHP, CFLFS, and C-terminal propeptide. Although trout VE proteins are synthesized in the liver and ZP proteins in the ovary, these common features suggest that both types of proteins are processed and polymerized by identical mechanisms *in vivo*. The mechanism involves proteolytic cleavage at the CFLFS and release of the C-terminal propeptide with its EHP, thereby exposing the ZP domain with its IHP and permitting polymerization to take place (Jovine et al., 2004, 2006). A similar proteolytic processing mechanism that regulates polymerization of proteins has been reported for several other proteins as well (Taylor et al., 1997; Bourne et al., 2000; Handford et al., 2000; Mosesson et al., 2001). In the case of trout VE proteins proteolytic processing at the CFLFS takes place after arrival of VE precursor proteins at the egg (ovary) (Darie et al., 2005).

It is of interest that disulfides C1-4 and C2-C3 are found in VE and ZP (Boja et al., 2003) proteins, suggesting that the N-terminal portion of the ZP domain is highly conserved in fish and mammals. This may be related in turn to the observation that the N-terminal half of the ZP domain can polymerize into long fibrils on its own and that some proteins contain only the N-terminal half of the ZP domain (Jovine et al., 2006). Several lines of evidence suggest that the ZP domain is a bipartite structure consisting of two subdomains (N- and C-domains) separated by a short, protease-sensitive linker (Jovine et al., 2004, 2006; Llorca et al., 2007).

Trout VEγ is most similar to mouse ZP protein ZP3, sharing ~25% sequence identity and having identical disulfide linkages in their ZP domains. On the other hand, VEα and VEβ, despite a relatively low sequence identity, are related to mouse ZP1 and ZP2. Under non-denaturing conditions both purified mouse ZP2 and ZP3 polymerize into higher order structures visualized by TEM as long inter-connected fibrils that consist of contiguous beads (Litscher et al., 2008). Similarly, under non-denaturing conditions purified trout VEβ and VEγ also polymerize into fibrils from homomeric dimers (Darie et al., 2007). It is likely that the fibrils arise, at least in part, due to interactions between ZP domains of the individual VE proteins. In this context, it is of interest that VEα and VEβ have the same disulfide linkages within their ZP domains, but VEγ has different linkages in the C-terminal half of its ZP domain (Darie et al., 2004). It has been suggested that the C-terminal portion of the ZP domain influences the specificity of protein-protein interactions, thereby determining whether homo- or hetero-dimers of ZP domain proteins form (Jovine et al., 2005; Darie et al., 2007).

Many questions about VE assembly and structure remain unanswered. For example, it is unclear how some fish VE protein precursors, such as those of the rainbow trout, are targeted specifically to the ovary. Perhaps, there is a receptor-mediated mechanism, responsible for the uptake of fish VE protein precursors, similar to that responsible for the uptake of the yolk precursor protein, vitellogenin, into the ovary in fish, birds, and amphians (Wallace, 1985) and the uptake of very-low-density lipoprotein, riboflavin-binding protein, and α-2-macroglobulin in the chicken (Schneider, 1996). Whether or not, a similar receptor-mediated mechanism applies to uptake of fish VE protein precursors remains to be determined.

Also of interest is the question of the nature and source of the furin-like enzyme that cleaves VE precursor proteins in the ovary. Possibly the enzyme is associated with oocyte plasma membrane that is close to the innermost layer of the VE into which nascent, processed VE proteins are incorporated. Furthermore, the enzyme could also serve as a receptor for VE proproteins, which would explain why rainbow trout VE proproteins lack a transmembrane domain.

Acknowledgments

References

Boja, E. S., T. Hoodbhoy, et al. (2003). "Structural characterization of native mouse zona pellucida proteins using mass spectrometry." *J Biol Chem* **278**(36): 34189–202.

Bourne, Y., M. H. Watson, et al. (2000). "Crystal structure and mutational analysis of the Saccharomyces cerevisiae cell cycle regulatory protein Cks1: implications for domain swapping, anion binding and protein interactions." *Structure* **8**(8): 841–50.

Brivio, M. F., R. Bassi, et al. (1991). "Identification and characterization of the major components of the Oncorhynchus mykiss egg chorion." *Mol Reprod Dev* **28**(1): 85–93.

Carr, M. D., C. J. Bauer, et al. (1994). "Solution structure of a trefoil-motif-containing cell growth factor, porcine spasmolytic protein." *Proc Natl Acad Sci U S A* **91**(6): 2206–10.

Conner, S. J. and D. C. Hughes (2003). "Analysis of fish ZP1/ZPB homologous genes--evidence for both genome duplication and species-specific amplification models of evolution." *Reproduction* **126**(3): 347–52.

Darie, C. C., M. L. Biniossek, et al. (2005). "Mass spectrometric evidence that proteolytic processing of rainbow trout egg vitelline envelope proteins takes place on the egg." *J Biol Chem* **280**(45): 37585–98.

Darie, C. C., M. L. Biniossek, et al. (2004). "Structural characterization of fish egg vitelline envelope proteins by mass spectrometry." *Biochemistry* **43**(23): 7459–78.

Darie, C. C., M. L. Biniossek, et al. (2005). "Isolation and structural characterization of the Ndh complex from mesophyll and bundle sheath chloroplasts of Zea mays." *Febs J* **272**(11): 2705–16.

Darie, C. C., W. G. Janssen, et al. (2007). "Purified trout egg vitelline envelope proteins VEbeta and VEgamma polymerize into homomeric fibrils from dimers in vitro." *Biochim Biophys Acta* doi:10.1016/j.bbapap.2007.10.011.

Dumont, J. N. and A. R. Brummett (1985). *Egg envelopes in vertebrates*. New York, Plenum press.

Fujita, T., M. Shimizu, et al. (2002). "Purification of serum precursor proteins to vitelline envelope (choriogenins) in masu salmon, Oncorhynchus masou." *Comp Biochem Physiol B Biochem Mol Biol* **132**(3): 599–610.

Greve, J. M. and P. M. Wassarman (1985). "Mouse egg extracellular coat is a matrix of interconnected filaments possessing a structural repeat." *J Mol Biol* **181**(2): 253–64.

Hagenmaier, H. E. (1973). "The hatching process in fish embryos. 3. The structure, polysaccharide and protein cytochemistry of the chorion of the trout egg, Salmo gairdneri (Rich.)." *Acta Histochem* **47**(1): 61–9.

Handford, P. A., A. K. Downing, et al. (2000). "Fibrillin: from domain structure to supramolecular assembly." *Matrix Biol* **19**(6): 457–70.

Hyllner, S. J. and C. Haux (1992). "Immunochemical detection of the major vitelline envelope proteins in the plasma and oocytes of the maturing female rainbow trout, Oncorhynchus mykiss." *J Endocrinol* **135**(2): 303–9.

Hyllner, S. J., D. O. Oppen-Berntsen, et al. (1991). "Oestradiol-17 beta induces the major vitelline envelope proteins in both sexes in teleosts." *J Endocrinol* **131**(2): 229–36.

Hyllner, S. J., L. Westerlund, et al. (2001). "Cloning of rainbow trout egg envelope proteins: members of a unique group of structural proteins." *Biol Reprod* **64**(3): 805–11.

Iwamatsu, T., N. Yoshizaki, et al. (1997). "Changes in the chorion and sperm entry into the micropyle during fertilization in the teleostean fish, Oryzias latipes." *Dev Growth Differ* **39**(1): 33–41.

Jovine, L., C. C. Darie, et al. (2005). "Zona pellucida domain proteins." *Annu Rev Biochem* **74**: 83–114.

Jovine, L., W. G. Janssen, et al. (2006). "The PLAC1-homology region of the ZP domain is sufficient for protein polymerisation." *BMC Biochem* **7**: 11.

Jovine, L., H. Qi, et al. (2002). "The ZP domain is a conserved module for polymerization of extracellular proteins." *Nat Cell Biol* **4**(6): 457–61.

Jovine, L., H. Qi, et al. (2006). *Features that affect secretion and assembly of zona pellucida glycoproteins during mammalian oogenesis*. Nottingham, U. Press.

Jovine, L., H. Qi, et al. (2004). "A duplicated motif controls assembly of zona pellucida domain proteins." *Proc Natl Acad Sci U S A* **101**(16): 5922–7.

Litscher, E. S., W. G. Janssen, et al. (2008). "Purified mouse egg zona pellucida glycoproteins polymerize into homomeric fibrils under non-denaturing conditions." *J Cell Physiol* **214**(1): 153–7.

Litscher, E. S., C. Liu, et al. (1999). "Zona pellucida glycoprotein mZP3 produced in milk of transgenic mice is active as a sperm receptor, but can be lethal to newborns." *Transgenic Res* **8**(5): 361–9.

Litscher, E. S. and P. M. Wassarman (2007). "Egg extracellular coat proteins: from fish to mammals." *Histol Histopathol* **22**(3): 337–47.

Llorca, O., A. Trujillo, et al. (2007). "Structural model of human endoglin, a transmembrane receptor responsible for hereditary hemorrhagic telangiectasia." *J Mol Biol* **365**(3): 694–705.

Mosesson, M. W., K. R. Siebenlist, et al. (2001). "The structure and biological features of fibrinogen and fibrin." *Ann N Y Acad Sci* **936**: 11–30.

Oppen-Berntsen, D. O., S. J. Hyllner, et al. (1992). "Eggshell zona radiata-proteins from cod (Gadus morhua): extra-ovarian origin and induction by estradiol-17 beta." *Int J Dev Biol* **36**(2): 247–54.

Sasanami, T., J. Pan, et al. (2002). "Secretion of egg envelope protein ZPC after C-terminal proteolytic processing in quail granulosa cells." *Eur J Biochem* **269**(8): 2223–31.

Schagger, H., W. A. Cramer, et al. (1994). "Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis." *Anal Biochem* **217**(2): 220–30.

Schagger, H. and G. von Jagow (1991). "Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form." *Anal Biochem* **199**(2): 223–31.

Schneider, W. J. (1996). "Vitellogenin receptors: oocyte-specific members of the low-density lipoprotein receptor supergene family." *Int Rev Cytol* **166**: 103–37.

Shibata, Y., T. Iwamatsu, et al. (2000). "Identification and cDNA cloning of alveolin, an extracellular metalloproteinase, which induces chorion hardening of medaka (Oryzias latipes) eggs upon fertilization." *J Biol Chem* **275**(12): 8349–54.

Sugiyama, H., K. Murata, et al. (1999). "Formation of mature egg envelope subunit proteins from their precursors (choriogenins) in the fish, Oryzias latipes: loss of partial C-terminal sequences of the choriogenins." *J Biochem (Tokyo)* **125**(3): 469–75.

Swamy, M., G. M. Siegers, et al. (2006). "Blue native polyacrylamide gel electrophoresis (BN-PAGE) for the identification and analysis of multiprotein complexes." *Sci STKE* **2006**(345): p. l4.

Taylor, K. M., A. R. Trimby, et al. (1997). "Mutation of recombinant complement component C9 reveals the significance of the N-terminal region for polymerization." *Immunology* **91**(1): 20–7.

Wallace, R. A. (1985). *Vitellogenesis and oocyte growth in nonmammalian vertebrates*. New York, Plenum Press.

Wassarman, P. M. and S. Mortillo (1991). "Structure of the mouse egg extracellular coat, the zona pellucida." *Int Rev Cytol* **130**: 85–110.

3. MALDI/MS COMPARISON OF FE-NTA IMMOBILIZED METAL AFFINITY CHROMATOGRAPHY AND COMMERCIALLY-AVAILABLE METAL OXIDE AFFINITY RESINS FOR PHOSPHOPEPTIDE ENRICHMENT

MATTHEW B. GATES, KENNETH B. TOMER, AND LEESA J. DETERDING
Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709, USA

Abstract Immobilized metal ion affinity chromatography in combination with mass spectrometry has been used to determine the extent of phosphorylation and the specific sites of phosphorylation on proteins. There are several advantages to this combined approach. Immobilized metal ion affinity chromatography is the use of resins with metal constituents and metal oxides to enrich and isolate phosphopeptides. By selectively enriching phosphopeptides on media prior to MS analyses, suppression effects can be greatly reduced. In this report, we have investigated several resins from various sources to assess the phosphopeptide enrichment capabilities of each prior to mass spectrometric analyses. The phosphopeptide enrichment capabilities of six different resins, Glygen TiO_2, ZrO_2, and mixed ZrO_2 and TiO_2 NuTips, Titansphere TiO_2 resin, PhosTrap magnetic titanium beads, and a Fe-NTA resin, are compared.

1.   Introduction

One of the most common and important means of biologically regulating protein activity is through reversible phosphorylation and dephosphorylation. Phosphorylation is involved in the regulation of gene expression and protein synthesis, which controls cell growth, division or differentiation. It has been reported that an estimated 30% of all eukaryotic proteins are phosphorylated at some point during their existence [1]. Also, deregulation of the phosphorylation process has been shown to be correlated with several types of cancers [2,3]. In order to better understand the molecular basis of these regulatory mechanisms, it is useful to identify the specific amino acid residues undergoing phosphorylation.

A number of different approaches to the identification of phosphorylation sites are in common use. Immunoblotting of proteins with antibodies that recognize phosphoamino acid epitopes is one of the more sensitive and widely-used techniques for detecting specific phosphorylation sites, but this approach requires prior knowledge of the phosphorylation site and the availability of an antibody that can recognize the antigenic site. Another commonly used approach is based on the

incorporation of a radioactive phosphate into the target protein. Incorporation of a radioactive phosphate, however, may alter the physiological state of the cells.

Mass spectrometry (MS) has been used to determine the extent of phosphorylation and the specific sites of phosphorylation on proteins [4]. There are several advantages to the MS-based approach, including generally rapid analyses and the fact that it does not require radiolabeling. But there are also some drawbacks, including suppression of low-level phosphopeptides in complex mixtures and the low sensitivity of phosphorylated peptides under positive ionization conditions due to the negative charge on the phosphate moiety. HPLC separation in combination with MS can be used to reduce suppression effects by introducing the phospho-peptide to the mass spectrometer as a pure peak or as a component of a less complex mixture. The HPLC eluent can be either directly analyzed using ESI/MS/MS to locate the phosphorylation sites or HPLC fractions can be collected and analyzed individually by MALDI/MS or ESI/MS. HPLC separations, however, can be time consuming.

Immobilized metal ion affinity chromatography (IMAC) is the use of resins with metal constituents to isolate phosphopeptides. By selectively enriching phospho-peptides on media prior to MS analysis, suppression effects can be greatly reduced. The enriched samples can then be subjected to MALDI/MS and/or ESI/MS analyses. In 1986, the free phosphate ion binding capability of Fe (III) was observed by Anderson and Porath [5]. However, the strong binding between phosphoserine and an $Fe(III)L_2$ complex has been known since the late 1950s [6]. Since these initial findings, reports of many metals showing affinity towards phosphate have been published, including: Lu (III), Th (III), Th (IV), and Sc (III) [5,7]. More commonly, phosphoproteins and peptides are bound with high specificity to immobilized metal ions, such as Fe^{3+} and Ga^{3+}. Enrichment techniques can be used with either on-line or off-line MS analysis. Tempst et al. [7] reported better selectivity with Ga^{3+} vs. Fe^{3+} for phosphopeptides. Our laboratory has previously shown that affinity-bound analytes, including phosphopeptides, can be directly analyzed by MALDI/MS without prior elution from both iron and gallium IMAC affinity media [8–11]. In addition, other studies have reported improved enrichment efficiency when using Al (III) rather than Fe (III) [12].

There are also many metal ion chelators used in traditional IMAC that bind in various motifs. Specifically, tridentate iminodiacetic acid (IDA), tridentate dipico-lylamine, tetradentate nitrilotriacetic acid (NTA), and pentadentate N,N,N'-tris (carboxymethyl)ethlenediamine are chelators that have been used in IMAC methods [13]. Two chelators, IDA and NTA, are widely used and commercially- available for a variety of affinity methods, and both have been used to enrich phosphopeptides. Fe-IDA bound to phosphate consists of a four member ring, consisting of two bonds between the metal center and the two oxygens of the phosphate group [5,7,14,15]. It has been reported that the NTA chelator is more selective than IDA [16]. The selectivity of NTA is due to the fact that the resin possesses a tetradentate metal binding ligand compared to IDA which binds to the metal ion through three coordination sites. Using an IDA resin may result in the loss of metal ions from the resin during

washing steps, not only contaminating the sample, but allowing more non-specific binding because of exposed carboxyl groups [7,17,18].

Recently metal oxides have received much attention after reports of improved isolation of phosphopeptides from protein digests using, for example, titanium dioxide [19,20] and zirconium dioxide [21]. Monodentate, bidentate and bidentate-bridging are possible binding motifs between phosphates and metal oxides [22,23]. Furthermore Håkansson, et al. demonstrated that titanium dioxide enriches predominantly multiphosphorylated peptides, while zirconium dioxide enriches mostly singly phosphorylated peptides [21].

Chemical modification of protein digests to enhance selectivity of enrichment media or to increase MS detection has been reported [13,24,25]. Methyl esterification of protein digests has also been used pre-enrichment to enhance the selectivity of IMAC resins by reducing the presence of carboxylic acid group. For example, it was observed in one study that [Glu^1]-fibrinopeptide shows affinity towards TiO$_2$, whereas, after methyl esterification, the peptide no longer binds to the titanium surface [19]. Methyl esterification, however, may result zn unwanted side reactions, and can complicate analyses in this regard. Also, esterification increases the time required to analyze a sample.

Enriched samples can also be treated with alkaline phosphatase to remove the phosphate [8,10]. A shift in mass of 80 Da or multiples of 80 Da that is observed in the MALDI/MS spectra after the phosphatase treatment confirms the loss of phosphate and the presence of a phosphopeptide. Subsequent MALDI/MS/MS analyses are more efficient without the presence of phosphate groups, and database searching can identify the peptides. However, this method is incomplete in that the specific sites of phosphorylation can not be determined.

During peptide binding to enrichment media, reagents can be added to enhance selectivity. Sodium chloride and phosphate buffers are often used in IMAC to reduce unwanted interactions that may hinder the efficiency of target protein or peptide(s) purification [5,26]. These salts and buffers, however, reduce ionization efficiency in both MALDI and ESI. In addition, it has been reported that high organic content in digest loading buffers, as well as low pH conditions, minimize non-specific binding to enrichment media [20,26].

Given the large number of metals and procedures reported in the literature, a comparison of IMAC resins using a standard digest mixture should be valuable. Furthermore, with reports of metal oxides as IMAC resins (e.g., titanium dioxide and zirconium dioxide), we have begun the development of the use of these resins for phosphopeptide enrichment prior to MALDI/MS and ESI/MS. In addition, the manufacturing process of metal oxides can lead to different crystal structures and surface chemistries (e.g., varying isoelectric points). Therefore, we have investigated several resins from various sources to assess the enrichment capabilities of each. Six different resins are compared: Glygen TiO$_2$ NuTip, ZrO$_2$ NuTip, mixed ZrO$_2$ and TiO$_2$ NuTip, Titansphere TiO$_2$, PhosTrap magnetic titanium beads, and a Fe-NTA resin.

2.   Methods

Materials. Ethylenediaminetetraaccetic acid (EDTA), α-cyano-4-hydroxy-cinnamic acid (α-cyano), 2,5-dihydroxybenzoic acid (DHB), ammonium bicarbonate (ABC), formic acid (FA), bovine α-casein, and bovine β-casein were purchased from Sigma-Aldrich Chemical Company (St. Louis, MO). The α-cyano was recrystallized from hot methanol and stored in the dark at room temperature. Also, 18 MΩ water was prepared on a model RO 40 water system (Hydro Service and Supplies, Durham, NC). Acetonitrile (ACN) and acetic acid were purchased from Caledon (Georgetown, Ontario, Canada). Trifluoroacetic acid (TFA) was purchased from Pierce (Rockford, IL), phosphoric acid (PA) was purchased from EMD Chemicals (Gibbstown, NJ), and ethanol (EtOH) was purchased from Warner-Graham (Cockeysville, MD). All solvents were HPLC grade. Ferric chloride ($FeCl_3$) was purchased from Allied Chemical (Brighton, UK). Ni-NTA resin was purchased from Qiagen (Valencia, CA). Compact reaction columns (CRC) and filters (10 μm pore size) were obtained from USB Corporation (Cleveland, OH). Titansphere TiO_2 resin was a gift from GL Sciences (Tokyo, Japan). The Phos-trap kit was purchased from Perkin Elmer (Waltham, Massachusetts). Metal oxide enrichment NuTips were purchased from Glygen Corp. (Columbia, MD).

Tryptic Digestion Conditions. The proteins were subjected to digestion with sequencing-grade modified porcine trypsin (Promega Corporation, Madison, WI). Enzyme was added at a protein:enzyme ratio of 20:1 and digests were allowed to proceed for at least 2 h at 37°C.

Iron IMAC. Iron IMAC columns were prepared as previously reported [10].

Titanium Dioxide IMAC. The TiO_2 IMAC procedure was adopted from Larsen, et al. [20]. Briefly, 1 mg of Titansphere TiO_2 resin was added to 10 μl bead buffer (80% ACN, 19.9% H_2O, 0.1% TFA) and applied to a CRC tube. Binding of phosphopeptides to the TiO_2 column was achieved by first draining the bead buffer, then loading the protein digestion mixture with 20 μl protein buffer (100 mg/ml DHB in 80% ACN, 19.9% H_2O, 0.1% TFA). The mixture was allowed to incubate at room temperature for 10 min, after which, the column was drained, and washed 1×30 μl protein buffer and 1 x 30 μl bead buffer. Elution of the bound peptides was accomplished by adding 5 μl of 200 mM ABC buffer (pH 10.5) to the column and incubated for 5 min at room temperature. The 5 μl elution solution was then drained and diluted with 5 μl water. 0.5 μl of the eluent solution was spotted on a MALDI target with 0.5 μl of DHB matrix (20 mg/ml DHB in 50% ACN, 49% H_2O, 1% PA).

Phos-Trap^{TM} Kit. The Phostrap magnetic titanium beads were prepared according to the manufacturer's instructions. Briefly, a 10 μl aliquot of magnetic beads was diluted to 200 μl in deionized water and placed in a 96-well plate. The beads were washed $3 \times$ with 200 μl binding buffer. After placing the plate on a separation magnet, which moved the beads to an inner o-ring of the 96-well plate, the supernatant was removed. The sample was diluted 1:10 in binding buffer and incubated

with beads for 10 min. Three 200 μl washes of binding buffer were followed by a single wash of 100 μl washing buffer. The peptides bound to the magnetic beads were eluted with 20 μl of elution buffer, lyophilized, and resuspended in 0.1% formic acid just prior to MS analyses.

Zirconium, Titanium, and Mixed Resin NuTips. Binding and washing of samples were performed similar to the manufacturer's guidelines, except the binding conditions were as reported previously for TiO$_2$ [20]. The tips were first equilibrated with a binding solution (80% ACN, 19.9% H$_2$O, 0.1% TFA) by washing 10 × 10 μl. Excess buffer was removed from the tip, and the digested sample was loaded (1:1 mixture of binding solution and sample) by aspirating and expelling 10 × 10 μl. H$_2$O washes followed (10 x 10 μl), and the bound peptides were eluted with 10 μl of 200 mM ABC (pH=10.5).

Matrix Assisted Laser Desorption Ionization. Positive ion MALDI MS analyses were performed using either a Voyager Super STR (Applied Biosystems, Framingham, MA) or a Waters Micro MX (Waters Corp., Beverly, MA). Both instruments are delayed-extraction time-of-flight (TOF) mass spectrometers and are equipped with a nitrogen laser (337 nm) to desorb and ionize the samples. A close external calibration, using two points to bracket the mass range of interest, was used. Either a 0.5 μl aliquot of the tryptic peptide elution or 0.5 μl of the IMAC resin was spotted with 0.5 μl MALDI matrix on a stainless steel sample target and allowed to dry at room temperature. A saturated solution of α-cyano in 45:45:10 EtOH:H$_2$O:FA (v/v) or 20 mg/ml DHB in 50:49:1 ACN:H$_2$O:PA (v/v) was used as the MALDI matrix. Spectra were obtained over the mass range of 800–4,000 Da in linear mode with an average 100 laser shots per spectrum.

MALDI MS/MS analyses were performed using an Applied Biosystems (Framingham, MA) 4,700 Proteomics Analyzer in the positive ion reflector mode. This instrument is equipped with a frequency tripled Nd:YAG laser (355 nm) to desorb and ionize the samples and a TOF-TOF mass spectrometer as the mass analyzer. For the MALDI analyses, 0.5 μl of the peptide elution samples were mixed with 0.5 μl MALDI matrix consisting of a 3:1 dilution of a saturated solution of α-cyano in 50:50 acetonitrile:water (0.1% FA) (v/v) on a 100-well MALDI sample target. Spectra were obtained over the mass range of 800–4,000 Da with 1,000 laser shots per spectrum

Electrospray Ionization. LC/ESI/MS/MS analyses were performed using an Agilent (Santa Clara, CA) XCT Ultra Ion Trap with an HPLC chip cube MS interface and an Agilent 1100 nanoHPLC. 30 μl injections were made onto a 40 nl enrichment column followed by a 43 mm × 75 μm analytical column, packed with ZORBAX 300SB C18 particles. Linear gradients of 3–50% (0.1% FA) over 50 min at a flow rate of 500 nl/min and automated data dependent acquisitions were employed. MS/MS data processing was completed by Spectrum Mill software (Agilent Corporation). The MS/MS data were searched against an in-house bovine casein database. Prior to database searching variable modifications were set for oxidized methionine, N-terminal pyroglutamic acid, and phosphorylated serine, threonine, and tyrosine residues using the Spectrum Mill software. Digestion specificity was

defined as trypsin with a maximum of four missed cleavages. Search results were interpreted manually to ensure correct assignment of modifications.

3.   Results and Discussion

The α- and β-casein tryptic digests were mixed together in solution and the resulting MALDI mass spectrum is shown in Figure 1. In this spectrum, only two ions which correspond in mass to phosphorylated tryptic peptides (T15 and T14-15) are observed (<10% relative abundance).
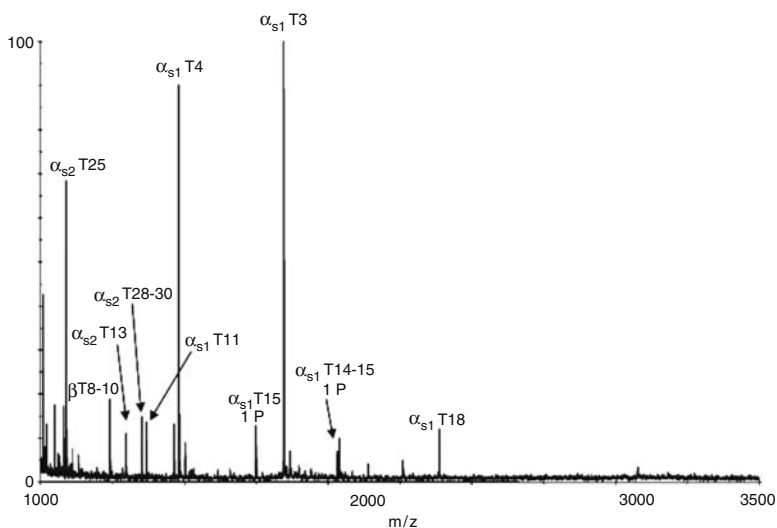
Figure 1. MALDI/MS of α- and β-casein tryptic digest before enrichment

Iron as an IMAC metal has been shown to efficiently enrich phosphopeptides from complex mixtures [7–11]. This is expected as the binding constant of Fe^{3+} is on the order of 10^{13} for a phosphorylated serine residue [5]. Using radiolabeled phosphate, it has been demonstrated that iron resins retain approximately 80–90% of phosphorylated species under optimum conditions [7]. Iron adducts on peptides have also been observed after elution from IMAC media [27].

We have previously shown successful enrichment of phosphopeptides using Fe-NTA IMAC and direct MALDI/MS analysis of the resin beads [8–11]. Therefore, the casein digestion mixture was loaded onto a Fe-NTA column for phosphopeptide enrichment. Acetic acid was used to minimize non-specific binding of acidic peptides. As a comparison to the MALDI analysis without enrichment, the resulting MALDI mass spectrum acquired directly from the Fe-NTA IMAC resin (Figure 2) following phosphopeptide enrichment shows eight ions corresponding in mass to phosphorylated tryptic peptides, T16, T15, T7, T14-15 (αS1), and T8 of α-casein and T6, T2, and T1-2 of β-casein. Ions corresponding in mass to the

nonphosphorylated tryptic peptides T20-22, T11, and T4 are also observed and resulted from nonspecific binding to the Fe-NTA resin.

To determine whether titanium dioxide resins were suitable for phosphopeptide enrichment, the α- and β-casein mixture was loaded onto an in-house prepared Titansphere TiO_2 CRC column for phosphopeptide enrichment. The resulting MALDI mass spectrum (Figure 3) showed nine ions corresponding in mass to the phosphorylated tryptic peptides T15, T7, oxidized T7, oxidized T16, T14-15, and T8 of α-casein and T6, T2, and T1-2 of β-casein. Ions corresponding in mass to tryptic peptides T13 and T28-30 were also observed and resulted from nonspecific binding to the TiO_2 resin.
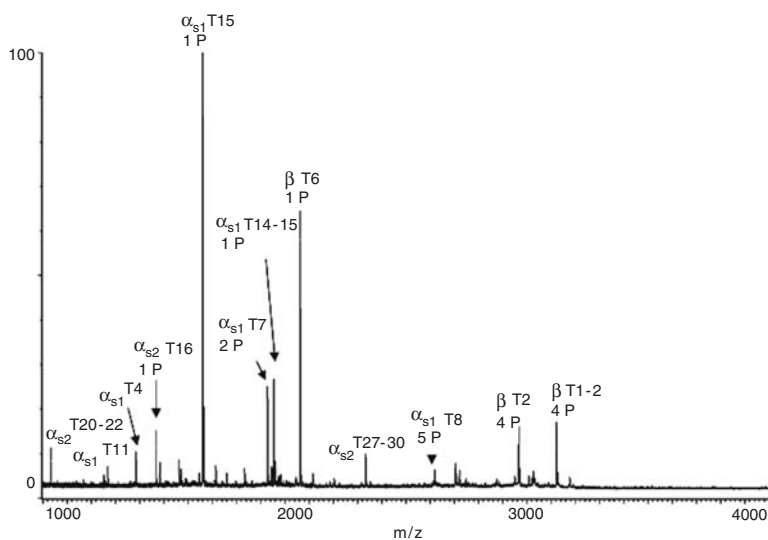
Figure 2. MALDI/MS of α- and β-casein peptides bound to Fe-NTA beads

In comparison to the Fe-IMAC enrichment where eight phosphopeptide ions and three nonphosphorylated peptide ions were observed, these data showed promise that the TiO_2 resin may prove beneficial for the enrichment of phosphorylated peptides.

The fact that fewer nonphosphorylated ions were observed with the TiO_2 resin suggests that perhaps less nonspecific binding occurs than with the previously reported Ga^{3+} and Fe^{3+} IMAC resins. The results shown here are similar to those previously reported for the titanium dioxide resin [20].

Reports of titania's affinity towards phosphates has been known for some time. At pH 2.3 McQuillan et al. found the Langmuir binding constant of titania to phosphate to be 3.8 (+/- 0.8) $\times 10^4$ dm^3 mol^{-1}. They also confirmed the bidentate binding motif of phosphate moieties to the titanium dioxide surface [28]. The iso-electric point of titanium dioxide has been found to be at a pH of 5, but this can vary depending on how the resin was manufactured [29].
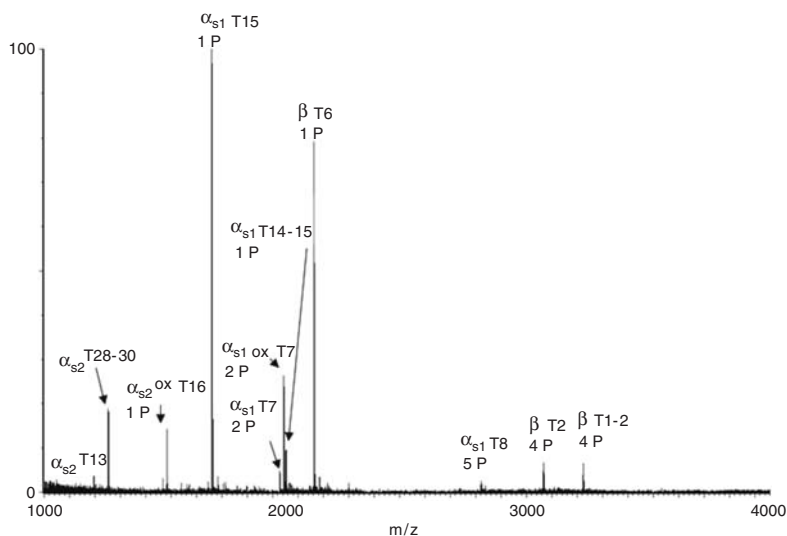
Figure 3. MALDI/MS of α- and β-casein peptides obtained from the Titansphere TiO_2 elution

It has been reported that the addition of acetonitrile can be effective for counter-acting hydrophobic peptides from binding to enrichment resins and the addition of acids, such as DHB and phthalic acid, can be used to further reduce nonphos-phorylated peptides from binding to titanium dioxide resins [30]. The use of substi-tuted aromatic carboxylic acids, specifically DHB, however, can create problems for LC systems and MS analyses due to contamination of the column and the mass spectrometer [30]. For the study reported here, the inclusion of DHB in the loading buffer to minimize nonspecific binding was investigated using the titanium dioxide CRC. Because of potential contamination to the LC system, the resulting eluent was subjected to MALDI/MS analyses only. The use of DHB at various concen-trations (25–300 mg/ml) in the loading buffer resulted in less non-specific binding as well as an increase in the number of phosphopeptides observed compared to enrichment on titanium media without DHB (data not shown). The use of 100 mg/ml DHB in the binding solution appeared to be optimum, where other reports suggest values as high as 300 mg/ml DHB for complex samples [20]. In our hands, quality MALDI/MS spectra could not be obtained when a concentration of 300 mg/ml DHB was used in the digest loading buffer (data not shown).

Reports of metal oxide nanoparticle synthesis for the purpose of phosphopep-tide enrichment have been reported in the literature [31–34]. For example, ZrO_2 nanoparticles (twenty nanometer particle size) have been reported as effective enrichment media [32]. The synthesis of these nanoparticles, however, can be very time consuming. Magnetic nanoparticles coated with titanium dioxide as well as zirconium dioxide have been synthesized and used for phosphopeptide enrichment [33]. The magnetic properties of enrichment media allows for the use of small

particle sizes with high surface area; thereby, making the process of elution simpler and contamination of the eluent less likely. Therefore, the commercially-available PhosTrap kit, which uses magnetic titanium beads, was evaluated for its utility to enrich for phosphopeptides. Following enrichment, the MALDI mass spectrum was acquired and is shown in Figure 4. Three ions were observed which correspond in mass to phosphorylated tryptic peptide T15 and T14-15 of α-casein and T6 from β-casein. In addition, many ions (>10) corresponding in mass to nonphosphorylated tryptic peptides were also observed (e.g., T25, T9-10, and T3 from α-casein, and T8-10 from β-casein). Also, several unidentified background ions were observed.
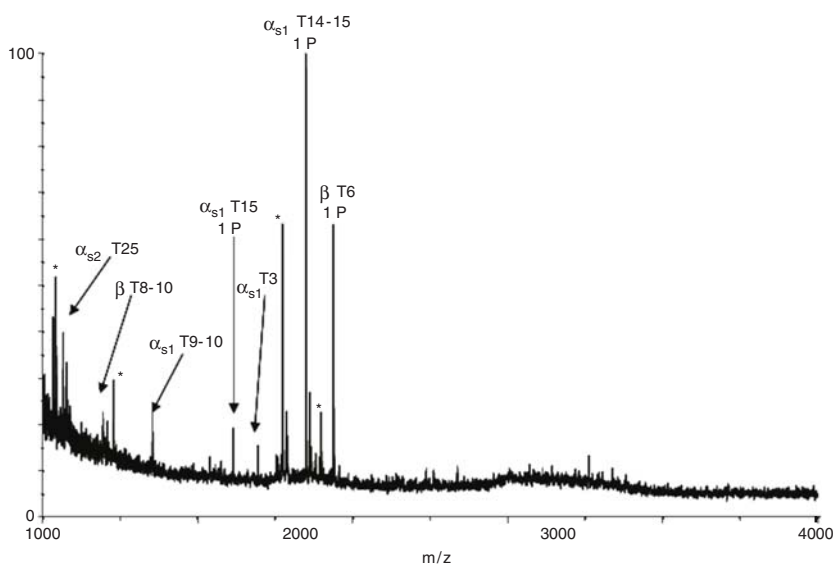
Figure 4. MALDI/MS of α- and β-casein peptides eluted from PhosTrap magnetic titanium beads. Background ions are labeled with an asterisk (*)

These MALDI MS data suggest that more non-specific binding occurs with the use of the PhosTrap kit compared to the Fe-NTA CRC or TiO_2 CRC methods. Also, these data suggest that possibly the magnetic titanium beads do not efficiently bind to phosphopeptides or phosphopeptides do not efficiently elute from the magnetic beads using the manufacturer's provided elution solution, which has a pH lower than the titanium elution buffers (pH 9 vs. 10.5).

Metal oxide enrichment products are also commercially-available from Glygen (NuTips) and, therefore, were employed to enrich the casein digestion mixture. The first NuTip investigated for phosphopeptide enrichment contained a TiO_2 resin. The resulting MALDI/MS data acquired from the eluent of the TiO_2 tip (Figure 5) shows ions which correspond in mass to eight phosphopeptides, tryptic peptide T16, T14-15 (from both α_{s1} and α_{s2}), T15, T7, and T8 from α-casein and tryptic

peptides T6 and T1-2 from β-casein. In addition, ions from nonphosphorylated tryptic peptides were observed including: T14, T10-11, and oxidized T15, from β-casein, as well as T17, T9-10, T3, and T27-30 from α-casein. These results indicate that this resin is not as selective as the Titansphere titanium dioxide resin.

Similar to the studies with the titanium dioxide CRC, the use of DHB was investigated to determine whether the amount of nonspecific binding could be reduced. Enrichment and MALDI/MS analyses were performed with and without the addition of 100 mg/ml DHB to the digest loading buffer. Based on a comparison of the result-ing MALDI spectra, the addition of the DHB did not enhance selectivity (data not shown). In contrast, the addition of DHB enhanced the selectivity of Titansphere titanium dioxide resin. These results may possibly be due to differences in manufacturing of the two resins; thereby, resulting in different surface chemistry [35].

Like iron, zirconia has been shown to retain 80–90% of phosphorylated species in radiolabeling experiments [7]. We, therefore, investigated the use of Glygen's ZrO_2 tip for phosphopeptide enrichment. The MALDI mass spectrum (Figure 6) acquired from the eluent of the ZrO_2 tip following enrichment was very similar to that obtained of the eluent following enrichment with Glygen's TiO_2 tip (Figure 5). Ions corresponding in mass to eight phosphopeptides, tryptic peptides T16, T16-17, T15, T7, T14-15 ($α_{s1}$), and T8 from α-casein and T6 and T1-2 from β-casein, were observed. Non-specific binding of peptides to the resin was very similar to that of the titanium tip, i.e., more than 10 non-specifically bound peptides were observed. These results are similar to those reported previously [30] in that zirconium dioxide was observed as being less specific for binding of phosphorylated peptides than titanium dioxide.

Zirconium dioxide has also been extensively studied as a chromatographic material for biopurifications due to its pH and thermal stability, compatibility with many buffers, as well as its mechanical strength [36–38]. Studies focusing on the removal of phosphate as a pollutant have concluded that zirconium dioxide exhibits large phosphate absorption capacity in an environmental setting [39]. The surface of zir-conium is primarily composed of exposed oxygen atoms and hydroxyl groups [38].

Lack of surface hydroxyls attributed to the manufacturing process from which the ZrO_2 was derived, has been correlated with low retention of phosphate [38]. Zirconium dioxide is amphoteric, and, depending on pH, can possess either a positive or negative charge on its surface [37]. Specifically, the isoelectric point of ZrO_2, where the surface of the resin has an overall positive charge, has been found to vary depending on the manufacturing process, and lies in the range of pH 4.9 to 8 [35]. Thus, depending on the source of the resin, the optimum binding and washing conditions can vary. An increase in phosphate absorption has been observed as the binding solution pH is lowered [39], and this binding is primarily due to interaction with surface hydroxyls [38]. While the titanium dioxide CRC out performed the zirconium dioxide tip, the titanium dioxide tip and zirconium dioxide tip performed similarly.
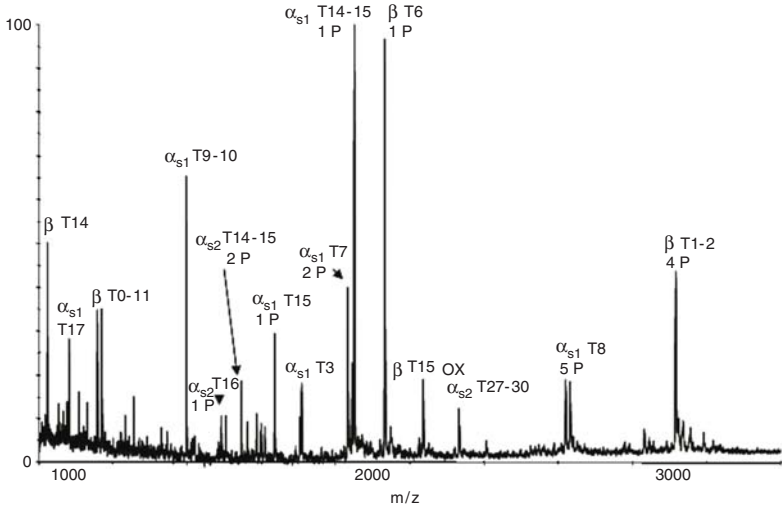
Figure 5. MALDI/MS of α- and β-casein peptides eluted from Titanium NuTip
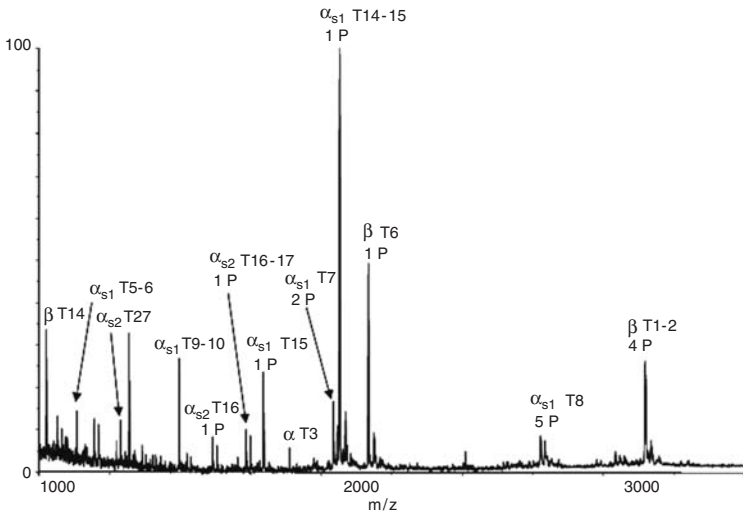
Figure 6. MALDI/MS of α- and β-casein peptides eluted from Zirconium NuTip
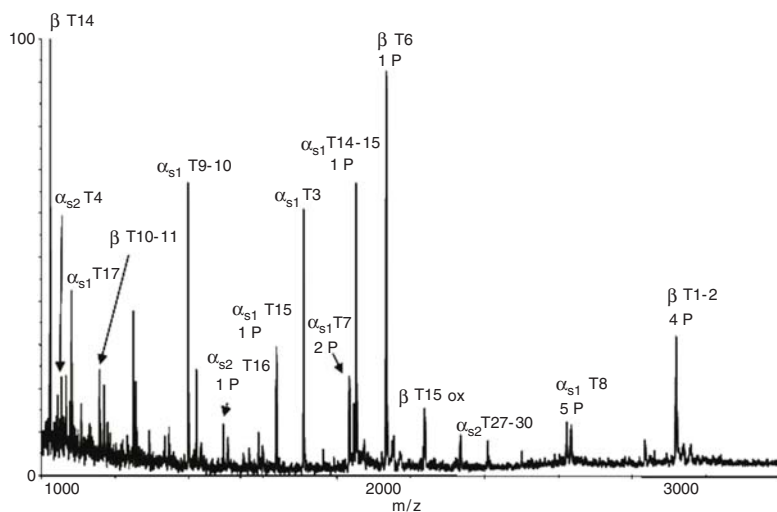
Figure 7. MALDI/MS of α- and β-casein peptides eluted from a mixed resin (TiO$_2$ & ZrO$_2$) NuTip

Combining different metal oxide resins in one tip has potential for improved recovery of phosphopeptides as well as enhanced selectivity. Therefore, a mixed resin tip (Glygen Corporation) consisting of titanium dioxide and zirconium dioxide was evaluated. The MALDI mass spectrum acquired from the mixed resin tip eluent shows seven ions corresponding in mass to phosphopeptides from the digest mixture (Figure 7). From the α-casein digest, phosphopeptides T16, T15, T14-15 (α$_{s1}$), T7, and T8 were observed while phosphopeptides T6 and T1-2 were observed from the β-casein digest. The number of non-specifically bound peptides observed in the spectrum was similar to that observed in the eluent spectra from Glygen's ZrO$_2$ and TiO$_2$ tips. However, the relative abundance of the non-specifically bound peptides is greater than was observed for either the zirconium or titanium tip.

From the MALDI/MS data it can be concluded that the titanium dioxide CRC provided the highest phosphopeptide binding with the lowest non-specific binding. The three Glygen NuTips provided similar enrichment, but with high non-specific binding when compared to the titanium dioxide CRC or the Fe-NTA CRC method. The MALDI/MS results of each enrichment method are summarized in Table 1.

In order to confirm the identity of the ions after enrichment as well as to locate the specific amino acid modified with the phosphate moiety tandem MS is necessary. Often, MALDI MS/MS results in the loss of the phosphate moiety (80 or 98 Da) making specific site determination of the phosphate difficult. A representative MALDI/MS/MS spectrum of the protonated molecule corresponding in mass to phosphorylated tryptic peptide T6 from β-casein (after titanium NuTip enrichment) is shown in Figure 8. The most abundant fragment ion observed corresponds in mass to the loss of the phosphate and water (−98 Da). A series of y ions, however, is observed which does allow for the determination of the site of phosphorylation in this peptide to Ser-35.

TABLE 1. Summary of MALDI/MS Results from Various Metal Affinity Enrichment Media

Tryptic peptide	Number of phosphate groups	CRC Fe-NTA	TiO$_2$	NuTip TiO$_2$	ZrO$_2$	TiO$_2$/ZrO$_2$	Phos-trap kit TiO$_2$
α_{S2} T16/ox.	1	X	X	X	X	X	
α_{S2} T16-17	1				X		
α_{S2} T14-15	2			X			
α_{S1} T15	1	X	X	X	X	X	X
α_{S1} T14-15	1	X	X	X	X	X	X
β T6	1	X	X	X	X	X	X
α_{S1} T8	5	X	X	X	X	X	
α_{S1} T7/ ox.	2	X	X	X	X	X	
β T2	4	X	X				
β T1-2	4	X	X	X	X	X	
#Nonspecific peptides		4	2	>10	>10	>10	>10
# Phosphopeptides		8	8	8	8	7	3
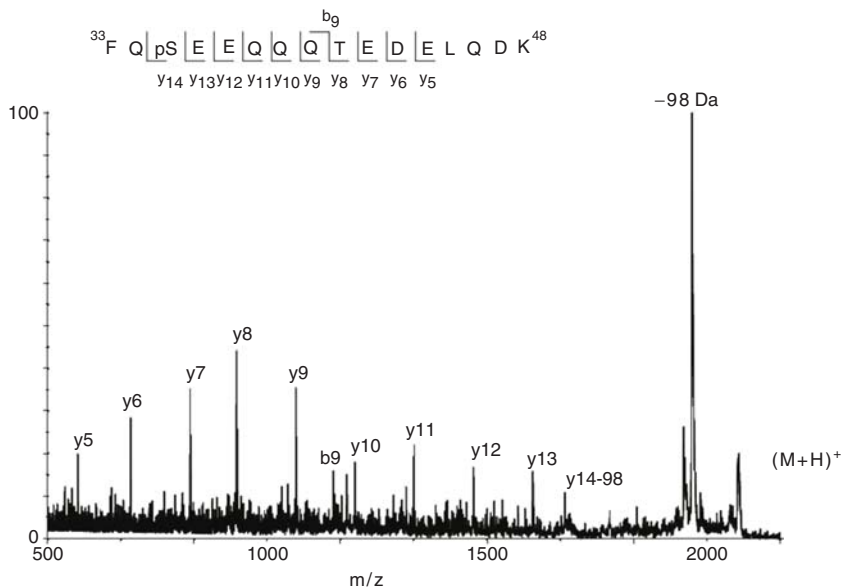
Figure 8. MALDI/MS/MS of β-T6 phosphopeptide from titanium NuTip eluent

In addition to facile cleavage of the phosphate moiety, the MALDI process suffers from suppression effects which are augmented by the negative charge of the phosphate group. To circumvent these issues, the eluents from the IMAC media were

analyzed by LC/ESI/MS/MS. The MS/MS data obtained following electrospray ionization should allow for an increase in the number of sites of phosphorylation determined because the loss of the labile phosphate group is reduced. In addition, the use of LC to separate the eluent components should reduce suppression effects. We have previously shown, however, that some phosphopeptides do not efficiently elute from Fe-NTA media [8–10]. In addition, analysis of the eluent from the titanium dioxide CRC, which contains DHB, has been reported to be problematic [30]. Therefore, LC/MS/MS analyses were performed on samples following PhosTrap and NuTip enrichment only.

Thus far, nineteen unique sites of phosphorylation have been determined using LC/MS/MS of NuTip and PhosTrap eluents. Figure 9 shows two MS/MS spectra in which the sites of phosphorylation can be identified in each peptide from α-casein. These MS/MS data confirm the masses identified by MALDI/MS as phosphorylated peptides. A summary of the results obtained from the LC/MS/MS analyses are shown in Table 2. The table shows the specific tryptic peptides and their respective phosphoamino acids.

TABLE 2. Sites of Phosphorylation Determined from LC/MS/MS Analyses following Phospho-peptide Enrichment

Tryptic peptide	# Phosphate groups	Sites of phosphorylation
α T16 (S2)	1	S 143
α T16 ox. (S2)	1	S 143
α T15 (S1)	1	S 115
α T15-16 (S2)	1	S 143
α T14-15 (S1)	1	S 115
β T6	1	S 35
β T5-6	1	S 35
α T13-15 (S2)	1	S 131
α T6-7 (S1)	3	S 41,46, 48
α T1-2 (S2)	4	S 8,9,10,16
α T7 (S2)	4	S 56,57, 58,61
β T2	4	S 15,17,18, 19
β T1-2	4	S 15,17,18 19

4. Feasibility Comparison

The cost, loading capacity, ease of use, and analysis time are all factors that determine the feasibility of an analytical method. Ni-NTA resin can be purchased for $271 for 25 ml, with each experiment requiring 30 μl of resin. The loading
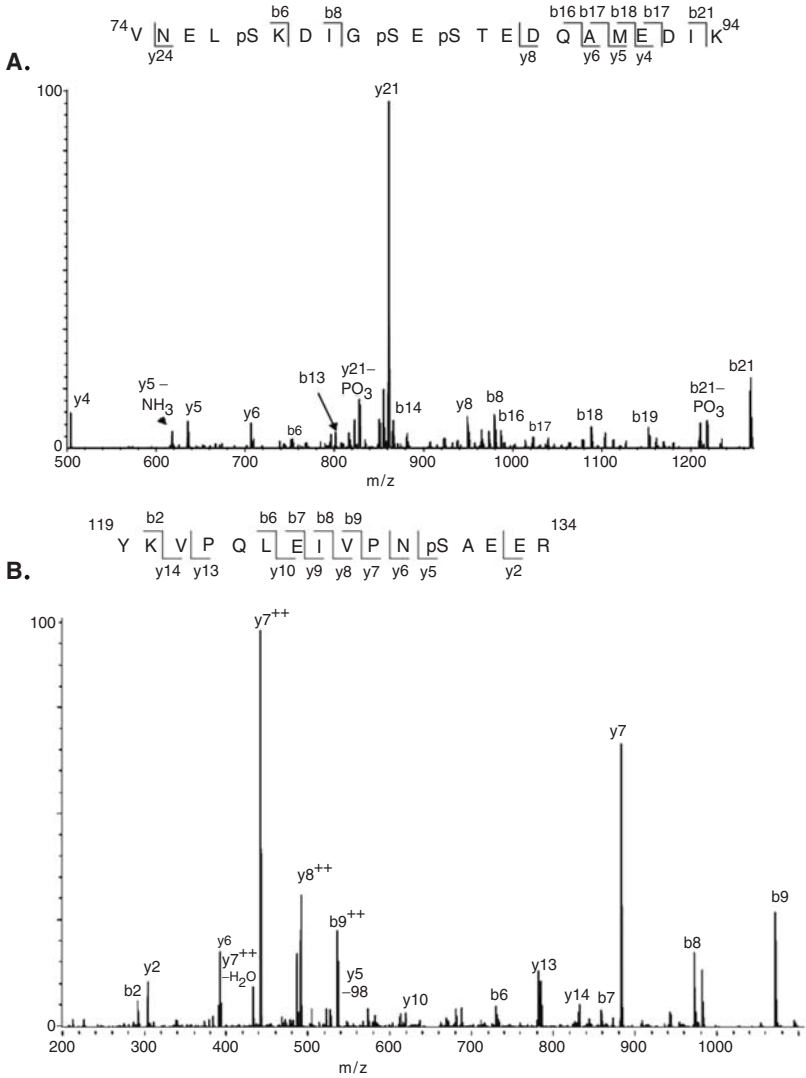
Figure 9. LC/MS/MS after PhosTrap TiO$_2$ enrichment. (A) α_{s1} T6-7 Phosphopeptide and (B) α_{s1} T14-15 Phosphopeptide

capacity of the Ni-NTA resin according to the manufacturer's literature is 20 mg/ml [18]. While the low efficiency elution from Fe-NTA media is a disadvantage, the beads can be analyzed directly by MALDI/MS. The time required to complete Fe-NTA enrichment is about 60 min.

The PhosTrap kit costs $336, and, according to the manufacturer, the limit of detection for the magnetic beads is 1 fmol to 100 nmol [40]. The analysis time is approximately 30 min, not including the time required for lyophilization which is recommended prior to LC/MS analysis. Titansphere titanium dioxide resin from GL Sciences costs $400 per 500 mg, and the analysis time is approximately 30 min. Each experiment using Titansphere resin requires 1 mg, thus approximately 500 samples can be enriched. The Titansphere eluent, however, is amenable only to MALDI/MS analysis due to the presence of large amounts of DHB (and possibly resin).

Glygen NuTips were by far the fastest enrichment method, requiring only 5–10 min per sample. In addition, these tips provide an estimated 50 μg loading capacity (according to the manufacturer) as well as a clean eluent that does not require lyophilization. The cost per pack of 96 tips is $155, which is also the least expensive up-front cost of the enrichment products. The limit of detection for the tips has previously been reported as 1 pmol [21]. From our experience, however, these tips suffer from high levels of nonspecific binding.

5.   Conclusions

Iron, titanium dioxide, and zirconium dioxide all provide enrichment of phosphopeptides from digest mixtures prior to mass spectrometric analyses. The selectivity of titanium dioxide was observed to be greater than that of iron. While Glygen Nutips were the most time efficient and required the least financial investment, these tips exhibited the highest levels of non-specific binding compared to the other enrichment methods. There was no evidence in this study that any given resin preferentially enriches mono- or multiphosphorylated peptides. In this study, the titanium CRC was more selective than Glygen's TiO_2 and ZrO_2 tips, and perhaps this difference in enrichment capability may be attributed to the effective-ness of DHB when using Titansphere resin. In our hands, the various Glygen NuTips performed very similarly in their enrichment capabilities and their level on non-specific binding. In conclusion, the Titansphere titanium dioxide resin was the most selective enrichment media when DHB was included in the loading buffer. Applicability of these methodologies to proteins with unknown sites of phospho-rylation is underway.

Acknowledgments

References

1. Hubbard, J.; Cohen, P. *Trends Biochem Sci* **1993**, *18*, 172–177.
2. Kelliher, M.A.; McLaughlin.; Witte, O.N.; Rosenberg N. *Proc Natl Acad Sci USA* **1990**, *87*, 6649–6653.
3. Daley, G.Q.; Ben-Neriah, Y. *Adv Cancer Res* **1991**, *57*, 151–184.
4. Reinders, J.; Sickmann, A. *Proteomics* **2005**, *5*, 4052–4061.
5. Andersson, L.; Porath, J. *Anal Biochem* **1986**, *154*, 250–254.
6. Osterberg, R. *Nature* **1957**, *179*, 476–477.
7. Posewitz, M.C.; Tempst, P. *Anal Chem* **1999**, *71*, 2883–2892.
8. Zhou, W.; Merrick, B.A.; Khaledi, M.G.; Tomer, K.B. *J. Am. Soc. Mass Spectrom.* **2000**, *11*, 273–282.
9. Merrick, B.A.; Zhou, W.; Martin, K.J.; Jeyarajah S.; Parker, C.E.; Selkirk, J.K.; Tomer, K.B.; Borchers, C.H. *Biochemistry* **2001**, *40*, 4053–4066.
10. Deterding, L.J.; Cutalo, J.M.; Tomer, K.B. Contribution to "Proteomic Methods for Phosphorylation Site Mapping" in Proteins and Proteomics: A Laboratory Manual (Ed. Richard Simpson), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, **2003**, pp. 597.
11. Cao, H.; Deterding, L.J.; Venable, J.D.; Kennignton, E.A.; Yates, J.R.; Tomer, K.B.; Blackshear, P.J. *Biochem J* **2006**, *394*, 285–297.
12. Andersson, L. *J Chromatogr* **1991**, *539*, 327–334.
13. Sun, X.; Chiu, J.F.; He, Q.Y. *Expert Rev Proteomics* **2005**, *2*, 649–657.
14. Chaga, G.; Andersson, L.; Porath, J. *J Chromatogr* **1992**, *627*, 163–172.
15. Muszynska, G.; Andersson, L; Porath, J. *Biochemistry* **1986**, *25*, 6850–6853.
16. Neville, D.C. A.; Rozanas, C. R.; Price, E. M.; Gruis, D. B.; Verkman, A. S.; Townsend, R. R. *Protein Sci* **1997**, *6*, 2436–2445.
17. Steinert, K.; Artz, C.; Fabis, R.; Ribbe, A. *QIAGEN News* **1996**, *5*, 12.
18. Steinert, K.; Wulbeck, M.; Ribbe, J. *QIAGEN News* **1997**, *6*, 11–15.
19. Pinske, J.W.H.; Uitto, P.M.; Hilhorst, J.J.; Ooms, B.; and Heck, A.J.R. *Anal Chem* **2004**, *76*, 3935–3943.
20. Larsen, M.R.; Thingholm, T.E.; Jensen O.N.; Roepstorff, P.; Jorgensen, T.J.D. *Mol Cell Proteomics* **2005**, *4*, 873–886.
21. Håkansson, K.; Kweon, H.K. *Anal Chem* **2006**, *78*, 1743–1749.
22. Kirwan, L.J.; Fawell, P.D.; van Bronswijk, W. *Langmuir* **1999**, *15*, 4324–4327.
23. Zhang, Q.L.; Du, L.C.; Weng, Y.X.; Wang, L.; Chen, H.Y.; Li, J.Q. *J. Phys. Chem. B*. **2004**, *108*, 15077–15083.
24. Ficarro, S.B.; McCleland, M.L.; Stukenberg, P.T.; Burke, D.J.; Ross, M.M.; Shabanowitz, J.; Hunt, D.F.; White, F.M. *Nature Biotechnol* **2002**, *20*, 301–305.
25. Brill, L.M.; Salomon, A.R.; Ficarro, S.B.; Mukheriji, M.; Stettler-Gill, M.; Peters, E.C. *Anal Chem* **2004**, *76*, 2763–2772.
26. Jensen, S.S.; Larsen, M.R. *Rapid Commun. Mass Spectrom.* **2007**, *21*, 3635–3645.
27. Hart, S.R.; Waterfield, M.D.; Burlingame, A.L.; Cramer, R. *J Am Soc Mass Spectrom* **2002**, *13*, 1042–1051.

28. Connor, P.A.; McQuillan, A.J. *Langmuir* **1999**, *15*, 2916–2921.
29. Dobson, K.D.; McQuillan, A.J. *Spectromchimica Acta Part A* **2000**, *56*, 557–565.
30. Jensen, S.S.; Larsen, M.R. *Rapid Commun Mass Spectrom* **2007**, *21*, 3635–3645.
31. Liang, S.S.; Makamba, H.; Huang, S.Y.; Chen, S.H. *J Chromatogr A* **2006**, *1116*, 38–45.
32. Zhou, H.; Tian, R.; Ye, M.; Xu, S.; Feng, S.; Pan, C.; Jiang, X.; Li, X.; Zou, H. *Electrophoresis* **2007**, *28*, 2201–2215.
33. Chen, C.T.; Chen, Y.C. *Anal Chem* **2005**, *77*, 5912–5919.
34. Liu, J.J.; Hartman, D.S.; Bostwick, J.R. *Anal Biochem* **2003**, *318*, 91–99.
35. Grun, M.; Kurganov, A.A.; Schacht, S.; Schuth, S.; Schuth, F.; Unger, K.K. *J Chromatogr A* **1996**, *740*, 1–9.
36. Nawrocki, J.; Rigney, M.P.; McCormick, A.; Carr, P.W. *J Chromatogr A* **1993**, *657*, 229–282.
37. Clausen, A.M.; Carr, P.W. *Anal Chem* **1998**, *70*, 378–385.
38. Wirth, H.J.; Hearn, M.T.W. *J Chromatogr* **1993**, *646*, 143–151.
39. Liu, H.; Sun, X.; Yin, C.; Hu, C. *J Hazard Mater* (**2007**), doi: 10.1016/j.jhazmat.2007.06.033
40. Phos-Trap Phosphopeptide Enrichment Kit Manual, Catalog Number PRT301001KT, Perkin Elmer, 2007, p. 11.

4. MOLECULAR RECOGNITION SPECIFICITY OF ANTI-3-NITROTYROSINE ANTIBODIES REVEALED BY AFFINITY- MASS SPECTROMETRY AND IMMUNOANALYTICAL METHODS

BRÎNDUŞA-ALINA PETRE, MIHAELA DRĂGUŞANU, AND MICHAEL PRZYBYLSKI[*3]

Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany

Abstract Nitration of tyrosine residues in proteins has been mainly characterised by immuno-analytical methods using anti-3-nitrotyrosine antibodies, and nitration sites and sequences have been hitherto identified only in a few cases using mass spectrometric methods. Immuno-analytical methods frequently suffer from low and poorly characterised detection specificity of anti-nitrotyrosine antibodies, while mass spectrometric methods for identification of Tyrosine nitration may be hampered by low levels of modification, and by possible changes of structure and proteolytic degradation of proteins introduced by the nitration. Moreover, no detailed, molecular characterisation of the specificity of anti-3-nitrotyrosine antibodies has been reported. In this study we describe a molecular study of the recognition specificities and affinities of two commercially available, monoclonal anti-nitrotyrosine antibodies by affinity-mass spectrometry, using different 3-nitrotyrosine containing peptides. Tyrosine-nitrated and non-nitrated substrate peptides of prostacyclin synthase (PCS), an enzyme inactivated by nitration of the active site Tyr-430 residue, were synthesised by solid-phase peptide synthesis (SPPS), purified by reversed phase-high performance liquid chromatography (RP-HPLC) and characterised by electrospray (ESI) and matrix-assisted laser desorption-ionisation (MALDI) mass spectrometry. Binding affinities and specificities of PCS peptides with different Tyr-nitration sites and sequence mutations adjacent to Tyr-430 were determined by evaluation of anti-nitrotyrosine antibodies using an affinity-mass spectrometry approach, compared to immuno-analytical determination using dot-blot and ELISA. The results showed that the antibodies may discriminate in the recognition of peptides with different N-terminal adjacent sequences to the nitrotyrosine residues, depending on the type of immunogen employed. A quantitative ELISA estimation was developed for the determination of antibody binding by Tyrosine-nitrated peptides.

1. Introduction

Tyrosine nitration is a covalent protein modification resulting from the addition of a nitro- (NO_2) group onto one of the two equivalent *ortho* carbons of the aromatic ring of tyrosine residues. Tyrosine nitration may be the result of an excessive generation of nitric oxide (NO) or its by-products as well as reactive oxygen species

---

[*]Corresponding author. Professor Dr. Michael Przybylski, Department of Chemistry, University of Konstanz; Tel:++49-7531-882249; Fax: ++49-7531-3097; e-mail: Michael.Przybylski@uni-konstanz.de

(ROS). Several mechanisms for tyrosine nitration have been proposed, the two of which most widely advanced to occur *in vivo* involve (i), the formation of peroxynitrite (PN), and (ii), catalytic nitration via heme-peroxidases [1]. Nitration thus may occur under physiological conditions, but may be substantially enhanced under pathophysiological conditions. Several proteins found to be modified by nitration, have been associated with different disorders such as lung infection [2], inflammation [3,4], Parkinson's [5] and Alzheimer's disease [6], and diabetes [7], all of which have been associated with enhanced oxidative stress.

Specific nitrated tyrosine residues or tyrosine-nitration motifs in a protein should be more specific oxidative biomarkers, than merely the characterisation of overall nitro-tyrosine levels. Several methods have been employed to detect and identify protein tyrosine nitration in both *in vitro* and *in vivo* studies. Anti-3-NT antibodies permit detection by immunohistochemistry [8], Western blot [9], or enzyme-linked immunosorbent assay (ELISA) [10,11]. However, these methods do not identify which specific tyrosine residue(s) have been nitrated; moreover, they are frequently limited by low detection specificities of commercial anti-3-NT-antibodies. Only little information on the characterization and properties of different antibodies against 3-nitrotyrosine is available from previous publications [12], and the identification and localization of the sites of nitration in proteins remains a significant methodological challenge. Of the currently available methods, only mass spectrometry (MS) is able to specifically locate 3-nitro-tyrosine sites in proteins due to its sensitivity and molecular specificity. High resolution Fourier-transform- ion cyclotron resonance mass spectrometry (FTICR-MS), in combination with immuno-analytical procedures has been recently developed as a powerful tool for the unequivocal, molecular identification of tyrosine-nitrations, such as upon peroxy-nitrite treatment at the active site Tyr-430 residue of bovine prostacyclin synthase (PCS) [13]. Detailed structural studies were performed with Tyr-nitrated peptides, prepared by solid phase peptide synthesis (SPPS), using FTICR-MS in combination with CD spectroscopy and structure modelling [14].

In the present study we have characterized two commercially available antibodies raised against different immunogens containing 3-nitrotyrosine (Table 1), with regard to their recognition specificities and affinities to different Tyr-nitrated peptides of prostacycline synthase. Tyr- nitrated and non-nitrated PCS peptides were compared using (i), conventional dot blot and ELISA, and (ii), a newly developed immunoaffinity – mass spectrometric approach as shown in Figure 1. Affinity – mass spectrometry methods, in combination with proteolytic digestion, have been previously developed and employed in our laboratory (i), for the identification of antigen- epitopes (epitope- excision and – extraction-MS) [15,16]; and (ii), in an affinity- proteomics approach which enables direct protein identification from biological material with unprecedented selectivity [17,18]. The affinity- MS procedure (see Figure 1) is based on the binding of antigen- peptides to an antibody immobilized on a micro-column; upon removal of unbound material by washings steps, the affinity- bound peptides are specifically dissociated from the column by acidification and analyzed by MS. Using this procedure we have characterized the recognition specificity of anti 3-NT antibodies by mass spectrometry.
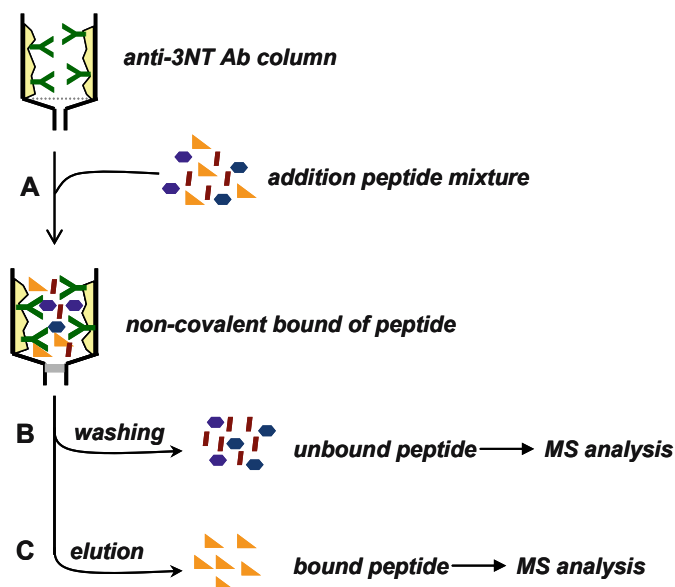
Figure 1. Schematic representation of an immuno-affinity experiment using immobilized antibody column. (A) Binding of peptides to the immobilized Ab column; (B) removal of unbound peptides; (C) dissociation of the antigen-antibody complex followed by washing steps and storage

2.   Materials and Methods

2.1.   ANTIBODIES

The anti-nitrotyrosine-antibodies employed in this work were obtained from Chemicon International (Bubendorf, Switzerland) and from Santa Cruz Antibodies (Santa Cruz, USA); characteristics of both antibodies as given by the anufacturers are summarized in Table 1.

The following commercially available reagents were used: N-α-Fmoc-3-Nitrotyrosine: Bachem (Bubendorf, Sitzerland); N-α-Fmoc protected amino acids, NovaBiochem TGA (Darmstadt, Germany); benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluoro-phosphate (PyBOP): NovaBiochem; Roti®-Block (10 × concentrate): Roth (Karlsruhe, Germany); ECL solution: Amersham Biosciences (München, Germany); activated CH-Sepharose 4B, poly-oxyethylenesorbitanmono-laureate (Tween20): Sigma (St.Louis, USA); α-cyano-4-hydroxycinnamic acid (HCCA): Bruker Daltonics (Bremen, Germany). Water was taken from an in-house high purity water system (Milli Q plus, Millipore, 18.2 MΩ). All reagents and solvents were of analytical grade or highest available purity.

TABLE 1. Investigated monoclonal antibodies against 3-nitro-tyrosine

Supplier	Antibody details
Chemicon International (1)	Mouse Anti-Nitrotyrosine Monoclonal Antibody (**MAB5404**) **Concentration:** 1 µg/µl **Immunogen:** Nitrated KLH (Keyhole-limpet hemocyanin)
Santa Cruz (2)	Mouse Monoclonal Nitrotyrosine (**39B6**) **Concentration:** 0.2 µg/µl **Immunogen:** 3-(2-(4-hydroxy-3-nitrophenyl)acetamido) propionic acid-bovine serum albumin conjugate

2.2. PEPTIDE SYNTHESIS

The 3-nitro-Tyrosine and Tyrosine containing peptides were synthesised on a semi-automated peptide synthesizer (EPS-221, Intavis, Langenfeld, Germany) by SPPS, using Fmoc/t-butyl protection chemistry with all chemicals and reagents of analytical grade or highest available purity. To synthesise peptides with C-terminal free carboxylate the TGA resin was employed with 40 min coupling and 5 min deprotection times in 20% piperidine solution in DMF.

2.3. HPLC PURIFICATION

Analytical RP-HPLC was performed on a BioRad (München, Germany) HPLC system using a 5 µm Nucleosil$^®$ 300-7 C_{18} column as a stationary phase. Linear gradient elution of peptides dissolved in eluent A (20% B, 5 min; 70% B, 55 min) was performed with eluent A (0.1% TFA in MilliQ water) and eluent B (0.1% TFA in acetonitrile-MilliQ water 80:20, v/v) at a flow rate of 1 ml/min. Peaks were detected at 365 nm for nitrated peptides and 220 nm for non-nitrated peptides. The crude products were purified on a 250 × 20 mm preparative C_{18} column Grom-Sil 120 OSD-4 HE, 10 µm (Grom, Herrenberg, Germany) at a flow rate of 10 ml/min, using identical eluent conditions as for analytical HPLC.

2.4. DOT BLOT ANALYSIS

Peptide solutions of 1-2 µl aliquots (1 µg/µl) in PBS (pH=7.4) were spotted on nitrocellulose (NC) membranes and dried for 5 min at room temperature. NC membranes were blocked for 2 h by using Roti®-Block solution in Milli-Q water. Membranes were then probed for 1 h with mouse monoclonal antibody against 3-nitro-tyrosine (1:2,500) in PBS-Tween buffer. The membranes were then washed three times in PBS-Tween buffer and probed again with a goat anti-mouse horse-radish peroxidase conjugate (Sigman; 1:5,000) in PBS-Tween for 45 min. After washing the membrane five times in washing buffer, the immuno-positive spots were visualised by using ECL- Plus system (Amersham Pharmacia).

2.5. PREPARATION OF AFFINITY COLUMNS AND AFFINITY- MASS SPECTROMETRY

The monoclonal anti 3-nitrotyrosine antibody (100 µg/µl) was dissolved in coupling buffer (0.2 M NaHCO$_3$, 0.5 M NaCl, pH 8.3). The solution was added to the appropriate amount of dry NHS-activated 6-aminohexanoic acid-coupled Sepharose (1 g sepharose- swell in approximately 3 ml of coupling buffer) and the coupling reaction was performed for 2 h at 25°C under vigorous stirring. The reaction mixture was then loaded into a micro-column and extensive washing steps (0.2 M NaOAc, 0.5 M NaCl, pH = 4.0) and blocking steps (0.1 M ethanolamine, 0.5 NaCl, pH 8.3) were carried out. Affinity experiments of the column prepared in this manner were performed by binding of the antigen peptides (solution prepared in PBS buffer, pH = 7.4) to the antibody column and subsequent reaction for 2 h. The unbound peptides were removed by washing steps using PBS buffer as described above, and the immune complex was dissociated by addition of 0.1% TFA (aqueous solution, pH = 2). The washing and elution fractions were collected, lyophilized, redissolved in 0.1% TFA and desalted using ZipTip procedure before MALDI-MS analysis.

MALDI-TOF (matrix assisted laser desorption-ionisation -time-of-flight) mass spectrometry was performed with a Bruker BiFlex-DE mass spectrometer equipped with a Scout MALDI source and video system, a nitrogen UV laser (337 nm), and dual channel plate detector. Sample preparation was performed with 0.7 µl of a saturated solution of HCCA in acetonitrile /0.1% TFA (2:1), which was mixed on the MALDI target with 1 µl of the sample solution. Spectra were recorded at an accelerating voltage of 25 kV and were averaged over forty single laser shots. Calibration was performed with a standard peptide mixture within an m/z range, 1,000–5,000.

2.6. ELISA

All assays were performed in 96-well CovaLink-NH plates which were activated for 1 h using disuccinimidyl suberate (DSS) solution. After three times washing with distilled water, 100 µl/well of the peptide solution (dilutions prepared in PBS buffer containing 5 mM NaHPO$_4$ and 150 mM NaCl, pH = 7.4) was added onto the plate and incubated over night. The non-specific adsorption sites were then blocked by treatment with 150 µl blocking buffer for 1 h (0.1 M ethanolamine). After washing steps, the plates were coated for 2 h at room temperature with 100 µl of the monoclonal 3-nitrotyrosine antibodies (dilution 1:2,500). The wells were washed four times with 200 µl/well of 0.05% Tween-20 (v/v) in PBS, and 100 µl of peroxidase labelled goat anti-mouse IgG diluted 5,000 times in 5% BSA was added to each well and incubated for 1 h at room temperature. The wells were washed three times with 200 µl/well PBS-Tween and two times with 200 µl/well 0.05 M sodium phosphate-citrate buffer, pH = 5, 100 µl of *o*-phenylenediamine dihydrochloride (OPD) in sodium phosphate-citrate buffer at 1 mg/ml, and 2 µl of 30% hydrogen-peroxide

per 10 ml of substrate buffer was added. After 5–10 min, the absorbance at $\lambda = 490$ nm was measured on a Wallac 1420 Victor[2] ELISA Plate reader (Perkin Elmer, Boston, USA).

3.   Results and Discussion

Two monoclonal 3-nitro-tyrosine antibodies (see Table 1) were characterised by dot blot, immunoaffinity chromatography – mass spectrometry and ELISA using different synthetic prostacycline synthase (PCS) peptides. A series of nitrated and non-nitrated tyrosine containing peptides within the domain PCS (419–432) were synthesised by SPPS as described in Materials and Methods. The use of 3-nitro-tyrosine within the standard Fmoc and side-chain protection chemistry (t.-butyl, trityl) provided high coupling yields and homogeneities already for the crude peptides following deprotection. All peptides were subjected to final purification by preparative RP-HPLC (see Table 2) and characterised by MALDI-TOF mass spectrometry. The UV-MALDI –TOF mass spectra showed abundant molecular ions for all peptides and a series of specific photochemical fragmentations only in the case of the Tyr-nitrated peptides was observed. Peaks occurring with 16 and 32 mass units lower than those corresponding to the nitrated peptides resulted from products formed by prompt fragmentation caused by the immediate loss of oxygen from the nitro group to yield a *nitroso* species, followed by loss of second oxygen to form a *nitrene*- type ion, as previously described [19]. The mass spectrometric characterisation of nitrated peptides by high resolution MALDI-FTICR- MS enabled the detection of the specific photochemical decomposition products due to the nitro group in 3-nitrotyrosine containing peptides [14].

Dot blot is a straightforward, simple technique for peptide and protein detection in which the samples are spotted directly onto a membrane through circular templates. The method is based on the antigen-antibody recognition by using a first antibody against the membrane- immobilized antigen, and a second label-conjugated antibody for detection of the first antibody. The PCS peptides were spotted on a nitrocellulose membrane, and the unspecific sites were blocked with Roti®-Block solution. After washing steps with PBS-Tween buffer, the membrane was incubated with anti-3-NT antibody for 1 h at room temperature. After further washing steps, a second enzyme-conjugated antibody (horse radish peroxidase (HRP) goatanti-mouse IgG) which recognizes the Fc (crystallizable fragment) fragment of the first antibody was added and incubated for 45 min. A mixture of ECL-solutions was added in order to develop the membrane and expose it on a film.

Figure 2 shows the dot blot binding responses of the 3-NT- antibodies to different PCS peptides. We observed that the MAB5404 (Chemicon; **1**) gave an intense positive response only for the authentic Tyr-430 –nitrated peptide containing the actual Tyr-nitration site in PCS (see Figure 2A); the 39B6 antibody (Santa Cruz; **2**) gave an intense response for the same peptide, however additional weaker positive

TABLE 2. HPLC and MALDI-TOF-MS data of synthetic PCS peptides

Peptide code	Sequence	HPLC retention time (min)[c]	MALDI-TOF[d] m/z [M+H]$^+_{calc/exp}$
PCS-Y[a]	DFYKDGKRLKN**Y**SL-OH	18.52	1,746.91/1,747.72
PCS-NO$_2$[a]	DFYKDGKRLKN**Y(NO$_2$)**SL-OH	18.37	1,791.90/1,792.96
PCS-NO$_2$(RR)[b]	DF**Y(NO$_2$)**KDG**R**RLKNYSL-OH	20.11	1,819.90/1,822.13
PCS-NO$_2$(Ala)[a]	DFYKDGKR**AAA**Y(NO$_2$)**A**L-OH	18.05	1,633.79/1,634.71
PCS-F[a]	DFYKDGKRLKN**F**SL-OH	20.47	1,730.92/1,732.71
PCS[a]	LKN**Y(NO$_2$)**-NH$_2$	12.25	581.64/582.97

[a]Tyr-430 of Prostacycline synthase.
[b]Tyr-421 of Prostacycline synthase, Lys^{425} was replaced by an Arg.
[c]RP-HPLC column: Vydac C$_{18}$ (250 × 4 mm, 300Å, 5 µm); 365 nm for nitro-Tyr containing peptides and 220 nm for Tyr/Phe containing peptides.
[d]MALDI-TOF mass spectra were recorded with a Bruker Biflex^{TM} linear TOF mass spectrometer.

responses were observed for the PCS peptide nitrated at Tyr-421 (**3**) and for the peptide (**4**) in which the amino acid residues around the nitrated Tyr-430 were replaced by alanine residues (see Figure 2B). Since no response was observed for the two non-nitrated Tyr- and Phe-containing peptides (**1**, **5**) of PCS (419–432) fragment, we can conclude that both antibodies show high specificity for nitro-tyrosine.

PCS-Y	(1)	^{419}DFYKDGKRLKNY**SL**^{432}
PCS-NO$_2$	(2)	^{419}DFYKDGKRLKN**Y(NO$_2$)**SL^{432}
PCS-RR	(3)	^{419}DF**Y(NO$_2$)**KDG**RR**LKNYSL^{432}
PCS-Ala	(4)	^{419}DFYKDGKR**AAAY(NO$_2$)A**L^{432}
PCS-F	(5)	^{419}DFYKDGKRLKN**F**SL^{432}
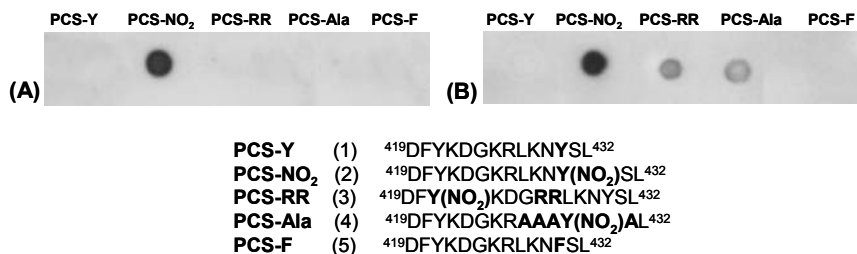
Figure 2. Dot blots analysis of (A) MAB5404 (Chemicon) antibody (1) and (B) 39B6 (Santa Cruz) antibody (2) to synthetic PCS (419–432) peptides

In order to ascertain the results of the dot blot experiments by a molecular recognition response, the 3-NT-antibody interactions were characterised by using the immunoaffinity- MS illustrated in Figure 1. Two antibody columns were prepared by coupling the mouse antibody MAB5404 (**1**) and the mouse antibody 39B6 (**2**) to NHS-activated Sepharose, as described in Materials and Methods. The PCS peptides were dissolved in PBS buffer, pH 7.4 and added to the antibody columns. After 2 h incubation, the unbound peptides were removed, the column was washed

several times with PBS buffer, and the peptide-antibody complexes were dissociated using 0.1% TFA. The first washing fractions containing unbound peptides, and the elution fractions which contain the peptides specifically bound to the 3-NT antibody column were collected, lyophilized and analysed by mass spectrometry.

Corresponding binding results for a mixture of the PCS (Y-430) (**1**) and the PCS-(Y(NO$_2$)-430) peptide (**2**) to the monoclonal MAB5404 (Chemicon) antibody column (**1**) are shown in Figure 3. The affinity-mass spectrometric data showed that only the Tyr-430- nitrated PCS- peptide bound to the MAB5404 (Chemicon) antibody column. In contrast, affinity–MS experiments performed with several PCS (419–432) peptides that (i), contained a nitration site at Tyr-421, or (ii) contained an alanine sequence adjacent to the Tyr-430 residue did not show binding affinity. These results suggest high recognition specificity for this nitration site that had been shown previously to cause specific inactivation of prostacyclin synthase upon treatment with peroxynitrite [13].

Furthermore, these results indicate that the affinity of the MAB5404 antibody (**1**) (raised against nitrated keyhole limpet hemocyanin carrier protein) is not merely directed by a nitrated tyrosine residue, but also by a recognition sequence encompassing amino acid adjacent to the nitration site.

Somewhat different results regarding a possible sequence-specific motif for recognition of nitro-Tyrosine were obtained for the 39B6 (Santa Cruz) antibody (**2**). The affinity-mass spectrometry data for this antibody generally showed binding affinity for the nitrated PCS peptides, consistent with the results obtained by the dot blot experiments. In order to compare the binding in affinity-MS analyses, equimolar mixtures (200 µmol) of nitrated PCS peptides were applied to the 3-NT antibody column. The MALDI- mass spectrum of the elution fraction of the equimolar mixture of nitrated Tyr-430 PCS peptides **2** and **4** (see Figure 4) showed affinity for both peptides, albeit with lower affinity for the peptide containing the
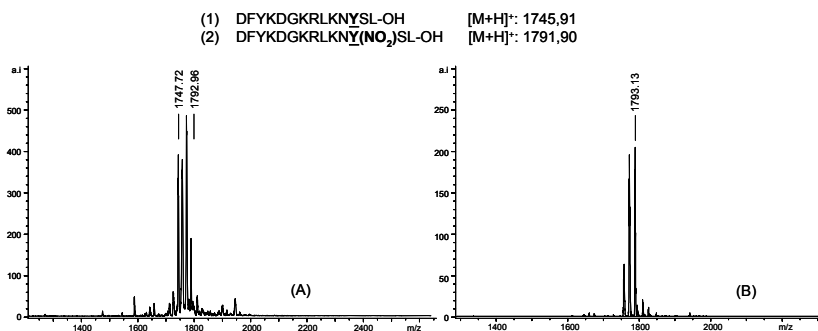
Figure 3. Affinity-mass spectrometry analysis of PCS (419–432) peptides **1** and **2** (nitrated at Tyr-430) with the monoclonal antibody MAB5404 (**1**); MALDI-TOF-MS of the washing (A) and elution (B) fractions.

alanine-sequence mutation adjacent to the nitration site. Similar results with the 39B6 antibody (**2**) column were obtained by affinity-MS of PCS peptides nitrated at different Tyr residues, such as using peptide **3**. Thus, the elution fraction of an equimolar mixture of nitrated Tyr-430 PCS-Ala (**4**) and nitrated PCS-RR (**3**); nitro-Tyr-421) indicate comparable affinities for both peptides (see Figure 5).

The binding affinity observed was significantly lower than that of the Tyr-430 nitrated peptide **2**; however, due to the photochemical decomposition of the nitro group in Tyr-nitrated peptides under UV-MALDI-MS conditions, no quantitative determination was carried out from the mass spectrometric data.
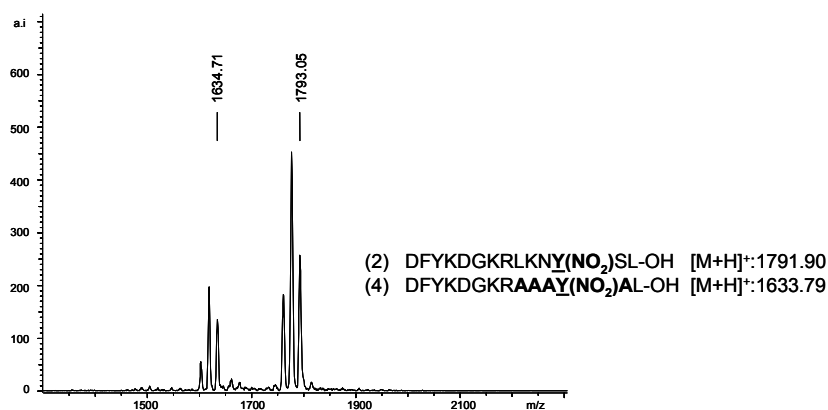
Figure 4. Affinity-mass spectrometry analysis: (MALDI-TOF-MS of elution fraction) of the nitrated Tyr-430 PCS- peptides 2 and 4 to the 39B6 antibody (**2**)
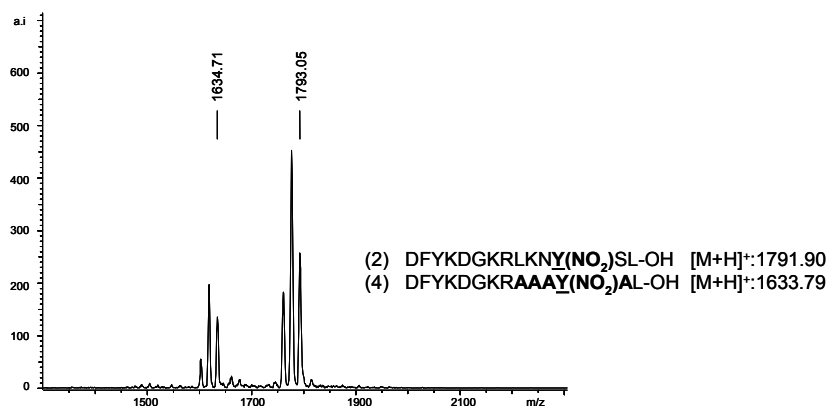
Figure 5. Affinity-mass spectrometry analysis: (MALDI-TOF-MS of elution fraction) of the mutated Tyr-nitrated peptides PCS-Ala (**4**) and nitrated Tyr-421 containing peptide, PCS-RR (**3**) with the 39B6 antibody (**2**)

A quantitative estimation of the bindining affinities of the PCS peptides was performed by ELISA as an established technique for comparison of biomolecular affinities. In this study, only a comparative semi-quantitative analysis was aimed at, with regard to the observation that antibody binding may be influenced by a hindered of access of 3-nitrotyrosine residues, and by the structural environment of a nitro-tryosine motif [10]. Corresponding ELISA experiments with the two anti-nitrotyrosine antibodies are compared in Figure 6 for the series of nitrated PCS peptides. Due to the substantial differences in surface adsorption of small peptides, and to minimize unspecific adsorption, covalent attachment was employed with CovaLink-NH plates [20], in which a polystyrene surface is grafted with secondary amino groups to orientate the immobilized peptide on the surface in order to expose the epitope. In a first step of the ELISA determination, the NH-linker was activated using an excess of disuccinimidyl suberate (DSS) and then the diluted peptides solution was added and incubated overnight.

After suitable washing and blocking steps, the CovaLink-NH plate was then incubated with the anti- 3-NT antibody, and horse radish peroxidase (HRP) goat anti-mouse IgG was used as a detection antibody with OPD/H_2O_2 used as the substrate for HRP as described in Materials and Methods. For comparison of the affinities of the PCS peptides, the concentrations needed for obtaining an $OD_{450} = 1$ were determined (see Figure 6).

According to the ELISA data with the MAB5404 antibody (see Figure 6A), only the Tyr-430 nitrated PCS- peptide **2** showed significant affinity to this antibody (lowest amount of PCS-NO_2 peptide to obtain an $OD_{450} = 1$: 0.37 µM). This result is consistent with the affinity-MS data and confirmed the requirement of a specific sequence motif around the 3-nitrotyrosine for antibody interaction. Thus, sequence mutation by alanine residues adjacent to the Tyr-430 nitration site, and change of the nitration position (Tyr-421), resulted in drastically diminished antibody binding. In contrast, the 39B6 antibody revealed significant binding for all nitrated PCS peptides (Figure 6B). The results of the ELISA determinations revealed some differences to the affinity-mass spectrometric analyses in which the Tyr-430 nitrated PCS-peptide (**2**) showed lower affinity than the alanine-mutated peptide **4** (lowest amount of PCS-NO_2 peptide to obtain an $OD_{450} = 1$: 0.095 µM). This effect might be explained by the increase in hydrophobicity introduced by the alanine mutation causing increased exposure of the nitrated Tyr-430. In the ELISA determination with the 39B6 antibody (**2**), significant affinity was also found for the short, nitrated PCS tetra-peptide, PCS (427–430) fragment which was not detected with the MAB5404 antibody (**1**).

The differences found in binding affinities of Tyr-nitrated peptides by affinity-mass spectrometry in comparison to dot blot and ELISA determinations, may be attributed to the different types of antigens used for producing the anti-3-NT antibodies have been raised. Thus, the 39B6 antibody (Santa Cruz) was raised against a 3-(2-(4-hydroxy-3-nitrophenyl) acetamido) propionyl-bovine serum albumin conjugate in which the (4-hydroxy)-3-nitrophenyl group is more exposed and freely accessible, compared to the nitrotyrosine residue in the polypeptide chain of
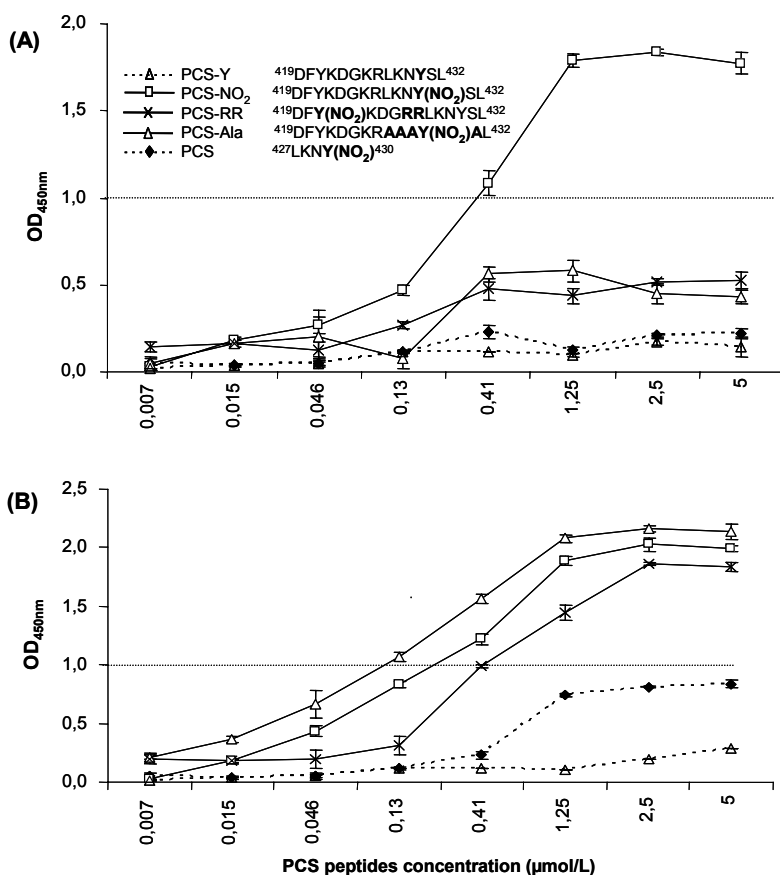
Figure 6. ELISA determination of the binding of the MAB5404 (Chemicon) antibody (A) and the 39B6 (Santa Cruz) antibody (B) to synthetic prostacycline synthase peptides

the nitrated KLH (Keyhole-limpet hemocyanin) carrier protein, employed as antigen for the MAB5404 (Chemicon) antibody. Hence, the 39B6 antibody showed generally high affinity to single 3-nitrotyrosine amino acid residues, in contrast to the epitope-peptide sequence motif determining the recognition specificity of the MAB5404 antibody. This recognition motif has been recently characterised in our laboratory to require a set of positively charged residues (Lys, Arg) at the N-terminal adjacent to the nitrotyrosine residue (Dragusanu et al.; *J. Peptide Sci.,* submitted for publication).

4. Conclusions

In complementing the conventional immuno-analytical techniques such as dot blot and ELISA, affinity- mass spectrometry is shown in this study as a powerful,

molecular tool for characterising recognition specificities of antibodies, especially for protein structure modifications such as tyrosine nitration. The comparison of the three methods essentially provided consistent results to reveal differences in binding affinities and specificities by the anti-nitrotyrosine antibodies, suggesting that antibody binding may be influenced by the peptide structure adjacent to the nitrotyrosine modification. The capability of anti-nitrotyrosine antibodies to discriminate between nitrotyrosine in different environments in proteins may be useful for producing antibodies to specific motifs containing tyrosine residues, and for the development of highly specific biomarkers.

Acknowledgments

References

1. Brennan ML, Wu W, Fu X, et al. (2002) A tale of two controversies: defining both the role of peroxidases in nitrotyrosine formation in vivo using eosinophil peroxidase and myeloperoxidase-deficient mice, and the nature of peroxidase-generated reactive nitrogen species. *J Biol Chem* **277:** 17415–17427.
2. Hopkins N, Cadogan E, Giles S, Bannigan J, McLoughlin P. (2003) Type 2 nitric oxide synthase and protein nitration in chronic lung infection. *J Pathol* **199:** 122–129.
3. Petersson AS, Steen H, Kalume DE, Caidahl K, Roepstorff P. (2001) Investigation of tyrosine nitration in proteins by mass spectrometry. *J Mass Spectrom* **36:** 616–625.
4. Ischiropoulos H. (2003) Biological selectivity and functional aspects of protein tyrosine nitration. *Biochem Biophys Res Commun* **305:** 776–783.
5. Good PF, Hsu A, Werner P, Perl DP, Olanow CW. (1998) Protein nitration in Parkinson's disease. *J Neuropathol Exp Neurol* **57:** 338–342.
6. Castegna A, Thongboonkerd V, Klein JB, Lynn B, Markesbery WR, Butterfield DA. (2003) Proteomic identification of nitrated proteins in Alzheimer's disease brain. *J Neurochem* **85:** 1394–1401.
7. Turko IV, Li L, Aulak KS, Stuehr DJ, Chang JY, Murad F. (2003) Protein tyrosine nitration in the mitochondria from diabetic mouse heart. Implications to dysfunctional mitochondria in diabetes. *J Biol Chem* **278:** 33972–33977.
8. ter Steege J, Buurman W, Arends JW, Forget P. (1997) Presence of inducible nitric oxide synthase, nitrotyrosine, CD68, and CD14 in the small intestine in celiac disease. *Lab Invest* **77:** 29–36.
9. Hinson JA, Michael SL, Ault SG, Pumford NR. (2000) Western blot analysis for nitrotyrosine protein adducts in livers of saline-treated and acetaminophen-treated mice. *Toxicol Sci* **53:** 467–473.
10. Khan J, Brennand DM, Bradley N, Gao B, Bruckdorfer R, Jacobs M. (1998) 3-Nitrotyrosine in the proteins of human plasma determined by an ELISA method. *Biochem J* **330(Pt 2):** 795–801.

11. Franze T, Weller MG, Niessner R, Poschl U. (2003) Enzyme immunoassays for the investigation of protein nitration by air pollutants. *Analyst* **128:** 824–831.
12. Franze T, Weller MG, Niessner R, Poschl U. (2004) Comparison of nitrotyrosine antibodies and development of immunoassays for the detection of nitrated proteins. *Analyst* **129:** 589–596.
13. Schmidt P, Youhnovski N, Daiber A, et al. (2003) Specific nitration at tyrosine 430 revealed by high resolution mass spectrometry as basis for redox regulation of bovine prostacyclin synthase. *J Biol Chem* **278:** 12813–12819.
14. Petre BA, Youhnovski N, Lukkari J, Weber R, Przybylski M. (2005) Structural characterisation of tyrosine-nitrated peptides by ultraviolet and infrared matrix-assisted laser desorption/ ionisation Fourier transform ion cyclotron resonance mass spectrometry. *Eur J Mass Spectrom (Chichester, Eng)* **11:** 513–518.
15. Suckau D, Kohl J, Karwath G, et al. (1990) Molecular epitope identification by limited proteolysis of an immobilized antigen-antibody complex and mass spectrometric peptide mapping. *Proc Natl Acad Sci USA* **87:** 9848–9852.
16. McLaurin J, Cecal R, Kierstead ME, et al. (2002) Therapeutically effective antibodies against amyloid-beta peptide target amyloid-beta residues 4–10 and inhibit cytotoxicity and fibrillogenesis. *Nat Med* **8:** 1263–1269.
17. Macht M, Marquardt A, Deininger SO, Damoc E, Kohlmann M, Przybylski M. (2004) "Affinity-proteomics": direct protein identification from biological material using mass spectrometric epitope mapping. *Anal Bioanal Chem* **378:** 1102–1111.
18. Stefanescu R, Iacob RE, Damoc EN, et al. (2007) Mass spectrometric approaches for elucidation of antigenantibody recognition structures in molecular immunology. *Eur J Mass Spectrom (Chichester, Eng)* **13:** 69–75.
19. Sarver A, Scheffler NK, Shetlar MD, Gibson BW. (2001) Analysis of peptides and proteins containing nitrotyrosine by matrix-assisted laser desorption/ionization mass spectrometry. *J Am Soc Mass Spectrom* **12:** 439–448.
20. www.nuncbrand.com.

Part II

LIPIDOMICS

5. MAPPING AND SEQUENCING OF GANGLIOSIDES FROM ANENCEPHALY BY ELECTROSPRAY IONIZATION HIGH CAPACITY ION TRAP MASS SPECTROMETRY

CRISTINA MOSOARCA[1], ŽELJKA VUKELIĆ[2], AND ALINA D. ZAMFIR[1,3]

[1]*Mass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania*

[2]*Department of Chemistry and Biochemistry, Faculty of Medicine, University of Zagreb, Croatia*

[3]*"Aurel Vlaicu" University of Arad, Romania*

Abstract. Congenital malformation referred to as anencephaly is a neural tube defect that occurs when the cephalic end of the neural tube fails to close, resulting in the absence of a major portion of the brain, skull, and scalp. Infants with this disorder are born without a forebrain—the largest part of the brain consisting mainly of the cerebrum, which is responsible for thinking and coordination. Although some individuals with anencephaly may be born with a rudimentary brain stem, the lack of a functioning cerebrum permanently rules out the possibility of ever gaining consciousness. Gangliosides (GGs) are sialylated glycosphingolipids present in the cell plasma membrane, responsible for the modulation of the cell signal transduction events. GGs act as receptors of interferon, epidermal growth factor, nerve growth factor and are differently expressed in various pathological states of central nervous system (CNS) acting as biomarkers of CNS disorders. In this study a native GG mixture extracted and purified from a histopathologically-defined anencephalic fetal brain remnant was analyzed by electrospray ionization high capacity ion trap mass spectrometry. Structural data upon disease-associated species were collected by multiple stage collision-induced dissociation of the molecular ions. As a control a native GG mixture from a normal fetal brain in the same developmental stage was used. Comparative screening and sequencing revealed the differential expression of the GGs in aberrant vs. healthy tissue and provided accurate information upon the structure of several anencephaly-associated species.

1. Introduction

Anencephaly, a severe congenital malformation, is frequently associated with other malformations not involving the central nervous system. Diagnosis is usually antenatal and expectation of life is very short [1,2]. Anencephaly is a cephalic disorder that results from a neural tube defect that occurs when the cephalic (head) end of the neural tube fails to close, usually between the 23rd and 26th day of pregnancy, resulting in the absence of a major portion of the brain, skull, and scalp. Infants with this disorder (Figure 1) are born without a forebrain, the largest part of the brain consisting mainly of the cerebral hemispheres (which include the isocortex, which is responsible for higher level cognition, i.e., thinking).

Figure 1. Morphoanatomic structure of an anencephalic fetal head

The remaining brain tissue is often exposed – not covered by bone or skin. Infants born with anencephaly are usually blind, deaf, unconscious, and unable to feel pain. Although some individuals with anencephaly may be born with a rudimentary brainstem, which controls autonomic and regulatory function, the lack of a functioning cerebrum permanently rules out the possibility of ever gaining consciousness [3]. Reflex actions such as breathing and responses to sound or touch may occur.

The disorder is one of the most common disorders of the fetal central nervous system. Anencephaly can often be diagnosed before birth through an ultrasound examination. The maternal serum alpha-fetoprotein (AFP screening) and detailed fetal ultrasound [2] can be useful for screening for neural tube defects such as spina bifida or anencephaly. There are many false diagnoses for anencephaly, as it is not a common diagnosis, often confused with exencephaly or microcephaly. Also, sometimes a false prognosis stating that an anencephalic baby can live for years is given, but this cannot occur because the brain is open, meaning that infection sets in rapidly. The anencephalic brain is also usually very disorganised on a cellular level.

There is no cure or standard treatment for anencephaly and the prognosis for affected individuals is poor. Most anencephalic babies do not survive birth, accounting for 55% of non-aborted cases. If the infant is not stillborn, then he or she will usually die within a few hours or days after birth from cardiorespiratory arrest. In almost all cases anencephalic infants are not aggressively resuscitated since there is no chance of the infant ever achieving a conscious existence. Instead, the usual clinical practice is to offer hydration, nutrition and comfort measures and to "let nature take its course". Artificial ventilation, surgery (to fix any co-existing congenital defects), and drug therapy (such as antibiotics) are usually regarded as futile efforts. Some clinicians and medical ethicists even view the provision of nutrition and hydration as medically futile, arguing that euthanasia is morally and clinically appropriate in such cases. In the United States, approximately 1,000 to 2,000 babies are born with anencephaly each year. Female babies are more likely to be affected by the disorder.

About 95% of women who learn that they will have an anencephalic baby choose to have an abortion. The cause of anencephaly is unknown. Neural tube defects do not follow direct patterns of heredity [4], and recent animal models indicating a possible association with deficiencies of the transcription factor TEAD2 [4]. Studies show that a woman who has had one child with a neural tube defect such as anencephaly, has about a 3% risk to have another child with a neural tube defect. This risk can be reduced to about 1% if the woman takes high dose (4 mg/day) of folic acid before and during pregnancy. It is known that women taking certain medication for epilepsy and women with insulin dependent diabetes have a higher chance of having a child with a neural tube defect. Genetic counseling is usually offered to women at a higher risk of having a child with a neural tube defect to discuss available testing.

Recent studies have shown that the addition of folic acid to the diet of women of child-bearing age may significantly reduce, although not eliminate, the incidence of neural tube defects. Therefore, it is recommended that all women of child-bearing age consume 0.4 mg of folic acid daily, especially those attempting to conceive or who may possibly conceive. It is not advisable to wait until pregnancy has begun, since by the time a woman knows she is pregnant, the critical time for the formation of a neural tube defect has usually already passed. A physician may prescribe even higher dosages of folic acid (4 mg/day) for women who have had a previous pregnancy with a neural tube defect.

Gangliosides [5–9] are acid glycosphingolipids widely distributed in most vertebrate tissues and fluids. They are present in mammalian milk, where they are almost exclusively associated with the membrane fraction of the fat globule. In human milk, the content and individual distribution of gangliosides changes during lactation, GD3 being the most abundant ganglioside in colostrum, while in mature milk, GM3 is the major individual species. Gangliosides function as "unintended" target receptors for bacterial adhesion in specific tissues. After oral administration, they can be putative decoys that interfere with pathogenic binding in the intestine, this being the main mechanism by which these compounds can prevent infection. Ganglioside-supplemented infant formula has been reported to modify the intestinal ecology of preterm newborns, increasing the Bifidobacteria content and lowering that of Escherichia coli. In addition, the influence of dietary gangliosides on several parameters related to the development of intestinal immune system, such as cytokine and intestinal IgA production, has also been described in animal models. Recently, the influence of GM3 and GD3 on dendritic cell maturation and effector functionalities has also been reported, suggesting a role for these milk gangliosides, especially GD3, in modulating the process of oral tolerance during first stages of life. Dietary gangliosides may have an important role in the modification of intestinal microflora and the promotion of intestinal immunity development in the neonate, and consequently in the prevention of infections during early infancy. [8] Gangliosides (GGs) are sialylated glycosphingolipids present in the cell plasma membrane, responsible for the modulation of the cell signal transduction events.

GGs act as receptors of interferon, epidermal growth factor, nerve growth factor and are differently expressed in various pathological states of central nervous system (CNS) acting as biomarkers of CNS disorders [10–17].

In this study a native GG mixture extracted and purified from a hystopatologically-defined anencephalic fetal brain remnant was analyzed by electrospray ionization high capacity ion trap mass spectrometry. Structural data upon disease-associated species were collected by multiple stage collision-induced dissociation of the molecular ions.

The anencephalic cerebral remnant, as a primitive brain structure, represents a model to study the ganglioside involvement in induction of aberrant brain developement. We have chosen human anencephalic brain-like formed structures, as a potentially useful physiological system to study the ganglioside invovement in brain developement and etiology of brain developmental arrest. Here we present the detailed ganglioside composition and quantification in histologically defined human brain-resembling anencephalic tissue structures in comparison to respective normal fetal brain regions .

A difference in the expression of ganglio-series gangliosides with GM1a core was found between anencephalic and normal fetal brain, with less expression of GM1a and GD1a in anencephaly compared with normal fetal brain, in which these gangliosides dominate. Small amounts of GM1b were detected in fetal brain whereas only traces were found in anencephalic brain. Lactosamine-containing gangliosides were present in fetal and in anencephalic brain as alpha 2-3 as well as alpha 2-6 sialylated nLcOse4Cer structures. A heterogeneous group of neolacto-series gangliosides was expressed in anencephalic brain in both the monosialo- and presumed disialoganglioside range. These findings demonstrate a significant change in ganglioside pattern in anencephaly where the process of cell differentiation and maturation has been severely disturbed [11].

In anencephalic tissue, GM1b, GD1alpha, nLM1 and nLD1 were expressed at a higher rate in relation to normal tissue. It can be demonstrated that the anencephalic cerebral remnant, as a primitive brain structure, represents a naturally-occurring model to study the ganglioside involvement in induction of aberrant brain development.

2.　　Materials and Methods

2.1.　MASS SPECTROMETRY

Mass spectrometry was performed on a High Capacity Ion Trap Ultra (HCT Ultra, PTM discovery) mass spectrometer from Bruker Daltonics, Bremen, Germany. HCT MS is interfaced to a PC running the CompassTM 1.2. integrated software package, which includes the HystarTM 3.2.37 module for instrument controlling and spectrum acquisition, Esquire 6.1.512 and Data Analysis 3.4.179 modules for

storing the ion chromatograms and processing the MS data. All mass spectra were acquired using on-line (–) microESI infusion.

2.2. GANGLIOSIDE SAMPLE

The native mixture of gangliosides from a hystopatologically-defined anencephalic fetal brain remnant (36 g.w.) was extracted and purified as described in detail elsewhere [14,18,19]. The GG stock solution at approximately 1 mg/ml was prepared by dissolving the dried material in MeOH. For MS analysis, the solution was evaporated (digital SpeedVac Concentrator SPD111V-230, Thermo Electron Corporation, Milford, MA, USA) and the resulting dry substrate was dissolved in MeOH (HPLC grade, Merck, Darmstadt, Germany) to the final concentration of approximately 20 pmol/µl (calculated for an average molecular weight of 2,000).

3.   Results and Discussions

In the past decade, the studies aiming at determination of GG composition, quantity, distribution and cell surface expression were almost exclusively conducted on chromatographic, immunochemical and immunohistochemical methods [18,19]; however, the information acquired by these techniques is based merely on comparisons and, because of the detection boundaries, it is restricted to the major species. Therefore, fast atom bombardment (FAB) mass spectrometry (MS) was introduced as the first MS method in human brain GG analysis, and showed its ability to provide closer insights into the expression pattern at nanomolar sensitivities [20].

Nowadays, modern protocols for detection and sequencing of different types of glycosphingolipids (GSLs) including GGs and sulfated GlcA-GSLs, based on either matrix assisted laser desorption/ionization (MALDI) MS [21–23] or electrospray ionization (ESI) in combination with different analyzers are available [24–27]. TLC-MALDI MS [28–30], liquid chromatography [31] and capillary electrophoresis (CE) ESI MS [32–34] couplings were also developed for efficient separation of complex mixtures followed by direct MS analysis. These innovative methods allowed not only detection of individually isolated fractions but also a reproducible mapping and MS/MS sequencing of single components in complex tissue extracts that are accessible only in very low amounts (pico to attomoles).

In earlier work, to improve the ESI process, increase the experiment throughput, sensitivity and reproducibility, we have implemented fully automated chip-based nanoESI MS in glycolipidomics [35,36] and multistage sequencing [37].

In the present work we applied the novel protocol based on HCT multistage sequencing for strucutral analysis of GG from fetal anencephalic brain-like residue.

Screening of the anencephaly GG mixture indicated the dominance of structures exhibiting short oligosaccharide chains, with sialylation degree 1 and 2. The most abundant species are GM3, GM1, GD1 and GD2. A notable aspect is that the dominant structures mono- and disialotetraoses are bearing different ceramide residues and that several structures appear associated to anencephaly. Under optimized solution and instrumental conditions, the abundances of the ions corresponding to mono- and disialylated GM and GD components dominate while asialo species were not found. We and others [38,39] have demonstrated before that a direct correlation between ganglioside sialylation degree and brain developmental stage exists. Higher sialylation extent was found specific for incipient developmental stages. Consequently, this elevated expression of sialylated structures represents a marker of brain development stagnation, which occurs in anencephaly. Additionally, GG chains modified by labile attachments such as Fuc or *O*-Ac were earlier reported as associated to the tissue in its later fetal developmental phase [38,39]. Present data render interesting information also from this point of view. While several modified structures were earler detected in normal fetal brain, only two O-Ac species exhibiting (d18:1/20:0) and (d18:0/22:0) ceramide compositions were identified in anencephalic remnant.

For structural elucidation GM1 (d18:1/18:0) detected as singly charged ion at *m/z* 1544.48 and GD2 (d18:1/18:0), singly charged ion at *m/z* 1,383.00 were submitted to multistage CID MS experiments (Figures 2–5).
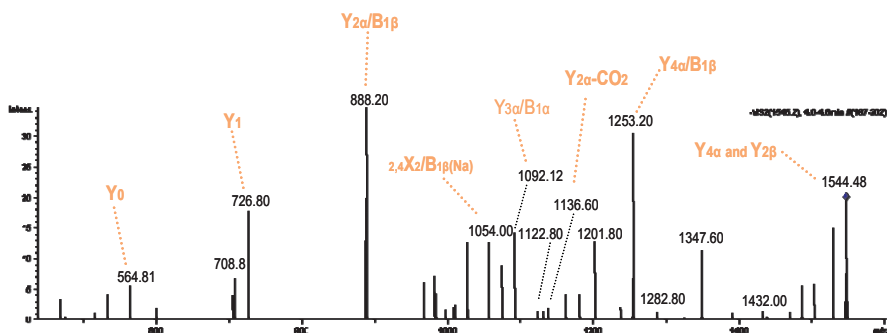
Figure 2. Negative ESI HCT MS^2 of the singly charged ion at *m/z* 1,544.48 corresponding to the GM1 (d18:1/18:0) ganglioside species detected in the GG mixture from anencephalic remnant. Assignment of the product ions is according to Ref [40,41]
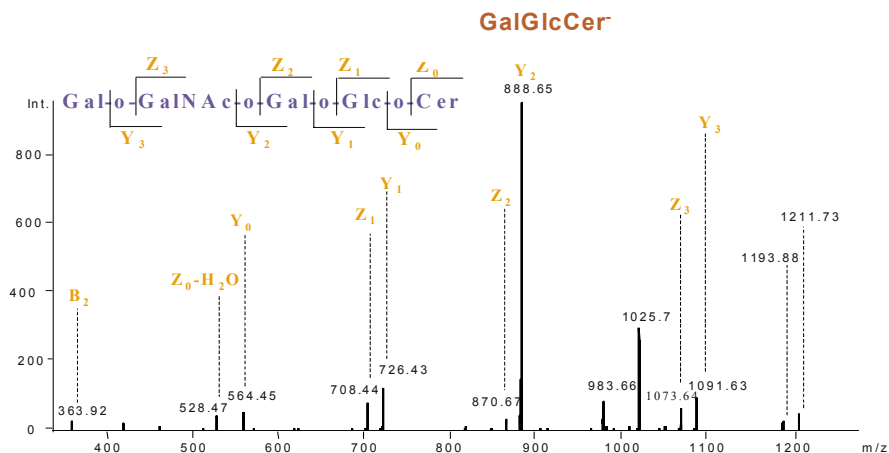
Figure 3. Negative ESI HCT MS^3 of the singly charged ion at *m/z* 1,253.82 corresponding to the desialylated GM1 fragment ion detected in the MS^2 mode Assignment of the product ions is according to Ref [40,41]
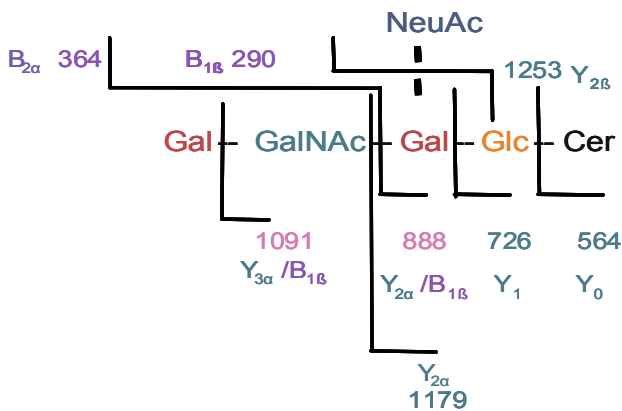
Figure 4. Fragmentation pathway of the GM1a species under low energy CID conditions in the negative ion mode
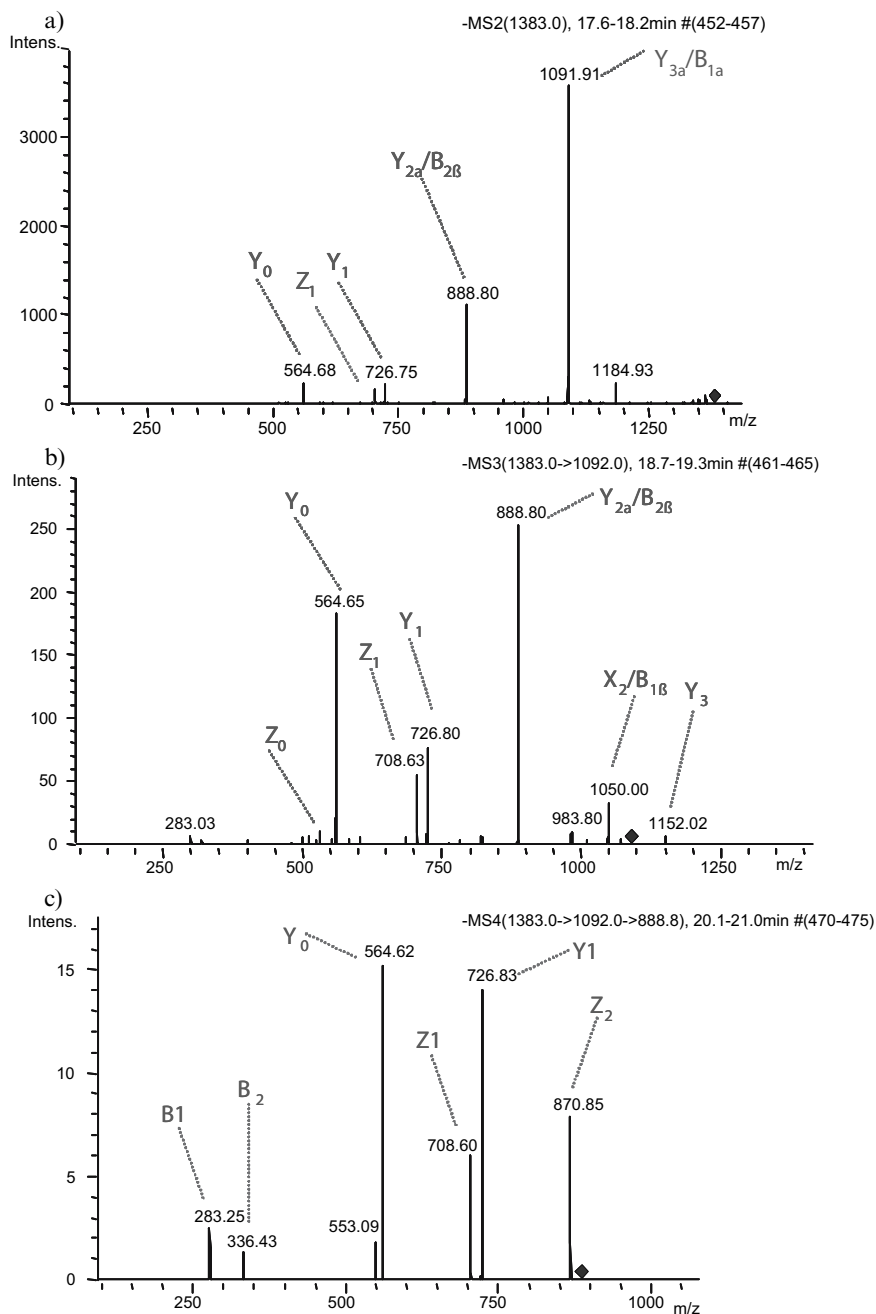
Figure 5. Negative ESI HCT multistage CID MS of the singly charged ion at *m/z* 1,383.00 corresponding to the GD2 (d18:1/18:0) ganglioside species detected in the GG mixture from anencephalic remnant: (a) MS[2]; (b) MS[3]; (c) MS[4] Assignment of the product ions is according to Ref [40,41]

4.   Conclusions

We developed here a novel strategy in glycolipidomics based on electrospray sample infusion followed by MS detection and multistage sequencing of ganglioside components in a native mixture extracted from anencephalic fetal brain.

ESI MS screening of the native GG mixture from anencephalic fetal brain remnant revealed a series of structures exhibiting shorter oligosaccharide chains as compared to healthy controls in the same developmental stage. GM1 and GD2 were found as dominant structures, a feature which is in agreement with the data collected by thin layer chromatography and immunochemical analyses.

Data upon these disease-associated species, obtained by multiple stage collision-induced dissociation of the molecular ions, indicated the structures of the sugar chains as well as the composition of the ceramide moieties. Moreover, the presence of the GM1a structural isomer could be discovered by this method.

Applied ganglioside nomenclature

Gangliosides and the precursor glycosphingolipids are abbreviated according to the system introduced by Svennerholm in 1980 [42] and the recommendations of IUPAC-IUB Commission on Biochemical Nomenclature [43] as follows: LacCer, Galβ4Glcβ1Cer; GA2, Gg$_3$Cer, GalNAcβ4Galβ4Glcβ1Cer; GA1, Gg$_4$Cer, Galβ3GalNAcβ4Galβ4Glcβ1Cer; nLc$_4$Cer, Galβ4GlcNAcβ3Galβ4Glcβ1Cer; Lc$_4$Cer, Galβ3GlcNAcβ3Galβ4Glcβ1Cer; GM3, II^3-α-Neu5Ac-LacCer; GD3, II^3-α-(Neu5Ac)$_2$-LacCer; GT3, II^3-α-(Neu5Ac)$_3$-LacCer; GM2, II^3-α-Neu5Ac-Gg$_3$Cer; GD2, II^3-α-(Neu5Ac)$_2$-Gg$_3$Cer; GM1a or GM1, II^3-α-Neu5Ac-Gg$_4$Cer; GM1b, IV^3-α-Neu5Ac-Gg$_4$Cer; GalNAc-GM1b, IV^3-α-Neu5Ac-Gg$_5$Cer; GD1a, IV^3-α-Neu5Ac,II^3-α-Neu5Ac-Gg$_4$Cer; GD1b, II^3-α-(Neu5Ac)$_2$-Gg$_4$Cer; GT1b, IV^3-α-Neu5Ac,II^3-α-(Neu5Ac)$_2$-Gg$_4$Cer; GQ1b, IV^3-α-(Neu5Ac)$_2$,II^3-α-(Neu5Ac)$_2$-Gg$_4$Cer; nLM1 or 3'-nLM1, IV^3-α-Neu5Ac-nLc$_4$Cer; LM1 or 3'-isoLM1, IV^3-α-Neu5Ac-Lc$_4$Cer; nLD1, disialo-nLc$_4$Cer

Acknowledgements

References

1. J.G. Joó, A. Beke, C. Papp, et al, Neural tube defects in the sample of genetic counselling, Prenat. Diagn. (2007) 27 912–921.
2. M. Cedergren, A. Selbing, Detection of fetal structural abnormalities by an 11-14-week ultrasound dating scan in an unselected Swedish population, Acta obstetricia et gynecologica Scandinavica 85 (2006) 912–915.
3. B. Merker, Consciousness without a cerebral cortex: a challenge for   neuroscience and medicine, The Behavioral and brain sciences 30 (2007) 63–81.
4. K.J. Kaneko, M.J. Kohn, C. Liu, M.L. Depamphilis, Transcription factor TEAD2 is involved in neural tube closure, Genesis 45 (2007) 577–587.
5. S. Liang, M. Wang, R.I. Tapping, V. Stepensky, H.F. Nawar, M. Triantafilou, K. Triantafilou, T.D. Connell, G. Hajishengallis, Ganglioside GD1a is an essential coreceptor for Toll-like receptor 2 signaling in response to the B subunit of type IIb enterotoxin, J. Biol. Chem. 282 (2007) 7532–7542.
6. M. Saito, K. Sugiyama, Tissue-specific expression of c-series gangliosides in the extraneural system, Biochim. Biophys. Acta 1474 (2000) 88–92.
7. L. Svennerholm, Identification of the accumulated ganglioside, Adv Genet. 44 (2001) 33–41.
8. I. Kracun, H. Rosner, V. Drnovsek, Z. Vukelić, C. Cosovic, M.Trbojevic-Cepe, M. Kubat, Gangliosides in the human brain development and aging, Neurochem. Int. 20 (1992) 421-431.
9. S. Ngamukote, M. Yanagisawa, T. Ariga, S. Ando, R.K. Yu, Developmental changes of glycosphingolipids and expression of glycogenes in mouse brains, J. Neurochem. 103 (2007) 327–341.
10. M. Saito, R.F. Mao, R. Wang, C. Vadasz, M. Saito, Effects of gangliosides on ethanol-induced neurodegeneration in the developing mouse brain, Alcohol Clin. Exp. Res. 31 (2007) 665–674.
11. T. Okada, M. Wakabayashi, K. Ikeda, K. Matsuzaki, Formation of toxic fibrils of Alzheimer's amyloid beta-protein-(1-40) by monosialoganglioside GM1, a neuronal membrane component., J. Mol. Biol. 371 (2007) 481–489.
12. M. Zappia, L. Crescibene, D. Bosco, G. Arabia, G. Nicoletti, A. Bagalà, L.   Bastone, I.D. Napoli, M. Caracciolo, S. Bonavita, A.Di. Costanzo, A. Gambardella, A. Quattrone, Anti-GM1 ganglioside antibodies in Parkinson's disease,.Acta Neurol. Scand.106 (2002) 54–57.
13. Z. Vukelić, S. Kalanj-Bognar, M. Froesch, L. Bîndila, B. Radić, M. Allen, J. Peter-Katalinić, A.D. Zamfir, Human gliosarcoma-associated ganglioside composition is complex and distinctive as evidenced by high-performance mass spectrometric determination and structural characterization, Glycobiology 17(2007) 504–515.
14. K.M.Hedberg, R.Mahesparan, T.A. Read, B.B. Tysnes, F. Thorsen, T. Visted, R. Bjerkvig, P. Fredman, The glioma-associated gangliosides 3'-isoLM1, GD3 and GM2 show selective area expression in human glioblastoma xenografts in nude rat brains, Neuropathol. Appl. Neurobiol. 27 (2001) 451–464.
15. A. Markowska-Woyciechowska, A. Bronowicz, M. Ugorski, E. Gamian, P. Jabłoński, Study on ganglioside composition in brain tumours supra- and infratentorial, Neurol. Neurochir. Pol. 34 (2000) 24–30.
16. M. Potapenko, G.V. Shurin, J. de León, Gangliosides as immunomodulators, Adv. Exp. Med. Biol. 601 (2007) 195–203.
17. S. Modak, N.K. Cheung, Disialoganglioside directed immunotherapy of neuroblastoma, Cancer Invest. 25 (2007) 67–77.
18. J.Müthing, Analyses of glycosphingolipids by high-performance liquid chromatography, Methods. Enzymol. 312 (2000) 45–64.
19. J. Müthing, TLC in structure and recognition studies of glycosphingolipids, Methods. Mol. Biol. 76 (1998) 183–195.

20. H. Egge, J. Peter-Katalinić, G. Reuter, R. Schauer, R. Ghidoni, S. Sonnino, G. Tettamanti, Analysis of gangliosides using fast atom bombardment mass spectrometry, Chem. Phys. Lipids 37 (1985) 127–141.
21. S.B. Levery, Glycosphingolipid structural analysis and glycosphingo-lipidomics, Methods Enzymol. 405 (2005) 300–369.
22. P.B. O'Connor, B.A. Budnik, V.B. Ivleva, P. Kaur, S.C. Moyer, J.L. Pittman, C.E. Costello, A high pressure matrix-assisted laser desorption ion source for Fourier transform mass spectrometry designed to accommodate large targets with diverse surfaces, J. Am. Soc. Mass Spectrom. 15 (2004) 128–132.
23. P.B. O'Connor, E. Mirgorodskaya, C.E. Costello, High pressure matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for minimization of ganglioside fragmentation, J. Am. Soc. Mass Spectrom. 13 (2002) 402–407.
24. Z. Vukelić, M. Zarei, J. Peter-Katalinić, A.D. Zamfir, Analysis of human hippocampus gangliosides by fully-automated chip-based nanoelectrospray tandem mass spectrometry, J. Chromatogr. A. 1130 (2006) 238–245.
25. Z.C. Tsui, Q.R. Chen, M.J. Thomas, M. Samuel, Z.A. Cui, Method for profiling gangliosides in animal tissues using electrospray ionization-tandem mass spectrometry, Anal. Biochem. 341 (2005) 251–258.
26. W. Metelmann, J. Peter-Katalinić, J. Müthing, Gangliosides from human granulocytes: a nano-ESI QTOF mass spectrometry fucosylation study of low abundance species in complex mixtures, J. Am. Soc. Mass Spectrom. 12 (2001) 964–973.
27. Z. Vukelić, A.D. Zamfir, L. Bindila, M. Froesch, J. Peter-Katalinić, S. Usuki, R.K. Yu, Screening and sequencing of complex sialylated and sulfated glycosphingolipid mixtures by negative ion electrospray Fourier transform ion cyclotron resonance mass spectrometry, J. Am. Soc. Mass Spectrom. 16 (2005) 571–580.
28. V.B. Ivleva, L.M. Sapp, P.B. O'Connor, C.E. Costello, Ganglioside analysis by thin-layer chromatography matrix-assisted laser desorption/ionization orthogonal time-of-flight mass spectrometry, J. Am. Soc. Mass Spectrom. 16 (2005) 1552–1560.
29. V.B. Ivleva, Y.N. Elkin, B.A. Budnik, S.C. Moyer, P.B. O'Connor, C.E. Costello, Coupling thin-layer chromatography with vibrational cooling matrix-assisted laser desorption/ionization Fourier transform mass spectrometry for the analysis of ganglioside mixtures, Anal. Chem. 76 (2004) 6484–6491.
30. K. Dreisewerd, J. Müthing, A. Rohlfing, I. Meisen, Z. Vukelić, J. Peter-Katalinić, F. Hillenkamp, S. Berkenkamp, Analysis of gangliosides directly from thin-layer chromato-graphy plates by infrared matrix-assisted laser desorption/ionization orthogonal time-of-flight mass spectrometry with a glycerol matrix, Anal. Chem. 77 (2005) 4098-4107.
31. L.K. Sorensen, A liquid chromatography/tandem mass spectrometric approach for the determination of gangliosides GD3 and GM3 in bovine milk and infant formulae, Rapid. Commun. Mass Spectrom. 20 (2006) 3625-3633.
32. A.D. Zamfir, J. Peter-Katalinić, Capillary electrophoresis-mass spectrometry for glyco-screening in biomedical research, Electrophoresis 25 (2004) 1949–1963.
33. A.D. Zamfir, Z. Vukelić, J. Peter-Katalinić, A capillary electrophoresis and off-line capillary electrophoresis/electrospray ionization-quadrupole time of flight-tandem mass spectrometry approach for ganglioside analysis, Electrophoresis 23 (2002) 2894-2903.
34. D.D. Ju, C.C. Lai, G.R. Her, Analysis of gangliosides by capillary zone electrophoresis and capillary zone electrophoresis-electrospray mass spectrometry, J. Chromatogr. A. 779 (1997) 195–203.
35. A.D. Zamfir, L. Bîndila, N. Lion, M. Allen, H.H. Girault, J. Peter-Katalinić, Chip electro-spray mass spectrometry for carbohydrate analysis, Electrophoresis 26 (2005) 3650–3673.
36. A.D. Zamfir, Z. Vukelić, L. Bîndila, J. Peter-Katalinić, R. Almeida, A. Sterling, M. Allen, Fully-automated chip-based nanoelectrospray tandem mass spectrometry of gangliosides from human cerebellum, J. Am. Soc. Mass Spectrom. 15 (2004) 1649–1657.

37. A.D. Zamfir, Z. Vukelić, A. Schneider, E. Sisu, N. Dinca, A. Ingendoh, A novel approach for ganglioside structural analysis based on electrospray multiple-stage mass spectrometry, J. Biomol. Tech. 18 (2007) 188–193.
38. C. Stoll, Y. Alembik, B. Dott, Associated malformations in cases with neural tube defects, Genet. Couns. 18 (2007) 209–215.
39. Z. Vukelić, W. Metelmann, J. Müthing, M. Kos, J. Peter-Katalinić, Anencephaly: structural characterization of gangliosides in defined brain regions, Biol. Chem. 382 (2001) 259–274.
40. B. Domon, C.E. Costello, A systematic nomenclature for carbohydrate fragmentations in FAB-MS spectra of glycoconjugates, Glycoconjugate J. 5 (1988) 397–409.
41. C.E. Costello, P. Juhasz, H. Perreault, New mass-spectral approaches to gangliosides structure determination, Progr. Brain Res. 101 (1994) 45–61.
42. L. Svennerholm, Ganglioside designation, Adv. Exptl. Med. Biol. 125 (1980) 125–211.
43. IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCNB). Eur. J. Biochem. (1998) 257– 293

Part III

GLYCOMICS

6. STRUCTURAL ANALYSIS OF CHONDROITIN SULFATE DISACCHARIDES BY ELECTROSPRAY IONIZATION HIGH CAPACITY ION TRAP MASS SPECTROMETRY

ADINA MURESAN[1], MIRELA GALUSCA[1], DANIELA G. SEIDLER[2], NICOLAE DINCA[1], AND ALINA D. ZAMFIR[1,3*1]

[1]*Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania*
[2]*Institute for Physiological Chemistry and Pathobiochemistry, University of Muenster, Germany*
[3]*Mass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania*

Abstract. Decorin is a small leucine-rich repeat proteoglycan linked to a single chondroitin or dermatan sulfate (CS/DS) glycosaminoglycan (GAG) chain depending on tissue expressed. Determination of molecular characteristics of CS/DS, is essential for understanding decorin functions. CS/DS analysis includes the identification of glycoforms obtained after detachment from coreprotein followed by collecting information upon epimerization, sulfate content and its distribution. For this purpose, one of the most efficient methods is electrospray ionization (ESI) tandem mass spectrometry (MS/MS). Here we present an approach for structural investigation of human skin decorin CS disaccharides based on combining a specific enzyme digestion with negative ion mode ESI MS-screening followed by fragmentation analysis by collision-induced dissociation (CID) MS/MS. The obtained results showed clearly that human decorin encompasses oversulfated CS motifs in which the site of oversulfation is at glucuronic acid. Moreover, it was possible to demonstrate that the presence of regular and/or irregular GAG domains, which play a critical biological role, can be directly documented by ion sequencing in tandem MS experiments.

Keywords: Electrospray, tandem mass spectrometry, decorin, glycosamino-glycans, chondroitin sulfate, structural analysis

1. Introduction

Proteoglycans are heavily glycosylated proteins attached to a particular type of polysaccharides called glycosaminoglycans (GAGs). In the connective tissue proteoglycans function as joint lubricants and structural components. In other tissues, they mediate the adhesion of cells to the extracellular matrix and bind factors that stimulate cell proliferation [1–3].

---

*Corresponding author. Plautius Andronescu Str. 1, RO-300224, Timisoara, Romania; Tel: +40-256-494413; Fax: +40-256-204698; e-mail: zamfir@uav.ro

The intrinsic properties of proteoglycans are determined primarily by the glyco-saminoglycan chains. GAGs consist of disaccharide repeating units which contain an uronic acid and a derivate of an amino sugar, either glucosamine or galactosamine. At least one of the sugars in the repeating unit has a negatively charged carbo-xylate or sulfate group. The major glycosaminoglycans in animals are chondroitin sulfate (CS), keratan sulfate (KS), heparin (HS), dermatan sulfate (DS) and hyaluronan (HA) (Figure 1).

D-Glucuronate   N-Acetyl-D-GlucosaminL-Iduronate   N-Acetyl-D-galactosamine-4- sulfate
(HYALURONATE)   **(DERMATAN SULFATE)**

D-Glucuronate   N-Acetyl-D-Galactosamine-4-   L-Iduronate-2-sulfate   N-sulfo-D-
glucosamine sulfate   -6-sulfate
(CHONDROITIN 4- AND 6- SULFATE)   (HEPARIN SULFATE)

D-Galactose   N-Acetyl-D-Glucosamine
(KERATAN SULFATE)
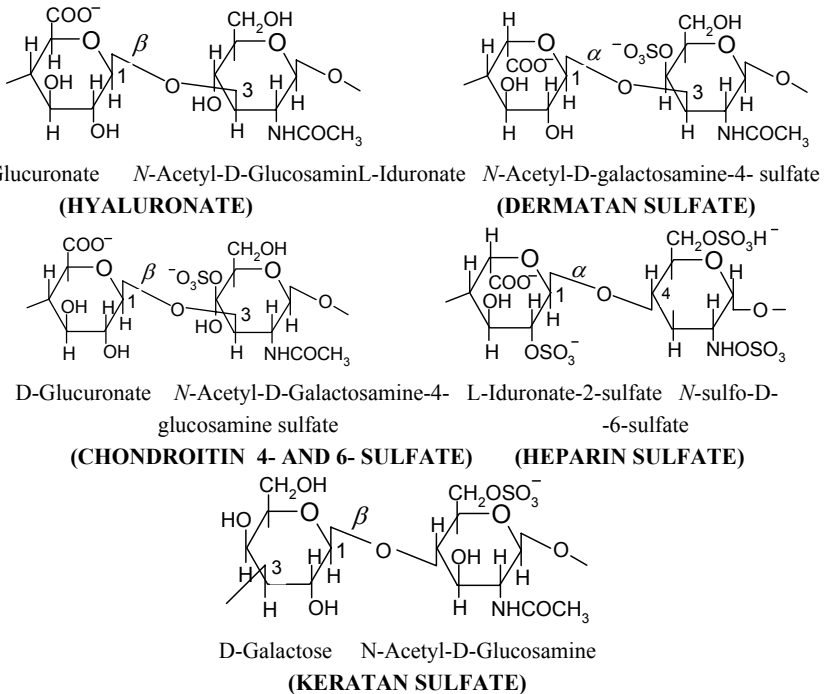
Figure 1. Repeating disaccharide units in glycosaminoglycans

The proteoglycans in the extracellular matrix of cartilage are the best-characterized members of this diverse class of glycoproteins. The proteoglycan aggrecan (Figure 2) and the protein collagen are key components of the cartilage. The triple helix of collagen provides structure and tensile strength, whereas aggrecan serves as a shock absorber [4]. The protein component of aggrecan is a large molecule composed of 2,397 amino acids. The protein has three globular domains, and the site of glyco-saminoglycan attachment is the extended region between globular domains 2 and 3. This linear region contains highly repetitive amino acid sequences, which are sites for the attachment of keratan sulfate and chondroitin sulfate. Many molecules of aggrecan are in turn noncovalently bound through the first globular domain to a very long filament form by linking together molecules of the glycosaminoglycan hyaluronan. Water is absorbed on the glycosaminoglycans, attracted by the many negative charges. Aggrecan can cushion compressive forces because the absorbed
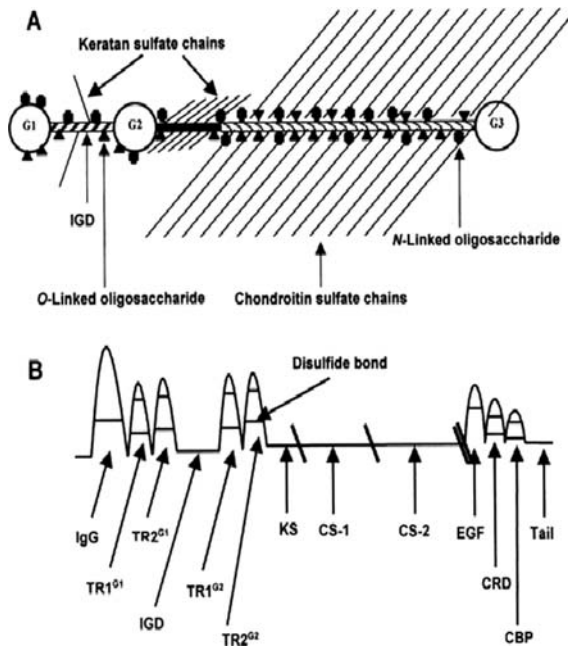
Figure 2. Aggrecan structure (A) Globular protein and attached GAG chain structure. (B) Protein domain structure. Folded modules: IgG, immunoglobulin fold: TR, tandem repeats: EGF, epidermal growth factor-like module: CRD, carbohydrate recognition domain: CBP, complement binding protein-like module. Extended domains: IGD, interglobular domain: KS, keratan sulfate attachment domain: CS-1 and CS-2, chondroitin sulfate attachment domains.

water enables it to spring back after having been deformed [4]. When pressure is exerted, as when the foot hits the ground while walking, water is squeezed from the glycosaminoglycan, cushioning the impact. When the pressure is released, the water rebinds. Osteoarthritis can result from the proteolitic degradation of aggrecan and collagen in the cartilage [4].

Decorin is a proteoglycan on average 90–140 kDa in size which belongs to the small leucine-rich proteoglycan (Figure 3) family and consists of an *N*-glycosylated protein core to which a single glycosaminoglycan chain of dermatan sulfate and/or hybridized with chondroitin sulfate is attached (Figure 4) [2,5].

Decorin influences fibrillogenesis [6], and also interacts with fibronectin, thrombospondin, the complement component C1q, epidermal growth factor receptor (EGFR) and transforming growth factor-beta-TGF-beta [7].

In some reports, decorin has been shown to enhance the bioactivity of TGF-beta 1, in other publications, TGF-beta 1 bioactivity has been shown to be inhibited [8–12]. Because of this, it is considered that the primary function of decorin lies in certain aspects of regulation during the cell cycle.

Decorin has been shown also to have anti-tumorigenic properties in an experimental murine tumor model and to be capable of suppressing the growth of various tumor cells [13,14].
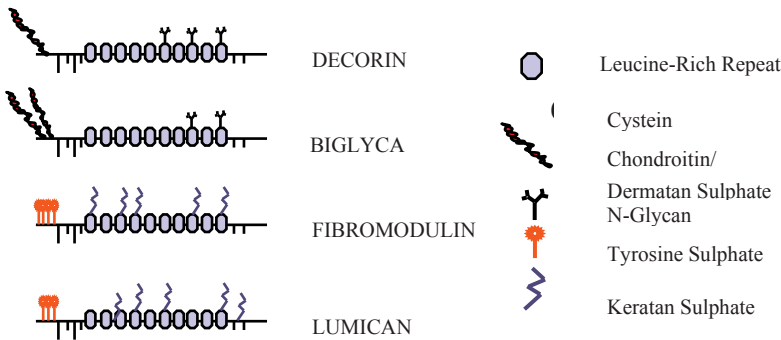
Figure 3. Proteoglycans with leucine-rich repeats
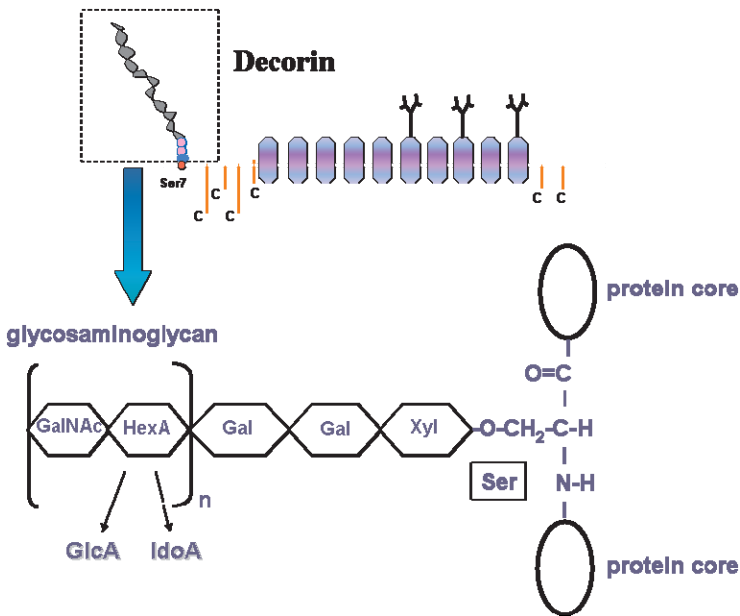
Figure 4. Structure of decorin

Chondroitin sulfate (CS) is composed of repeating disaccharide units of [GlcA (1-3)GalNAc(1-4)] bound through a xylosyl linker to Ser residues of decorin core protein. Most of the disaccharide units are sulfated at either the 4- or 6- position of GalNAc residues, with a small percentage modified at both positions. In dermatan sulfate, a large percentage of the uronic acid residues is epimerized to IdoA and may be sulfated at the 2-position [15]. CS sulfation occurs less commonly at the 2-position of Uro-A. The determination of 4- and 6-sulfation is more difficult since both positions lie on the same residue.

In the past few years mass spectrometry (MS) started to represent an option in the field of CS/DS compositional and structural analysis [16–23]. Beside the sensitivity and speed of analysis, a major advantage of MS over other analytical tools is that modern instruments exhibit tandem MS facility, which gives the possibility to sequence the precursor ion and obtain, in a single experiment, straightforward information upon the repeating disaccharide unit, number of sulfate groups and their precise location along the CS/DS chain and/or within a certain monomer ring. Recently, systematically optimized methods based on electrospray (ESI) collision-induced dissociation (CID tandem mass spectrometry (MS/MS) provided accurate information on unsulfated and regularly sulfated chondroitin oligomers [20,23]. CID MS/MS ion signatures allowed quantification of CS-like repeats in DS chains and identification of the general epimerization and sulfation status [17]. Discrimination between L-IdoA and D-GlcA epimers in GAGs has been also recently accomplished by electron detachment dissociation (EDD) MS/MS [24].

The structural investigation of CS/DS from human skin fibroblast decorin was accomplished also by capillary electrophoresis (CE) for separating oversulfated species in off-line combination with nanoESI quadrupole time-of-flight tandem mass spectrometry (nanoESI–QTOF–MS/MS) [21–23]. By this protocol it was possible to elucidate the structure of a previously unknown pentasulfated CS/DS hexasaccharide.

The capability of tandem MS to determine the patterns of sulfation and epimerization of CS/DS oligosaccharides in a single step [23] and even to differentiate uronic acid epimers was also recently demonstrated [24]. These results indicate that almost exhaustive information on CS/DS structure may be obtained using tandem MS.

CHONDROITIN LYASES

Di-6S   Δ Di-6S
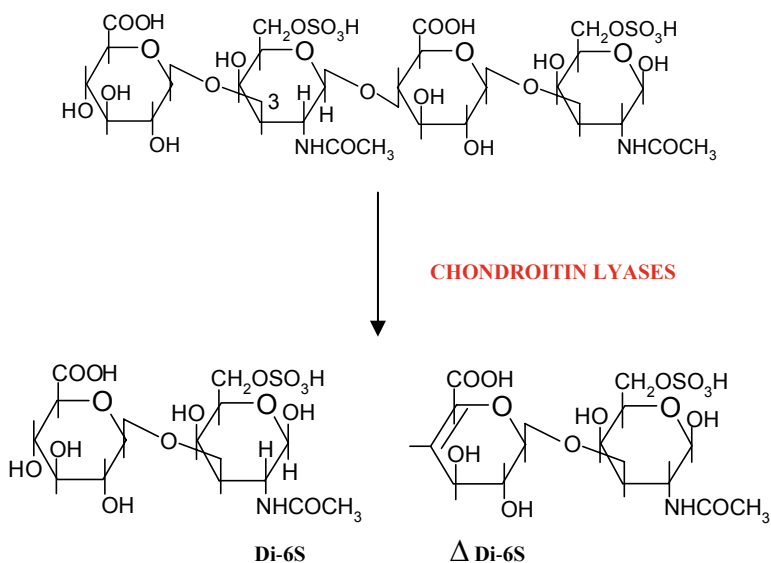
In the present work we developed a straightforward method to analyze atypical sulfation patterns in human fibroblast decorin CS/DS-disaccharides. We took advantage of chondroitin lyase recognition specificity (Figure 5) and produced disaccharides with known hexuronate epimerization. By employing ESI MS and tandem MS, we were able to determine oversulfation of the fragment ions and unequivocally identify the oversulfation site. The obtained structural data are consistent with sulfation of glucuronate in human decorin.

Chondroitin AC lyase   **Chondroitin B lyase**

...-GlcA-GalNAc-GlcA-GalNAc-IdoA-GalNAc-...
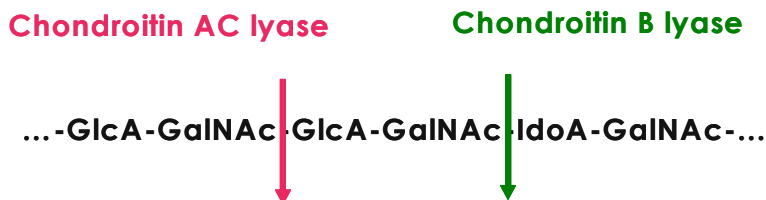
Figure 5. Recognition specificity of chondroitin lyases used as a tool for identification of GlcA- and IdoA-containing domains in CS-/DS GAG chains

2.   Experimental

2.1.   PREPARATION OF DECORIN

1 l of conditioned medium was collected from quiescent human skin fibroblasts and fractionated by chromatography on DEAE-Tris-Acryl M followed by concentration with Aquacide I (Calbiochem, Bad Soden, Germany). Thereafter, 0.5 mg proteoglycan was dialyzed against 20 mM Tris/HCl pH 7.4, containing 150 mM NaCl and separated by high performance liquid chromatography on a SEC-DEAE column (Phenomenex) by a discontinuous NaCl gradient to obtain the decorin fraction. For purity, the decorin (proteoglycan and core) was visualized on a silver stained 12.5% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE).

2.2.   EXTRACTION AND PREPARATION OF GLYCOSAMINOGLYCANS

To obtain free CS/DS chains, β-elimination was carried out in 0.15 M NaOH and 1 M NaBH$_4$ (final concentration) for 20 h at 37°C. The reaction mixture was neutralized with 50% acetic acid, diluted with 50 mM NaCl, 20 mM Tris/HCl, pH 7.4 and applied to a 0.5 ml column DEAE-Tris-Acryl-M. The column was washed and the CS/DS chains were eluted with 1 M NaCl and desalted with a PD-10 column. A pool (0.5 µmol of HexA) was exhaustively digested with 2 × 50 mU/assay chondroitin AC I lyase (Seikagaku Kogyo, Tokyo, Japan). Size fractionation of the oligosaccharides was performed on a Superdex Peptide HR10/30 column (Amersham-Pharmacia, Freiburg, Germany), equilibrated and eluted in 150 µl

fractions with 0.5 M $(NH_4)HCO_3$ at a flow rate of 0.5 ml/min and continuous UV detection at 232 nm. The disaccharide fraction was pooled and desalted on a pre-packed D-Salt column (MWCO 5000) (Pierce, Rockford, IL). In order to maximize the yield of the analyte, the collected disaccharides were not recromatographed.

2.3. MASS SPECTROMETRY

For MS analysis, the aqueous solution of the disaccharide fraction obtained from the gel filtration chromatography separation was evaporated (digital SpeedVac system SPD111V, Thermo Electron Corporation, Asheville, NC, USA) coupled o a vacuum pump PC 2002 Vario with CVC 2000 Controller from Vaccubrand GmbH (Wertheim, Germany). The resulting dry substrate was dissolved in 20 µl of 20 mM ammonium acetate in deionized water/MeOH (HPLC grade, Merck, Darmstadt, Germany) 40:60 [v/v] solution, pH 12.0. Prior to MS analysis, the sample solution was centrifuged in a SIGMA 2-16 model centrifuge from Sartorius (Göttingen, Germany).

Mass spectrometry was performed on a High Capacity Ion Trap (HCT) Ultra PTM mass spectrometer (Bruker Daltonik, Bremen, Germany). The HCT mass spectrometer is interfaced to a PC running the Compass integrated software package under Windows[XP], which includes EsquireControl and Hystar modules for instrument tuning, control and spectrum acquisition, and DataAnalysis for storing the ion chromatograms and processing the MS data. The samples were infused into MS by on-line syringe pump electrospray at a constant flow rate of 0.5 µl/min. Nitrogen at a flow rate of 2–5 l/min was employed at 250°C for desolvation and as a nebulizer gas at 7 p.s.i. The instrument was set to operate in the negative ion mode under 3.2 kV ESI potential. All spectra were acquired in the mass range 100–1,500 m/z, with a scan speed of 8,100 m/z per second. Tandem mass spectrometry was performed in the manual MS^2 mode, by CID using He as the collision gas. For all tandem MS experiments, the precursor ion was selected within an isolation width of 2u. The MS^2 was carried out using a fragmentation amplitude of 0.8 V.

The m/z scale of the mass spectrum was calibrated by use of an external calibration standard G2421A electrospray ''tuning mix,'' from Agilent Technologies (Santa Rosa, CA, USA). The reference provided in negative ion mode a spectrum with a fair ionic coverage of the m/z range scanned in both MS and CID MS^2 experiments. The obtained mass accuracy was situated within the normal range of an HCT MS instrument.

3. Results

Decorin CS disaccharide fraction obtained after depolymerization using chondroitin AC I lyase and separation by GFC was submitted to screening by ESI HCT MS. The ESI MS screening is presented in Figure 6. As visible in Figure 6, by optimized MS screening conditions, the in-source loss of the labile sulfate group was hampered.

The spectrum revealed the presence of an unsaturated monosulfated species documented by the singly charged ions at *m/z* 458.17 (disaccharide monomer) and *m/z* 916.89 (disaccharide dimer) and of an atypical sequence detected as a mono-deprotonated and sodiated ion at *m/z* 538.12 and 560.14 respectively (Figure 6). In analogy to previous reports [21–23], the ion at *m/z* 458.15 was assigned by mass calculation to the monosulfated unsaturated 4,5-Δ-[GlcA-GalNAc]. The ion assignment to a structure containing unsaturated GlcA was based on the known mechanism of chondroitin AC I lyase cleavage reaction which leads to the formation of a 4,5 double bond in the GlcA moiety at the non-reducing end.

Since the 79.96 Da shift exhibited by the singly charged ion at *m/z* 538.11 is matching, with fair accuracy, the mass of sulfate ester group, this ion was attributed to the bisulfated form of 4,5-Δ-[GlcA-GalNAc].

In order to elucidate the disaccharide structure, bring solid evidence upon the oversulfation and determine the site of oversulfation, the ion at *m/z* 538.11 was isolated within an isolation window of 2 u and submitted to CID MS/MS experiment. ESI CID MS/MS (Figure 7) gave rise to a fragmentation pathway consistent with GalNAc monosulfation. The cleavage of the glycosydic GlcA-GalNAc bond generated the Y_1^- at *m/z* 300.08 and Z_1^- at *m/z* 282.07 ions, related to the mono-sulfated GalNAc$^-$ moiety and its dehydrated form. Further on, the C_1-SO_3^- ion at *m/z* 175.06 and the low intensity B_1^- at *m/z* 157.06 though both attributable to CID-induced GlcA$^-$ desulfation are structurally interesting since the latter one is diagnostic for
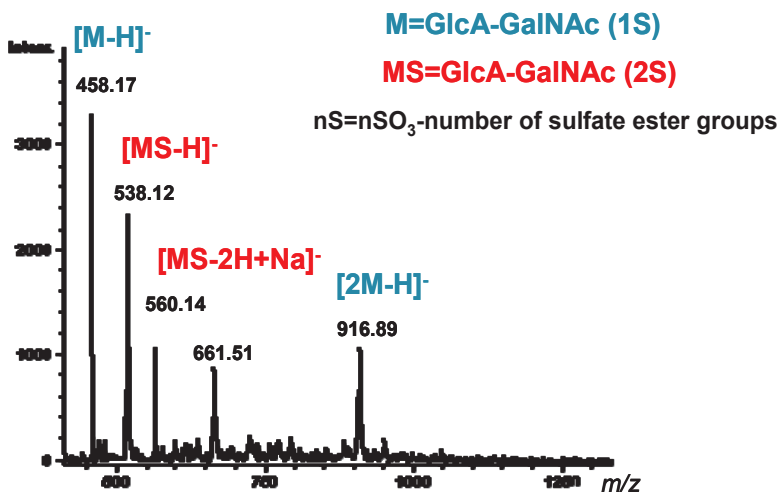
Figure 6. (–) ESI HCT MS of the decorin CS disaccharide mixture obtained after deploymerization with chondroitin AC I lyase and separation on GFC-Superdex peptide column
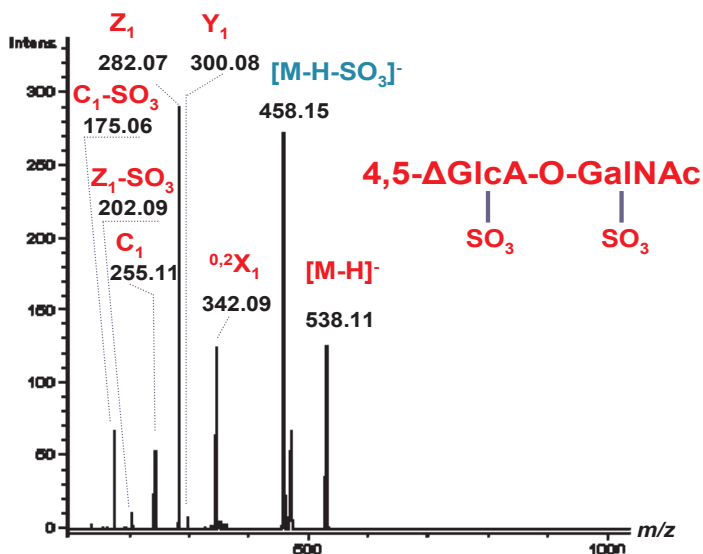
Figure 7. (–) ESI HCT CID MS/MS of the singly deprotonated ion at *m/z* 538.11 assigned according to mass calculation to oversulfated CS disaccharide having the composition 4,5Δ-GlcA-GaNAc (2S). Nomenclature of fragment ion is according to published recommendations [25]

double bond formation and implicit unsaturation of the glucuronic acid. The spectrum provides clear evidence upon the generation of C_1^- ion at *m/z* 255.11, which is diagnostic for 4,5-Δ-GlcA sulfation (inset Figure 7). The reduced loss of sulfate in this CID experiment is documented by the absence of an ion at nominal *m/z* 378 which would correspond to desulfated disaccharide precursor and the low intensity of the ion at *m/z* 202.09, characteristic for the non-sulfated GalNAc⁻.

According to previous findings [21–23] the relative intensity of the regularly sulfated Y⁻ and Z⁻ sequence ions is correlated with the sulfation site.

In this context, a closer inspection of the spectrum in Figure 7, reveals that Z_1^- ion exhibits a higher relative abundance than Y_1^- ion. This feature leads to the conclusion that GalNAc is sulfated in position 6, which additionally corroborates the monosulfation of 4,5-Δ-GlcA.

4.   Discussion

We developed here a method to allow the structural analysis of oversulfated regions in CS/DS chains from the proteoglycan decorin. This was achieved by a combination of specific digestion with chondroitin AC I lyase and fragmentation analysis by tandem mass spectrometry.

In a previous attempt [23] to analyze the CS/DS chain of the ubiquitously expressed decorin by tandem MS, while the sulfation of IdoA was unambiguously

demonstrated, adequate data to allow clear-cut discrimination between the structural motifs with either bisulfated D-GalNAc or monosulfated D-GlcA could not be obtained. Here, we generated by AC I -lyase digestion disaccharides with unsaturated GlcA, which served as a tag of the non-reducing end, easily recognizable by MS. Further, chromatographically purified disaccharides were screened and sequenced in MS/MS. By optimized screening and CID MS/MS conditions, the loss of labile sulfate group was minimized, therefore it was possible to identify regular and bisulfated decorin CS disaccharide species. MS/MS brought solid evidence upon additional sulfate group location and implicitly the first evidence upon sulfation of D-GlcA. Hence, we could identify for the first time in human decorin a novel motif with single sulfated D-GlcA unit interspersed within adjacent blocks of regularly sulfated GalNAc.

To complete the detailed structural investigation of decorin GAG oversulfation, further studies involving multiple stage fragmentation for the determination of sulfate group location within GlcA monomer ring will be undertaken.

Acknowledgments

References

1. Berg JM, Tymoczko JL, Stryer L. *Biochemistry* 6-th Ed., W.H. Freeman and Company, 2007, New York.
2. Seidler DG, Peter-Katalinić J, Zamfir AD. Galactosaminoglycan function and oligosaccharide structure determination. *Sci World J.* 2007, 7, 233–241.
3. Iozzo RV. Matrix proteoglycans: from molecular design to cellular function. *Annu Rev Biochem.* 1998, 67, 609–652.
4. Charni-Ben Tabassi N, Garnero P. Monitoring cartilage turnover. *Curr Rheumatol Rep.* 2007, 25, 16–24.
5. Zamfir AD, Seidler DG, Kresse H, Peter-Katalinić J. Structural investigation of chondroitin/dermatan sulfate oligosaccharides from human skin fibroblast decorin. *Glycobiology.* 2003, 13, 733–742.
6. Rühland C, Schönherr E, Robenek H, Hansen U, Iozzo RV, Bruckner P, Seidler DG. The glycosaminoglycan chain of decorin plays an important role in collagen fibril formation at the early stages of fibrillogenesis. *FEBS J.* 2007, 274, 4246–4255.
7. Zutter MM, Edelson BT. The alpha2beta1 integrin: a novel collectin/C1q receptor. *Immunobiology.* 2007, 212, 343–353.
8. Zanotti S, Saredi S, Ruggieri A, Fabbri M, Blasevich F, Romaggi S, Morandi L, Mora M. Altered extracellular matrix transcript expression and protein modulation in primary Duchenne muscular dystrophy myotubes. *Matrix Biol.* 2007, 26, 615–624.
9. Zhu J, Li Y, Shen W, Qiao C, Ambrosio F, Lavasani M, Nozaki M, Branca MF, Huard J. Relationships between transforming growth factor-beta1, myostatin, and decorin: implications for skeletal muscle fibrosis. *J Biol Chem.* 2007, 282, 25852–25863.

10. Zhang Z, Li XJ, Liu Y, Zhang X, Li YY, Xu WS. Recombinant human decorin inhibits cell proliferation and downregulates TGF-beta1 production in hypertrophic scar fibroblasts. *Burns*. 2007, 33, 634–641.
11. Schnabel LV, Mohammed HO, Miller BJ, McDermott WG, Jacobson MS, Santangelo KS, Fortier LA. Platelet rich plasma (PRP) enhances anabolic gene expression patterns in flexor digitorum superficialis tendons. *J Orthop Res.* 2007, 25, 230–240.
12. Shi YF, Zhang Q, Cheung PY, Shi L, Fong CC, Zhang Y, Tzang CH, Chan BP, Fong WF, Chun J, Kung HF, Yang M. Effects of rhDecorin on TGF-beta1 induced human hepatic stellate cells LX-2 activation. *Biochim Biophys Acta*. 2006, 1760, 1587–1595.
13. Goldoni S, Iozzo RA, Kay P, Campbell S, McQuillan A, Agnew C, Zhu JX, Keene DR, Reed CC, Iozzo RV. A soluble ectodomain of LRIG1 inhibits cancer cell growth by attenuating basal and ligand-dependent EGFR activity. *Oncogene.* 2007, 26, 368–381.
14. Nash MA, Loercher AE, Freedman RS. In vitro growth inhibition of ovarian cancer cells by decorin: synergism of action between decorin and carboplatin. *Cancer Res.* 1999, 59, 6192–6196.
15. Seidler DG, Breuer E, Grande-Allen KJ, Hascall VC, Kresse H. Core protein dependence on epimerization of glucuronosyl residues in galactosaminoglycans. *J Biol Chem*. 2002, 277, 42409–42416.
16. Zaia J. Mass spectrometry of oligosaccharides. *Mass Spectrom Rev.* 2004, 23, 161–227.
17. Miller MJ, Costello CE, Malmstrom A, Zaia J. A tandem mass spectrometric approach to determination of chondroitin/dermatan sulfate oligosaccharide glycoforms. *Glycobiology.* 2006, 16, 502–513.
18. Desaire H, Leary JA. Detection and quantification of the sulfated disaccharides in chondroitin sulfate by electrospray tandem mass spectrometry, *J Am Soc Mass Spectrom.* 2000, 11, 916–920.
19. Henriksen J, Hoffmeyer Ringborg L, Roepstorff P. On-line size-exclusion chromatography/mass spectrometry of low molecular mass heparin. *J Mass Spectrom.* 2004, 39, 1305–1312.
20. Hitchcock AM, Costello CE, Zaia J. Glycoform quantification of chondroitin/dermatan sulfate using a liquid chromatography-tandem mass spectrometry platform. *Biochemistry.* 2006, 45, 2350–2361.
21. Zamfir AD, Peter-Katalinić J. Capillary electrophoresis-mass spectrometry for glycoscreening in biomedical research. *Electrophoresis.* 2004, 25, 1949–1963.
22. Zamfir AD, Seidler DG, Kresse H, Peter-Katalinić J. Structural characterization of chondroitin/dermatan sulfate oligosaccharides from bovine aorta by capillary electrophoresis and electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *Rapid Commun Mass Spectrom*. 2002, 16, 2015–2024.
23. Zamfir AD, Seidler DG, Kresse H, Peter-Katalinić J. Structural investigation of chondroitin/dermatan sulfate oligosaccharides from human skin fibroblast decorin. *Glycobiology.* 2003, 13, 733–742.
24. Wolff JJ, Chi L, Linhardt RJ, Amster IJ. Distinguishing glucuronic from iduronic acid in glycosaminoglycan tetrasaccharides by using electron detachment dissociation. *Anal Chem*. 2007, 79, 2015–2022.
25. Domon B, Costello CE. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 1988, 5, 397–409.

7. APPLICATION OF HIGH PERFORMANCE MASS SPECTROMETRY TO STRUCTURAL ANALYSIS OF GLYCOSAMINOGLYCAN OLIGOSACCHARIDES

DANIELA G. SEIDLER

Institute of Physiological Chemistry and Pathobiochemistry, University of Münster, Waldeyer Str. 15, D-49149, Münster, Germany;
Tel: +49-251-83-55578; Fax: +49-251-83-55596;
e-mail: dgseidle@uni-muenster.de

Abstract. Glycosaminoglycans, including those attached to core proteins of proteoglycans, are linear polyanions composed of hexuronate residues alternating with either glucosamine- or galactosamine derivatives. In galactosaminoglycans the disaccharide subunits can be modified in varying locations within the chains to a variable extent. The modification reaction of sulfation and epimerization lead to specific saccharide sequences controlling a wide range of recently recognized, pivotal functions in tissue development, homeostasis, and repair, and in tumor metastasis. To understand the biological function of galactosaminoglycans it is important to know their biosynthesis and the fine structure.

1. Proteoglycans

Proteoglycans (PGs) are structural components of extracellular matrix aggregates or of cell surfaces to which they are anchored by membrane-spanning domains or by inositol phospholipids. PGs contain core proteins covalently linked to linear sulfated polysaccharides, the glycosaminoglycan chains (GAGs). PGs are involved in maintaining the transparency of the cornea, the viscoelasticity of blood vessels, tensile strength of skin and tendon and in addition they play a role in depots for growth factors and cytokines. PGs are organized in several families, for example with respect of their core proteins, like lecticans or small leucine-rich repeat proteoglycan family (SLRP). Decorin is a member of the SLRP and widely expressed in the connective tissues. It is considered to be a multifunctional proteoglycan [1]. The proposed functions of decorin include the regulation of collagen fibrillogenesis, maintenance of tissue integrity via binding with fibronectin and thrombospondin, and a reservoir of transforming growth factor-β (TGF-β) [1,2]. Decorin is highly conserved among species, with the exception of the hypervariant region near the N-terminus. The mature proteoglycans of the human decorin consists of a 36-kDa protein core. The *N*-terminal domain contains a Ser-Gly sequence with a single GAG chain [3], four cysteine residues forming intrachain disulfide bonds, and 2-3 asparagine-linked oligosaccharides [4]. Biglycan, another member of the SLRPs shares a 55% homology and is linked at the N-terminus with two GAG chains.

Like for decorin the GAG chain can be chondroitin or dermatan sulfate depending on the tissue expressed.

According to the structural type of the disaccharide repeating unit, GAGs and their proteoglycans are categorized into chondroitin sulfate (CS), dermatan sulfate (DS), heparan sulfate (HS), heparin, hyaluronic acid, and keratan sulfate. CS/DS and HS are O-linked to a core protein via a linkage-region consisting of 4 sugar residues: GlcA-β1,3-Gal-β1,3-Gal-β1,4-Xyl [5] which is formed by specific gly-cosyltransferases [6–9]. The decision which *O*-linked GAG is synthesized onto the core protein is not completely resolved, yet. The transfer of either GlcNAc or GalNAc residue decides whether HS or CS/DS is polymerized [5]. A number of studies have focused on the structure of the linkage-region as a possible determent of the type of GAG chain. This resulted in the discovery of several structural modifications, for example the Xyl can be phosphorylated at C2 position which is found in DS and HS. Interestingly, the sulfation of the two Gal residues is only found in CS [for review: 10].

CS is consisting of repeating disaccharide units of D-glucuronate (D-GlcA) and *N*-acetylgalactosamine (D-GalNAc) with usually sulfated 4′- or 6′-OH groups. The glycosidic linkage stereochemistry in CS is (βD-GlcA 1′->3′ βD-GalNAc1′->4′-)$_n$. A part of D-GlcA in CS can be epimerized to L-iduronate (L-IdoA) giving rise to dermatan sulfate (DS). So called over-sulfation occurs by generation of GalNAc-4′,6′-disulfate- and/or sulfated uronic acid residues. Thus, there are reports of 2′-sulfation in L-IdoA or D-GlcA in DS or 3′-sulfation in D-GlcA in squid cartilage CS-E. These ester sulfates contribute to the structural and functional diversity of GAGs [11]. Conformational studies revealed L-IdoA being more flexible which is prone to exhibit energetically more favorable conformations [11], which explains the implication of the DS oligosaccharides in biologically active complexes.

To understand the complexity and function of CS and DS it is important to know the biosynthesis of these two GAGs.

2.   The Biosynthesis of CS/DS

The biosynthesis of CS/DS occurs on a linkage region with the transfer of a βGal-NAc followed by D-GlcA carried out by a specific enzyme complex [5]. After the polymerization of chondroitin further modification occur on the polymer level (Figure 1). To obtain chondroitin sulfate the D-GalNAc can be modified with sul-fate groups at 4 or 6 sulfate. Furthermore, the D-GlcA can be converted into the sterioisomer L-IdoA by the chondroitin-glucuronate 5-epimerase (E.C. 5.3.1.19). The inversion at C-5 of D-GlcA to L-IdoA occurs on the polymer level [13] and involves an abstraction of the C-5 hydrogen of the uronosyl residues [14]. A two-base reaction mechanism was postulated, involving a monoprotic L-ido-specific base and a polyprotic D-gluco-specific base, during the action of the enzyme on an *E. coli*-derived capsular polysaccharide [15]. The chondroitin-glucuronate 5-epimerase was the last enzyme cloned and characterized [16], however, described earlier as

SART-2 over-expressed in squamous carcinoma cells [17]. One single L-IdoA in a GAG chain and the chain is then called DS. 4-*O*-sulfation of GalNAc residues is a typical modification found in CS/DS and was postulated to be a prerequisit for IdoA stabilization [13] carried out by three different sulfotransferases [18]. These properties enable DS to perform a variety of biological functions that are different from those of CS.
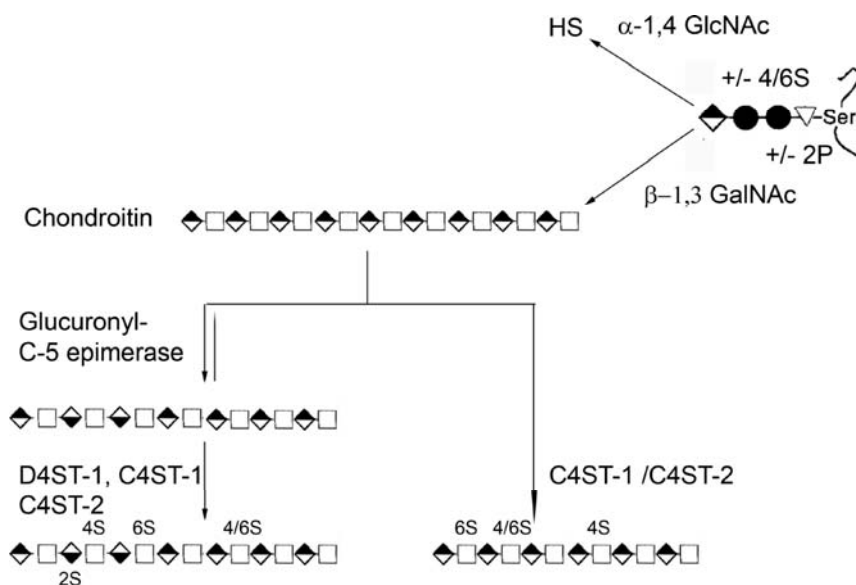
Figure 1. Biosynthesis of chondroitin/dermatan sulfate [62]. Symbols: ▽ Xyl; ● : Gal; ◆ GlcA; ▢ : GalNAc; ◆ IdoA. Abbreviations: D4ST-1: dermatan 4-sulfotransferase-1; C4ST: chondroitin 4-sulfotransferase; heparan sulfate: HS; P: phosphate; S: sulfate

3.   Function of DS

The biological function of the GAG chains is extensively studied for HS [19]. However, a growing body of evidence has identified DS, as playing a role in a large number of biological functions. DS may interact in place of heparin or HS with matrix proteins, like fibronectin [20], cell membrane receptors and platelet factor 4 [21]. It is known that decorin, biglycan and the truncated form of decorin interact with type I collagen. Binding of decorin to type I collagen is mediated by specific binding of its core protein because the DS-depleted core protein shows binding activity, too [22], but with a reduced affinity [23]. Although DS is not essential for binding, it has been demonstrated that DS is influencing the size of the collagen fibrils [24].

DS chains can stimulate thrombin inhibition by heparin cofactor II, a plasma proteinase inhibitor [25]. Another important aspect of CS/DS was reported in studies on tumor metastasis. There are indications that CS/DS may regulate angiogenesis and melanoma cell invasion and proliferation [26]. Furthermore, DS function in immune response [27] by inducing ICAM-1 expression on endothelial cells and recruiting leukocytes via selectin interactions [for review: 28]. CS/DS is controlling organ and tissue development in general [5,29], for example proliferation of neural stem cell/progenitor cells are regulated by CS and CSPGs intervene in fate decisions between the neuronal and glial lineage [30]. Also neurite outgrowth [31] and axon regeneration [32], or leukocyte or tumor cell transmigration [33] are regulated by CS/DS. DS is abundant in the wound environment and binds and activates growth factors such as fibroblast growth factor-2 (FGF-2) and FGF-7, which are present during the wound repair process [34]. The minimum size required for the activation of FGF-2 was an octasaccharide and for FGF-7 a decasaccharide enrich in monosulfated, primarily 4-*O*-sulfated, disaccharides and L-IdoA. This indicated a preference for specific DS oligosaccharides capable of promoting FGF-2- and FGF-7-dependent cell proliferation [35]. CS/DS can modu-late even more transduction of extracellular signals, e.g., of hepatocyte growth/ scatter factors [36] and FGF-1 (unpublished results).

Among the GAGs, HS, heparin and HS show considerable conformational flexibility due to the presence of L-IdoA residues which change easily between chair and skew conformations [12]. Crystal structures of heparin tetra- and hexasaccharides in the presence of FGF-2 indicated that FGF binding stabilizes the $^1C_4$ conformation of the L-IdoA2S residue involved in binding [37].

4.   Ehlers-Danlos Syndrome and CS/DS

Ehlers-Danlos syndrome (EDS) is a heterogenous group of heritable connective tissue disorders. It mostly affects skin, joints, blood vessels and generalized connective tissue fragility [38]. The latest classification of EDS contains 6 subtypes with major and minor diagnostic criteria. The classic type of EDS is charactierized by skin hyperextensibility and abnormal wound healing. The molecular basis was shown by ultrastructural studies revealing that heterotypic collagen I and V fibrils showed increased diameter and highly variable in width and shape [39,40]. Mutations in *COLVA1* and *COLVA2* gene were responsible for the classical EDS [41]. Several other genes have been implicated to result in EDS phenotypes. The progeroid type of EDS is characterized by wrinkled, loose skin on the face, curly fine hair, scanty eyebrows and eyelashes, in addition to the classical features of EDS.

In humans, aberrant GAG substitution of decorin was observed in a progeroid syndrome patient [42,43] who carried compound heterozygous mutations in *B4GALT7* gene encoding galactosyltransferase I (ß4GalT-7; E.C. 2.4.1.133), which catalyses the transfer of galactose to the xylose residue forming the linkage region of proteoglycans [6]. This patient's fibroblasts displayed reduced galactosyltransferase

activity [43] and while biosynthesis of large CS/DS-PGs and HS-PGs was normal, 20–80% of the decorin core-protein was substituted with a GAG chain [7,42–44]. In a recent report, two patients affected with EDS exhibited a phenotype [45] similar to the previously described patient [42]. The patients showed the typical EDS features including craniofacial appearance and skeletal abnormalities. However, their skin was not remarkably loose and no progeroid appearance was apparent. Sequence analysis of the *B4GALT7* gene from affected patients identified a homozygous C808T substitution causing an Arg270Cys change in the C-terminal catalytic region of galactosyltransferase I [45]. Biochemically, it could be demonstrate that cultured skin fibroblasts from one of these patients exhibit reduced galactosyltransferase activity. This resulted in a defective biosynthesis of mature decorin and biglycan. Only 50% of decorin was linked with a GAG chain. Furthermore, the fibroblasts exhibited a reduced epimerization of the GAG chain of decorin and slightly increased length (K_{av} 0.43/wt vs 0.40/mutant) [46]. The GAG chains of biglycan were not affected by the ß4GalT-7[Arg270Cys] substitution. Since the degree of epimerization of the GAGs of DSPGs can have a profound influence on its functional properties [28], epimerization of decorin and biglycan was investigated. Although previously described [47], the function of monoglycanated biglycan remains elusive. Epimerization of the biglycan glycosaminoglycan chain was only moderately affected by the *B4GALT7* mutation. This could have resulted from changes in extracellular environment of the glycosaminoglycan synthesizing cells, as has been suggested before [27]. Pulse-chase analysis and confocal microscopy studies suggest that biosynthesis and secretion of decorin are not severely perturbed by the ß4GalT-7 mutation. Thus, we can currently only speculate if the changes in glycosaminoglycan structure observed for decorin and biglycan may lead to a delay in enzymatic degradation and subsequent accumulation in intracellular degradative compartments. These changes in decorin and biglycan glycosylation might be responsible for several aspects of the EDS phenotype in ß4GalT-7 deficient patients [46].

5.   Mass Spectrometry as a Tool to Determine Biological Active Structures in Oligosaccharides

Taken together, it is obvious that a detailed knowledge of the fine structure of GAG chains would help to interpret the mechanisms required for the manifestation of their biological properties. Usually, specific GAGs of a cell population or tissue are analysed by enzymatic digestion using chondroitin ABC lyase which cleaves both CS/DS type into disaccharide units. On the other hand, the patients [42,45] described above shows that in some cases the biosynthesis of only a single proteoglycan [42,46] is affected. Previously, we could demonstrate that electrospray coupled with mass spectrometry is an appropriate tool to determine the structure of oligosaccharides with a chain length $\geq$ than hexasaccharides derived from different

sources, for example bovine aorta [48] and keeping the liable sulfate ester groups. Unfortunately, only elaborate chemical saccharide sequencing methods [49,50] are available rather than simpler procedures analogous to Edman-degradation of polypeptides. For the function of GAGs it is important to know the location of the sulfate groups, therefore, a strategy of analyzing a specific proteoglycan, decorin was developed. We were interested in the oligosaccharide structures with a size longer then tetrasacchairdes of GAGs derived from decorin. As a model system skin fibroblasts were used, because these cells express endogenous decorin linked with a DS chain. Decorin was purified from the conditioned medium [51–53] of skin fibroblasts and the free DS chain obtained by β-elimination were further analyzed.

The first step was to determine biological active sequences within the GAG chain from skin fibroblast decorin. GAGs derived from skin fibroblast decorin were capable to bind FGF-2 in a dose dependent manner [52]. Oligosaccharides were obtained from the chondroitin B lyase digests. Chondroitin B lyase is cleaving DS chains at the 1,4-β-D-hexosaminyl and a 1,3-α-L-iduronosyl to saccharides containing at the non-reducing end a unsaturated desoxy-β-D-gluc-4-enuronosyl group which results from L-IdoA of the DS. The obtained oligosaccharides were fractionated by gel permeation chromatography and it revealed that the abundance of octa- and longer saccharides is very low. Hexasccharides and larger oligosaccharides were still capable to bind to immobilized FGF-2, albeit at lower capacity than the intact GAG chain. The heterogeneity of the oligosaccharide mixture obtained after chondroitin B lyase digestion was first demonstrated by capillary electrophoresis (CE) with UV detection. A peak of octasaccharides resulted after CE in 13 different species [52]. These different species were separated in two fractions: first fraction as oversulfated and the second one as undersulfated oligosaccharides. The oversulfated oligosaccharides were applied to ESI-MS. MS analysis revealed an oversulfated hexasaccharide either with a sulfation at the d-GlcA or bissulfated GalNAc [52]. At this stage the exact location of the last sulfate group was not possible, due to the limited fragmentation.

By optimizing the CE and ESI-MS conditions we succeeded in the structural characterization of up to 12-mer CS/DS oligosaccharides expressing different degrees of sulfation. The combination of enzymatic degradation, capillary electrophoresis and mass spectrometry allowed the determination of building blocks, the chain size and type of the repeating HexA-HexNAc, HexA-HexNAc(S) units, and the identification of oversulfation patterns.

In addition, collision induced dissociation conducted at variable energies (CID-VE) is a novel potent tool for GAG oligosaccharide sequencing and determination of sulfation patterns. The application of tandem MS under CID-VE conditions to the structural analysis of the pentasulfated hexasaccharide allowed the identification of oversulfated domains with potential relevance for carbohydrate/protein interactions. The limitation by this method was the determination and location of the sulfate ester group in a trisaccharide with three sulfates.

6. Sequencing of Oligosaccharides with Known Unsaturated HexA at the Non-reducing End

The increasing awareness of the biological importance of GAGs calls for the availability of sensitive and straightforward methods for their isolation and structural characterization, and for the determination of their noncovalent interactions. However, progress has been achieved recently by mass spectrometry (MS) [54–59]. Systematically optimized methods based on electrospray (ESI) collision-induced dissociation (CID) in a single fragmentation stage (MS/MS) provided accurate information on unsulfated and regularly sulfated chondroitin oligomers [48,57,58]. CID MS/MS ion signatures allowed quantification of CS-like repeats in DS chains and identification of the general epimerization and sulfation status [58]. Discrimination between L-IdoA and D-GlcA epimers in GAGs has been accomplished by electron detachment dissociation (EDD) MS/MS [60,61].

To overcome this limitation a three-step-methodology based on enzymatic digestion, chromatographic separation and multiple stage MS introduced here represents a viable alternative in glycosaminoglycomics and the prerequisite for further studies aiming at defining the sulfation in a specific site of the monomer and its role in mediating biological activities through noncovalent glycosaminoglycan-growth factor interactions. In this novel strategy developed here we used first an enzymatic approach to generate an unsaturated HexA derived from D-GlcA or L-IdoA, which acted further as a MS-recognizable tag of the non-reducing end. Subsequently, following size-exclusion chromatographic separation and collection of hexasaccharides, we employed multiple stage MS which, unlike tandem MS in a single dissociation phase, provided a strict control upon the fragmentation process and the possibility to re-sequence small oversulfated fragments until unequivocal elucidation of sulfate position.

The fragmentation in MS^2 followed by stepwise re-fragmentation within CID MS^3- MS^4 processes revealed that the non-reducing end of the CS-rich hexasaccharide is characterized by a monosulfated L-IdoA and a monosulfated d-GlcA. CID MS^2-MS^3 experiments on DS-rich hexasaccharide provided concrete evidence upon the presence of a DS structural motif containing sulfated d-GlcA at the non-reducing end. Interestingly, MS-MS^n results proved that both pentasulfated CS-rich and tetrasulfated DS-rich hexamers exhibit a regular sulfation of the d-GalNAc moiety. Hence, data collected by the new methodology corroborate for the first time the existence within the human decorin of a CS/DS motif incorporating one sulfated L-IdoA unit and one sulfated D-GlcA in a sequence of regularly sulfated adjacent and remote GalNAc and HexA blocks.

References

1. Kresse, H., Hausser, H., Schönherr, E., & Bittner, K. Biosynthesis and interactions of small chondroitin/dermatan sulphate proteoglycans. *Eur. J. Clin. Chem. Clin. Biochem.* **32**, 259–264 (1994).
2. Yamaguchi, Y., Mann, D.M., & Ruoslahti, E. Negative regulation of transforming growth factor-beta by the proteoglycan decorin. *Nature* **346**, 281–284 (1990).
3. Iozzo, R.V. The family of the small leucine-rich proteoglycans: key regulators of matrix assembly and cellular growth. *Crit Rev. Biochem. Mol. Biol.* **32**, 141–174 (1997).
4. Glössl, J., Beck, M., & Kresse, H. Biosynthesis of proteodermatan sulfate in cultured human
    fibroblasts. *J. Biol. Chem.* **259**, 14144–14150 (1984).
5. Sugahara, K. Mikami, T., Uyama, T., Mizuguchi, S., Nomura, K., & Kitagawa, H. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **13**, 612–620 (2003).
6. Götting, C., Kuhn, J., Zahn, R., Brinkmann, T., & Kleesiek, K. Molecular cloning and expression of human UDP-d-Xylose:Proteoglycan core protein beta-d-xylosyltransferase and its first isoform XT-II. *J. Mol. Biol.* **304**, 517–528 (2000).
7. Almeida, R., Levery, S.B., Mandel, U., Kresse, H., Schwientek, T., Bennett, E.P., & Clausen, H. Cloning and expression of a proteoglycan UDP-galactose:beta-xylose beta1,4-galactosyltransferase I. A seventh member of the human beta4-galactosyltransferase gene family. *J. Biol. Chem.* **274**, 26165–26171 (1999).
8. Bai, X., Zhou, D., Brown, J.R., Crawford, B.E., Hennet, T., & Esko, J.D. Biosynthesis of the linkage region of glycosaminoglycans: cloning and activity of galactosyltransferase II, the sixth member of the beta 1,3-galactosyltransferase family (beta 3GalT6). *J. Biol. Chem.* **276**, 48189–48195 (2001).
9. Kitagawa, H., Tone, Y., Tamura, J., Neumann, K.W., Ogawa, T., Oka, S., Kawasaki, T., & Sugahara, K. Molecular cloning and expression of glucuronyltransferase I involved in the biosynthesis of the glycosaminoglycan-protein linkage region of proteoglycans. *J. Biol. Chem.* **273**, 6615–6618 (1998).
10. Krishna, N.R., & Agrawal, P.K. Molecular structure of the carbohydrate-protein linkage region fragments from connective-tissue proteoglycans. *Adv. Carbohydr. Chem. Biochem.* **56**, 201–234 (2000).
11. Kusche-Gullberg, M., & Kjellen, L. Sulfotransferases in glycosaminoglycan biosynthesis. *Curr. Opin. Struct. Biol.* **13**, 605–611 (2003).
12. Casu, B., Petitou, M., Provasoli, M., & Sinay, P. Conformational flexibility: a new concept for explaining binding and biological properties of iduronic acid-containing glycosaminoglycans. *Trends Biochem. Sci.* **13**, 221–225 (1988).
13. Malmström, A., & Fransson, L.-Å. Biosynthesis of dermatan sulfate. I. Formation of L-iduronic acid residues. *J. Biol. Chem.* **250**, 3419–3425 (1975).
14. Malmström, A. Biosynthesis of dermatan sulphate. Loss of C-5 hydrogen during conversion of D-glucuronate to L-iduronate *Biochem. J.* **198**, 669–675 (1981).
15. Hannesson, H.H., Hagner-McWhirter, A., Tiedemann, K., Lindahl, U., & Malmström, A. Biosynthesis of dermatan sulphate. Defructosylated Escherichia coli K4 capsular polysaccharide as a substrate for the D-glucuronyl C-5 epimerase, and an indication of a two-base reaction mechanism *Biochem. J.* **313**, 589–596 (1996).
16. Maccarana, M. Olander, B., Malmstrom, J., Tiedemann, K., Aebersold, R., Lindahl, U., Li, J.P., & Malmstrom, A. Biosynthesis of dermatan sulfate: chondroitin-glucuronate C5-epimerase is identical to SART2. *J. Biol. Chem.* **281**, 11560–11568 (2006).
17. Nakao, M. Shichijo, S., Imaizumi, T., Inoue, Y., Matsunaga, K., Yamada, A., Kikuchi, M., Tsuda, N., Ohta, K., Takamori, S., Yamana, H., Fujita, H., & Itoh, K. Identification of a gene coding for a new squamous cell carcinoma antigen recognized by the CTL. *J. Immunol.* **164**, 2565–2574 (2000).

18. Mikami, T., Mizumoto, S., Kago, N., Kitagawa, H., & Sugahara, K. Specificities of three distinct human chondroitin/dermatan N-acetylgalactosamine 4-O-sulfotransferases demonstrated using partially desulfated dermatan sulfate as an acceptor: implication of differential roles in dermatan sulfate biosynthesis. *J. Biol. Chem.* **278**, 36115–36127 (2003).

19. Bernfield, M., Götte, M, Park, P.W., Reizes, O., Fitzgerald, M.L., Lincecum, J., & Zako, M. Functions of cell surface heparan sulfate proteoglycans *Annu. Rev. Biochem.* **68**, 729–777 (1999).

20. Walker, A., & Gallagher, J.T. Structural domains of heparan sulphate for specific recognition of the C-terminal heparin-binding domain of human plasma fibronectin (HEPII). *Biochem..J.* **317**, 871–817 (1996).

21. Cella, G., Boeri, G., Saggiorato, G., Paolini, R., Luzzatto, G. & Terribile, V.I. Interaction between histidine-rich glycoprotein and platelet factor 4 with dermatan sulfate and low-molecular-weight dermatan sulfate. *Angiology.* **43**, 59–62 (1992).

22. Schönherr, E., Hausser, H., Beavan, L., & Kresse, H. Decorin-type I collagen interaction. Presence of separate core protein- binding domains. *J. Biol. Chem.* **270**, 8877–8883 (1995).

23. Nareyeck, G., Seidler, D.G., Troyer, D., Rauterberg, J., Kresse, H., & Schönherr, E. Differential interactions of decorin and decorin mutants with type I and type VI collagens. *Eur. J. Biochem.* **271**, 3389–3398 (2004)

24. Rühland, C., Schönherr, E., Robenek, H., Hansen ,U., Iozzo, R.V., Bruckner, P., Seidler, D.G. The glycosaminoglycan chain of decorin plays an important role in collagen fibril formation at the early stages of fibrillogenesis. *FEBS J.* **274**, 4246–4255 (2007).

25. Mascellani, G., Liverani, L., Bianchini, P., Parma, B., Torri, G., Bisio, A., Guerrini, M., & Casu, B. Structure and contribution to the heparin cofactor II-mediated inhibition of thrombin of naturally oversulphated sequences of dermatan sulphate. *Biochem. J.* **296**, 639–648 (1993).

26. Denholm, E.M., Lin, Y.Q., & Silver, P.J. Anti-tumor activities of chondroitinase AC and chondroitinase B: inhibition of angiogenesis, proliferation and invasion. *Eur. J. Pharmacol.* **416**, 213–221 (2001).

27. Lee, P.H., Trowbridge, J.M., Taylor, K.R., Morhenn, V.B., & Gallo, R.L. Dermatan sulfate proteoglycan and glycosaminoglycan synthesis is induced in fibroblasts by transfer to a three-dimensional extracellular environment. *J. Biol. Chem.* **279**, 48640–48646 (2004).

28. Taylor, K.R., & Gallo, R.L. Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation. *FASEB J.* **20**, 9–22 (2006).

29. Hitchcock, A.M., Costello, C.E., & Zaia, J. Glycoform quantification of chondroitin/dermatan sulfate using a liquid chromatography-tandem mass spectrometry platform. *Biochemistry* **45**, 2350–2361 (2006).

30. Sirko, S., von, H.A., Wizenmann, A., Gotz, M., & Faissner, A. Chondroitin sulfate glycolsaminoglycans control proliferation, radial glia cell differentiation and neurogenesis in neural stem/progenitor cells. *Development* **134**, 2727–2738 (2007).

31. Gilbert, R.J., McKeon, R.J. Darr, A., Calabro, A., Hascall, V.C., & Bellamkonda, R.V. CS-4,6 is differentially upregulated in glial scar and is a potent inhibitor of neurite extension. *Mol. Cell Neurosci.* **29**, 545–558 (2005).

32. Koprivica, V., Cho, K.S., Park, J.B., Yiu, G., Atwal, J., Gore, B., Kim, J.A., Lin, E., Tessier-Lavigne, M., Chen, D.F., & He, Z. EGFR activation mediates inhibition of axon regeneration by myelin and chondroitin sulfate proteoglycans. *Science* **310**, 106–110 (2005).

33. Kawashima, H., Atarashi, K., Hirose, M., Hirose, J., Yamada, S., Sugahara, K., & Miyasaka, M. Oversulfated chondroitin/dermatan sulfates containing GlcAbeta1/IdoAalpha1-3GalNAc (4,6-O-disulfate) interact with L- and P- selectin and chemokines. *J. Biol. Chem.* **277**, 12921–12930 (2002).

34. Penc, S.F., Pomahac, B., Winkler, T., Dorschner, R.A., Eriksson, E., Herndon, M., & Gallo, R.L. Dermatan sulfate released after injury is a potent promoter of fibroblast growth factor-2 function. *J. Biol. Chem.* **273**, 28116–28121 (1998).

35. Taylor, K.R., Rudisill, J.A., & Gallo, R.L. Structural and sequence motifs in dermatan sulfate for promoting fibroblast growth factor-2 (FGF-2) and FGF-7 activity. *J. Biol. Chem.* **280**, 5300–5306 (2005).

36. Villena, J., & Brandan, E. Dermatan sulfate exerts an enhanced growth factor response on skeletal muscle satellite cell proliferation and migration. *J. Cell Physiol* **198**, 169–178 (2004).

37. Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J., & Rees, D.C. Heparin structure and interactions with basic fibroblast growth factor. *Science* **271**, 1116–1120 (1996).

38. Steinmann, B., Royce, P., & Superti-Furga, A. The Ehlers-Danlos syndrome: In Royce, P., Steinmann,B. editors. Connective tissue and heritable disorders. 2nd edn. New York: Whiley-Liss, Inc. pp. 431–523 (2002).

39. Vogel, A., Holbrook, K.A., Steinmann, B., Gitzelmann, R., & Byers, P.H. Abnormal collagen fibril structure in the gravis form (type I) of Ehlers-Danlos syndrome. *Lab Invest* **40**, 201–206 (1979).

40. Sokolov, B.P., Prytkov, A.N., Tromp, G., Knowlton, R.G., & Prockop, D.J. Exclusion of COL1A1, COL1A2, and COL3A1 genes as candidate genes for Ehlers-Danlos syndrome type I in one large family. *Hum. Genet.* **88**, 125–129 (1991).

41. Nicholls, A.C., Oliver, J.E., McCarron, S., Harrison, J.B., Greenspan, D.S., & Pope, F.M. An exon skipping mutation of a type V collagen gene (COL5A1) in Ehlers-Danlos syndrome. *J. Med. Genet.* **33**, 940–946 (1996).

42. Kresse, H., Rosthoj, S., Quentin, E., Hollmann, J., Glössl, J., Okada, S., & Tonnesen, T. Glycosaminoglycan-free small proteoglycan core protein is secreted by fibroblasts from a patient with a syndrome resembling progeroid. *Am. J. Hum. Genet*. **41**, 436–53 (1987).

43. Quentin, E., Gladen, A., Rodèn, L., & Kresse, H. A genetic defect in the biosynthesis of dermatan sulfate proteoglycan: galactosyltransferase I deficiency in fibroblasts from a patient with a progeroid syndrome. *Proc. Natl. Acad. Sci. USA*. **87**, 1342–1346 (1990).

44. Götte, M. , Kresse, H. Defective glycosaminoglycan substitution of decorin in a patient with pro-geroid syndrome is a direct consequence of two point mutations in the galactosyltransferase I (ß4Gal-T7) gene. *Biochem. Genet.* **43**, 67–79 (2005).

45. Faiyaz-Ul-Haque, M., Zaidi, S.H., Al Ali, M., Al Mureikhi, M.S., Kennedy, S., Al Thani, G., Tsui, L.C., & Teebi, A.S. A novel missense mutation in the galactosyltransferase-I (B4GALT7) gene in a family exhibiting facioskeletal anomalies and Ehlers-Danlos syndrome resembling the progeroid type. *Am. J. Med. Genet*. **128**A, 39–45 (2004).

46. Seidler, D.G., Faiyaz-Ul-Haque, M., Hansen, U., Yip, G.W., Zaidi, S.H., Teebi, A.S., Kiesel, L., Götte, M. Defective glycosylation of decorin and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers-Danlos syndrome patient carrying the novel Arg270Cys substitution in galactosyltransferase I (beta4GalT-7). *J. Mol. Med.* **84**, 583–594 (2006).

47. Neame, P.J., Choi, H.U., & Rosenberg, L.C. The primary structure of the core protein of the small, leucine-rich proteoglycan (PG I) from bovine articular cartilage. *J Biol Chem* **264**, 8653–8661 (1989).

48. Zamfir, A., Seidler, D.G., Kresse, H., & Peter-Katalinic, J. Structural characterization of chondroitin/dermatan sulfate oligosaccharides from bovine aorta by capillary electrophoresis and electrospray ionization quadrupole time-of-flight tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **16**, 2015–2024 (2002).

49. Calabro, A., Benavides, M., Tammi, M., Hascall, V.C., & Midura, R.J. Microanalysis of enzyme digests of hyaluronan and chondroitin/dermatan sulfate by fluorophore-assisted carbohydrate electrophoresis (FACE). *Glycobiology* **10**, 273–281 (2000a).

50. Calabro, A., Hascall, V.C., & Midura, R.J. Adaptation of FACE methodology for microanalysis of total hyaluronan and chondroitin sulfate composition from cartilage. *Glycobiology* **10**, 283–293 (2000b).

51. Seidler, D.G., Breuer, E., Grande-Allen, K.J., Hascall, V.C., & Kresse, H. Core protein dependence of epimerization of glucuronosyl residues in galactosaminoglycans. *J. Biol. Chem.* **277**, 42409–42416 (2002).

52. Zamfir, A., Seidler, D.G., Kresse, H., & Peter-Katalinic, J. Structural investigation of chondro- itin/ dermatan sulfate oligosaccharides from human skin fibroblast decorin. *Glycobiology* **13**, 733–742 (2003).

53. Zamfir, A., Seidler, D.G., Schönherr, E., Kresse, H., & Peter-Katalinic, J. On-line sheath- less capillary electrophoresis/nanoelectrospray ionization-tandem mass spectrometry for the analysis of glycosaminoglycan oligosaccharides. *Electrophoresis* **25**, 2010–2016 (2004).

54. Zhang, Z., Xie, J., Zhang, F., & Linhardt, R.J. Thin-layer chromatography for the analysis of glycosaminoglycan oligosaccharides. *Anal. Biochem.* **371**, 118–120 (2007).

55. Laremore, T.N., & Linhardt, R.J. Improved matrix-assisted laser desorp tion/ionization mass spectrometric detection of glycosaminoglycan disaccharides as cesium salts. *Rapid Commun Mass Spectrom.* **21**, 1315–1320 (2007).

56. Hitchcock, A.M., Yates, K.E., Shortkroff, S., Costello, C.E., & Zaia, J. Optimized extrac- tion of glycosaminoglycans from normal and osteoarthritic cartilage for glycomics profiling. *Glycobiology* **17**, 25–35 (2007).

57. Zaia, J., Miller, M.J., Seymour, J.L., & Costello, C.E. The role of mobile protons in nega- tive ion CID of oligosaccharides. *J. Am. Soc. Mass Spectrom.* **18**, 952–960 (2007).

58. Miller, M.J., Costello, C.E., Malmström, A., & Zaia, J. A. tandem mass spectrometric ap- proach to determination of chondroitin/dermatan sulfate oligosaccharide glycoforms. *Gly- cobiology* **16**, 502–513 (2006).

59. Schenauer, M.R., Yu, Y., Sweeney, M.D., & Leary, J.A. CCR2 chemokines bind selectively to acetylated heparan sulfate octasaccharides. *J Biol Chem.* **282**, 25182–25188 (2007).

60. Wolff, J.J., Chi, L., Linhardt, R.J., & Amster, I.J. Distinguishing glucuronic from iduronic acid in glycosaminoglycan tetrasaccharides by using electron detachment dissociation *Anal Chem.* **79**, 2015–2022 (2007b).

61. Wolff, J.J., Amster, I.J., Chi, L., & Linhardt, R.J. Electron detachment dissociation of gly- cosaminoglycan tetrasaccharides. *J. Am. Soc. Mass Spectrom.* **18**, 234–244 (2007a).

62. Seidler, D.G., & Dreier, R. Decorin and its glycosaminoglycan chain: regulator for cellular function. IUBMB *Life* submitted.

8. SITE SPECIFIC IDENTIFICATION OF N-LINKED GLYCOSYLATION IN PROTEINS BY LIQUID CHROMATOGRAPHY–ELECTROSPRAY IONIZATION TANDEM MASS SPECTROMETRY

IRINA PERDIVARA[1,2], ROXANA ELENA IACOB[2,3], MICHAEL PRZYBYLSKI[1], AND KENNETH B. TOMER[2]

[1]*Laboratory of Analytical Chemistry, University of Konstanz, Germany*
[2]*Laboratory of Structural Biology, Mass Spectrometry group, NIEHS/NIH, Research Triangle Park, NC, USA*
[3]*The Barnett Institute, Northeastern University, Boston, MA, USA*

Abstract. Recently, we reported the characterization of the glycans attached at the 11 N-glycosylation sites of Hepatitis C virus E2 envelope glycoprotein by tandem mass spectrometry. Infections caused by Hepatitis C virus represent the main cause of liver diseases such as hepatitis, cirrhosis and hepatocellular carcinoma. The N-linked sugars consist primarily of high mannose glycans, with structures ranging from the minimal core structure, $Man_3GlcNAc_2$ (Man3) up to 12 hexose residues attached to the GlcNAc-ß(1-4)-GlcNAc core (depicted as $Hex_3Man_9GlcNAc_2$). Furthermore, the site N41 *(N423)* was observed to contain complex type glycans with the structures Man3-GlcNAc and Man3-GlcNAcFuc, in addition to the high mannose population Man3 through Man6, while the site N48 *(N430)* was occupied exclusively with complex type glycans (Man3-Fuc, Man3-GlcNAcFuc and $Man3-GlcNAc_2Fuc$). The present contribution summarizes our experimental observations upon the factors which may have an impact on the CID tandem mass spectra of glycopeptides.

1.   Introduction

Protein glycosylation represents an important post-translational event involved in cellular localization of proteins, folding, proteolytic stability and biological lifetime, signal transduction and protein-protein interactions [1–4]. Most of cell-surface and secreted proteins are modified by N-linked glycosylation [5] which occurs at the consensus sequence Asn-Xxx-Ser/Thr/Cys, where Xxx can be any amino acid except for proline [6,7]. Unlike other post-translational modifications (e.g., phosphorylation), N-glycans are expressed as a set of structural variations on a conserved pentasaccharide core ($Man_3GlcNAc_2$). Thus, a specific N-consensus site may be occupied by several different glycans (site microheterogeneity), whereas macroheterogeneity arises from various such N-motifs contained in a polyprotein sequence. Structural characterization of individual glycoforms and their relative site occupancies in glycoproteins has become increasingly important for understanding biological functions exhibited by this class of biomolecules, as well as for quality control of recombinant protein pharmaceuticals [8].

Recently, mass spectrometry (MS) using soft ionization methods [9,10] has become an invaluable tool for the analysis of post-translational modifications, in particular for protein glycosylation [11–13]. Several approaches have become commonly used to address this issue: (i) analysis of the N-linked carbohydrates released from the protein by enzymatic procedures followed by derivatization of the glycan pool [14], or (ii) analysis of the glycopeptides formed by proteolytic degradation of the protein [15,16]. While the former provides a global picture of the glycans decorating the protein, the second strategy discloses information about the carbohydrate attachment site, which can be obtained using various tandem mass spectrometric methods (reviewed in [17]).

Recently, we reported the mass spectrometric characterization of the N-glycans attached on the Hepatitis C virus (HCV) E2 glycoprotein [18] (DOI: 10.1016/j.jasms.2007.11.022). Hepatitis C viral infections represent the main cause of liver diseases in humans, including chronic hepatitis, cirrhosis or hepatocellular carcinoma. HCV is a small, enveloped positive-strand RNA virus belonging to the Flaviviridae family [19]. The polyprotein encoded in the HCV genome is co- and post-translationally processed by host and viral peptidases into four structural (Core, E1, E2 and p7) and five non-structural proteins [20]. The envelope proteins E1 and E2, containing 6 and 11 Asn-consensus motifs, respectively, are heavily glycosylated [21]. Using a combination of proteases (trypsin and α-chymotrypsin) and nano-scale liquid chromatography in combination with collision induced dissociation (CID) tandem MS, it was possible to characterize the glycoforms associated with each of the 11 glycosylation sites. The major glycans at 9 of the 11 sites are exclusively of high mannose type, comprising species ranging from the minimal core structure $Man_3GlcNAc_2$ (Man3) up to 12 hexose residues elongating the chitobiose core ($Hex_3Man_9GlcNAc_2$), whereas two sites, N41 *(N423)* and N48 *(N430)* proved to be occupied with complex type oligosaccharides of the following sugar composition: Man3-GlcNAc and Man3-GlcNAcFuc at the site N41 and Man3-Fuc, Man3-GlcNAcFuc, $Man3-GlcNAc_2Fuc$ at the site N48 [18]. Digestion by α-chymotrypsin led to formation of glycopeptides of different amino acid compositions containing the same glycosylation site, whose identity was confirmed by tandem MS. However, at specific collision energies, the MS/MS analyses of these glycopeptides revealed that specific fragmentation pathways may be preferred depending on: (i) amino acid composition; length of the peptide; and (ii) nature of the glycan. The present contribution summarizes our experimental observations upon the factors which may have an impact on the CID tandem mass spectra of glycopeptides.

2.   Experimental

2.1.   MATERIALS

HCV E2 envelope glycoprotein (recombinant) was purchased from Austral Bio-logicals (San Ramon, CA). Urea, dithiothreitol, iodoacetamide, 96% formic acid, ammonium bicarbonate and ethanol were purchased from Sigma-Aldrich (St. Louis, MO). Sequencing grade-modified trypsin was obtained from Promega (Madison, WI). Chymotrypsin was purchased from Roche Diagnostics Corporation (Indianapolis, IN). Acetonitrile was purchased from Caledon Laboratories, Ltd. (Georgetown, Ontario). Purified water (17.8 MΩ) was obtained from an in-house Hydro Picopure 2 system. All chemicals were used without further purification unless otherwise specified.

2.2.   SAMPLE PREPARATION

HCV E2 protein expressed in Chinese hamster ovary (CHO) cells was reduced, alkylated and digested with either trypsin or chymotrypsin following the procedure described in [18].

2.3.   MASS SPECTROMETRY

nanoLC/MS/MS analyses were performed on a Waters Q-Tof Premier mass spectrometer equipped with a nanoAcquity UPLC system and a NanoLockspray source (Waters, Milford, MA). Separations were performed using a 3 μm nanoAquity Atlantis column dC18 100 μm x 100 mm (Waters) at a flow rate of 300 nl/min. A nanoAquity trapping column 5 μm C18 180 μm x 200 mm (Waters) was positioned in-line with the analytical column. Trapping of a 2 μl aliquot of the digested sample was performed for 3 min at 5 μl/min flow rate. Peptides were eluted using a linear gradient from 98% A (water/ 0.1% formic acid (v/v)) and 2% B (acetonitrile/ 0.1% formic acid (v/v)) to 95% B over 60 min. Mass spectrometer settings for MS analyses were as follows: capillary voltage of 3.5 kV, cone voltage of 30 V, collision energy of 8.0 V, and scan range of 200–2,000 Da. The tandem mass spectra were obtained in the data dependant acquisition mode, and a fixed range of collision energies between 30 and 40 V were applied in order to ensure good fragmentation of the glycopeptides. Peptides were identified by mass and charge measurement compared to the *in silico* digestion using the BioLynx Protein/ Peptide Editor, feature of MassLynx V4.0 (Micromass, UK).

3.   Results and Discussion

The amino acid sequence of the HCV E2 glycoprotein is shown in Figure 1, with
the numbering starting from Ala1 to Lys333 (which corresponds to the numbering
Ala383 to Lys715 within the entire HCV polyprotein of reference strain H, Gen-
Bank accession number 009606: this numbering is shown in parentheses). The 11
glycosylation sites, highlighted in bold, were found to be primarily occupied by
high-mannose glycans at nine sites and by complex type glycans at the remaining
two (N41 *(N423)* and N48 *(N430)*). The population of the observed high-mannose
glycans was highly heterogeneous, containing structures ranging from $Man_3GlcNAc_2$
(Man3) up to $Hex_3Man_9GlcNAc_2$ (Hex3Man9). The primary sequence of E2 con-
tains only a few trypsin cleavage sites, and, therefore, longer proteolytic fragments,
containing multiple glycosylation sites (which were not observed in MS), were
formed when trypsin was used to generate the glycopeptides. Consequently, only
4 out of 11 sites were observed in the MS analysis of the tryptic digest. In order to
elucidate the glycosylation pattern at the remaining sites, chymotrypsin was used
to produce shorter proteolytic fragments, which would contain a single glycosyla-
tion position within each peptide.

```
  1  AETHVTGGNA  GRTTAGLVGL  LTPGAKQNIQ  LINTNGSWHI  NSTALNCNES

 51  LNTGWLAGLF  YQHKFNSSGC  PERLASCRRL  TDFAQGWGPI  SYANGSGLDE

101  RPYCWHYPPR  PCGIVPAKSV  CGPVYCFTPS  PVVVGTTDRS  GAPTYSWGAN

151  DTDVFVLNNT  RPPLGNWFGC  TWMNSTGFTK  VCGAPPCVIG  GVGNNTLLCP

201  TDCFRKYPEA  TYSRCGSGPR  ITPRCMVDYP  YRLWHYPCTI  NYTIFKVRMY

251  VGGVEHRLEA  ACNWTRGERC  DLEDRDRSEL  SPLLLSTTQW  QVLPCSFTTL

301  PALSTGLIHL  HQNIVDVQYL  YGVGSSIASW  AIK
```

Figure 1. Amino acid sequence of the envelope glycoprotein E2, starting from amino acid 1
to 333 (383-715) from strain HCV-1a (GenBank accession number AF 009606). The 11
glycosylation sites are shown in bold

3.1.   CID OF GLYCOPEPTIDES WITH IDENTICAL PEPTIDE BACKBONE
CONTAINING DISTINCT GLYCANS

The site N41 *(N423)*, observed in the chymotryptic peptide 39–45 (HIN*STAL),
was found to be occupied by both high mannose and complex type glycans. The
high mannose sugars comprise structures ranging from Man3 to Man6, whereas
the structures determined for the two complex type glycans were Man3-GlcNAc
and Man3-GlcNAcFuc, respectively.
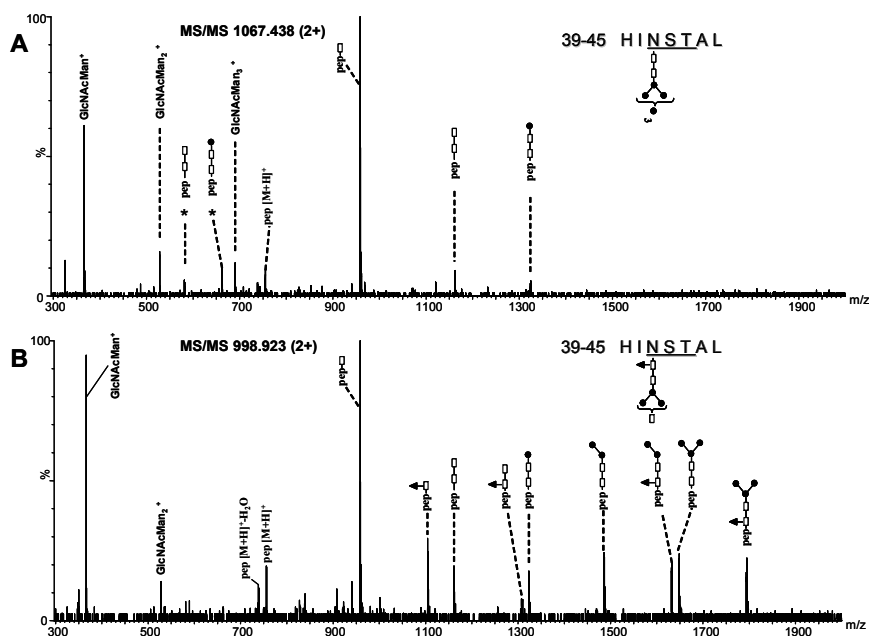
Figure 2. MS/MS spectra of the peptide 39–45 containing: A. a high mannose glycan of the structure Man6, and B. a complex type glycan of the structure Man3GlcNAcFuc. The spectra were obtained in the data dependent mode with a collision energy ramp from 30 to 40 V. The ions indicated with an asterisk (*) are doubly charged, all the other ions are singly charged

Figure 2 shows the MS/MS analyzes of the peptide 39–45 containing the Man6 carbohydrate (Figure 2A) and the complex type glycan Man3-GlcNAcFuc (Figure 2B). The precursor ions of *m/z* 1,067.438 and 998.923 respectively, are doubly charged while the majority of the fragment ions observed in these spectra are singly charged. The ion of *m/z* 958.480, representing the base peak in both tandem mass spectra, was assigned to the intact peptide 39–45 retaining the first *N*-acetyl glucosamine of the chitobiose core and allowed for the identification of the peptide carrying the glycan 39–45 with a the mass accuracy of 5 ppm. Complete loss of the glycan moiety was observed, as the singly charged ion *m/z* 755.398 represents the singly protonated aglycone.

The ions observed in the spectra shown in Figure 2 result from B/Y-type fragmentations of the oligosaccharide (nomenclature after Domon and Costello [22]) and are accompanied by charge reduction of the precursor ion. An additional fragmentation pathway, arising primarily from the loss of the core fucose with charge reduction, is observed in the MS/MS of the glycopeptide containing the complex glycan.
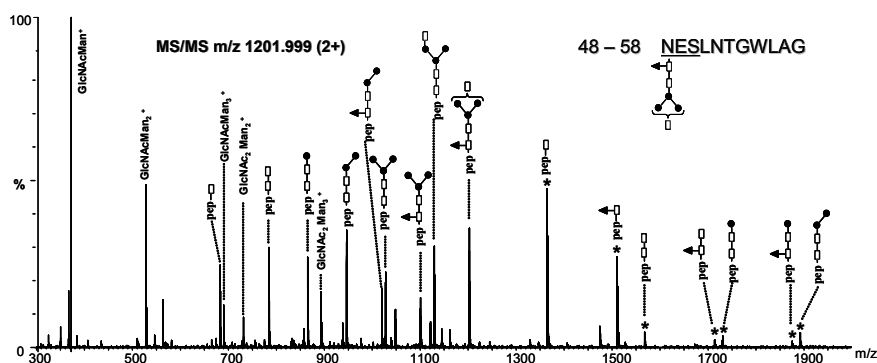
Figure 3. MS/MS spectrum of the doubly protonated precursor of *m/z* 1,201.95, representing the peptide 48–58 containing the glycan Man3GlcNAcFuc. The spectrum was obtained in the data dependent mode by ramping the collision energy from 30 to 40 V. The singly charged ions containing the peptide backbone are indicated with an asterisk (*)

The MS/MS spectrum of the doubly protonated glycopeptide *m/z* 1,201.999, which corresponds to the amino acid sequence 48–58 containing the glycosylation site N48 *(N430)* which is occupied by complex glycan of the same composition as that on peptide 39–45, is shown in Figure 3. This glycopeptide was formed by non-specific chymotrypsin activity, which cleaved the protein backbone at the alkylated Cys47 and Gly58. As in the MS/MS spectrum of glycopeptide 39–45, singly charged fragments arise from B/Y-type fragmentations and allow for the identification of the glycan composition. Another series, however, of doubly charged ions is observed that arise by the sequential loss of neutral monosaccharides from the non-reducing end with the charge on the remaining Y-fragment. These results suggest that the decomposition of glycopeptides under identical fragmentation conditions can be dependent on: (i) the composition and identity of the glycan, because the presence of the labile fucose on the chitobiose core introduces a new fragmentation channel which competes with the fragmentation from the non-reducing end, and (ii) amino acid composition and length of the glycopeptide, as the longer peptides more readily retain multiple charges.

3.2. CID OF GLYCOCONJUGATES WITH IDENTICAL GLYCANS ATTACHED TO THE SAME SITE IN DISTINCT PEPTIDES

The glycosylation site N263 *(N645)*, carrying exclusively high mannose glycans of composition Man5 to Man9, was observed in two glycopeptides, 246–272 and 251–264, respectively. Figure 4 shows the MS/MS spectra obtained for the triply charged precursor ions of *m/z* 1,547.855 and 992.743 of these glycopeptides, both having the carbohydrate Man6 attached at the asparagine residue. The dominating fragments in Figure 4A are the sugar oxonium ions resulting from processing of

the glycan from the non-reducing end and the doubly charged glycopeptide ion containing a single *N*-acetyl glucosamine (which is the base peak in this spectrum). The complete series of singly charged sugar fragments provides structural information about the oligosaccharide attached to the peptide.
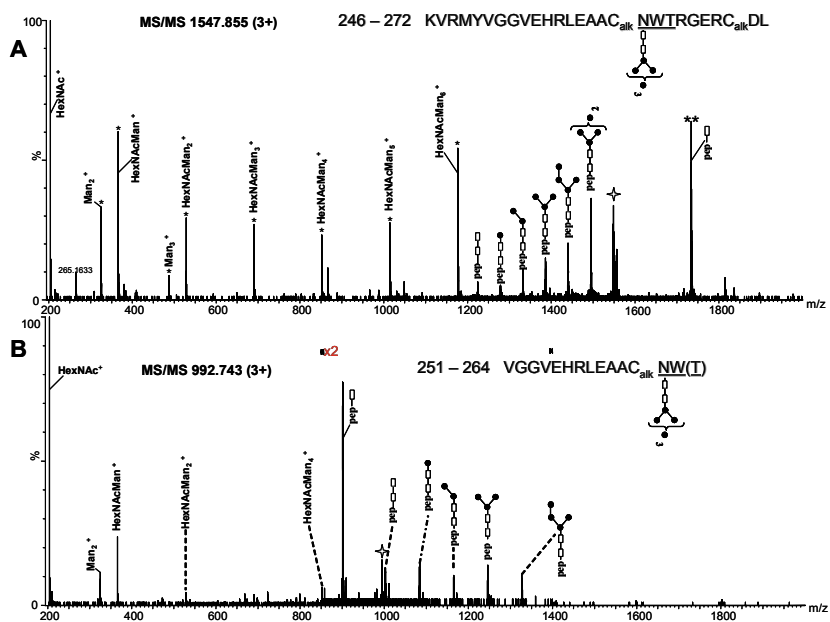
Figure 4. MS/MS spectra of glycopeptides containing the site N263 (N645) and a high mannose glycan of composition Man6: (A) triply charged precursor of *m/z* 1,547.855 assigned to the glycopeptide 246–272. The precursor ion is highlighted with a star. Singly charged sugar oxonium ions are indicated with an asterisk (*). Doubly charged ions are depicted with a double asterisk (**). The ions in the range *m/z* (1,200–1,600) are triply charged; (B) triply charged precursor of *m/z* 992.840 assigned to the glycopeptide 251–264. The precursor is indicated with a star. The ions in the range *m/z* (900–1,400) (two times magnified) are doubly charged. The spectra were obtained with a collision energy ramp from 30 to 40 V

The triply charged fragment ions in the range *m/z* (1,200–1,600) arise from stepwise neutral loss of the individual monosaccharides from the free oligosaccharide end, with the minimal glycan retained on the peptide backbone being chitobiose. Except for the most abundant doubly protonated molecular species in this spectrum, observed at *m/z* 1,733.430, no significant fragmentation with charge reduction of the precursor was observed. By comparison, the MS/MS shown in Figure 4B contains almost exclusively doubly charged fragments resulted from the charge reduction fragmentation pathway. Compared with Figure 4A, only a few sugar oxonium ions of low abundance were observed. In the MS/MS spectrum of

glycopeptide 246–272, the fragment ion containing the first N-acetyl glucosamine residue, formed by B/Y-type fragmentation and charge reduction, represents the most abundant ion in this spectrum.

Peptide 246–272, compared to peptide 251–264, extends five amino acids longer at the N-terminus (two of which are basic residues) and seven amino acids longer at the C-terminus. This may explain the enhanced probability for retention of three charges on the fragment ions observed in Figure 4A in the mass range 1,200–1,600. Thus, it appears that the amino acid composition and the length of the peptide containing the glycan determine the dominant decomposition pathways of glycoconjugates.

3.3. PEPTIDE BACKBONE FRAGMENTATION (IN THE CID OF GLYCOCONJUGATES)

A major characteristic of the CID spectra of glycopeptides is that fragment ions resulting from processing of the glycan [17] are primarily observed, whereas the formation of backbone fragments is less common. While fragmentation of the glycan moiety is achieved using low collision energies (typically 15–25 V), decomposition of the peptide requires elevated collision energies. Wuhrer *et al.* reported the observation of abundant backbone fragmentation using collision energies of 30 V in a quadrupole-FTICR instrument. Under these conditions, however, fragmentation of the glycan was complete and fragment ions due to successive loss of glycans were no longer observed [17]. Other groups used collision energies up to 80 V in order to observe fragmentation of the peptide [15].

The glycosylation site N174 *(N556)*, observed in the chymotryptic peptide 173–178 and in the peptide 171–180, formed by non-specific activity of the enzyme, is occupied with a heterogeneous population of high mannose glycans, ranging from the minimal core structure Man3 up to nine mannoses attached to the chitobiose core (Man9). Representative MS/MS spectra of the glycopeptides 173–178 and 171–180, containing the Man6 oligosaccharide are shown in Figure 5. These spectra were obtained with a collision energy ramp from 30 to 40 V and in both the most abundant ions were assigned the structure of the peptide retaining the first GlcNAc rest. The MS/MS spectrum in Figure 5A displays an abundant ion corresponding to the protonated peptide 173–178, which was formed by the complete loss of the oligosaccharide chain. Remarkably, the MS/MS spectrum of the shorter peptide 173–178 (Figure 5A) contains abundant b and y peptide fragments, while no such fragments were observed in the MS/MS spectrum of the peptide 171–180. The latter peptide contains four additional amino acids compared to peptide 173–178 (Figure 5B). The MS/MS spectrum of the precursor ion of *m/z* 1,275.970 (Figure 5B) consists primarily of ions derived from fragmentation of the glycan retained on the intact peptide. The majority of the observed of the b-type backbone fragment ions (b_3–*m/z* 333.100, b_4–*m/z* 434.151, b_5–*m/z* 491.172), which contain the N-glycosylation site, were also found accompanied by ions retaining
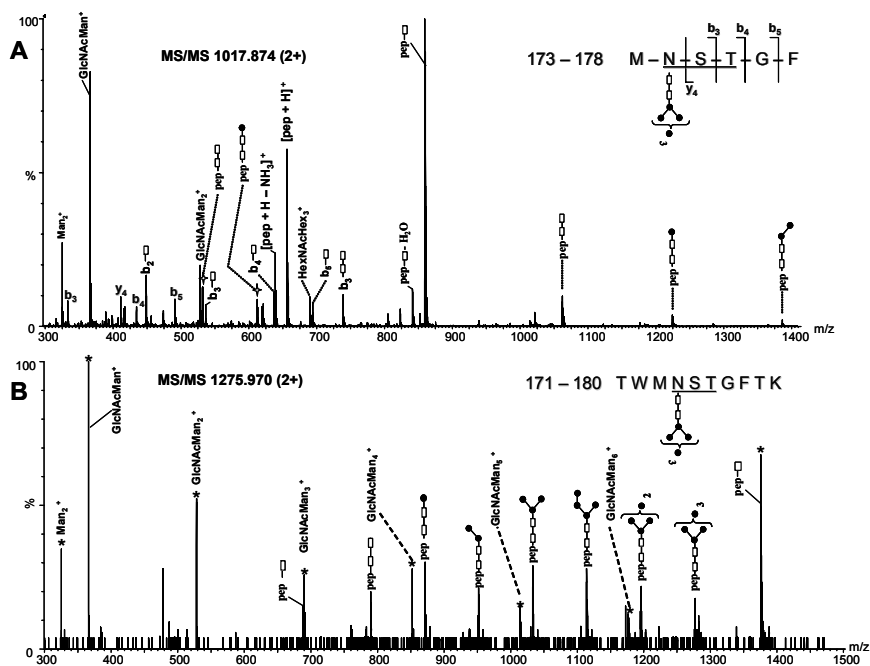
Figure 5. MS/MS spectra of doubly protonated glycopeptide species containing the glycosylation site N174 (N556); (A) precursor ion *m/z* 1,017.83. The ions highlighted with a star are doubly charged; (B) precursor ion *m/z* 1,276.48. The ions indicated with an asterisk (*) are singly charged. The spectra were obtained with a collision energy ramp from 30 to 40 V

the first *N*-acetyl glucosamine residue (b_3–GlcNAc *m/z* 536.221, b_4–GlcNAc *m/z* 637.210, b_5–GlcNAc *m/z* 694.262) and, to a lesser extent, the second GlcNAc (b_3–GlcNAc$_2$–*m/z* 739.242). These ions confirmed the sugar attachment site.

The observation of peptide fragments which partially retain *N*-acetyl glucosamine is in agreement with data published by other groups [15,17], which showed that the GlcNAc was retained on the y-ions. Their data suggests that the formation of such fragments, however, is dependent on the amino acid composition and location of the glycosylation site within the peptide.

Another example in which backbone fragmentation was observed is illustrated in the MS/MS spectrum of the glycopeptide 235–242 containing the site N241 *(N623)* (Figure 6). Although the length of this peptide is comparable with that of the peptide 173–178, only two peptide fragments were observed, b_6 at *m/z* 715.351 and y_6-GlcNAc at *m/z* 913.381, probably because the formation of the y_6 ion is enhanced by the presence of the amino acid motif YP [23].
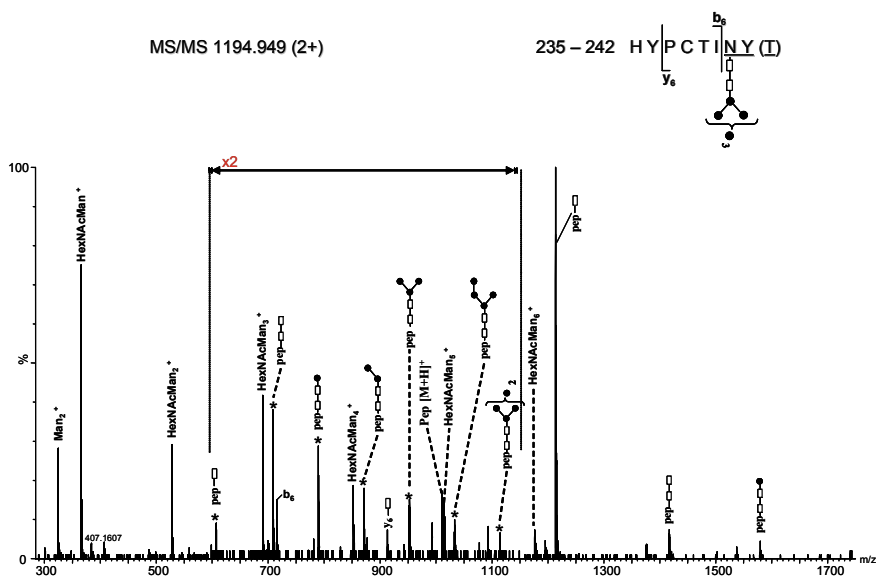
Figure 6. MS/MS spectrum of the doubly charged precursor ion of *m/z* 1,194.898, corre-
sponding to the glycopeptide 235–242 containing the Man6 glycan. For clarity, the *m/z*
region 600–1,150 was magnified by two. Doubly charged ions are indicated with an aster-
isk (*). The spectrum was obtained with a collision energy ramp from 30 to 40 V

    Thus, under identical collision energy conditions, the observation of backbone
frag-mentation appears to be dependent on the amino acid composition of the gly-
copeptide.

4.   Conclusions

Several examples of tandem mass spectra of N-glycopeptides derived from the
HCV E2 envelope glycoprotein, which contain neutral oligosaccharides, were pre-
sented and all contain abundant ions resulted from the fragmentation of the glycan
from the non-reducing end. Optimal decomposition of the glycopeptides was
achieved using a collision energy ramp from 30 to 40 V. The most abundant ion in
all spectra corresponds, almost exclusively, to the intact peptide retaining one
N-acetyl glucosamine residue. Four major pathways for fragmentation of the gly-
copeptides were observed: (i) neutral loss of the individual monosaccharides from
the free end of the sugar with charge retention on the remaining Y-ion; (ii) subse-
quent loss of the monosaccharide residues with charge reduction of the precursor
ion and/or characteristic formation of sugar oxonium ions (B-ions); (iii) if the gly-
can contains a fucose, neutral loss of the fucose residue competes with neutral loss
from the non-reducing end; (iv) cleavage of the peptide backbone. From our data

and from the results published by other groups, the rate of each process was observed to depend on several factors, such as: composition of the glycan, number and amino acid sequence of the peptide and the amount of collision energy and conditions must be optimized for each individual experiment in order to obtain optimal information.

References

1. Marsh, M. and A. Helenius, Virus entry: open sesame. Cell, 2006. **124**(4): 729–40.
2. Slawson, C., M.P. Housley, and G.W. Hart, O-GlcNAc cycling: how a single sugar post-translational modification is changing the way we think about signaling networks. J Cell Biochem, 2006. **97**(1): 71–83.
3. Smith, A.E. and A. Helenius, How viruses enter animal cells. Science, 2004. **304**(5668): 237–42.
4. Wormald, M.R. and R.A. Dwek, Glycoproteins: glycan presentation and protein-fold stability. Structure, 1999. **7**(7): R155–60.
5. Apweiler, R., H. Hermjakob, and N. Sharon, On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta, 1999. **1473**(1): 4–8.
6. Kornfeld, R. and S. Kornfeld, Assembly of asparagine-linked oligosaccharides. Annu Rev Biochem, 1985. **54**: 631–64.
7. Gavel, Y. and G. von Heijne, Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. Protein Eng, 1990. **3**(5): 433–42.
8. Srebalus Barnes, C.A. and A. Lim, Applications of mass spectrometry for the structural characterization of recombinant protein pharmaceuticals. Mass Spectrom Rev, 2007. **26**(3): 370–88.
9. Fenn, J.B., et al., Electrospray ionization for mass spectrometry of large biomolecules. Science, 1989. **246**(4926): 64–71.
10. Hillenkamp, F., et al., Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal Chem, 1991. **63**(24): 1193A–1203A.
11. Wuhrer, M., A.M. Deelder, and C.H. Hokke, Protein glycosylation analysis by liquid chromatography-mass spectrometry. J Chromatogr B Analyt Technol Biomed Life Sci, 2005. **825**(2): 124–33.
12. Zaia, J., Mass spectrometry of oligosaccharides. Mass Spectrom Rev, 2004. **23**(3): 161–227.
13. Huddleston, M.J., M.F. Bean, and S.A. Carr, Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests. Anal Chem, 1993. **65**(7): 877–84.
14. Ciucanu, I., Per-O-methylation reaction for structural analysis of carbohydrates by mass spectrometry. Anal Chim Acta, 2006. **576**(2): 147–55.
15. Harazono, A., et al., Site-specific N-glycosylation analysis of human plasma ceruloplasmin using liquid chromatography with electrospray ionization tandem mass spectrometry. Anal Biochem, 2006. **348**(2): 259–68.
16. Wuhrer, M., et al., IPSE/alpha-1, a major secretory glycoprotein antigen from schistosome eggs, expresses the Lewis X motif on core-difucosylated N-glycans. Febs J, 2006. **273**(10): 2276–92.
17. Wuhrer, M., et al., Glycoproteomics based on tandem mass spectrometry of glycopeptides. J Chromatogr B Analyt Technol Biomed Life Sci, 2007. **849**(1-2): 115–28.

18. Iacob, R., Perdivara, I., Przybylski, M., and Tomer, K.B., Mass spectrometric characteri-zation of glycosylation of Hepatitis C virus E2 envelope glycoprotein reveals extended micro-heterogeneity of N-glycans. JASMS, 2007.

19. Rehermann, B. and M. Nascimbeni, Immunology of hepatitis B virus and hepatitis C virus infection. Nat Rev Immunol, 2005. **5**(3): 215–29.

20. Randall, G. and C.M. Rice, Hepatitis C virus cell culture replication systems: their potential use for the development of antiviral therapies. Curr Opin Infect Dis, 2001. **14**(6): 743–7.

21. Duvet, S., et al., Hepatitis C virus glycoprotein complex localization in the endoplasmic reticulum involves a determinant for retention and not retrieval. J Biol Chem, 1998. **273**(48): 32088–95.

22. Domon, B. and Costello, C.E., A Systematic Nomenclature for Carbohydrate Fragmenta-tions in FAB-MS/MS Spectra of Glycoconjugates. Glycoconjugate J, 1988. **5**: 397–409.

23. Paizs, B. and S. Suhai, Fragmentation pathways of protonated peptides. Mass Spectrom Rev, 2005. **24**(4): 508–48

.

Part IV

IMMUNOLOGY

9. CHARACTERIZATION OF IMMUNE RESPONSES TO PATHOGEN CHALLENGE BY MS-BASED EPITOPE MAPPING

JASON G. WILLIAMS[*], LEESA J. DETERDING[*],
AND KENNETH B. TOMER
Laboratory of Structural Biology, Mass Spectrometry Group, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC, USA

Abstract. Mass spectrometry has long proven to be an outstanding bioanalytical technique for identifying proteins, defining their amino acid sequences, and for identifying sites of protein modifications. The application of mass spectrometry to more complicated biological structures, however, is less well established. We have been applying mass spectrometry to structural studies protein, especially to the determination of epitopes on proteins from pathogens that are recognized by monoclonal antibodies as part of the body's immune system. As part of the defense mechanism against the invading pathogens, the body produces a cadre of antibodies to various epitopes on the pathogens. Here, we outline the development of epitope mapping techniques by mass spectrometry for the identification of linear epitopes and complex discontinuous epitopes using examples from our studies of epitopes on HIV proteins.

1.   Introduction

Over the past decade or more, we have been developing and applying mass spectrometry-based techniques to the characterization of the response of the immune system to challenge by pathogens. The broad goals of our research program include: enhanced understanding of the immune system; the identification of antigenic determinants (epitopes), that is, the areas of pathogen-produced proteins that are recognized by antibodies expressed in response to the presence of the pathogen; and clinical aspects, such as detection of the presence of pathogens and vaccine development. This report will focus on the identification of antigenic determinants recognized by monoclonal antibodies.

Antibodies are produced on the surface of B lymphocytes in response to immunogens, primarily foreign proteins and polysaccharides and these antibodies are often secreted into the blood and other body fluids. The most prevalent antibody in human plasma or other bodily fluids is immunoglobin G, IgG. IgG is a glycoprotein consisting of two heavy chains and two light chains bound by disulfide bonds, and it has a molecular weight of approximately 150 kD. The region of the antibody that recognizes foreign substances is called the paratope and is formed by the variable regions on the heavy and light chains. Monoclonal antibodies recognize a

---

[*] Jason G. Williams and Leesa J. Deterding contributed equally

123

specific epitope while polyclonal antibodies are a mixture of antibodies, often times recognizing different regions on a protein.

Protein epitopes recognized by antibodies are usually around 12–15 residues in length, though much of the binding affinity may reside in a shorter sequence. Structurally, epitopes can be linear continuous, linear conformational, or discontinuous and conformational (Figure 1).
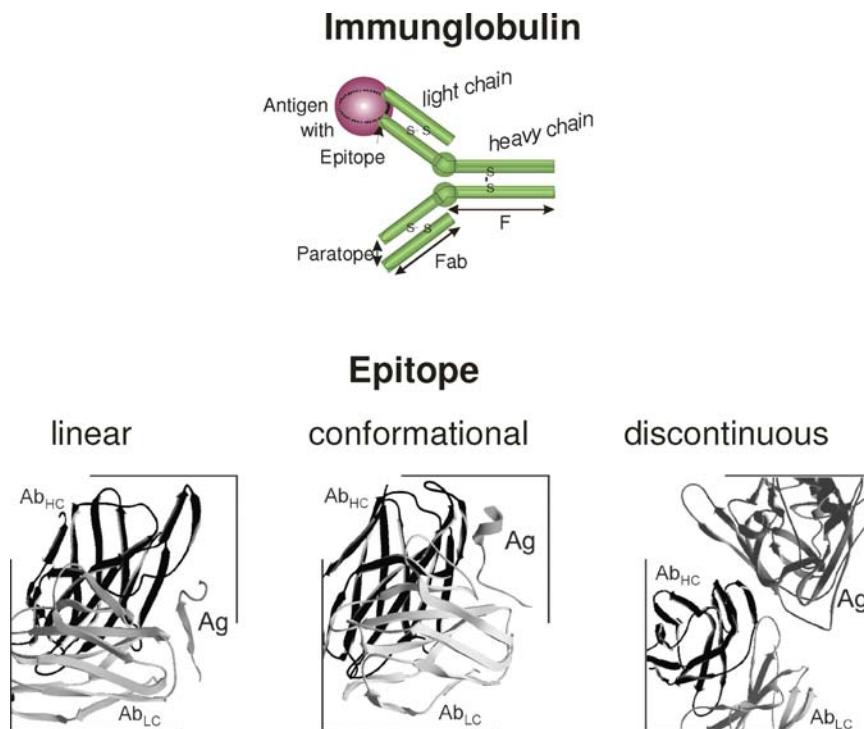
Figure 1. Cartoon of immunoglobulin and the three types of epitopes showing the interaction between the antibody (Ab) and the antigen (A) (Reproduced from Expert Rev. Proteomics. 2(5), 745–756 (2005) with permission of Future Drugs Ltd [1])

Linear continuous epitopes are the simplest and least prevalent. Linear conformational epitopes are epitopes with defined secondary structures, such as helices, in which not all of the amino acids contact the antibody, but are required to maintain the contacting residues in a conformation where they can be recognized by the antibody. Discontinuous conformational epitopes are epitopes comprised of residues sequentially distant that are brought into close proximity in the native form of the protein. Most naturally occurring antibodies recognize epitopes that are conformational in nature.

Defining the antigenic determinant on pathogen proteins provides insight into how the immune system reacts to the presence of pathogens and also provides

information that can be useful in vaccine development. The most common method for identifying epitopes recognized by monoclonal antibodies is the use of short (6-7 amino acids in length) peptide sequences in ELISA assays. Although highly sensitive, this approach is only valid for linear epitopes and some linear conformational epitopes. Additionally, the technique can identify sequences that are not on the surface of the protein and, therefore, not accessible to the antibody under native conditions. These are termed crypto-epitopes. The most accurate picture of the interaction of an antibody with the protein that it recognizes is by X-ray crystallography, but this approach has been limited by the potential difficulties in obtaining crystals from sufficient material in sufficient purity. In 1986, Jemmerson and Paterson published an article in which they used limited proteolysis, HPLC separation and amino acid analysis to define epitopes on a native protein recognized by monoclonal antibodies[2]. More recently, gene deletion has also been used to localize epitopes on native proteins.

2.   Mass Spectrometry-Based Epitope Mapping

In 1990, Przybylski and co-workers published the use of immobilized antibodies, limited proteolysis, elution and mass spectrometric characterization of the products to elucidate continuous epitopes [3], and in 1994, our laboratory introduced the use of direct analysis of affinity-bound analytes using matrix-assisted laser desorption/ionization (MALDI) mass spectrometry to identify continuous epitopes [4]. This methodology is outlined in Figure 2. Briefly, during proteolysis of the antigen bound to the antibody, potential cleavage sites protected by the antibody are not cleaved. Moreover, cleavage sites whose access to proteolytic enzymes are sterically hindered by the presence of the antibody are cleaved slower in the presence of the antibody than on the native, unbound protein. For continuous epitopes, the epitope-containing peptide remains bound to the antibody after proteolysis. An aliquot of the immobilized antibody beads is placed on the MALDI target and matrix added. The epitope-containing peptide is then released by the MALDI matrix and characterized by mass spectrometry.

3.   Applications to Linear Epitopes on HIV Proteins

We initially probed continuous epitopes on HIV p24 [5] and HIV gp120 [6]. We then collaborated with Prof. John Moore [5] of the Cornell University Weill Medical School and Prof. Hermann Katinger [7] of the Institute of Applied Microbiology/University of Agriculture in Vienna, Austria on identifying epitopes on HIV gp41 recognized by two of the relatively few known human neutralizing anti-HIV antibodies, 2F5 and 4E10. The glycoprotein gp41 is initially expressed as gp160, a construct of the protein that contains both gp41 and another surface glycoprotein gp120. After expression, gp160 is rapidly cleaved to gp120 and gp41, but the two proteins remain as a complex (three molecules of each) in the

virus. This complex is involved in the syncitium formation during the infection process. ELISA had been used to identify a short, six amino acid sequence in the pre-transmembrane region of gp41 as the epitope recognized by mAb 2F5, but all attempts to use this sequence to produce neutralizing anti-HIV antibodies proved futile [8–11]. Professor Moore hypothesized that the epitope was more complex than just the six amino acid sequence, possibly including regions from HIV gp120.
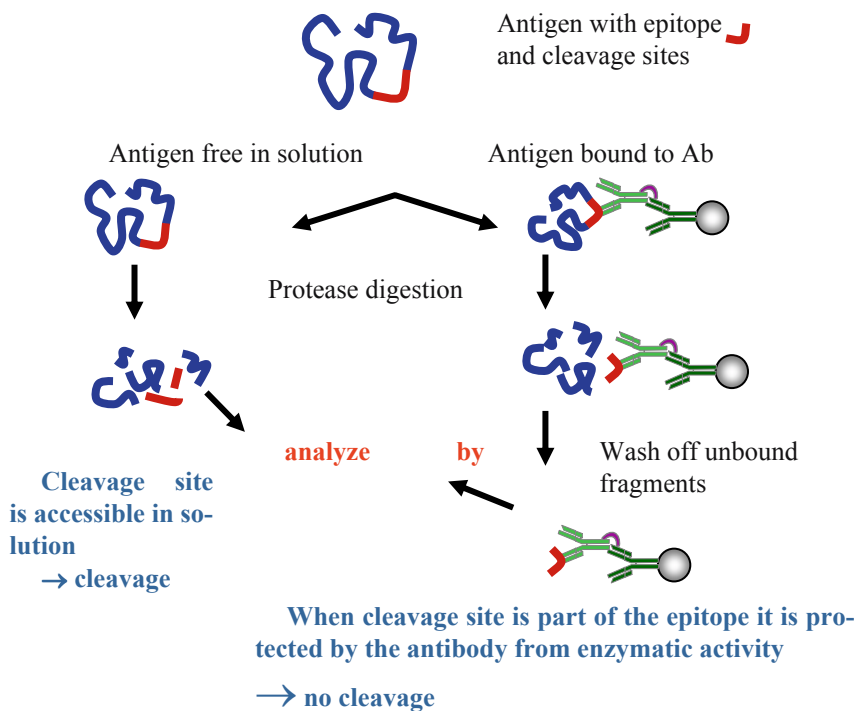
Figure 2. Cartoon schematic of epitope excision with direct MALDI analysis

To mimic the gp41:gp120 complex found on the virus, Moore's lab synthesized a stable HIV gp160 construct, SOS gp160 (Figure 3A), containing a disulfide bond that maintained the connectivity between gp120 and gp41 after expression [12] and the goal of our collaborative effort was to attempt to map the epitope recognized by 2F5 [5].

Using the epitope excision methodology, a 16 amino acid sequence extending both the N- and C-terminal of the ELISA determined epitope was identified (Figures 3A and B). Thus, our results implicated an extended epitope that was probably necessary to maintain the recognized conformation. This result was later confirmed by crystallography that showed there was an almost 90 degree bend in

the molecule. Unfortunately, attempts to synthesize the epitope with this bend still did not yield a highly neutralizing antibody [13–16].

Another human, broadly neutralizing anti-HIV antibody isolated and expressed by the Katinger group is 4E10. Like 2F5, this antibody recognizes a six amino acid sequence in the pre-transmembrane region of gp41 and also like 2F5, the ELISA determined epitope, NWF(N/D)IT [17,18], does not produce neutralizing antibodies [10]. Additionally, some phylogenetic clades of HIV containing the NWF(N/D)IT are not neutralized, while some clades with NWF(N/D)IT variants are neutralized by 4E10 Furhtermore, a crystal structure of 4E10: KGWNWF(N/D)ITNW shows a helical structure with the peptide interacting only with the base and central residues of the CDR H3 loop, the tip of which bends away from epitope binding site, hinting at possible additional interactions [19].
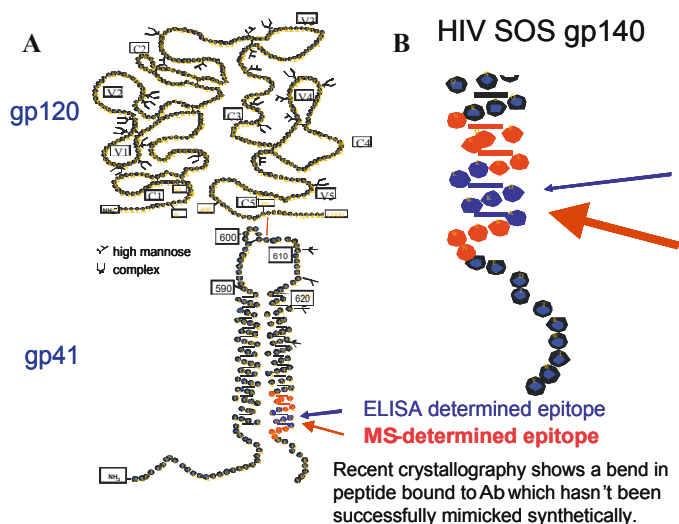
Figure 3. (A) Cartoon of HIV SOS gp140 [12]. ELISA defined epitope recognized by mAb 2F5 is shown in blue, additional residues identified by MS shown in red; (B) expansion of epitope region

Using the intact SOS gp160 protein and several variants of gp41, we found by epitope excision that there was a 30 amino acid region in the fusogenic region of gp41 that remained bound to the Ab (Figure 4) in any epitope excision experiment on any expressed gp41 protein that contained the fusogenic region (Figure 5). Additionally an ion arising from the gp120 moiety of the SOS gp160 construct was also seen. We hypothesized that this fragment may be in complex with the gp41 fragment [7].
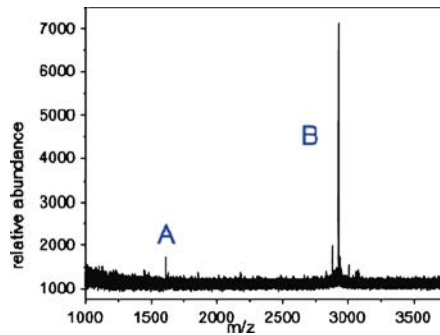
Figure 4. MALDI MS spectrum obtained from epitope excision (trypsin followed by carboxypeptidase Y) of SOS-gp160 bound to mAb 4E10. A= *m/z* 1609 from aa 34-46 of gp120; B=*m/z* 2925.6 of gp41, aa AVGIGAVFLGFLGAASTMGAASMTLTVQAR (Reproduced with permission of the Journal of Immunology [7])
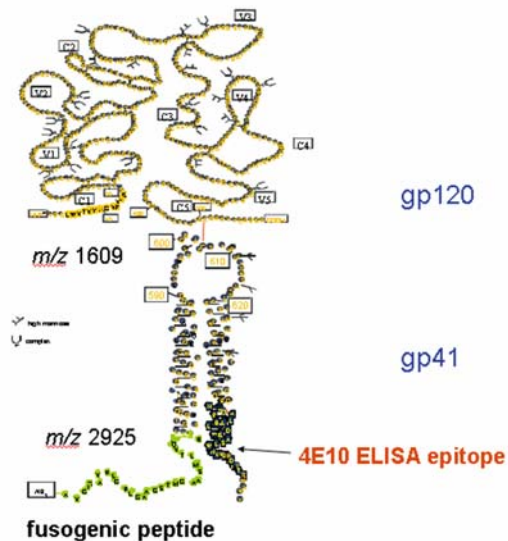
Figure 5. Cartoon of SOS gp160 showing location of the fusogenic peptide identified as the native epitope on SOS gp160 recognized by mAb4E10 in contrast to the ELISA identified epitope

Only when this region was not expressed in the construct was NWF(N/D)IT found bound to the Ab. Based on further MS and ELISA experiments, we hypothesize that the fusogenic peptide on gp41 is recognized by the native protein and that the pre-transmembrane region is only recognized in the absence of the fusogenic region or on denatured gp41.

These results are totally consistent with the observed neutralizing capability of the antibody with gp41 clade variants [18,20]. These results also point out the importance of investigating mAb interactions with the native protein, not just a series of short peptides.

4.   Conformational Epitopes

Although we have had significant success with the epitope excision approach, the majority of epitopes are conformational in nature. This requires approaches that incorporate tertiary and quaternary protein structures. Mass spectrometry is widely recognized as one of the most useful tools for the determination of chemical structures. This capability is especially relevant in the biochemical / biological sciences where it plays a central role in both small molecule mass determination and structural elucidation as well as in protein identification, and identification of posttranslational modifications. These applications rely on the determination of both the mass of analyte and on structural information obtained by tandem mass spectrometric interrogation of the fragment ions of each precursor ion of the analyte.

The capabilities of mass spectrometry in higher order structural studies, especially of proteins and protein complexes, are, however, less widely recognized. As shown in Figure 6, mass spectrometry can be used to probe secondary, tertiary and quaternary protein structures. Underlying all of these higher order structural studies, however, are the same techniques that are used for determination of peptide amino acid sequences and post-translational modifications, i.e., mass determination and MS/MS analyses. The MS approaches are almost universally-based on various differential reactivity experiments, such as H/D exchange or chemical derivatization.

There is an extensive literature about the use of chemical reactivity of protein amino acids to probe the surface accessibility of these residues.

In the mass spectrometry-based approach, isotopic labeling can be used to precisely determine changes in reactivity of specific amino acids in a protein between an unbound and a complex-bound state. It is important to determine differential reactivity because steric effects as well as electronic effects, such as electrostatic interactions, can reduce the reactivity of a residue, not just the surface accessibility. The changes in differential reactivity can subsequently be used, optimally in combination with existing structural data, to identify the recognition surface. This is shown schematically in Figure 7.

Thus, as we became more involved with exploring the determination of discontinuous epitopes, we began to use differential surface modification to provide structural information.

Figure 6. The role of mass spectrometry in protein structural studies showing applications under specific structural aspects with epitope mapping in red

Figure 7. Cartoon of differential surface modification. Triangles represent residues that are modified, red in first step, blue in second. The ratio of red to clear is the extent of modification in the first step. Unmodified (clear) are modified in second step (blue)

Our group has primarily used acylation, typically acetylation, of lysine side chains [21,22] and either arginine phenylglyoxylation [23] or reaction with cyclohex-anedione [24] to determine differential chemical reactivities of surface residues.

Our strategy, as well as a demonstration of the utility of mass spectrometry to help characterize discontinuous epitopes, is illustrated in a study of recognition surfaces of enhancing and non-enhancing antibodies and HIV p24 [22]. A strong $CD4^+$ T cell response to p24 has been shown to correlate with lack of progression of HIV infection to AIDS. In addition, high levels of anti-HIV antibodies recognizing highly-conserved regions of p24 are often present in infected individuals, but the antibodies are not known to possess antiviral capabilities. In fact, several antibodies directed against p24 have been found to inhibit the $CD4^+$ T cell response. Our collaborators, Catarina Hioe and Susan Zolla-Pazner at the NY Veterans Affairs Medical Center and Philip Norris at UCSF, hypothesized that one explanation for the behavior of antibodies that inhibit T cell response relative to non-inhibiting antibodies is that the inhibiting antibodies bind to or are close to the T cell epitope, thereby preventing that epitope from being efficiently presented to the T cells.

Our initial experiments used the epitope excision approach to probe the epitope recognized by the inhibiting mAb 1751. This mAb is known to inhibit recognition of the AC-25 Tcell epitope PEVIPMFSALSEGATP, aa 34–49. For these studies, antihuman IgG Fc was first immobilized on Sepharose ® 4B beads, which in turn were then used to capture the monoclonal antibody 1751. The antibody complex formed was then covalently cross-linked with BS^3 to eliminate background from mAb 1751 in the mass spectra. p24 was bound to then bound to the immobilized cross-linked mAb 1751. Both trypsin and/or Glu C were used to digest the p24 for varying times while it was bound to the antibody. After washing any unbound peptides from the beads, an aliquot of the beads was taken and placed on a MALDI target. Addition of the MALDI matrix (α-cyano-4-hydroxycinnamic acid) released the affinity-bound peptide(s) from the beads. The spectrum obtained from a 1 h trypsin digestion is shown in Figure 8A. A number of peptides were observed to be still bound to the mAb, with amino acids 30–71 being present in each fragment. Longer trypsin or Glu C digestion times resulted in the absence of any shorter peptides bound to the mAb. This is shown in Figure 8B for an overnight trypsin digestion, at which time, no peptides remain bound to the antibody. This suggests that the epitope is discontinuous in that any further cleavage within the 42 amino acid sequence destroys the recognition sequence.

Differential chemical modification was then used to provide additional structural information. Because acetylation of primary amines (lysines and N-term amino acid) was successful in previous differential surface modification studies on p24 [21], differential acetylation was again used in this study. Acetylation offers a significant advantage over other modifying reagents in differential chemical modification studies because of the availability of isotopically labeled forms. Thus, after initial derivatization under native conditions, exhaustive derivatization with a labeled isotopologue was performed.
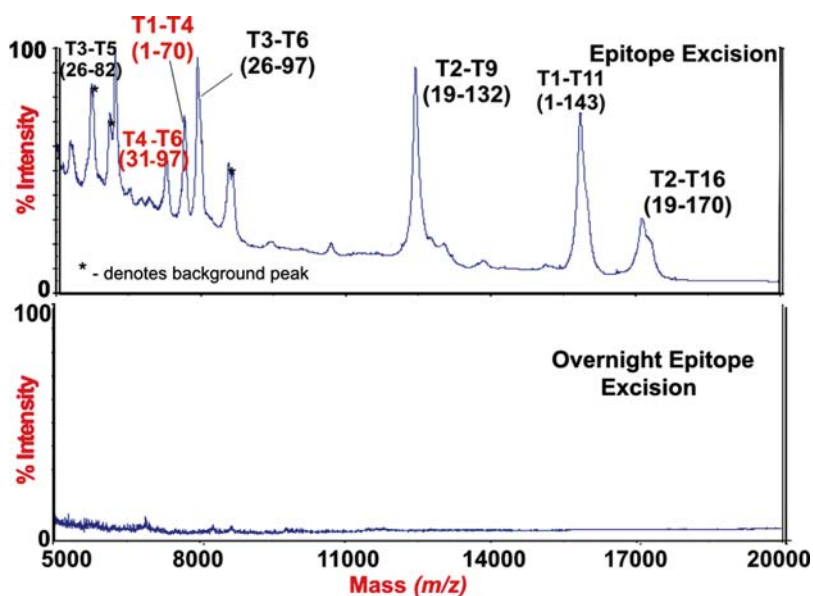
Figure 8. A (top), MALDI spectrum of 1 h trypsin digestion of mAb1571:p24. B (bottom), MALDI spectrum of overnight trypsin digestion of mAb1571:p24 (Reprinted with permission from the journal of the American Society of Mass spectrometry, published by Elsevier [22])

The ratio of unlabeled to labeled forms of peptides containing acetylated residues is a direct measurement of the chemical reactivity of a residue while the ratio of the chemical reactivities of residues in a bound form to that of the unbound form is a direct measure of changes in reactivity. Another advantage of the use of a lysine-reactive isotopologue is that it simplifies the proteolysis results using the most commonly used enzyme, trypsin. In the absence of exhaustive modification, trypsin will not cleave at modified lysine residues while still cleaving at unmodified residues. In this case the determination of the differential reactivity becomes more complex due to the requirement that the abundances of peptides originating from cleavage at the lysine and of peptides formed by not cleaving at the modified lysine need to be considered. With exhaustive modification of all lysines, only cleavage at arginine should occur. The observed changes in reactivity can be due to a specific residue being part of the recognition surface or being sterically hindered in the complex due to close spatial proximity close to the non-covalently bound component or due to differences in the conformation of the protein in the bound and free state. Relative abundance data from full scan mass spectra can be used to obtain the ratios of labeled to unlabeled acetylation at specific sites while MS/MS data is used to verify the site of acetylation. Examples of these data are shown in Figure 9A and B.
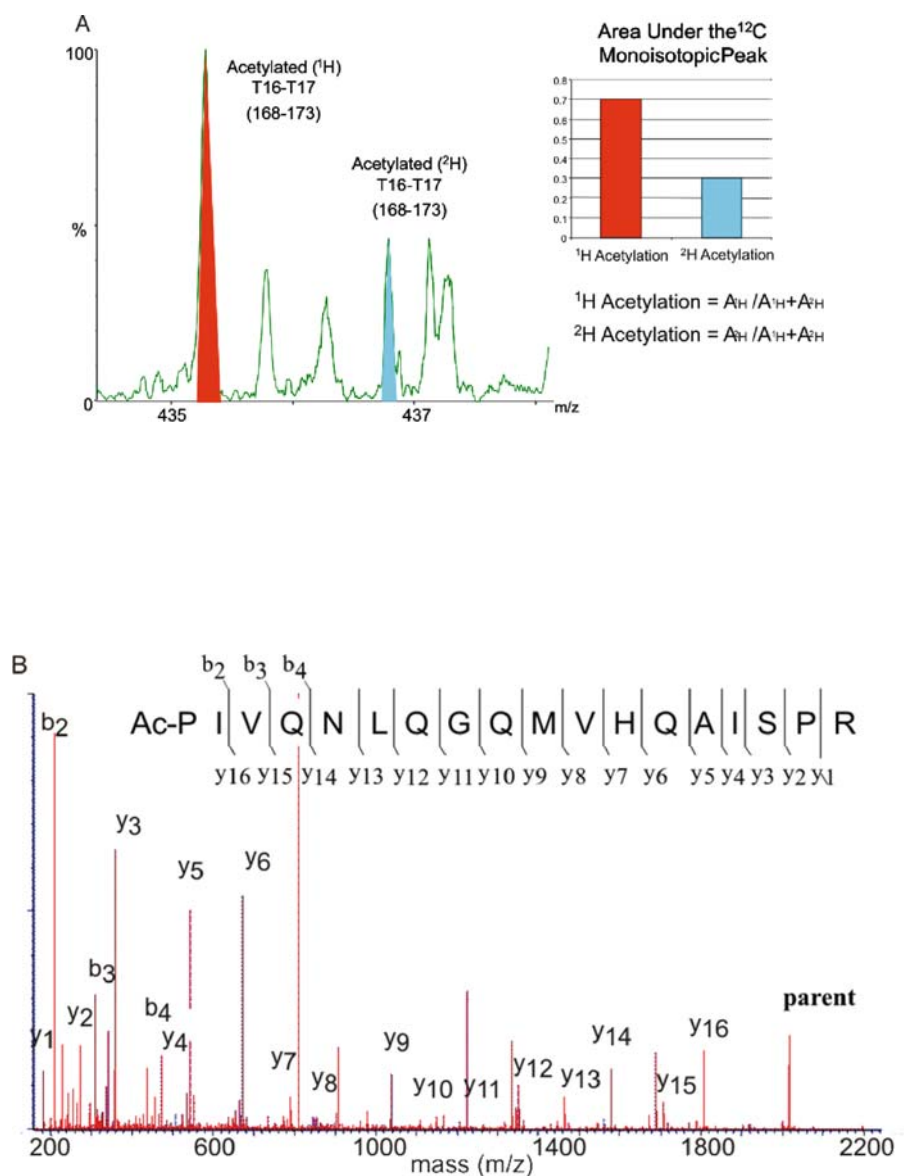
Figure 9. (A) Full scan MS (ESI) showing doubly protonated molecule region of acetylated peptide 168-173. H3-acetylated isotope peak in red and D3 acetylated isotope peak in blue. (B) MS/MS spectrum (deconvoluted) of peptide T1 (Reprinted with permission from the journal of the American Society of Mass spectrometry, published by Elsevier [22])

Based on the MS and MS/MS data, the ratio of reactivities obtained for the lysine side chains of p24 is given in Table 1.

The four lysines with significant changes in relative reactivities are K131 > K25 > K30 > K140. K131 shows a complete loss of reactivity in the complex, indicating that it is part of the recognition site. K30 shows a significant decrease in reactivity, but not a total loss, indicating that reactivity of acetic anhydride with this residue is probably sterically hindered. K140 also shows a decrease in reactivity, but in the crystal structure of p24 (PDB 1E6J) (Figure 10), the side chains are oriented nearly 180° from each other. This indicates that there is probably a conformational change upon mAb binding as well as steric interference with the antibody itself. K25 shows an increase in reactivity in the complex. From the crystal structure, the K25 side chain is pointing in the opposite direction of the K131 side chain. This would also indicate a conformational change upon mAb binding leaving the K25 side chain more accessible.

Molecular modeling of the known CD4+ T-cell epitope recognized onto the crystal structure shows that the epitope is centered in the region shown in green in Figure 10. This region bisects the two lysines, K131 and K30, with the most

TABLE 1. Changes in p24D3:H3 Acetylation Ratio from Unbound to Complexed

Amino acid	p24:mAb1571	Amino acid	p24:mAb1571
N-Terminus	0.97±0.05	Lys140	0.73±0.02
Lys25	1.37	Lys170	0.89±0.10
Lys30	0.67±0.04	Lys182	1.02±0.22
Lys70	0.97±0.09	Lys227	1
Lys131	0±0.02		
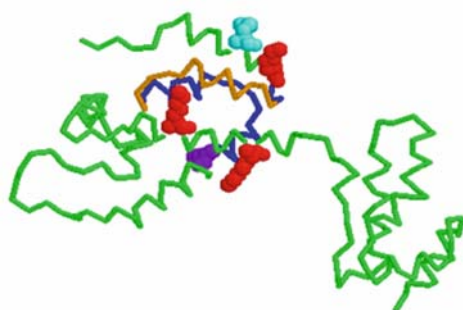
Figure 10. Ribbon structure of p24 with the known CD4+ Tcell epitope in orange; region of aa 31–70 that extends beyond the Tcell epitope in dark blue. Side chains of lysines 30, 70, 131, that are protected in the differential modification experiments are in red and the side chain of lysine25 that is more reactive is in cyan. Lysine 170 in purple is unchanged

significantly decreased reactivity. Thus, one can readily see how the antibody, when bound, is in a position to effect the decreased reactivity of the three lysines, 30, 131 and 140 and can still affect K30 and 170.

5.   Conclusions

In this report, we have shown the development of mass spectrometry for the elucidation of epitopes from simple linear epitopes to complex discontinuous epitopes using studies of anti-HIV proteins. The approaches employed here, limited proteolysis and differential chemical modification, are applicable to a wide range of protein structural studies, not only epitope elucidation. With the basic tools of mass spectrometry, the range of molecular structures and interactions that can be probed is wide open and we confidently expect a tremendous number of structural applications to human health and safety to appear in the near future.

Acknowledgments

References

1.   Hager-Braun, C., and Tomer, K. B. (2005) Determination of protein-derived epitopes by mass spectrometry. *Expert Review of Proteomics 2*, 745–756.
2.   Jemmerson, R., and Paterson, Y. (1986) Mapping Epitopes on a Protein Antigen by the Proteolysis of Antigen-Antibody Complexes. *Science 232*, 1001–1004.
3.   Suckau, D., Koehl, J., Karwath, G., Schneider, K., Casaretto, M., Bitter-Suermann, D., and Przybylski, M. (1990) Molecular epitope identification by limited proteolysis of an immobilized antigen-antibody complex and mass spectrometric peptide mapping. *Proc. Natl. Acad. Sci. U. S. A. 87*, 9848–52.
4.   Papac, D. I., Hoyes, J., and Tomer, K. B. (1994) Epitope mapping of the gastrin-releasing peptide/anti-bombesin monoclonal antibody complex by proteolysis followed by matrix-assisted laser desorption ionization mass spectrometry. *Protein Sci 3*, 1485–92.
5.   Parker, C. E., Deterding, L. J., Hager-Braun, C., Binley, J. M., Schulke, N., Katinger, H., Moore, J. P., and Tomer, K. B. (2001) Fine definition of the epitope on the gp41 glycoprotein of human immunodeficiency virus type 1 for the neutralizing monoclonal antibody 2F5. *J. Virol. 75*, 10906–10911.
6.   Jeyarajah, S., Parker, C. E., Sumner, M. T., and Tomer, K. B. (1998) Matrix-assisted laser desorption ionization/mass spectrometry mapping of human immunodeficiency virus-gp120 epitopes recognized by a limited polyclonal antibody. *J. Am. Soc. Mass Spectrom. 9*, 157–165.
7.   Hager-Braun, C., Katinger, H., and Tomer, K. B. (2006) The HIV-neutralizing monoclonal antibody 4E10 recognizes N-terminal sequences on the native antigen. *J. Immunol. 176*, 7471–7481.

8. Coeffier, E., Clement, J. M., Cussac, V., Khodaei-Boorane, N., Jehanno, M., Rojas, M., Dridi, A., Latour, M., El Habib, R., Barre-Sinoussi, F., Hofnung, M., and Leclerc, C. (2000) Antigenicity and immunogenicity of the HIV-1 gp41 epitope ELDKWA inserted into permissive sites of the MalE protein. *Vaccine 19*, 684–693.

9. Liang, X. P., Munshi, S., Shendure, J., Mark, G., Davies, M. E., Freed, D. C., Montefiori, D. C., and Shiver, J. W. (1999) Epitope insertion into variable loops of HIV-1 gp120 as a potential means to improve immunogenicity of viral envelope protein. *Vaccine 17*, 2862–2872.

10. Eckhart, L., Raffelsberger, W., Ferko, B., Klima, A., Purtscher, M., Katinger, H., and Ruker, F. (1996) Immunogenic presentation of a conserved gp41 epitope of human immunodeficiency virus type 1 on recombinant surface antigen of hepatitis B virus. *J. Gen. Virol. 77*, 2001–2008.

11. Muster, T., Guinea, R., Trkola, A., Purtscher, M., Klima, A., Steindl, F., Palese, P., and Katinger, H. (1994) Cross-Neutralizing Activity against Divergent Human-Immunodeficiency-Virus Type-1 Isolates Induced by the Gp41 Sequence Eldkwas. *J. Virol. 68*, 4031–4034.

12. Binley, J. M., Sanders, R. W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D. J., Maddon, P. J., Olson, W. C., and Moore, J. P. (2000) A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. *J. Virol. 74*, 627–643.

13. Sanchez-Martinez, S., Lorizate, M., Katinger, H., Kunert, R., and Nieva, J. L. (2006) Membrane association and epitope recognition by HIV-1 neutralizing anti-gp41 2F5 and 4E10 antibodies. *AIDS Res. Hum. Retroviruses 22*, 998–1006.

14. Ho, J., Uger, R. A., Zwick, M. B., Luscher, M. A., Barber, B. H., and MacDonald, K. S. (2005) Conformational constraints imposed on a pan-neutralizing HIV-1 antibody epitope result in increased antigenicity but not neutralizing response. *Vaccine 23*, 1559–1573.

15. Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J. R., Wyatt, R., and Kwong, P. D. (2004) Structure and mechanistic analysis of the anti-human immunodeficiency virus type 1 antibody 2F5 in complex with its gp41 epitope. *J. Virol. 78*, 10724–10737.

16. McGaughey, G. B., Barbato, G., Bianchi, E., Freidinger, R. M., Garsky, V. M., Hurni, W. M., Joyce, J. G., Liang, X. P., Miller, M. D., Pessi, A., Shiver, J. W., and Bogusky, M. J. (2004) Progress towards the development of a HIV-1 gp41-directed vaccine. *Curr. HIV Res. 2*, 193–204.

17. Stiegler, G., Kunert, R., Purtscher, M., Wolbank, S., Voglauer, R., Steindl, F., and Katinger, H. (2001) A potent cross-clade neutralizing human monoclonal antibody against a novel epitope on gp41 of human immunodeficiency virus type 1. *AIDS Res. Hum. Retroviruses 17*, 1757–1765.

18. Zwick, M. B., Labrijn, A. F., Wang, M., Spenlehauer, C., Saphire, E. O., Binley, J. M., Moore, J. P., Stiegler, G., Katinger, H., Burton, D. R., and Parren, P. (2001) Broadly neutralizing antibodies targeted to the membrane-proximal external region of human immunodeficiency virus type 1 glycoprotein gp41. *J. Virol. 75*, 10892–10905.

19. Cardoso, R. M. F., Zwick, M. B., Stanfield, R. L., Kunert, R., Binley, J. M., Katinger, H., Burton, D. R., and Wilson, I. A. (2005) Broadly neutralizing anti-HIV antibody 4E10 recognizes a helical conformation of a highly conserved fusion-associated motif in gp41. *Immunity 22*, 163–173.

20. Binley, J. A., Wrin, T., Korber, B., Zwick, M. B., Wang, M., Chappey, C., Stiegler, G., Kunert, R., Zolla-Pazner, S., Katinger, H., Petropoulos, C. J., and Burton, D. R. (2004) Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J. Virol. 78*, 13232–13252.

21. Hochleitner, E. O., Borchers, C., Parker, C., Bienstock, R. J., and Tomer, K. B. (2000) Characterization of a discontinuous epitope of the human immunodeficiency virus (HIV) core protein p24 by epitope excision and differential chemical modification followed by mass spectrometric peptide mapping analysis. *Protein Sci. 9*, 487–496.

22. Williams, J. G., Tomer, K. B., Hioe, C. E., Zolla-Pazner, S., and Norris, P. J. (2006) The antigenic determinants on HIV p24 for CD4(+) T cell inhibiting antibodies as determined by limited proteolysis, chemical modification, and mass spectrometry. *J. Am. Soc. Mass Spectrom. 17*, 1560–1569.

23. Hager-Braun, C., and Tomer, K. B. (2002) Characterization of the tertiary structure of soluble CD4 bound to glycosylated full-length HIVgp120 by chemical modification of arginine residues and mass spectrometric analysis. *Biochemistry 41*, 1759–1766.

24. Iacob, R. E., Perdivara, I., Przybylski, M., and Tomer, K. B. (2008) Mass spectrometric characterization of glycosylation of Hepatitis C virus E2 envelope glycoprotein reveals extended micro-heterogeneity of N-glycans. *J. Am. Soc. Mass Spectrom. 19*, 428–444.

10. CHEMICAL SURFACE MODIFICATION AND CHEMICAL CROSSLINKING COMBINED WITH MASS SPECTROMETRY FOR PROTEIN TERTIARY STRUCTURAL INFORMATION

LEESA J. DETERDING AND KENNETH B. TOMER
Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709, USA

Abstract. In an effort to gain tertiary structural information of proteins, chemical modification of surface exposed residues and chemical crosslinking have been used in combination with mass spectrometry. For this work, the acetylation of lysine residues was used as a chemical modification reagent. The lysine residues that are determined to be acetylated are, thus, assumed to be surface accessible. The chemical crosslinker DTSSP, 3,3′-dithiobis-[sulfosuccinimidylpropionate], was used for the crosslinking experiments. DTSSP is a homobifunctional amine reactive cross-linker that is cleavable. Therefore, an aliquot of the crosslinked protein was then subjected to conditions that would result in cleavage of the linker. After performing all modification and crosslinking experiments, the protein was digested and analyzed by both MALDI and LC/ESI mass spectrometry. For proof-of-principle, bovine β-lactoglobulin was used for these analyses, and the results are compared to the X-ray crystal structure of the protein. As an extension of this work, these biochemical and mass spectrometric techniques have been applied to a protein of unknown structure.

1. Introduction

The unique three dimensional structures of proteins are encoded into the primary amino acid sequence of the protein. The biological functions of proteins are dependent upon their tertiary and quaternary structures; therefore, it is important to study and determine the higher order structures of proteins. Commonly used techniques to investigate protein structures include x-ray crystallography [1–3] and NMR spectroscopy [3–6]. These approaches can provide very detailed structural information; however, these techniques suffer from various limitations, including insufficient protein quantity and purity, inability to crystallize certain proteins, and poor solubility of many proteins. In addition, these techniques tend to be very time consuming and tedious. Molecular modeling has also been used to obtain three-dimensional structural information of proteins [7–10]. The success of this approach, however, relies on prior structural information of a homologous protein. Due to these limitations, there are increasing demands for additional methods to probe protein structures or conformations.

In recent years, the combined use of protein surface modification and/or protein crosslinking and mass spectrometry have been used to probe protein tertiary and quaternary structures [11,12]. The surface modification approach involves the chemical modification of amino acids to determine which residues in a protein are surface accessible. On the other hand, chemical crosslinking of amino acid residues in close proximity can determine which regions within a protein are adjacent. Chemical crosslinking can be used to join two molecules or regions within a molecule by means of a chemical crosslinker. The use of crosslinkers with different spacer arm lengths in combination with mass spectrometry can be used to determine distance constraints within and between proteins. These combined techniques have been used to determine a variety of protein structures, including the determination of protein surface topology, protein-peptide interactions, protein-protein interactions, and protein-DNA interactions [12].

In this work as "proof-of-concept", bovine β-lactoglobulin was subjected to surface modification. For these studies, the acetylation of lysine residues was used [13,14]. In addition, the protein was chemically crosslinked using 3,3′-dithiobis-[sulfosuccinimidylpropionate], DTSSP (a homobifunctional amine reactive crosslinker) [15]. Because DTSSP is a cleavable linker, an aliquot of the crosslinked protein was then subjected to conditions that would result in cleavage of the linker. After performing the modification reactions, the protein solutions were digested with trypsin and analyzed by MALDI/MS/MS and LC/MS/MS. From these data, acetylated lysines and crosslinked peptides were identified, and these results were compared to the x-ray crystal structure of β-lactoglobulin. As an extension of this work, these biochemical and mass spectrometric techniques have been applied to a protein of unknown structure. The results are presented here.

2.   Methods

Materials. Purified bovine β-lactoglobulin, acetic anhydride, and dithiothreitol were purchased from Sigma-Aldrich Chemical Co. (St. Louis, MO). 3,3′-Dithiobis-[sulfosuccinimidylpropionate] was obtained from Pierce Chemical Company (Rockford, Ill). Sequencing-grade modified porcine trypsin was purchased from Promega Corporation (Madison, WI). All solutions and buffers were prepared using deionized water (18 MΩ) from an in-house Hydro Services (Durham, NC) Picopure 2 water system.

Surface Modification Reactions. Acetylation of the protein was performed at a protein concentration of 10 μM in 100 mM sodium phosphate buffer (pH 7) by adding acetic anhydride at a concentration of 1,000-fold molar excess (per primary amino group). Reactions were performed at pH 7 (pH maintained by the addition of 5 M NaOH) for 30 min at room temperature. Reactions were quenched by adding an equal volume of 1 M Tris.

Chemical Crosslinking Reactions. 3,3′-Dithiobis(sulfosuccinimidylpropionate) (DTSSP) is a water-soluble, homobifunctional, thiol-cleavable and membrane

impermeable crosslinker. It contains an amine-reactive *N*-hydroxysulfo-succinimide (sulfo-NHS) ester at each end of an 8-carbon spacer arm; thereby, resulting in a spacer arm length of 12 Å. Crosslinking of the proteins was performed at a protein concentration of 10 μM in 100 mM sodium phosphate buffer (pH 7). DTSSP was added to the protein solutions at either a 20- or 50-fold molar excess of the cross-linker. Crosslinking reactions were allowed to proceed for 30 min at room temperature. Reactions were quenched by adding 1 M Tris to a final concentration of 50 mM. For the reduction of the DTSSP crosslinker, 1 M dithiothreitol (DTT) was added to an aliquot of the crosslinked protein solution at a final concentration of 50 mM. The reduction reactions were allowed to proceed for 30 min at 37˚C.

Tryptic Digestion Conditions. The protein samples were subjected to tryptic digestion by adding sequencing-grade modified porcine trypsin at a protein:enzyme ratio of 20:1. The reactions were allowed to proceed overnight at 37˚C.

Mass Spectrometry. MALDI analyses were performed using an Applied Biosystems 4700 Proteomics Analyzer (Framingham, MA) in the positive ion reflector mode. Just prior to the MALDI analyses, the samples were desalted using C18 ziptips (Millipore Corporation, Billerica, MA). The MALDI matrix is prepared initially as a saturated solution of α-cyano-4-hydroxycinnamic acid in 50:50 acetonitrile:water containing 0.1% formic acid (v/v). This saturated solution of α-cyano-4-hydroxycinnamic acid is then diluted 1:2 (v/v) with 50:50 acetonitrile:water containing 0.1% formic acid, of which 0.3 μl is mixed with 0.3 μl of the peptide digestion solution on a 100-well MALDI sample target. Spectra were obtained over the mass range of 800–4,000 Da with 1,000 laser shots per spectrum. For each sample spot, data dependent acquisitions were acquired in a fully automated mode such that a MALDI mass spectrum is acquired followed by MS/MS of the five most abundant ions in the spectrum (excluding ions from matrix and trypsin autolysis products). Ions corresponding in mass to trypsin autolysis products were used to internally calibrate the mass spectra; thereby, allowing a routine mass accuracy of 10 ppm or less. Following the analyses, mass spectra were processed and analyzed using a Global Proteome Server Explorer™ workstation and software (Applied Biosystems, Framingham, MA).

Either a Waters Q-Tof Premier (Beverly, MA) or an Agilent Technologies XCT Ultra ion trap (Santa Clara, CA) was used for the acquisition of the nanoLC/MS/MS analyses. The Q-Tof Premier is a hybrid tandem mass spectrometer equipped with a nanoflow lockspray electrospray source and consists of a quadrupole mass filter and an orthogonal acceleration time-of-flight mass spectrometer. For the LC/MS/MS analyses, the needle voltage was 4,200 V and automated data dependent acquisition software was employed with the four most abundant ions from each spectrum selected for MS/MS analysis. The collision energy used for the MS/MS experiments was set according to the charge state and the *m/z* of the precursor as determined from a charge state recognition algorithm. A Waters NanoAcquity Ultra Performance LC system was used to deliver the gradients. Injections of 2 μl were made onto a Waters Symmetry C18 trapping column (20 mm x 180 μm id) and a linear gradient of 2–40% (0.1% formic acid) over 60 min was used for the

chromatographic separations. The column used was a 100 mm × 100 μm id Waters Atlantis™ dC18 column at a flow rate of 300 nl/min. Data analysis was accomplished with a MassLynx data system, MaxEnt deconvolution software, and ProteinLynx software supplied by the manufacturer. The chemical crosslinkers and acetylation were entered as a user-defined modifier agent on lysine amino acids.

The Agilent XCT Ultra ion trap is equipped with an HPLC-Chip Cube MS interface and an Agilent 1100 nanoLC system. Injections of 30 μl from a 1:10 dilution of the peptide digests were made onto a 40 nl enrichment column followed by a 43 mm × 75 μm analytical column, packed with ZORBAX 300SB C18 particles. Peptides were separated and eluted using a linear gradient of 3–50% acetonitrile (0.1% formic acid) over 40 min, followed by a linear gradient of 50–95% acetonitrile over 7 min at a flow rate of 500 nl/min. The ion trap mass spectrometer was operated in the positive ion mode, standard enhanced mode using the following settings: capillary voltage, –2,150 V; mass range, 300–1,500; ICC smart target (number of ions in the trap prior to scan out), 100,000 or 200 milliseconds of accumulation; and MS/MS fragmentation amplitude, 1.0 V. During the LC/MS/MS analyses, automated data dependent acquisition software was employed with the six most abundant ions (threshold requirement of 10,000 counts) from each spectrum selected for MS/MS analysis. Following the analyses, the MS/MS data were extracted and analyzed using Spectrum Mill MS Proteomics software (Agilent Technologies, Inc). To generate peak lists, the raw data files were processed using the Data Extractor function with the following parameters: deconvoluted ions of 300–6,000 Da and a retention time of 10 to 60 min. MS scans with the same precursor *m/z* were merged based on a +/– 1.4 *m/z* window and a +/– 15 s retention time window. Using the extracted data, searches were performed against an in-house database using the MS/MS search function where the chemical crosslinkers and acetylation were entered as a user-defined modifier agent on lysine amino acids. All sites of modification or crosslinking assignments were manually validated.

3.   Results and Discussion

In the present study, the utility of chemical surface modification and chemical crosslinking in combination with mass spectrometric analyses has been investigated for the characterization of protein tertiary structures. For this work and as proof-of-principle, bovine β-lactoglobulin was initially subjected to chemical modification to determine which residues were surface exposed. This procedure involved the acetylation of lysine residues, followed by tryptic digestion and MS analyses. Both the native protein tryptic digests and the acetylated protein tryptic digests were analyzed by MALDI/MS and LC/MS. During these analyses, data dependent MS/MS acquisitions were acquired to determine sites of modification. From the DDA data, many peptides containing an acetylated lysine residue were identified. As an example, the MALDI MS/MS data of the (M+H)$^+$ ion of *m/z*

2,892.53 corresponding in mass to tryptic peptide T15–18 (amino acids 141–164) plus three acetyl groups of bovine β-lactoglobulin is shown in Figure 1. In this spectrum, a complete series of y ions are observed which allows for the identification of the sites of modification in this peptide to be determined. Within this peptide, Lys151, Lys154, and Lys157 are all observed as containing the acetyl modification; implying that these residues are surface accessible. Additional lysine residues in β-lactoglobulin observed as being acetylated include K2, K24, and K63.
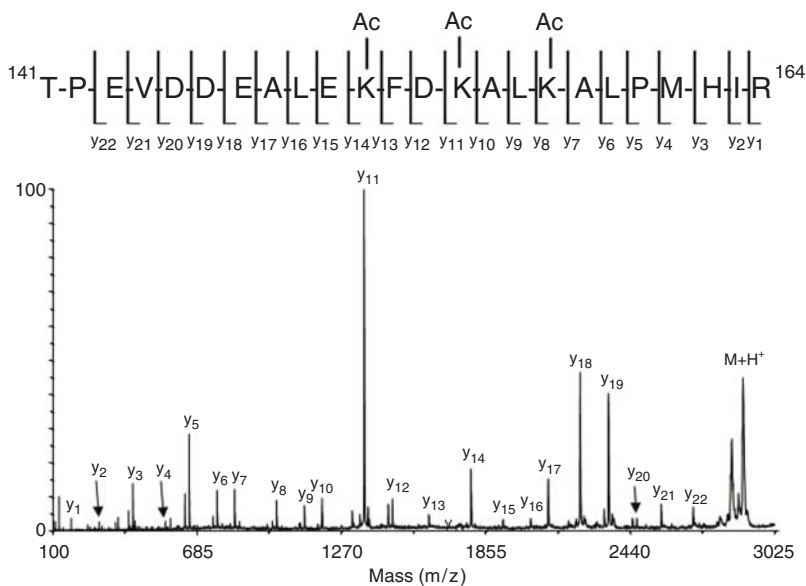
Figure 1. MALDI/MS/MS of the $(M+H)^+$ ion of *m/z* 2892.53 corresponding in mass to tryptic peptide T15-18 (amino acids 141–164) plus three acetyl groups of bovine β-lactoglobulin

For the chemical crosslinking reactions, the protein was first crosslinked with the 12 Å DTSSP crosslinker. Following completion of the reaction, an aliquot of the crosslinked protein was subjected to conditions that would result in cleavage of the linker. Then, the native protein, crosslinked protein, and the linker-cleaved protein solutions were all digested with trypsin and analyzed by MALDI/MS/MS and LC/MS/MS. These data were compared to identify crosslinked peptides. For the DTSSP crosslinker, several reaction products have been previously identified [16]. If a given peptide contains two reactive residues within the amino acid sequence, an intramolecular crosslink can be formed. The observed mass increase for a DTSSP intrapeptide crosslink is 174 Da. Figure 2 shows the LC/MS/MS of the $(M+2H)^{2+}$ ion of *m/z* 816.4 corresponding in mass to tryptic peptide T8-10

(amino acids 87–99) plus an intrapeptide DTSSP crosslink (i.e., plus 174 Da) of β-lactoglobulin. Upon fragmentation, both b and y ions are observed which allows for the confirmation of an intrapeptide crosslink between lysine 91 and lysine 93 in this peptide. Additional intrapeptide crosslinks in β-lactoglobulin were confirmed between K93-K99 and K151-K154.

In addition to the intrapeptide crosslinks, peptides containing a free linker may be observed in the analyses of the tryptic digests. Because DTSSP hydrolyzes very rapidly in water, the reaction product results in a mass increase of the peptide of 192 Da per hydrolyzed DTSSP. The LC/MS/MS data of the $(M+3H)^{3+}$ ion of m/z 981.74 is shown in Figure 3A. This ion corresponds in mass to tryptic peptide T11-12 (amino acids 100–116) plus 192 Da. Both b and y fragmentations along the peptide backbone are observed which allows for the determination of the site of free linker at lysine 107. Because this amino acid contains a free linker, one can assume that this residue is surface exposed.

Analyses of the crosslinked tryptic digest subjected to a reducing agent revealed many peptides which contain a reduced DTSSP linker. Some of these peptides result from reduction of the free hydrolyzed linker; thereby, resulting in an observed mass increase of the peptide of 88 Da per DTSSP linker. For example, tryptic peptide T11-12, which was observed as having a hydrolyzed free linker at lysine 107 (Figure 3A) in the crosslinked digest, is observed as having a reduced crosslinker (Figure 3B) following DTT treatment. This peptide is observed as an $(M+3H)^{3+}$ ion of m/z 726.99 which corresponds in mass to tryptic peptide T11-13 (amino acids 100–177) plus 88 Da. The LC/MS/MS data of this peptide ion shows a series of y fragmentations which allows for the assignment of the reduced DTTSP linker to lysine 107.
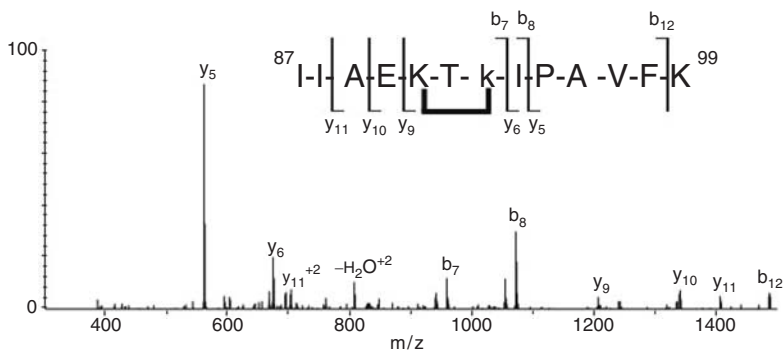
Figure 2. LC/MS/MS of the $(M+2H)^{2+}$ ion of m/z 816.4 corresponding in mass to tryptic peptide T8-10 (amino acids 87–99) plus an intrapeptide DTSSP crosslink (i.e., plus 174 Da) of β-lactoglobulin
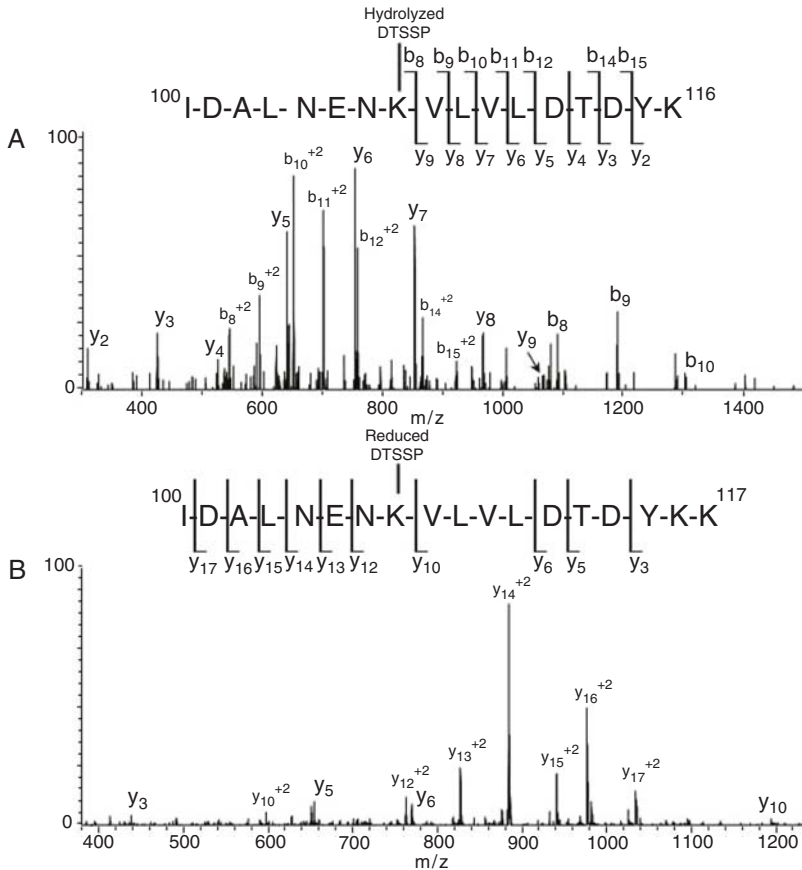
Figure 3. (A) LC/MS/MS data of the (M+3H)$^{3+}$ ion of *m/z* 981.74 which corresponds in mass to tryptic peptide T11-12 (amino acids 100–116) plus 192 Da (hydrolyzed DTSSP) of β-lacto-globulin. (B) LC/MS/MS of the (M+3H)$^{3+}$ ion of *m/z* 726.99 which corresponds in mass to tryptic peptide T11-13 (amino acids 100–177) plus 88 Da (reduced DTSSP) of β-lactoglobulin

If more than one amino acid within a peptide contains a hydrolyzed linker or the peptide contains an intrapeptide crosslink, then the peptide can be observed with more than one reduced DTSSP linker attached. The theoretical mass of the peptide will be observed plus 88 Da per linker. As an example, an (M+3H)$^{3+}$ ion of *m/z* 981.7 was observed in the reduced tryptic digest of the 50x DTSSP cross-linked β-lactoglobulin. This mass corresponds to tryptic peptide T15-18 plus 164 Da (two reduced DTSSP linkers). The LC/MS/MS data of this ion shows a nearly

complete series of y fragment ions as well as the b_{13} and b_{23} fragment ions (Figure 4). These data allow the assignment of the two reduced DTSSP linkers within this peptide to lysine 154 and lysine 157.

If DTSSP reacts with amino acids from two different peptides, an intermolecular crosslink can form. The mass of the resulting product would be the mass of each peptide plus 174 Da. Due to limitations of existing software programs and the complexity of these data, these interpeptide crosslinks must be manually assigned. From the analyses of the digest of crosslinked β-lactoglobulin, no ions were found which correspond in mass to an interpeptide crosslink.
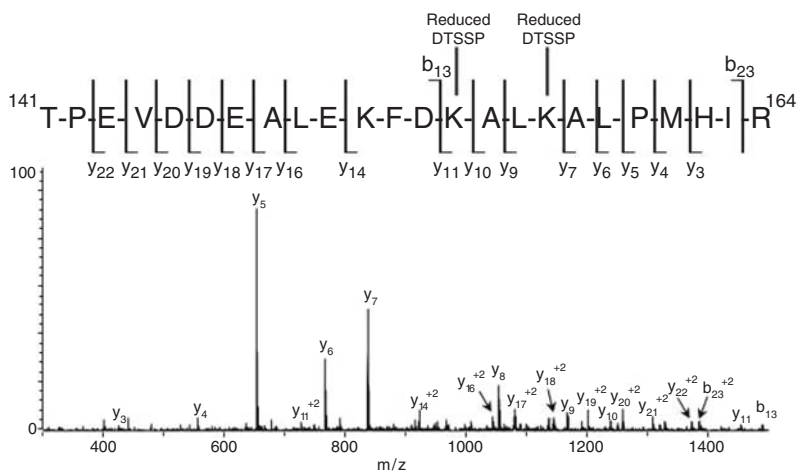
Figure 4. LC/MS/MS of the $(M+3H)^{3+}$ ion of *m/z* 981.7 which corresponds in mass to tryptic peptide T15–18 plus 164 Da. (i.e., two reduced DTSSP linkers) of β-lactoglobulin

The results obtained from the MS experiments were compared to the X-ray crystal structure (1qg5) of a variant of bovine β-lactoglobulin [17]. This variant corresponds to a truncated form of β-lactoglobulin (amino acids 17–178) that was used for these structural studies. Figure 5 shows the space-filled X-ray crystal structure of the β-lactoglobulin variant. The lysine residues which were observed in the MS analyses as either acetylated or determined to have a free linker (i.e., surface accessible) are shown in yellow.

These lysine residues include: K2, K24, K63, K86, K93, K107, K151, K154, and K157. Because the crystal structure is of a variant which does not contain residues 1–17, the lysine 2 residue is not observed in this figure. All lysine residues observed as modified in the MS analyses are surface accessible in this structure. These data demonstrate the ability to determine the surface accessibility of selected residues based on chemical modification.

In addition to surface accessibility, distance constraints between residues can be obtained from the chemical crosslinking data. From the MS analyses, three intrapeptide crosslinks were confirmed. These crosslinks were between lysine 91 and lysine 93, lysine 93 and lysine 99, and lysine 151 and lysine 154. The length of the DTSSP linker is 12 Å; therefore, the distance between the crosslinked residues in the tertiary structure of the protein should be 12 Å or less.
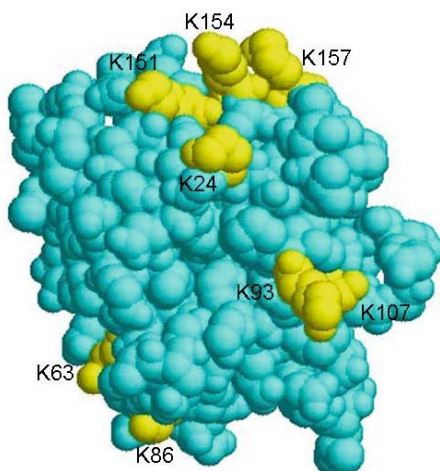
Figure 5. X-ray crystal structure of β-lactoglobulin (1qg5) [17] with the lysine residues observed as either acetylated or determined to have a free DTSSP linker shown in yellow

Figure 6 shows the distance between lysine 93 and lysine 99 in the X-ray crystal structure of β-lactoglobulin (1qg5) [17]. The distance between these residues is 10.4 Å. The distance between lysine 91 and lysine 93 is 9.51 Å while the distance between lysine 151 and lysine 154 is 7.7 Å.

These results demonstrate how the crosslinking experiments can provide distance constraints between residues in the tertiary structure of proteins.

As an extension of this work and in an effort to increase our understanding of the pathogenesis of Sjogren's syndrome, chemical modification and chemical crosslinking studies of the most common antigens associated with this autoimmune disease are under investigation.

Primary Sjogren's syndrome [18] may be defined as a chronic inflammatory autoimmune disease which typically affects the salivary and lachrymal glands; thereby, leading to dry eyes and dry mouth. By current estimations, approximately 1% of the population (mean age of onset 45–55 years and male/female ratio of 1:9) is affected by Sjogren's syndrome, thus, making it one of the most common autoimmune rheumatic diseases. Most autoimmune rheumatic diseases are characterized by the presence of autoantibodies. Autoantibodies commonly observed in
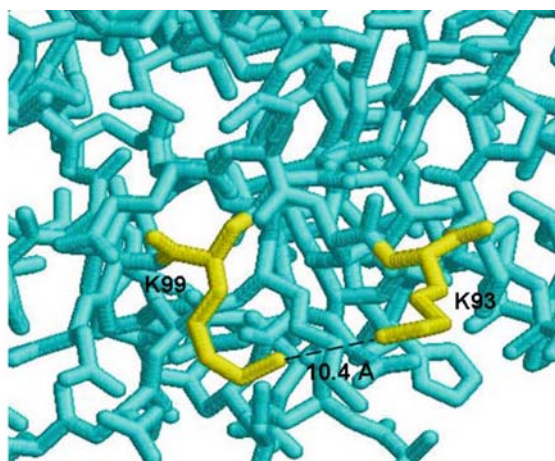
Figure 6. X-ray crystal structure of β-lactoglobulin (1qg5) [17] showing the distance between lysine 93 and lysine 99. These residues were observed to be crosslinked via an intra-peptide crosslink in the LC/MS/MS analyses

patients with Sjogren's syndrome include those against a 52 kDa Ro/SSA antigen, a 60 kDa Ro/SSA antigen, an SSB/La antigen, and an α-fondrin antigen [19]. The role of Ro/SS-A (Ro52 and Ro60) and Ro/SS-B (La) proteins in the development of autoimmune conditions is thought to be critical for understanding the mechanism of the development of autoimmunity.

To understand why LaSSB becomes the target of an autoimmune response, a better understanding of this protein's structure and function is needed. We have used chemical modification of surface exposed residues (acetylation of lysines) and crosslinking experiments (DTSSP) in combination with mass spectrometry to try to gain structural information of the full-length LaSSB protein. Representative data from these experiments are shown in Figure 7. The deconvoluted MS/MS data of the $(M+2H)^{2+}$ ion of *m/z* 686.4 corresponding in mass to tryptic peptide T33-36 (amino acids 197–206) plus three acetyl groups is shown in Figure 7A. A complete series of y ions are observed in this spectrum. These data allow the de-termination of the sites of acetylation in this peptide to lysine 197, 200, and 204. Therefore, these residues should be on the surface of this protein. Figure 7B shows the MS/MS data of the $(M+2H)^{2+}$ ion of 1,010.3 which corresponds in mass to tryptic peptide T58-60 (amino acids 318–332) plus an intrapeptide DTSSP crosslink. Both b and y fragment ions are observed which allows the confirmation of the intrapeptide crosslink in this peptide to be between lysine 328 and lysine 330. In the tertiary structure of LaSSB, these two residues should be no more than 12 Å apart. As these MS data are further interrogated, a molecular model for the full-length LaSSB will be developed.
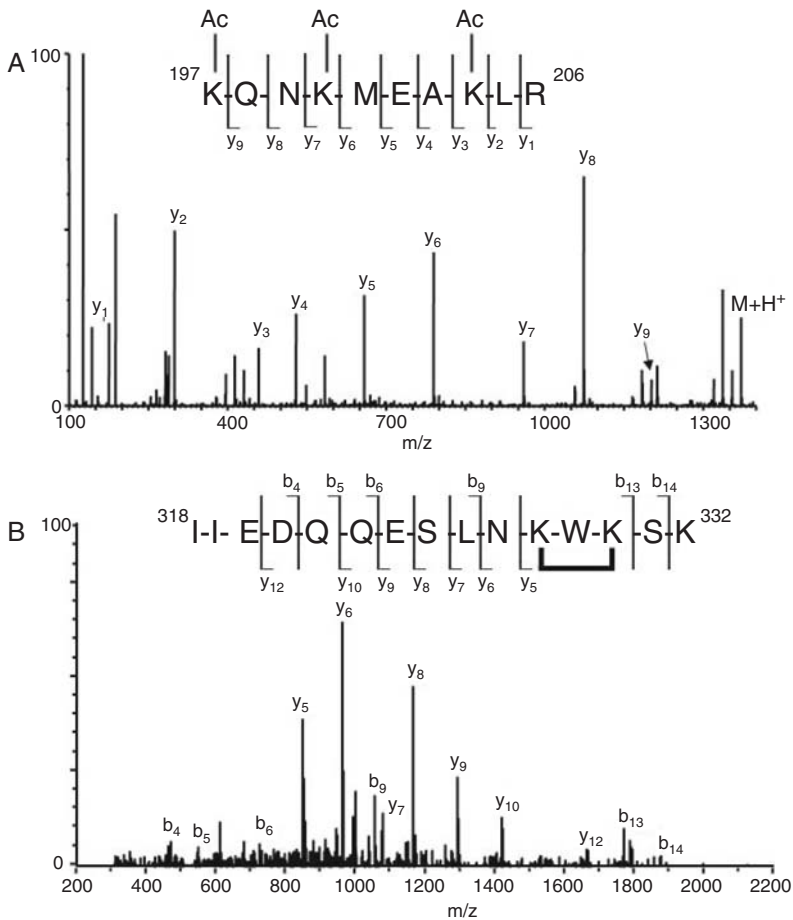
Figure 7. (A) Deconvoluted LC/MS/MS of the (M+2H)$^{2+}$ ion of *m/z* 686.4 corresponding in mass to tryptic peptide T33–36 (amino acids 197–206) plus three acetyl groups. (B) LC/MS/MS of the (M+2H)$^{2+}$ ion of *m/z* 1,010.3 which corresponds in mass to tryptic peptide T58–60 (amino acids 318–332) plus an intrapeptide DTSSP crosslink

4. Conclusions

Chemical modification and chemical crosslinking followed by mass spectrometric analyses are powerful approaches for obtaining residue-specific tertiary information about amino acids within a protein. The results obtained from the protein β-lactoglobulin compare nicely with the x-ray structure of this protein. As an extension

of this work, structural information regarding the LaSSB autoantigen associated with Sjogren's syndrome has been obtained, and these data will be used to develop a structural model of this protein.

Acknowledgments

References

1. P.C. Weber, *Adv. Protein Chem.*, **1991**, *41*, 1.
2. Schmidt and V.S. Lamzin, *Cell. Mol. Life. Sci.*, **2007**, *64*, 1959.
3. H.-L. Liu and J.-P. Hsu, *Proteomics*, **2005**, *5*, 2056.
4. K. Wüthrich, *Acct. Chem. Res.*, **1989**, *22*, 36.
5. G.T. Montelione, D. Zheng, Y.J. Huang, K.C. Gunsalus, and T. Szyperski, *Nat. Struct. Biol.*, **2000**, *Suppl.*, 982.
6. A. Yee, A. Gutmanas, and C.H. Arrowsmith, *Curr. Opin. Struct. Biol.*, **2006**, *16*, 611.
7. D. Xu and Y. Xu, *Curr. Protocols in Protein Sci.*, **2000**, *Supplement 19*, 2.7.1.
8. R. Sanchez, U. Pieper, F. Melo, N. Eswar, M.A. Marti-Renom, M.S. Madhusudhan, N. Mirkovic, and A. Sali, *Nat. Struct. Biol.*, **2000**, *Suppl.*, 986.
9. G.L. Butterfoss and B. Kuhlman, *Annu. Rev. Biophys. Biomol. Struct.*, **2006**, *35*, 49.
10. G. Mayr, F.S. Domingues, and P. Lackner, *BMC Struct. Biol.*, **2007**, *7*, 50.
11. A. Sinz, *Mass Spectrom. Rev.*, **2006**, *25*, 663.
12. K. Bennett, R. Matthiesen, and P. Roepstorff, *Methods Mol. Biol.*, **2000**, *146*, 113.
13. J.F. Riordan and B.L. Vallee, *Methods Enzymol.*, **1972**, *25*, 494.
14. D. Suckau, M. Mak, and M. Przybylski, *Proc. Natl. Acad. Sci. U. S. A.*, **1992**, *89*, 5630.
15. A.J. Lomant and G. Fairbanks, *J. Mol. Biol.*, **1976**, *104*, 243.
16. C.L. Swaim, J.B. Smith, and D.L. Smith, *J. Am. Soc., Mass Spectrom.*, **2004**, *15*, 736.
17. K.M.G. Oliveira, V.L. Valente-Mesquita, M.M. Botelho, L. Sawyer, S.T. Ferreira, and I. Polikarpov, *Eur. J. Biochem.*, **2001**, *268*, 477.
18. N. Delaleu, R. Jonsson, and M.M. Koller, *Eur. J. Oral Sci.,* **2005**, *113*, 101.
19. K.B. Elkon, A.E. Gharavi, G.R.V. Hughes, J.M. Moutsoupoulos, *Ann. Rheum. Dis.*, **1984**, *43*, 243.

Part V

SMALL MOLECULES

11. BROMINATED FLAME RETARDANTS: ANALYTICAL, TOXICOLOGICAL AND ENVIRONMENTAL ASPECTS

ADRIAN COVACI[1][*1] AND ALIN C. DIRTU[1,2]

[1]*Toxicological Centre, Department of Pharmaceutical Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium*
[2]*Department of Inorganic and Analytical Chemistry, "Al. I. Cuza" University of Iassy, Carol I Bvd. No 11, 700506 Iassy, Romania*

Abstract. Brominated flame retardants (BFRs), such as polybrominated diphenyl ethers (PBDEs), hexabromocyclododecanes (HBCDs) and tetrabromobisphenol-A (TBBP-A), have routinely been added to consumer products for several decades in a successful effort to reduce fire-related injury and property damage. Recently, concern for this emerging class of chemicals has risen because of their occurrence in the environment and in human biota. Here we briefly review scientific issues related to analytical, toxicological and environmental aspects of these BFRs and discuss data gaps.

1. General Information Concerning Flame Retardants

Flame retardants are materials that inhibit or resist the spread of fire that are added to polymers which are used in plastics, textiles, electronic circuitry or other materials (WHO/ICPS, 1994, 1997). The different classes of flame retardants include naturally occurring substances (asbestos), synthetic inorganic materials, such as antimony oxides, aluminium hydroxide, magnesium hydroxide, and borates, organic phosphate esters with or without halogens and chlorinated and brominated organic compounds (WHO/ICPS, 1997). The most used brominated flame retardants (BFRs) are polybrominated diphenyl ethers (PBDEs), hexabromocyclododecane (HBCD), tetrabromobisphenol-A (TBBP-A) and polybrominated biphenyls (PBBs).

Flame retardants act by different mechanisms depending on the respective chemical class: some compounds break down through endothermic processes when subjected to high temperatures (e.g., aluminium and magnesium hydroxides), while other compounds act as diluents of fuel or of the gas phase and thus lowering the combustible portion of the material (calc, calcium carbonate or inert gases, most often carbon dioxide or water). Another way to stop the flame spreading is to create a thermal insulation barrier between the burning and unburned parts of a material (e.g., intumescent additives).

Chlorinated and brominated flame retardants under thermal degradation conditions release hydrogen chloride and hydrogen bromide. These reaction products react with highly reactive $H\bullet$ and $HO\bullet$ radicals present in the flame resulting in the formation of inactive molecules and of $Cl\bullet$ or $Br\bullet$ radicals. The halogen radical has

---

*Corresponding author. E-mail: adrian.covaci@ua.ac.be

much lower energy than $H\bullet$ and $HO\bullet$, and therefore much lower potential to propagate the radical oxidation reaction and therefore the flame.

Despite of their benefits for reducing fire-related injury and property damage, growing concern for BFRs has risen because of their occurrence and persistence in the environment, biota and humans, in a similar way to other persistent organic pollutants (de Wit, 2002; Birnbaum and Staskal, 2004).

2.   Brominated Flame Retardants: Uses and Production Levels

2.1. POLYBROMINATED DIPHENYL ETHERS

PBDEs are flame retardant additives which are used in a wide array of household products in concentrations up to 30% by weight, typically between 2% and 6%. They are structurally related to polychlorinated biphenyls (PCBs) and are produced commercially as mixtures. However, PBDE mixtures contain fewer congeners than the commercial PCB mixtures.

The three commercially mixtures of PBDEs are Penta-BDE, Octa-BDE and Deca-BDE according to the number of bromine atoms in the dominating congeners of the mixtures. The three PBDE mixtures have different applications:

– Penta-BDE mixture is primarily used in foams, such as seat cushions and other household upholstered furniture, as well as in rigid insulation.
– Octa-BDE is used in high impact plastic products, such as housing for fax machines and computers, automobile trim, telephone handsets and kitchen appliance casings.
– Deca-BDE is used in plastics, such as wire and cable insulation, adhesives, coatings and textile coatings. Typical end products include housing for television sets, computers, audiotape cassettes stereos and other electronics. Deca-BDE is also used as a fabric treatment and coating on carpets and draperies, but it is not used in clothing.

The European Union has banned since August 2004, the use of Penta- and Octa-BDE technical mixtures, but the use of Deca-BDE mixture is unrestricted following favorable risk assessment in May 2004 (BSEF, 2007). In U.S., only California has banned the use, by the end of 2008, of Penta- and Octa-BDE mixtures and other U.S. states are in the phase out legislation for PBDEs.

2.2. HEXABROMOCYCLODODECANE

Structurally, HBCD is a cyclic aliphatic ring consisting in twelve carbon atoms with six bromine atoms tied to the ring. The commercial HBCD consists in a mixture of the α–, β– and γ–HBCD diastereomers, with the γ–HBCD isomer being dominant ($\geq$70%). With a worldwide production of 16,700 t in 2001, HBCD is the third most widely used BFR in the world and on the second place in the European

Union. HBCD is considered to be a high-production-volume chemical and a priority pollutant under the "Existing Substance Regulation" of the European Chemicals Bureau. It is mostly used in extruded (XPS) and expanded (EPS) polystyrene foams, but also is used as insulation material in construction industry. HBCD is highly efficient so that very low levels are required to reach the desired flame retardancy. Typical HBCD levels in EPS are 0.7% and in XPS 2.5%. At present, HBCD is the only suitable flame retardant for these applications. Other uses of HBCD are upholstered furniture, automobile interior textiles, car cushions and insulation blocks in trucks, packaging material, video cassette recorder housing and electric and electronic equipment.

In Europe and the United States, HBCD is not subject to regulatory restriction, whereas in Japan it is classified as a type I monitoring substance together with PBDEs and TBBP-A.

2.3. TETRABROMOBISPHENOL-A (TBBP-A)

With a global production of 170,000 t in 2004 (BSEF, 2007), TBBP-A is the largest used BFR. TBBP-A is mainly used as reactive BFR in laminates for printed wiring boards which are commonly used in electronic devices. Additionally, TBBP-A is used as an additive BFR in acrilonitrile-butadiene-styrene (ABS) polymers plastic housings, but it is also used as an intermediate in the production of other BFRs, such as TBBP-A derivatives and brominated epoxy oligomers. Following favorable risk assessments, the use of TBBP-A is not restricted in any country (BSEF, 2007).

2.4. OTHER BROMINATED FLAME RETARDANTS

There are in total a number of 75 different types of BFRs. Besides the above mentioned compounds, the most important group of BFRs according to its impact on environment are polybrominated biphenyls (PBBs). As a BFR, PBB was used in epoxy and phenolic resins, industrial plastics, such as high-impact polystyrene. Combusted PBB-containing materials may produce highly toxic brominated dioxins and furans. Since they have been found to be persistent, bioaccumulative toxins, being also classified as potential carcinogens, most of the production of PBBs was ceased in 1970s (BSEF, 2007).

3. Analytical Methodologies

Due to the observed increasing temporal trends in humans or biota, BFRs are being determined in a growing number of laboratories. Analytical methods for the determination of BFRs have shown a rapid development and they were in most of the cases based on protocols previously established for persistent organic pollutants

(POPs), such as organochlorine pesticides, PCBs or polychlorinated dioxins and furans (PCDD/Fs). Even if different properties of BFRs, such as polarity or vapor pressure, suggest that different procedures should be applied for their analysis from environmental samples, some common approaches can still be underlined depending also on the type of sample or the detection method (de Boer et al., 2001; Covaci et al., 2003, 2007; de Boer et al., 2007). Indeed, due to particular physico-chemical properties, the determination of individual HBCD diastereomers and TBBP-A may require specific analytical approaches.

In general, the methods used for the determination of BFRs in different matrices are very sensitive and thus able to detect extremely low amounts of these compounds. The methods described in the literature have been recently reviewed (Covaci et al., 2003; Stapleton, 2006; Covaci et al., 2007). Some basic steps of the BFR determination are sample pre-treatment, extraction, clean-up and instrumental analysis. However, being present in all environmental compartments, laboratory contamination during each analysis step can easily occur.

3.1. SAMPLE PRE-TREATMENT

Various sample pre-treatments are used depending on the employed extraction method. For solid samples (sediment, soil, dust, biological tissues), sample pre-treatment involves usually the drying of the sample. Dry samples are more effectively homogenized, allowing accurate sub-sampling for parallel analyses for other determinants (e.g., organic carbon). In addition, storage and transport may be easier. The absence of water in the samples avoids laborious extraction with separation funnels and makes the sample matrix more accessible to organic solvents. As an alternative to drying through evaporation, several methods can be applied for water binding and the most easily to perform is chemical drying by grinding of the sample with anhydrous Na_2SO_4. Freeze-drying (water evaporation below 0°C under vacuum conditions) can also serve for sample drying (Smedes and de Boer, 1997).

3.2. EXTRACTION

The extraction procedure of the analytes is dependent on the sample matrix and thus different methods are used for solid or liquid samples.

In the case of solid samples, including soil, sediment, sewage sludge, adsorbent materials used for air sampling and biological samples (tissues, eggs), the extraction efficiency depends on few major factors: the analyte solubility in the extraction solvent, the accessibility of the extraction solvent to the matrix and the extraction time. The use of non-polar organic solvents is advantageous because of the high solubility of most BFRs, but there is a low accessibility of such solvents to the inner part of biotic tissues or to the sediment/sewage sludge organic matter (containing many polar groups like amines, phenols and carboxylic acids). Therefore, the use

of binary solvent mixtures (combination of an non-polar and polar solvent) proved to be the ideal solution in terms of obtaining the best recoveries for extraction of BFRs from the solid samples (de Boer et al., 2001).

An attractive extraction technique, due to its general robustness and low cost, is the Soxhlet liquid-solid extraction. The Soxhlet extraction may be also performed in a classical way or in an automated way, so-called hot Soxhlet. In the later case, the extraction solvent is distilled into the extraction chamber, which is heated below the boiling point of the solvent keeping in this way the sample in permanent contact with hot, but no boiling solvent, thus accelerating the desorption and finally the extraction process.

Other extraction methods are also applied including ultra-sonication, used for the extraction of PBDEs from polyurethane foams used for air sampling or from high-impact polystyrene, but lower extraction recoveries compared with Soxhlet extraction were obtained (Covaci et al., 2003). Extraction of BFRs with organic solvents from a chromatographic column filled with homogenized sample has been used, but despite of its simplicity, the method uses large volumes of organic solvents that have to be further evaporated and disposed off (Manchester-Neesvig et al., 2001). Other new extraction techniques, such as accelerated solvent extraction (ASE) or microwave assisted extraction (MAE), are currently applied on BFRs analysis (Covaci et al., 2003) and although higher costs which are involved compared with Soxhlet extraction, these techniques have the advantage of lower solvent consumption, which makes the long-term costs lower and the procedures more environmentally friendly.

In the case of liquid samples, liquid-liquid extraction (LLE) was applied by using a binary mixture of solvents for BFRs analysis in river and seawater samples, but also from human milk or serum. Solid phase extraction (SPE) has been used mostly for the analysis of neutral type compounds from human serum or milk, but also for the analysis of phenolic BFRs. In this later case, an additional clean-up step is necessary due to higher amounts of co-extracted lipids.

3.3. CLEAN-UP

The non-selective nature of the exhaustive extraction procedures and the complexity of the sample matrices result in complex extracts that require further purification prior to the gas or liquid chromatographic separation of BFRs. The clean-up techniques for BFR analysis were recently reviewed by Covaci et al. (2003, 2007). Depending on the type of sample analyzed, the extracts may contain sulfur that has to be removed (the case of sediment, sewage sludge or soil samples) or high concentrations of lipids (usually the case of biological samples). Similar to other organic contaminants, the BFR concentrations are usually lipid-normalized and therefore, the lipid content has to be determined. This can be done gravimetrically before clean-up, determined on a separate sample aliquot a total lipid determination or on separate aliquots by enzymatic tests (for serum and plasma samples).

For abiotic samples, elemental sulfur can be removed by simple treatments of either the sample or the sample extract with copper powder or by gel permeation chromatography (GPC). There are also other procedures, including treatment with AgNO$_3$-modified silica or with tetrabutylammonium sulfide, but these techniques are less frequently used.

The lipids contained in biological samples extracts may be eliminated using either destructive or non-destructive methods. The common non-destructive methods of lipids removal are GPC and adsorption chromatography on selected sorbents. However, for complex matrices, two serially connected GPC columns or GPC followed by further clean-up by adsorption chromatography are often required for complete fat removal and/or isolation of BFRs from other organo-halogenated compounds. For adsorption chromatography, silica gel, alumina and Florisil with different degrees of activation have been widely used for lipid removal.

Similarly to other organohalogenated compounds, the most used destructive treatment in the BFR analysis is the sulfuric acid treatment since they are stable under strong acid conditions. The simplest approach consists of direct addition of the acid to the sample extract dissolved in *n*-hexane. However, this treatment requires several sequential LLE and centrifugation steps, which results in a multi-step and time-consuming procedure. The dispersion of sulfuric acid onto the surface of activated silica gel results in an adsorbent which can be easily loaded into a column. The use of acidified silica avoids the emulsion problems of the LLE approach, reduces the sample handling and solvent consumption and increases the sample throughput. Silica gel can also be modified with alcoholic NaOH or KOH, but such treatment may cause losses of Br atoms from highly brominated BFRs, such as HBCDs, PBBs and PBDEs (de Boer et al., 2001).

3.4. INSTRUMENTAL ANALYSIS

Because of their physico-chemical properties (like vapor pressure or polarity), gas-chromatography has become a standard analytical separation method employed for the analysis of PBDEs, while a different approach may require the analysis of HBCDs and to some extent TBBP-A, for which LC-based separation techniques are often used.

3.4.1. *Gas-Chromatographic Analysis of BFRs*

An important step for the accurate and optimal determination of compounds with relatively high boiling points, such as PBDEs, is the sample injection into the gas chromatographic (GC) system. Various injection methods were reported, but the most common are split/splitless injection, on-column injection and programmable temperature vaporization injection. All mentioned methods possess advantages and disadvantages that derive primarily from their availability, price, acceptable detection methods and discrimination of congeners on the basis of molecular weight.

As a consequence of their selective uptake, metabolism or degradation, the profiles of BFRs in samples are commonly estimated by determining of selected individual congeners. This means that single-capillary column GC may offer sufficient resolution for a congener specific BFR determination. Therefore, in order to achieve enough separation between BDE congeners particularly and possible interferences, there is a need for using sufficiently long columns (30–50 m) and small diameters (≤0.25 mm). Good resolution may also be obtained by using narrow bore columns (internal diameter = 0.1 mm) (Covaci et al., 2002). Prior of using a certain capillary column for a specific type compound analysis, it has to be tested first for possible co-elutions of target compounds, internal standards and other compounds present in the sample, especially if electron capture detector is used. The impact of the co-elutions depends on the sample clean-up, pollutant load in the sample, chromatographic system and detection method. Several co-elutions are reported in the literature for BFR analysis (Figure 1), but most of them can be solved by MS detection either in selective ion monitoring of full scans (reviewed by Covaci et al., 2003).

The characteristics of the GC system significantly influence the PBDE determination by GC-MS. If not selected properly, the column brand, type of retention gap, press-fit connector and stationary phase as well as column length and injection technique may have a very strong influence on the accuracy and precision of the PBDE analysis (Björklund et al., 2004). By selecting an erroneous GC set-up,

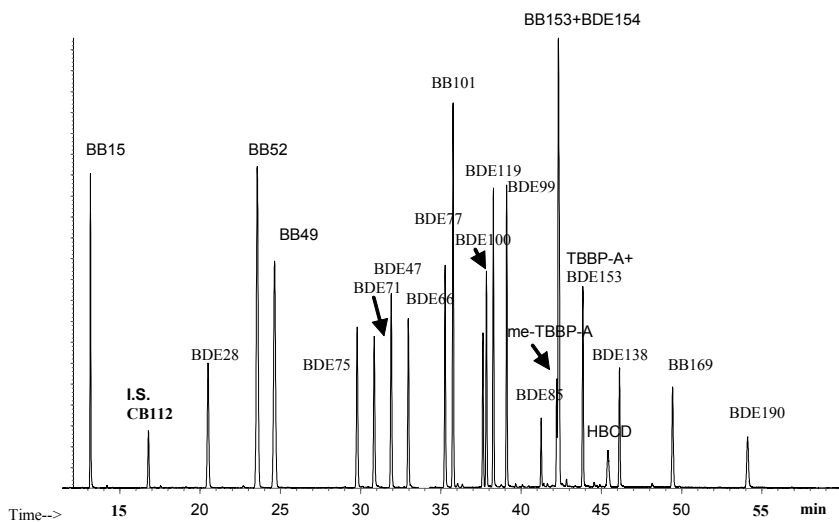*Figure 1*. ECNI-MS chromatogram of a standard mixture of BFRs on a 50 m × 0.2 mm × 0.25 μm CP-Sil 8 column (1 μl injected in pulsed splitless, injector temperature: 275°C, oven program: 90°C (3 min), with 30°C/min to 210°C (20 min), with 5°C/min to 315°C) (From Covaci et al., 2003)

the yield of nona- and deca-BDE congeners can be reduced to zero and the precision of the determination of congeners with more than five bromine atoms can be strongly decreased. An overview on GC-MS parameters used in the analysis of PBDEs, TBBP-A and other BFRs is presented in Table 1.

The use of comprehensive two-dimensional gas chromatography (GC×GC) has been evaluated also for the analysis of PBDEs. It was previously shown that even on the most efficient stationary phases, single dimension GC cannot separate all or nearly all PBDE congeners. From the 125 BDE congeners that were tested, 55 congeners were involved in 22 co-elutions for the most efficient phase, DB-XLB (Korytár et al., 2005a). As a consequence, Korytár et al. (2005b) evaluated six column combinations for the GC × GC separation of PBDEs with µ-ECD or time of flight-mass spectrometry (TOF-MS) as detectors. They concluded that a DB-1 × 007-65HT (Quadrex) combination was the most suitable because of: various reasons including the highest number of PBDE congeners separated, the least decomposition of higher brominated congeners, and the most suitable maximum operating temperature. With this set-up, there were only 17 co-eluting pairs involving 35 congeners. Since other BFRs or MeO-BDEs may be present in environmental samples, it was shown that the second dimension column improves the separation of Me-TBBP-A, TBBP-A, BB 169, 6-OH-BDE 47 and 6-MeO-BDE 47, while most of other compounds were already separated in the first dimension (Figure 2).

A different approach has to be used in the case of BDE 209 analysis because of its sensitivity for higher temperatures and the higher susceptibility for degradation in the GC system. The GC column should be relatively short (10–15 m) to reduce as much as possible the residence time (de Boer et al., 2001). Based on the ability of the low pressure-GC (LP-GC) technique to elute compounds at lower temperatures than conventional GC techniques, BDE 209 was analyzed at temperatures below the degradability limit. The low elution temperature of BDE 209, combined with its short residence time in GC has lead to minimal thermal degradation (Dirtu et al., 2007).

The analysis of more polar BFRs, especially phenolic-type compounds, such as TBBP-A, would require derivatization prior to injection into the GC system. This procedure involves a more complex sample preparation procedure compared to LC methods, but the methods limits of quantification obtained by GC analysis are more appropriate for measuring contaminants like TBBP-A presents at lower concentrations in biota compared to PBDEs and HBCDs due to its lower bioaccumulation potential.

TABLE 1. Overview of GC-MS Parameters Used for the Analysis of PBDEs, TBBP-A and Other BFRs

Compound	Column	Dimensions	Injection mode	Source temp (°C)	Derivati- sation	Ionisa- tion	Instrument	Ion	References
PBDEs	VF5-MS (Factor Four, Varian)	55 m × 0.25 mm × 0.25 μm	Splitless	220	–	EI	IT	M$^+$ and [M–2Br]$^+$	(Gómara et al., 2006)
PBDEs	DB-5 ms (J&W Scientific)	15–30 m × 0.25 mm × 0.25 μm	Splitless	250	–	EI	QTrap	M$^+$ and [M–2Br]$^+$	(Wang et al., 2005)
PBDEs	CP-Sil8 CB low bleed (Chrom-Pack)	30 m × 0.25 mm × 0.25 μm	Splitless	250	–	EI	IT	BDE 47: 486-324,326,328; BDE 100:566-404,406; BDE 154:644-482,484,486	(Salgado-Petinal et al., 2006)
PBDEs	DB-1 (J&W Scientific) × 007-65HT (Quadrex)	30 m × 0.25 mm × 0.25 μm × 1.0 m × 0.1 mm × 0.10 μm	Splitless	–	–	–	μ-ECD	–	(Korytár et al., 2005b)
PBDEs	DB-1 (J&W Scientific) × HT-8 (SGE)	15 m × 0.25 mm × 0.25 μm × 1.2 m × 0.1 mm × 0.1 μm	Splitless	250	–	EI	TOF	Scan	(Focant et al., 2004)

(Continued)

(Continues)

Compound	Column	Injection	Temperature	Derivatization	Ionization	Analyzer	Ions	Reference
DeBDethane	DB-5 ms (J&W Scientific) 12 m × 0.25 mm × 0.12 μm	Splitless	220	–	ECNI	Q	79, 81	(Kierkegaard et al., 2004)
DeBDethane	DB-5 ms (J&W Scientific) 14 m × 0.25 mm × 0.10 μm	on-column	260	–	EI	HRMS	$[C_6Br_5CH_2]^+$: 482,484,486,488 $[C_{12}Br_{10}C_2H_4]^+$: 969,971,973	(Kierkegaard et al., 2004)
DeBDethane	DB-5 ms (J&W Scientific) 15 m × 0.25 mm × 0.25 μm	Splitless	–	–	–	ECD	n.a.	(Kierkegaard et al., 2004)
TBBP-A	UB5-P (Interchim) 15 m × 0.25 mm × 0.25 μm	Splitless	230	MSTFA	EI	HRMS	682.8509, 684.8489	(Cariou et al., 2005)
TBBP-A	DB-5 ms (J&W Scientific) 30 m × 0.25 mm × 0.1 μm	Splitless	275	methyl chloroformate	EI	HRMS	556.7608, 554.7629	(Berger et al., 2004)

IT – Ion trap, Q – Quadrupole, TOF – Time-of-flight

Figure 2. Overlaid GC×GC–μECD chromatograms on DB-1×007-65HT column combination of (●) BDEs, (●) fluorinated BDEs, (●) other brominated flame retardants and (●) BDE metabolites. Temperature programme: 90°C (2 min), at 20°C/min to 190°C, then at 2°C/min to 325°C (5 min). Modulation period, 8 s; constant flow of helium carrier gas, 1.2 ml/min [From Korytár et al., 2005b]

3.4.2. *Detectors*

The most widely used detectors for the BFR determination are mass spectrometers, classified into low-resolution (LR) or high-resolution (HR) mass spectrometric (MS) instruments. The LR-MS instruments are operated either in electron impact (EI) or electron capture negative ionization (ECNI) mode. For specific applications, electron capture detectors (ECD) have been also used, but for the use of such detector, the possibility of interferences to occur is very high.

In case of PBDE analysis using EI-MS, the major ions formed are M^+ and the $[M\text{-}2Br]^+$ which can be used for their identification and quantification (Sellström, 1999) allowing also their determination in the presence of possible co-eluted compounds (such as PCBs). Because of these fragments formed, the use of LR-EI-MS provides a higher selectivity compared to LR-ECNI-MS. However, the use of LR-EI-MS (with quadrupoles as mass analyzer) is not routinely used for the PBDE analysis because of its relatively low sensitivity, especially when measuring BDE congeners with more than six bromine atoms.

In contrast to EI, ECNI is a "soft" ionization technique that takes advantage of the interactions between thermal energy electrons and electrophilic molecules, such as PBDEs. In ECNI, the low-energy electrons (thermal electrons) generated by interactions between a high-energy electron beam and a moderating or reagent gas, react with the analytes to form negative ions. The electron energy should be very low to facilitate electron capture, and the specific energy required for electron capture depends on the molecular structure of the analyte. Therefore, ECNI-MS is usually preferred for the analysis of PBDEs, because it is more selective towards aromatic brominated compounds.

Compared to LR-EI-MS, the use of LR-ECNI-MS for PBDE analysis is less selective because of monitoring of the bromide ions $[Br]^-$ for all homologue group, but instead is a much sensitive method with one order of magnitude lower limits of detection. Therefore, this technique proves to be suitable for the analysis of low-concentration samples such as human serum and plasma. However, selectivity can be retained when using LR-ECNI-MS under optimized conditions. Optimizing of the electron energy, emission current, source temperature and system pressure it was noticed that relative abundance of the molecular fragment $[M\text{-}x\text{H-}y\text{Br}]^-$ is increasing and therefore it can be used for the monitoring of each homologue group in place of the nonspecific bromide ions (Ackerman et al., 2005).

The use of GC-MS (operated in EI or in ECNI mode) was also applied for the analysis of other BFRs, such as HBCD, in fish samples (Roosens et al., 2008). The results were compared to those obtained by LC-MS. A good agreement has been found between the different analysis techniques. The GC-MS methods do not allow having individual isomer data, but they give a very good estimation of the total HBCD concentrations. In parallel with the "traditionally" monitored peak at $m/z = 79$, other ions, such as the $[M\text{-}Br]^-$ ions at $m/z = 561$, could be also monitored, but this is accompanied by a decrease in sensitivity as it can be seen in the full scan ECNI-MS spectra of ^{12}C-α-HBCD (Figure 3). This procedure
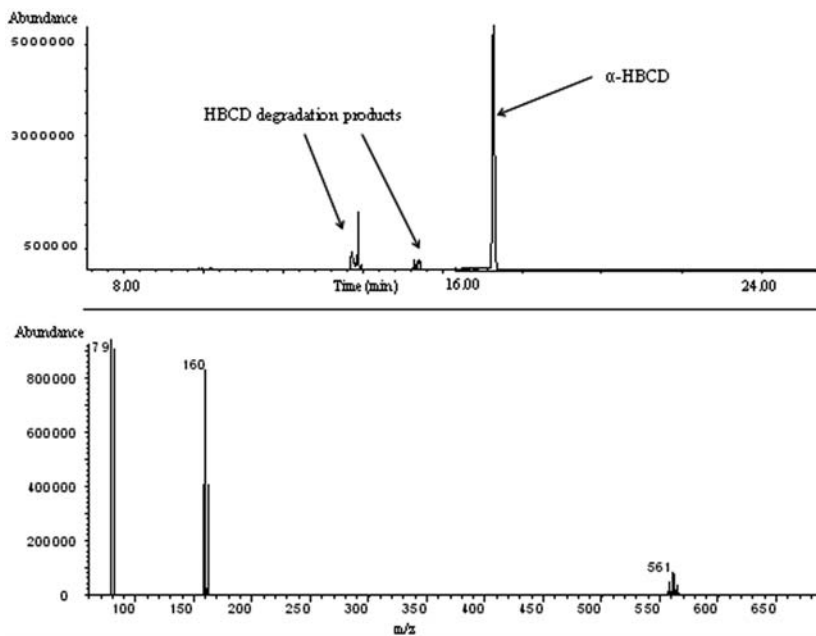
Figure 3. GC–ECNI/MS total ion chromatogram and full scan mass spectra of α–HBCD standard (conc. 1 ng/µl in iso-octane)

allows the use of ^{13}C-α-HBCD as internal standard and thus improving greatly the quality of GC-MS measurements. However, due to the 50–100 loss in sensitivity when other ions than $m/z = 79$ are used, this method can be employed only to samples with high concentrations of HBCDs (Roosens et al., 2008).

The use of ECD for BFR analysis was applied only when concentrations were relatively high (Allchin et al., 1999; Manchester-Neesvig et al., 2001). Its relatively good sensitivity for compounds with four or more bromine atoms, combined with its relatively low purchase and maintenance cost, could make it very attractive to be used in this field, but despite of such advantages, several drawbacks have reduced its application area. One problem is unequal responses for the different congeners caused by the influence on the detector sensitivity of the substitution pattern of the rings (Sellström, 1999). Furthermore, ECD is known for being linear only over a limited concentration range and also it is known for its lack of selectivity.

3.4.3. *Liquid-Chromatography Analysis of BFRs*

Despite the limited chromatographic resolving power of LC, methods employing LC/MS and LC/MS-MS offer very good results for the analysis of some BFRs, such as HBCDs and TBBP-A. These methods have been applied due to a number of inconveniences recorded for GC analysis of compounds such HBCD

or TBBP-A. At temperatures higher than 160°C, the interconversion of HBCD diastereomers make impossible the GC determination of HBCD as single isomers and moreover, HBCD degrades at temperatures above 240°C and partially break-down in dirty GC systems. For TBBP-A, because of its higher polarity compared to other BFRs generated by the presence of HO groups, LC-MS methods were also intensively applied. The use of GC-based methods for this compound would require the derivatization of the hydroxyl groups, which makes sample preparation a much complex process compared with sample preparation for LC analysis. However, the method limits of quantification are much lower in the case of GC analysis, being possible using such methodology the monitoring of TBBP-A concentrations in environmental compartments in which the exposure is at a normal level. An overview of LC-MS parameters used for in the analysis of HBCDs, TBBP-A and other BFRs is presented in Table 2.

The analysis of PBDEs by atmospheric pressure photoionization (APPI) and LC/MS-MS was also tested (Debrauwer et al., 2005). It is known that PBDEs do not ionize well using the most traditional LC/MS-MS methods, electrospray ionization (ESI) or atmospheric chemical ionization. Using APPI, PBDEs ionize well in both negative (higher sensitivity for penta- through deca-BDE congeners) and positive (higher sensitivity for di- through penta-BDE congeners) modes, depending on the degree of bromine substitution. APPI is a softer ionization technique compared to electron impact (EI) since M^+ ions are the most intense ions produced by the interaction of PBDEs with dopants charged by photons and MRM in the MS-MS system follows the M^+ to $[M\text{-}Br_2]^+$ transition. The most intense ions in EI are the $[M\text{-}Br_2]^+$ ions.

For HBCD isomer-specific analysis, reversed-phase LC coupled to ESI or atmospheric pressure chemical ionization (APCI-MS) is a versatile tool for its determination in environmental samples. However, the use of LC-ESI-MS/MS results in better performances than LC–APCI-MS/MS when a single MRM for the transition $[M\text{–}H]^-$ (*m/z* 640.6) $\rightarrow$ $[Br]^-$ (*m/z* 79) is used (Budakowski and Tomy, 2003).

Column selectivity towards HBCD diastereomers was evaluated for C_{30} and C_{18} stationary phases under different mobile phase conditions and column temperatures. The HBCD elution order was dependent on the shape selectivity of the stationary phase and the mobile phase composition (Figure 4). Greater resolution, on columns with reduced shape selectivity, of β-HBCD and γ-HBCD was achieved with the use of an acetonitrile/water (compared with a methanol/water) mobile phase composition (Dodder et al., 2006).

TABLE 2. Overview of LC-MS Parameters Used for in the Analysis of BFRs

Compound	Column	Dimensions	Mobile phase (gradient)	Flow (ml/min)	Mobile phase modifiers	Ionisation	Instrument	Ion	Source temp (°C)	References
HBCDs	Luna C_{18} (Phenomenex)	150 × 2 mm; 5 μm	AcN:MeOH:H_2O (y)	0.2	Ammonium acetate	ESI	IT or Q	640.7	160	(Morris et al., 2004)
HBCDs	Genesis C_{18} (Chromatogr. specialties)	50 × 2.1 mm; 4 μm	MeOH:H_2O (y)	0.3	–	ESI	QqQ	MRM (640.6 => 79) MRM	500	(Budakowski and Tomy, 2003)
HBCDs	Symmetry C_{18} (Waters)	150 × 2.1; 5 μm	AcN:MeOH:H_2O (y)	0.25	–	ESI	QqQ	(640.6 => 79) MRM		(Janak et al., 2005)
HBCDs	Symmetry C_{18} (Waters)	150 × 2.1 mm; 3.5 μm	AcN:MeOH:H_2O (y)	0.25	Acetic acid	ESI	QqQ	(640.6 => 79)	–	(Cariou et al., 2005)
PBDEs	Ultrabase RP$_{18}$ (SFCC)	250 × 2 mm; 5 μm	MeOH:Toluene:H_2O (n)	0.2	–	APPI	QTrap	Scan	n.a.	(Debrauwer et al., 2005)
TBBP-A	Nucleodur 100-C_8 (Interchim)	250 × 4 mm; 5 μm	AcN:H_2O (y)	1[a]	Acetic acid	APPI	QTrap	Scan	n.a.	(Debrauwer et al., 2005)
TBBP-A	Luna C_{18} (Phenomenex)	150 × 2 mm; 5 μm	AcN:H_2O (y)	0.25	Ammonium acetate	ESI	Q	540.9	150	(Morris et al., 2004)

(Continued)

(Continued)

TBBP-A	Ace 3 C$_{18}$ (Advanced chromatography technologies)	150 × 2.1 mm; 3.0 μm	MeOH:H$_2$O (y)	0.2	Ammonium acetate	ESI	TOF	Scan (230–550)	130	(Berger et al., 2004)
TBPP-A	Genesis C$_{18}$ (Chromatographic specialties)	150 × 2.1 mm; 4 μm	MeOH:H$_2$O (y)	0.2	–	ESI	QqQ	MRM (543 => 81)	130	(Chu et al., 2005)

[a]Post-column splitting 1:10; IT – ion trap, Q – quadrupole, QqQ – triple quadrupole; TOF – time-of-flight
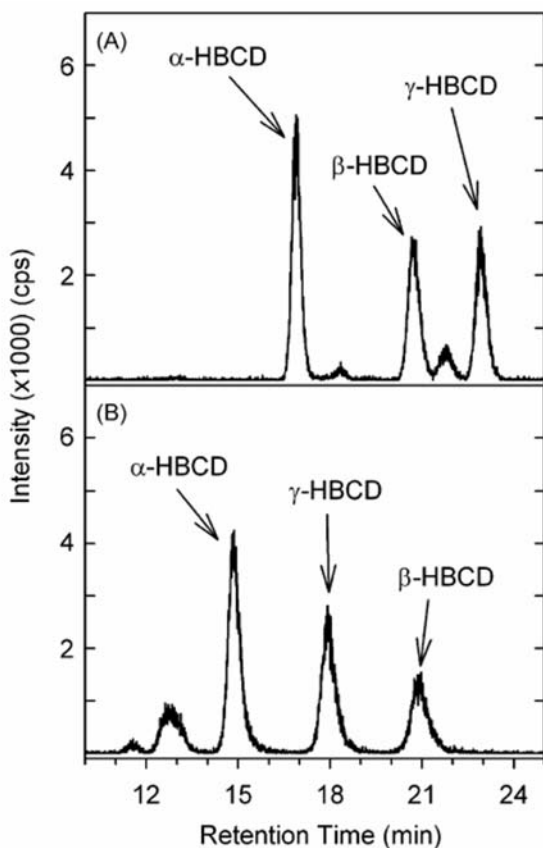
Figure 4. Atmospheric particle-phase MRM chromatograms. The same sample was analyzed using an Eclipse C18 column (A) and a Carotenoid C30 column (B), both 250×4.6 mm. In both cases, the mobile phase was 90% methanol, 10% water (From Dodder et al., 2006)

Although TBBP-A is the most widely used BFR, this compound is not frequently measured, due to its lower concentrations in biota compared to PBDEs and HBCDs and due to its lower bioaccumulation potential. TBBP-A is also more polar than PBDEs and HBCDs, which demands therefore more complicated methods for a proper determination. Acidification and derivatization are compulsory before GC analysis can be carried out, while LC has the advantage that no derivatization step is required (de Boer and Wells, 2006). The LC separation of TBBP-A from other composunds and matrix components is greatly dependent on the mobile phase used. For example, one-third higher response with methanol instead of acetonitrile was reported (Chu et al., 2005). For optimized chromatographic separation and/or ionization response, mobile phase additives, such as formic acid, tris(hydroxymethyl)aminomethane and ammonium acetate, are often used. However, while the two former products were found to give a decreased

ESI-response, ammonium acetate significantly increased the response. Further-more, using a methanol mobile phase alone resulted in a more stable detector baseline and thus a lower LOQ. Therefore, using methanol and water as a mobile phase can be more advantageous for the quantitative analysis of TBBP-A.

3.5. QUALITY ASSURANCE/QUALITY CONTROL

Quality assurance (QA) is a set of procedures, which include the quality control (QC) activities, and which are undertaken to confirm the quality of obtained data. As a general rule, approximately 15% of the analysis time should be spent on the QA procedures. To assure sufficient quality, a number of measures should be taken during the pre-analysis quality control (or validation) and in-process quality control. These measures can be divided into three major areas: calibrants, analytical procedure control, system performance/long-term stability.

A proper dilution (working with volumes sufficiently large to minimize the accuracy errors) and storage of calibration standards (usually cool, dark place) should be employed in such way that after regular checking the weight loss is maximum 2% for a 6–9 month period. Quantification procedures based on internal standard addition (the use of ^{13}C-labelled compounds is recommended) should be addressed to compensate for the losses throughout the analytical procedure, combined with the use of syringe standards for inter-injection fluctuations compensation.

Because all essential steps of analytical procedure are matrix-specific, the analyte recovery, the use of procedural blanks and the determination of limits of detection and quantification should be performed for each compound and matrix to be investigated.

The analytical characteristics of the method should be also considered as an internal quality control by determining the following parameters: repeatability (same operating conditions over a short time), intermediate precision (within-laboratory variation), reproducibility (precision between laboratories) and accuracy (estimated through the use of certified reference materials). The external quality control is usually assessed through the participation in interlaboratory tests which facilitates the evaluation and assessment of the overall method performance.

4.   Toxicity

Although several BFRs are found in quantifiable levels in wildlife and in humans and have been extensively investigated in the last decade, we are still lacking information on the health effects caused by these compounds (Birnbaum et al., 2004; McDonald, 2002).

Generally, HBCD, TBBP-A and PBDEs are absorbed from the gastrointestinal tract and accumulate in fatty tissues. It seems that they don't cause immediate symptoms from acute toxicity at average doses, but their health effects from

chronic exposure are of more concern, especially when they are related to the exposure of developing infants and wildlife. However, based on the available data, it is known that BFRs are associated with several health effects in animal studies, including neurobehavioral toxicity, thyroid hormone disruption and possibly cancer, only for some PBDE congeners. Even if limited information is published in this field, there is some evidence that BFRs can cause developmental effects (Darnerud, 2003), endocrine disruption (McDonald, 2002; Darnerud, 2003), immunotoxicity (Birnbaum et al., 2004; Darnerud, 2003), reproductive and long term effects, including second generation effects (Birnbaum et al., 2004; Kuriyama et al., 2005). For PBDEs and TBBP-A, there is some evidence available for estrogenic activity (Meerts et al., 2001; Legler and Brouwer, 2003), but more studies have to be undertaken to determine if low dose exposures have estrogenic activity in humans or other species. The penta-BDE congeners have been shown to cause toxicity at lower doses than the octa- through deca-BDE congeners (Darnerud, 2003). There are no data on the relative toxicity of the different HBCD isomers or TBBPA derivatives (Birnbaum et al., 2004). Furthermore, there are no data on the toxicity of exposure to BFR mixtures.

One case in history caused the removal from the market in the early 1970s of a class of flame retardants, namely PBBs because of the poisonings in Michigan by mixing a bag containing a commercial PBB mixture, into animal feed. This resulted into long-term impacts on the health of exposed farm families (Dunckel, 1975).

5.   Environmental Levels

Despite of their societal benefits, BFRs seem to migrate from the products in which they are used and entering the environment and people. An increasing numbers of papers, including several reviews, have been published in the last decades and it was shown that BFRs may be measured now in a lot of variety of samples, including air, water, fish, birds, marine mammals and humans. In many cases, it was revealed that the concentrations of these compounds are increasing over time.

5.1. TEMPORAL TRENDS

5.1.1. *Aquatic Environment*

Most data related to the BFR levels presented in the literature focuses on the aquatic environment. A good example of showing time trends in aquatic environment species was underlined by Hites (2004) based on the determination of PBDE concentrations in tissues of marine mammals (seals and porpoises). One data series is composed by samples from the Canadian Arctic, in which low PBDE levels less than 5 ng/g lipid weight (lw) (Ikonomou et al., 2002), while the second series various species all over the world. A significant increasing of ΣPBDE levels
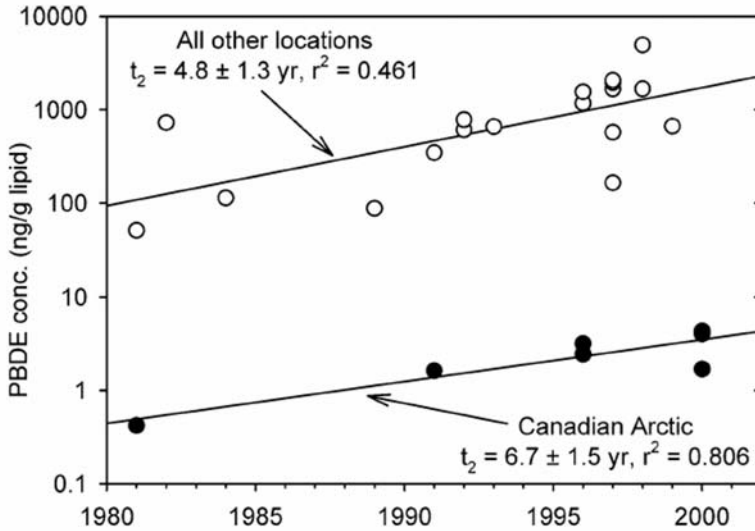
Figure 5. ΣPBDE concentrations in marine mammals (in ng/g lipid weight) shown as a function of the year in which the samples were collected. The bottom line with filled symbols represents samples from the Canadian Arctic (Ikonomou et al., 2002) and the top line with open symbols is for all other samples. The regressions for the two data sets are shown separately; the doubling times of the types of samples are not significantly different (From Hites, 2004)

with the time could be observed for each series (Figure 5). For the Arctic samples, a doubling in the PBDE concentration over a 7 year-period was calculated, while for the rest of the samples, a doubling time of 5 years was estimated (Hites, 2004).

Similar to other organohalogenated pollutants, the BFR levels found in adult marine mammals are significantly higher compared to juvenile animals showing that they are able to bioaccumulate BFRs with time (Law et al., 2003, 2006).

Except few studies, when fish samples are analyzed, the BFR concentrations did not vary systematically with the sampling date either in Europe, in North America or for both combined (Hites, 2004) indicating that temporal trends can not be assessed using such samples. However, a recent study (Batterman et al., 2007) on trends of four PBDE congeners in fish species from Great Lakes, showed large increases in concentrations of PBDEs that started in the early to mid-1980s with fairly consistent doubling times (generally 2–4 years, except in Lake Erie smelt where levels increased very slowly). Furthermore, concentrations and trends show differences by congener, fish species and lake. It was also shown that in fish samples collected recently, the accumulation rates are slowing and concentrations of penta- and hexa-BDE congeners in trout from Lakes Ontario and Michigan and smelt from Lake Ontario started to decrease in the mid-1990s.

Other studies focused on determination of PBDEs from marine sediment cores, showed also time trends for these compounds in collected samples. In sediment cores taken from Baltic Sea (Nylund et al., 1992), a doubling in the PBDE concentration over a 3–4 years period was shown compared to sediment core samples from Drammenfjord, Norway, where a doubling time of approximately 3 years was estimated until the mid-1980s at which time the concentrations seem to have leveled off (Zegers et al., 2003).

Temporal trends for HBCDs were also investigated (Figure 6) through studies applied on different bird eggs (guillemot eggs from the Baltic Sea, peregrine falcon eggs from Greenland or other marine bird eggs from Norway) and also marine mammals (sea lions from USA). HBCD measured in individual and pooled archived guillemot eggs from the Baltic Sea indicated an increase in concentrations between 1969 and 1995 and this increase has leveled-off between 1995 and 2001 when concentrations of HBCDs seem to have stabilized, whereas PBDE concentrations were decreasing (Sellström et al., 2003). A clear and remarkably consistent increasing of HBCD levels since 1983 was reported for the bird egg sampled in the Northern Norway (Knudsen et al., 2005). In another study, Stapleton et al. (2006b) pointed to an exponential increase in the HBCD concentrations with a doubling time of approximately 2 years in California sea lions stranded between 1993 and 2003 (Figure 6). It is unclear at this time why HBCD concentrations increase in the sea lions, while PBDE levels were highly variable without a significant
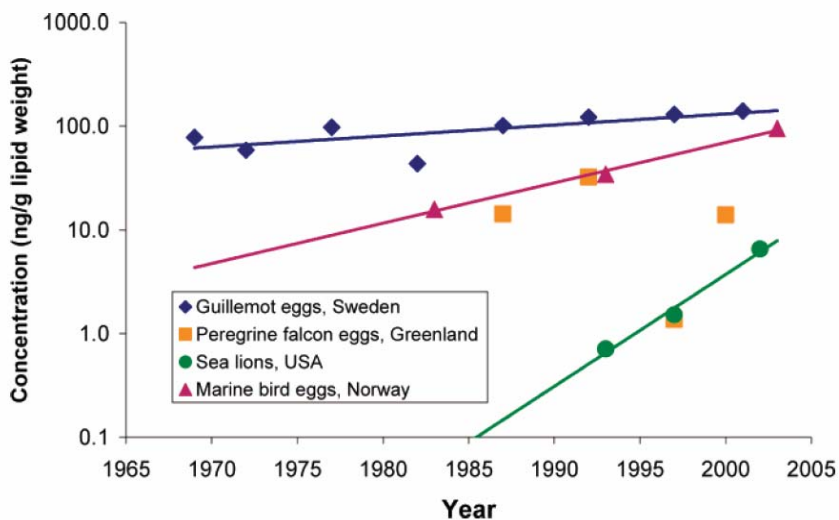

Figure 6. Temporal trends of HBCDs in different bird eggs and in sea lions. Exponential regressions of the complete HBCD concentration data vs. time were significant ($p < 0.05$) for the guillemot eggs, the marine birds eggs and the sea lions data. The peregrine falcon data did not show a significant change over the observed time period (From Covaci et al., 2006)

temporal trend. In male juvenile gray seals from the Baltic Sea sampled in the 1990s, HBCD concentrations were higher compared to those sampled between 1980 and 1985 (Roos et al., 2001).

5.1.2. *Terrestrial Environment*

In contrast to aquatic species, no time trends could be established for HBCDs in terrestrial birds, such as peregrine falcon and sparrowhawk tissues from the UK sampled between 1973 and 2002, due to a low detection frequency and biased sampling (de Boer et al., 2004) or in peregrine falcon eggs from South Greenland sampled between 1986 and 2003 (Vorkamp et al., 2005). Interestingly, for the latter species, a 10% increase per year in the PBDE levels has been observed throughout the investigated time period.

In summary, time trends are not clear yet as the data obtained so far showed either an increasing or no significant trend (Covaci et al., 2006). However, there are no indications available that industry's measures to limit emissions of HBCDs at production and handling sites have led to decreasing concentrations in the environment on a global scale. No study has found parallel time trends for HBCDs and PBDEs. This likely reflects the different regional production and application history for these two BFRs.

5.2. GEOGRAPHICAL TRENDS

When available data regarding the environmental levels of BFRs has to be related to the geographical position of the samples taken into study, few general remarks may be deducted:

(a) Concentrations of BFRs in general (HBCD in particular (Covaci et al., 2006) are often elevated by at least one order of magnitude in the vicinity of plants either producing or using these compounds.
(b) Detection of BFRs in air samples from remote sites, strongly suggest that these compounds undergo long-range transport. Although the influence of possible local sources cannot be ruled out completely, the human activities of these areas are probably not sufficient to explain the environmental levels observed.
(c) In general, the different continental market demands seem to be reflected in different environmental residue levels.

5.2.1. *Aquatic Environment*

When fish samples are analyzed, a high variability of BFR concentrations is recorded, even for the same species, suggesting that these concentrations are related to the proximity of the feeding grounds to BFR sources. These results indicate that fish samples may be used as local indicators of exposure to BFRs. In general, the concentrations of PBDEs in fishes from Europe are significantly

lower than in fishes from North America, the arithmetic and geometric averages for ΣPBDEs are 120 and 50 ng/g lipid for the European fishes, respectively, and 1,050 and 310 ng/g for the North American fishes, respectively (Hites, 2004). Contrarily, HBCD concentrations found in fish (Tommy et al., 2004), dolphins (Peck et al., 2005) and sea lions (Stapleton et al., 2006b) from the North American environment appear to be lower than levels in similar samples from Europe (Janák et al., 2005; Morris et al., 2004; Zegers et al., 2005).

The idea of local monitoring of BFR exposure through fish contamination is sustained by the results of the study performed by Voorspoels et al. (2003). A variety of suspected BFR sources present in the Western Scheldt estuary, such as BFR manufacturing plant, Antwerp harbor or textile industry, was pointed by measuring eight PBDE congeners concentrations in samples of biota, including crab, shrimp, starfish, benthic fish and gadoid fish from the Scheldt Estuary and afterwards compared to those in samples from the Belgian North Sea beyond the mouth of the estuary. Concentrations observed in the Scheldt Estuary samples were up to 30 times higher than in those from the Belgian North Sea, with an increasing gradient towards Antwerp.

5.2.2. *Terrestrial Environment*

Different biota samples have been used to study the geographical distribution of BFRs and other persistent organic pollutants in terrestrial environment, but sometimes it is difficult or expensive to sample these materials. Tree bark has been used successfully as a passive sampler to monitor such contaminants being easy and inexpensive to sample and having a high lipid content and large surface area. Such matrix was used to determine the BFR levels in 29 locations from USA and furthermore, the PBDE levels from US environment were compared with those from few European and Asiatic countries (Zhu and Hites, 2006). The average ΣPBDE concentrations from Italy (13 ng/g lw), Germany (72 ng/g lw) and South Korea (140 ng/g lw) were found to be comparable to those from the northern part of the US and Canada (median concentrations of 83 ng/g lw). For the samples from Germany and Italy, BDE-183, which is a marker of Octa-BDE, contributed around 30% to the total PBDE load, indicating more of the Octa-BDE commercial product was used in Europe than in North America. This is consistent with the known PBDE market demand; in 2001, the ratio of the demand for the Penta- to Octa-BDE was 1:4 in Europe, but it was 5:1 in North America (BSEF, 2007).

Indoor dust analysis was also used as a monitoring tool for BFR contamination, but compared to tree bark samples which are used for external monitoring, dust samples are suitable to determine the in-house exposure to such compounds and finally, human exposure to BFRs. If the use of tree bark samples did not show a very high variability of the data, this is not the case of indoor dust samples. Indeed most of the samples felt within a relatively small range, but some samples were significantly more contaminated (Harrad et al., 2006; Harrad et al., 2007; Wilford et al., 2005). A comparison study regarding the levels of PBDEs in indoor dust samples form several countries (Canada, New Zealand, United Kingdom and

United States) was performed by Harrad et al. (2007) and the median concentrations found for the main BDE congeners and ΣPBDE are presented in Table 3. The results show that North American dusts are contaminated by both Deca- and Penta-BDE commercial formulations, UK dusts are contaminated predominantly by Deca-BDE. The Octa-BDE formulation appears of minimal importance in accordance with available market demand figures. An interesting idea suggested by this study was that despite the commercial PBDE formulations have never been manufactured in nor imported into New Zealand, their presence in dusts from that country suggests international trade in PBDE-containing goods is an important pathway effecting their global distribution.

Reports on BFR concentrations from terrestrial biota samples are generally scarce, compared to aquatic species, with most data being available for top-predator species. de Boer et al. (2004) showed that most samples from predatory water birds had no detectable BDE 209 concentrations, while samples from terrestrial birds of prey had a relative high number of positive results. The extremely low water solubility of BDE 209, the large size of this molecule, combined with a very low uptake and a possible fast metabolism in fish may be explanations for this phenomenon. BDE 209 in dust and other fine particles can apparently be taken up by small terrestrial animals, which are prey for some birds of prey. Terrestrial birds of prey were also intensively studied in order to monitor the contamination with PBDEs (Voorspoels et al., 2006a; Jaspers et al., 2006).

5.3. BIOMAGNIFICATION THROUGH THE FOOD CHAIN

According to above presented data regarding BFRs, it is obvious that these compounds are chemically and biologically persistent and furthermore lipophilic, which results in their bioaccumulation in fatty tissues of organisms and enrichment throughout food chains (Law et al., 2003). As a consequence of bio-magnification, increasing concentrations of PBDEs can be found with increasing trophic level, leading to highest concentrations in top-predators (Law et al., 2006).

Until this date, only few studies have discussed biomagnification of BFRs and they were oriented toward aquatic biota. Based on biomagnification models used by Broman et al. (1992) and Rolff et al. (1993), biomagnification of a persistent substance is independent on the concentration of the substance at the base of the food chain, but it is dependent on the species position in the food chain. Therefore, Burreau et al. (2000) conducted a study on the biomagnification of PBDEs in food chains from the Baltic Sea and the Northern Atlantic Sea and showed that biomagnification occurred similarly, meaning that the ratio between a prey and its predator is the same in spite of different concentrations. The lipid-normalized levels of the major congeners (BDEs 47, 99, and 100) were up to two times higher in large herring than in zooplankton, whereas the levels in salmon were again 2–3 times higher than in large herring.

TABLE 3. Summary of Concentrations (ng/g) of Selected PBDE Congeners in Indoor Dust Samples From Different Countries

Location	N	Median concentrations (ng/g)									ΣPBDEs[a]	Reference
		BDE 28	BDE 47	BDE 99	BDE 100	BDE 153	BDE 154	BDE 183	BDE 207	BDE 209		
Toronto, Canada	–[b]	4.1	140	330	65	43	39	9.0	29	560	970	(Harrad et al., 2008)
Wellington, New Zealand	–[c]	0.65	24	51	8.9	5.4	5.1	–	–	–	–	(Harrad et al., 2008)
Birmingham, UK	–[d]	0.53	13	23	4.2	5.2	3.3	13	57	2800	3000	(Harrad et al., 2008)
Amarillo and Austin, TX, US	–[e]	14	410	820	160	110	89	16	71	1300	4000	(Harrad et al., 2008)
Ottawa, Canada, n = 68	68	3.0	300	430	73	49	37	19	–	630	1800	(Wilford et al., 2006)
Various regions, UK	10	0.35	24.8	44	–	23	–	–	–	7100	–	(Santillo et al., 2003)
Newcastle, UK	10	–	22	28	4	5	3	–	–	10000	10000	(Sjödin et al., 2006)
Various locations, US	17	14.8	644	676	119	64.4	72.8	17.6	19.1	1350	4250	(Stapleton et al., 2005)
Various locations, Romania	18	–	2.5	2.6	0.4	2.2	0.4	–	–	482	490	(Dirtu and Covaci, unpublished data)

[a]Sum of PBDEs 28, 47, 49, 66, 99, 100, 153, 154, 183, 196, 197, 203, and 209.
[b]Ten samples analysed for tri-hexa BDEs; seven samples analysed for tri-deca BDEs.
[c]Twenty samples analysed for tri-hexa BDEs; hepta-deca BDEs not analysed.
[d]Twenty eight samples analysed for tri-hexa BDEs (eight reported previously [Harrad et al., 2006]); sixteen samples analysed for tri-deca BDEs.
[e]Twenty samples analysed for tri-hexa BDEs; seventeen samples analysed for tri-deca BDEs.

It was also shown that the major biomagnification step in the aquatic food chain occurs from fish to marine mammals based on the fact that the lipid normalized PBDE levels in blubber and liver were similar and generally higher with more than one order of magnitude compared to levels found in invertebrates and fish (Law et al., 2003). Additionally, Muir et al. (2006) demonstrated that some PBDE congeners show substantial biomagnification from seals to polar bears with biomagnification factors ranging from 3.9 to 71.

The biomagnification potential of the most commonly reported PBDE-congeners was recently assessed in three small terrestrial food chains (Voorspoels et al., 2007a). To achieve this, PBDE data on passerines and rodents were combined with previously published data on terrestrial predators from the same area, i.e., birds of prey (Voorspoels et al., 2006a; Jaspers et al., 2006) and red fox (Voorspoels et al., 2006b). Because birds of prey occupy top positions in the food chain, it is most likely that biomagnification of persistent and lipophilic PBDEs through their preys would be substantial. Therefore, BMFs (the ratio of lipid normalized concentrations in the same tissues (if available) of predator and prey) were calculated by authors using the median concentrations. The BMF values for PBDEs in buzzard and sparrowhawk ranged from 2 to 34, depending on the congener, thus evidencing biomagnification. All PBDE congeners that could be determined in both prey and predator were biomagnified (BMF > 1). Surprisingly, no biomagnification could be observed from rodents to foxes. The median levels of sum PBDEs measured in fox were even lower than those in rodents. This was not in accordance with the biomagnification hypothesis and the observations made for BMFs in birds.

5.4. HUMAN EXPOSURE TO BFRS

A large number of samples from people (including human tissues, serum and milk) have been analyzed for PBDEs and, it could be concluded that the PBDE concentrations have increased in people by a factor of ~100 during the last 30 years (Hites, 2004). The regression of these data as a function of year is good, despite the disparate sample types, the different continents of origin, and the various congeners measured (Figure 7). This analysis shows that samples coming from North America are always above the regression line (in recent years by a factor of >10) and that the Japanese samples are usually below the regression line (by a factor of ~5) showing that the results are related to the exposure level of the general population from studied areas.
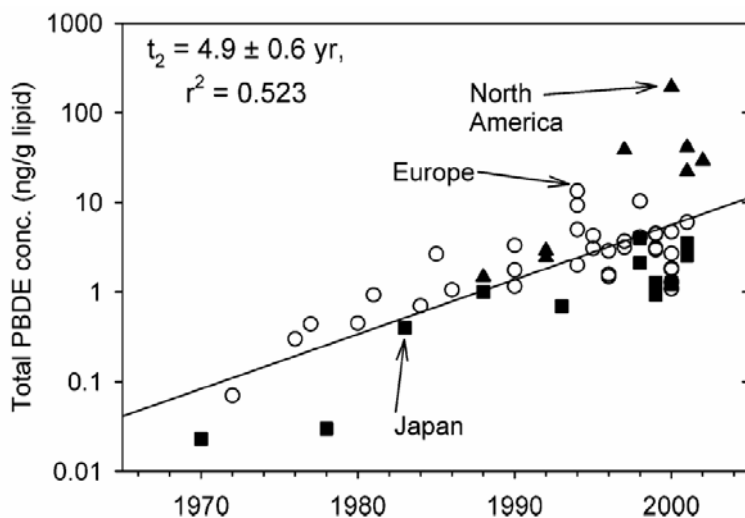
Figure 7. Total PBDE concentrations (ΣPBDE) in human blood, milk and tissue (in ng/g lipid) shown as a function of the year in which the samples were taken. The three symbol types indicate the location from which the samples were collected. The overall regression is shown (From Hites, 2004)

Considering that BFRs are used in plastics or other materials which end up in consumer electronics, it would be expected that workers involved in assembling or disassembling of these products would have an increased level of some BFRs in their blood. Indeed, occupational exposure has been reported to result from the repair and maintenance of computers (Jakobsson et al., 2002), dismantling electronics (Sjödin et al., 1999; Bi et al., 2007) and recycling printed circuit boards (Thuresson et al., 2002), commercial Deca-BDE flame retarded rubber manufacture and handling electric cables using the same rubber (Thuresson et al., 2005). Furthermore, different PBDE profiles were observed comparing occupationally exposed (with higher levels of PBDEs with 7–10 bromine atoms) and non-occupationally exposed populations (with higher levels of PBDEs with 3–6 bromine atoms) (reviewed by Sjödin et al., 2003).

Another pathway for human exposure to BFRs is the dietary intake (Voorspoels et al., 2007b). Due to the lipophilic nature of these chemicals, BFRs are mostly found in lipid-rich food of animal origin, such as meat, fish and diary products, which are a part of our daily diet. It has been shown that food, and more in particular food of animal origin, is responsible for more than 90% of the average human intake of polychlorinated biphenyls (PCBs) (Liem et al., 2000) and therefore for BFRs, this should be similar.

Acknowledgments

References

1. Ackerman LK, Wilson GR, Simonich SL. Quantitative analysis of 39 polybrominated diphenyl ethers by isotope dilution GC/low-resolution MS. *Anal Chem* **2005**, 77, 1979–1987.
2. Allchin CR, Law RJ, Morris S. Polybrominated diphenyl ethers in sediments and biota downstream of potential sources in the UK. *Environ Pollut* **1999**, 105, 197–207.
3. Batterman S, Chernyak S, Gwynn E, Cantonwine D, Jia C, Begnoche L, Hickey JP. Trends of brominated diphenyl ethers in fresh and archived Great Lakes fish (1979–2005). *Chemosphere* **2007**, 69, 444–457.
4. Berger U, Herzke D, Sandanger TM. Two trace analytical methods for determination of hydroxylated PCBs and other halogenated phenolic compounds in eggs from Norwegian birds of prey. *Anal Chem* **2004**, 76, 441–447.
5. Bi X, Thomas GO, Jones KC, Qu W, Sheng G, Martin FL, Fu J. Exposure of electronics dismantling workers to polybrominated diphenyl ethers, polychlorinated biphenyls, and organochlorine pesticides in South China. *Environ Sci Technol* **2007**, 41, 5647–5653.
6. Birnbaum LS, Staskal DF. Brominated flame retardants: Cause for concern? *Environ Health Perspect* **2004**, 112, 9–17.
7. Björklund J, Tollbäck P, Hiarne C, Dyremarck E, östman C. Influence of the injection technique and the column system on gas chromatographic determination of polybrominated diphenyl ethers. *J Chromatogr A* **2004**, 1041, 201–210.
8. Broman D, Näf C, Rolff C, Zebühr Y, Fry B, Hobbie J. Using ratios of stable nitrogen isotopes to estimate bioaccumulation and flux of polychlorinated dibenzo-p-dioxins (PCDDs) and dibenzofurans (PCDFs) in 2 food-chains from the Northern Baltic. *Environ Toxicol Chem* **1992**, 11, 331–345.
9. Bromine Science and Environmental Forum (BSEF). Website: http://www.bsef.com last accessed 15th December 2007.
10. Budakowski W, Tomy G. Congener-specific analysis of hexabromocyclododecane by high-performance liquid chromatography/electrospray tandem mass spectrometry. *Rapid Commun Mass Spectrom* **2003**, 17, 1399–1403.
11. Burreau S, Zebühr Y, Ishaq R, Broman D. Comparison of biomagnification of PBDEs in food chains from the Baltic Sea and the Northern Atlantic Sea. *Organohalogen Compd* **2000**, 47, 253–255.
12. Cariou R, Antignac JP, Marchand P, Berrebi A, Zalko D, Andre F, Le Bizec B. New multiresidue analytical method dedicated to trace level measurement of brominated flame retardants in human biological matrices. *J Chromatogr A* **2005**, 1100, 144–153.
13. Chu SG, Haffner GD, Letcher RJ. Simultaneous determination of tetrabromobisphenol A, tetrachlorobisphenol A, bisphenol A and other halogenated analogues in sediment and sludge by high performance liquid chromatography-electrospray tandem mass spectrometry. *J Chromatogr A* **2005**, 1097, 25–30.

14. Covaci A, de Boer J, Ryan JJ, Voorspoels S, Schepens P. Determination of polybrominated diphenyl ethers and polychlorinated biphenyls in human adipose tissue by large-volume injection-narrow-bore capillary gas chromatography/electron impact low-resolution mass spectrometry. *Anal Chem* **2002**, 74, 790–798.

15. Covaci A, Voorspoels S, de Boer J. Determination of brominated flame retardants, with emphasis on polybrominated diphenyl ethers (PBDEs) in environmental and human samples–a review. *Environ Int* **2003**, 29, 735–756.

16. Covaci A, Gerecke AC, Law RJ, Voorspoels S, Kohler M, Heeb NV, Leslie H, Allchin CR, de Boer J. Hexabromocyclododecanes (HBCDs) in the environment and humans: A review. *Environ Sci Technol* **2006**, 40, 3679–3688.

17. Covaci A, Voorspoels S, Ramos L, Neels H, Blust R. Recent developments in the analysis of brominated flame retardants and brominated natural compounds. *J Chromatogr A* **2007**, 1153, 145–171.

18. Darnerud PO. Toxic effects of brominated flame retardants in man and wildlife. *Environ Int* **2003**, 29, 841–853.

19. de Boer J, Allchin C, Law R, Zegers B, Booij JP. Method for the analysis of polybrominated diphenyl ethers in sediments in biota. *Trends Anal Chem* **2001**, 20, 591–599.

20. de Boer J, Leslie LA, Leonards PEG, Bersuder P, Morris S, Allchin CR. Screening and time trend study of decabromodiphenylether and hexabromocyclododecane in birds. In: Proceedings of the Third International Workshop on Brominated Flame Retardants BFR 2004, Toronto, Canada, June 6–9, **2004**, pp. 125–128.

21. de Boer J, Wells DE. Pitfalls in the analysis of brominated flame retardants in environmental, human and food samples–including results of three international interlaboratory studies. *Trends Anal Chem* **2006**, 25, 364–372.

22. de Wit C. An overview of brominated flame retardants in the environment. *Chemosphere* **2002**, 46, 583–624.

23. Debrauwer L, Riu A, Jouahri M, Rathahao E, Jouanin I, Antignac JP, Cariou R, Le Bizec B, Zalco D. Probing new approaches using atmospheric pressure photo ionization for the analysis of brominated flame retardants and their related degradation products by liquid chromatography–mass spectrometry. *J Chromatogr A* **2005**, 1082, 98–109.

24. Dirtu AC, Ravindra K, Roosens L, van Grieken R, Neels H, Blust R, Covaci A. Fast analysis of decabrominated diphenyl ether using low-pressure gas chromatography–electron-capture negative ionization mass spectrometry. *J Chromatogr A* **2008**, 1186, 295–301.

25. Dodder NG, Peck AM, Kucklick JR, Sander LC, Analysis of hexabromocyclododecane diastereomers and enantiomers by liquid chromatography/tandem mass spectrometry: Chromatographic selectivity and ionization matrix effects. *J Chromatogr A* **2006**, 1135, 36–42.

26. Dunckel AE. An updating on the polybrominated biphenyl disaster in Michigan. *J Am Vet Med Assoc* **1975**, 167, 838–841.

27. Focant JF, Sjödin A, Turner WE, Patterson Jr. DG. Measurement of selected polybrominated diphenyl ethers, polybrominated and polychlorinated biphenyls, and organochlorine pesti-cides in human serum and milk using comprehensive two-dimensional gas chromatography isotope dilution time-of-flight mass spectrometry. Anal Chem 2004, 76, 6313–6319.

28. Gómara B, Herrero L, Bordajandi LR, González MJ. Quantitative analysis of polybrominated diphenyl ethers in adipose tissue, human serum and foodstuff samples by gas chromatography with ion trap tandem mass spectrometry and isotope dilution, Rapid Commun Mass Spectrom 2006, 20, 69–76.

29. Harrad S, Hazrati S, Ibarra C. Concentrations of polybrominated diphenyl ethers in indoor air and dust and polychlorinated biphenyls in indoor air in Birmingham, UK: Implications for human exposure. *Environ Sci Technol* **2006**, 40, 4633–8.

30. Harrad S, Ibarra C, Diamond M, Melymuk L, Robson M, Douwes J, Roosens L, Dirtu AC, Covaci A. Polybrominated diphenyl ethers in domestic indoor dust from Canada, New Zealand, United Kingdom and United States. *Environ Int* **2008**, in press.

31. Hites RA. Polybrominated diphenyl ethers in the environment and in people: A meta-analysis of concentrations. *Environ Sci Technol* **2004**, 38, 945–956.
32. Ikonomou MG, Rayne S, Addison RF. Exponential increases of the brominated flame retardants, polybrominated diphenyl ethers, in the Canadian Arctic from 1981 to 2000 *Environ Sci Technol* **2002**, 36, 1886–1892.
33. Jakobsson K, Thuresson K, Rylander L, Sjödin A, Hagmar L, Bergman A. Exposure to polybrominated diphenyl ethers and tetrabromobisphenol A among computer technicians. *Chemosphere* **2002**, 46, 709–716.
34. Janak K, Covaci A, Voorspoels S, Becher G. Hexabromocyclododecane in marine species from the Western Scheldt Estuary: Diastereomer- and enantiomer-specific accumulation. *Environ Sci Technol* **2005**, 39, 1987–1994.
35. Jaspers VLB, Covaci A, Voorspoels S, Dauwe T, Eens M, Schepens P. Brominated flame retardants and organochlorine pollutants in aquatic and terrestrial predatory birds of Belgium: Levels, patterns, tissue distribution and condition factors. *Environ Pollut* **2006**, 139, 340–352.
36. Kierkegaard A, Björklund J, Friden U. Identification of the flame retardant decabromodiphenyl ethane in the environment. *Environ Sci Technol* **2004**, 38, 3247–3254.
37. Knudsen LB, Gabrielsen GW, Verreault J, Barrett R, Skare JU, Polder A, Lie E. Temporal trends of brominated flame retardants, cyclododeca-1,5,9-triene and mercury in eggs of four sea bird species from Northern Norway and Svalbard; Report 942/205; Norwegian Pollution Control Authority, Oslo, **2005.**
38. Korytár P, Covaci A, de Boer J, Gelbin A, Brinkman UATh. Retention-time database of 126 polybrominated diphenyl ether congeners and two Bromkal technical mixtures on seven capillary gas chromatographic columns. *J Chromatogr A* **2005a**, 1065, 239–251.
39. Korytár P, Covaci A, Leonards P E G, de Boer J, Brinkman UATh. Comprehensive two-dimensional gas chromatography of polybrominated diphenyl ethers. *J Chromatogr A* **2005b**, 1100, 20–31.
40. Kuriyama SN, Talsness CE, Grote K, Chahoud I. Developmental exposure to low dose BDE 99. Effects on male fertility and neurobehavior in rat offspring. *Environ Health Perspect* **2005**, 113, 149–154.
41. Law RJ, Alaee M, Allchin C, Boon JP, Lebeuf M, Lepom P, Stern GA. Levels and trends of Polybrominated diphenylethers and other brominated flame retardants in wildlife. *Environ Int* **2003**, 29, 757–770.
42. Law RJ, Allchin CR, de Boer J, Covaci A, Herzke D, Lepom P, Morris S, Tronczynski J, de Wit CA. Levels and trends of brominated flame retardants in the European environment. *Chemosphere* **2006**, 64, 187–208.
43. Legler J, Brouwer A. Are brominated flame retardants endocrine disruptors? *Environ Int* **2003**, 29, 879–885.
44. Liem AKD, Fürst P, Rappe C. Exposure of populations to dioxins and related compounds. *Food Addit Contam* **2000**, 17, 241–259.
45. Manchester-Neesvig JB, Valters K, Sonzogni WC. Comparison of polybrominated diphenyl ethers (PBDEs) and polychlorinated biphenyls (PCBs) in Lake Michigan salmonids. *Environ Sci Technol* **2001**, 35, 1072–1077.
46. McDonald TA. A perspective on the potential health risks of PBDEs. *Chemosphere* **2002**, 46, 745–755.
47. Meerts I, Letcher RJ, Hoving S, Marsh G, Bergman A, Lemmen J, van der Burg B, Brouwer A. In vitro estrogenicity of polybrominated diphenyl ethers, hydroxylated PBDEs, and poly-brominated bisphenol A compounds. *Environ Health Perspect* **2001**, 109, 399–407.
48. Morris S, Allchin CR, Zegers BN, Haftka JJH, Boon JP, Belpaire C, Leonards PEG, Van Leeuwen SPJ, de Boer J. Distribution and fate of HBCD and TBBP-A flame retardants in North Sea estuaries and aquatic food webs. *Environ Sci Technol* **2004**, 38, 5497–5504.

49. Muir DCG, Backus S, Derocher A, Dietz R, Evans TJ, Gabrielsen GW, Nagny J, Norstrom R, Sonne C, Stirling I, Taylor MK, Letcher RJ. Brominated flame retardants in polar bears (*Ursus maritimu*s) from Alaska, the Canadian Arctic, East Greenland, and Svalbard. *Environ Sci Technol* **2006**, 40, 449–455.

50. Nylund K, Asplund L, Jansson B, Jonsson P. Analysis of some polyhalogenated organic pollutants in sediment and sewage-sludge. *Chemosphere* **1992**, 24, 1721–1730.

51. Peck AM, Tuerk KJS, Keller J, Kucklick JR, Schantz MM. Hexabromocyclododecane diastereomers and enantiomers in white-sided dolphin blubber and liver tissue. *Organohalogen Compd* **2005**, 67, 1259–1262.

52. Rolff C, Broman D, Näf C, Zebühr Y. Potential biomagnification of PCDD/Fs–new possibilities for quantitative assessment using stable-isotope trophic position, *Chemosphere* **1993**, 27, 461–468.

53. Roos A, Nylund K, Haggberg L, Asplund L, Bergman A, Olsson M. Brominated flame retardants (BFR) in young Grey Seal Males (*Halicoerus grypus*) from the Baltic Sea. Proceedings of the Second International Workshop on Brominated Flame Retardants, BFR 2001, Stockholm, Sweden, 14–16 May, **2001,** 337–341.

54. Roosens L, Dirtu AC, Goemans G, Belpaire C, Gheorghe A, Neels H, Blust R, Covaci A. Brominated flame retardants and organochlorine contaminants in fish from the Scheldt River, Belgium. *Environ Int* **2008**, doi: 10.1016/j.envint.2008.02.009.

55. Salgado-Petinal C, Garcia-Chao M, Llompart M, Garcia-Jares C, Cela R. Headspace solid-phase microextraction gas chromatography tandem mass spectrometry for the determination of brominated flame retardants in environmental solid samples. *Anal Bioanal Chem* **2006**, 385, 637–347.

56. Santillo D, Labunska I, Davidson H, Johnston P, Strutt M, Knowles O. Consuming chemicals: hazardous chemicals in house dust as an indicator of chemical exposure in the home: Part I–UK. Greenpeace Research Laboratories Technical Note 01/2003; **2003**.

57. Sellström U. Determination of some polybrominated flame retardants in biota, sediment and sewage sludge. PhD Thesis. University of Stockholm, Sweden, **1999**.

58. Sellström U, Bignert A, Kierkegaard A, Haggberg L, de Wit CA, Olsson M, Jansson B. Temporal trend studies on tetra and pentabrominated diphenyl ethers and hexabromocyclododecane in guillemot egg from the Baltic Sea. *Environ Sci Technol* **2003,** 37, 5496–5501.

59. Sjödin A, Hagmar L, Klasson-Wehler E, Kronholm-Diab K, Jakobsson E, Bergman A. Flame retardant exposure: Polybrominated diphenyl ethers in blood from Swedish workers. *Environ Health Perspect* **1999**, 107, 643–648.

60. Sjödin A, Patterson DG, Bergman A. A review on human exposure to brominated flame retardants–particularly polybrominated diphenyl ethers. *Environ Int* **2003**, 29, 829–839.

61. Sjödin A, Päpke O, Focant J-F, Jones RS, Pless-Mulloli T, Leontjew Toms L-M. Concentration of polybrominated diphenyl ethers (PBDEs) in household dust from various countries-is dust a major source of human exposure? *Organohalogen Compd* **2006**, 68, 2181–2185.

62. Smedes F, de Boer J. Determination of PCBs in sediments–analytical methods. *Trends Anal Chem* **1997**, 16, 503–517.

63. Stapleton HM, Dodder NG, Offenberg JH, Schantz MM, Wise SA. Polybrominated diphenyl ethers in house dust and clothes dryer lint. *Environ Sci Technol* **2005**, 39, 925–931.

64. Stapleton HM. Instrumental methods and challenges in quantifying polybrominated diphenyl ethers in environmental extracts: a review. *Anal Bioanal Chem* **2006a**, 386, 807–817.

65. Stapleton HM, Dodder NG, Kucklick JR, Reddy CM, Schantz MM, Becker PR, Gulland F, Porter BJ, Wise SA. Determination of HBCD, PBDEs and MeO-BDEs in California sea lions (*Zalophus californianus*) stranded between 1993 and 2003. *Mar Pollut Bull* **2006b**, 52, 522–531.

66. Thuresson K, Jakobsson K, Hagmar L, Englyst V, Bergman Å. Work related exposure to bro-minated flame retardants when recycling metals from printed circuit boards. *Organohalogen Compd* **2002**, 58, 249–252.

67. Thuresson K, Bergman A, Jakobsson K. Occupational exposure to commercial decabromo-diphenyl ether in workers manufacturing or handling flame-retarded rubber. *Environ Sci Technol* **2005**, 39, 1980–1986.

68. Voorspoels S, Covaci A, Schepens P. Polybrominated diphenyl ethers in marine species from the Belgian North Sea and the Western Scheldt estuary: levels, profiles and distribution. *Environ Sci Technol* **2003**, 37, 4348–4357.

69. Voorspoels S, Covaci A, Lepom P, Jaspers VLB, Schepens P. Levels and distribution of polybrominated diphenyl ethers in various tissues of birds of prey. *Environ Pollut* **2006a**, 144, 218–227.

70. Voorspoels S, Covaci A, Lepom P, Escutenaire S, Schepens P. Remarkable findings concerning PBDEs in the terrestrial toppredator red fox (*Vulpes vulpes*). *Environ Sci Technol* **2006b**, 40, 2937–2943.

71. Voorspoels S, Covaci A, Jaspers VLB, Neels H, Schepens P. Biomagnification of PBDEs in three small terrestrial food chains. *Environ Sci Technol* **2007a**, 41, 411–416.

72. Voorspoels S, Covaci A, Neels H, Schepens P. Dietary PBDE intake: A market basket study in Belgium. *Environ Int* **2007b**, 33, 93–97.

73. Vorkamp K, Thomsen M, Falk K, Leslie H, Møller S, Sørensen PB. Temporal development of brominated flame retardants in peregrine Falcon (*Falco peregrinus*) eggs from South Greenland (1986–2003). *Environ Sci Technol* **2005**, 39, 8199–8206.

74. Wang D, Cai Z, Jiang G, Wong MH, Wong WK. Gas chromatography/ion trap mass spectrometry applied for the determination of polybrominated diphenyl ethers in soil. *Rapid Commun Mass Spectrom* **2005**, 19, 83–88.

75. WHO/ICPS. Environmental Health Criteria 162: Brominated Diphenyl Ethers. Geneva: World Health Organization; 1994.

76. WHO/ICPS. Environmental Health Criteria 192: Flame Retardants–General introduction. Geneva: World Health Organization; 1997.

77. Wilford BH, Shoeib M, Harner T, Zhu J, Jones KC. Polybrominated diphenyl ethers in indoor dust in Ottawa, Canada: implications for sources and exposure. *Environ Sci Technol* **2005**, 39, 7027–7035.

78. Zegers BN, Lewis WE, Booij K, Smittenberg,RH, Boer W, de Boer J, Boon JP. Levels of polybrominated diphenyl ether flame retardants in sediment cores from Western Europe. *Environ Sci Technol* **2003**, 37, 3803–3807.

79. Zegers BN, Mets A, van Bommel R, Minkenberg C, Hamers T, Kamstra JH, Pierce G, Reid B, Patterson T, Boon JP. Levels of hexabromocyclododecane in harbour porpoises and common dolphins from Western European Seas, with evidence for stereoisomer-specific biotransformation by cytochrome P450. *Environ Sci Technol* **2005**, 39, 2095–2100.

80. Zhu L, Hites RA. Brominated flame retardants in tree bark from North America. *Environ Sci Technol* **2006**, 40, 3711–3716.

12. STEREOCHEMISTRY STUDIES OF SOME 1,3-DIOXANE DERIVATIVES BY DIFFERENTIAL MASS SPECTROMETRY AND COMPUTATIONAL CHEMISTRY

FLORIAN HARJA[1], CHRISTINE BETTENDORF[2],
ION GROSU[3], AND NICOLAE DINCA[1]
[1]*Mass Spectrometry Laboratory, Department of Chemistry and Biology, University of Arad, Romania*
[2]*BET2 Software, Konigsbrunn, Germany*
[3]*Department of Organic Chemistry, "Babes-Bolyai" University of Cluj-Napoca, Romania*

Abstract. In this work mass spectrometry is for the first time used in the stereo-chemical study of and conformational equilibriums of the saturated 6 atom rings, namely of 1,3 dioxans: *cis-* and *trans-*2-phenyl-4-methyl-1,3-dioxane and *cis-* and *trans-*2-phenyl-5-methyl-1,3-dioxane. The technique involved is based on the correlation of the experimental differential mass spectra (diff ms) with the enthalpies of formation obtained through quantum-chemical calculations for the main fragmentation ions from the mass spectra. The lists with the correlation probabilities were obtained using the computer program *Chemical Structure Identification by Differential Mass Spectra* (CSI Diff ms). The obtained results are in agreement with NMR and X-Ray studies made of these compounds. This also highlighted the preference of the aromatic substituent for the axial position, when the molecule is isolated (in vacuum).

1. Introduction

The elaboration of new techniques for the stereochemistry study of six atoms saturated rings is a must as these structures are present in numerous biomolecules and synthetic compounds. Establishing the conformational structure is very important when the biological role of some compounds or their quality is assessed. The stereo-chemical study of the 1,3 dioxan rings is a field where magnetic nuclear resonance and the X-ray diffraction have proved their efficiency. Until now mass spectrometry was not used to obtain this kind of information because it was considered that the energy used for ionization, which was much higher than the energy difference of the conformers, would have a hindering effect in obtaining relevant information.

Due to the fact that in the past years differential mass spectra were used to distinguish the structural differences that appear between isomer substances possessing similar mass spectra, we studied how can this way of approach could be used
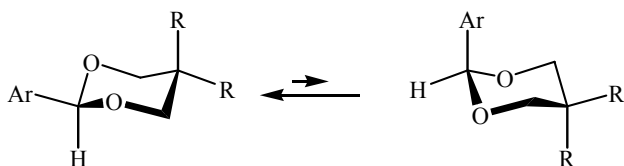
Figure 1. Conformational equilibrium of 2-aryl-1,3-dioxanes revealed by NMR analysis [1–4]

to establish the conformation of 1,3 dioxan rings. In order to obtain differential mass spectra does not mean that one must use a special type of mass spectrometer or new techniques to process the samples or data acquisition.

Studies [1–4] on the stereochemistry of 2-aryl-1,3-dioxanes revealed the high preference of the aromatic substitute for the equatorial position, the characteristic conformational equilibrium being shifted toward the conformer exhibiting the aryl group in equatorial orientation (see Figure 1). The high A-value for aryl groups (e.g., $A_{Ph} = 3.12$ kcal/mol) suggest that these substitutes are efficient "holding groups"[5].

On the other hand, quantum-chemical calculations of the standard forming heats for isolated molecules indicate a higher stability of the structures that have the aromatic substitute axial oriented (Figure 2).
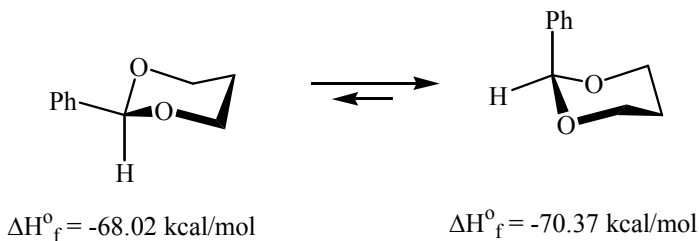

$\Delta H^o_f = -68.02$ kcal/mol $\Delta H^o_f = -70.37$ kcal/mol

Figure 2. Conformational equilibrium of 2-phenyl-1,3-dioxane indicated by quantum chemical calculations

Taking in consideration that the GC/MS analysis is performed by ionizing the isolated molecule, one should expect that there is a correlation between the intensity of the peaks from the fragmentation spectra and the stability of the substituted 2-Ar-1,3-dioxans diastereoisomers. This paperwork presents the way in which correlation degree between the differential mass spectra and calculated heats of formation obtained from quantum chemistry calculations applied to the ions (Diff MS-CQC), can be used in discriminating the diastereoisomers possessing this structure.

2.   Experimental

2-phenyl-4-methyl-1,3-dioxane and 2-phenyl-5-methyl-1,3-dioxane obtained by acetalization of benzaldehyde with butane-1,3-diol and respectively iso-butane-1,3-diol were used. The diastereoisomers mixture of obtained compounds (Figures 3 and 4) was characterized by ^{1}H-NMR, ^{13}C-NMR in other papers [6,7]. The mass spectra were acquired with a HP EI-GC/MSD instrument in the positive ion mode.

2.1. STRATEGY OF ΔH^0_f DATABASE CALCULUS

The heats of formation (ΔH^0_f) of molecules and fragment ions database were calculated. The geometries of the molecules and radicals were optimised with the MM+ force field and re-optimised with the semi-empirical method AM1 [8–10], using the RHF operators for molecules or ions and UHF for the radical ions. For each diastereoisomer and the corresponding ions a ponderated heat of formation was calculated using the equilibrium constant (K_{eq}) of the two conformers, A and B, from equation (1).

$$\Delta H^0_f = \Delta H^0_{fB} \frac{K_{eq}}{(K_{eq}+1)} + \Delta H^0_{fA} \frac{1}{(K_{eq}+1)}$$
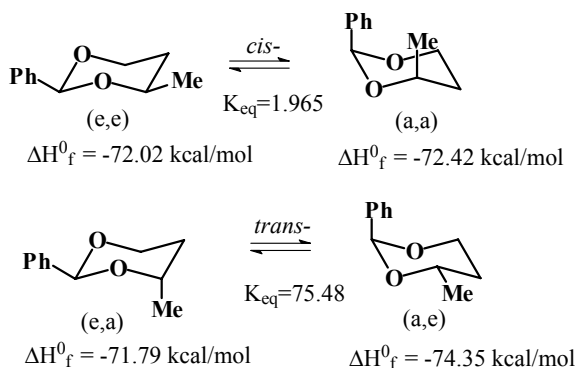
(1)

Figure 3 The conformers of 2-phenyl-4-methyl-1,3-dioxane diastereoisomers

The calculated ponderated heats of formation were introduced in the database of Chemical Structure Identification by Differential Mass Spectra–CSI Diff ms program (see Tables 1 and 2).
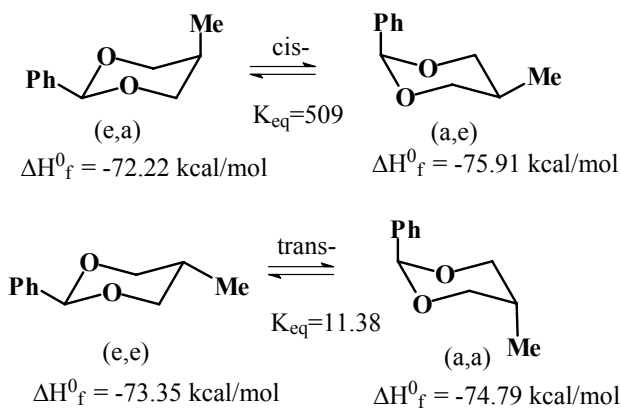
Figure 4. The conformers of 2-phenyl-5-methyl-1,3-dioxane diastereoisomers

TABLE 1. Heat of Formation for 2-phenyl-4-methyl-1,3-dioxane Molecules and Ions

ΔH^0_f (kcal/mol)	Conformer	M	M^+	$[M-H]^+$	$[M-Me]^+$	$[M-Ph]^+$	$[Ph]^+$
cis-isomer	(e, e)	−72.02	129.72	100.35	203.78	81.17	
K_{eq} = 1.965	(a, a)	−72.42	132.06	100.62	182.58	81.45	
$\Delta H^0_{f\ cis\ ponderal}$		−72.28	131.27	100.52	189.72	81.35	283.6
trans-isomer	(a, e)	−74.35	131.42	100.35	203.78	81.17	
K_{eq} = 0.01325	(e, a)	−71.79	129.69	100.62	153.57	81.45	
$\Delta H^0_{f\ trans\ ponderal}$		−74.31	131.39	100.35	203.12	81.17	283.6

TABLE 2. Heat of Formation for 2-phenyl-5-methyl-1,3-dioxane Molecules and Ions

ΔH^0_f (kcal/mol)	Conformer	M	M^+	$[M-H]^+$	$[M-Me]^+$	$[M-Ph]^+$	Ph^+
cis-isomer	(e, a)	−72.22	129.78	101.36	158.87	82.65	
K_{eq} = 509	(a, e)	−75.91	130.26	100.41	153.57	81.78	
$\Delta H^0_{f\ cis\ ponderal}$		−75.90	130.25	100.41	153.58	81.78	283.6
trans-isomer	(e, e)	−73.35	128.55	100.41	158.87	81.78	
K_{eq} = 11.38	(a, a)	−74.79	131.47	101.36	153.57	82.65	
$\Delta H^0_{f\ trans\ ponderal}$		−74.67	131.23	101.28	153.99	82.57	283.6

2.2. CORRELATION OF DIFFERENTIAL MASS SPECTRA WITH THE HEATS OF FORMATION DATABASE

The files of the analyses were imported in CSI Differential Mass Spectrometry Data Analysis 2.5.1 software, produced by Bet2 Software Company, prior to which a subtract background correction was performed. For differentiation ms scans having approximately the same TIC on the rising side of each chromatographic peak were considered. The plots of the differential spectra are presented in Figure 5 [11].

The thermodynamic data of principal ions are loaded automatically by the input of 2-phenyl-4-methyl-1,3-dioxane and respectively 2-phenyl-5-methyl-1,3-dioxane diastereoisomers using the command "Select Ion Database Record" of CSI-Diff-ms. The establishment of the structure are made using "Calculate Probability" command. The results are given as spreadsheets containing the probabilities of the possible variants (see Figure 6). The correct structural choice is corresponding to the highest probability. (Tables 3 and 4)

TABLE 3. The Probability List Resulted by GC/MS and CSI Diff ms–CQC Analysis for *cis*- and *trans*- 2-phenyl-4-methyl-1,3-dioxane Isomers

Probability	Isomer X	Isomer Y
	Retention time 24.25 min	Retention time 24.33 min
80%	*cis*-2-Phenyl-4-methyl-1,3-dioxane	*trans*-2-Phenyl-4-methyl-1,3-dioxane
20%	*trans*-2-Phenyl-4-methyl-1,3-dioxane	*cis*-2-Phenyl-4-methyl-1,3-dioxane

TABLE 4. The Probability List resulted by GC/MS and CSI Diff ms–CQC Analysis for *cis*- and *trans*- 2-phenyl-5-methyl-1,3-dioxane Isomers

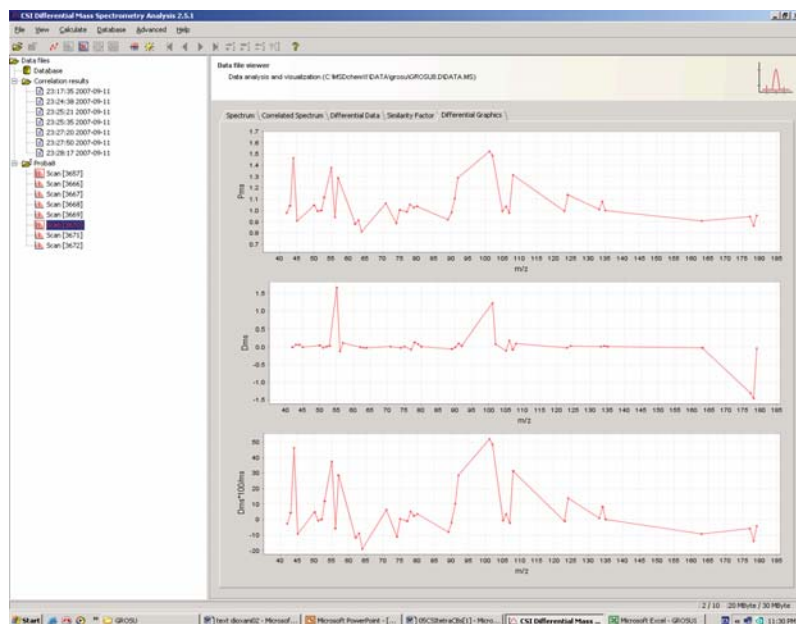| Probability	Isomer Z	Isomer T
	Retention time 24.99 min	Retention time 26.26 min
75%	*trans*-2-Phenyl-5-methyl-1,3-dioxane	*cis*-2-Phenyl-5-methyl-1,3-dioxane
25%	*cis*-2-Phenyl-5-methyl-1,3-dioxane	*trans*-2-Phenyl-5-methyl-1,3-dioxane

Figure 5. Graphic form of differential mass spectra of *cis-* and *trans-*2-phenyl-5-methyl-1,3-dioxane: the ratios of spectra (Pms), the difference of the spectra (Δms) and the percentage difference of the spectra (Δ%ms)
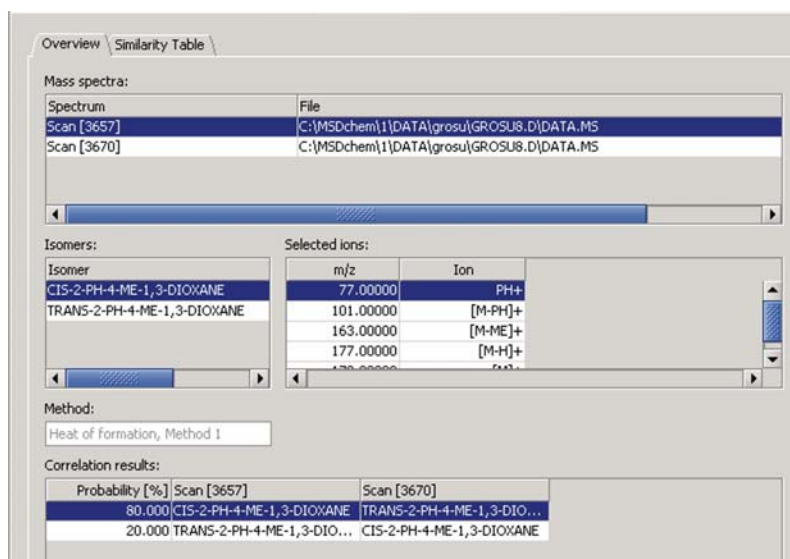

Figure 6. Snapshot of window with compared ms scans, isomers, ions and the list of structures probabilities for *cis-* and *trans-* 2-phenyl-5-methyl-1,3-dioxane isomers

Four out of five ions used in correlation for *cis*- and *trans*-2-phenyl-4-methyl-1,3-dioxane have intensities that are correlated with the heats of formation obtained from the AM1 semi empirical method. At the same time for *cis*- and *trans*-2-phenyl-5-methyl-1,3-dioxane the correlation is present for three of the four used ions. The structures found by this way for the X-Y and Z-T diastereoisomers pairs are in concordance with the studies carried out through NMR and X-ray analysis.

3.   Conclusions

The results obtained by CSI-Diff-ms on discrimination of these diastereoisomers of 2-phenyl-4-methyl-1,3-dioxane and 2-phenyl-5-methyl-1,3-dioxane presented above are in agreement with the results obtained by conformational analysis [6,7]. The CSI Diff ms–CQC analysis proved that under vacuum the molecules of the studied substances adopt conformations exhibiting the phenyl in axial orientation. The solvation and the energy of the crystalline network, which are involved in NMR and X-Ray analysis, modify this orientation. The 75–80% agreement between the experimental values Pi from the differ-ential spectrum Pms and the values estimated by calculation, ΔPi, in the CSI Diff ms program proves the ability of the semiempirical method AM1 to evaluate the heats of formation, to be coupled with Diff MS in order to describe the energy profile of the ionization processes from mass spectrometry.

References

1. I. Grosu, S. Mager, G. Ple, N. Ple, A. Toscano, E. Mesaros and R. Martinez, *Liebigs Annalen*, **1997**, 2371.
2. I. Grosu, S. Mager, L. Toupet, G. Ple, E. Mesaros and A. Mihis, *Acta Chem. Scand.*, **1998**, *52*, 366.
3. I. Grosu, S. Mager, E. Mesaros and G. Ple, *Heterocyclic Commun.*, **1998**, *4*, 53.
4. I. Grosu, S. Mager, G. Ple, E. Mesaros, A. Dulau and C. Gego, *Tetrahedron*, **1998**, *54*, 2905.
5. I. Grosu, M. Balogh, C. Paisz, G. Ple, F.D. Irimie, S. Mager and R. Podea, *Rev. Roum. Chim.*, **2000**, *45*, 877.
6. L. Muntean, M. Balog, C. Florian, A. Terec, I. Grosu , S. Mager, D.Margineanu, *Stud. Univ. Babes-Bolyai, Chemia,* **2002**, *47*, 195.
7. L. Muntean, E. Mesaros, G. Plé, I. Grosu and S. Mager, *Stud. Univ. Babes-Bolyai, Chemia,* **2000**, *45*, 47.
8. M.J.S. Dewar, G.E. Zoebisch, F.E. Healy, J.J.P. Stewart, *J. Amer. Chem. Soc.,* **1985**, *107*, 3902.
9. J.J.P. Stewart, *J. Comput. Aided Mol. Design,* **1990**, *4*, 1.HyperChemTM Release 5. Professional for Windows, Hypercube, Inc. **1999**, Gainesville, Fl 32601, USA.www.bet2-soft.de

13. MALDI-TOF MASS SPECTROMETRY IN TEXTILE INDUSTRY

FLORENTINA-DANIELA MUNTEANU[1], NICOLAE DINCA[1], AND ARTUR CAVACO-PAULO[2]

[1]*University "Aurel Vlaicu" Arad, Faculty of Food Engineering, Tourism and Environmental Protection, Department of Chemical and Biological Sciences, Elena Dragoi 2, 310330, Arad, Romania*
[2]*University of Minho, Department of Textile Engineering, 4800-058 Guimarães, Portugal*

Abstract. In this paper are presented the possibilities of using matrix assisted laser desorption ionization time of flight (MALDI-TOF) mass spectrometry in textile industry. MALDI-TOF mass spectrometry it is a convenient, versatile method for characterization and identification of dyes and pigments, and also for characterization of fibers and contaminants of the fabrics.

1. Introduction

Mass spectrometry is a technique that has been known since the early twentieth century. In 1913, based on the growing understanding of electromagnetic forces, J.J. Thomson was able to separate the two neon isotopes of mass 20 and 22 with his parabola mass spectrograph [1]. Mass spectrometry takes advantage of the force balance that is established when a charged species, in a gas phase, travels through an electric or magnetic field [1]. The trajectory of this species will depend on its mass-to-charge ratio (*m/z*). Since the early experiments of Thomson, who analyzed small gaseous atoms and compounds, strategies to vaporize and/or ionize large non-volatile molecules have been developed.

Matrix-assisted laser desorption/ionization (MALDI) and fast-atom bombardment (FAB) are some common desorption/ionization methods, whereas electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) are some examples of spraying techniques.

For the separation and to record and store the results, a mass spectrometer contains the following components:

(i) *The ion source*: where the analyte is brought into the gas phase and ionized.
(ii) *The mass filter*: necessary for the ions' separation on basis of their mass-to-charge ratios.
(iii) *The detector*: different *m/z* ions hit the detector and their impacts are turned into electrical signals.
(iv) *The system of data acquisition*: for the storage of all the electrical signals that are generated from the detector.

Matrix Assisted Laser Desorption/Ionisation–Time of Flight–Mass Spectrometry (MALDI-TOF-MS) was developed in the middle to late 1980s (Nobel Prize winner 2002) [2–5] for the analysis of molecules with large molar weights.

The first reports demonstrating successful MALDI-TOF-MS biochemical analysis were published in the late 1980s from the labs of Tanaka et al. [2] and Karas and Hillenkamp [5]. Instruments have subsequently become commercially available since 1991 [6]. This implies that the general utilization of MALDI-TOF-MS has emerged as an effective analytical tool to study biomolecules only within the last 15 years. Although relatively young compared to other analytical techniques using mass spectrometry, there has been an enormous increase in the publication of MALDI-TOF-MS methods and applications in the literature [7–9].

MALDI is an ionization technique that, as the name implies, makes use of a photon-absorbing matrix for the desorption and ionization of the analyte [10]. In preparation for a MALDI analysis the analyte is mixed with matrix and co-crystallized on an electrically conducting target, e.g., a stainless steel plate (see Figure 1).

In MALDI the sample is prepared by mixing matrix and analyte in solution at a ratio usually in the range from 100:1 to 50,000:1. The matrix is commonly a small aromatic organic solid or liquid with acidic and/or basic functional groups, and has often highly light-absorbing characteristics.
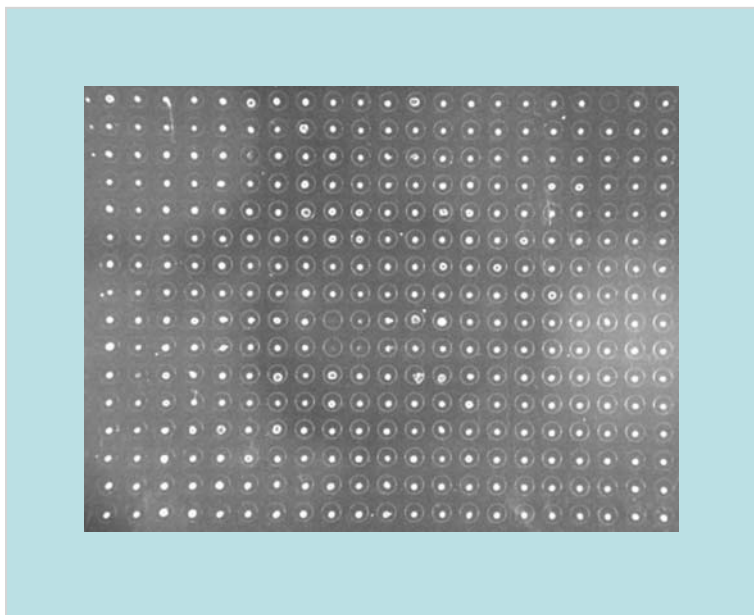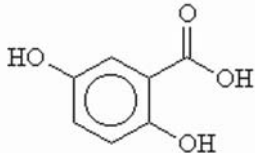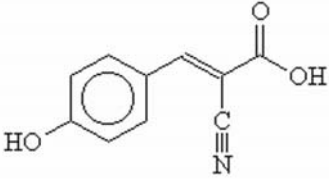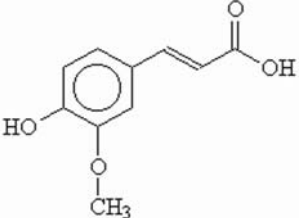

Figure 1. MALDI-TOF stainless steel target plate

The matrix is often dissolved in solvents such as water, ethanol, methanol, acetone, or mixtures of these [11,12].

The steel plate is then placed in a vacuum and energy in the form of a short laser pulse is applied to the sample and matrix crystal.

The purpose of adding the matrix molecules is that when they are exposed to a laser, usually a nitrogen laser at 337 nm, they become excited and their absorbed energy causes translational motion and ionization of the analyte molecule [10,13]. The matrix strongly reduces decomposition of the analyte by absorbing most of the laser pulse energy. As a consequence of this, an "explosion" at the surface causes vaporization of the matrix and the analyte molecules are thereby ejected into the gas phase under vacuum or atmospheric pressure. The ejected material contains both neutral and charged species that are often analyzed with a time-of-flight analyzer but can also be analyzed with other mass analyzers. It should be noted that the nature of the ionization mechanism is still being discussed [13,14].

TABLE 1. Common MALDI Matrices

Name	Molecular structure	Molecular formula	Average mass $[M+H]^+$
2,5-dihydroxy-benzoic acid (gentisic acid)		$C_7H_6O_4$	155.130
α-cyano-4-hydroxy-cinnamic acid		$C_{10}H_7NO_3$	190.178
4-hydroxy-3-methoxycinnamic acid (ferulic acid)		$C_{10}H_{10}O_4$	195.195

(Continued)

(Continued)

3,5-dimethoxy-4-hydroxy-cinnamic acid (sinapinic acid)		$C_{11}H_{12}O_5$	225.222
3-hydroxy-2-pyridine-carboxylic acid (3-hydroxy-picolinic acid)		$C_6H_5NO_3$	140.119
nicotinic acid-N-oxide		$C_6H_5NO_3$	140.119
2'-6'-dihydroxyaceto-phenone		$C_8H_8O_3$	153.158

The mixture of matrix and sample (normally 0.1–2 μl) is placed onto a target plate, where the solvent is allowed to evaporate, whereby crystals are formed. The matrix and sample may also be deposited in separate individual steps. The drying of the solvent can either be performed at ambient pressure or at reduced pressure. The method used can influence the formation of crystals to a significant extent, which can impact performance and ease of use [11,12]. It is important that the matrix is relatively stable and not evaporate under low pressure conditions. The mechanisms of the MALDI-ionization are not yet fully understood.

Many different approaches have been proposed and it is most likely a combination of several mechanisms that give rise to the ions observed in the spectrometer [10]. Some of the routes proposed are the photoionization, thermal ionization and excited-state proton transfer.

It has even been suggested that the ions are preformed and hence are already present in ionic form before the laser impact and the in-plume processes are initiated [10].

One great advantage with MALDI is that the most abundant species are usually single charged molecular ions such as $[M+H]^+$ (positive ions) or $[M–H]^-$ (negative ions), where M corresponds to the molar mass of the neutral molecule. The interpretation of the spectrum thereby becomes easier (no multiply charged species interfering). Doubly-charged species may occur but this is not as common as in other MS techniques such as ESI-MS.

As stated above, MALDI is often used together with a time-of-flight (TOF) mass analyzer (Figure 2).
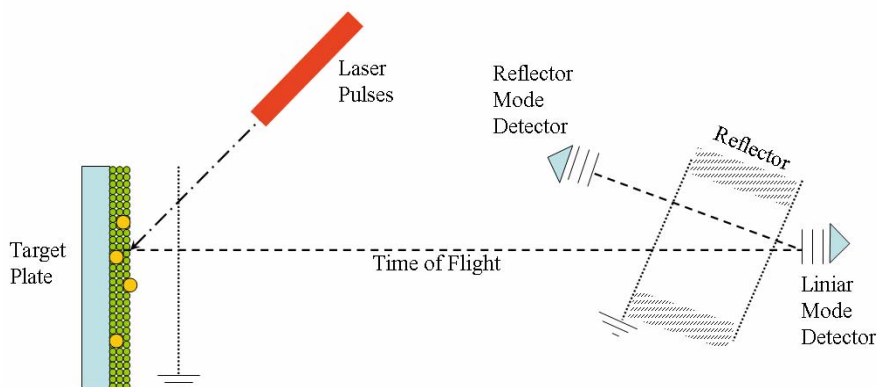

Figure 2. The principle for matrix-assisted laser desorption/ionization with time-of-flight separation

During the desorption process of a sample containing a mixture of several substances, there may be a discrimination against analytes that are less prone to be ionized. This phenomenon is referred to as ion suppression [15], and is often seen in the MALDI analysis of a mixture of peptides. If different peptides are mixed in equimolar amounts and analyzed, the intensities of the corresponding peaks will vary. Easily ionized peptides are known to suppress the ionization of peptides that are less prone to be ionized. This is often the phenomenon that limits the use of internal calibration procedures in MALDI-TOF-MS.

MALDI-TOF-MS has become a powerful tool in a variety of fields [16–21] because it enables molecular weight determination and structural characterization of macromolecules with the characteristics of speed, good mass accuracy, good resolution, excellent sensitivity and high throughput capability. Nevertheless, when searching for low abundant proteins the demands on high mass accuracy and low detection limits are extensive. A sensitive and accurate mass analysis is crucial for obtaining a well-defined peptide mass fingerprint containing as many peptide fragments, stemming from the original protein, as possible.

2.   Applications of MALDI-TOF-MS in Textile Industry

MALDI-TOF can be a useful tool to elucidate the structural details of bast fiber peels of developing flax (*Linum usitatissimum L.*) stem [22].

The polymers that contain galactose are the major matrix components in fiber cell walls of important crops, like flax (*Linum usitatissimum L.*). The percent of galactose containing polymers represents up 40–80% of noncellulosic polysaccharides [23,24]. Gur'janov et al. [22] were using an Ultraflex instrument (Bruker Daltonics) equipped with a nitrogen laser of 337 nm. The experiments were done in positive mode and after a delayed extraction time of 200 ns, the ions were accelerated to a kinetic energy of 12,000 V. Thereafter, the ions were detected using the reflector mode. For the preparation of the matrix solution were used 10 mg 2,5-dihydroxybenzoic acid (Bruker Daltonic's) in 700 µl water and 300 µl acetonitrile. The samples (approximately 20–40 µg sugar) were dissolved in 10 µl water. After desalting the sample, 1 µl of sample solution was placed on a MALDI-TOF-plate together with 1 µl of matrix solution and allowed to dry under a constant stream of warm air.

The authors show that the flax galactan is a complex polysaccharide with variable side chain structures. Moreover, the backbone is composed of the common galactan linked to rhamnogalacturonan repeats with a high degree of branching.

Moreover, treatment with purified and well-characterized galactanase does not change the hydrodynamic volume of flax galactan, suggesting a complex "secondary" structure of the polymer.

In 2000, Holland et al. used MALDI-TOF to obtain mass spectra of bacterial proteins from cotton cloth samples contaminated with *Shigella flexneri, Escherichia coli,* and *Aeromonas hydrophila* [25].

For this purpose a saturated matrix solution was prepared by dissolving α-cyano-4-hydroxycinnamic acid (HCCA) in 2:1 water:acetonitrile (0.1% trifluoroacetic acid) solution. The experiments were done in positive mode, the acceleration voltage was set at 28 kV and the instrument's laser (Nd-YAG) output frequency was tripled (355 nm) for sample desorption.

The authors came to the conclusion that MALDI-TOF mass spectrometry is a rapid method with high potential for detection of biomarker proteins recovered directly from clothing samples contaminated with *E. coli*, *Shigella flexneri*, and *Aeromonas hydrophila*. All three of these species could cause illness through the fecal-oral routes.

Polyamide (nylon) represents a family of synthetic polymers first produced in 1935 by Gerard J. Berchet of Wallace Carother's research group at DuPont.

The first product was a nylon-bristled toothbrush (1938), followed more famously by women's nylon's stockings (1940). Nylon 6.6 was the first commercially successful polymer and the first synthetic fibre to be made entirely from coal, water and air. It was intended to be a synthetic replacement for silk and substituted for it in parachutes after the USA entered World War II in 1941, making stockings hard to find until the war's end.

Commonly, commercial nylons are manufactured using processes related to either of three basic approaches, namely:

(i) Poly condensation of diamines and dibasic acids, as exemplified by hexa-methylene diamine and adipic acid (nylon 6.6)
(ii) Poly condensation of ω-amino acids, as typified by 11-aminoundecanoic acid (nylon 11)
(iii) Ring-opening polymerisation of lactams, such as ε-caprolactam in nylon 6

The structure of polyamide fibres is defined by both chemical and physical parameters. The chemical ones are related mainly to the constitution of the poly-amide molecule and are concerned primarily with its monomeric units, end-groups and molecular weight. The physical parameters are related essentially to chain conformation, orientation of both polymer molecule segments and aggregates and to crystallinity.

In a review by Montaudo et al. are presenting the fundamentals and practice in MALDI of polymers (such as the laser, ion source, ion optics, reflectron, detector, ionization efficiency) as well as to some basic concepts of sample preparation (such as the MALDI matrix and cationization agents). Then, the authors are focus on the measurable quantities of polymers: average molar masses, the chemical formula and the structure of the monomer (actually of the repeat unit), the masses of the chain end groups, etc [26].

Simplicity of sample preparation, rapid spectrum acquisition, high sensitivity, and relative tolerance to impurities make MALDI-TOF attractive for analysis of small molecule analytes. Soltzberg and colab [27] have used this advantages of MALDI-TOF and used it for the identification of dyes and pigments with mole-cular masses between 200 and 1,100 Da.

For the experiments was used a Bruker Omniflex benchtop MALDI-TOF mass spectrometer equipped with a 337 nm nitrogen laser. Pulsed ion extraction was employed, with an extraction delay of 300 ns. The experimentation was done in positive and negative mode for each dye and in reflectron mode at the same laser power. The accelerating potential was 14.8 kV for positive ion spectra and 15.8 kV for negative ion spectra. Each spectrum was the average of 50 laser shots. The reason for which the experiments were run in both negative and positive mode was to quickly classify the dyes as basic or acid based on the comparison of the intensity of negative and positive ion spectra for a particular dye. The base peak data for each dye are listed with negative- or positive-ion spectra, depending on which instrument polarity shows the strongest signal (Table 2).

The results presented in this research paper [27] give information about an unknown dye sample and represents an important aid in identification.

The results obtained by this group show that MALDI-TOF mass spectrometry affords a rapid, versatile, and reliable method for identifying dyes from samples with femtomolar concentrations. Sample preparation involves minimal handling and, therefore, minimal sample loss. The comparison of positive and negative ion

spectra provides structural information about the dye or pigment class and, in most cases, the identity of the substance. In favorable cases, the spectra can distinguish between isomers.

TABLE 2. Dyes Showing Negative and Positive Ions Mass Spectra

Dye	Formula
Negative ions mass spectra	
Picric acid	$C_6H_3N_3O_7$
Acid yellow 24	$C_{10}H_5N_2O_5Na$
Acid yellow 1	$C_{10}H_4N_2SO_8Na_2$
Acid yellow 36	$C_{18}H_{14}N_3SO_3Na$
Acid orange 5	$C_{18}H_{14}N_3SO_3Na$
Acid red 74	$C_{16}H_{11}N_4SO_5Na$
Mordant yellow 1	$C_{13}H_8N_3O_5Na$
Acid orange 6	$C_{12}H_9N_2SO_5Na$
Acid orange 20	$C_{16}H_{11}N_2SO_4Na$
Acid brown 6	$C_{20}H_{13}N_2SO_4Na$
Acid orange 7	$C_{16}H_{11}N_2SO_4Na$
Acid red 88	$C_{20}H_{13}N_2SO_4Na$
Acid red 9	$C_{20}H_{13}N_2SO_4Na$
Acid orange 12	$C_{16}H_{11}N_2SO_4Na$
Acid red 13	$C_{20}H_{12}N_2S_2O_7Na_2$
Acid red 25	$C_{20}H_{12}N_2S_2O_7Na_2$
Acid orange 14	$C_{16}H_{10}N_2S_2O_7Na_2$
Acid red 26	$C_{18}H_{14}N_2S_2O_7Na_2$
Acid red 17	$C_{20}H_{12}N_2S_2O_7Na_2$
Acid red 27	$C_{20}H_{11}N_2S_3O_{10}Na_3$
Acid orange 10	$C_{16}H_{10}N_2S_2O_7Na_2$
Acid red 44	$C_{20}H_{12}N_2S_2O_7Na_2$
Acid red 18	$C_{20}H_{11}N_2S_3O_{10}Na_3$
Acid red 41	$C_{20}H_{10}N_2S_4O_{13}Na_4$
Acid red 33	$C_{16}H_{11}N_3S_2O_7Na_2$
Acid red 1	$C_{18}H_{13}N_3S_2O_8Na_2$
Acid violet 7	$C_{20}H_{16}N_4S_2O_9Na_2$
Acid yellow 23	$C_{16}H_9N_4S_2O_9Na_3$
Acid black 1	$C_{22}H_{14}N_6S_2O_9Na_2$
Direct red 28	$C_{32}H_{22}N_6S_2O_6Na_2$
Direct yellow 4	$C_{26}H_{18}N_4S_2O_8Na_2$
Acid red 151	$C_{22}H_{15}N_4SO_4Na$
Acid red 115	$C_{24}H_{18}N_4S_2O_7Na_2$
Acid red 73	$C_{22}H_{14}N_4S_2O_7Na_2$
Acid blue 3	$C_{27}H_{31}N_2S_2O_7Na$

Acid green 5	$C_{37}H_{34}N_2S_3O_9Na_2$
Acid yellow 73	$C_{20}H_{12}O_5$
Acid red 87	$C_{20}H_6O_5Br_4Na_2$
Acid red 51	$C_{20}H_6O_5I_4Na_2$
Acid yellow 3	$C_{18}H_8NS_3O_{11}Na_3$
Murexide	$C_8H_8N_6O_6$
Mordant red 11	$C_{14}H_8O_4$
Alizarin red PS	$C_{14}H_8SO_8$
Mordant red 2	$C_{14}H_7SO_8Na$
Acid blue 83a	$C_{45}H_{44}N_3S_2O_7Na$

Positive ions mass spectra

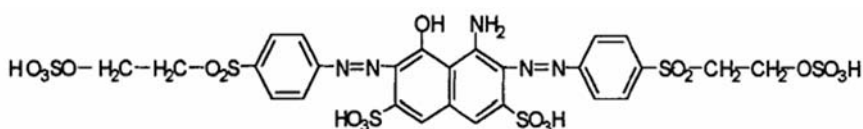|	
Basic orange 2	$C_{12}H_{13}N_4Cl$
Basic brown 4	$C_{21}H_{26}N_8Cl_2$
Basic yellow 2	$C_{17}H_{22}N_3Cl$
Basic green 4	$C_{23}H_{27}N_2OCl$
Basic green 1	$C_{27}H_{33}N_2$
Basic violet 14	$C_{20}H_{20}N_3Cl$
Basic violet 3	$C_{25}H_{30}N_3Cl$
Basic blue 11	$C_{29}H_{32}N_3Cl$
Basic red 1	$C_{28}H_{31}N_2O_3Cl$
Solvent red 49	$C_{28}H_{31}N_2O_3Cl$
Basic red 2	$C_{20}H_{19}N_4Cl$
Basic blue 9	$C_{16}H_{18}N_3SCl$
Basic blue 26	$C_{33}H_{32}N_3Cl$

Figure 3. Chemical structure of the reactive dye Ostazine Black V-B

Reactive dyes are a class of colorants and are designed to form a covalent bond with the substrate of application, in contrast to other classes, where the coloration is dependent on physical adsorption or mechanical retention. The capability of forming chemical bonds is based on the dualistic structure of the molecule which is composed of a colored part (chromophore) and a reactive system. The colored part consists of a chromogenic system with groups that encourage solubility.

Chroma-Keull and coworkers [28] used MALDI-TOF-MS for the determination of reactive dyes and monitoring of their hydrolysis. As a model for the class of reactive dyes was chosen Ostazine Black V-B (see Figure 3).

In their experiments was used a Kratos Kompact MALDI III mass spectrometer (Manchester, UK) controlled by Kratos Kompact V5.2.0 software, equipped with a nitrogen laser VSL-337D-10-TTL from Laser Science Inc. (Franklin, MA, USA) operating at 337 nm with pulse energy 250 mJ, average power at 10 Hz 2.5 mW and peak power 85 kW.

After the optimization of the working conditions it was found that the reflectron negative mode is optimal, because under any conditions there were no signals obtained in positive mode for the substances studied. A laser power of 100 relative power units was used (maximum 180) with 2- (4-hydroxyphenylazo) benzoic acid as the matrix. The matrix was diluted in acetonitrile/0.5% ammonium citrate 1:1 (v/v). The ammonium citrate was found to not only suppress a number of sodium adducts, but also enhance the homogeneity of the crystals of dried sample.

As a result of these experiments the authors concluded that Using MALDI-TOFMS in reflectron negative mode, given sample of Ostazine Black V-B and its derivatives were determined with the LOD approximately 100 mg/l, and the method is suitable for immediate estimation of the dye-bath content during the dyeing process. Moreover using MALDI-TOF-MS it was possible to run 60 samples per h, while using HPLC it is possible to carry out just four experiments per h.

In conclusion, MALDI-TOF-MS it is a useful tool in textile industry. Its applications can be found again for characterization of textile fibers, identification of toxic compounds in cloth, but also for identification and characterization of textile dyes.

References

1. J. Roboz, Introduction to Mass Spectrometry Instrumentation and Techniques, Wiley, New York, **1968**.
2. K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, and T. Yohida, *Rapid Commun. Mass Spectrom.*, **1988,** *2*, 151.
3. M. Karas, D. Bachmann, U. Bahr, and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Proc.*, **1987,** *78*, 53.
4. M. Karas, D. Bachmann, and F. Hillenkamp, *Anal. Chem.*, **1985,** *57*, 2935.
5. M. Karas and F. Hillenkamp, *Anal. Chem.*, **1988,** *60*, 2299.
6. J. T. Stults, *Curr. Opin. Struct. Biol.*, **1995,** *5*, 691–8.
7. D. J. Harvey, *J. Chromatogr. A*, **1996,** *720*, 429–46.
8. T. Bonk and A. Humeny, *Neuroscientist*, **2001,** *7*, 6–12.
9. E. J. Zaluzec, D. A. Gage, and J. T. Watson, *Protein Expr. Purif*, **1995,** *6*, 109–23.
10. R. Zenobi and R. Knochenmuss, *Mass Spectrom. Rev.*, **1999,** *17*, 337.
11. D. Momcilovic, B. Wittgren, K.-G. Wahlund, J. Karlsson, and G. Brinkmalm, *Rapid Commun. Mass Spectrom.*, **2003,** *17*, 1116.
12. D. Momcilovic, B. Wittgren, K.-G. Wahlund, J. Karlsson, and G. Brinkmalm, *Rapid Commun. Mass Spectrom.*, **2003,** *17*, 1107.
13. M. Karas, M. Gluckmann, and J. Schafer, *J. Mass Spectrom.*, **2000,** *35*, 1.
14. D. J. Harvey, *Mass Spectrom. Rev.*, **1999,** *18*, 349.
15. R. Knochenmuss, A. Stortelder, K. Breuker, and R. Zenobi, *J. Mass Spectrom.*, **2000,** *35*, 1237.

16. S. J. Wetzel, C. M. Guttman, and J. E. Girard, *Int. J. Mass Spectrom.*, **2004,** *238*, 215.

17. J. Schiller, R. Suss, J. Arnhold, B. Fuchs, J. Lessig, M. Muller, P. M., H. Spalteholz, O. Zschornig, and K. Arnold, *Prog. Lipid Res.*, **2004,** *43*, 449.

18. Y. Q. Shen, H. D. Tang, and S. J. Ding, *Prog. Polym. Sci.*, **2004,** *29*, 1053.

19. L. H. Cohen and A. I. Gusev, *Anal Bioanal. Chem.*, **2002,** *373*, 571.

20. S. F. Macha and P. A. Limbach, *Curr. Opin. Solid State Mat. Sci.*, **2002,** *6*, 213.

21. P. Jungblut and B. Thiede, *Mass Spectrom. Rev.*, **1997,** *16*, 145.

22. 2O. P. Gur'janov, T. A. Gorshkova, M. Kabel, H. A. Schols, and J. E. G. van Dam, *Carbohydr. Polym.*, **2007,** *67*, 86.

23. R. Girault, F. Bert, C. Rihouey, A. Jauneau, C. Morvan, and M. Jarvis, *Int. J. Biol. Macromol.*, **1997,** *21*, 179–188.

24. C. Mooney, T. Stolle-Smits, H. Schols, and E. de Jong, *J. Biotechnol.*, **2001,** *89*, 205–216.

25. R. D. Holland, F. Raffi, T. M. Heinze, J. B. Sutherland, K. J. Voorhees, and J. O. Lay Jr., *Rapid Commun. Mass Spectrom.*, **2000,** *14*, 911.

26. G. Montaudo, F. Samperi, and M. S. Montaudo, *Prog. Polym. Sci.*, **2006,** *31*, 277–357.

27. L. J. Soltzberg, A. Hagar, S. Kridaratikorn, A. Mattson, and R. Newman, *J. Am. Soc. Mass Spectrom.*, **2007,** *18*, 2001–2006.

28. H. Chroma-Keull, J. Havlis, and J. Havel, *Rapid Commun. Mass Spectrom.*, **2000,** *14*, 40–43.

14. MALDI MS IN ANALYSIS OF KERATIN FIBRE PROTEINS

ANDREA KÖRNER
DWI an der RWTH Aachen, Germany

Abstract. The high thermal stability and shape memory properties of keratin fibres has been always a matter of interest for protein chemists willing to understand how the structure can withstand 200°C without any significant loss of properties. Partly these properties are explained by the presence of cystine as a cross-linking amino-acid, but to some extent they may be also due to the secondary and tertiary structure of the proteins inside. This is indicated by the strong variations in the mechanical behaviour of different keratin fibres, with similar chemical composition and geometrical features (length and diameter). There is a large range of fibres from soft, like those of cashmere, or yak, to stiff, like wool fibres of high diameter. Similar variations are met on samples of human hair from various people. Understanding the secondary and tertiary structure of keratin fibres is, therefore, an important matter for textile and cosmetic industry, both for scientific interest and more applied purposes. New mass spectrometric analysis techniques such as MALDI MS or ESI MS promise further inside into the variations of keratins. Several characteristics of keratins make them a challenge for mass spectrometric characterization and are the reasons why they have been less well studied than other protein families. These are associated with their solubility, recovery from their keratinized substrates, high degree of sequence homology and the presence of numerous post-translational modifications. Recent approaches to apply MALDI MS focus on monitoring changes in the keratin composition due to chemical influences, the characteristic MALDI MS peptide mass fingerprint as a tool for fibre species identification and on the identification of proteins by matching mass spectrometric data with known sequences presented in protein data bases. The purpose of this paper is to summarise past work in this area.

1. Introduction on Keratin Fibres

Keratin fibres like wool, cashmere, angora, mohair, alpaca, human hair etc. consist of two main cell types, the external layer of flat, overlapping cuticle cells and the spindle-shaped cells of the fibre core, the cortex cells. The cells of cuticle and cortex are separated from each other by the cell membrane complex (CMC). An overview on the morphology of a wool fibre according to Fraser et al. [1] providing an example for keratin fibres is given in Figure 1.

The cortical cells are mainly composed of macrofibrils, which consist of a composite network of 500–800 highly organized microfibrils, also addressed as keratin intermediate filaments (KIF). The α-helical keratin intermediate filament proteins (KIF or KRT), or low-sulfur proteins with respect to their comparatively low amount of cystine, are embedded into a matrix of amorphous interfilament material from keratin associated proteins (KAP or KRTAP).

The KAP/KRTAP comprise the protein families of type 1–10, rich in cystine (HS proteins), the ultrahigh sulphur proteins and the high glycine/high tyrosine
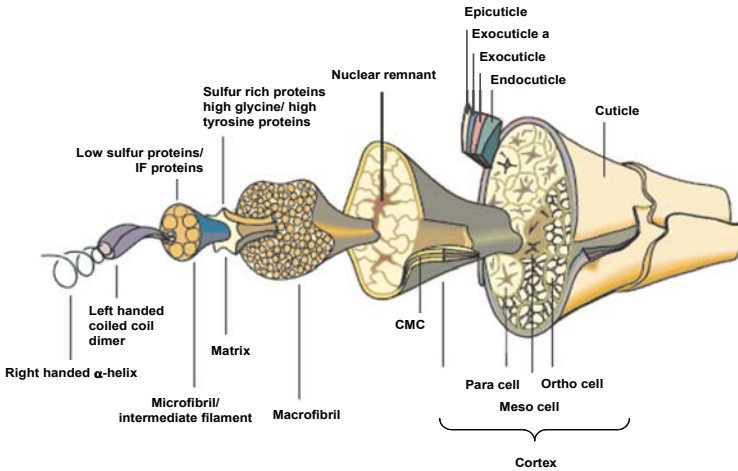
Figure 1. Schematic presentation of the morphology of keratin fibres for the example of a fine wool fibre according to Fraser et al. [1] (Reproduced with the permission of CSIRO Textile and Fibre Technology)

proteins (HGT proteins). There are, however, also proteins high in sulphur and glycine.

The keratin intermediate filament proteins in wool, human hair and other keratin fibres are made from about 400 to 500 amino acid residues and their molecular weights lie in the range between 40,000 to 58,000 Da. The monomers consist of a globular head region, a helical rod domain and a globular tail domain; the rod domain itself is composed of helical segments, named segment Coil 1A, 1B and Coil 2A and 2B, which are separated from each other by globular linkers. There are two groups of keratins, the acid type I keratins with isoelectric points between pH 4 and 6, and the neutral to basic type II keratins with isoelectric points between pH 6 and 8.

By parallel packing of two strands of keratin molecules, one of the type-I- and one of the type-II- keratin family, the heterodimer or coiled coil is formed. Two heterodimers form the tetrameric protofilament, which then aggregates to the octameric protofibril. Further aggregation leads to protofibrils with 4.5 nm diameter, two of which build the 10 nm filament [2].

The two strands of the coiled coil dimer are attached to each other by hydrophobic interactions of the non polar amino acid side chains and interhelical ionic interactions of cationic and anionic side chains. Disulfide bonds between the 2B segments in the helical parts of adjacent dimers provide the stability of the tetramer [3]. A recent review of Plowman [4] on the proteomic database of wool components gives an excellent detailed overview on the diversity of keratin fibre proteins.

The analysis of the α-helical keratin intermediate filaments (KIF) and the keratin associated HS and high glycine/high tyrosine proteins by methods such as polyacrylamide gel electrophoresis, liquid chromatography and mass spectrometric techniques like MALDI-MS and ESI-MS depends on the dissolution of the proteins and extraction from the fibre. The prerequisite for the dissolution of keratin proteins, usually in solutions with high urea concentrations, is the cleavage of the intermolecular and intramolecular disulfide bonds. There are three main approaches for this: (1) reductive cleavage, typically with mercaptoethanol, dithiothreitol or other reducing agents and subsequent alkylation to prevent reoxidation [5], (2) oxidative sulphitolysis with sodium sulphite in 8 M urea to form the so called "Bunte salts" [5], and (3) oxidative cleavage e.g., with peracetic acid and dissolution in ammonium hydroxide solution [6].

2. MALDI MS or Related Techniques in the Analysis of Keratin Fibre Proteins

Only a few publications deal with the analysis of KIFs and KAPs of keratin fibres by mass spectrometric techniques of MALDI MS or ESI MS and quite a few of them will refer to this topic in the tenor of Chamrad et al [7] "Typical sample contaminants observed are Keratin proteins from hair, nails, and skin…. ." Furthermore, there are several characteristics of keratins which make them a challenge for mass spectrometric characterization and are the reason why they have been less well studied than other protein families. These are, as Plowman [8] listed them recently, difficulties associated with their extraction from biological samples, high degree of sequence homology and the presence of numerous post-translational modifications.

There are some recent approaches, however, to apply MALDI MS for the characterization of keratins with three main aims:

(i) To monitor **changes in the keratin composition due to chemical influences on a molecular basis**
(ii) To profit from the characteristic MALDI MS peptide mass fingerprint as a tool for **fibre species identification**, and
(iii) To apply MALDI MS/ESI MS as a tool for the **identification of proteins** by matching mass spectrometric data with known sequences presented in protein data bases

I. Reports on the use of **MALDI MS to monitor changes in the keratin pattern** due to chemical impacts on the fibre have been published by two teams [9,10]. In their work on the effect of alkaline oxidative dehairing formulations on the degradation of keratins from wool and bovine hair, Marmer et al. [9] applied MALDI ToF/ToF spectrometry to determine the progress and extent of cysteine oxidation in the solubilised keratins and showed that some of the cysteine groups in the hair shaft remained unoxidised.

In a recent study Körner et al. [10] applied kinetically controlled reverse pro-teolysis aiming for an enrichment in lysine (2,6-diaminohexanoic acid) with proteases such as trypsin and papain as catalysts. They determined the molecular weight pattern of chymotryptic peptides obtained from the α-keratose fraction of wool before and after reverse proteolysis with lysine ethylester by means of MALDI-ToF-MS. The water soluble α-keratose peptide hydrolysate was prepared by enzymatic degradation with α-chymotrypsin from a wool keratin intermediate filament (KIF) fraction, which had been solubilised after oxidation of the disulfide bonds with formic acid/hydrogen peroxide in the form of α-keratoses according to Alexander and Smith [6].

Figure 2 shows the MALDI-ToF-MS of the chymotryptic peptides of wool α-keratose proteins. The desired decrease in the molecular weight of the KIF proteins from molecular weights between 39 up to 61 kDa (obtained by one-dimensional electrophoresis) due to incubation with α-chymotrypsin is obvious. As expected from virtual digestion of known KIF sequences with the peptide cutter tools of the ExPASy Proteomics Server, peptides ranging from $m/z = 500$ to $m/z = 3,700$ are present in the hydrolysate, with the major part of peptides within a range of 800–2,500 Da.

The success of the incorporation of additional lysine was controlled by monitoring shifts in the molecular weight distribution of the peptides with MALDI-ToF-MS and by determination of the amino acid composition after acid hydrolysis with GC-MS. With both enzymes a shift to higher peptide masses was observed. The number of additionally bound lysine molecules in the modified peptides was calculated from the MALDI-ToF-MS data to be in the range of 2 to 7 lysine units for both trypsin and papain catalyzed reactions.
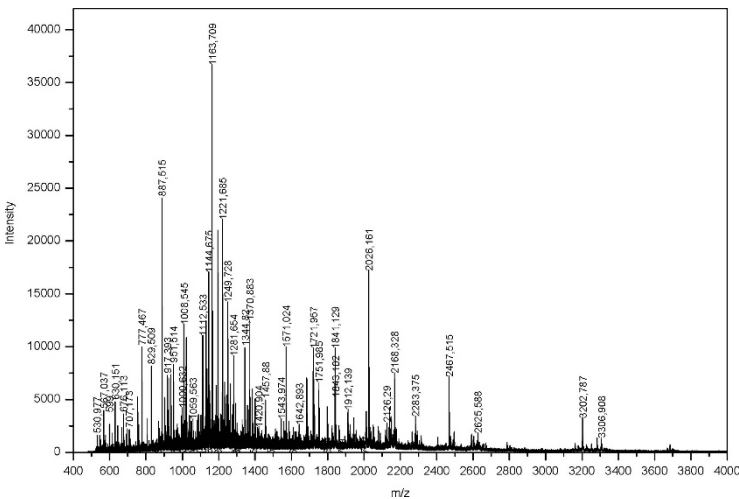

Figure 2. MALDI-ToF-MS of a chymotryptic peptide map from wool α-keratose proteins

II. In 2002, Hollemeyer et al. reported on the use of **MALDI-ToF-MS peptide mass fingerprinting** for the identification and quantification of feathers, down, and hair of avian and mammalian origin [11]. This technique also suggests an economically appalling path for discriminating cashmere goat fibres from those of yak in textile blends by their protein pattern and thus provides an interesting extension of the scanning electron microscopic reference procedure ISO/DIS 17751 [12]. Hollemeyer et al. [13] later extended their hair analysis technique to the identification of furs of domestic dog, raccoon dog, rabbit and domestic cat and established a method for determining genus, species, race and/or geographical origin of biological materials[14]. For this purpose the fibrillary proteins are first reduced with mercaptoethanol and then subjected to enzyme digestion with trypsin. Without further separating and isolating techniques the resulting peptides are subjected to MALDI-TOF mass spectroscopy; selected specific mass peaks from the mass spectrograms are used for identification of feathers, down, scales horn and/or hair and for quantification in blends by aid of reference mass spectrograms.

III. MALDI MS/ESI MS as a tool for the **identification of proteins** by matching mass spectrometric data with known sequences presented in protein data bases has been applied by several teams.

Yokohama et al. [15] made a mass spectrometric approach with the aim of proteome analysis of cashmere. For this purpose, the dissolved fiber proteins were separated by two-dimensional gel electrophoresis, eluted and then their molecular masses determined by mass spectrometry. By matching the molecular masses to those from known theoretical sequences of KIFs in web-based databases the major part of the isolated components was assigned to keratin type I and type II microfibrillar proteins of sheep, owing to the high degree of sequence homology within the KIFs. One component with a molecular weight of less than 17 kDa was identified as the keratin associated protein high-sulfur matrix protein IIIB2, also named KRA31_CAPHI (Capra hircus, Goat, accession no. P02447 in UniProtKB/Swiss-Prot), the sequence of which had been determined by Parris and Swart from reduced mohair [16].

Bryson and coworkers [17–20] performed detailed studies on the proteomics of wool keratins applying two-dimensional electrophoretic (2-DE), peptide mass spectral fingerprinting, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry, high-resolution nanoLC-MS/MS and nanoelectrospray on a quadrupole time-of-flight instrument. By means of mass spectrometry they studied the speed of cysteine alkylation, necessary to prevent reoxidation to disulphide links, and found that this process was rather fast, with 70–95% of the cysteines in the keratin proteins being alkylated within the first 10 min. In the case of the HGTPs, however, there was evidence for noncysteine alkylation occurring within this period, showing that long alkylation times do not suit all protein types or classes [18].

By combining two-dimensional electrophoresis (2DE) and peptide mass spectrometry fingerprinting Bryson and coworkers were able to match tryptic peptide digests from distinct KIF protein spots obtained by 2DE to the known theoretical sequences of some KIfs in web-based databases. This was possible, since tryptic digestion generated sufficient numbers of peptides from each spot to achieve this; the type I KIFs K1M1 and K1M2 as well as the type II KIFs K2M2 and K2M3 were identified despite 77% sequence identity [17].

The adoption of this protocol for the analysis of the keratin associated matrix proteins (KAP or KRTAP) proved to be more difficult. Due to their lower molecular weight and the absence of basic residues in the first two-thirds of the sequence the cystine rich HS proteins and the high glycine/high tyrosine proteins (HGT proteins) generate only a small number of peptides on tryptic digestion. Combined with their high sequence homology generally only a few of these peptides could be considered to be unique identifiers for each protein. Furthermore, particularly the HSPs turned out to be difficult to extract from the gels, even after tryptic digestion. In combination with the effect of a small number of high molecular weight peptides, this results in poor detectability of HS and HGT proteins by MALDI-TOF MS. Thus only a very limited number of high molecular weight peptides, present in low concentration only, were available for the peptide mass fingerprinting approach. For a tryptic digest of high sulphur proteins from wool to the KAP 1 family, Bryson [17] and Plowman et al. [8] obtained mass spectra with two dominating peaks at m/z of 1,049.2 and 1,274.4 in the MALDI-ToF-MS corresponding to two peptides common to all family members. The family consists of 4 different proteins, KRTAP 1.1, KRTAP 1.2, KRTAP 1.3, and KRTAP 1.4, which differ in the number of repeats of the decapeptide QTSCCQPT(I)SI. This renders the identification of distinct proteins within this family difficult.

With regard to HGT proteins, there are two known families: type I and type II consisting of 10 members with two types of peptide repeats [21]. The heterogeneous type I subfamily [22] consists of two major groups: a single protein in group F (KAP 7), for which both gene and amino acid sequences are available [23,24] and several components in group C. Both gene and amino acid sequences for component C2 or KR62 (KAP8) are available, though amino acid analysis for component C3 has so far shown it to be identical to C2 [22,24]. At least nine genes have been located for the type II subfamily [25], for which one full gene sequence (KAP6.1) [26] and one partial amino acid sequence (KAP6.2) [27] are available.

The mass spectra for the tryptic peptides of four HGT proteins excised from the 2DE wool protein map all produced a single, strong peptide peak at 1,700.27 Da. This mass matched to that of the 15-residue peptide R35-R49 from the KRHA KAP6.1 protein in SWISS-PROT, but not to Type I F (KAP 7) or KR62 (KAP 8). Bryson et al. [17] pointed out, however, that this peptide might theoretically also be derived from the KAP6.2 protein.

Bryson et al. [17] conclude that MALDI-ToF mass spectroscopy has apparent limitations in the application to tryptic peptides obtained from low molecular

weight proteins and suggest the use of further techniques such as post-source decay in MALDI-TOF or collision induced dissociation in electrospray tandem mass spectrometry to obtain sequence information on the unidentified peptides from these wool proteins.

Similar to intact KIF proteins a large part of functional proteins, such as membrane proteins, in their native forms show insufficient solubility in the matrix solution due to the absence of strong denaturing agents. Chaotropic agents, such as 8 M urea, do not absorb typical desorption laser wavelengths and impede transient laser heating of the embedded analytes in the conventional MALDI process. This limits the conventional MALDI–ToF-MS analysis. Aiming to overcome the difficulty, Naito et al. [28] have developed a novel MALDI technique (UV/ FEL-MALDI). The technique is based on simultaneous irradiation of a nitrogen laser (UV) and a free electron laser (FEL) for performing a selective ionization, improving sensitivity and extending the upper limit of detectable molecular weight. By this method, multimers of keratin molecules were observed up to clusters of approximately 900 kDa. These results suggest that UV/FEL-MALDI can be a more effective method for the functional analysis of macromolecules and supramolecules of proteins than conventional MALDI- MS techniques [29,30].

References

1. R.D.B. Fraser, J.M. Gillespie, T.P. MacRae, R.C. Marshall, *Internat report of CSIRO, Division of Protein Chemistry,* Parkville, Australia, **1981**.
2. H. Zahn, F.J. Wortmann, H. Höcker, Chemie in unserer Zeit, **1997**, *31. Jahrgang Nr. 6*, 280.
3. L.G. Sparrow, L.M. Dowling, V.Y. Loke, P.M. Strike, "Amino Acid Sequences of Wool Keratin IF Proteins" in "The Biology of Wool and Hair", Ed. G.E. Rogers, P.J. Reis, K.A. Ward and R.C. Marshall, Chapman & Hall, London, **1988**, 145.
4. J.E. Plowman, *J. Chromatogr. B*, **2003**, *787*, 63.
5. W.G. Crewther, *Proc. Int. Wool Tex. Res. Conf. Aachen*, **1975**, *I*, 1.
6. P. Alexander, L. Smith, *Proc. Int. Wool Text. Res. Conf. Aust.* **1955**, B-58.
7. D.C. Chamrad, G. Koerting, J. Gobom, H. Thiele, J. Klose, H.E. Meyer, M. Blueggel, *Anal. Bioanal. Chem.* **2003**, *376*, 1014.
8. J.E. Plowman, The proteomics of keratin proteins, *J. Chromatogr. B*, **2007** *849*(1–2), 181.
9. W.N. Marmer, R.L. Dudley, Journal of American Leather Chemists Association, **2006**, *101*(11), 408.
10. A. Körner, J. Thill, H. Keul, M. Möller, Proc. 1st Aachen-Dresden International Textile Conference, Aachen, 29–30. November 2007; ISSN 0942-301X; P50, 1–9.
11. K. Hollemeyer, W. Altmeyer, E. Heinzle, *Anal. Chem.* **2002,** *74*(23), 5960.
12. ISO/DIS 17751 Textiles–Quantitative analysis of animals fibres by microscopy–Cashmere, sheep's wool, speciality fibres and their blends, **2003**.
13. K. Hollemeyer, W. Altmeyer, E. Heinzle, *Spectroscopy Europe*, **2007,** *19*(2).
14. W. Altmeyer, H. Ewen, E. Heinzle, K. Hollemeyer, Method for qualitative and/or quantitative determination of gender, species, race and/or geographical origin of biological materials, **2002**, European Patent EP1388007, WO02093166 (A1); WO02093166 (A1); EP1388007 (A1); EP1388007 (A1); US2004142383 (A1); EP1388007 (A0); DE10123711 (A1).

15. M. Yokohama, T. Masuda, T. Amano, H. Hirayama, T. Manabe, *J. Anim. Sci.* **2004**, *75*(5), 401.
16. D. Parris, L.S. Swart, *Biochem. J.* **1975**, *145*, 459.
17. W.G. Bryson, J.E. Plowman, L.M. Flanagan, T.W. Jordan, *Wool Tech. Sheep Breed,* **2001**, *49*(4), 246.
18. J.E. Plowman, L.M. Flanagan, L.N. Paton, A.C. Fitzgerald, N.I. Joyce, W.G. Bryson, Proteomics, **2003**, *Jun. 3*(6), 942.
19. J.E. Plowman, Warren G. Bryson, Leanne M. Flanagan, T. William Jordan, Analytical Biochemistry, *300*(2), **2002**, 221.
20. St. Clerens, J.E. Plowman, 21st Australian and New Zealand Society for Mass Spectrometry (ANZSMS) Conference, University of Canterbury, Christchurch N.Z., January **2007**, Poster MoP-01.
21. B.C. Powell, *Wool Tech. Sheep Breed,* **1996,** *44*, 100.
22. R.C. Marshall, J.M. Gillespie, A.S. Inglis, M.J. Frenkel, in: Proceedings of the Xth International Wool Textile Research Pretoria, **1980**, Vol. *II*, p. 147.
23. T.A.A. Dopheide, Eur. J. Biochem. **1973**, *34,* 120.
24. E.S. Kuczek, G.E. Rogers, Eur. J. Biochem. **1987,** *166,* 79.
25. B.C. Powell, M.J. Sleigh, K.A. Ward, G.E. Rogers, *Nucleic Acids Res*. **1983**, *11*, 5327.
26. A. Fratini, B.C. Powell, G.E. Rogers, *J. Biol. Chem.* **1993,** *26*, 4511.
27. J.M. Gillespie, in: R.D. Goldman, P.M. Steinert (Eds.), Cellular and Molecular Biology of Intermediate Filaments, Plenum Press, New York, **1990**, p. 95.
28. S. Yoshihashi-Suzuki, Y. Naito, K. Ishii, K. Awazu, The Review of Laser Engineering, Special Issue on Mid-Infrared Free Electron Laser and Its Application Research, **2003,** *31*, 835.
29. Y. Naito, S. Yoshihashi-Suzuki, K. Ishii, K. Awazu, Proceedings of the 25th International Free Electron Laser Conference, and the 10th FEL Users Workshop, **2004**, *528*(1–2), 609.
30. Y. Naito, S. Yoshihashi-Suzuki, K. Ishii, T. Kanai, K. Awazu, *Int. J. Mass Spectrom.* **2005,** *241*(1), 49.

15. ELECTROSPRAY IONIZATION TANDEM MASS SPECTROMETRIC INVESTIGATION OF ESSENTIAL OILS FROM *MELISSA OFFICINALIS* (*LABIATAE* FAMILY) AND *PELLARGONIUM* SSP. (*GERANIACEAE* FAMILY)

CLAUDIA C. TOMA[1], IOAN B. PANCAN[2], MARIUS CHIRIȚĂ[3], FLORINA M. VATA[1], AND ALINA D. ZAMFIR[2,3]

[1]*Department of Pharmacognosy & Phytotherapy "Victor Babes" University of Medicine and Pharmacy Timisoara, Faculty of Pharmacy, 2 Eftimie Murgu Street, 300041 Romania*
[2]*Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania*
[3]*BioMass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania*

Abstract. In the present work we report upon the development of a novel methodology based on electrospray (ESI) high capacity ion trap (HCT) multistage mass spectrometry (MS/MS), for assessing the composition and structure of essential volatile oils. The method was particularly applied to a native terpenoid mixture extracted from *Melissa officinalis* and different species of *Pellargonium* genus. Optimized ESI HCT MS and MS/MS in positive ion mode allowed the detection of three major components without the need of chromatographic off- or on-line separation prior to MS. Linalol, citronellol, and citral species could be reliably identified based on accurate mass measurement of their molecular and related sequence ions.

1.   Introduction

Melissa officinalis L. *(Labiatae)*–with the common name lemon balm or sweet balm- is an herbaceous specie from *Lamiaceae* family, originating from Northern Africa (Egypt, Tunisia, Marocco) and Southern Europe. For its specific medicinal qualities, lemon balm exists in culture (crops) in countries such as France, Italy, Germany, Spain, USA and especially Hungary.

In Romania, lemon balm is a spontaneous plant often observed in the regions of Caraș-Severin, Mehedinți, Gorj and Vâlcea, where it prefers the rocky hills with oak, beech or acacia forests. *Melissa officinalis* exists around Baia de Aramă locality.

Botanically, *Melissa officinalis* is a perennial, herbaceous plant. The root-stock is short 25–30 cm, the stem square and branching, grows 60–120 cm high, and has at each joint pairs of broadly ovate or heart-shaped, crenate or toothed leaves which emit a fragrant lemon odor when bruised. They also have a distinct lemon taste. The flowers, white or yellowish, are in loose, small bunches from the axils of the leaves and bloom from June to October. The plant dies down in winter, but the root is perennial [1,2].

Figure 1. Areal of spontaneous distribution of Melissa *officinalis* L. *(Labiatae)* in Romania

Lemon Balm grows freely in any soil and can be propagated by seeds, cuttings or division of roots in spring or autumn. If propagated in fall, then it is preferably before October, so that the offsets may be established before the frosts come on. The roots may be divided into small pieces, with three or four buds to each, and planted in ordinary garden soil. *Melissa* grows best in alluvial soil, excess water is harmful, while yields diminish in light and dry soils [3,4].

The genus *Melissa* is widely diffused, having representatives in Europe, Middle Asia and North America. The name is from the Greek word signifying "bee", indicative of the attraction the flowers have for those insects, on account of the honey they produce. The word Balm is an abbreviation of Balsam, the chief of sweet-smelling oils. It is so called from its honeyed sweetness. *Melissa* was highly esteemed by Paracelsus, who believed it would completely revivify the man. It was formerly esteemed of great use in all complaints supposed to proceed from a disordered state of the nervous system. Formerly a spirit of Balm, combined with lemon-peel, nutmeg and angelica root, enjoyed a great reputation under the name of Carmelite water, being deemed highly useful against nervous headache and neuralgic affections [5,6].

Many reports exist regarding the chemical composition of *Melissa officinalis* L subsp. officinalis. *Oleum Melissae* is a pale yellow oil having a woody-green aroma [7].

Pelargonium capitatum Ehrh. ex Willd., *Geraniaceae* Family was the second species investigated.

The plants originated from South Africa as well as Reunion, Madagascar, Egypt, Tunis and Morocco. They were introduced to European countries such as

Italy, Spain and France in the seventeenth century. In early times geraniums were planted around the house to help keep evil sprits at bay.

Rose geranium is a hairy perennial shrub, often used in hedgerows, and will stand up to about one meter high with pointed leaves, serrate at the edges and pinkish-white flowers. The leaves and stalks are used for extraction, and the oil is obtained through steam distillation.

Bothanical description: Shrub to 100 cm high with an aromatic odour, harshly hairy; stems covered non-glandular and some glandular hairs; taproot ± fleshy. Leaves alternate or opposite, pinnatisect or palmatisect with lamina ± ovate in outline, to 7 cm long and 6 cm wide, deeply 5–7-lobed, lobes toothed, hirsute; petiole 4–10 cm long.Umbels 5–10-flowered; peduncle 3–6 cm long; pedicels 0–6 mm long. Calyx lobes 6–9 mm long, sepal spur 2–4 mm long. Petals are 10–20 mm long, pink, posterior ones with deeper pink or purple markings.

Fertile stamens are usually six. Fruit 15–20 mm long; mericarps hirsute [8–10]. Annual or perennial herbs or shrubs; leaves simple or almost compound, dentate or lobed, glabrous or variously pubescent and often aromatic.

Flowers irregular, arranged in cymose umbels, the inflorescences themselves arranged in a spiral sequence; sepals connate towards the base and the posterior one prolonged downwards into a spur which is adnate to the pedicel; petals usually convolute, clawed, free, white to deep pink; posterior ones usually larger and often marked with darker spots and lines; stamens usually more or less connate towards the base, 3–8 bearing anthers.

Mericarps dehiscing on the ventral suture, without a pronounced tuft of funicular hairs; the awn is curving upwards, villous inside [10].

From the chemical point of you, the biochemical compounds are: Citronellol, Geraniol, Citronellyl-Formiate, Geranyl-Formiate, linalol and borneol [8,9].

Rose Geranium oil can be used to help in the treatment of the following: acne, bruises, burns, cuts, dermatitis, eczema, hemorrhoids, lice, mosquito repellant, ringworm, ulcers, breast engorgement, edema, poor circulation, sore throat, tonsillitis, menopausal problems, stress and neuralgia. It's reputation as an extremely versatile essential oil is well documented [8].

Rose geranium is useful in many formulas for applying to the skin. It blends well with other floral oils, citrus notes and woody oils such as patchouli and sandalwood and makes a great bath oil.

In the traditional medicine of muslims from Tunisia, Egypt and Morocco, called "ethnopharmacologia", the users called "tradipracticians", utilised a suspension of Geranium volatile oil in distilled water for headteache, toothtache and articular pains. He consume the natural remedium by ingestion "a jeun" (before meals) [8,10] .

Although Rose geranium is not a substitute for real rose oil, it can be added to rose formulae as a way to enhance the rose notes and reduce the amount of true rose oil used. Rose Geranium is also great when added to massage oils.

2.   **Materials and Methods**

The *Melissa officinalis(Labiatae)* essential oil was produced by Fares Laboratory–Orăştie county (Romania) series 03.03.2007. For MS analysis, 10 µl oil was dissolved in 200 µl EtOH (HPLC grade, Merck, Darmstadt, Germany).

The *Pelargonium capitatum(Geraniaceae)* essential oil was produced by Omega Pharma Laboratory–France series 03.12.2006. For MS analysis, 10 µl oil was dissolved in 100 µl EtOH and consecutively dilutioned at 1/100,000 v/v and respective 10 µl suspension of oil in distilled water was dissolved in 200 µl EtOH (HPLC grade, Merck, Darmstadt, Germany).

2.1.   MASS SPECTROMETRY

Mass spectrometry was performed on a High Capacity Ion Trap (HCT) Ultra Post-translational Modification (PTM) mass spectrometer from Bruker (Bruker Daltonik, Bremen, Germany). HCT mass spectrometer is interfaced to a PC running the Compass integrated software package under Windows^{XP}, which includes EsquireControl and Hystar modules for instrument tuning, control and spectrum acquisition, and Data Analysis for storing the ion chromatograms and processing the MS data. For *Melissa officinalis* L. *(Labiatae)* the sample was infused into MS by on-line syringe pump electrospray at a constant flow rate of 0.5 µl/min. Nitrogen at 0.5 l/min flow rate and 300°C temperature was employed for desolvation and as a nebulizer gas at 2 p.s.i. The instrument was set to operate in the positive ion mode under 3.0 kV ESI potential. All spectra were acquired in the mass range 100–1,000 *m/z*, with a scan speed of 8,100 *m/z* per s. Tandem mass spectrometry was performed in the manual MS/MS mode, by collision-induced dissociations (CID) at low energies using He as the collision gas. The precursor ion was selected within an isolation width of 2u. Sequencing was carried out using fragmentation amplitude of 0.8 V.

For *Pelargonium capitatum (Geraniaceae)* volatile oil dilution 1/10, the sample was infused into MS by on-line syringe pump electrospray at a constant flow rate of 0.7 µl/min. Nitrogen at 0.5 l/min flow rate and 300°C temperature was employed for desolvation and as a nebulizer gas at 2 p.s.i. The instrument was set to operate in the negative ion mode under 3.5 kV ESI potential. Capillary exit was seted at 250 V. The spectrum was acquired in the mass range 100–2,000 *m/z*, with a scan speed of 8,100 *m/z* per sec.

For *Pelargonium capitatum (Geraniaceae)* volatile oil dilution 1/100,000, the sample was infused into MS by on-line syringe pump electrospray at a constant flow rate of 0.4 µl/min. Nitrogen at 0.5 l/min flow rate and 300°C temperature was employed for desolvation and as a nebulizer gas at 2 p.s.i. The instrument was set to operate in the positive ion mode under -4.5 kV ESI potential. Capillary exit was seted at 10 V. The spectrum was acquired in the mass range 50–2,800 *m/z*, with a scan speed of 8,100 *m/z* per sec.

For *Pelargonium capitatum (Geraniaceae)* aquous suspension of volatile oil, dilution 1/20 the sample was infused into MS by on-line syringe pump electrospray at a constant flow rate of 0.5 µl/min. Nitrogen at 1 l/min flow rate and 300°C temperature was employed for desolvation and as a nebulizer gas at 2 p.s.i. The instrument was set to operate in the positive ion mode under –4.5 kV ESI potential. Capillary exit was seted at 10 V. The spectrum was acquired in the mass range 100–2,800 *m/z*, with a scan speed of 8,100 *m/z* per s. [11–14]

3.   Results and Discussion
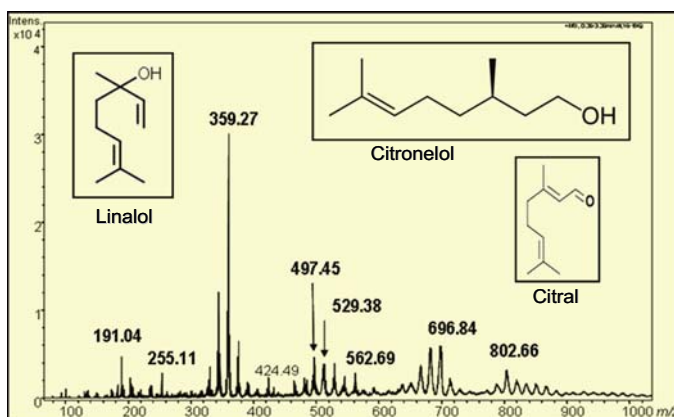
Figure 2. Positive ESI HCT MS of the native terpenoid mixture extracted from Melissa officinalis oil

The (+) ESI HCT MS analysis of the ethanolic solution of the native terpenoid mixture extracted from *Melissa officinalis* is presented in Figure 2. The spectrum was summed over 50 scans (at 2.1 s/scan), which, at the flow rate of 0.5 µl/min, is equivalent to a consumption of 0.87 µl solution and only 0.043 µl essential oil for a mass screening experiment. This value of analyte consumption is situating the experiment sensitivity in the lower nanoliter range.

As visible in Figure 1, by positive ESI HCT MS of the terpenoid mixture, three dominant components: linalol, citronelol, and citral could be identified with high sensitivity and accuracy through the corresponding singly charged ions of their mono- to pentamer complexes (Figure1). According to mass calculation the most abundant ion, detected at *m/z* 359.27, was assigned to the disodiated dimer of citronellol. Additional confirmation of the structure corresponding to the ion at *m/z* 359.27 was achieved via (+) ESI MS/MS by employing collision induced-dissociations at low energy with He as collision gas. The ESI MS/MS is depicted in Figure 3.
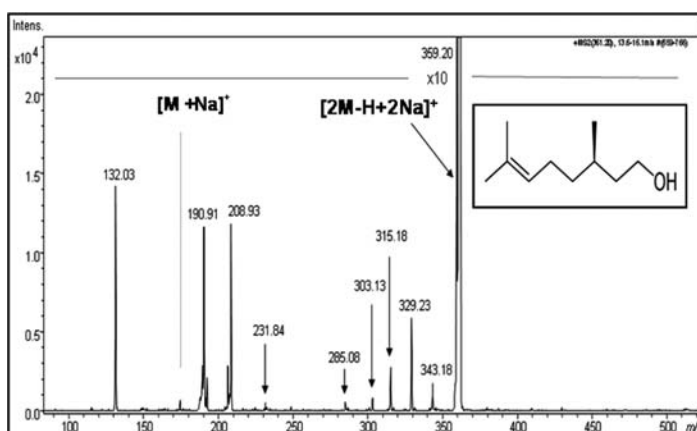
Figure 3. Positive ESI HCT MS^2 of the singly charged ion at *m/z* 359.20 corresponding to the disodiated citronellol dimmer
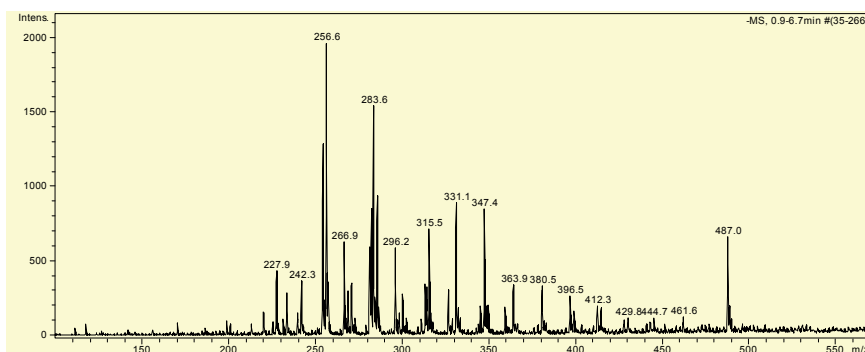

Figure 4. Negative ESI HCT MS of the native terpenoid mixture extracted from *Pelargonium capitatum* volatile oil (dilution 1/10)
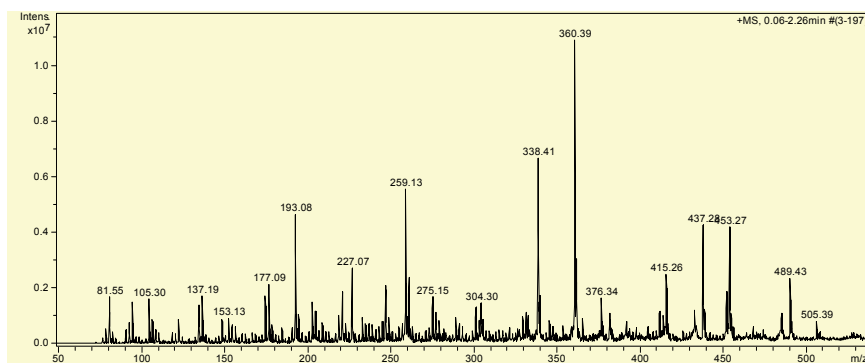

Figure 5. Positive ESI HCT MS of the native terpenoid mixture extracted from *Pelargonium capitatum* volatile oil (dilution 1/10^4)
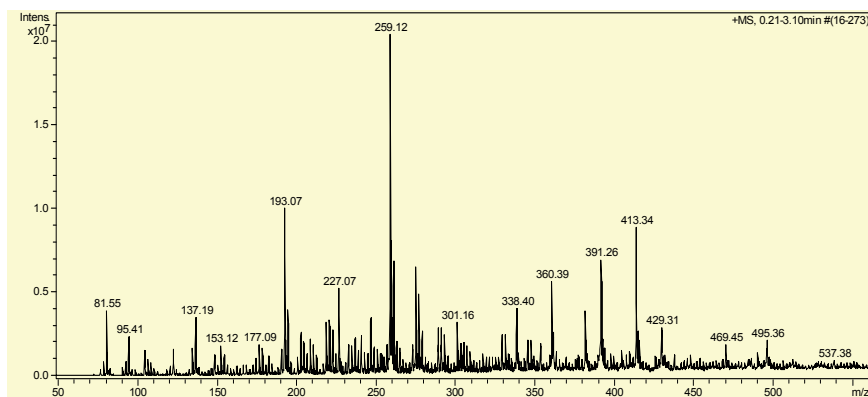
Figure 6. Positive ESI HCT MS of the native terpenoid mixture extracted from *Pelargonium capitatum* volatile oil in aqueous suspension (dilution 1/20)

The spectrum was acquired for 25 scans at the same constant flow rate of ESI infusion, which yielded a sample consumption of 0.021 µl. Figure 3 encompasses as an inset the citronellol structure deduced from diagnostic ions that are representative for successive dehydration and stepwise neutral stripping of $CH_3 (CH_3)_2$, CO, CO_2 and OH of its corresponding disodiated dimer and monosodiated monomer, at nominal *m/z* 179, respectively.

4.   Conclusions

This study was designed to assessing the potential of the ESI HCT MS and MS/MS for screening and sequencing of volatile oils. The methodology was successfully applied for the analysis of a native mixture of terpenoids extracted from *Melissa officinalis* L (*Labiatae*) and *Pelargonium capitatum* vegetal volatile oils. NanoESI HCT MS and MS/MS methodology enabled a highly sensitive, detailed and accurate mass analysis, and subsequently the first ESI MS identification of a number of different terpenoid compounds in the investigated mixtures. Supplementary structural and compositional data upon citronellol expressed in *Melissa officinalis* was achieved by ion trapping and their fragmentation in multistage CID MS. As one of the most modern methods in structural analysis, HCT CID tandem MS was here for the first time employed in volatile oil investigation. The obtained data permit a positive estimation regarding the method perspectives and its general applicability in volatile oil analysis. Moreover, we consider it also as an important step forward in the implementation of advanced mass spectrometry in phitochemistry-aromatherapy research area.

References

1. Carnal A.P., Camat A., Fraisse D. and Lamaison J.L., The aromatic and polyphenolic composition of lemon balm (Melissa officinalis L. subsp. officinalis) tea, *Pharma. Acta Helv.*1998, 72, 301–305.
2. Dawson B.S.W., Franich R.A and Medler R, Essential Oil of Melissa officinalis 1. subsp altissima. (Sibthr. et Smith), *Flavour and fragrance journal,* 1988, 33, 167–170.
3. Miron A., Aprotosoaie C., Hancianu M., Tanasescu V., Stanescu U.,Influenta tratamentului fitosanitar cu Topsin M asupra potentialului antioxidant al frunzelor speciei *Melissa officinalis* L., *Farmacia,* 2005, 5, 26–34.
4. Kreis P. and Mosandl A., Chiral compounds of essential oils. Pan XVI. Enantiosetectlve multidimensional gas chromatography in authentic control of lemon balm., *Flavour and fragrance journal,* 1994, 9, 249–256.
5. Pank P.F., Eichholz E., Ennet D. and Zygmunt B., "Chemical Weed Control in medicinal plant crops" Part 8 Balm (Melissa officinalis L.) *Pharmazie,* 1987,42, 191–195.
6. Tittel G., Wagner H. and Bos R., Chemical composition of the essential oil from Melissa. *Planta Med*ica., 1982, uh 46, 91–98.
7. Zgorka G. and GIowniak K., Variation of free phenolic acids in medicinal plants belonging to the Lamiaceae family, *J. Pharm. Biomed. Anal.,* 2001, 26, 79–87.
8. Bellakhdar J.. La pharmacopee marocaine traditionnelle, Ibis Press, Paris, 1997, 328–330.
9. Bruneton J., Pharmacognosie, phitochimie, plantes medicinales, Ed. Tec & Doc, Paris, 1999, 286–312.
10. Chevalier A., Encyclopedie des plantes medicinales, Larousse, Paris, 2001, 215–217.
11. Zamfir A.D., Vukelic Z., Schneider A., Sisu E., Dinca N. and Ingendoh A., A novel approach for ganglioside structural analysis based on electrospray multiple-stage mass spectrometry., *J Biomol Tech.*, 2007, 18, 188–193.
12. Balen B., Krsnik-Rasol M., Zamfir A.D., Milosević J., Vakhrushev S.Y. and Peter-Katalinić J., Glycoproteomic survey of Mammillaria gracilis tissues grown in vitro, *J Proteome Res.*, 2006, 5, 1658–1666.
13. Cederkvist F.H., Zamfir A.D., Bahrke S., Eijsink V.G., Sørlie M., Peter-Katalinić J. and Peter M.G.,Identification of a high-affinity-binding oligosaccharide by (+) nanoelectrospray quadrupole time-of-flight tandem mass spectrometry of a noncovalent enzyme-ligand complex., *Angew Chem Int Ed Engl.*, 2006, 45, 2429–2434.
14. Zamfir A.D., Bindila L., Lion N., Allen M., Girault H.H. and Peter-Katalinić J., Chip electrospray mass spectrometry for carbohydrate analysis, *Electrophoresis*, 2005, 26, 3650–3673.

16. CHEMICAL STRUCTURE IDENTIFICATION BY DIFFERENTIAL MASS SPECTRA

NICOLAE DINCA

"Aurel Vlaicu" University of Arad, Romania

Abstract. In this paper a new technique FOR mass spectra interpretation based on chemical structure identification by differential mass spectra (CSI Diff-ms) is presented. CSI Diff-ms is broadening the applicability of the mass spectrometry in the area of identification of isomer structure for analytes exhibiting similar mass spectra. By this technique are correlated differential mass spectra (diff ms) of the isomers and formation enthalpies that are calculated quantum-chemical for the molecules and the formed ions in the mass spectrum for the structure identification. Here are presented types of differential spectra the theoretical justification of the correlation between the spectra ratio and formation enthalpies of the ions and three applications which allow to establish the chemical structures of some aromatic position isomers' and of configuration (*endo*- si *exo*-) using the fragmentation spectra. The CSI Diff-ms software allows to import mass spectra, is calculating the differential mass spectra, the spectra similarity, and makes the correlation diff ms with the formation enthalpies calculated in order to obtain the structures probabilities list. The knowledge of the intimate structure of the isomers with help of this technique represents an important progress in the knowledge of the relation structure-activity or toxicity of the natural and synthetic chemical compounds.

1. Introduction

Mass and mass difference are well known ways to interpret mass spectra in order to find out chemical structures and their sequences. Ion intensity is used for quantification of analyte and rarely in qualitative analysis. The similarity of the analyte mass spectrum with the etalon mass spectrum is useful for the recognition of the structures using mass spectra libraries.

Both methods of qualitative analysis have a commune limitation: these cannot precisely establish the chemical structures, usually isomers, which have similar mass spectra. This limitation is a serious disadvantage since in the living cells and not only, these isomers can have a different role or action [1].

This boundary cannot be overcome without having a method to emphasize small differences that appear between similar mass spectra.

1.1. DIFFERENTIAL MASS SPECTRA (DIFF MS)

The graphic or tabulated differential mass spectra can be used for this purpose because they highlight, with sensitivity and accuracy, both the fluctuations of the little peaks as well as the fluctuations of the medium and high peaks. Each differential spectrum is calculated using two fragmentation spectra. It can be obtained few types of differential mass spectra using similar mass spectra [2]:

– The difference of spectra (Δms)

– The percentage difference (Δ%ms)– The ratio of spectra (Pms)These differential spectra are some representation of the difference, percentage difference and respectively of the intensity of the isobar ions from the two differentiated spectra, as a function of *m/z*.

The following snapshots shows the soft Diff-ms graphical visualization features used for the analysis of two pesticides, which are aromatic isomers, having similar mass spectra: 2,2',3,6'-tetrachlorobiphenyl and 2,2',3,5'-tetrachlorobiphenyl (Figures 1, 2 and 3) [3].

The reasons for which are appearing differences between the similar spectra can be: different analytes, same analyte in different matrices, different conditions of recording, fluctuation of the measuring conditions, etc …

In order to have a differential mass spectrum that shows only the structure differences, it is necessary to eliminate all the other causes that influence mass spectra. Thus, the mass spectra that are differentiated have to be recorded under the same condition, to be corrected (e.g., background subtraction and the correction of the analyte concentration's variation [2]) and bring them to a comparative form using TIC normalization.

In this condition the different intensity of the isobar ions from the isomers spectra is due to the different activation energies of the similar fragmentation or to the different stabilities of the formed isobar ions.

In an endothermal transformation, like fragmentations, the structures and the energies of the activated complexes are more similar to the ones of the reaction products that are formed (ions) than to the ones of the reactants (molecules). Thus, when $Ea / Ea' > 1$, then $\Delta H^0_f / \Delta H^0_{f'} > 1$ and the intensities rapport $P_i = F+ / F' + < 1$ (Figure 4).
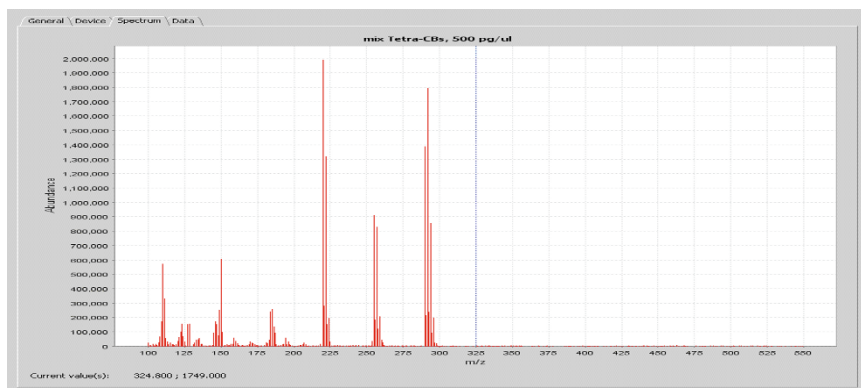

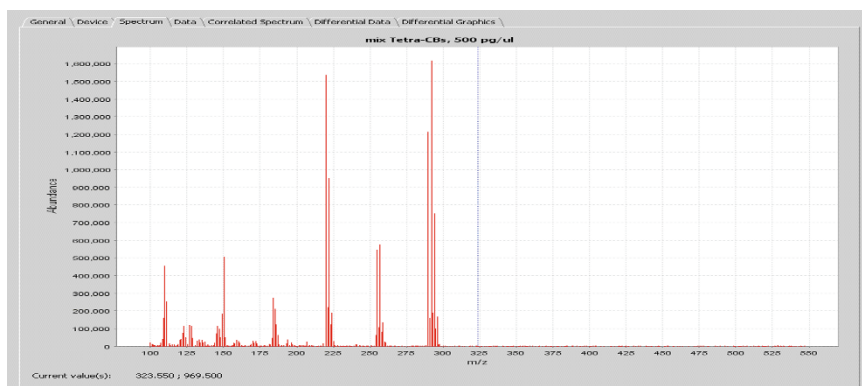*Figure 1.* Mass spectrum of 2,2',3,6'-tetrachlorobiphenyl

Figure 2. Mass spectrum of 2,2',3,5'-tetrachlorobiphenyl

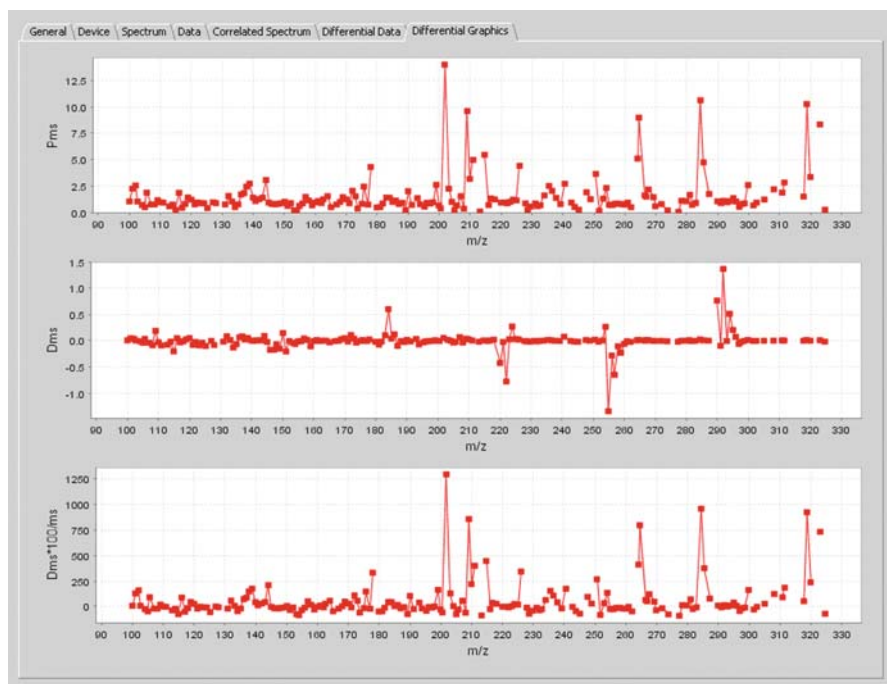*Figure 3.* The graphical differential mass spectra: the difference of spectra (Δms), the percentage difference (Δ%ms) and the ratio of spectra (Pms) of 2,2',3,6'-tetrachlorobiphenyl and 2,2',3,5'-tetrachlorobiphenyl [3]

The heats of formation of the isomer molecules with similar spectra, of the ions and radical ions that are produced from these are accessible using semi-empirical quantum-chemical calculation methods and allow only the estimation of the theoretical ionic intensities rapport, P_i.

If these theoretical rapports, P_i, and the experimental ones from the differential spectrum Pms are both bigger or smaller than 1 for the most important ions from the spectrum, then the structures of the isomer molecules, of the ions and the radical ions resulted from these and used in quantum-chemical calculations are correct. When there are two isomers with similar mass spectra then the calculations volume is reduced and the comparison of the theoretical and experimental P_i can be done relatively easy. When the group of isomers with similar mass spectra it is larger then it is necessary to make all the possible pairing of the spectra in order to obtain the corresponding rapports. In this case, in order to obtain the optimum alternative of structure ascription, the comparison of the theoretical and experimental P_i values it is done through more series and implies a bigger volume of work.
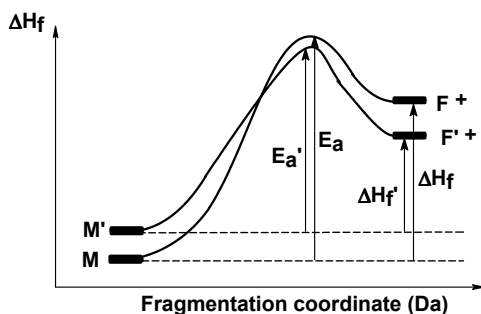

Figure 4. The most common energetic fragmentation diagram of similar structures. Usually the stable molecules are more difficult to be ionized and fragmented

All operations described above can be simplified if it is used the software named Chemical Structure Identification by Differential mass spectra (CSI-Diffms) [3]. The software has the advantage that allows inputting the formation heats, calculated with quantum-chemical software, in a database, and it is also calculating the enthalpy variation and is doing an estimation of the all-theoretical P_i. Moreover the mass spectra of the isomers can be imported and the calculation of the all-possible differential spectra and of the all-experimental P_i values it is automatically done. The software it is also searching the structures based on the comparison of the theoretical and experimental values of P_i for all the spectra series and possible isomers having as final purpose the obtaining of the probabilities. It is considered correct that assignation of the structures for which the probability is highest [2].

2.   Experimental

In the case of the analysis for the recognition through GC/EI-MS of six of the most important congeners of tetra-chloro-biphenyls (**1-6**), the search in the spectra databases it did not allow the identification of the structures, so that the settlement of the position of the four chlorine atoms on the benzene nuclei is not possible.

In this way, in the obtained structures list for the 98–99% probability appears a series of more than ten isomers of position for each analysed spectrum (Table 1). Moreover, just in the case of two isomers of six appear the correct structures (underlined in the table) out of other with the same probability. It is clear that the qualitative analysis of the isomers with similar mass spectra using the spectra databases has serious problems.
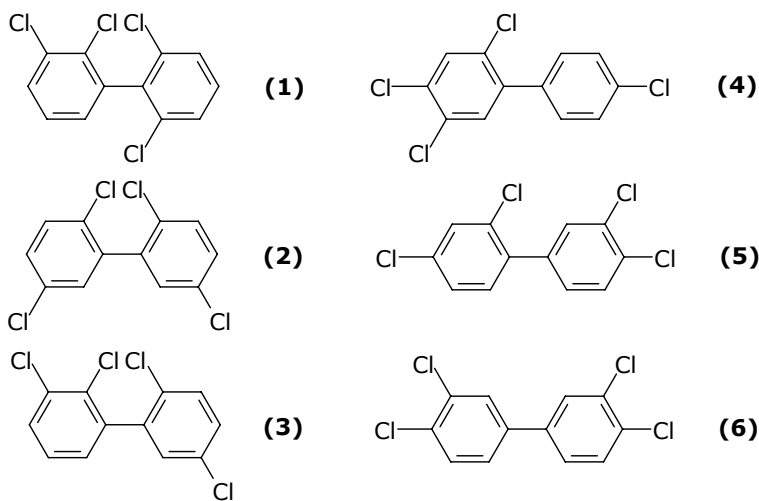

Figure 5. 2,2′,3,6′-tetra-CB (PCB 46) **(1)**; 2,4,4′,5 -tetra-CB (PCB 74) **(4),** 2,2′,5,5′-tetra-CB (PCB 52) **(2)**; 2,3′,4,4′-tetra-CB (PCB 66) **(5),** 2,2′,3,5′-tetra-CB (PCB 44) **(3)**; 3,3′,4,4′-tetra-CB **(**PCB 77) **(6)**

TABLE 1. Probability of Chemical Structure Identification with Spectra Database

*Isomer		Retention time (min)	Probability
X	13.44	1,1'-Biphenyl, 2,3',4,4'-tetrachloro	99%
		1,1'-Biphenyl, tetrachloro-	99%
		1,1'-Biphenyl, 2,2',5,6'-tetrachloro	99%
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',3,4-tetrachloro	99%
		1,1'-Biphenyl, 2,2',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',3,5'-tetrachloro	99%
		1,1'-Biphenyl, 3,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',6,6'-tetrachloro	99%
Y	13.64	1,1'-Biphenyl, 2,3,4',6-tetrachloro	99%
		1,1'-Biphenyl, tetrachloro	99%
		1,1'-Biphenyl, 2,2',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',6,6'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',4,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,3,3',5'-tetrachloro	98%
		1,1'-Biphenyl, 2,2',3,4-tetrachloro	98%
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro	98%
		1,1'-Biphenyl, 2,3',4',5-tetrachloro	98%
		1,1'-Biphenyl, 2,4,4',6-tetrachloro	98%
Z	14.20	1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',3,4-tetrachloro	99%
		1,1'-Biphenyl, tetrachloro	99%
		1,1'-Biphenyl, 2,3,4',6-tetrachloro	99%
		1,1'-Biphenyl, 3,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,3',4,4'-tetrachloro	99%
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',6,6'-tetrachloro	98%
		1,1'-Biphenyl, 2,2',4,4'-tetrachloro	98%
T	15.33	1,1'-Biphenyl, 3,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,2',3,4-tetrachloro	99%
		1,1'-Biphenyl, 2,3,4',6-tetrachloro	99%
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99%
		1,1'-Biphenyl, 2,3',4,4'-tetrachloro	98%
		1,1'-Biphenyl, 2,3',4',5-tetrachloro	98%
		1,1'-Biphenyl, 2,3,3',4-tetrachloro	98%
		1,1'-Biphenyl, 2,4,4',6-tetrachloro	98%
		1,1'-Biphenyl, 2,2',4,4'-tetrachloro	98%

		1,1'-Biphenyl, 2,3',4',6-tetrachloro	98%
W	15.52	1,1'-Biphenyl, 2,3',5,5'-tetrachloro	99
		1,1'-Biphenyl, 3,3',5,5'-tetrachloro	99
		1,1'-Biphenyl, 2,3,4',6-tetrachloro	99
		1,1'-Biphenyl, 2,2',3,4-tetrachloro	99
		1,1'-Biphenyl, 2,4,4',6-tetrachloro	99
		1,1'-Biphenyl, 2,3,3',4'-tetrachloro	99
		1,1'-Biphenyl, 3,3',4,5'-tetrachloro	99
		1,1'-Biphenyl, 3,3',4,4'-tetrachloro	99
		1,1'-Biphenyl, 2,3',4',5-tetrachloro	99
U	17.43	1,1'-Biphenyl, 2,3,3',4'-tetrachloro-	99
		1,1'-Biphenyl, 2,2',3,4-tetrachloro-	99
		1,1'-Biphenyl, 3,3',5,5'-tetrachloro-	99
		1,1'-Biphenyl, 2,3,4',6-tetrachloro-	99
		1,1'-Biphenyl, 3,3',4,5'-tetrachloro-	99
		1,1'-Biphenyl, 2,3',4',5-tetrachloro-	99
		1,1'-Biphenyl, 3,3',4,4'-tetrachloro-	99
		1,1'-Biphenyl, 2,2',4,4'-tetrachloro-	99
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro-	99
		1,1'-Biphenyl, 2,3',4,4'-tetrachloro-	99
		1,1'-Biphenyl, 2,3',5,5'-tetrachloro-	99

TABLE 2. The Similarity of the Mass Spectra for the Isomers. The Shadings are Indicating High Similarities, Over 90%

Isomer	X	Y	Z	T	W	U
X	100%	92.78	97.79	80.83	81.12	77.71
Y	92.78	100%	93.73	88.12	88.47	86.09
Z	97.79	93.73	100%	81.84	82.14	82.06
T	80.83	88.12	81.84	100%	97.94	86.55
W	81.12	88.47	82.14	97.94	100%	90.04
U	77.71	86.09	82.06	86.55	90.04	100%

The high similarity of the six spectra, calculated with a severe algorithm of the CSI-Diff-ms explains why the search in the spectra databases does not give results (Table 2). In this way, can be remarked the fact that the X, Y, Z, isomers and respectively T, W, U have similarities of over 90% even in the situation of using a severe algorithm specially elaborated to differentiate similar mass spectra [2]. The usage of the differential mass spectra and of the formations enthalpies for the molecules and ions, M^+; $[M-Cl]^+$, $[M-2Cl]^+$, $[M-3Cl]^+$, $[M-4Cl]^+$, within the CSI Diff-ms software has allowed the calculations of the probabilities' (Table 3).

TABLE 3. The Probability List Obtained by the CSI Diff-ms Program for Six Congeners of Tetra-chloro-biphenyls. The Shadings are Indicating the Accurate Structures

Probability	Isomer X	Isomer Y	Isomer Z	Isomer T	Isomer W	Isomer U
93%	22'36'-tetraCB	22'55'-tetraCB	22'35'-tetraCB	244'5-tetraCB	23'44'-tetraCB	33'44'-tetraCB
88%	22'36'-tetraCB	22'35'-tetraCB	22'55'-tetraCB	244'5-tetraCB	23'44'-tetraCB	33'44'-tetraCB
85%	22'36'-tetraCB	22'55'-tetraCB	22'35'-tetraCB	23'44'-tetraCB	244'5-tetraCB	33'44'-tetraCB
85%	22'36'-tetraCB	22'55'-tetraCB	22'35'-tetraCB	244'5-tetraCB	33'44'-tetraCB	23'44'-tetraCB
84%	22'55'-tetraCB	22'35'-tetraCB	22'36'-tetraCB	244'5-tetraCB	23'44'-tetraCB	33'44'-tetraCB
84%	22'36'-tetraCB	22'55'-tetraCB	22'35'-tetraCB	33'44'-tetraCB	244'5-tetraCB	23'44'-tetraCB
82%	22'35'-tetraCB	22'55'-tetraCB	22'36'-tetraCB	244'5-tetraCB	23'44'-tetraCB	33'44'-tetraCB

The maximum probability (93%) obtained corresponds to the real structure of the standards used for the verification of this new method for the establishment of the chemical structures using the differential mass spectra and the formation enthalpies of the molecules and ions from the mass spectra. The following probabilities' series are containing inversions in regard to the structures ascription. The fact that are easily mixed up the structures that have spectra with high similarity, X, Y and Z or T, W and U, is also demonstrating that the correlation algorithm of the differential mass spectra with the formation enthalpies is veridical.

The appropriation of the structure through this method it is becoming more accurate when the matching of the theoretical and practical P_i ratios is done for a bigger number of ions from the spectrum. The minimum number of ions that has to be used in order to discriminate the structures of a certain number of isomers with similar spectra it is indicate din Table 4.

TABLE 4.The Minimum Number of Ions Used for Search Structures with CSI Diff-ms

Number of identified isomers	Minimum number of selected ions
2	1
4	2
8	3
……	…..
2^g	g

In the case of the aromatic position isomers 3- nitro-benzophenone-dimethylacetal (**7**) and 4- nitro-benzophenone-dimethylacetal (**8**) (Figure 6) were used six ions of the spectrum, ions that are mentioned in the window "Selected Ions" from Figure 7, which it six times more than the necessary minimum [4]. The heats of formation (ΔH^0_f) of molecules and fragment ions were calculated with the semi-empirical method AM1 [6–8], using the RHF operators for molecules or ions and UHF for the radical ions. The calculated heats of formation were introduced in the database of CSI Diff-ms (Table 5).

The appropriation of the isomers structure was correct and sure (Figure 7). The probability of 100% with what is done shows that for the all considered ions was realised a correlation between the intensities from the spectrum and the formation enthalpies.
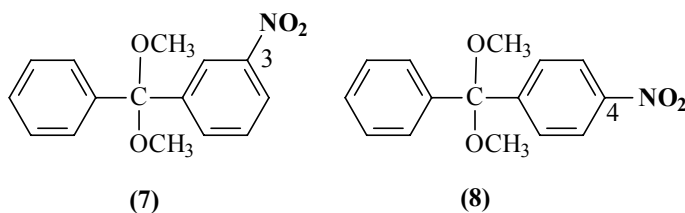
Figure 6. 3- nitro-benzophenone-dimethylacetal (**7**) and 4- nitro-benzophenone-dimethylacetal (**8**)

Figure 7. The snapshot of probability list window obtained for (**7**) and (**8**) aromatic position isomers by CSI Diff-ms

TABLE 5. Heats of Formation (ΔH_f) for the Molecules and Ions from the Mass Spectra of (**7**) and (**8**), Calculated with the Semi-Empirical Methods AM1

Molecule, ion	*m/z*	ΔH°_f (kcal/mol)	
		(**7**)	(**8**)
M	273	−29.04	−29.08
M+	273	+178.81	+194.55
$[M-CH_3O]^+$	242	+197.76	+199.68
$[M-CH_3O-O]^+$	226	+254.94	+256.88
$[M-C_6H_5]^+$	196	+123.08	+125.22
$[M-C_6H_5-CH_3]^+$	181	+176.37	+186.88
$[M-C_6H_5-CH_3O]^+$	165	+226.02	+230.18
$[M-C_6H_5-CH_3O-CH_3]^+$	150	+205.82	+207.73

By M have been denoted the molecules and by M^+ and [ ]$^+$ the ions.

Good results have been obtained, too, by using the algorithm of configure-tion isomer identification, e.g., *exo-* and *endo-*5,10-methylene-10,11-dihydro-5H-dibenzo [a,d] cyclohepten-11-ols (**9**) and (**10**) (Figure 8) [5].
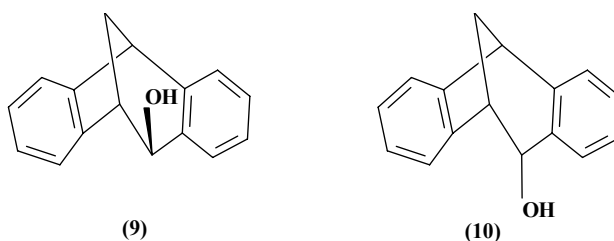

(**9**)   (**10**)

Figure 8.   exo- (**9**) and *endo-*(**10**) isomers of 5,10-methylene-10,11-dihydro-5H- dibenzo [a,d] cyclohepten-11-ols

In this work, too, the heats of formation (ΔH^0_f) of molecules and fragment cations were calculated with the semi-empirical method AM1 [6–8]. Since in the case of these diasteroisomers are formed groups of isobar ions a ponderated heat of formation was calculated for each group of the individual ionic enthalpies. The resulted values were introduced in the database of CSI Diff-ms (Figure 9).

For this correlation were used the most important of the spectrum ions: M^+. $[M-H]^+$, $[M-2H]^+$, $[M-3H]^+$, $[M-OH]^+$, $[M-H_2O]^+$, that is six times more than the necessary minimum [5].

In this case of isomers, too, the appropriation of the isomer structurues was correct and sure (Figure 10). The probability of 100% with what this was done shows that for all the considered ions were realized the correlation between the theoretical and practical P_i intensities.
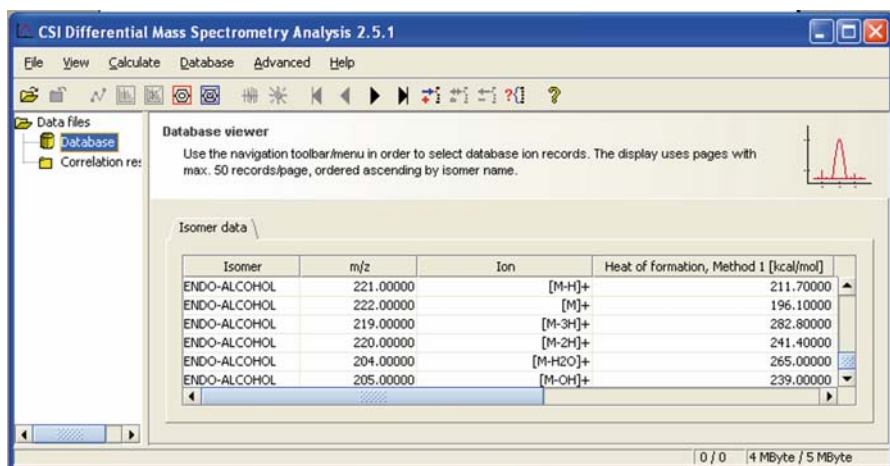
Figure 9. The snapshot of CSI Diff-ms database window containing **(9)** and **(10)** diastereoi-somers heat of formations

Figure 10. The snapshot of probability list window obtained for **(9)** and **(10)** configuration isomers by CSI Diff-ms

3.   Conclusions

The applications presented in here and the ones that refers to the studies of some 1,3-dioxane derivatives publicated in this book, demonstrates that the relation between the ions intensity from the similar mass spectra and their formation

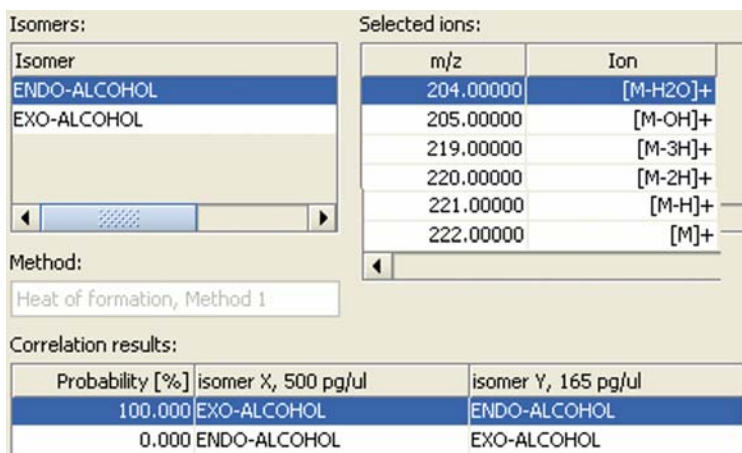enthalpies is permanent, logic and justifies by the chemical kinetic and thermo-dynamic laws.

The **CSI Diff-ms** technology and software do not include and not replace the well-known *"library search"* but enhance this method. So, "library search" is easily indicating the isomers group with similar spectra, and the CSI Diff-ms is establishing the structures within the isomers group using the ions formation enthalpies (Figure 11).
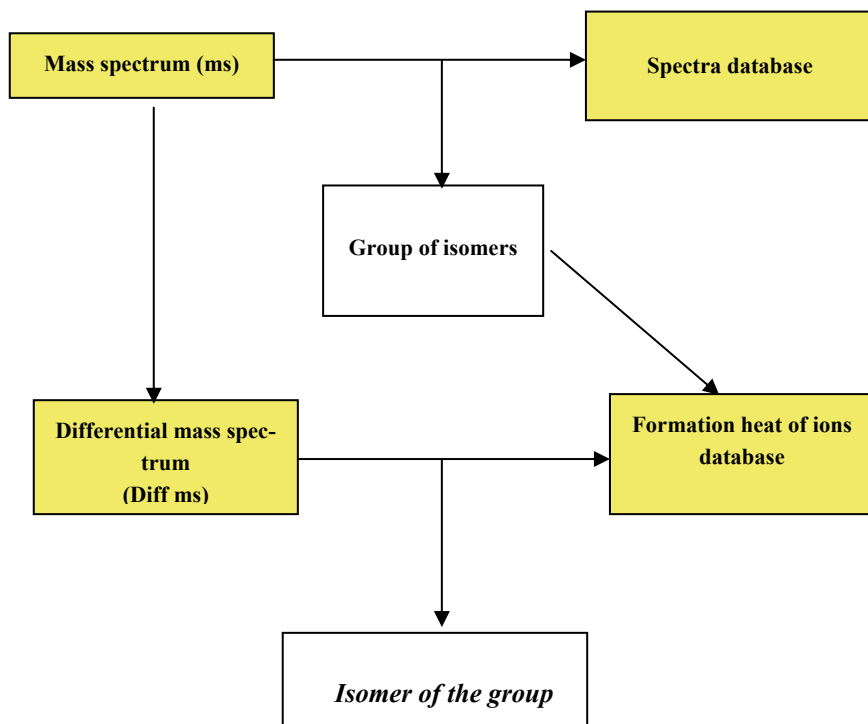

Figure 11. The relationship between the establishment technique of the chemical structure through the search in spectra database and the one through the correlation of the differential mass spectra with calculated formation enthalpies

The chemical structure identification technique by differential mass spectra (CSI Diff-ms) is part of a relatively new mass spectrometry field, differential mass spectrometry (Diff MS). Through this new technique it is put in value the possibility of using the ionic intensities in order to obtain qualitative information, aspect that is not often met in the mass spectrometry of the isomers. The actual stage that CSI Diff ms is passing through it is one of verification using the etalon isomers, and of generalization too in regards to the classes of substances and all types of isomers for which is applied as well as for the fragmentation spectra which can be used. One of the actual limitations of the method it is due to the absence of some

software adapted for the mass spectrometer which can allow the fast quantum-chemical calculus for the thermodynamic and kinetic data that characterize the ions and the fragmentations from the mass spectrometer.

The knowledge of the intimate structure of the isomers with help of the CSI Diff ms represents an important progress in the knowledge of the relation isomer structure-activity or toxicity, factors that are important in the assurance of life safety.

References

1. Smith, R.M., Bush, K.L. *Understanding Mass Spectrometry,* J. Wiley & Sons, **1999,** 76–170.
2. Deutsche Patent 10 2005 028944.4 / 21.06.**2005**, pending homologation.
3. www.bet2-soft.de
4. E. Sisu, I. Sisu, N. Dinca, C. Csunderlik, I. Oprean, V. Rusu, M. Mracec, *Rev. Roum.Chim.*, **2002**, *47*, 339–342.
5. Stănescu M.D., Florea C., Filip P., *Proc. Rom. Acad. Series B*, **2000**, *2*, 7.
6. M.J.S. Dewar, G.E. Zoebisch, F.E. Healy, J.J.P. Stewart, *J. Amer. Chem. Soc.,* **1985**, *107*, 3902.
7. J.J.P. Stewart, *J. Comput. Aided Mol. Design,* **1990**, *4*, 1–103.
8. HyperChemTM Release 5. Professional for Windows, Hypercube, Inc. **1999**, Gainesville, Fl 32601, USA.

LIST OF CONTRIBUTORS

Name	Address	Chapter
Christine Bettendorf	BET2 Software, Konigsbrunn, Germany	12
Steven Blais	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Helene L. Cardasis	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Artur Cavaco-Paulo	University of Minho, Department of Textile Engineering, 4800-058 Guimarães, Portugal	13
Marius Chiriță	BioMass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania	15
Adrian Covaci	Toxicological Centre, Department of Pharmaceutical Sciences, University of Antwerp, Universiteitsplein 1, B-2610 Antwerp, Belgium	11
Costel C. Darie	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1, 2
Leesa J. Deterding	Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709, USA	3, 9, 10
Nicolae Dinca	Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania	6, 12, 13, 16
Alin C. Dirtu	Department of Inorganic and Analytical Chemistry, "Al. I. Cuza" University of Iassy, Carol I Bvd. No 11, 700506 Iassy, Romania	11
Mihaela Drăgușanu	Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany	4

David Fenyo	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Mirela Galusca	Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania	6
Matthew B. Gates	Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709, USA	3
Ion Grosu	Department of Organic Chemistry, "Babes-Bolyai" University of Cluj-Napoca, Romania	12
Florian Harja	Mass Spectrometry Laboratory, Department of Chemistry and Biology, University of Arad, Romania	12
Roxana Elena Iacob	The Barnett Institute, Northeastern University, Boston, MA, USA	8
Andrea Koerner	DWI an der RWTH Aachen, Pauwelstr. 8, D-52056 Aachen, Germany	14
Eveline S. Litscher	Department of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029-6574, USA	2
Cristina Mosoarca	Mass Spectrometry Laboratory, National Institute for Research and Development in Electrochemistry and Condensed Matter, Timisoara, Romania	5
Florentina-Daniela Munteanu	University "Aurel Vlaicu" Arad, Faculty of Food Engineering, Tourism and Environmental Protection, Department of Chemical and Biological Sciences, Elena Dragoi 2, 310330, Arad, Romania	13
Adina Muresan	Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania	6
Thomas A. Neubert	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Ioan B. Pancan	Department of Chemical and Biological Sciences, "Aurel Vlaicu" University of Arad, Romania	15

Irina Perdivara	Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany	8
Brînduşa-Alina Petre	Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany	4
Michael Przybylski	Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany	4, 8
Daniela G. Seidler	Institute for Physiological Chemistry and Pathobiochemistry, University of Muenster, Germany	6, 7
Vivekananda Shetty	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Daniel S. Spellman	Department of Structural Biology, Skirball Institute of Biomolecular Medicine, New York University School of Medicine, 540 First Avenue, Lab 5-18 New York, NY 10016, USA	1
Claudia C. Toma	Department of Pharmacognosy & Phytotherapy "Victor Babes" University of Medicine and Pharmacy Timisoara, Faculty of Pharmacy, 2 Eftimie Murgu Street, 300041 Romania	15
Kenneth B. Tomer	Laboratory of Structural Biology, National Institute of Environmental Health Sciences, National Institutes of Health, PO Box 12233, MD F0-03, Research Triangle Park, NC 27709, USA	3, 8, 9, 10
Florina M. Vata	Department of Pharmacognosy & Phytotherapy "Victor Babes" University of Medicine and Pharmacy Timisoara, Faculty of Pharmacy, 2 Eftimie Murgu Street, 300041 Romania	15
Željka Vukelić	Department of Chemistry and Biochemistry, Faculty of Medicine, University of Zagreb, Croatia	5
Paul M. Wassarman	Department of Molecular, Cell and Developmental Biology, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, NY 10029-6574, USA	2

Jason G. Williams   Laboratory of Structural Biology, Mass Spectrometry   9
                       Group, National Institute of Environmental Health
                       Sciences, National Institutes of Health, Research
                       Triangle Park, NC, USA

Chongfeng Xu   Department of Structural Biology, Skirball Institute   1
                       of Biomolecular Medicine, New York University
                       School of Medicine, 540 First Avenue, Lab 5-18
                       New York, NY 10016, USA

Alina D. Zamfir   Mass Spectrometry Laboratory, National Institute for   5, 6, 15
                       Research and Development in Electrochemistry and
                       Condensed Matter, Timisoara, Romania

Guoan Zhang   Department of Structural Biology, Skirball Institute   1
                       of Biomolecular Medicine, New York University
                       School of Medicine, 540 First Avenue, Lab 5-18
                       New York, NY 10016, USA