Protein Identification and Characterization through Peptide Mass Spectrometry

Method Development for Improved Ricin and Botulinum Neurotoxin Analysis

Tomas Bergström
Faculty of Forest Sciences
Department of Forest Genetics and Plant Physiology
Umeå

Doctoral Thesis
Swedish University of Agricultural Sciences
Umeå 2016

Acta Universitatis agriculturae Sueciae
2016:99

ISSN 1652-6880
ISBN (print version) 978-91-576-8700-5
ISBN (electronic version) 978-91-576-8701-2

Print: Arkitektkopia, Umeå 2016

Protein Identification and Characterization through Peptide Mass Spectrometry. Method Development for Improved Ricin and Botulinum Neurotoxin Analysis

Abstract

The illicit usage of toxic substances is increasing rapidly around the world, creating a need for comprehensive chemical methods for detecting and quantitating toxic agents that might be used in criminal activities or terror attacks. There are several toxic proteins that could potentially be used in bioterrorism or biocrime. For example, the plant toxins ricin and abrin are readily available, very toxic, and easily produced, and are therefore frequently involved in biocrime incidents. Another threat agent is the extremely poisonous botulinum neurotoxin, which is among the most toxic substances known. Forensic analyses of samples potentially containing these diverse and very toxic agents therefore require analytical methods capable of detecting trace amounts of the target analyte in complex mixtures.

This thesis describes the development of new methods for peptide mass spectrometry that offer improved performance in forensic toxin analysis. A galactose affinity method was developed for the enrichment and subsequent analysis of ricin, abrin and botulinum neurotoxin. The method's applicability was confirmed during a forensic investigation into illegal toxin preparations. Additionally, the investigations generated new information on the structural properties of ricin that will facilitate the forensic matching of samples to sources. Details of botulinum neurotoxin's sophisticated intoxication mechanisms were determined by using a broad analytical approach to study the importance and heterogeneity of SV2C glycosylation.

Overall, the results and procedures presented in this thesis will help to improve analytical capabilities relevant to the detection and prevention of biocrime and bioterrorism. More generally, it provides methodological guidance and useful strategies for researchers in peptide mass spectrometry.

Keywords: biocrime, bioterrorism, mass spectrometry, toxin, *Ricinus communis*, ricin, abrin, Type 2 ribosome-inactivating protein (RIP-II), *Clostridium botulinum*, botulinum neurotoxin (BoNT), synaptic vesicle glycoprotein 2C (SV2C), deamidation, affinity chromatography

Author's address: Tomas Bergström, SLU, Department of Forest Genetics and Plant Physiology, SE-901 83 Umeå, Sweden or Swedish Defence Research Agency (FOI), SE-901 82 Umeå, Sweden
E-mail: tomas.i.bergstrom@slu.se, tomas.bergstrom@foi.se

Till Alfred. Min stolthet. Min hjälte.

…till Kanté, med vänsterfoten framåt, å vad bra, till skottläge för Payet, det är ett skott, I KRYSSET!! VILKET MÅL!! DET ÄR FULLSTÄNDIGT FANTASTISKT!! Det är 2-1 för Frankrike, det är konstnären och tavlan, solen, vinden och vattnet, hela havet fullt med dansande laxar. Det är alla trix i alla trollkarlars hattar genom alla tider. DIMITRI PAYET, WAOW-WAOW-WAOW! UPP I KRYSSET! Ah de är ju inte klokt. Och matchen är slut i stort sett och så gör han ett sånt mål för Frankrike, oh la la…

Christian Olsson, Radiosporten

Contents

List of Publications

This thesis is based on the work contained in the following papers, referred to by Roman numerals in the text. The order of papers is scientific rather than chronological.

I Bergström T, Fredriksson S-Å, Nilsson C, Åstot C (2015). Deamidation in ricin studied by capillary zone electrophoresis- and liquid chromatography-mass spectrometry. *Journal of Chromatography B* 974, 109–117.

II Fredriksson S-Å, Artursson E, Bergström T, Östin A, Nilsson C, Åstot C (2015). Identification of RIP-II toxins by affinity enrichment, enzymatic digestion and LC-MS. *Analytical Chemistry* 87(2), 967-74.

III Bergström T, Fredriksson S-Å, and Åstot C. Selective botulinum neurotoxin enrichment based on progenitor complex chemistry followed by LC-MS identification. *Manuscript.*

IV Mahrhold S, Bergström T, Stern D, Dorner BG, Åstot C, Rummel A (2016). Only the complex N559-glycan in synaptic vesicle glycoprotein 2C mediates high affinity binding to botulinum neurotoxin serotype A1. *Biochemical Journal* 473(17), 2645-2654.

The contribution of Tomas Bergström to the papers included in this thesis was as follows:

I Developed the main ideas, planned and performed experiments. Played the major role in drawing the conclusions. Wrote most of the manuscript and made all figures and tables.

II Assisted in planning. Participated in some of the method development steps. Took an active part in in drawing conclusions and in laying out and revising the manuscript.

III Helped to develop some of the ideas and planned most of the work. Performed all experiments and played the major role in drawing the conclusions. Wrote most of the manuscript and made all figures and tables.

IV Contributed to the overall planning of the study. Planned and performed the LC-MS analysis. Participated in the construction and purification of some of the SV2C mutants. Wrote the text and made all figures and tables in the LC-MS section of the manuscript. Participated actively in revising the manuscript.

Other papers by the author not included in this thesis:

Weisemann J, Krez N, Fiebig U, Worbs S, Skiba M, Endermann T, Dorner MB, Bergström T, Muñoz A, Zegers I, Müller C, Jenkinson SP, Avondet MA, Delbrassinne L, Denayer S, Zeleny R, Schimmel H, Åstot C, Dorner BG, Rummel A (2015). Generation and Characterization of Six Recombinant Botulinum Neurotoxins as Reference Material to Serve in an International Proficiency Test. *Toxins* 7(12), 5035-54.

Ovenden SP, Fredriksson S-Å, Bagas CK, Bergström T, Thomson SA, Nilsson C, Bourne DJ (2009). De novo sequencing of RCB-1 to -3: peptide biomarkers from the castor bean plant *Ricinus communis*. *Analytical Chemistry* 81(10), 3986-96.

Östin A, Bergström T, Fredriksson S-Å, Nilsson C (2007). Solvent-assisted trypsin digestion of ricin for forensic identification by LC-ESI MS/MS. *Analytical Chemistry* 79(16), 6271-8.

Abbreviations

All abbreviations are explained as they first appear in the text.

Preface

This thesis summarizes the research work that I have done as a postgraduate student working in a collaboration between the Swedish University of Agricultural Sciences (SLU) and the Swedish Defence Research Agency (FOI). As a student employed and supervised by SLU in its core plant biology facility the Umeå Plant Science Centre (UPSC), I have had access to a very productive research environment with excellent support from fellow PhD students, researchers and supervisors. I have also taken most of my courses through SLU, and received both strategic and technical support from that institution.

Although I have been educated under SLU's Faculty of Forest Sciences, my research tasks and objectives have been aligned with FOI's purpose of conducting "research for a safer and more secure world". FOI is an applied research institute working in the areas of defence and security, and the division for CBRN Defence and Security is located in Umeå. As part of this FOI facility, I have had the opportunity to work on research questions and tasks at the field's frontier. In particular, working at FOI has granted me access not just to the toxic and regulated substances necessary for my research, but also to the institution's vast research experience and extensive competences. Additionally, FOI's active collaborations with other government agencies, defence research organisations, and academic research groups have provided excellent opportunities for presentations, dialogues, and educational visits.

The primary goal of summarizing research in a thesis is to show that one fulfil the criteria for a PhD degree. My secondary goal was to provide methodological guidance and useful strategies for researchers in the field of peptide mass spectrometry.

1 Introduction

Societal concerns about acts of biocrime and bioterrorism necessitate the development of improved chemical analysis capabilities to detect chemical and biological agents that may be used in such acts. The illicit usage of toxic substances has increased globally (Roxas-Duncan and Smith, 2012, Pita and Romero, 2014, Worbs *et al.*, 2011), emphasizing the need for comprehensive analytical methods to detect and quantitate a range of toxic substances in diverse matrices.

An act of bioterrorism is the deliberate release of viruses, bacteria or other agents of biological origin intended to cause illness or death in people, animals or plants, while an act of biocrime implies to kill or make ill single individuals or a small group of people. Biocrime is usually motivated by hate, revenge or monetary gains, while bioterrorists are driven by political, ideological or religious beliefs (Jansen *et al.*, 2014, CDC).

In suspected cases of biocrime or bioterrorism, analytical results (i.e. information on a sample's composition) must be acquired and delivered as quickly as possible. If a sample contains toxic or dangerous substances, a rapid response involving the delivery of suitable medical treatments or other safety measures is obviously desirable (Moran *et al.*, 2008). Similarly, if a sample can be rapidly proven safe, prompt dissemination of this information may prevent public alarm and avoid the need to enact costly precautionary measures.

Two of the most important threat agents in these contexts are ricin and botulinum neurotoxin. Ricin is extracted from the seeds of the castor bean plant (*Ricinus communis*), and is a substance of concern due to its high toxicity and availability. The ricin plant is grown worldwide for the castor oil content of its seeds, but it is also common as an ornamental garden plant. Additionally, numerous recipes on the extraction and "recommended" use of the toxin are available on the internet. Consequently, it is frequently used in acts of biocrime

(Roxas-Duncan and Smith, 2012) and has become recognized as “the poor man’s toxic weapon”.

Botulinum neurotoxin is extremely toxic and is considered the most lethal protein known to mankind. Fortunately, its illicit use is rare. Because it is produced by the common anaerobic spore-forming bacterium *Clostridium botulinum*, most cases of botulinum intoxication cases have been due to natural causes including food poisoning from badly preserved food. However, the extreme toxicity and global presence of *C. botulinum* spores has prompted the US Centre for Disease Control to designate the botulinum neurotoxin as one of the agent most likely to be used in acts of bioterrorism (Arnon *et al.*, 2001, CDC).

Methods for identifying these extremely toxic substances must be highly sensitive due to the small quantities that are sufficient to intoxicate an individual or a group of people. Although ricin and botulinum neurotoxin are well known and widely studied proteins, deeper knowledge about their structures and mechanisms of intoxication facilitate the development of faster, more sensitive, more accurate, more specific, and more robust methods for their detection.

2 Background

The general background information on proteins presented in sections 2.1-2.4 is based on the following references in cases where no other references are provided: (Encyclopædia Britannica, Van Holde and Mathews, 1996, Campbell, 1996, Creighton, 1992, Branden and Tooze, 1999, Raven *et al.*, 2005).

2.1 Protein Fundamentals

Proteins are essential components of all living organisms. They are referred to as macromolecular polypeptides because they are large molecules composed of amino acids linked together by peptide bonds. There are about 20 different amino acids that occur naturally in proteins, and in a polypeptide they are arranged like wild strawberries on a straw. The sequence of amino acids in a protein is defined by the sequence of the corresponding gene in the organism's genetic code. Each type of protein has a unique sequence of amino acids; this sequence, known as its primary structure, determines the protein's shape and function. Proteins are involved in a wide array of functions within an organism, including metabolic reactions, DNA replication, immune responses, structural and mechanical functions, and molecule transportation. The photosynthetic systems of green plants are also based on a complex battery of proteins.

2.2 Protein Synthesis

In eukaryotes, the DNA sequence of a protein is transcribed inside the cell nucleus into messenger RNA (mRNA) and transported to the cell's cytoplasm. The mRNA sequence is then translated into a peptide sequence by transfer RNA (tRNA) and the ribosome as illustrated in Figure 1. The finished polypeptide chain is then translocated into the endoplasmic reticulum (ER) and

subsequently into the Golgi apparatus, two sites at which the polypeptide can undergo various post-translational modifications to become a complete protein.

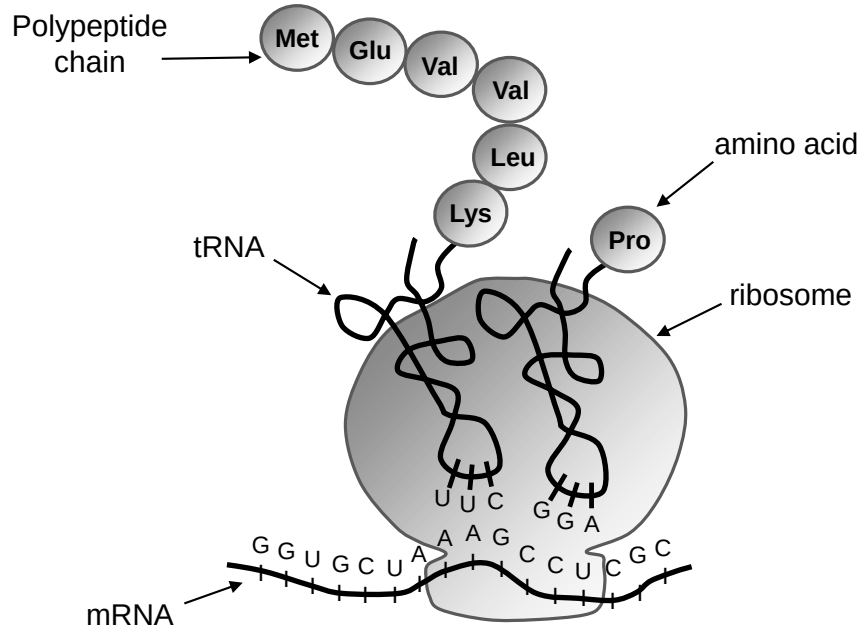

Figure 1. Illustration of the principles of protein synthesis. The ribosome and tRNA translate mRNA into a polypeptide chain. A set of three ribonucleases (codons) in mRNA matches three anticodons in the tRNA, leading to the addition of a specific amino acid to the polypeptide sequence.

The 20 protein-building amino acids encoded by triplet codons in the genetic code are displayed in Figure 2. The different sidechains determine the amino acid properties. These properties classify amino acids as nonpolar, polar or charged. The white panel in Figure 2 shows how amino acids can be linked together by a peptide bond to form the polypeptide chain.

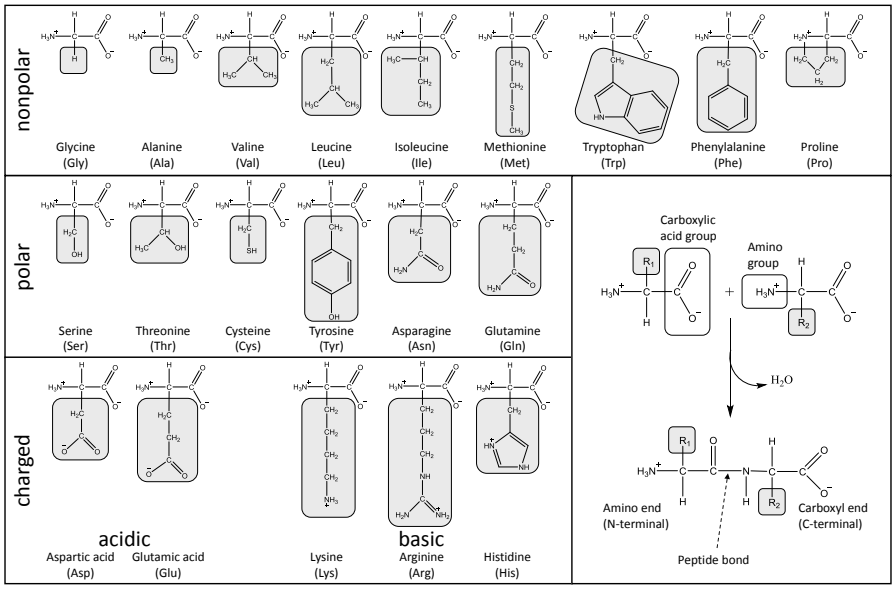

Figure 2. In left panels, the 20 protein-building amino acids are displayed with their sidechains outlined. They are grouped according to their side chain properties, and full name and three-letter code is presented below. The bottom right panel displays the mechanism of peptide bond formation linking amino acids into a polypeptide chain. Amino acids and peptides are always illustrated with the amino end (N-terminal) to the left and the carboxyl end (C-terminal) to the right.

2.3 Protein Structures

The polypeptide sequence is a protein's primary structure, but interactions between different amino acids, their sidechains and their posttranslational modifications give rise to higher structures. The secondary structure is determined by hydrogen bonds between amino acids in the backbone chain. The most common secondary structures are α-helices and β-sheets. Proteins also have an overall tertiary structure, wherein the different secondary structures are arranged in specific ways. This tertiary structure, generally termed a protein's fold or native conformation, can be stabilized by different interactions, including side-chain hydrogen and disulphide bonds. A fourth level of structure (quaternary) exists when several proteins, subunits, are arranged in a complex linked by non-covalent interactions. In such complexes, different subunits can have different functions, as later described for the botulinum neurotoxin complexes.

2.4 Post-Translational Modifications

After a protein has been synthesized, its residues can be chemically modified by post-translational modification (PTM). These can alter the protein's size and its physical and chemical properties, including stability, activity, folding and function. There are over 100 different PTMs, whereof phosphorylation is the most common (Khoury *et al.*, 2011). Many eukaryotic proteins also get carbohydrate molecules (or "glycans") attached to them in a PTM process of glycosylation. This is known to promote protein folding, improve stability and to have regulatory functions (Khoury *et al.*, 2011). Two other very common modifications are peptide bond cleavage, a process in which polypeptides are transformed into mature proteins by removing a part of its sequence, and the formation of disulphide bonds between cysteine residues.

2.5 Toxic Proteins: Structure and Mechanism of Action

A toxin is defined as a poisonous substance produced within living cells or organisms (Dorland's Medical Dictonary). Toxins can be small molecules, peptides or proteins, and may be produced by bacteria, animals or plants. The structure and mechanism of the studied toxins ricin and botulinum neurotoxin are described below.

2.5.1 Ricin

Ricin is a highly toxic 60 kDa glycosylated protein produced in the seeds of the castor oil plant, *Ricinus communis*. In addition to being cultivated extensively for oil production (UN statistics), the plant is decorative as seen in Figure 3, and frequently used in gardening and public parks. Ricin is a type 2 ribosome-inactivating protein (RIP-II), a group that also includes the toxins abrin and viscumin, and kills target cells by disrupting their protein synthesis.

All RIP-II toxins are dimeric and initially produced as single polypeptide chains that are cleaved into two polypeptide chains (A and B) after translation. The B-chain functions as a lectin and has an affinity for galactose-terminated receptors on cell surfaces. The A chain acts as an enzymatic ribosomal inhibitor once incorporated into the targeted cell, interrupting protein synthesis and thereby killing the cell (Endo *et al.*, 1987, Endo and Tsurugi, 1987, Lord *et al.*, 1994, Lord *et al.*, 2003). There are two different ricin isoforms, D and E, which differ in their B-chain sequences. Ricin E was suggested to be a hybrid between ricin D and the co-existing agglutinin described below (Ladin *et al.*, 1987).

The toxicity of RIP-II toxins depends on their mode of administration. The disruption of cell machinery, the actual toxicity, is most lethal when executed

on vital organ cells, including those of the liver and kidney. The ricin and abrin dose required to kill 50% of mice (LD_{50}) was reported to be in the range of 0.2-10 µg/kg when injected or inhaled. Due to the tough protein environment in the stomach, toxicity by ingestion is about three orders of magnitude lower, with LD_{50} values of around 1-20 mg/kg. Consequently, consumption of the toxin content in a handful of seeds (about 5 mg each) is enough to cause significant toxic effects (Schep *et al.*, 2009). Because of ricin's high toxicity and availability, and its historical use in bioweapon programs (Smart, 1997), it has been regulated under the chemical weapons convention. As such, all production, storage or use of ricin must be declared to relevant authorities (CWC). The natural function of ricin's toxicity is to protect the seeds from fungi or other microbes. To prevent toxicity towards itself, the plant uses a strategy of first producing a precursor toxin inside the ER. This proricin is then converted into an active toxin inside protein storage vacuoles via the removal of an N-terminal propeptide and the A-B linker peptide (Frigerio and Roberts, 1998).

There are multiple cultivars of the castor plant that differ in their geographic origin, appearance, and oil content. Unlike ricin E, ricin D has been found in all cultivars that have been studied to date, with an identical protein sequence in each case. Efforts to map ricin's cultivar or geographic origin based on analyses of crude seed extraction procedures that might be used in bioterrorism have revealed differences in the contents of small molecules and other proteins, including ricin E (Despeyroux *et al.*, 2000, Ovenden *et al.*, 2009, Fredriksson *et al.*, 2005, Stern *et al.*, 2016). One frequently co-extracted protein is the related and very similar *Ricinus communis* agglutinin (RCA120). RCA120, possessing orders of magnitude lower toxicity compared to ricin, is a heterodimeric protein with two A- and two B-chains and an intact mass of 120 kDa; its amino acid sequence exhibits approximately 89% identity with that of ricin (Kalb *et al.*, 2015).

Figure 3. To the left, a *Ricinus communis* plant with its big characteristic leaves and spiky bright red seed capsules. The inset shows a split capsule with the embedded seeds. The top right image shows a handful of seeds from various ricin cultivars together with a neckless, popular to create from the beautiful seeds (purchased on Cuba). The bottom right image shows ricin's dimeric structure, with the A-chain in red, B-chain in green, β-sheets as arrows, and α-helices as spirals.

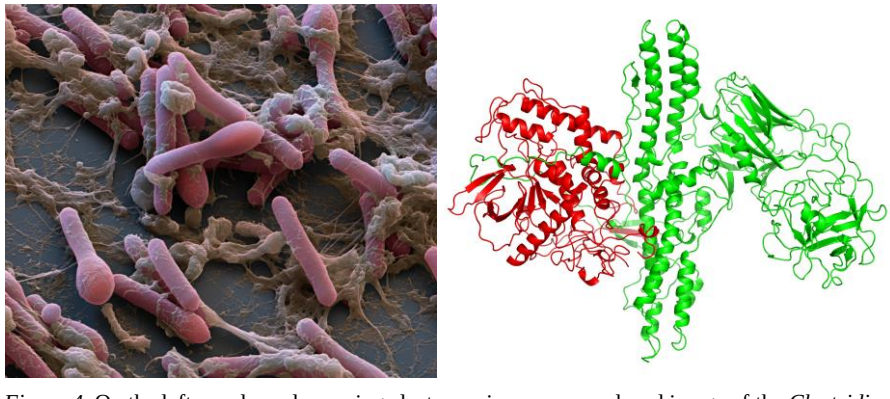

Figure 4. On the left, a coloured scanning electron microscope produced image of the *Clostridium botulinum* bacteria at a magnification of 4000 times (from Science Photo Library with permission). Rod shaped bacteria are in vegetative, growing state, while drumstick shaped are beginning to form spores, a non-reproductive tough and resistant survival state triggered by starvation. On the right, the structure of BoNT/A neurotoxin; the light chain is shown in red, the heavy chain in green, the β-sheets represented as arrows, and the α-helices as spirals.

2.5.2 Botulinum Neurotoxins

The botulinum neurotoxin (BoNT) produced by *Clostridium botulinum* (both displayed in Figure 4) and a few other clostridia strains, is a family of proteins consisting of seven established serotypes termed A to G (BoNT/A-G) (Rossetto *et al.*, 2014). This toxin causes the neuromuscular disease botulism, which is deadly if not treated in time. The toxin acts at the interface between neuronal and muscle cells, where it permanently disrupts the signal system for muscle contraction leading to long term paralysis. This is deadly when affecting vital muscles including those involved in breathing.

There are three major forms of botulism. Foodborne botulism is caused by consuming food containing botulinum neurotoxin, usually from bacterial growth in poorly prepared home-canned food. A more rare kind is wound botulism, caused by toxin produced from *Clostridium botulinum* which has infected a wound. This type is mostly occurring among injection drug users with weakened immune system. The most common type is infant botulism, caused by consumption of botulinum bacteria spores. These are able to colonize the small intestine of infants as their intestinal microflora is not fully developed. Botulinum bacteria then starts to grow in the intestines and released neurotoxin is subsequently taken up into the circulation system.

Lethal amounts of BoNT for adult humans are estimated to 1 µg/kg taken orally, 10 ng/kg inhaled, and 1 ng/kg intravenously, but there is reportedly some variation in LD_{50} between the different human-affecting serotypes (A, B, E, F and G) (Arnon *et al.*, 2001). Despite their extreme toxicity, both serotype A and B are successful, licensed drugs for to treatment of multiple medical and cosmetic conditions characterized by hyperactivity of peripheral synapses (Bigalke, 2013, Simpson *et al.*, 2016).

Like ricin, BoNTs are AB-type proteins that are synthesized as single polypeptide chains and post-translationally modified to form disulphide linked toxins with heavy and light chains (Figure 4). The light chain (LC) represents the enzymatically active part, and the heavy chain (HC) mediates the neurotoxin's binding and uptake into neuronal cells (Rossetto *et al.*, 2014).

The mechanism by which BoNTs mainly interact with their gastrointestinal target cells resemble that of ricin. All BoNT serotypes are secreted along with several other proteins in quaternary complexes of various sizes that are known as progenitor toxin complexes (PTCs). In serotype A, B, C, D, and G, the neurotoxin (NT) can associate into large progenitor toxin complex, L- PTC, containing NT, NTNHA (non-toxic non-hemagglutinin), HA70 (hemagglutinin, size 70 kDa), HA17 and HA33. In serotype E and F only the

medium-sized complex, M-PTC, can be formed containing NT and a partly disrupted NTNHA (115 kDa instead of 130 kDa). Serotype A can form both L- and M-PTC, and the assembled BoNT/A L-PTC complex is shown in Figure 5. These associated proteins protect the complex against degradation in the harsh environment of the stomach and confer multivalent sugar affinity that is important for its uptake by the epithelial cells of the small intestine (Gu *et al.*, 2012, Rummel, 2015). It was also proposed that the primary role of NTNHA is to protect NT from proteases and denaturation in decaying biological material rather than protection in the gastrointestinal tract (Rossetto *et al.*, 2014). After entering the intercellular space, complexes dissociate via pH-induced conformational changes and released neuro-toxins proceeds to target neuronal cells (Lam and Jin, 2015).

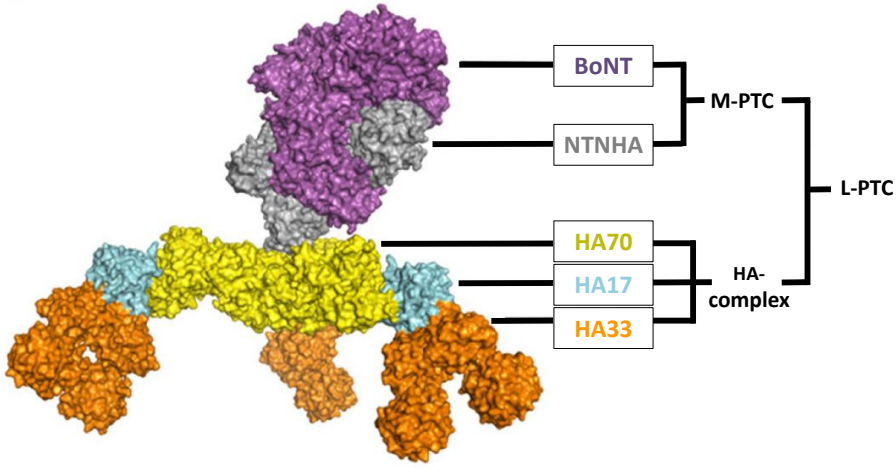

Figure 5. The BoNT A L-PTC complex consisting of one NT, one NTNHA, three HA70, three HA17 and six HA33 proteins. The six HA33 proteins have affinity for terminal galactoses and facilitates intestinal toxin uptake by a multivalent action. Figure adapted from (Yao *et al.*, 2014) with permission.

A major difference between BoNTs and ricin relates to the way the former interact with target cell molecules. BoNTs selectivity against neuronal cells is mediated by two different neuron-specific receptor molecules. Gangliosides in the synaptic cell membrane accumulate neurotoxin molecules. This enrichment greatly facilitates the subsequent interaction with a second receptor which then leads to the toxin's uptake into small synaptic vesicles (Rossetto *et al.*, 2014). The second receptor for BoNT/A is the luminal domain of synaptic vesicle glycoprotein 2 isoforms A-C (SV2A-C). This domain becomes extracellularly exposed as synaptic vesicles fuse with the synaptic membrane and release their contents as shown in Figure 6 (Janz and Sudhof, 1999).

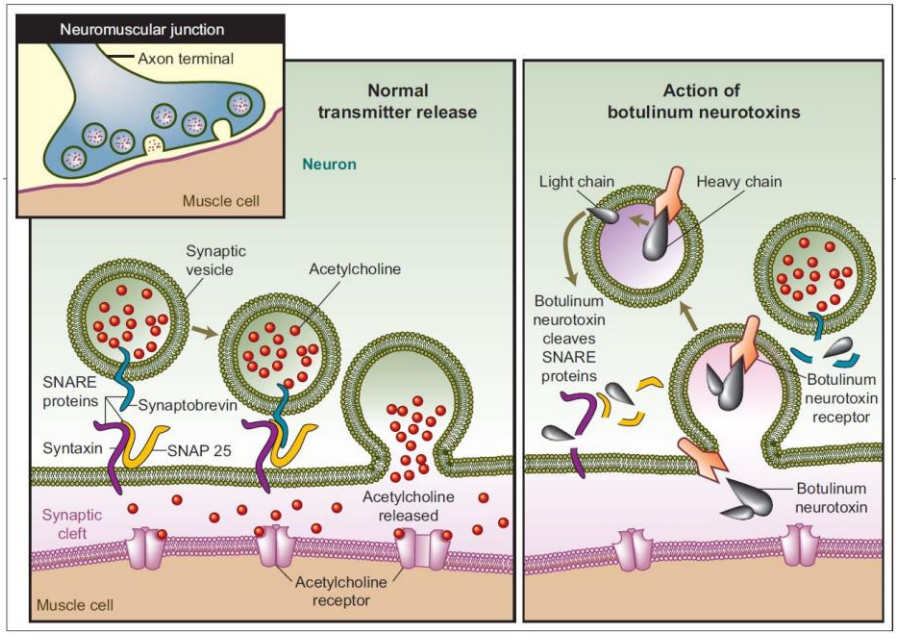

Figure 6. The left panel shows the normal release of the neurotransmitter acetylcholine to induce muscle contraction. The right panel shows how the neurotoxin's LC cleaves vesicle receptors, disabeling acetylcholine release and hence muscle contraction. Picture from (Dickerson & Janda, 2006) with permission.

Recently it was shown that asparagine glycosylation in these luminal domains promote BoNT/A binding and uptake (Yao *et al.*, 2016, Mahrhold *et al.*, 2016). After uptake into the synaptic vesicles, the neurotoxin dissociates and the light chain exits the vesicle to specifically cleave proteins involved in the fusion of synaptic vesicles to the synaptic membrane. This prevents the release of acetylcholine at the neuron-muscle interface, leading to paralysis of the affected muscle and hence the symptoms of botulism (Rossetto *et al.*, 2014). The general mechanism of this process is shown in Figure 6.

2.6 Sample Preparation

Sample preparation refers to the way in which a sample is processed before being subjected to a specific analysis. In protein studies, particularly those involving mass spectrometry (MS), sample preparation is a crucial step. Many chemicals are incompatible with MS, and sample matrix components could cause signal suppression or other instrumental problems. A large part of this thesis deals with development of sample preparation methods, so no further discussion of the broader field will be provided here. For more information on

this topic, I recommend the reviews of (Guerrier and Boschetti, 2007, Wang *et al.*, 2014, Olszowy and Buszewski, 2014, Li and Franz, 2014).

2.7 Proteolytic Digestion

A key step when performing peptide mass spectrometry for protein identification and characterization is the transformation of target proteins into peptides, which is known as proteolytic digestion. All of the sample preparation methods described in the papers appended to this thesis involve digestion, and it was found that the quality of the results obtained depended strongly on the digesting enzyme's efficiency and specificity, together with the digestion time.

An analytical procedure may be performed to identify a specific protein, to elucidate details of its modification or structure, or (in a proteomics project, for example) to detect and possibly quantify multiple proteins simultaneously. Regardless of the objective, the transformation of target proteins into peptides is a critical step. If the proteolytic digestion is ineffective, the sensitivity of the analysis will decline steeply: in principle, an analysis in which the target protein is completely degraded to the desired peptides (corresponding to a 100% yield) will be 20 times more sensitive than one in which the yield is only 5%. Higher efficiencies can be achieved by using high quality enzyme batches with high purity and stability that achieve a high number of enzymatic reactions (cleavages) per time unit.

Trypsin is a widely used proteolytic enzyme with well-defined target motifs. It cleaves with high specificity on the carboxyl side of the basic amino acids arginine and lysine (Vandermarliere *et al.*, 2013). These properties make it an excellent tool in peptide mass spectrometry because they ensure that every newly formed peptide terminates with a positively charged amino acid, which facilitates ionisation. Moreover, its high specificity (which gives high reproducibility) makes digestions predictable, and also prevents bias among the resulting peptides because it, together with efficiency, prevent target sequences or amino acids from being "diluted" in terms of signal intensity. Additionally, trypsin specificity is almost independent of the reaction time, temperature, or enzyme concentration, making trypsin well suited for digestion protocol development.

The kinetics of enzymatic reactions are described by the Michaelis-Menten (MM) equation (Figure 7), and the influence of trypsin digestion kinetics on proteomic coverage was recently analysed by Fonslow and co-workers (Fonslow *et al.*, 2013). At low substrate concentrations ([S]), the reaction rate (v) has a linear relationship to [S], but as [S] increases, enzyme performance

increases and reaction kinetics approaches maximum rate, (V_{max}). The importance of substrate concentration can be visualized by presenting the equation graphically as in Figure 7. In keeping with the predictions of the MM equation, higher substrate concentrations yielded faster digestion reactions in all of the four studies summarized in this thesis.

$$v = \frac{V_{max}\,[S]}{K_M + [S]}$$

v = reaction rate

V_{max} = maximum reaction rate

[S] = substrate concentration

K_M= The Michaelis-Menten constant is substrate concentration at which the reaction rate is half V_{max}

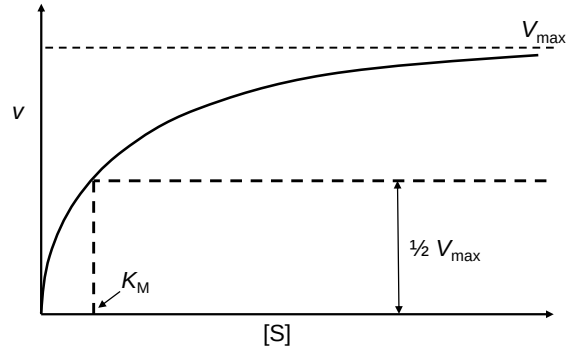

Figure 7. The kinetics of the digestion of proteins into peptides by proteases are modelled by the Michaelis-Menten equation: the digestion rate (v) is a function of the substrate concentration ([S]), the maximum reaction rate (V_{max}), and the Michaelis-Menten constant (K_M). The resulting graph clearly indicates that high substrate concentrations yield high digestion rates: as [S] increases, v rises towards V_{max}.

The speed of proteolytic reactions is somewhat temperature dependent: modest temperature increases lead to higher reaction turnover per time unit. Higher turnover rates enable shorter digestion protocols and hence shorter total sample preparation times. This is particularly desirable in the context of bioterrorism and biocrime analysis, where rapid responses are preferred. Shorter total sample preparation times also reduce the risk of introducing side-effects and undesired reactions, such as deamidation reactions (further described in section 5.2 and in paper I), which are known to be time- and temperature-dependent (Ren *et al.*, 2009).

2.8 LC-MS and CE-MS Analysis of Proteins

Mass spectrometry (MS) is very widely used in protein and peptide analysis, and mass spectrometric methods are under constant development. Separation techniques such as liquid chromatography (LC) and capillary electrophoresis (CE) can be used in conjunction with many different kinds of mass spectrometers, giving access to a great variety of analytical systems with different features. The extent of this field will not be covered in this thesis, but it has been comprehensively reviewed in several recent publications: (Aharoni

and Brandizzi, 2012, Lu *et al.*, 2008, Wei *et al.*, 2013, Mikami *et al.*, 2012, Yates *et al.*, 2009, Wang *et al.*, 2014). The most important characteristics of the CE- and LC-MS setups presented in this thesis are described in the experimental section and additionally in each paper.

2.9 Protein Forensics

In peptide mass spectrometry, target proteins are identified by analysing the molecular masses or amino acid sequences of their derived peptides. The former approach, which is known as peptide mass fingerprinting, will not be discussed further in this thesis but has been reviewed at length by multiple authors (Cottrell, 1994, Thiede *et al.*, 2005).

Criteria for sequence level (product ion) identification has been used since first suggested in the late 70's on confirmation of animal drug residues (Sphon, 1978). The development of increasingly powerful and capable mass spectrometers has been a major driving force behind criteria development because the use of instruments with higher resolutions and mass accuracies generally reduces the number of product ions that must be detected to achieve a given level of identification. This is exemplified by drug testing in sports, where the use of higher mass accuracy instruments increases the number of identification points achieved per product ion, allowing specified identification criteria to be achieved with fewer product ions (Thevis *et al.*, 2007).

In the context of monitoring compliance with the Chemical Weapons Convention, efforts to reliably identify ricin by peptide mass spectrometry have been complicated by the frequent co-occurrence of the closely related (but not regulated) protein RCA120, which is almost identical to ricin at the sequence level as noted previously. This sequence similarity necessitates careful peptide selection (Fredriksson *et al.*, 2005). A set of peptides unique to ricin has been identified, and it has been suggested that the presence of ricin can be identified unambiguously on the basis of sequence information from at least two of these peptides (Kalb *et al.*, 2015). In sports drug testing, unequivocal protein identification is based on both the uniqueness of the detected peptides, which must be demonstrated, and a total protein sequence coverage of at least 10% (WADA, 2010). A similar criterion has been proposed for CWC-related analysis because the sequence specificity of species such as ricin can change depending on the number of identified sequences of related species and the continuing expansion of protein databases.

3 Objectives

The overall aim of the work presented in this thesis was to develop improved methods for the chemical analysis of toxins likely to be used in biocrime and bioterrorism. More specifically, the objective was to develop improved mass spectrometry methods for peptide based protein identification and characterization.

The first study conducted along these lines sought to combine in-silico prediction, a thorough peptide investigation, and a broad analytical approach to study the extent of deamidation in ricin. An additional objective was to determine whether deamidation analysis could be useful in the forensic matching of ricin samples to sources. (Paper I)

The second study focused on developing a chromatography-based affinity method for sample preparation with subsequent identification through peptide mass spectrometry in order to improve the capacity for detecting ricin and other RIP-II toxins. (Paper II)

The third study utilized the natural affinity of BoNT L-PTC for galactosyl moieties and explored the possibility of applying the methods described in paper II to this toxin complex. An additional objective was to develop an NT-specific method by activating the natural pH-triggered NT release mechanism on a column format. (Paper III)

The final study in this work was conducted to determine whether a broad approach to analysing SV2C, a BoNT intoxication related protein, could reveal the effect of different glycosylation patterns on interactions with BoNT A at neuromuscular junctions. By applying a range of collision energies in MS analyses of different SV2C sequence mutants, we also sought to characterize each glycan's structure and heterogeneity. (Paper IV).

4 Experimental

This section describes the two mass spectrometry setups used in the studies presented in the appended papers. Practical details relating to sensitive protein identification through peptide analysis are also highlighted.

4.1 Capillary Electrophoresis

Capillary electrophoresis (CE) is a technique for separating various analytes, including proteins and peptides. The technique is using electric fields to separate different analytes based on their charges and capacity to migrate through an electrolyte solution under an applied electrical potential. The most widely used CE technique (and that used in paper I) is capillary zone electrophoresis (CZE), which employs narrow bore open tubular fused silica capillaries with lengths of 40-100 cm. These are filled with a suitable solvent and a small sample plug is injected, either hydrodynamically (by pressure) or electrokinetically (by potential). An electrical field (typically of 10-30 kV) is then applied across the capillary to drive separation. The CZE-MS setup presented in paper I used an applied potential of 350 V/cm; due to the MS hyphenation, separation was performed in a mixture of volatile solvents acetonitrile and formic acid. The analytes that migrate most rapidly through the capillary towards the cathode will be the smallest (i.e. those with the greatest mobility) and the most positively charged. Because the basis of separation by CZE differs from that for LC, the two techniques are complementary.

A big advantages of using CZE rather than LC separations, especially in peptide analysis, is the ability to separate and detect very small peptides and even single amino acids, that are not retained in reversed phase LC. The most apparent drawback is the difficulty of automating CZE-MS setups, especially the reconditioning and voltage onset steps. Figure 8 illustrates the principles of CE and the specificities of the CZE-MS setup described in paper I.

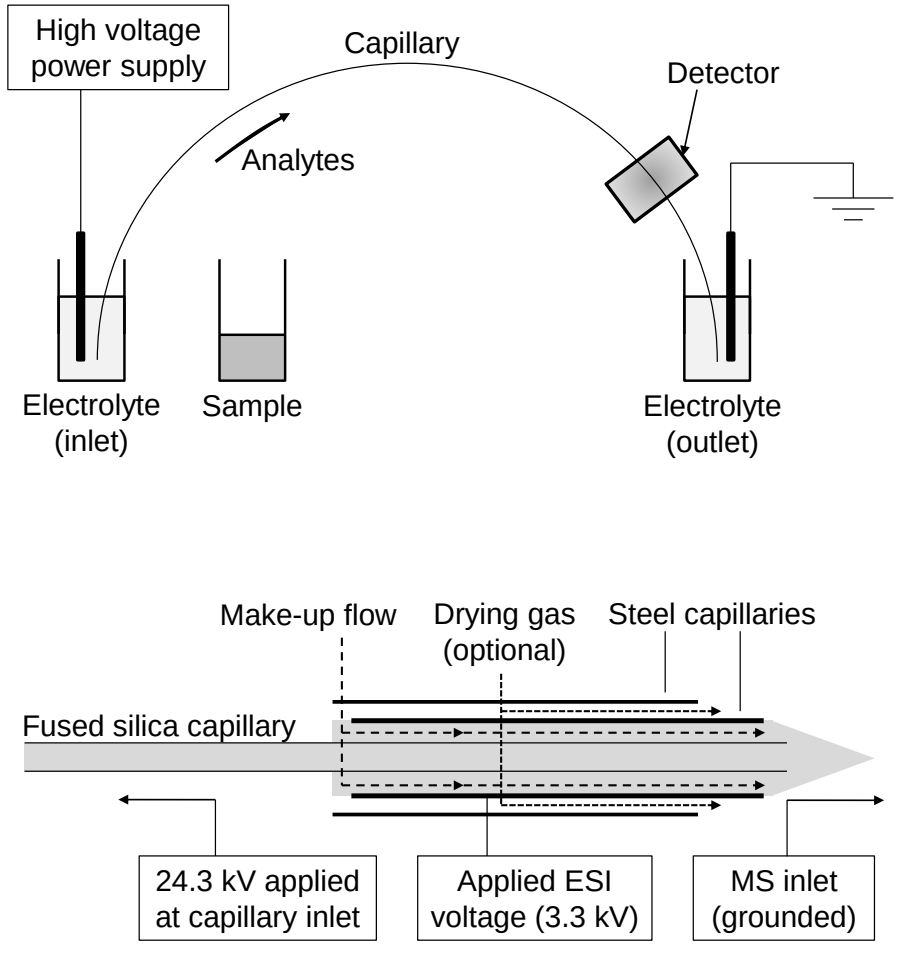

Figure 8. Schematic picture of capillary electrophoresis (CE) separation setup on top. A sample plug is injected from the sample vial, and the separation of injected analytes is driven by the electrical potential between the inlet electrolyte and the grounded outlet electrolyte. Below is the CZE-MS sheath flow interface and applied potentials from paper I, generating a separation potential of 21 kV. The methanol and acid containing sheath flow stabilizes the electrospray function.

4.2 Liquid Chromatography

In liquid chromatography (LC), analytes are separated based on their interactions with a column's stationary phase and the solvent (mobile phase) that is used. The stationary phase is generally hydrophobic and so interacts most strongly with non-polar compounds; this is referred to as reversed phase chromatography. However, hydrophilic stationary phases that interact most strongly with polar compounds (known as normal phase chromatography) are

also common, as are affinity phases that interact specifically with more complex chemical moieties on the analytes to be separated. The latter approach is discussed at greater length section 5.1 and in papers II and III. In reversed phase setups in general, and LC-MS setups in particular, the most frequently used stationary phase is C18, an alkyl bonded silica phase that interacts most strongly with hydrophobic analytes including peptides (Krokhin and Spicer, 2009).

All of the peptide-based mass spectrometry work presented in this thesis was performed using a C18 nano-UHPLC (ultra-high pressure LC) setup. In brief, this means that a mobile phase flow in the nanoliter range was used over a 150 mm x 75 µm analytical column packed with sub-2 µm sized C18 particles, enabling high resolution and high sensitivity peptide separations. The packing material affects the LC column's resolution: smaller particles lead to tighter column packing and thus better resolution. However, small particles also generate higher backpressures, necessitating the use of an UHPLC pump. The column length and diameter also affect performance because the number of analytes that can be separated (the plate number) is proportional to the column length, and the sensitivity is inversely proportional to the square of the column's cross sectional area (Rieux *et al.*, 2011, Köcher *et al.*, 2011). The setup also includes a small trap column, typically 10-20 mm long and with slightly larger inner diameter than the analytical column. The trap column is connected to the system via a flow valve, enabling it to be switched off-line. This allows an elevated flow rate to be used during initial sample loading. Samples are injected and loaded onto the trap-column using a high flow of 3% isocratic acetonitrile. This initial acetonitrile content increases system stability without lowering peptide detection ability. The main drawback of using high resolution and high sensitivity columns is the total analysis time. A sample-to-sample time of around 45 minutes was necessary to permit sample trapping, the application of a suitable separation gradient, and column washing and regeneration. Still, non-targeted analysis of complex samples could gain from a slower increase in the LC gradient, creating much longer sample-to-sample times. The LC time program from paper III is visualized in Figure 9, with the different chromatography steps indicated.

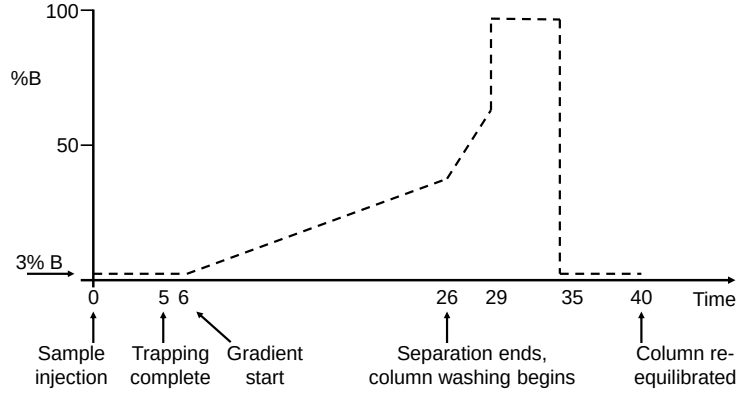

Figure 9. The LC time program used in paper III. Samples were injected at time zero and thereafter loaded onto the trap-column using a high flow of 3% isocratic acetonitrile (B). The analytical column is switched in-line after 5 minutes and the peptide separating gradient goes up to 40% B in 20 minutes. Column washing and re-equilibration is important for reproducibility. This program generated a total sample to sample time of 45 minutes.

4.3 Mass Spectrometry

4.3.1 Electrospray Ionisation

The field of mass spectrometry has evolved rapidly since the invention of electrospray ionisation, ESI, source in the 1980's (Fenn, 2002). ESI is a soft ionization technique because it generates charged analytes with relatively low internal energies and thereby does not induce structural changes in the ionized products. Ionization is achieved by spraying a solution of the analytes through a needle-type tip (emitter) under the influence of differences in electrical potential between the emitter and the mass spectrometer. The resulting spray is known as a Taylor cone (Taylor, 1964). In positive mode ESI, the solvent is an acidic mobile phase (we used a water/acetonitrile gradient acidified with formic acid), which results in the formation of positively charged droplets at the tip of the Taylor cone. These droplets move towards the low potential (grounded) instrument, and repeatedly split into smaller and smaller sub-droplets due to evaporation and analyte repulsion within the droplet, as illustrated in Figure 10.

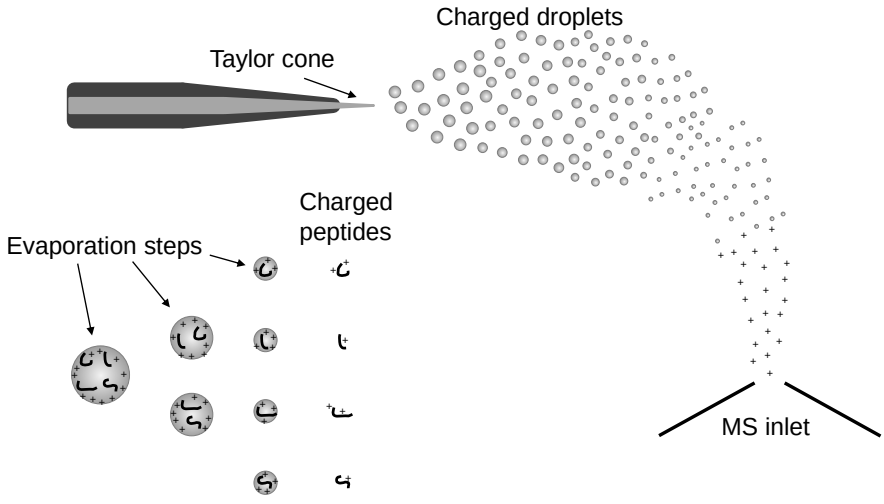

Figure 10. The principle of electrospray ionisation, ESI: The Taylor cone emits multiple droplets towards the low potential (grounded) instrument. Due to evaporation and charge repulsion they repeatedly split into smaller and smaller drops, eventually creating charged analytes (peptides) that may enter the MS inlet.

As the droplets shrink they become more highly charged, until eventually the embedded protonated peptides enters the gas phase (Herbert and Johnstone, 2002). The use of nanoscale ESI increases the sensitivity of the analysis because the use of low flow rates reduces the size of the droplets leaving the spray emitter and thus promotes gas phase desorption (Abian *et al.*, 1999). Peptides entering the gas phase typically carry multiple charges because their basic amino acids and sites - i.e. the N-terminal amine and the side-chains of lysine (K), arginine (R), and histidine (H) – all attract charges

4.3.2 Mass Detection

Two different Q-TOF mass spectrometers were used in the MS analyses discussed in this thesis. A Q-TOF instrument is a hybrid or tandem mass spectrometer that features both triple quadrupole and Time-of-Flight (TOF) mass analysers. The first mass analyser, the quadrupole (Q) part, acts as an active filter and collision cell. When operated in filter mode, settings enable ions with specific masses (i.e. mass-per-charge, *m/z*) to pass through to the TOF analyser while excluding others. The TOF component operates under near-vacuum conditions and analytes entering are given kinetic energy in the pusher that is inversely proportional to their *m/z* value The TOF then separates analytes based on their drift tube flight times (Mirsaleh-Kohan *et al.*, 2008). In the flight tube, small ions with low *m/z* travel more quickly than high *m/z* ions.

The longer the flight tube, the better the separation and the resolution increases. The resolution can be further increased by introducing reflectors into the flight length, as illustrated in Figure 11.

When operating a Q-TOF instrument in tandem mode (MS/MS), one selects a precursor ion and fragments it in the quadrupole part, and then analyses the resulting fragments (product or daughter ions) in the TOF part. Fragmentation is achieved by collision induced dissociation (CID). In CID, peptides are fragmented by collision with inert gases (N_2, He or Ar); higher collision energies yield more fragments.

Peptides primarily fragment along the backbone, allowing their sequence to be read from the fragments (Roepstorff and Fohlman, 1984, Steen and Mann, 2004, Kinter and Sherman, 2005). When the precursors are glycosylated peptides, as in paper IV, the attached glycans are brittle and so fragment first. Fragmentation starts occurring at rather low collision energies (CE), so structure elucidation can be achieved through stepwise increases in the CE. As the CE is increased, N-glycosylated peptides are gradually stripped of all their glycans (except for the last asparagine-bound GlcNAc moiety) before the peptide backbone starts to break.

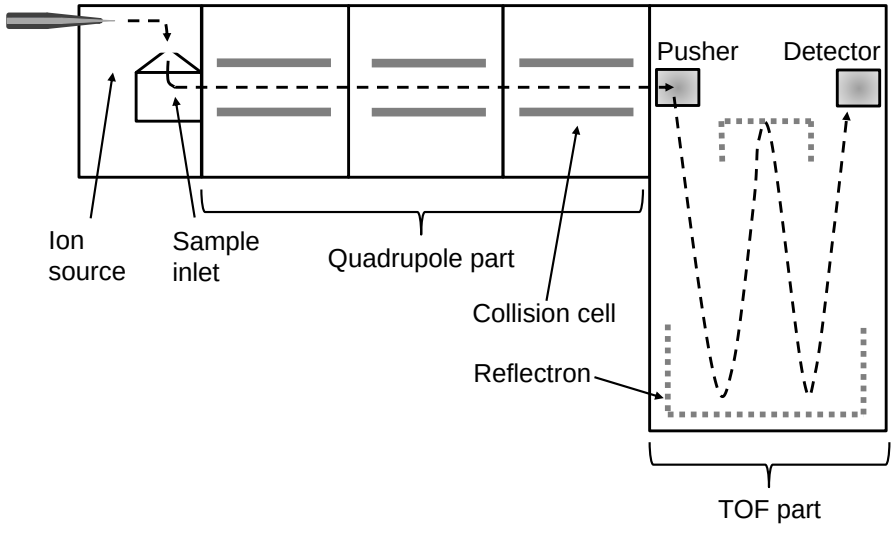

Figure 11. Schematic overview of a Q-TOF mass spectrometer. The quadrupole part acts as an active mass filter and collision cell. Fragments or intact analytes entering the TOFs pusher are given kinetic energies and are then separated based on their difference in drift tube flight time. Separation can be increased by reflectors in the flight tube increasing the total flight length. The flight time is registered at the detector and translated into mass-per-charges, *m/z*.

In MS/MS mode, the precursor ion can be selected manually or automatically. In automatic data dependent acquisition, which is frequently used in proteomics studies, the most intense ions at each time slot are selected for MS/MS analysis. Depending on the desired or undesired masses that are present (e.g. peptides or matrix components), this selection can be tampered with using include or exclude *m/z* lists. Ions that are both included and detected are selected for MS/MS. In the MS/MS method development of paper III, the include list approach was initially tested to increase sensitivity, but did not provide the low detection levels that were sought. Automatic data dependent MS/MS analysis generally offers limited performance when dealing with low level samples in complex background. Sample matrix components then supress and mask target peptide signals, preventing any identification based on detectable *m/z* signals. In recent years, a new type of approach suitable for dealing with samples of this kind has evolved, taking advantage of the technical development of high resolution full scanning mass spectrometers, such as QTOF instruments. This approach, termed PRM (parallel reaction monitoring), was first defined in 2012 by Peterson and co-workers (Peterson *et al.*, 2012) and is a targeted strategy where full product ion spectra of specific precursor ions are generated. As demonstrated in paper III, the instrument is set to alter between a set of precursor ions throughout a time segment. Consequently, peptides with a targeted *m/z* value can be collided to generate TOF-detectable product ions even if they are 'hidden' behind co-eluting matrix components or peptides. The co-eluting components fragment simultaneously, but specificity is achieved by selecting multiple specific fragment ions from the full scan data (Rauniyar, 2015, Schilling *et al.*, 2015, Dupré *et al.*, 2015). Also a recent comparison of different mass spectrometry setups for targeted MS/MS highlighted the benefits of using high resolution full scanning instruments to successfully combine sensitivity and specificity (Mbasu *et al.*, 2016).

In addition to its sensitivity, specificity, and high mass accuracy, the PRM method is also easily changeable. Adding or removing parent or daughter ions, changing time segments, or even switching to data-dependent full scan analysis can all be done easily on a run-to-run manner. In biocrime- and bioterrorism-related analyses, this allows samples to be rapidly analysed for their general peptide content before performing a specific and sensitive toxin-targeting PRM analysis. The principle of PRM is illustrated in Figure 12.

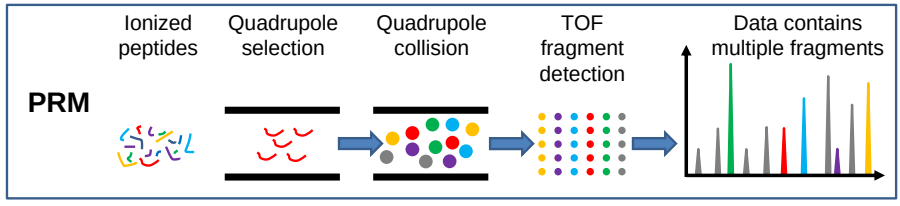

Figure 12. Principle description of parallel reaction monitoring (PRM). Full product ion spectra of selected precursors are generated, offering both specificity and flexibility in daughter ion monitoring.

4.4 Combining Sensitivity and Specificity

When developing a sensitive protein identification method, all of the steps in the protocol must be developed accordingly. In the peptide mass spectrometry step, sensitivity must be combined with specificity so that the relevant identification criteria are satisfied. To ensure that this was achieved, the following approach was adopted in paper III. Starting with a relatively concentrated stock solution of the pure target protein, serial dilutions were created to give solutions with concentrations ranging from 2-3 orders of magnitude above the LC-MS instrument's presumed detection limit, down to 2-3 orders of magnitude below the limit. Samples were then denatured and precipitated using organic solvents, and trypsin digestion buffer was added after evaporation. To ensure complete digestion, all samples were digested with a quantity of trypsin sufficient to achieve complete digestion of the most concentrated sample. After terminating the reactions, equal volumes of each sample were injected and analysed on the LC-MS system. The results obtained for the most concentrated samples were then taken to represent the greatest possible range of tryptic peptides that could be generated from the toxin. These peptides were evaluated to determine their specificity and collective sequence coverage. The evaluation of these initial results also indicated whether it was possible to retrieve protein sero- and subtype information. BoNT serotypes differ significantly in sequence, but BoNT/A subtypes differ by only a few peptides (Kull *et al.*, 2015). Although we only used BoNT subtype A1, we were still eager to identify peptides bearing subtype information. The results for the more dilute samples were then evaluated to see which peptides that supported highly sensitive analysis by remained identified, and which were gradually lost due to low signal intensities

5 Results and Discussion

5.1 Natural Mechanism Based Affinity Enrichments

In the work with recombinant SV2C proteins presented in paper IV, a specific constructed affinity tag was incorporated into the polypeptide sequence. The eight amino acid *Strep*-tag® II sequence (Trp-Ser-His-Pro-Gln-Phe-Glu-Lys), one of the most common affinity tags used in life science, generates a special motif with high affinity for streptavidin structures, including the developed *Strep*-Tactin® (Schmidt and Skerra, 2007). All samples used in paper IV were purified using this incorporated tag and the purification principle is illustrated in Figure 13.

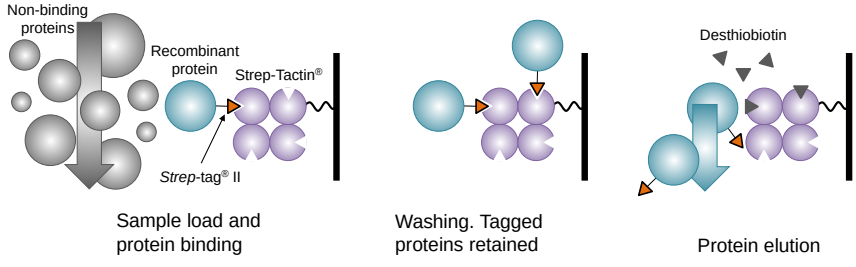

Figure 13. The *Strep*-tag® II peptide sequence exhibits affinity towards engineered streptavidin (*Strep*-Tactin®). Elution of enriched proteins are done using biotin structures, replacing bound proteins.

Another approach is to make use of natural mechanisms that do not rely on recombinant tools, as is done in the protocol developed in paper II and applied in paper III. As part of their natural mechanisms of intoxication, both BoNT complexes and RIP-II toxins exhibit strong affinity for carbohydrate structures that are exposed on targeted cell surfaces. Such mechanism-linked natural affinities have previously been exploited in a variety of ways for different

proteins (Lyimo *et al.*, 2012, Jimenez *et al.*, 2013, Sun *et al.*, 2005). A natural mechanism, developed through the selective pressure of evolution, is thereby adopted to its role. Although such a mechanism may be well suited to its function in vivo, to be useful analytically it must be successfully transferred to an in vitro system. Various column materials have been developed for use with carbohydrate binding proteins (lectins) and are commercially available. We routinely use these materials for preparative scale purification of castor bean extracts. The challenge we faced in paper II was to create a pressure-tolerant miniaturized column system that enabled high sample throughput, and offered greater sensitivity than existing alternatives.

During the method's development, the selected galactosyl-modified Poros® material, Figure 14, was found to successfully bind ricin. However, desorption of the bound ricin using lactose, which is a standard technique in preparative purifications, yielded unsatisfactory results. Lactose, galactose and similar structures containing a terminal β1-4-linked galactosyl moiety all exhibit affinity for ricin's carbohydrate binding sites, so injecting a competitive amount of one of these sugars onto the column system should release the bound toxin. However, when developing a natural mechanism based approach, it is important to reflect on the mechanism's purpose and origin. The cell surface-binding RIP-II toxins are not released from the glycans that they recognize after attachment; instead, they are taken up into the cell and transported between its compartments, with their two chains ultimately being separated inside the Golgi apparatus. Consequently, strong (and not readily reversible) binding to carbohydrates is beneficial for their evolved purpose, which may be why they are not successfully displaced from bound sugars using lactose. Therefore, an alternative approach was tested. Acidic conditions have been shown to change the three dimensional structures of ricin in parallel with a decrease in galactoside affinity (Frenoy, 1986), so the column was washed with a low pH (<3) solution resulting in a rapid and complete ricin desorption. Acidic desorption was tested with a range of different lectins, including crude ricin, crude abrin, viscumin and peanut agglutinin (PNA). All lectins except PNA bound strongly to the galactosyl material, and acidic desorption was successful in all cases. It is likely that PNA was poorly retained because all of the other tested lectins have high affinities for the Galβ1-4 structures with which the column was modified, whereas PNA has an affinity for Galβ1-3GalNAc structures.

The subsequent sample analysis was favoured by the low pH desorption, as it only introduced the mass spectrometer compatible chemicals as TFA and methanol. The constructed miniaturized column based on peek tubing packed with galactosyl-modified Poros® material is displayed in Figure 14.

Conclusively, the developed affinity method resulted in a successful enrichment of galactosyl binding lectins.

Figure 14. The miniaturized affinity columns were based on Poros® chromatography material (b), modified with a galactose terminated ligand (c), and packed into conventional peek tubing (a).

After having developed and evaluated the miniaturized column setup with the galactosyl-modified Poros® stationary phase for RIP-II toxins, its use in the separation and analysis of botulinum neurotoxin was investigated (paper III). This was because the large progenitor complex of BoNT also has an affinity for terminal galactose structures through its associated HA33 proteins (Lee *et al.*, 2013, Yao *et al.*, 2014). While recent antidote experiments have indicated that it has a general affinity for a range of carbohydrate structures, terminal galactose moieties are the natural intestinal target of the large progenitor complex, L-PTC (Lee *et al.*, 2015).

Experiments using non-associated HA33 proteins and a HA complex verified the existence of a special multivalent binding feature of L-PTC complex, which is known as the velcro effect (Zopf and Roth, 1996, Arimitsu *et al.*, 2008). Because non-associated HA33 proteins did not bind to the column material, but the HA complex (holding six HA33 subunits) and the complete L-PTC complex did, this verified the need for a multivalent galactose binding process to associate L-PTC to galactosyl exposing surfaces. In paper II we showed that it can be difficult to exploit ricin's full galactose affinity when other carbohydrates are present. The studies presented in paper III did not reveal any affinity problems caused by BoNTs attraction to matrix sugars. It seems likely that the velcro effect is enabling efficient BoNT L-PTCs affinity enrichment as competing sugar structures caused no significant enrichment problems.

In addition to their glycan binding mechanism, BoNT progenitor complexes also exhibit a toxin release mechanism: an increase in pH, as follows when entering a host's circulation system, induces conformational changes in NTNHA that release the associated neurotoxin (Gu *et al.*, 2012, Eisele *et al.*,

2011, Matsui *et al.*, 2014). This phenomenon was exploited in our developed experimental protocol, in which a high pH borate solution is injected onto the column bearing the bound L-PTC to release the NT from the complex (Figure 15). Paper III describes the use of this approach for toxin detection in crude culture media.

After specifically triggering NT release in this way, the remaining bound complex was eluted using the low pH desorption protocol presented in paper II. In contrast to the two-step method developed for RIP-II method, in which the first step was based on the toxin's natural mechanism of action, the BoNT enrichment protocol takes a three-step approach in which the first two steps exploit the toxin's natural mechanism. In both protocols, the final step is a low pH injection to release all remaining bound proteins from the column, including components of the original toxin complex and any other galactosyl-bound proteins. Overall, the affinity method developed in paper III enabled robust and specific enrichment of NT, resulting in highly sensitive MS analyses even when the toxin is present in a complex background. Moreover, the protocol is compatible with other protocols for enriching and identifying other species such as RIP-II or other galactosyl binding proteins.

5.1.1 Affinity, Toxicity and Milk

Several papers and reports have discussed the positive effect of administering milk or other solutions with a high lactose or galactose content after exposure to or intoxication by ricin (Nagatsuka *et al.*, 2010, Lumor *et al.*, 2013, Rasooly *et al.*, 2012). In addition, it was concluded that pouring a lactose containing solution into a ricin exposed eye could prevent toxic effects (Strocchi *et al.*, 2005). No such protective effect is expected in cases of botulinum neurotoxin poisoning because multiple in vitro experiments with BoNT have demonstrated that the progenitor complex's HA proteins bind strongly to various carbohydrates without in any way preventing intoxication in vivo (Lee *et al.*, 2015).

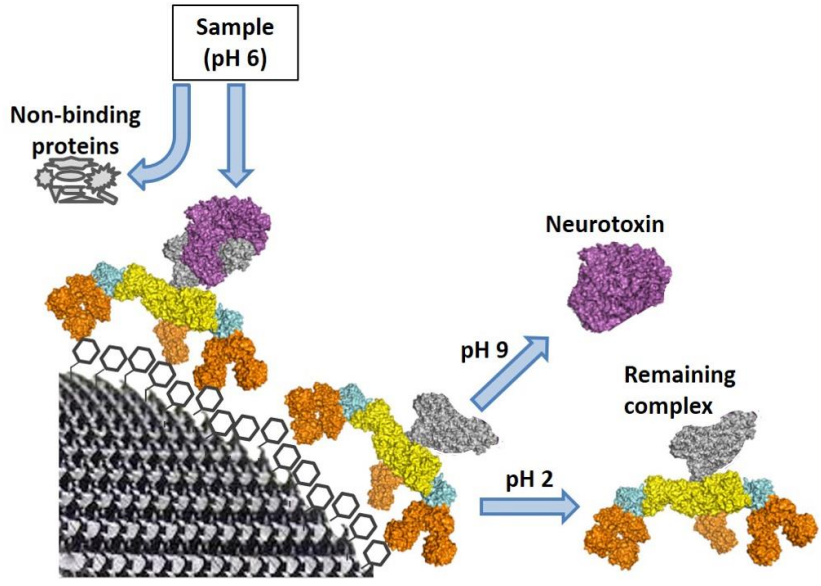

Figure 15. Schematic display of the progenitor complex chemistry relocated to a galactosyl modified chromatography material. BoNT samples are loaded at pH 6 keeping the toxin associated in complex form. Multivalent interactions, here illustrated by the velcro surface, efficiently enrich galactose binding progenitor complexes. The pH induced neurotoxin release mechanism is triggered by a high pH borate solution, successfully releasing complex bound neurotoxin. The subsequent low pH injection efficiently releases remaining complexes. Figure adapted from (Yao *et al.*, 2014) with permission.

Based on the results reported in paper III, this finding can be attributed to the strength of the velcro effect; the individual affinity of the HA33 proteins for free carbohydrates is not strong enough to overcome the complex's strong multivalent binding to the precisely oriented carbohydrates displayed on the surfaces of the epithelial cells of the small intestine. The importance of this combined action was also stressed in a recent paper on the development of a BoNT antidote (Lee *et al.*, 2015). The inability of lactose to suppress the toxic effects of botulinum neurotoxin has also prompted concerns about a potential bioterrorism attack on the US milk supply chain (Wein and Liu, 2005). Because lactose binds competitively to ricin, this toxin could not be used in such an attack; if poisoning by ricin or another RIP-II toxin is suspected, it would be advisable to give the victims milk to drink as a readily available and easily administered antidote.

5.2 Turning Proteins Into Peptides

5.2.1 The Denature-Through-Precipitation Approach

In the deamidation work (paper I) a short digestion time was necessary, because we only wished to monitor protein-level deamidation without any contribution from deamidation occurring during digestion.

Deamidation is a posttranslational process that converts asparagine to aspartic and isoaspartic acid. This transformation alters the charge and can affect protein functionality and the standard model for the reaction is outlined in Figure 16. The mechanism and background of the deamidation study is further described in paper I, but by mapping its existence in ricin, we wanted to conclude if there were differences between different cultivars as well as between different preparations, and if this difference could contribute in forensic sample-to-source matching.

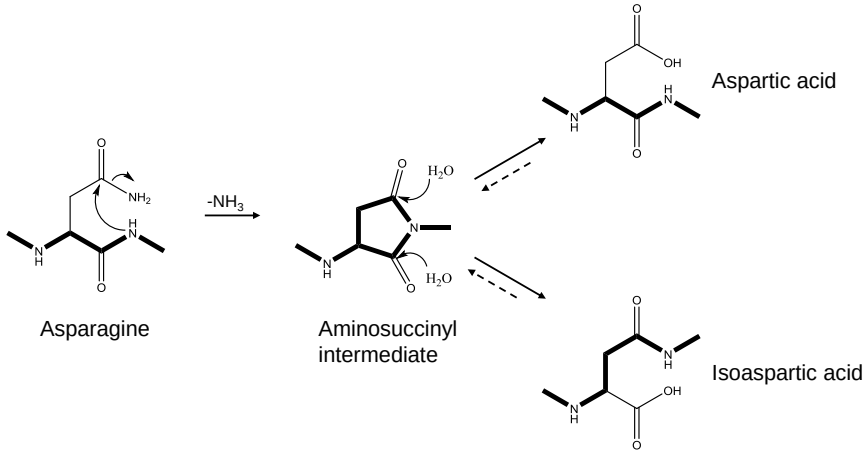

Figure 16. Standard model for non-enzymatic deamidation of asparagine. The peptide backbone (thick line) is extended by one carbon when isoaspartic acid is formed.

Deamidation during sample preparation steps performed at elevated temperatures (i.e. reduction and digestion) has been studied extensively in the context of proteomics research and requires special attention (Ren *et al.*, 2009). Common sample preparation steps such as sulphur bridge reduction and subsequent alkylation of the free thiols were therefore not included. Additionally, we added a reference peptide that is known to have a high asparagine deamidation rate in the reactions as a control with which to monitor deamidation originating from sample preparation. A short digestion protocol was ultimately developed by using high quality trypsin and a denature-through-precipitation approach, which makes the proteins readily accessible to the

digestive enzymes. We have previously explored the use of organic solvents to aid proteolytic digestion of various proteins including ricin (Östin *et al.*, 2007). This approach was effective, but necessitated the use of high enzyme concentrations because the organic solvents reduced the enzyme's digestion efficiency. We subsequently introduced an evaporation step to replace the organic solvent with a suitable digestion buffer. This enables optimal reaction conditions using smaller quantities of the digestive enzyme without sacrificing efficiency. Using small quantities of the enzyme also reduces the risk of intact protein interference during the LC-MS analysis. Modifications of this denature-through-precipitation method were used in the sample preparation step prior to proteolytic digestion in paper I-IV.

In Paper II, the desorbed fraction mainly consisted of the target toxin and the desorption buffer (50 mM TFA/50% methanol). Samples intended for mass spectrometry were precipitated by drying before the enzyme was added. No reduction and subsequent alkylation was needed as it was not critical either for the general protein identification objective, or for the quantitative digestion down to low attomole amounts (Fig. 4 in paper II). The protocol presented in paper III includes organic solvent precipitation, and the final digestion efficiency was carefully monitored using time-intensity plots of the desired peptides. The developed digestion method proved to be stable and yielded near-quantitative digestion (fig S-6 in paper III). As further explained in the peptide identification section below, the reduction and alkylation steps were absolutely necessary for the work reported in paper IV. The denature-through-precipitation protocol was therefore further refined by performing three sequential reactions on the precipitate. First, the protein was precipitated to remove the protein storage matrix (which may contain detergents, salts, etc.). Then a reduction reaction was performed directly on the pellet, followed by an alkylation reaction (further described in the experimental section of paper IV). Finally, the pellet was subjected to enzymatic digestion. This novel protocol provided high yields of reduced, alkylated and enzymatically digested peptides within three hours, demonstrating both the accessibility of the SV2C-protein to the enzyme under these conditions and the effectiveness of the reaction sequence. Visual inspections of the vials during these steps confirmed that the precipitate persisted until the enzyme's degradative action converted the protein into peptides. Visual clarification of precipitated solutions has previously been shown to indicate successful peptide production (Östin *et al.*, 2007).

5.3 Mass Spectrometry Analysis

5.3.1 Specific Peptide Identification

Considerable time and effort was saved during the development of mass spectrometric methods by performing thorough background studies on the target proteins and conducting comprehensive theoretical preparations before starting the experimental work. In particular, some important conclusions and workflow decisions could be made even at this preliminary stage simply by listing the desired peptides in a table together with brief summaries of their properties. In the following section I will describe this process in more detail and explain how these background studies influenced and guided each work.

In paper I, we used a public available tool (Robinson and Robinson) to predict which amino acids in ricin might be predisposed to deamidate based on the protein's sequence and tertiary structure. This yielded a list of all the protein's asparagine (Asn) residues, ranked according to their predicted deamidation halftime (see tables 1 and 2 in paper I). As the prediction tool only gave theoretical probabilities for the different sequence motifs, we decided to map all the deamidation sites. After combining *in silico* and LC-MS data, we concluded that both an alternative enzyme (chymotrypsin) and an alternative separation method (CE-MS) would be needed to cover all of the asparagines in ricin. The use of chymotrypsin enabled mapping of motifs not accessible with trypsin, while CE-MS enabled detection and separation of shorter peptides than was possible with LC-MS.

Some of the peptides were more interesting than others. The short peptides T_A22 (NGSK) and T_B9 (SNGK), which both contain highly ranked potential deamidation sites, were not accessible using the nano-LC column switching setup. Instead their detection required CZE separation combined with MS analysis; the deamidation of Asn60 in T_B9 was confirmed in this way. Asn141 in peptide T_A12, the candidate identified as being most likely to undergo deamidation according to Table 2 of paper I, was not found to be deamidated in preliminary experiments. This was further confirmed using the synthetic T_A12 peptide with Asp substitution at Asn141. Careful analysis of this result suggests that this asparagine is stabilised by a hydrogen bond to a nearby threonine residue, Thr13, an effect that was not fully recognized by the deamidation prediction tool. The location of Asn141 at the very beginning of a rigid alpha helix structure provides additional stability, adding up to a stability disabling the proposed deamidation reaction. The position of Asn141 and the proposed hydrogen bond is highlighted in Figure 17. As seen in Table 2 of paper I, the comprehensive use of two enzymes and two separation techniques was necessary to recover all asparagine-containing peptides.

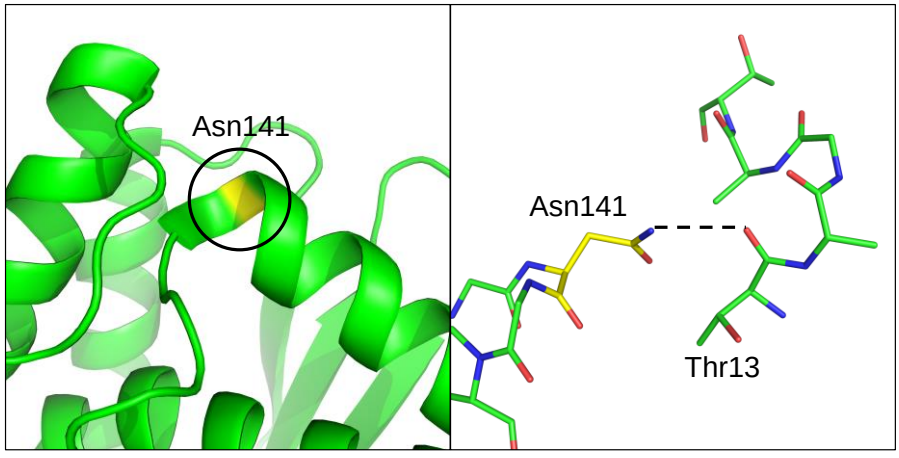

Figure 17. Highlighted on the left is the position of Asn141 at the very end of a rigid alpha helix structure. On the right the proposed hydrogen bond connecting Asn141 to the nearby Thr13.

Another successful use of this peptide-level in silico preparation was in the glycan mapping of the BoNT/A receptor SV2C in paper IV. The final peptide selections may seem very straightforward since all of the results were from trypsin peptides. However, we were confronted with glycans of unknown sizes and numbers, sulphur bridges and very hydrophobic regions within the sequence of interest. When initially summarizing the theoretical tryptic peptides (Table 1), we identified several problematic issues. First, the glycosylation positions of highest interest, N559 and N565 (outlined in paper IV, and illustrated in Figure 1 therein), contained a possible sulphur bridge involving a cysteine (C*) located between them. It was necessary to reduce this bridge to detect the (NC*SFFHNK) peptide. However, this octa-peptide was expected to be extensively glycosylated, and some uncertainty remained as to whether the fully glycosylated peptide would be detectable by LC-MS or not. Another notable issue is that the upstream and downstream trypsin digestion sites are very close to the N559- and N565 glycosylation sites respectively, so the corresponding N-glycans could potentially affect the digestion efficiency at these sites. We therefore searched for glycosylated peptides resulting from unsuccessful cleavage at one or both sites, elongating the octa-peptide in both the C- and N-terminal directions.

In the end, analysis of single and double mutants in the targeted sequence revealed the retention time of the octa-peptide core and the size of each glycan, thereby supplying the key information needed for complete glycan mapping as shown in paper IV. In the final analysis it was also shown that the reduced and alkylated short tryptic peptide bearing two large glycans was retrievable using our chosen LC-MS setup. Additionally, we concluded that one of the glycans

did indeed suppress its downstream digestion; notably, this missed-cleavage peptide was the most suitable option for displaying the glycan complexity at N559 and N565.

Table 1. *Theoretical tryptic peptides from the wild type SV2C construct covering the luminal domain sequence of interest including the glycosylation positions N559 and N565*.

A.A. Start	A.A. End	M.C.	Sequence
519	533	0	(K) SC*TFEDVTSVNTYFK (N)
534	551	0	(K) **N$_{glyco}$**C*TFIDTVFDNTDFEPYK (F)
552	558	0	(K) FIDSEFK (N)
559	566	0	(K) **N$_{glyco}$**C*SFFH**N$_{glyco}$**K (T)
567	579$^{\#}$	0	(K) TGC*QITFDDDYSA*PGSAWSHPQFEK$^{\#}$* (*S*)
534	558	1	(K) NCTFIDTVFDNTDFEPYKFIDSEFK (N)
552	566	1	(K) FIDSEFK**N$_{glyco}$**C*SFFH**N$_{glyco}$**K (T)
559	579$^{\#}$	1	(K) NCSFFHNKTGC*QITFDDDYSA*PGSAWSHPQFEK$^{\#}$* (*S*)

Amino acid (A.A.) numbering is according to SV2C wild-type sequence, (X) denotes up- and downstream amino acids, presumed N-glycosylated asparagines are in bold, C*=carbamidomethylated cysteine, $^{\#}$residues in italic are part of the C-terminally fused *Strep*-tag® II sequence and not included in the A.A. numbering.

5.3.2 Protein Identification

As described in the experimental section, sensitive identification of target proteins requires the selection of peptides that can satisfy low level detection demands and the desired identification criteria, which can include subtype-, serotype- or cultivar-specific sequences.

In paper II, ricin peptides were chosen for maximum sensitivity while retaining an emphasis on specificity. The selected peptides needed to be unique for ricin and abrin sequences respectively, because these proteins exhibit extensive sequence similarities with their respective agglutinins. Consequently, the peptide showing the strongest response in the analysed mixture was common to both ricin and its agglutinin (figure 4, paper II). Its dual origin made this peptide unsuitable as an explicit ricin identifier, although it is still worth monitoring in forensic sample analysis because its presence is a sensitive general indicator of castor bean preparations. After selecting appropriate peptides, an high resolution accurate mass-based LC-MS approach was developed in which identification is supported by data on both the retention times and masses of reference sample peptides. Quantification of ricin was made against reference digests of ricin standards. For RIP-II toxins not available as quantitative reference standards a “top 3 quantification” approach was employed using the summed response of three high response peptides in the ricin reference digests as standards (Silva *et al.*, 2006, Bygdell, 2013).

In paper III the most sensitive BoNT/A1 peptides specific for the parent protein were selected based on the results obtained using the dilution procedure described in the preceding sections, and the subtype information presented by these peptides was annotated. Because the mass spectrometric technique used in this case was PRM, data on the peptides' chromatographic distributions were also considered. The selected peptides originated from both BoNT/A1 chains, as shown in Table 2 below. Additionally, subtype indication could be extracted from this set of peptides because different subtypes of the toxin would present different subsets of these peptides. However, it should be noted that the development of a procedure capable of verified rather than indicative subtyping would require the development of tailored MS methods for each targeted subtype because the behaviour of individual peptides is difficult to predict.

Table 2. *Selected PRM peptides from BoNT/A1listed according to their elution order in LC-MS.*

Peptide	Amino Acid[1]	Chain	Sequence	Present in Subtypes[2]
1	721-729	HC	VNTQIDLIR	A1, A2, A3, A5 and A6
2	232-241	LC	LYGIAINPNR	A1-A3, A5-A8
3	273-283	HC	FIDSLQENEFR	A1, A2, A5-A8
4	344-356	LC	M^{ox}LTEIYTEDNFVK	A1, A5-A8
5	1141-1156	HC	GSVM^{ox}TTNIYLNSSLYR	A1
6	566-581	HC	IALTNSVNEALLNPSR	A1 and A8
7	582-592	HC	VYTFFSSDYVK	A1, A6*, A7* and A8
8	382-393	LC	VNYTIYDGFNLR	A1, A5*, A6 and A7*
IS	Leu-Enk	-	YGGFL	-

[1]Amino acid sequence and residue numbering from strain Hall A1 (ATCC 3502). [2]Subtype comparison is according to previously published alignments (Kull *et al.*, 2015). *indicates identical sequences but the correct upstream trypsin motif (R or K) is missing. M^{ox}=oxidized methionine. IS= internal standard

5.3.3 Glycosylation Screening

In paper IV, a collision energy alteration approach was used to solve the glycan structure puzzle. The use of high collision energies generated specific glycopeptide fragments and extracted ion chromatograms of these fragments, such as HexNAc 204.1 *m/z* (Conboy and Henion, 1992), indicated glycopeptide retention times as shown in Figure 18. Increasing the collision energy in a stepwise fashion made it possible to identify the glycans associated with a given peptide, and the intact glycopeptide could then be selected for specific LC-MS/MS analysis. In summary, this collision energy stepping approach fully resolved the details of SV2C's glycan structures, as shown in Figure 5 and Figure 6 in paper IV.

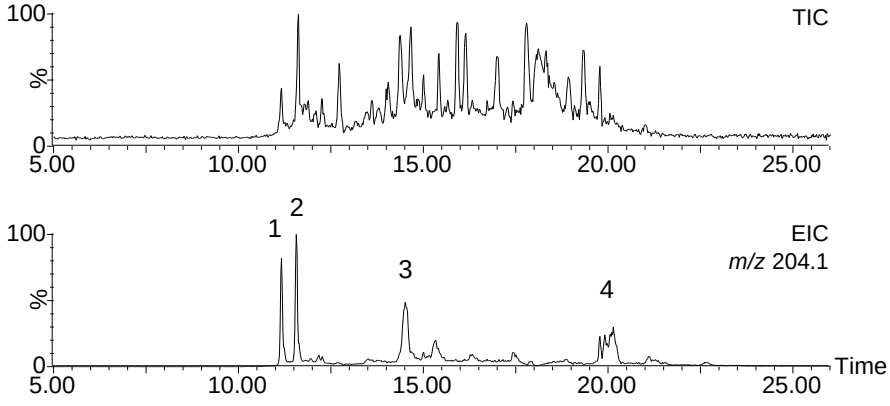

Figure 18. Total ion chromatogram and extracted ion chromatogram (204.1 *m/z*) from high collision energy screening for glycopeptides in a SV2C N565Q mutant sample. Four distinct peaks indicating the presence of glycopeptides. In subsequent analysis peak 1 and peak 2 were confirmed as glycosylated peptides from the IgG1-part of the construct, whereas peak 3 was confirmed as the N559-A579 peptide glycosylated on N559, and peak 4 was the N534-K551 peptide, glycosylated on N534. The two latter from the luminal domain as outlined in Table 1.

5.4 Method Performance Evaluation

The methods developed in papers II and III were evaluated in two different ways. In paper II, the affinity method was developed and then applied to toxin identification in various sample matrixes. This called for two types of evaluations: evaluations of the affinity column's general capacity to attract galactose binding lectins, and evaluations of matrix effects on the system. The first series of evaluations was performed using pure and crude RIP-II preparations, and revealed that the system was capable of high capacity, throughput and sensitivity while retaining galactose specificity. The second series of evaluations was performed by calculating recoveries from a set of different matrices, and revealed the complete method to be matrix-tolerant, RIP-II-comprehensive, and compatible with both enzymatic processing and MS. Special attention should though be made regarding the matrix content, as the affinity could be suppressed by matrix present galactose or similar carbohydrates. Additionally, as the method enrich other galactose binding proteins besides RIP-IIs, this has to be considered in the subsequent digestion and LC-MS steps.

In paper III we applied the developed and validated method to a botulinum neurotoxin. Method evaluation was performed with the pure toxin complex and a spiked bacterial growth matrix. The pure toxin complex was used for the initial proof of concept experiments, while the method's linearity and

sensitivity were tested using the spiked bacterial growth matrix. In addition to the strengths inherited from paper II, the experiments with BoNT/A samples revealed the method to be NT-specific, sensitive, quantitative in the tested matrix, and amenable to parallelization with the previously developed RIP-II approach. One identified drawback in analysing BoNT is the method's specificity for galactose binding complex's, which discriminates against non-L-PTC containing sub- and serotypes, including the human toxic E serotype (Rossetto *et al.*, 2014).

5.5 Applications

The different methods developed in this work were shown to be useful in several contexts.

The development of rapid and effective enzymatic digestion protocols to reduce sample preparation times was initiated by a requirement for rapid responses to alleged illegal use of toxins (Östin *et al.*, 2007). Later experiments including the removal organic solvent before digestion significantly improved the protocol's effectiveness because it allowed samples to be digested at high concentrations (as precipitates) using a low volume reaction mixture in the absence of organic solvents (Bergström *et al.*, 2007). In paper I a high kinetic rate was achieved through the precipitation approach, enabling a short protocol to minimize peptide level deamidations. In papers II and III, the use of the developed method helped to ensure that the new protocols could be performed rapidly while still achieving complete protein digestion. As a proof of concept, the denaturation-through-precipitation protocol was successfully used for efficient and rapid digestion of the SV2C proteins in paper IV. Although in this case a short response time was not strictly required, the 1 h digestion time reduced the total time needed for sample analysis, and allowed useful results to be generated more quickly than would otherwise have been possible. The benefits of handling the target protein in pellet form were further highlighted by the finding that the pellets underwent efficient reduction and alkylation reactions in low volume reactions.

The complete RIP-II toxin identification method was applied to samples from a forensic investigation into illegal production and possession of toxins. As described in paper II, a 32-year old man was suspected of handling toxins, and samples found at his apartment were submitted for analysis at FOI (Figure 18).

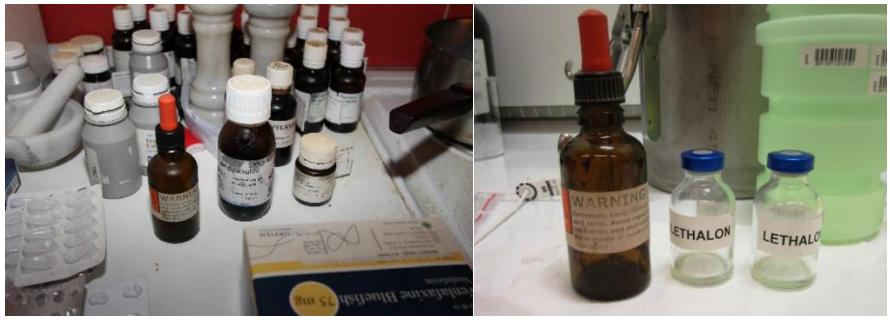

Figure 19. Various chemicals found at the scene (left) and samples analysed regarding illegal production and possession of toxins (right). Photo: FOI and Swedish Police (with permission).

The method was applied to low volume samples (10 µL) and both non-retained and low pH desorbed fractions were analysed using LC-MS. Peptides from ricin, RCA120, abrin and its corresponding agglutinin were all indicated in the full scan analysis (fig 5 in paper II). Diagnostic ricin biomarkers identified in the non-retained fractions further indicated ricin content. For unambiguous toxin identification, samples were reanalysed using LC-MS/MS together with ricin and abrin reference digests. Retention times, *m/z* measurements, and product ion spectra matched the reference material, providing conclusive evidence of ricin and abrin content. Finally, the method's capacity for quantitation was exploited to determine the toxin concentrations in the samples.

The widespread demand for ricin analysis capabilities, together with the diversity of techniques and reference materials used, prompted the adoption of a collaborative approach to develop reference materials and organize an international proficiency test for ricin and RCA120 analysis (Worbs *et al.*, 2015b). During the proficiency test, the developed RIP-II toxin identification method was applied to the analysis of nine blinded samples potentially containing ricin and/or RCA120 of unknown concentrations. Sample matrixes were of high protein content, including a meat extract, organic fertilizer (solid sample) and milk, and were all intended to mimic real sample scenarios. The successful detection of ricin and RCA120 at all spiking levels and in all matrixes except the intermediate level of ricin in the milk matrix, highlighted some important features of the method. Specifically, it was found that if the sample's lactose content is high, lectin binding is suppressed and so RIP-II detection becomes impossible. However, the method proved to be viable for all of the other tested matrices, suitable for RIP-II detection, and less costly than alternative immunoaffinity methods (Kalb *et al.*, 2015).

In addition, a proficiency test was also organized regarding botulinum neurotoxin analysis (Worbs *et al.*, 2015a). During the characterization of the

reference material for this test (Weisemann *et al.*, 2015), the denaturation-through-precipitation protocol was applied before enzymatic digestion and sequence-confirming mass spectrometric analysis of four different BoNT serotypes. As seen in Table 3, the digestion was successful, yielding almost 50% peptide coverage for serotypes A, B and E. However, high BSA levels reduced the coverage achieved for serotype F peptides (Weisemann *et al.*, 2015).

Table 3. *Summary of LC-MS/MS results on the four characterized BoNT reference samples.*

Sample	Number of peptides identified (of total)	Identified peptides from heavy chain	Identified peptides from light chain
BoNT/A	68 (48%)	39	29
BoNT/B	74 (47%)	44	30
BoNT/E	64 (50%)	48	16
BoNT/F	43 (33%)	30	13

6 Conclusions

The aim of this thesis was to improve society's capacity to deal with threats from biocriminal and terrorist activities, involving toxins such as ricin and botulinum neurotoxin. The objective was to develop novel analytical methods with improved performance, thereby providing tools for forensic toxin analysis.

Paper I describes studies on the structural details of ricin deamidation. The developed analytical procedures and tools enabled mapping of all the deamidation sites in ricin. It was also shown that ricin's previously reported isoelectric diversity is related to its level of deamidation, which may be useful in overcoming the forensic attribution challenge of matching a sample to a specific source. The outlined methodology should be applicable in other protein characterisation studies and the method's potential use in forensic profiling could be useful in forensic studies of other proteins.

The objective in paper II was to exploit galactose affinity for the trace enrichment of ricin and its RIP-II analogues. A successful RIP-II enrichment method utilizing the toxin's natural toxic mechanism was developed based on a galactosyl modified material. The LC-MS protocol was developed for use with the resulting enriched samples that enables sensitive detection of ricin preparations and unambiguous ricin and abrin identification. The method's applicability in investigations into bioterrorism and biocrime was confirmed during a forensic investigation of a "home-made" toxin preparation.

The developed galactosyl affinity material was then successfully used for BoNT enrichment in paper III. BoNT L-PTC's natural galactosyl binding capacity, and the pH-triggered NT release mechanism were successfully transferred to the column format, resulting in NT-specific enrichment.

Analyses of samples related to bioterrorism and biocrime incidents must be conducted quickly and accurately. The newly developed affinity methods satisfy these criteria: they are highly sensitive for the RIP-II toxin and BoNT,

and have high throughputs in the affinity step. Even if analysed samples are found to be from hoax letters or of non-toxic content, it is important to retrieve general sample knowledge. In this setup, that is achieved by analysing non-retained fractions.

In paper IV, a broad analytical approach revealed the importance of SV2C glycosylation for BoNT/A interaction at the neuron-muscle interface. Mass spectrometric analysis of various SV2C sequence mutants revealed their glycan heterogeneity. The results obtained provided an improved understanding of the interaction between BoNT/A and its receptor.

The overall objective for all four papers was to develop peptide mass spectrometry methods for protein identification and characterization. This objective was achieved by the development of four distinct fit-for-purpose methods based on new approaches to both sample preparation and mass spectrometry analysis.

7 Future Perspectives

The results presented in this thesis highlight some new ideas and potentially fruitful topics for future investigation.

First, while the deamidation data presented in paper I suggest that deamidation patterns could be useful in sample matching, more thorough investigations will be needed to determine whether such an approach is genuinely viable. It is possible that the seeds' origin (cultivar), age (maturity) and storage conditions all affect the extent of deamidation, and that sample matching based on deamidation might be limited.

Second, when analysing suspicious samples, the affinity method presented herein does not perform well when small quantities of the target substance are present together with a strongly galactosyl binding background. Further investigations will be needed to determine at which levels the target substance can be reliably detected, and whether it is the affinity method or the MS analysis that is limiting. The RIP-II method and the BoNT approach should also be integrated and the expanded three step approach evaluated to verify that the high pH-induced release of BoNT does not affect galactosyl bound ricin.

The analysis of BoNT samples could be further expanded to clarify the proposed benefits of the velcro effect in various matrixes, and the complete method including the MS identification could be developed and refined to achieve serotype and subtype resolution.

Finally, while the results presented in paper IV on the glycosylation of SV2C and its influence on interactions with BoNT are interesting, more studies on native proteins are needed. Interaction studies with glycosylated SV2s have been performed, but there has been no comprehensive glycan mapping of native SV2C combined with interaction studies. Such an investigation would be required to draw definitive conclusions regarding the role of glycosylation and glycan structure.

References

ABIAN, J., OOSTERKAMP, A. J. & GELPÍ, E. 1999. Comparison of conventional, narrow-bore and capillary liquid chromatography/mass spectrometry for electrospray ionization mass spectrometry: practical considerations. *Journal of Mass Spectrometry,* 34**,** 244-254.

AHARONI, A. & BRANDIZZI, F. 2012. High-Resolution Measurements in Plant Biology. *The Plant Journal,* 70**,** 1-4.

ARIMITSU, H., SAKAGUCHI, Y., LEE, J. C., OCHI, S., TSUKAMOTO, K., YAMAMOTO, Y., MA, S., TSUJI, T. & OGUMA, K. 2008. Molecular properties of each subcomponent in *Clostridium botulinum* type B haemagglutinin complex. *Microb Pathog,* 45**,** 142-9.

ARNON, S. S., SCHECHTER, R., INGLESBY, T. V., HENDERSON, D. A., BARTLETT, J. G., ASCHER, M. S., EITZEN, E., FINE, A. D., HAUER, J., LAYTON, M., LILLIBRIDGE, S., OSTERHOLM, M. T., O'TOOLE, T., PARKER, G., PERL, T. M., RUSSELL, P. K., SWERDLOW, D. L. & TONAT, K. 2001. Botulinum toxin as a biological weapon: medical and public health management. *Jama,* 285**,** 1059-70.

BERGSTRÖM, T., ÖSTIN, A., FREDRIKSSON, S.-Å. & NILSSON, C. Rapid Verification of Ricin by Mass Spectrometry 13th Nordic Mass Spectrometry Congress 28-31/8 2007 2007 Savonlinna, Finland.

BIGALKE, H. 2013. Botulinum toxin: application, safety, and limitations. *Curr Top Microbiol Immunol,* 364**,** 307-17.

BRANDEN, C. & TOOZE, J. 1999. *Introduction to Protein Structure,* New York, Garland Publishing, Inc.

BYGDELL, J. 2013. *Plant Biology Through Quantitative Proteomics.* SLU.

CAMPBELL, N. 1996. *Biology,* San Francisco, CA, The Benjamin/Cummings Publishing Company.

CDC. *Centers for Disease Control and Prevention. Emergency Preparedness and Response. Specific hazards. Bioterrorism.* [Online]. Available: http://emergency.cdc.gov/bioterrorism/overview.asp [Accessed August 31 2016].

CONBOY, J. J. & HENION, J. D. 1992. The determination of glycopeptides by liquid chromatography/mass spectrometry with collision-induced dissociation. *J Am Soc Mass Spectrom,* 3**,** 804-14.

COTTRELL, J. S. 1994. Protein identification by peptide mass fingerprinting. *Pept Res,* 7**,** 115-24.
CREIGHTON, T. E. 1992. *Proteins: Structures and Molecular Properties,* New York, W.H. Freeman.
CWC. *Convention on the Prohibition of the Development, Production, Stockpiling and Use of Chemical Weapons and on Their Destruction* [Online]. Available: https://www.opcw.org/chemical-weapons-convention/ [Accessed 31 August 2016].
DESPEYROUX, D., WALKER, N., PEARCE, M., FISHER, M., MCDONNELL, M., BAILEY, S. C., GRIFFITHS, G. D. & WATTS, P. 2000. Characterization of ricin heterogeneity by electrospray mass spectrometry, capillary electrophoresis, and resonant mirror. *Anal Biochem,* 279**,** 23-36.
DORLAND'S MEDICAL DICTONARY. Available: www.dorlands.com [Accessed August 31 2016].
DUPRÉ, M., GILQUIN, B., FENAILLE, F., FERAUDET-TARISSE, C., DANO, J., FERRO, M., SIMON, S., JUNOT, C., BRUN, V. & BECHER, F. 2015. Multiplex quantification of protein toxins in human biofluids and food matrices using immunoextraction and high-resolution targeted mass spectrometry. *Anal Chem,* 87**,** 8473-80.
EISELE, K. H., FINK, K., VEY, M. & TAYLOR, H. V. 2011. Studies on the dissociation of botulinum neurotoxin type A complexes. *Toxicon,* 57**,** 555-65.
ENCYCLOPÆDIA BRITANNICA. Available: https://www.britannica.com/ [Accessed August 31 2016].
ENDO, Y., MITSUI, K., MOTIZUKI, M. & TSURUGI, K. 1987. The mechanism of action of ricin and related toxic lectins on eukaryotic ribosomes. The site and the characteristics of the modification in 28 S ribosomal RNA caused by the toxins. *J Biol Chem,* 262**,** 5908-12.
ENDO, Y. & TSURUGI, K. 1987. RNA N-glycosidase activity of ricin A-chain. Mechanism of action of the toxic lectin ricin on eukaryotic ribosomes. *J Biol Chem,* 262**,** 8128-30.
FENN, J. B. 2002. Electrospray ionization mass spectrometry: How it all began. *J Biomol Tech,* 13**,** 101-18.
FONSLOW, B. R., STEIN, B. D., WEBB, K. J., XU, T., CHOI, J., PARK, S. K. & YATES, J. R., 3RD 2013. Digestion and depletion of abundant proteins improves proteomic coverage. *Nat Methods,* 10**,** 54-6.
FREDRIKSSON, S.-Å., HULST, A. G., ARTURSSON, E., DE JONG, A. L., NILSSON, C. & VAN BAAR, B. L. 2005. Forensic identification of neat ricin and of ricin from crude castor bean extracts by mass spectrometry. *Anal Chem,* 77**,** 1545-55.
FRENOY, J. P. 1986. Effect of physical environment on the conformation of ricin. Influence of low pH. *Biochem J,* 240**,** 221-6.
FRIGERIO, L. & ROBERTS, L. M. 1998. The enemy within: ricin and plant cells. *Journal of Experimental Botany,* 49**,** 1473-1480.
GU, S., RUMPEL, S., ZHOU, J., STROTMEIER, J., BIGALKE, H., PERRY, K., SHOEMAKER, C. B., RUMMEL, A. & JIN, R. 2012. Botulinum

neurotoxin is shielded by NTNHA in an interlocked complex. *Science,* 335**,** 977-81.
GUERRIER, L. & BOSCHETTI, E. 2007. Protocol for the purification of proteins from biological extracts for identification by mass spectrometry. *Nat Protoc,* 2**,** 831-7.
HERBERT, C. G. & JOHNSTONE, R. A. W. 2002. *Mass Spectrometry Basics*, CRC Press.
JANSEN, H. J., BREEVELD, F. J., STIJNIS, C. & GROBUSCH, M. P. 2014. Biological warfare, bioterrorism, and biocrime. *Clin Microbiol Infect,* 20**,** 488-96.
JANZ, R. & SUDHOF, T. C. 1999. SV2C is a synaptic vesicle protein with an unusually restricted localization: anatomy of a synaptic vesicle protein family. *Neuroscience,* 94**,** 1279-90.
JIMENEZ, P., CABRERO, P., BASTERRECHEA, J. E., TEJERO, J., CORDOBA-DIAZ, D. & GIRBES, T. 2013. Isolation and molecular characterization of two lectins from dwarf elder (Sambucus ebulus L.) blossoms related to the Sam n1 allergen. *Toxins (Basel),* 5**,** 1767-79.
KALB, S. R., SCHIELTZ, D. M., BECHER, F., ÅSTOT, C., FREDRIKSSON, S.-Å. & BARR, J. R. 2015. Recommended Mass Spectrometry-Based Strategies to Identify Ricin-Containing Samples. *Toxins (Basel),* 7**,** 4881-94.
KHOURY, G. A., BALIBAN, R. C. & FLOUDAS, C. A. 2011. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci Rep,* 1.
KINTER, M. & SHERMAN, N. E. 2005. *Frontmatter*, John Wiley & Sons, Inc.
KULL, S., SCHULZ, K. M., WEISEMANN, J., KIRCHNER, S., SCHREIBER, T., BOLLENBACH, A., DABROWSKI, P. W., NITSCHE, A., KALB, S. R., DORNER, M. B., BARR, J. R., RUMMEL, A. & DORNER, B. G. 2015. Isolation and functional characterization of the novel *Clostridium botulinum* neurotoxin A8 subtype. *PLoS One,* 10**,** e0116381.
KÖCHER, T., SWART, R. & MECHTLER, K. 2011. Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Anal Chem,* 83**,** 2699-704.
LADIN, B. F., MURRAY, E. E., HALLING, A. C., HALLING, K. C., TILAKARATNE, N., LONG, G. L., HOUSTON, L. L. & WEAVER, R. F. 1987. Characterization of a cDNA encoding ricin E, a hybrid ricin-Ricinus communis agglutinin gene from the castor plant Ricinus communis. *Plant Mol Biol,* 9**,** 287-95.
LAM, K. H. & JIN, R. 2015. Architecture of the botulinum neurotoxin complex: a molecular machine for protection and delivery. *Curr Opin Struct Biol,* 31**,** 89-95.
LEE, K., GU, S., JIN, L., LE, T. T., CHENG, L. W., STROTMEIER, J., KRUEL, A. M., YAO, G., PERRY, K., RUMMEL, A. & JIN, R. 2013. Structure of a bimodular botulinum neurotoxin complex provides insights into its oral toxicity. *PLoS Pathog,* 9**,** e1003690.

LEE, K., LAM, K. H., KRUEL, A. M., MAHRHOLD, S., PERRY, K., CHENG, L. W., RUMMEL, A. & JIN, R. 2015. Inhibiting oral intoxication of botulinum neurotoxin A complex by carbohydrate receptor mimics. *Toxicon,* 107**,** 43-9.
LI, X. & FRANZ, T. 2014. Up to date sample preparation of proteins for mass spectrometric analysis. *Arch Physiol Biochem,* 120**,** 188-91.
LORD, J. M., ROBERTS, L. M. & ROBERTUS, J. D. 1994. Ricin: structure, mode of action, and some current applications. *Faseb j,* 8**,** 201-8.
LORD, M. J., JOLLIFFE, N. A., MARSDEN, C. J., PATEMAN, C. S., SMITH, D. C., SPOONER, R. A., WATSON, P. D. & ROBERTS, L. M. 2003. Ricin. Mechanisms of cytotoxicity. *Toxicol Rev,* 22**,** 53-64.
LU, W., BENNETT, B. D. & RABINOWITZ, J. D. 2008. Analytical strategies for LC-MS-based targeted metabolomics. *J Chromatogr B Analyt Technol Biomed Life Sci,* 871**,** 236-42.
LUMOR, S. E., DEEN, B. D., RONNINGEN, I., SMITH, K., FREDRICKSON, N. R., DIEZ-GONZALEZ, F. & LABUZA, T. P. 2013. Assessment of the inhibition of ricin toxicity by lactose in milk. *J Food Prot,* 76**,** 2037-9.
LYIMO, B., FUNAKUMA, N., MINAMI, Y. & YAGI, F. 2012. Characterization of a new alpha-galactosyl-binding lectin from the mushroom Clavaria purpurea. *Biosci Biotechnol Biochem,* 76**,** 336-42.
MAHRHOLD, S., BERGSTRÖM, T., STERN, D., DORNER, B. G., ÅSTOT, C. & RUMMEL, A. 2016. Only the complex N559-glycan in the synaptic vesicle glycoprotein 2C mediates high affinity binding to botulinum neurotoxin serotype A1. *Biochem J,* 473**,** 2645-54.
MATSUI, T., GU, S., LAM, K. H., CARTER, L. G., RUMMEL, A., MATHEWS, II & JIN, R. 2014. Structural basis of the pH-dependent assembly of a botulinum neurotoxin complex. *J Mol Biol,* 426**,** 3773-82.
MBASU, R. J., HEANEY, L. M., MOLLOY, B. J., HUGHES, C. J., NG, L. L., VISSERS, J. P., LANGRIDGE, J. I. & JONES, D. J. 2016. Advances in quadrupole and time-of-flight mass spectrometry for peptide MRM based translational research analysis. *Proteomics,* 16**,** 2206-20.
MIKAMI, T., AOKI, M. & KIMURA, T. 2012. The application of mass spectrometry to proteomics and metabolomics in biomarker discovery and drug development. *Curr Mol Pharmacol,* 5**,** 301-16.
MIRSALEH-KOHAN, N., ROBERTSON, W. D. & COMPTON, R. N. 2008. Electron ionization time-of-flight mass spectrometry: historical review and current applications. *Mass Spectrom Rev,* 27**,** 237-85.
MORAN, G. J., TALAN, D. A. & ABRAHAMIAN, F. M. 2008. Biological terrorism. *Infect Dis Clin North Am,* 22**,** 145-87, vii.
NAGATSUKA, T., UZAWA, H., OHSAWA, I., SETO, Y. & NISHIDA, Y. 2010. Use of lactose against the deadly biological toxin ricin. *ACS Appl Mater Interfaces,* 2**,** 1081-5.
OLSZOWY, P. & BUSZEWSKI, B. 2014. Urine sample preparation for proteomic analysis. *J Sep Sci,* 37**,** 2920-8.
OVENDEN, S. P., FREDRIKSSON, S.-Å., BAGAS, C. K., BERGSTRÖM, T., THOMSON, S. A., NILSSON, C. & BOURNE, D. J. 2009. De novo

sequencing of RCB-1 to -3: peptide biomarkers from the castor bean plant Ricinus communis. *Anal Chem,* 81**,** 3986-96.
PETERSON, A. C., RUSSELL, J. D., BAILEY, D. J., WESTPHALL, M. S. & COON, J. J. 2012. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics,* 11**,** 1475-88.
PITA, R. & ROMERO, A. 2014. Toxins as Weapons: A Historical Review. *Forensic Sci Rev,* 26**,** 85-96.
RASOOLY, R., HE, X. & FRIEDMAN, M. 2012. Milk inhibits the biological activity of ricin. *J Biol Chem,* 287**,** 27924-9.
RAUNIYAR, N. 2015. Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int J Mol Sci,* 16**,** 28566-81.
RAVEN, P. H., EVERT, R. F. & EICHHORN, S. E. 2005. *Biology of Plants,* New York, W.H. Freeman.
REN, D., PIPES, G. D., LIU, D., SHIH, L. Y., NICHOLS, A. C., TREUHEIT, M. J., BREMS, D. N. & BONDARENKO, P. V. 2009. An improved trypsin digestion method minimizes digestion-induced modifications on proteins. *Anal Biochem,* 392**,** 12-21.
RIEUX, L., SNEEKES, E.-J. & SWART, R. 2011. Nano LC: Principles, Evolution, and State-of-the-Art of the Technique. *LCGC North America,* 29**,** 926-934.
ROBINSON, N. E. & ROBINSON, A. B. *Deamidation* [Online]. Available: www.deamidation.org [Accessed 31 January 2013].
ROEPSTORFF, P. & FOHLMAN, J. 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom,* 11**,** 601.
ROSSETTO, O., PIRAZZINI, M. & MONTECUCCO, C. 2014. Botulinum neurotoxins: genetic, structural and mechanistic insights. *Nat Rev Microbiol,* 12**,** 535-49.
ROXAS-DUNCAN, V. I. & SMITH, L. A. 2012. Ricin Perspective in Bioterrorism. *In:* MORSE, S. (ed.) *Bioterrorism.* InTech.
RUMMEL, A. 2015. The long journey of botulinum neurotoxins into the synapse. *Toxicon,* 107**,** 9-24.
SCHEP, L. J., TEMPLE, W. A., BUTT, G. A. & BEASLEY, M. D. 2009. Ricin as a weapon of mass terror--separating fact from fiction. *Environ Int,* 35**,** 1267-71.
SCHILLING, B., MACLEAN, B., HELD, J. M., SAHU, A. K., RARDIN, M. J., SORENSEN, D. J., PETERS, T., WOLFE, A. J., HUNTER, C. L., MACCOSS, M. J. & GIBSON, B. W. 2015. Multiplexed, Scheduled, High-Resolution Parallel Reaction Monitoring on a Full Scan QqTOF Instrument with Integrated Data-Dependent and Targeted Mass Spectrometric Workflows. *Anal Chem,* 87**,** 10222-9.
SCHMIDT, T. G. & SKERRA, A. 2007. The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins. *Nat Protoc,* 2**,** 1528-35.

SILVA, J. C., GORENSTEIN, M. V., LI, G. Z., VISSERS, J. P. & GEROMANOS, S. J. 2006. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. *Mol Cell Proteomics,* 5**,** 144-56.

SIMPSON, D. M., HALLETT, M., ASHMAN, E. J., COMELLA, C. L., GREEN, M. W., GRONSETH, G. S., ARMSTRONG, M. J., GLOSS, D., POTREBIC, S., JANKOVIC, J., KARP, B. P., NAUMANN, M., SO, Y. T. & YABLON, S. A. 2016. Practice guideline update summary: Botulinum neurotoxin for the treatment of blepharospasm, cervical dystonia, adult spasticity, and headache: Report of the Guideline Development Subcommittee of the American Academy of Neurology. *Neurology,* 86**,** 1818-26.

SMART, J. K. 1997. Medical Aspects of Chemical and Biological Warfare. *In:* ZAJTCHUK, R. (ed.) *Medical Aspects of Chemical and Biological Warfare.* Department of the Army, United States of America: Office of The Surgeon General.

SPHON, J. A. 1978. Use of mass spectrometry for confirmation of animal drug residues. *J Assoc Off Anal Chem,* 61**,** 1247-52.

STEEN, H. & MANN, M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol,* 5**,** 699-711.

STERN, D., PAULY, D., ZYDEK, M., MULLER, C., AVONDET, M. A., WORBS, S., LISDAT, F., DORNER, M. B. & DORNER, B. G. 2016. Simultaneous differentiation and quantification of ricin and agglutinin by an antibody-sandwich surface plasmon resonance sensor. *Biosens Bioelectron,* 78**,** 111-7.

STROCCHI, P., DOZZA, B., PECORELLA, I., FRESINA, M., CAMPOS, E. & STIRPE, F. 2005. Lesions caused by ricin applied to rabbit eyes. *Invest Ophthalmol Vis Sci,* 46**,** 1113-6.

SUN, X. L., HALLER, C. A., WU, X., CONTICELLO, V. P. & CHAIKOF, E. L. 2005. One-pot glyco-affinity precipitation purification for enhanced proteomics: the flexible alignment of solution-phase capture/release and solid-phase separation. *J Proteome Res,* 4**,** 2355-9.

TAYLOR, G. 1964. Disintegration of Water Drops in an Electric Field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences,* 280**,** 383-397.

THEVIS, M., LOO, J. A., LOO, R. R. & SCHANZER, W. 2007. Recommended criteria for the mass spectrometric identification of target peptides and proteins (<8 kDa) in sports drug testing. *Rapid Commun Mass Spectrom,* 21**,** 297-304.

THIEDE, B., HOHENWARTER, W., KRAH, A., MATTOW, J., SCHMID, M., SCHMIDT, F. & JUNGBLUT, P. R. 2005. Peptide mass fingerprinting. *Methods,* 35**,** 237-47.

UN STATISTICS. *FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS* [Online]. Available: http://faostat3.fao.org/home/E [Accessed 31 August 2016].

WADA 2010. Identification criteria for qualitative assays incorporating column chromatography and mass spectrometry, WADA Technical Document, TD2010-DCR, 2010.
VAN HOLDE, K. E. & MATHEWS, C. K. 1996. *Biochemistry*, Menlo Park, California, Benjamin/Cummings Pub. Co., Inc. .
VANDERMARLIERE, E., MUELLER, M. & MARTENS, L. 2013. Getting intimate with trypsin, the leading protease in proteomics. *Mass Spectrom Rev,* 32**,** 453-65.
WANG, K., HUANG, C. & NICE, E. 2014. Recent advances in proteomics: towards the human proteome. *Biomed Chromatogr,* 28**,** 848-57.
WEI, H., TYMIAK, A. A. & CHEN, G. 2013. High-resolution MS for structural characterization of protein therapeutics: advances and future directions. *Bioanalysis,* 5**,** 1299-313.
WEIN, L. M. & LIU, Y. 2005. Analyzing a bioterror attack on the food supply: the case of botulinum toxin in milk. *Proc Natl Acad Sci U S A,* 102**,** 9984-9.
WEISEMANN, J., KREZ, N., FIEBIG, U., WORBS, S., SKIBA, M., ENDERMANN, T., DORNER, M. B., BERGSTROM, T., MUNOZ, A., ZEGERS, I., MULLER, C., JENKINSON, S. P., AVONDET, M. A., DELBRASSINNE, L., DENAYER, S., ZELENY, R., SCHIMMEL, H., ASTOT, C., DORNER, B. G. & RUMMEL, A. 2015. Generation and Characterization of Six Recombinant Botulinum Neurotoxins as Reference Material to Serve in an International Proficiency Test. *Toxins (Basel),* 7**,** 5035-54.
WORBS, S., FIEBIG, U., ZELENY, R., SCHIMMEL, H., RUMMEL, A., LUGINBUHL, W. & DORNER, B. G. 2015a. Qualitative and Quantitative Detection of Botulinum Neurotoxins from Complex Matrices: Results of the First International Proficiency Test. *Toxins (Basel),* 7**,** 4935-66.
WORBS, S., KOHLER, K., PAULY, D., AVONDET, M. A., SCHAER, M., DORNER, M. B. & DORNER, B. G. 2011. Ricinus communis intoxications in human and veterinary medicine-a summary of real cases. *Toxins (Basel),* 3**,** 1332-72.
WORBS, S., SKIBA, M., BENDER, J., ZELENY, R., SCHIMMEL, H., LUGINBUHL, W. & DORNER, B. G. 2015b. An International Proficiency Test to Detect, Identify and Quantify Ricin in Complex Matrices. *Toxins (Basel),* 7**,** 4987-5010.
YAO, G., LEE, K., GU, S., LAM, K. H. & JIN, R. 2014. Botulinum neurotoxin A complex recognizes host carbohydrates through its hemagglutinin component. *Toxins (Basel),* 6**,** 624-35.
YAO, G., ZHANG, S., MAHRHOLD, S., LAM, K.-H., STERN, D., BAGRAMYAN, K., PERRY, K., KALKUM, M., RUMMEL, A., DONG, M. & JIN, R. 2016. N-linked glycosylation of SV2 is required for binding and uptake of botulinum neurotoxin A. *Nat Struct Mol Biol,* 23**,** 656-662.
YATES, J. R., RUSE, C. I. & NAKORCHEVSKY, A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng,* 11**,** 49-79.

ZOPF, D. & ROTH, S. 1996. Oligosaccharide anti-infective agents. *Lancet,* 347**,** 1017-21.
ÖSTIN, A., BERGSTRÖM, T., FREDRIKSSON, S.-Å. & NILSSON, C. 2007. Solvent-assisted trypsin digestion of ricin for forensic identification by LC-ESI MS/MS. *Anal Chem,* 79**,** 6271-8.

Acknowledgements

Först några efterord

Sådär ja.
Avhandlingen är klar och man närmar sig ett forskarkörkort. Ja det är i alla fall så jag tänker mig att en doktorsexamen bäst beskrivs för den som inte är i branschen. I likhet med ett körkort så behöver man öva, läsa på en hel del, och få respons och support på vägen fram. I båda fall testas man teoretiskt såväl som praktiskt: teoriprov följt av uppkörning, avhandling följt av disputation.

Vad innebär ett forskarkort då? Förutom att man utbildats, övat, och förhoppningsvis blivit en både bättre och mer erfaren forskare, så får man också ett intyg på att man kan sin sak. Precis som att man genom körkortet erkänns som förare, redan på papperet så att säga, på samma sätt fungerar även en doktorsexamen. Kan man då forska utan forskarkort? Självklart. Själv har jag gjort det i ett bra tag innan jag bestämde mig för att försöka ta forskarkortet. Här kanske jämförelsen med bilkörkortet haltar något. Mer likt ett förarbevis på snöskoter i det fallet. Inget måste (i alla fall om man är tillräckligt gammal…), utan själva utövandet ger färdigheten. En annan bra jämförelse är ett motorsågskörkort. Jo, det finns faktiskt. Men det kortet har jag kvar att ta. Men det kommer.

Åter till själva avhandlingen. Under arbetets gång har jag insett att jag nog är en hyfsad problemlösare. Kanske riktigt bra i vissa stunder. Men sämre på att skriva ner resultat. Det lär visst finnas de som har strukturen för en text klar för sig redan innan de börjat skriva. Själv tillhör jag är mer de tvångsmässiga knådarna. Ner med texten, baka om, flytta, ta bort och lägg till. Flytta om igen, få respons av andra, baka om på ett annat sätt. Ett hårt arbete helt enkelt. Men ”jämnen dra”, ni som vet vad det betyder, och till slut är jag något på spåret. Och till slut så sitter den där. En bra text helt enkelt. Jag vet dock inte om jag hann ända dit med kappan. Men ”fit-for-purpose” funkar förhoppningsvis på kappan också.

Jag hade tidigt i skrivandet en målsättning att kappan skulle vara användbar på längre sikt. Med bra tips och rekommendationer på peptidanalysernas snåriga väg fram. Hm, väldigt naivt tänkt där, Tomas. Vad jag har förstått så blir inte många avhandlingar speciellt vältummade (förutom acknowledgement då, den delen läser väl alla) eller ens öppnade efter att de blivit klara. Men jag tänkte att försöker säkra upp detta på mitt eget sätt. Och det redan på sidan 4 som ni kanske sett. Så om ni i framtiden inte kommer att öppna denna avhandling av vetenskapliga anledningar, så kan ni i alla fall göra det när ni sitter och klurar på hur det egentligen var med sporthändelserna där under mitten av 2016. Hur gick det i fotbolls-EM? Och hur gick det för oss under OS i Rio? (Ja, att skriva avhandling samtidigt som dessa stora sportevent har satt väldigt tydliga spår).

Börja så här då: Ta ner denna avhandling från hyllan och bläddra upp sidan 4. Läs radiosportens okrönte referatkonung Christian Olssons referat av fransmannen Didier Payets kanonmål och njut av bilderna som spelas upp framför dig. Detta otroligt snygga mål, men framför allt den ohämmade glädje som Payet visade upp, kommer i alla fall för mig att vara bilden av fotbolls EM 2016. Den där glädjen smittade verkligen av sig, ända ner i TV-soffan (jodå, klart att EM prioriterades). Och ett antal gånger under mina skrivande stunder tillsammans med radiosporten har jag haft möjlighet att återuppleva, glädjas och uppmuntras av, det där målet igen. Tack min vän i etern, Christian Olsson.

För övrigt så vanns EM av Portugal efter en prestation av Eder (här återkommer nyttan igen; vad hette han nu, han som avgjorde finalen…) trots att **Alfreds** favorit Cristiano Ronaldo utgick tidigt ur finalen. Skyttekung blev Antoine Griezmann på 6 mål. Och Zlatan avslutade sin landslagskarriär med en förlust, men sa i en av intervjuerna; *I'm only warming up!* Och 34-åringen har vid detta verks pressläggning levererat med råge i Manchester United och leder skytteligan.

Och hur var det med OS då? Sara Sjöström höll. Många andra favoriter föll. Ja nästan alla topptippade svenska deltagare. Men damfotbollen levererade ett silver, Jenny Rissveds på sin mountainbike levererade ett guld. Och en del stora internationella namn levererade: Bolt vann tre guld. Phelps fem. Och förvånande så blev Michael Jonsson av med sitt världsrekord på 400m till en Sydafrikan vid namn van Niekerk. Fotbollsmässigt bjöds det på frisparksgodis av Neymar som dessutom satte avgörande straffen för ett efterlängtat brasilianskt OS-guld i fotboll.

World Cup i hockey spelas ju också i slutfasen av detta arbete. Tyvärr blir det svårt att återge på något trovärdigt sätt då det börjar efter ”pressläggningen” och avslutas innan disputationen. Och gissa resultat har inte varit min styrka i år (referens: EM tipset 2016). Men Canada känns troliga i en

finalserie. Och hjärtat säger mot Sverige. Och med Henke L i toppform och Loui Eriksson i formation med Sedinarna kan det hända. Men förnuftet säger förlust i semin mot just Canada. Loui blir bäste poängplockare i Sverige. Håll förresten ett öga på Pavel Datsiuk, den ynglingen kommer att leverera.

På tal om Henrik Lundqvist, så kanske vi ska åka och kolla på Rangers snart, **Alfred**. Innan han funderar på att trappa ner. Men det verkar som om han är väldigt hungrig på en Stanley Cup seger, så han kanske är med i några år till i alla fall. Men i år spelar ju fortfarande min favorit #68 Jagr i ligan. Hm, jag tror jag tar och kollar upp spelschemat redan nu.

Nej, nog surrat om sport nu.

Över till några varma tack istället:

Tack alla härliga superkollegor på jobbet! (utan inbördes ordning)

Håkan (du fick ändå stå först, kom ihåg det!), korridorens glädjespridare. På ditt sätt. En skön blandning av Kramer, Per Elofsson och en nördig forskare. När ska jag komma ikapp dig i spåret egentligen? Du är trots allt 50 snart.. Ska du med till Sälen i mars förresten? Du har fina vasaloppstider kvar i kroppen. Jag vågar ju knappt åka rullskidor med dig längre. **Roger**, dina bilars underhåll underhåller. Inget ont menat, vi skrattar inte åt dig. Snarare med… Men nu är det väl lugnt på den fronten ett tag. Och hoppas att det blir ordning på foten snart, så att vi kan köra några trailpass igen. **Karin**; en klippa i mycket som alltid levererar. Och trevlig att jobba med såklart. **Johanna**, glad och positiv, oavsett om det gäller jobb eller fritid. Den enda av här nämnda som också kollar golf på småtimmarna. Kanske jag kan hålla undan för dig på golfbanan, för när det gäller fallhöjdshöjdmeter är jag nog chanslös. **Petrus** och **Tobias,** ni har verkligen hamnat rätt. Kul att ni kom med på båten. Stänger ni festen? **Lars H**, pålitlighet och kunnande så det blir över. Och med ett riktigt lövenhjärta! **Daniel** och **Jenny**, två friska projektledarfläktar. Jag tycker ni bidrar till gruppen på ett imponerande sätt. **Linda**, håller ordning på fler och fler lösa trådar och gör det förträffligt. **Anders**, du har alltid tid över. Och bra svar på det mesta där i din skrubb. Det var ju vi som knäckte fram den där A-B peptiden på ricin en gång i tiden. Du hade tänkt rätt men räknat fel. Sedan hittade jag den och vi fick till ett snyggt MS/MS. Och ricinfällningarna som klarnade. Där klarnade ju mycket av det som är grunden i dessa pek; fällning och effektiv klyvning. **Sten-Åke**. Vi är rätt lika du och jag. Gillar problemlösningar. Kanske lite för noggranna ibland. Sådant tar sin tid. Men vi är måna om instrumenten. Det ska vara rent och snyggt, även på insidan så att säga. Men oj vad vi fått rengöra ibland. Ingen nämnd ingen glömd i det fallet, men jag ska hålla fanan högt på det området medan du sitter på en veranda i

Skåne och klunkar Bandol. **Yvonne,** du håller gruppen och enheten i toppform genom en skön blandning av ledarskap och fredagsbubbel. Tack också för allt stöd runt tid och upplägg av detta projekt. Men din egen toppform påverkar en del medarbetare lite väl mycket. **Håkan** t.ex. Han tror på allvar att han kan nå låga 70kg… **Crister**, handledaren, mannen med väldigt många verktyg i sin låda. Jag tror tyskarna funderade ordentligt när du kom sent till det där telefonmötet för att du hade varit och gjort snö hela natten :) Men för oss skidåkare en anledning så god som någon att vara sen till jobbet. Och lagar du inte bilen så bygger du väl om din Hi-Fi anläggning eller styckar en ren i garaget. Men till kaffet surrar jag hellre om vallor och strukturer än digital streaming, men det har du nog förstått vid det här laget. Men du har ett imponerande djup inom många fält även utanför kemin. Det skulle vara kul att följa med dig upp till Jokkmokk och jaga ihop dina renar någon gång, en upplevelse kan jag tänka mig. **Calle**, du lärde mig det mesta om bra och pålitliga data under din tid på FOI. Du har varit ett bra stöd genom detta arbete också. Ni andra ska veta att när man får kommentarer som ”jag är imponerad” från **Calle**, då känns det riktigt bra. **Henrik**, som redan tidigt såg mig som ”FOIs nya påläggskalv”... Det tog ett tag, men nu blir jag granskad in på benet, så vi får se om jag blir pålägg eller köttfärs. Men vi har en hel del gemensamt. Fäblessen för tyska kvalitetsbilar bland annat. ”because the germans don’t know how to make bad cars”. Jag måste nog skaffa en Quattro® igen. Skidinstruktören **Christian** och cykelfixaren och stabile grannen **Markus**. Efterfesternas hedersmedlemmar: **Oskar**, **Susanne**, **Lina**, **Ante** m.fl. På den här festen får ni stanna så länge ni orkar. Och jag glömmer inte ölen och påsen med Polly i första taget, de satt bra! **FUSK** i alla dess former som inspirerar till träning på många olika sätt. Och **Signar**. Den som sett till att jag i alla fall tränat något under slutveckorna. **Anders A** och **Fredrik** för superb assistans i proteinstrukturernas värld. **Majken, Petra** och ni andra som ger stöd när det som bäst behövs. **Mats S** som står bakom oss på många sätt. **Mona**, **Johanna T**, **Lillemor**, **Rikard**, **Sofia & Bettan,** och alla andra härliga FOI:ter, ingen nämnd ingen glömd, som gör stället så trevligt.

Gunnar W, handledare och gisslan (som du själv uttryckte det). Men du har varit mycket mer än gisslan. Och du var den som en gång i tiden släppte fram mig till en Q-TOF för första gången. Tack för det också. Övriga SLU:are och UPSC folk: **Hasse** för ovärderliga strategiska tips inför skrivandet av kappan. Vi får paratas vid på festen huruvida det fungerade. **Jocke**, **Jonas G**, **Maria A**, **Mr Moritz**, tack för ert jobb och stöd på vägen fram.

Tack till många utanför jobbet också. Till att börja med alla härliga vänner som jag/vi borde umgås mer med: De Umeå-nära **Hanna**&**David**, **Emma**&**Peter**, **Karin**&**Micke**, **Erik**&**Maria**, **Anna**&**Jonas, Sandra**&**Frasse**.

Öviksborna **Mia**&**Johan**, **Mogge**&**Lina**. Klintsjöborna **Erik** och **Jonas** med familjer. Knivstaborna **Sara**&**Rune**, stockholmarna **Henke**&**Sofia,** Ångeborna **Malin**&**Kent.** Jag passade även på att bjuda **s**kidåkarna och gamla gänget på festen; **Olof**, **Anders**, **Mats**, **Wangler**, **Johan S**, **David E** med familjer. Får se om någon av dom har möjlighet att åter dyka upp i Umeå. Men visst borde vi ta oss iväg på lite mer skidåkning, grabbar?

Och så några extra varma tack till de som står mig nära: **Per**&**Carin** och deras härliga grabbgäng i Täby, alltid lika kul att umgås med er. **Jenny**, det blir alltid trevligt när du är med. Ska du med till fjälls i år? **Anita**&**S-G,** omtänksamma, gästvänliga, hjälpsamma, och ett bra umgänge helt enkelt. Vårt stöd på distans när det behövs. **Håkans**, **Bosses**, **Gun-Britt**&**Helge** i Enånger, kusiner och kusinbarn. Morbror **Seth. Micke**&**Jenny**, alltid nära, alltid supportande. **Viggo**, **Hilda** och **Sally**, jag ska lära er köra det mesta och allt bus jag kan. Ni håller mig ung i pulkbacken. **Johanna**, du är mångas favorit, vi saknar dig härborta. **Mamma**, **morsan**, **Gulli**, **farmor**. Kärt barn har många namn. Du betyder väldigt mycket för mig och för oss alla tre. Tack för allt du gjort och allt du gör. Aldrig omöjlig, alltid govillesam.

Alfred! Mitt stora driv här i livet. Min älskade grabb, min stolthet, kämpe och min bästa kompis. Nu börjar innebandysäsongen och det är dags att vi börjar träna in några finter. Jag ska t.ex. lära dig den enda dragning jag kan. Och så ska jag visa dig hur man skjuter för höga skott. Och så ska vi bygga klart din koja såklart. Förresten tror jag att det kan bli en liten crossmotorcykel för din del till våren. Det vore väl något? Men säg inget till mamma och mormor, då blir dom bara oroliga…

Annika! Mitt fundament i livet. Tack för att du är min kärlek, bästa vän och största supporter på samma gång. Och dessutom min nära-till-hands kemiportal och figurproducent... I kemi är jag nöjd tvåa.

Nu stundar bättre tider.