Route attribution of chemical warfare agents

Retrospective classification of unknown threat samples
Route attribution of chemical warfare agents
Retrospective classification of unknown threat samples

Karin Höjer Holmgren
Abstract

Although chemical warfare agents (CWAs) are prohibited under international law, there have been numerous crimes that violate the 1997 Chemical Weapons Convention (CWC) during the last decade, especially in the civil war in Syria where sarin, mustard gas and chlorine have all been used. CWAs have also been used in political assassinations and attempts thereof. In such situations, it is important to identify the deployed CWA and to find information on how it was produced, as this information is potentially of considerable value for any ensuing judicial process. The development and use of advanced analytical methods and multivariate data analysis methods are required to produce this kind of robust forensic evidence and intelligence.

This thesis describes conducted research that aims at retrospectively tracing the synthesis methods applied in the production of CWAs. In three studies, methods for the determination of the employed synthetic route have been assessed. The relative distribution of the impurities gave a unique profile – in effect a “chemical fingerprint” - that was used for retrospective determination of the production method of a specific CWA. The study in paper I was done on the nerve agent Russian VX, S-[2-(diethylamino)ethyl] O-isobutyl methylphosphonothioate, while paper II and III focused on sulfur mustard, bis(2-chloroethyl)sulfide. This thesis discusses the study set up, the choice of analytical methods, methods for data processing and the manner in which classification methods have been employed. The studies shows that the classification models could clearly separate the six production routes used in paper I and the five routes used in paper II. In paper III, a novel non-targeted approach in combination with high-resolution mass spectrometry allowed detection of additional low-concentration compounds in the sulfur mustard samples. This method produced data with sufficient information for classifying samples according to the production method of the precursor thiodiglycol (TDG).

The performance of the classification models was successfully validated with test set samples. All test set samples were correctly assigned in paper I and paper II. The classification of TDG in paper III was more demanding, but still as much as 56-89% of the test set samples were correctly assigned. In addition to the established classification models, compounds with importance for route differentiation were identified, which gave enhanced information on the chemicals formed during the employed synthesis conditions. Their stability has also been investigated, and the results showed that the majority of the chemical attribution signatures (CASs) were stable at room temperature.

The fourth study in this thesis (paper IV) is an international inter-laboratory comparison jointly conducted by eight defence research laboratories
based in Europe, North America, Asia and Australia respectively. All participating laboratories analysed the same samples prepared at the Swedish Defence Research Agency (FOI). The impurity profiles in nerve agent precursor methylphosphonyl dichloride samples were compared by a gas chromatography mass spectrometry (GC/MS) method using a retention index to facilitate data comparison. Retention indices of 16 CASs were calculated and compared, and this showed that the between-laboratory variation was low. This work is a first step towards a harmonised laboratory method for the profiling of CWA samples. The methods developed in this thesis will enhance accurate source attribution of CWAs and could potentially be used when alleged use of a CWA is being investigated.

Keywords: impurity profiling, chemical attribution signature, route sourcing
Preface

This thesis is a result of a collaboration between Linköping University and the Swedish Defence Research Agency, FOI.

Linköping University, with Johan Dahlén at the Department of Physics, Chemistry and Biology, IFM, has broadened the chemical attribution projects at the FOI with a more forensic perspective and a deeper knowledge of source attribution. In addition to the collaboration with the FOI, Linköping University conducts research together with the National Forensic Centre and the National Board of Forensic Medicine. This institutional network has created a strong research platform that is beneficial for all partners involved. Linköping University has also contributed with its excellent PhD courses, and support during the research period and when actually writing this thesis. At Linköping University I have also met other PhD students working in the same or similar fields and exchanged useful ideas.

The FOI is a governmental research agency mainly working for the Swedish armed forces. The FOI is one of the Europe’s leading research institutes in the area of defence and security. The FOI performs applied research and studies within specific areas such as crisis management, IT security, sensor technology, underwater technology, aero systems, and protection against hazardous substances. The FOI division for CBRN Defence and Security have unique laboratory facilities and staff competence for handling toxic and radiological chemicals as well as highly infectious micro-organisms. The FOI also has the legal framework to conduct research on these threat agents. The work included in this thesis was performed at the FOI and could not have been done elsewhere in Sweden. Research on CWAs and related chemicals is done within dedicated laboratories with highly educated and specially trained personnel using well-adapted procedures for sample handling, storage and decontamination.
List of publications


III. Karin Höjer Holmgren, Lina Mörén, Linnea Ahlinder, Andreas Larson, Daniel Wiktelius, Rikard Norlin, Crister Åstot, Route determination of sulfur mustard using non-targeted chemical attribution signature screening, Analytical chemistry, (2021), 93, 11, 4850-4858

IV. Karin Höjer Holmgren, Hanna Hakulinen, Rikard Norlin, Mirjam Bruin-Hoegee, Samantha Qi Shu See, Renee Webster, Karen L. Jacques, Ang Lee Hwi, Christopher P. Evans, Simon Ovenden, Daan Noort, Gregoire Delaporte, Johan Dahlén, Carlos Fraga, Paula Vanninen, Crister Åstot, Interlaboratory comparison study of a chemical profiling method of methylphosphonic dichloride, a nerve agent precursor, manuscript

Paper I, II and III are reprinted with permission from the publisher Elsevier.
Contribution to included publications:

I. Performed the chemical analysis, the data processing and interpretation of the data. Did the multivariate data analysis with co-authors. Made most the figures and tables and contributed to the writing of the paper.

II. Performed the chemical analysis and the data processing, did the multivariate data analysis with co-authors. Made most of the figures and tables and contributed to the writing of the paper.

III. Planned the study, performed the chemical analysis and performed part of the data processing, and wrote part of the paper.

IV. Planned the study and instructed the participating laboratories how they should perform the samples preparation and chemical analysis. Performed the data evaluation and wrote most of the paper.
Publications not included in the thesis:


Daniel Wiktelius, Linnea Ahlinder, Andreas Larsson, Karin Höjer Holmgren, Rikard Norlin, Per Ola Andersson, On the use of spectra from portable Raman and ATR-IR instruments in synthesis route attribution of a chemical warfare agent by multivariate modeling, Talanta, (2018), 186, 622–627

**Abbreviations**

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMDIS</td>
<td>Automated mass spectral deconvolution and identification system</td>
</tr>
<tr>
<td>CAS</td>
<td>Chemical attribution signature, one analyte</td>
</tr>
<tr>
<td>CAS profile</td>
<td>The profile of CAS/analytes in a chromatogram</td>
</tr>
<tr>
<td>CI</td>
<td>Chemical ionisation</td>
</tr>
<tr>
<td>CWA</td>
<td>Chemical warfare agent</td>
</tr>
<tr>
<td>CWC</td>
<td>Chemical Weapons Convention</td>
</tr>
<tr>
<td>DC</td>
<td>Methylphosphonic dichloride</td>
</tr>
<tr>
<td>DCM</td>
<td>Dichloromethane</td>
</tr>
<tr>
<td>DMMP</td>
<td>Dimethyl methylphosphonate</td>
</tr>
<tr>
<td>ED</td>
<td>Euclidian distance</td>
</tr>
<tr>
<td>EI</td>
<td>Electron impact/ electron ionisation</td>
</tr>
<tr>
<td>EIC</td>
<td>Extracted ion chromatogram</td>
</tr>
<tr>
<td>GC</td>
<td>Gas chromatography</td>
</tr>
<tr>
<td>HD</td>
<td>Sulfur mustard, mustard gas, bis(2-chloroethyl) sulfide</td>
</tr>
<tr>
<td>HR</td>
<td>High resolution</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared spectroscopy</td>
</tr>
<tr>
<td>IRMS</td>
<td>Isotope ratio mass spectrometry</td>
</tr>
<tr>
<td>LC</td>
<td>Liquide chromatography</td>
</tr>
<tr>
<td>LLNL</td>
<td>Lawrence Livermore National Laboratory</td>
</tr>
<tr>
<td>MDMA</td>
<td>N-Methyl-3,4-methylenedioxyamphetamine</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass to charge</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance spectroscopy</td>
</tr>
<tr>
<td>OPCW</td>
<td>Organisation for the Prohibition of Chemical Weapon</td>
</tr>
<tr>
<td>OPLS-DA</td>
<td>Orthogonal projection to latent structures discriminant analysis</td>
</tr>
<tr>
<td>PC</td>
<td>Principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal component analysis</td>
</tr>
<tr>
<td>PLS-DA</td>
<td>Projection to latent structures discriminant analysis</td>
</tr>
<tr>
<td>RF</td>
<td>Random forest</td>
</tr>
<tr>
<td>S/N</td>
<td>Signal to noise ratio</td>
</tr>
<tr>
<td>TDG</td>
<td>Thiodiglycol</td>
</tr>
<tr>
<td>TIC</td>
<td>Total ion chromatogram</td>
</tr>
<tr>
<td>TMS</td>
<td>Trimethylsilyl</td>
</tr>
<tr>
<td>VERIFIN</td>
<td>Finnish Institute for Verification of the Chemical weapons convention</td>
</tr>
<tr>
<td>VR</td>
<td>Russian VX, Isobutyl S-2-diethylaminoethyl methylphosphonothiolate</td>
</tr>
<tr>
<td>VX</td>
<td>Ethyl S-2-diisopropylaminoethyl methylphosphonothiolate</td>
</tr>
</tbody>
</table>
# Table of Contents

Abstract ............................................................................................................. iii  
Preface ................................................................................................................ v  
List of publications ........................................................................................... vii  
Contribution to included publications: ........................................................... viii  
Publications not included in the thesis: ............................................................ ix  
Abbreviations .................................................................................................... xi  
Table of Contents ............................................................................................ xiii  

1. Introduction to chemical warfare agents ...................................................... 1  
   1.1 CWAs ................................................................................................... 1  
   1.2 OPCW-designated laboratories ............................................................ 2  
   1.3 Recent illicit use of CWAs ................................................................. 3  
2. Theory and background in chemical forensics ........................................ 7  
   2.1 Profiling studies at forensic laboratories .......................................... 7  
   2.2 Chemical attribution signatures in CWAs ....................................... 9  
   2.3 Chemical analysis of CASs ............................................................... 10  
   2.4 Targeted and non-targeted methods ............................................... 13  
   2.5 Statistical classification methods ..................................................... 15  
   2.6 Summary ........................................................................................... 18  
3. Objectives ................................................................................................... 19  
4. Experimental ............................................................................................. 21  
   4.1 Study design ..................................................................................... 21  
   4.2 Chemical analysis ........................................................................... 24  
   4.3 Data processing .............................................................................. 24  
   4.4 Classification methods ................................................................... 26  
5. Results and discussion ............................................................................. 27  
   5.1 The samples ...................................................................................... 28  
   5.2 Stability of the samples .................................................................. 30  
   5.3 Chemical analysis of CASs .............................................................. 32  
   5.4 Data processing ............................................................................... 35  
   5.5 Identification of CASs .................................................................... 40  
   5.6 Multivariate classification models ................................................... 44  
   5.7 Recommendations for future studies .............................................. 51  
6. Conclusions ............................................................................................... 53  
7. Future perspectives and reflections ......................................................... 55  
8. Populärvetenskaplig sammanfattning ....................................................... 57  
9. References ................................................................................................. 59  
10. Acknowledgements ............................................................................... 65
1. Introduction to chemical warfare agents

Chemicals have historically been used in many wars, even in the remote past, for example arsenic smoke was used by the Spartans, and the Romans used toxic smoke that might have been phosgene [1]. In World War 1, chlorine, phosgene and sulfur mustard (i.e. HD – the two terms are used interchangeably throughout the thesis) were used, and this resulted in many casualties – both deceased and severely wounded soldiers. Some areas in Germany, Belgium and France are still affected to this day by the intense use of toxic chemicals in 20th-century wars, for example barrels of sulfur mustard are still found in the ground [1].

Chemical warfare agents (CWAs) are defined as chemicals used with the deliberate intention to harm or kill humans or animals [2]. CWAs are often divided into groups of chemicals according to their effect on humans; thus there are nerve agents, blister agents, riot control agents and incapacitating agents. The highly toxic nerve agents, such as sarin or soman, cause harm or death by preventing the enzymatic degradation of acetylcholine in the synapse and the result will be catastrophic muscle cramps. A blistering agent, such as sulfur mustard, causes damage to the skin, lungs, and eyes. High concentrations can result in death.

The Chemical Weapon Convention (CWC) came into force in 1997, implemented by the Organisation for the Prohibition of Chemical Weapons (OPCW). OPCW was established in 1997 and has today 193 member states. The CWC prohibits chemical weapon use, production, development, and storage facilities. The OPCW regularly inspects stockpiles, destruction plants, chemical industry plants and scheduled chemical facilities. It has been notably successful since 1997 more than 96.5% of the world’s chemical weapon stockpiles have been destroyed [3].

1.1 CWAs

CWAs and related compounds are named scheduled compounds within the CWC. The scheduled chemicals are put in classes, according to their structure. The schedule 1 compounds are toxic chemicals and their precursors; the chemicals within the schedule 1 have no peaceful applications. For example 1.A.1 are the O-alkyl alkyl phosphonofluoridates, 1.A.2 are the O-alkyl N,N-dialkyl phosphoramidocyanidates, 1.A.3 are the O-alkyl S-2-dialkyl-aminoethyl alkyl phosphonothioates and sulfur mustards are in the 1.A.4 group. The schedule 2 chemicals are precursors and degradation products of schedule 1 chemicals.
such as phosphonic acids and thiodiglycol. The schedule 3 chemicals are toxic chemicals such as hydrogen cyanide and phosgene but also precursors such as phosphorous trichloride and triethanolamine. Many of the schedule 3 chemicals are freely used around the world, with legitimate applications in agriculture, the pharmaceutical and chemical industries.

1.2 OPCW-designated laboratories

The OPCW was established to implement the CWC, and the chemical analyses carried out on- and off-site aims at showing the presence or absence of scheduled chemicals. To be able to investigate alleged use of CWAs at off-site laboratories, the OPCW has established collaborative processes with a number of national laboratories around the world. These laboratories are obligated to perform annual proficiency testing to prove their ability. Three years with high grades and a national accreditation will give a laboratory a designation. This designation means that the OPCW can use the laboratory for investigative purposes if crimes with CWAs are suspected to have been committed. Laboratories can be designated for environmental analysis and/or biomedical analysis [4]. In Sweden, it is the FOI that conducts research on CWAs. The FOI has an OPCW-designated laboratory for both environmental and biomedical analysis of CWA samples. The chemical analysis of environmental and biomedical samples is one important part of verifying alleged use. Verification analysis is performed on identical sets of samples, since samples from the crime scene are split into at least two sets of samples. The analysis is performed at two independent OPCW-designated laboratories. The samples are coded, and the scope of analysis is to search for any CWC-scheduled compounds. Detection of CWAs or degradation products using two different techniques at the laboratories, and reporting according to the strict criteria, results in unambiguous proof of CWA use. When trying to identify those responsible for the crimes, even more information from the chemical analysis is beneficial. In those cases, the chemical analysis has been extended to find and report chemicals that are not present on the CWC list, but are still possibly related to the CWA attack. The differences between verification analysis and profiling analysis are shown in Figure 1. Chemical attribution signatures (CASs) are the by-products and impurities seen in the chromatogram, besides the CWA.
1.3 Recent illicit use of CWAs

CWAs have certainly been used during this last decade, with most of the known examples coming from the war in Syria [5], although there are also notable cases of usage from Malaysia and the United Kingdom.

When the CWC was agreed not many foresaw that terrorists or non-state actors would use CWAs. Earlier, the mission for the OPCW had been to determine whether CWAs had been used in cases of alleged use. In the civil war in Syria, there have been attacks with sarin [6, 7], chlorine [8] and sulfur mustard [5, 9, 10] (Figure 2). The OPCW was given a new and extended mandate in 2017 to find those responsible for the CWA attacks in Syria [11]. This additional task of the OPCW has also affected the research within the CWA field, and new knowledge about CWAs from a forensic perspective is needed.

Syria joined the CWC in 2013; the country’s declared CWA stockpiles had been destroyed. (Syria had known stockpiles of sarin, mustard gas and precursors for other chemical weapons [12].) The United Nations (UN) and the OPCW have investigated the alleged use of CWAs in Syria. The OPCW’s fact-finding mission has been responsible for collecting evidence after incidents and has interviewed witnesses. Samples for chemical analysis have also been collected. The Joint Investigative Mechanism was a collaborative undertaking between the UN and the OPCW (2015-2017) and aimed to identify the perpetrators of the attacks. From 2019, it has been the investigation and identification team at the OPCW that investigates present and previous unsolved cases of suspected use of CWAs in Syria. An investigation of alleged use of CWAs involves gathering all
possible information from the incident, utilising, among other things, photographs, interviews with victims and sampling (Figure 3).

![Figure 2. The number of confirmed attacks with CWAs used in the civil war in Syria. This figure is from the report ‘Nowhere to hide’ by the Global Public Policy Institute [5].](image)

In the environmental samples (for example soil, feathers and rocks) collected after the chemical attack in the town of Khan Shaykhun in Syria, the presence of many by-products and impurities such as pyrophosphonates, phosphorofluoridates, methylphosphonates, phosphates and also hexamine was conclusively identified [6]. The sarin precursor methylphosphonic difluoride from Syrians stockpiles were analysed prior to destruction and found to have characteristic impurities that were similar to those found in environmental samples from Khan Shaykhun. Laboratory experiments have been performed to prove the hypothesis that the sarin came from the precursor stock of methylphosphonic difluoride belonging to the Syrian state [13]. The OPCW and the UN Security Council jointly concluded that the sarin came from the Syrian regime´s stockpiles. The fact that sarin has been used in Syria after Syria had entered the CWC and their known stockpiles had been destroyed, has led to the suspension of rights for Syria within the OPCW [14].
Figure 3. Environmental samples from attacks with CWAs carried out in Syria. © Spiez Laboratory 2013.

Other recent use is the attack with VX at the airport in Kuala Lumpur. Kim Jong-nam, the half-brother of Kim Jong-un, the leader of North Korea, was killed in Malaysia in 2017. The chemical analysis from the investigation showed that he was killed by Ethyl S-2-diisopropylaminoethyl methylphosphonothioate (VX) \[15\]. It was also concluded that the VX was most probably formed after the addition of chemicals from two sources. Two women were seen at the crime scene rubbing something on the victim. Chemical analysis of Mr Kim Jong-Nam’s body and clothes, and chemical analysis of the women’s clothes showed different compounds on all three individuals. This strengthens the hypothesis that the assassination was done by the addition of two components, and that VX was formed on the skin.

In the 2018 nerve agent attack on Sergei Skripal in Salisbury, UK, several other people were poisoned \[16\]. The UK requested help from the OPCW to investigate the attack \[17\]. The chemical analysis of biomedical and environmental samples from the Salisbury incident showed high levels of purity and indeed almost no impurities were detected \[17\]. After this attack, the CWC schedule 1 chemicals were extended to include three new subgroups of chemicals, namely 1.A.13, 1.A.14, 1.A.15 and 1.A.16 \[18, 19\].
2. Theory and background in chemical forensics

2.1 Profiling studies at forensic laboratories

There are not many profiling studies published within the field of CWA analysis [20]. It is therefore useful to learn from studies performed in other areas, such as illicit drug profiling. In the field of forensic science, profiling studies of impurities in illegal drugs and/or narcotics have been conducted for a long time. For example, amphetamine impurities has been studied in a series of EU projects [21-25]. The focus for these projects was on the different production methods, and the goal to discover batch and route differences. A harmonised sample preparation method and analytical method was developed, and a joint database was established. Well-characterised target analytes were selected and produced, and their stability was examined.

Amphetamine impurity profiles have previously been studied by Nielsen et al. [26]. They aimed to distinguish linked and unlinked samples. Linked samples are samples from the same source, for example the same production batch of amphetamine. The study investigated how the impurity profiles changed when samples were stored, and conducted by stressing the importance of having stable CASs.

A work done on N-methyl-3,4-methylenedioxyamphetamine (MDMA) [27] is also highly relevant. Different batches, temperature conditions and reaction times were examined, and it was evaluated how reactions conditions affected the CAS distribution. Morelato et al. [28] showed in their publication how impurities in MDMA could be analysed using several analytical techniques covering both organic and inorganic compounds (Figure 4). For intelligence purposes, gas chromatography mass spectrometry (GC/MS) data gave sufficient information and enabled classification of MDMA seizures as linked or unlinked.

Impurity profiles of cocaine are used to link one seizure of the drug to another [29, 30]. The stability of the impurity profiles has been studied [31, 32], demonstrating that the impurity profiles changed after 12 months of storage and that the storage time and temperature affected the profiles. The alkaloids were more stable than the residual solvent profiles. Another study by the same group [32] showed that the residual solvent profiles can give additional information compared to alkaloid profiles in cocaine seizures.
Chemical profiling of MDMA is done with several analytical techniques, adapted from [28].

Synthesis routes of fentanyl analogues have recently been studied [33-35], showing that route-related impurities are present in fentanyl and that fentanyl samples could be classified according to their synthesis and how the impurities survive on surfaces [36].

There are few published studies on CWAs with an explicitly chemical forensic perspective, which might partly be a result of fact that scientific results of this kind are often considered unsuitable for publication and also because only a small number of laboratories around the world are actually allowed to do research on CWAs.

Earlier studies have focused on degradation products formed during long-term storage, for example sulfur mustard degradation products and impurities in abandoned chemical weapons [37], or sulfur mustard degradation products found in blocks from dumped munition sites in the Baltic Sea [38]. More recently, Ovenden et al. [39] studied impurities in three VX batches after storage, Webster et al. [40] studied sarin impurities, and Fraga et al. have studied nitrogen mustard impurities related to reagent impurities [41]. The recovery of VX and its degradation products and impurities in soil after fire or decontamination has also been studied [42], and the results showed that traces of VX and related chemicals could be recovered from soil.

Despite a relative paucity of older studies, the research area of chemical forensics applied to CWAs has developed rapidly over the last decade, with most of the relevant publications dating from 2010 or later [20]. Common to all of these publications is that CWC-related research have been put in a forensic context. Relevant chemicals to study are all of those chemicals related to the CWC, or chemicals suspected to have been used in crimes that aim to hurt or kill humans. In total, the CWC covers more than one million of chemicals [4].
2.2 Chemical attribution signatures in CWAs

By the term CAS(s), we essentially mean chemical markers. These markers can be either intrinsic or extrinsic. Intrinsic markers are found within a molecule, for example functional groups or specific atoms within a molecule. Intrinsic markers are studied by using isotope ratios using nuclear magnetic resonance spectroscopy (NMR) or isotope ratio mass spectrometry (IRMS). IRMS has been used for instance when links between precursor and home-made explosives were investigated [43]. NMR analysis using carbon 13 isotope ratios has been used in a study of methylphosphonic dichloride (DC) [44]. That particular study showed that the ratios remain constant after fluorination to methylphosphonic difluoride and after being hydrolysed to the degradation product methyl phosphonic acid (MPA). A study of soman, its precursor and degradation product showed that they could be linked to its source by position specific isotope analysis by NMR [45].

Extrinsic markers are impurities or by-products, often present at low concentrations together with the CWA in question. These are the markers studied within this thesis. A CAS is a compound, and the CAS profile is the term used for all CASs found in one sample. Another term often used is impurity profiling. The CAS can be chemicals related to a route/batch/start material or variation in synthesis conditions such as time, heat and so on, depending on the study. Lu et al. [20] listed categories of CAS, namely impurities, unreacted starting materials, additives, by-products formed during synthesis, degradation of impurities, and reactions of by-products. They also stated that a suitable CAS should be stable, unique, and reproducible.

Few precursors of CWAs and their impurities have been studied thoroughly. However, examples of more deeply studied precursors are the nerve agent precursors DC [44, 46, 47] and dimethyl methyl phosphonate (DMMP) [48, 49]. The work by Fraga and co-workers showed how different DC impurities end up as impurities in the sarin stock. The CASs used in the DC studies were detected by liquid chromatography (LC)-HRMS [46] and GC/MS [50]. The studied CASs were stable and were well suited for batch separation.

Impurities in cyanide bought from different suppliers have been investigated [51]: cyanides produced in Germany could be distinguished from US-produced cyanide by IRMS analysis.

CASs related to batches of DMMP have been studied by Hoggard et al. [48] using two-dimensional GC in combination with time-of-flight MS, GCxGC-TOFMS. The stability of the CASs was further explored by adding DMMP batches to indoor material; it was shown that the most volatile CASs are lost if
the time from incident to sampling is long [49]. The recovery of CASs of DMMP from wallboard surfaces has also been studied [52]. DMMP was also added to painted wallboards, and the extraction recovery of a selection of impurities was evaluated [53].

Impurities related to production methods of a tabun precursor have been studied by GC/MS [54] and 20 potential CASs have been identified.

Sampling after a CWA incident will probably involve taking samples from the environment (such as soil, concrete and/or water) if an outdoor attack has occurred. Samples from indoor surfaces such as walls, textile and flooring material can be of interest if an indoor attack has happened. Several studies of CWAs or related compounds on matrices have been made. Mixtures of VX degradation products were applied to dust, and the chemical profiles were measured after exposure to different sampling times and pH [55]. The CAS profiles from VX degradation products were to a high degree recovered from dust, showing that dust can be used as a matrix to sample after a CWA attack. Sampling of phosphonates, phosphonic acids, and pesticides from surfaces has been studied [56], and the results showed that volatility of the chemicals in question affected the sampling recovery. Another study, using LC/MS/MS, sampling indoor materials such as carpets, paper, painted drywalls, and nylon, showed that CWAs could be recovered from these matrices [57]. The authors also showed that tabun impurities could be detected using headspace sampling with SPME. Isobutyl S-2-diethylaminoethyl methylphosphonothiolate (VR) produced by different synthesis methods was applied to food and the CASs were analysed [58, 59]. Based on the recorded CASs it was possible to separate spiked matrix samples by the employed synthetic route. The chemicals in home-made pepper sprays were studied by Mörén et al. [60]. They investigated how the profile was affected by the type of chilli pepper and vendor. Analysis of extracts added to textiles showed that the CAS profiles related to particular species of chilli could be recovered.

2.3 Chemical analysis of CASs

GC or LC together with a MS is commonly used in profiling studies [20], [61]. Within the family of MS, the high-resolution (HR) MS provide better sensitivity and selectivity than the quadrupole instruments. There are many other techniques that can be used, and there follows a brief selection and discussion of these.

Infrared (IR) and Raman spectroscopy can also be used for characterisation and classification of samples. Published applications of Raman spectroscopy
include route separation of sulfur mustard samples [62] and differentiation of uranium ore samples [63], while IR spectroscopy has been used in vehicle paint analysis [64]. Both IR and Raman are non-destructive methods that are easy to use in field measurements. The analysis time is short, and often samples are analysed without any sample preparation. Raman or IR techniques are often combined with chemometric methods to enable their use in attribution/sourcing studies.

The use of several analytical techniques in various combinations is interesting since one technique alone may not show all characteristics of the sample. There are compounds which cannot be detected by GC/MS, such as metals, inorganic compounds, ions and thermo-labile compounds. Mayer et al. [65] analysed fentanyl CASs with GC/MS, LC-MS/MS, inductively coupled plasma mass spectrometry (ICP-MS) and gave better route classification models when combining data from LC and GC analysis.

Mirjankar et al. [51] analysed cyanide impurities with IRMS, inductively coupled optical emission spectroscopy (ICP-OES) and high performance ion chromatography (HPIC). The data from the three analytical techniques were evaluated separately, but altogether they contributed to the separation of samples according to batch, country and factory. Ovenden et al. [39] used NMR, GC/MS and LC/HRMS to identify a number of relevant VX impurities.

GC/MS is a separation and detection technique suitable for use on small neutral organic analytes [66]. The analytes are dissolved in a solvent and the subset of the sample is vaporised in the injector. The analytes are separated in the GC mostly due to boiling point but also by interaction with the stationary phase within the GC. The temperature in the GC is increased during analysis, and the most volatile analytes reach the detector first. The analytes are ionised and fragmented in the MS. The result from a GC/MS analysis is a chromatogram with the analytes separated in time. The second dimension is the mass spectra. Analytes are fragmented into characteristic parts that have a weight because of the number and type of atoms. The analytes get characteristic mass spectra. A quadruple MS in EI mode 70 eV gives identical mass spectra for an analyte. This makes it possible to create spectra libraries and search unknown spectra in such libraries. HR MS determines the weights with better accuracy. GC/MS are used in two modes, electron ionisation (EI) or chemical ionisation (CI). Generally, EI gives a strong fragmentation pattern while CI is softer and often gives less fragmentation and is frequently used to confirm molecular ions. CI can be used with several gases such as methane or isobutane. Methane will produce a lower number of fragments than EI but more than CI with isobutane. In CI, the ionisation is done by chemicals instead of electrons but the gas is ionised
by electrons emitted from the filament. CI can be used in positive or negative mode. The reactions that take place are mainly proton transfer. The isobutane gets positively charged by loss of an electron. This ion reacts with the analyte and forms [A+H]^+ + gas molecule. If an unknown compound is analysed by GC/MS and interpretation of the EI spectrum has resulted in a preliminary structure, then CI can be used to see if the molecular weight matches the proposed structure. With high-resolution data, with more than unit resolution, then the different atoms (and the number) can be given from a molecular ion in CI or EI if seen. This really helps when identifying unknown analytes.

Retention index (RI) was invented by Kovatz in 1958 [67] and further developed after that; and is today often used in CWA analysis [4] when GC data from different instruments are to be compared. The RI is more suitable than retention time, because retention time will depend on column length and the applied oven temperature programme, which will not be the case for the RI value. When applying an RI, an alkane series is analysed with a GC/MS or GC/FID system and every alkane is given a value according to the number of carbons (Figure 5). Decane with ten carbons gets RI value 1000, tetradecane with 14 carbons gets RI 1400 and so one. Retention times of the alkanes together with the RI values are used in the RI calibration. The RI calibration is then used for calculation of the RI of compounds according to their retention time. An RI can be calculated manually by using this formula [4]:

\[
RI = 100C_n + 100(C_{n+i} - C_n)\frac{t_{R(x)} - t_{R(n)}}{t_{R(n+i)} - t_{R(n)}}
\]

Equation 1

where \(C_{n+i}\) is the alkane eluting after the target compound and \(C_n\) is the alkane eluting prior the target compound, \(t_R\) is the retention time of the target analyte (x), alkane prior target analyte (n) and after target analyte (n+i). It is important to remember is that an RI is dependent on the column stationary phase. An RI can also be used for selection of the most probable structure of an unknown compound if several structures with different RIs are proposed [4].
2.4 Targeted and non-targeted methods

Targeted analytical methods have a pre-defined set of analytes to analyse. The methods can be qualitative or quantitative. The verification work done at OPCW-designated laboratories is mainly based on qualitative methods, and chemicals listed by the CWC define the target analytes. The number of chemicals is more than one million, if all isomers are included. Both the sample preparation methods and analytical methods are broad enough to cover as many of the listed chemicals as possible. For GC/MS EI there are dedicated spectra library, the OPCW chemical analysis database (OCAD) and the validation group database (VGWD). These spectra libraries are extremely useful when analysing a CWA, and together with the deconvolution software Automated Mass Spectral Deconvolution and Identification System, AMDIS, GC/MS EI data can be efficiently searched for CWAs. For LC/MS analysis where less fragmentation occurs, ions related to a class of chemicals are often searched for. An interesting approach is presented by Zhang et al. [68]: they used non-targeted screening in combination with target analysis with an alerting database with fragments related to CWAs. A thorough mapping of the fragmentation pattern in LC/HRMS of nerve agents and degradation products resulted in the database. Another example with a combined methodology is blood sample analysis of narcotic drugs. The data processing method was a combination of targeted and
non-targeted methods [69] thought to be able to detect new psychoactive substances. A visualisation of the ideas behind the targeted and non-targeted methods is shown in Figure 6.

True non-targeted screening methods involve a sample preparation method and an analytical technique that enables detection of a wide range of chemicals. The data processing is then performed without limiting the compounds to be detected. Sample profiles are then compared, for example serum analysis of exposed and non-exposed animals. This methodology has been used in environmental analysis [70-72] and in metabolomics [73].

<table>
<thead>
<tr>
<th>Targeted analysis</th>
<th>Non-targeted analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detect chosen compounds</td>
<td>Detect &quot;all&quot; compounds</td>
</tr>
<tr>
<td>Detect compounds in a target library</td>
<td>Search data files for differences between samples and identify differentiating compounds</td>
</tr>
</tbody>
</table>

**Figure 6.** The different thinking behind targeted and non-targeted methods.

Large and complex data sets are often handled, and the processing methods are highly automated. This work can be compared to the old saying about finding a needle in a haystack (see Figure 7). One goal in such studies can be to find markers that are higher or unique in samples from exposed patients, and identify those markers. The markers can later on be used for identifying an exposed patient (or patients) who is in an early stage of a disease.

**Figure 7.** How to find a needle in a haystack? Allegorical illustration of the data processing problems faced when handling complex data sets: how do we find the interesting analytes when there are so many compounds that may not be relevant for the topic. Photo from www.freeimages.com
2.5 Statistical classification methods

Complex data sets with many variables, samples or both many samples and a large number of variables can be difficult to evaluate. Looking at one variable at a time may not give a good overview of the data, and there may be variables that are correlated to each other. Multivariate data analysis is one method where the complexity of the data is reduced, and the data is transferred into fewer dimensions and is thus more easily visualised and analysed. Source attribution studies of CWAs are not the most complex samples but they are still too complex to study manually. Statistical classification methods are often necessary to extract information that enables differentiating of samples and evaluating differences between samples. The data often needs to be pre-treated prior to the statistical analysis. Since the response of the instrument can vary, or the concentrations of the analytes differ, normalisation of the peak areas is commonly used. Normalisation can be done to an internal standard or calculated according to relative peak areas by dividing peak area by the total peak areas within the sample. How to handle zeros in the data matrix is also important: either they can be kept, or a low value can be added. The data matrix can also be transformed, for example by logarithm. The data is then further scaled, with unit variance scaling often used.

Non-supervised methods, such as principal component analysis (PCA), gives an overview of the data and possible groups can be seen. PCA reduces the data and visualises it by principal components (PC). PC1 and PC2 shows the largest variation in the data, and the sample distribution can be visualised in a few dimensions. Variables (the CASs) can be studied in loading plots and in coefficient plots showing variables correlated to a class.

Examples of supervised methods are projections to latent structures discriminant analysis or partial least squares discriminant analysis, PLS-DA and orthogonal OPLS-DA. These classification methods have a category with information of sample class, and classification models are built by using the data matrix (for example peak areas used as variables, the X matrix) and the classes (Y matrix). OPLS-DA is a development from PLS and separates variation in X that are orthogonal to Y, and that part does not contribute to class separation, but shows within class variation [74]. PCA, PLS-DA and OPLS-DA are linear methods, and the data needs to be normally distributed for the methods to perform well.

A non-linear classification method is, for example, random forest (RF), a tree base method [75]. RF has been used for the classification of CASs used in the characterisation of essential oils [76] and in chemical forensic applications [77, 78].
Independent of the applied method, the developed classification model must be validated. Cross validation is done by removing a subset of the data and thereafter recalculating the model; the procedure is repeated when omitting different subsets of the data until all subsets have been omitted. The cross validation gives a measurement of the robustness of the model. The model quality is expressed in $R^2$, a measurement of the model’s ability to describe the sample variance, and $Q^2$, the model’s ability to predict samples. A measurement of the prediction error is shown by the root-mean-square error of prediction (RMSEP). Models are also validated by test set samples. The optimal test set samples are samples that are similar to those samples that the models are constructed of, and they are also similar to samples that the models will be used for in the future. The test set samples are samples with a known class, so we can see how well the model performs by comparing the predicted class to the true class of the test set samples. If the test set samples work well and the models will be used for similar samples, the models will most likely perform as well as they did with the test set samples.

There are other classification methods apart from OPLS-DA and RF. Mirjankar et al. [51], compared five classification methods: two unsupervised methods, i.e. hierarchical cluster analysis (HCA) and PCA, and three supervised methods, i.e. PLS-DA, K nearest neighbour (KNN) and support vector machines discriminant analysis (SVMDA). Data from several analytical techniques were used; the classification methods were evaluated and the methods worked well. Mirjankar et al. [51] also showed that reduction of variables gave better predictions in their data set. Reduction of variables was also used in a study on impurities in home-made pepper sprays [60]. The study showed that the target analytes could be reduced from 1122 to 273 without lowering the performance of the models. Another method that can be used is a hierarchical decision tree with several models. Jansson et al. [58] successfully implemented this method on the VR in food data. This is a good method to use if the data is complex. Strozier et al. [79] used two approaches for the data: in the first method they divided variables as present and absent in the samples (In/Out variables), while the second method used the relative peak areas (Oval Area). The In/Out variables and the Oval Area were then used as in-data in classification tree analysis (random forest); the authors concluded that both methods worked well.

Many studies have used distances to compare how similar profiles are – this is often referred to as linked and unlinked samples [22], [26], [28] and [32]. There are different types of distances that can be used, for example euclidian distance, ED, and pearson distances. ED was calculated using the formula [22]:

$$ED = \frac{\sum_{i=1}^{n} (x_{i} - y_{i})^2}{\sum_{i=1}^{n} (x_{i})^2 + \sum_{i=1}^{n} (y_{i})^2 - 2 \sum_{i=1}^{n} (x_{i}y_{i})}$$
\[ ED(kl) = \sqrt{\sum_{j=1}^{n} (x_{kj} - x_{lj})^2} \] Equation 2

where \( n \) is the dimensions and \( x_l \) is one profile and \( x_k \) is another profile. The distances between individual profiles (samples) are calculated to show the similarity within and between classes.

The study of CWAs from a forensic perspective, for example sample matching of a confiscated CWA, is not currently done to any great extent. Ahlinder et al. [80], however, have shown an example of how chemical forensic research papers can be used in court. Their study transformed model data into evidence values. In forensic science, the value of a piece of evidence is expressed through likelihood ratios, and in Sweden a 9-graded scale is used to communicate the strength of evidence. Bayes’ theorem describes how to calculate the likelihood ratios [81].
2.6 Summary

In summary, source attribution of CWA-samples is not currently well investigated. Studies on how synthesis conditions affect CASs are scarce, which is why the research opportunities in this area are enormous. The stability of the CASs is not often known, which is essential if the CASs are to be used in real cases. The field of chemical analysis is well developed for the analysis of CASs and data processing methods for detection and peak area integration do exist. However, the methods can be improved and made more flexible, precise, and user-friendly. From my point of view, as an analytical chemist, there are enough classification methods for solving the profiling questions, but the implementation of a forensic perspective remains under-represented in many studies.

The route differences of CWAs is one particularly relevant area to study. States, non-state actors and terrorists are expected to have very different resources, knowledge level and available time that will affect their choice of equipment and chemicals. This means that information on how any given CWA was produced will give information about the perpetrator and will therefore be of importance in crime scene investigations. It follows that the knowledge of how a CWA was produced may therefore be useful for intelligence purposes or in forensic investigations after CWA use.
3. Objectives

The overall objective of this thesis was to develop suitable methods for source attribution of CWAs, with a focus on route attribution/production method differences. The chemical attribution profiles of two CWAs and one precursor were studied (Figure 8).

Paper I
The objective of this work was to investigate if route attribution was possible by analysing impurities in VR samples produced by six different methods.

Paper II
The objective of the study was to challenge the methodology developed in paper I with a more demanding sample set consisting of sulfur mustard samples. The sulfur mustard samples had a high degree of purity and were synthesised using similar synthetic routes.

Paper III
The objective of paper III was to further explore the profiling methods by analysing the sulfur mustard sample by HRMS and using a non-targeted data approach for the data processing. Two statistical classification methods were evaluated.

Paper IV
The objective of this paper was to evaluate a method for impurities profiling. This was done as an international inter-laboratory comparison study. The performance of the analytical method and the data processing method were studied.
Figure 8. Visualisation of papers I, II, III and IV. The structures of VR, HD and DC and hydrolysis product methyl phosphonic acid are shown.
4. Experimental

The details of each of the studies are described in the papers. This experimental section gives an overview of the performed studies that hopefully facilitates comparison of the studies and the methods used. Table 1 summarises the aims of the studies, as well as the samples, samples preparation methods and analytical methods. Table 2 shows the data processing methods used in the four papers.

4.1 Study design

The first study, paper I, was done on the nerve agent Russian VX (VR), and the aim was to investigate if there existed any measurable route differences, and see whether the CAS could be used to classify samples to a specific route. Paper I was done in collaboration with the Lawrence Livermore National Laboratory (LLNL). Six synthesis methods were selected for the study and VR were produced at the FOI and LLNL using identical synthesis protocols. The methods started with either methylphosphonic dichloride (DC) or methylphosphonous dichloride, and VR was formed through 2 to 4 steps. Samples produced at the FOI and LLNL were exchanged and analysed at both laboratories. An overview of the collaboration is shown in Figure 9.

In paper II and paper III, the CASs of HD was studied. Paper II was also produced through collaboration with the LLNL. HD was the study choice as a complement to the VR study that had given us proof that the concept of performing identical synthesis at two laboratories and sharing samples and data was beneficial when investigating potential route differences. Eleven different synthetic routes were applied in the synthesis of HD. Nine of these were similar as they had thiodiglycol (TDG) as a common intermediate. TDG was produced by three methods and TDG was chlorinated by three chlorinating reagents, hence giving the nine related methods for HD synthesis. Two additional methods, which used ethylene as the starting material, were also employed. All 11 routes gave HD with high purity; this was fully intended since the aim was to get a more homogenous data set and to see if there were still measurable route differences and if the concept used in paper I could be applied. HD was produced at the FOI and LLNL and samples were exchanged.
Figure 9. Description of the cooperation between the partners FOI and LLNL in paper I. Information was exchanged at all points in the research work, from planning of the synthesis to analysis, identification and evaluation of the data. Raw data and samples have also been exchanged.

Paper III was conducted on the FOI part of the produced HD samples. The analytical method and the processing method were changed from GC quadrupole MS to GC orbitrap HRMS. The data processing method was changed from making a target library to a non-targeted approach in combination with isotope ratio filtration. The methodology was applied to spiked soil and textile samples to further explore the formed CASs.

Paper IV was an inter-laboratory study where a selection of CASs related to DC impurities was studied. DC was produced by two methods at the FOI and distributed to international laboratories together with sample preparation instructions. The participating laboratories were Innovation for life (TNO) from the Netherlands, Direction générale de l’armement (DGA) from France, VERIFIN from Finland, National Laboratories (DSO) from Singapore, the Defence Science and Technology Laboratory (DSTL) from the UK and the Defence Science and Technology Group (DSTG) from Australia. The aim in Paper IV was to improve the method by using instrumental settings more similar to OPCW-designated laboratories methods using a common target library and calculate RI for data comparison.
<table>
<thead>
<tr>
<th>Aim of study</th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
<th>Paper IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigation of route differences</td>
<td>Investigation of route differences</td>
<td>Investigation of route differences using HRMS</td>
<td>Compare analysis result from 8 laboratories</td>
<td></td>
</tr>
<tr>
<td>CWA</td>
<td>VR</td>
<td>HD</td>
<td>HD</td>
<td>DC, precursor of nerve agents</td>
</tr>
<tr>
<td>Number of routes</td>
<td>6</td>
<td>11</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Synthesis laboratories</td>
<td>2 (FOI, LLNL)</td>
<td>2 (FOI, LLNL)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Replicates /route</td>
<td>2-6</td>
<td>6</td>
<td>4 + 2 (test set samples)</td>
<td>1</td>
</tr>
<tr>
<td>Number of samples</td>
<td>37/57</td>
<td>66</td>
<td>44</td>
<td>48 (6 samples x 8 labs)</td>
</tr>
<tr>
<td>Stability of CAS</td>
<td>-</td>
<td>0, 7 days’, 6 months’ storage at room temperature</td>
<td>-</td>
<td>0, 1 and 2 months storage in room temperature</td>
</tr>
<tr>
<td>Sample preparation</td>
<td>- Dilution in DCM to 0.1 mg/mL - TMS derivatisation</td>
<td>- Dilution in DCM 0.1 mg/mL - TMS derivatisation</td>
<td>- Dilution in DCM 0.5 mg/mL - Ethylacetate extraction of spiked HD matrix samples</td>
<td>Crude DC dissolved in DCM, extraction of hydrolysed DC</td>
</tr>
<tr>
<td>Analytical techniques</td>
<td>- GC/MS EI - GC/MS CI - GC/HRMS EI - NMR</td>
<td>- GC/MS EI</td>
<td>- GC/HRMS EI</td>
<td>- GC/MS EI - NMR</td>
</tr>
</tbody>
</table>

DCM – dichloromethane, DC - methylphosphonic dichloride, CAS - chemical attribution signature, CWA - chemical warfare agent, TMS - trimethylsilyl
4.2 Chemical analysis

The aim of the chemical analysis in all four papers was to detect CASs. GC/MS was used as the major analytical technique in all four publications. The MS used in papers I, II and IV was a quadrupole with unit resolution. GC/MS CI was used for molecular ion confirmation of preliminarily identified analytes in paper I and paper II. In paper I, an HRMS was used as a tool for identification of the CASs. In paper III, the analysis of all samples was done using a GC/HRMS. NMR was used in paper I and paper IV. The purpose of the NMR analysis was to analyse the purity of the intermediates and final products.

4.3 Data processing

Data processing involved peak detection and integration of CAS. In paper I and paper II, a target library in AMDIS was built. In paper I, all compounds with an intensity higher than 0.5% of the sum of the total ion chromatogram (TIC) area in the chromatogram were included in the target library. The target library was constructed of mass spectra with additional information such as RI, and the molecular weight determined by GC/MS with CI. In paper II, the aim was to include all detectable compounds in the chromatograms with a few exceptions for substances present only in a single sample or common laboratory contaminants clearly not related to HD synthesis.

In paper III, a non-targeted strategy was used, with no initial selection of compounds, but with area thresholds, retention time alignment and isotope ratio filtration that allowed only sulfur and/or chlorine compounds to be selected.

In paper IV, the CASs were extracted by AMDIS and peak areas were used when calculating ED.

Dibenzothiophene was used as an internal standard in paper II and paper III. In paper II, the IS was used during the data processing work to be enable a comparison of samples. In paper III, the IS was used as a quality control on the processing method. If IS was found in the sample, it was assumed that the processing setting was correct.
**Table 2. Overview of data processing methods.**

<table>
<thead>
<tr>
<th></th>
<th>Paper I</th>
<th>Paper II</th>
<th>Paper III</th>
<th>Paper IV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CAS selection criteria</strong></td>
<td>Targeted</td>
<td>Targeted</td>
<td>Non-targeted + isotope ratio filtration</td>
<td>Targeted</td>
</tr>
<tr>
<td><strong>Data processing software</strong></td>
<td>AMDIS</td>
<td>AMDIS</td>
<td>TraceFinder</td>
<td>AMDIS, Chemstation</td>
</tr>
<tr>
<td><strong>CASs/Number of CASs</strong></td>
<td>35/49</td>
<td>50/103</td>
<td>Total: 2713, Filtered: 714 S and/or Cl</td>
<td>16/16</td>
</tr>
<tr>
<td><strong>Peak area</strong></td>
<td>- TIC area</td>
<td>- TIC area</td>
<td>- EIC area</td>
<td>- TIC area</td>
</tr>
<tr>
<td></td>
<td>- Normalised, log, scaled</td>
<td>- Normalised, log, scaled</td>
<td>- Normalised, log, scaled</td>
<td>- Normalised peak areas</td>
</tr>
<tr>
<td><strong>Multivariate data analysis method</strong></td>
<td>PCA</td>
<td>PCA</td>
<td>PCA</td>
<td>PCA</td>
</tr>
<tr>
<td></td>
<td>PLS-DA</td>
<td>OPLS-DA</td>
<td>OPLS-DA</td>
<td>ED</td>
</tr>
<tr>
<td><strong>Models</strong></td>
<td>6 classes model</td>
<td>Models of the HD forming step with 5 classes</td>
<td>Hierarchal tree models (OPLS-DA and RF) 11 classes model (RF)</td>
<td>-</td>
</tr>
<tr>
<td><strong>Validation</strong></td>
<td>Subset of the VR samples from both LLNL and FOI</td>
<td>Subset of the HD samples from both LLNL and FOI</td>
<td>Newly produced HD samples of 9 routes</td>
<td>-</td>
</tr>
</tbody>
</table>

AMDIS - Automated mass spectral deconvolution and identification system, ED - Euclidian distance, EIC - extracted ion chromatogram, RI – retention index, CAS – chemical attribution signature, TIC – total ion chromatogram, PCA – principle component analysis, RF – random forest
4.4 Classification methods

The aim of papers I, II and III was to study the differences in the formed impurities and by-products when using different synthesis methods, and see if the samples could be classified according to their production method based on their impurity profiles - the chemical attribution profiles. To evaluate this, statistical classification methods have been used since the data was too complex to evaluate manually. The data used in the classification models were relative peak areas; the peak areas were normalised to the peak area sum of all target compounds in the sample. Relative peak areas were used as this enables comparison of low- and high-concentration samples.

In paper I, the relative peak areas were log transformed to make them more normally distributed prior to modelling. The log transformation \((X+ 0.01)\) also turned all non-detected peak areas into a very low value instead of a zero-value. Both unit variance and Pareto scaling were evaluated during model work; Pareto scaling was chosen since it gave slightly better models in the multivariate data analysis.

In paper II, the peak areas were normalised to a peak area sum and log transformed prior to modelling. The data was mean centered.

In paper III, the data was normalised, log transformed and scaled to unit variance. Unit variance reduces the importance of a variable due to its size, so both low- and high-concentration analytes should have equal impact when modelling route differences. By using PCA, an overview of the data was obtained and detection of outliers was possible. In paper I, PLS-DA was used to establish classification models, and in paper II and paper III, OPLS-DA models were employed.

In paper III, RF models and OPLS-DA models were tested and their performances were compared. The data was complex since there were nine routes and a hierarchical classification tree was used. The first model separated the samples according to the TDG-method or ethylene-method. The second model separated the samples into the three chlorination method that had been used, and finally there were three separate models for classifying samples based on how the TDG was produced. The models were validated using test set samples. By using classification models, CASs important for route separation could be detected.

In paper IV euclidian distance was used as a toll for comparison of impurity profiles, the differences between the two calluses were also visualised by dendrograms.
5. Results and discussion

Paper I, paper II and paper III are source attribution studies. They focus on finding and characterising differences in the CAS profiles that depend on the applied production method. The method employed to do this differs between the papers. When starting this research we did not know to what extent samples would differ or how to evaluate these potential differences. At this early stage there were discussions about the possibility of finding unique markers, having a set of analytes with a known structure and even understanding why these analytes are formed. After analysing the VR samples in paper I, it was obvious that many analytes were present in several routes and that it was the differences between the relative proportions of these analytes that made it possible to discriminate between samples, rather than the presence of unique markers. This was the most important result from paper I.

In paper II, the most important result was that the methodology developed in paper I could be used for the more difficult sample set where the CAS profiles were more similar. HD samples produced with different methods could be differentiated due to their CAS profiles. Every chlorination method leaves unique traces, and many CASs could also be identified.

In paper I and paper II, dedicated target libraries were built and used for data processing of the samples, while in paper III a non-targeted approach was used in which the defined data processing settings decided which peaks were selected. This is a non-targeted method, where no preselection of analytes is done; instead, all detected analytes passing the set criteria were included in the classification analysis. The result from paper III showed that by using this method, the HD sample could be classified according to the HD forming method equally as well as in paper II. Additional information was also collected, the samples could be classified according to the production method of the TDG-intermediate; this was not possible in paper II. The method was also applied to spiked matrix samples and worked well. The CAS profiles of HD in spiked matrix samples differed from the CAS profiles of crude HD. However, there were route-specific CAS profiles also in the spiked matrix samples.

Paper IV is an application study within the field of chemical forensics, a step towards sample matching. The knowledge developed in paper I, paper II, and paper III, is not easily transferred to cases where alleged use of CWAs is being investigated. After a CWA attack, samples will be taken for chemical analysis. Initially the question will be if a CWA is present or not. Then, at a later stage, some kind of sample matching or origin investigation to determine how the CWA was produced may be required. Since it is forbidden to produce, stockpile and use CWAs, there will be investigations of where the CWA came
from; was it stolen from somewhere, for instance? If not stolen, how was it produced? To respond to such searching questions, chemical analysis followed by some kind of sample matching is likely to be required. For sample matching to work there needs to be an accurate measure of how similar two samples are, and how similar they need to be in order to conclude definitively whether or not they originated from the same source. This has been done with the sarin used in Syria [82]. In that particular case, the clinching evidential point was the presence of specific chemicals in the stored sarin precursor sample and the same impurities found and in the used sarin.

The study done in paper IV was a sample matching study in the sense that reproducibility between laboratories was studied and the ability of the laboratories to produce consistent CAS profiles was evaluated. The most important result from paper IV was the fact that eight laboratories produced highly similar CAS profiles. The ability of designated laboratories to produce reproducible analytical results is a vital prerequisite for reliable investigations of the impurity profiles of CWA samples and origin determination.

What are the most important parts of the studies included in this thesis? I would like to highlight three areas. First of all, the samples. The ability to produce CWA samples for research purposes is obviously necessary when conducting this kind of work. The samples cannot be achieved in any other way than by in-house synthesis, and all four papers are dependent on the quality of the samples. Secondly, the chemical analysis of the samples must be done correctly with suitable instrumentation and analytical methods. The third important area is that of data processing and interpretation. Work in all three of these crucial areas needs to be done with the utmost care in order to draw correct conclusions from the studies.

5.1 The samples

In contrast to forensic studies on illicit drugs, where several thousands of samples may be available, only a limited number of samples were available for the current studies discussed here. In paper I and paper II, the number of samples and the variation within each class were increased by the sample exchange with the LLNL. In the established classification models in paper I and paper II, samples from both laboratories were used.

In paper II, there were 66 unique samples, which is a high number considering the time required to make these samples and the elaborate infrastructure needed in terms of dedicated laboratories and personnel. In paper III, HD from 54 batches from nine production methods were used.
In paper IV, focus was set on the analytical results generated at the participating laboratories and the inter-laboratory reproducibility. The data set consisted of two DC extracts with different CAS profiles. The DC samples in paper IV were used as a model substance for chemical attribution studies.

The synthesis methods used in paper I and paper II were selected to cover both well-established methods and less sophisticated methods. The starting materials used in these syntheses are carefully regulated compounds and their accessibility is highly restricted. Both states and terrorists/non-state actors have the ability to produce CWAs. Terrorists, if they produce the CWAs themselves, are assumed to use methods with starting materials which they can readily access. Terrorists/non-state actors may also steal from state CWA stores. CWAs produced in state weapon programmes have historically been produced in larger production facilities using more advanced equipment.

Real-world CWA samples collected from crime scenes will have different CWA concentrations and the matrices in which the CWA is present will vary, which is an analytical challenge. The work in paper I, paper II, and parts of paper III, was carried out on crude synthesis samples with no background chemicals except for the solvent in which the analytes were dissolved. In paper III, the crude HD samples were applied to two different matrices, i.e. to textile and soil (Figure 10). Although the matrix of the soil and textile samples contained many chemical compounds, the applied data processing method could successfully extract CAS profiles from the HD samples present in these matrices.

Figure 10. Left: soil and textile samples used in paper III. Right: the spiking procedure. Photo: Karin Höjer Holmgren.

In conclusion, the number of samples was limited in our studies, but given the work and resources needed to produce them, the sample sets were adequate
for the current research purposes. Thanks to the co-operation with the LLNL we were able to increase the number of samples.

5.2 Stability of the samples

The CWA samples were dissolved in DCM and stored at -20°C to minimise degradation. These conditions were considered suitable for long-term storage of the samples. Ideally, CAS profiles should be stable over significant periods of time. However, the work conducted in all four papers (Papers I-IV) had issues regarding CAS stability or CAS profile stability. CWAs are highly reactive compounds, and a CWA in combination with CWA-related impurities or by-products may cause reactions resulting in changes in abundance of various impurities over time. The stability of the CAS is of vital importance and was investigated in all four papers. In paper I, the samples were analysed at two different laboratories, situated in Sweden and on the US west coast respectively. The shipping of samples was done without cooling, and thus the transport time is a parameter that could affect the CASs. This study showed that the synthesis variation was larger than the analysis variation.

In paper II, the CAS reactivity varied depending on the specific CAS and the used synthetic route. After six months of storage at room temperature, the CAS profiles of samples made by routes 1, 4 and 7 were nearly unaffected, while samples produced using routes 2, 5 and 8 exhibited some changes to their CAS profiles. Most significant were the CAS profile changes in samples produced by routes 3, 6 and 9. Although some of the profiles changed over time, data interpretation using OPLS-DA could still enable proper separation of the classes in the score plots. Models with samples from different storage times were made, but not validated with test set samples. Some CASs increased with time, for example 1,4-oxathiane and bis(2-chloroethylthiomethyl)ether, while others disappeared. 1,4-oxathiane and bis(2-chloroethylthiomethyl)ether have also been found in HD from long-term storage [37, 38], [83]. One interesting question to consider is if the route differences will decrease over time and gradually become more similar. Is there a time limit when all stored HD samples will have identical CAS profiles?

The stability of the 16 CASs selected for analysis was also examined in paper IV. The DC CAS profile was measured at the start of the study and again after one and two months of storage at room temperature, without any change in their relative distribution.

In the performed study in paper III, both textile and soil samples spiked with crude HD gave changed CAS profiles compared to the crude HD samples. New CASs were formed and the CAS profiles were changed. The classification
models built on the CASs found in the crude samples could not be used for the CASs found in matrix samples. The explanation may lie in both the reactivity of the CASs and the low spiking level. The samples were stored at room temperature overnight before extraction of the CASs. This is a rather short time from addition to extraction, and real samples may be rather more difficult to extract.

My reflection after having performed these four studies is that stability of a CAS is crucially important. The reactivity of CWAs and related impurities will affect the CAS profiles. More stable, high-boiling CASs with higher RI will be easier to detect in samples when a longer time from a CWA attack to sampling has passed, because such compounds are less likely to evaporate. Both outdoor and indoor samples will be affected by weathering and the time from attack to sampling. The most volatile CASs will evaporate, and the most reactive CASs will react with the surfaces on which they sit. Most of the HD CASs in paper II have an RI in the range of 1000 to 2000. This fraction of the chromatogram contains relatively high boiling compounds, whereas more volatile compounds would have lower RI values. Sulfur mustard has an RI of 1178. The CASs in paper I have RI values ranging from 901 to 2457, and VR has an RI of 1732. The distribution of CASs according to their RI is shown in Figure 11. How the CWA is spread to the environment also affects the compounds found after an attack. Aircraft barrels, explosions or leakage from tanks are examples of dispersive methods. If the CWA is spread to the environment by an explosion, the heat will affect the CWA itself as well as its CAS profile. If profiling of CWA samples is needed, the best samples for profiling will be those with the highest concentration of the target analytes (CASs). Attacks with sarin (RI= 820) often give environmental samples where the hydrolysis product isopropyl methyl phosphonate is found when sampling at crime scenes [6]. The reason for this is that sarin is both volatile and highly reactive.
5.3 Chemical analysis of CASs

It is important to understand that the choice of analytical technique will determine the category of compounds that will be detected. When analysing CWAs and related compounds GC/MS is suitable to use since many CWAs are small neutral organic compounds. The aim of the GC/MS analysis in all four papers was the detection of CASs. One major advantage with GC/MS is that all quadrupole instruments with electron ionisation use an electron energy of 70 eV that will give very similar fragmentation patterns of a compound. This enables use of large spectral libraries that are useful for identification of analytes. However, as some compounds are not present in these databases, mass spectral interpretations are often useful as a complement.

In paper I, the VR samples had many by-products and impurities. The selection of CASs was limited by the size of the peak areas, and more potential CASs could have been added. The detection limit of the GC/MS analysis did not limit the CAS detection in the VR sample set; overall 67 CASs were selected.

In paper II, it was decided to include all detectable peaks in the target library. The HD samples were of high purity and therefore many CASs were present at low concentrations. The samples were analysed using two different dilutions, and it was seen that a larger number of compounds were detected in the less diluted sample. For laboratory safety reasons (for both the personnel and the potential contamination of the instruments), samples with even higher CWA concentrations could not be analysed. A chromatogram of a typical HD sample is shown in Figure 12.
In paper IV, the DC samples were hydrolysed and extracted with DCM. The more hydrophobic fraction of the impurities was successfully extracted. This sample preparation method included a concentration step where the volume of the DCM was reduced. This enabled detection of the low-abundance analytes but saturated the detector for the more abundant CASs (Table 3). These results emphasise the difficulties when analysing CAS profiles where the target analytes are present in a wide concentration range. For samples collected after CWA use, the concentration of the CWA and related products as well as other relevant chemicals, will not be known prior to the chemical analysis, and there will be a need to analyse both diluted samples and highly concentrated samples.

![Figure 12. GC chromatogram of an HD sample, and a 50 x zoomed chromatogram. This figure illustrates that the CASs have low abundance compared to HD, but also that the intensities of the target analytes vary.](image)

In paper III, all samples were analysed with orbitrap HRMS. For GC/HRMS there are only minor spectra libraries commercially available, and MS data collected from an orbitrap MS differs from those collected by a quadrupole MS. The advantage with HRMS is the mass resolution, which results in lower detection limits and better selectivity. An example of the improvement in S/N obtained by GC/HRMS in comparison to conventional single quadrupole GC/MS is illustrated in Figure 13. As can be seen, many compounds that could not be detected by quadrupole MS were clearly detected by HRMS.
Table 3. DCM extract of hydrolysed DC sample analysed by GC/MS and concentrated from 2.5 mL to 50 μL. Retention index (RI) and signal to noise (S/N) for DC CASs. ND means not detected, Sat means saturated and “_” that S/N cannot be calculated.

<table>
<thead>
<tr>
<th>DC CAS</th>
<th>Concent. DC extract</th>
<th>DC extract</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RI</td>
<td>S/N</td>
</tr>
<tr>
<td>Pyridiene</td>
<td>738-772</td>
<td>Sat</td>
</tr>
<tr>
<td>Pyridine, 3-chloro</td>
<td>890</td>
<td>465</td>
</tr>
<tr>
<td>Butane, 1,4-dichloro</td>
<td>914</td>
<td>91</td>
</tr>
<tr>
<td>Bis(2-chloroethyl) ether</td>
<td>988</td>
<td>237</td>
</tr>
<tr>
<td>Benzyl chloride</td>
<td>1020</td>
<td>125</td>
</tr>
<tr>
<td>Thiophene, tetrachloro-</td>
<td>1271</td>
<td>149</td>
</tr>
<tr>
<td>Tetradecane</td>
<td>1400</td>
<td>79</td>
</tr>
</tbody>
</table>

¹ Indicates peak area close to the detection limit.

HD was also applied to soil and textiles and the CASs were analysed (paper III). The amount of added HD was low (0.5 mg HD/g matrix), and the study showed that even in such complex sample sets, the route differences - especially compounds related to the chlorination step - were still detectable with GC/HRMS and the used data processing methodology with isotope ratio filtration. This matrix study would not have been possible with ordinary low-resolution single quadrupole GC/MS. Attempts were made to analyse the samples using triple quadrupole GC/MS (GC/MS/MS). However, these experiments were unsuccessful in two ways: (i) it was difficult to select compounds related to route separation in crude samples, and (ii) to detect these compounds in the matrix samples. In the GC/HRMS work (paper III) it was also shown that many new compounds were formed in the matrix samples.
In conclusion, these studies showed that single quadrupole GC/MS analysis successfully separated and detected the target analytes (CASs) present in the CWA samples. The possibility of identifying CASs was also good using spectral libraries and interpretation of mass spectra. For even more challenging tasks, such as detection of different production methods of the HD intermediate or the analysis of HD samples present at low concentrations in various matrices, the use of GC/HRMS was required. For future studies, I would consider using both GC/HRMS and LC/HRMS to detect a wider set of analytes. Samples with high purity have to be analysed at high concentration to be able to detect CASs.

5.4 Data processing

The process for identifying and selecting CASs was a set of important steps in all four papers. In terms of analytical concept and methodology, there has been a shift throughout this work from a targeted (papers I and paper II) to a more non-targeted data processing method (paper III), (Figure 14, Figure 15). The main reason why a non-targeted approach would be more successful is that we do not need to select the target analytes prior to classification analysis. This
advantage is fundamental as we cannot select important CASs for route classification before the samples have been analysed. Hence, by adding as many analytes as possible to the multivariate classification model, the risk of losing the important ones is reduced. On a broader perspective, one can reflect on how the sample preparation method and analytical technique will influence the selection of analytes.

In paper IV, the relevant DC CASs were known beforehand [50], but the DC extracts were still analysed at the beginning of the study to verify their presence. Two more DC CASs were included after discussions within the group of participating laboratories: the aim was to have a limited number of CASs and in the end, it was decided to use 16 CAS.

In paper I we wanted to include a larger, but still manageable, number of CASs. They were selected due to their intensity in the chromatograms and thereby their abundance in the samples. As a selection criterion, all CASs with a chromatographic peak exceeding 0.5% of the total peak area sum were included: this resulted in a total of 67 CASs. This meant that the CAS were selected according to specific criteria (see table 2); the CASs were also added to a target library. The target library contained spectra and RI, and the molecular formula was included, if known.

In paper II, the employed selection criterion required that a CAS should be detected in at least two of the analysed samples in order to be included. The concept of using an area threshold (like in paper I) was not used in paper II. This approach resulted in 103 potential CASs.

In paper III, a non-targeted approach was used. This resulted in 2713 potential CASs and after isotope filtration there were 714 CASs with S and/or Cl. The idea with the non-targeted processing was that the peak selection would be less time-consuming and more automatic. To set up a target library takes time, and every mass spectrum needs to be controlled and accepted as a good representative for the compound in question. With the non-targeted method, criteria for how to define a peak was determined and then the software, TraceFinder, did the data selection, including retention time alignment and integration. The peak areas were extracted ion peak areas. An isotope ratio filter was then applied to the data and only the compounds (in this case defined by one mass and retention time) containing sulfur and/or chlorine CASs were included. This workflow was in theory better than the targeted method since it would be more automated and reduced the risk of manual errors. However, when applied to the data it was more demanding than expected; many processing tests were needed before it worked correctly. For example, the retention time alignment needed chromatography to be run under almost identical conditions to ensure that the retention time variation was within ten seconds.
The criteria for judging the presence of a compound were in practice difficult to apply, and a lot of time was spent on controlling and evaluating the employed parameters. However, the processed data was more informative than the low-resolution data generated by the quadrupole GC/MS. There was sufficient information in the data to distinguish between the synthesis methods of the precursor TGD. For the HD spiked matrix samples this data processing method worked well, and route differences could be detected despite the low spiking levels. Examples of CAS distribution in spiked matrix samples are shown in Figure 15. For a method to be truly non-targeted, the whole process from sample preparation to data processing has to be broad. This was not the case with the work presented in paper III. There might have been better choices of analytical techniques such as LC HRMS to detect polar compounds better. Another choice could have been to TMS-derivatised the HD extracts to include more polar analytes. This approach was investigated in paper II, and the conclusion was not to include these derivatives since they would not give enough extra information and also be very time consuming to include and difficult to process. TMS-derivatives of HD related compounds will have very similar mass spectra since loss of the TMS group is often the base peak. With the changed processing method of the data in paper III, inclusion of TMS-derivatives might have been of value. GC/MS of DCM extracts of HD will detect neutral, thermally stable organic compounds. The use of the isotope filtration is also a selection of a compound and not a true non-targeted method.

**Data processing strategy**

<table>
<thead>
<tr>
<th>Paper I</th>
<th>Peaks &gt; 0.5% TIC intensity from every analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper II</td>
<td>All detectable peaks in chromatograms</td>
</tr>
<tr>
<td></td>
<td>Common laboratory contaminants omitted</td>
</tr>
<tr>
<td>Paper VI</td>
<td>16 pre-selected CASs</td>
</tr>
<tr>
<td>Paper III</td>
<td>Area threshold</td>
</tr>
<tr>
<td></td>
<td>Retention time alignment</td>
</tr>
<tr>
<td></td>
<td>Isotope filtrations</td>
</tr>
</tbody>
</table>

**Figure 14.** The data processing strategies used in paper I, paper II, paper III and paper IV.
In papers I, II and IV, the software AMDIS was used for the data processing. AMDIS detects compounds based on mass spectra and RI, and in this work AMDIS was also used for integration. Integration in AMDIS is not a transparent process and there are no settings to adjust. A comparison of the integrations by AMDIS and the software Chemstation showed that they were almost identical for tested analytes at 2 ppm level (Table 4). AMDIS was developed for compound identification and works well for detection of CASs. The time-consuming step is to set up the library, as all spectra must be added manually. Hence, setting up libraries with many analytes will be time-consuming. However, for a smaller number of compounds, for example the 16 CASs used in paper IV, AMDIS is very convenient. The RI algorithm in AMDIS works well. The advantages of RI were clearly shown in paper IV, where data from eight laboratories were compared. RI calculation of the CASs found in the DC samples

Figure 15. CAS distribution in routes in spiked matrix samples for two CASs from the data processing software TraceFinder. Some of the CASs were present in all spiked matrix samples, upper figure, and other CASs were route-specific and were only present in some of the routes, lower figure. Routes are named R1-R9. Data from paper III. The bars shows the mean peak areas in each route. The coloured circles, triangles and stars show peak areas of individual samples including blank samples.
showed that the different laboratories reported almost identical RI values. Sixteen compounds were analysed in 46 DC extracts by eight laboratories. The between-laboratory RI differences were calculated for each of the sixteen compounds in each of the 46 DC extracts, and the results showed that all RI differences ranged from 2 (minimum) to 27 (maximum). The key subject in the RI calibration in AMDIS is the identification of the alkanes. If one alkane is incorrectly assigned as being another alkane, then the RI calibration will fail. In paper IV, both the RI calibration library and the target library was constructed from FOI data. These libraries worked well when analysing data from the participating laboratories, but one RI calibration failed due to a slightly different fragment pattern resulting in some alkanes being wrongly identified.

**Table 4.** Within-day repeatability of a reference mix (n = 5), and comparison of integration done in the software AMDIS or the software Chemstation using RTE integration. Peak areas calculated as area relative to hexachlorobenzene (HCB). Data from paper IV. RSD - Relative standard deviation.

<table>
<thead>
<tr>
<th>Chemical</th>
<th>AMDIS Peak area relative to HCB</th>
<th>AMDIS RSD (%)</th>
<th>Chemstation Peak area relative to HCB</th>
<th>Chemstation RSD (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyridine</td>
<td>2.37</td>
<td>2.3</td>
<td>2.40</td>
<td>2.4</td>
</tr>
<tr>
<td>Chlorobenzene</td>
<td>1.26</td>
<td>2.1</td>
<td>1.35</td>
<td>1.3</td>
</tr>
<tr>
<td>Bis(2-chloroethyl) ether</td>
<td>0.90</td>
<td>1.6</td>
<td>0.86</td>
<td>0.5</td>
</tr>
<tr>
<td>Benzyl chloride</td>
<td>1.36</td>
<td>0.8</td>
<td>1.34</td>
<td>0.9</td>
</tr>
<tr>
<td>Ethane, hexachloro-</td>
<td>2.18</td>
<td>0.8</td>
<td>2.11</td>
<td>0.7</td>
</tr>
<tr>
<td>Dodecane</td>
<td>1.61</td>
<td>1.2</td>
<td>1.54</td>
<td>0.8</td>
</tr>
<tr>
<td>Tridecane</td>
<td>1.79</td>
<td>1.2</td>
<td>1.73</td>
<td>1.1</td>
</tr>
<tr>
<td>Tetradecane</td>
<td>1.93</td>
<td>1.0</td>
<td>1.91</td>
<td>0.5</td>
</tr>
</tbody>
</table>

My view, after having used several methods for data processing, is that a combination of targeted/non-targeted methods is probably the best approach. It takes time to set up a target library, but it also takes time to do the non-targeted data processing. In cases where additional samples are added at a late stage of an attribution study, it will cause additional work with either of the two strategies. With the target library set up, there might be new markers in the new set of samples that need to be added to the library and the old samples need to be reprocessed after the update. For the non-targeted data processing, new data will be demanding since the retention time will differ and the retention time alignment might need manual adjustments. In conclusion, it is beneficial to use a non-targeted approach in the discovery phase of a study and then select CASs to be included in a target library for future use.
5.5 Identification of CASs

The identity of all CASs must not be known when applying multivariate data analysis to the data set. The CASs, the variables, can be named after retention time, RI or by retention time and fragment. However, even if identification is not necessary when constructing classification models, it is always beneficial to have the structure of the CASs. By knowing the chemical structure of the CASs, their presence in the samples can be understood. Are the CASs related to impurities in the solvents used? Is it found in any of the starting material used in the synthesis, or is it a laboratory contaminant and therefore might it vary independently of the applied synthetic routes?

To fully identify all CASs in large sample sets is probably not a realistic goal, mainly due to the time needed to do so. Instead, a selection of the most relevant CASs must be done and then these can be identified. Although mass spectra libraries consist of many common compounds, there are still many substances that are not included in the libraries. Identification can be done to various degrees. The first step is establishing the elemental composition of the analyte. A full identification involves determining a structure confirmed by analysis of a reference analyte [84]. In the work included in this thesis, a range of different tools have been used to identify the CASs. Initially, spectra libraries were used. A match in a spectra library was further evaluated by comparing the calculated RI with the reference value. Furthermore, the suggested structure was also checked for relevance with respect to the synthesis set up. GC/MS CI often gave molecular ions and that could be used for confirmation of the proposed structure. When available, reference compounds were also analysed to confirm the identification. For compounds with no match in spectra libraries, mass spectra interpretation was done. Ideally, the GC/MS CI gave a molecular ion weight to start with. Isotopic pattern was also evaluated.

To unambiguously identify a CAS, a reference sample needs to be analysed under identical analytical conditions. If retention time and mass spectra are identical compared to the reference sample, then the CAS is identified, although with the exception that isomers and stereoisomers are normally not specified.

The identification of CASs has been successfully done to varying degrees in the papers in this thesis. In paper I, there were a limited number of compounds due to the imposed intensity restriction of only looking at chemicals with a higher intensity than 0.5% of the total TIC area. This resulted in a sample set of 49 compounds where the majority were given a proposed structure by matches in spectra libraries or interpretation of mass spectra in combination with CI data (35/49). All samples were also analysed by GC/MS CI to confirm
the proposed molecule ion. GC/HRMS was used on one representative sample from each synthetic route and GC/HRMS improved the identification since the elemental composition of fragments can be calculated with a high degree of certainty. An example of this is given in Figure 16. The position of one of the sulfur was unclear prior to HRMS analysis, since the fragment m/z 130 was not clearly seen in low resolution MS. The presence of fragment 130.0681 in HRMS strengthened the hypothesis that the sulfur would be positioned at the diethyl amino side of the molecule rather than the isobutyl side.

Figure 16. Example of how elemental composition from HRMS improved structure determination O-isobutyl S-(2-diethylaminoethyl) methylphosphonodithioate.

The production of VR involved several synthesis steps, and each step can give rise to side reactions or have impurities. Some by-products that were formed in the early steps of the synthesis remain throughout the synthesis and can be detected as impurities in the VR sample. The presence of the impurities carries information about the applied reaction steps (Figure 17). Each reaction step might not have 100% recovery and unreacted starting material and/or intermediates are often present.

\[ \text{Characteristic fragments for all VR chemicals} \]

- \( m/z = 86 = \text{C}_6\text{H}_{12}\text{N} \) (base peak)
- \( m/z = 99 = \text{C}_6\text{H}_{13}\text{N} \)

\[ \text{Low-intensity fragments from GC/HRMS analysis} \]

- \( m/z = 154.97502 = \text{C}_3\text{H}_8\text{POS}_2 \) (option 1 and 2)
- \( m/z = 183.00609 = \text{C}_6\text{H}_{12}\text{OPS}_2 \) (option 1 and 2)
- \( m/z = 130.06861 = \text{C}_6\text{H}_{12}\text{NS} \) (option 1)
- \( m/z = 114, \text{not seen} \)
Figure 17. Examples of how identification of CAS gave information about synthesis steps. All chemicals in this Figure, except methylphosphonic dichloride and methylphosphonothionic dichloride, were detected in the crude VR samples.

In paper II, all detectable CASs were included in the target library, which resulted in 103 CASs. The identification process was here much more demanding; 24 CASs were identified by spectra library hits, and 26 CASs were preliminarily identified using spectral interpretations and Cl isotope pattern. HD-related analytes are not well suited for Cl analysis; many of the analytes will not give strong molecular ions, and instead the same base peak as seen in EI was formed. The remaining 53 CASs were not assigned any structure or molecular formula, but could be retrieved in the chromatograms based upon their RI, mass spectra and, in some cases, their molecular weight that was determined by Cl. In paper III, where the analysis was performed with GC/HRMS, some of the preliminary structure of the tentatively identified CASs from paper II, could be verified. An overview of CASs related to HD production routes are shown in Figure 18.

In paper I, the VR-related impurities that were available in-house were used for confirmation of identity. No references were run in the HD work, mainly because the majority of the CASs were not present as references and the work to produce them would have been too time consuming.

In paper IV, 16 CASs were used and 11 were available as references at the FOI or VERIFIN. These 11 compounds were analysed at the FOI and VERIFIN,
and a spiked extract was sent to the participating laboratories along with the DC samples.

**Figure 18.** Examples of CASs related to HD routes. Simplified Figure from supplementary material in paper III. R1 to R11 are the production methods.

Unknown substances can be identified by interpretation of their mass spectra. The isotopic pattern and fragmentation processes such as α cleavage, σ cleavage, inductive cleavage and hydrogen rearrangement [85] are used in the interpretation work. To study mass spectra of similar compounds is also helpful. One example of how the interpretation can be done is seen in Figure 19 (paper II).
My conclusions are that even if identification can be difficult and time-consuming, it does give valuable additional information. To know the structure of analytes formed under specific conditions or with starting materials gives a strength to their importance. I would recommend for future studies carrying out more work on the identification of the CASs. By knowing their structures, it is easier to discuss how and why they are formed and how stable they will be. The use of HRMS improves the possibility of identifying CASs since the elemental composition can be determined in both the molecular ion and major fragments.

5.6 Multivariate classification models

The aim of the statistical analysis was to evaluate whether or not samples produced by different routes had unique impurity profiles. This was done using PCA, PLS-DA, OPLS-DA and RF. A comparison of a selection of PLS models established in the publications is shown in Table 5. In the VR study the data was complex, and the PLS model needed ten PCs to explain the variation within the data. In paper II, the applied OPLS-DA model removed the parts of the variation in the X matrix (i.e. pre-treated peak area values of target analytes) that were not correlated to route separation (Y matrix), which resulted in models with fewer components. A PCA of the data used in paper II is shown in Figure 20.
Table 5. Comparison of models for VR synthesis method (paper I), final step HD synthesis (papers II and III) established in the papers.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>57</td>
<td>66</td>
<td>36</td>
<td>108</td>
</tr>
<tr>
<td>Variables</td>
<td>49</td>
<td>103</td>
<td>714</td>
<td>514</td>
</tr>
<tr>
<td>Classes</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Classification method</td>
<td>PLS-DA</td>
<td>OPLS-DA</td>
<td>OPLS-DA</td>
<td>OPLS-DA</td>
</tr>
<tr>
<td>PCs</td>
<td>10</td>
<td>4+3+0</td>
<td>2+2+0</td>
<td>2+3+0</td>
</tr>
<tr>
<td>R²X</td>
<td>0.89</td>
<td>0.71</td>
<td>0.61</td>
<td>0.46</td>
</tr>
<tr>
<td>R²Y</td>
<td>0.93</td>
<td>0.95</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>Q²</td>
<td>0.85</td>
<td>0.89</td>
<td>0.95</td>
<td>0.80</td>
</tr>
</tbody>
</table>

Figure 20. Score plot of PC1 and PC2 from a PCA model of sulfur mustard impurities. Sample coloured according to synthesis route.

In paper III, many models were formed, starting with the separation of sample according to intermediate, then separation according to the chlorination method and lastly according to the production method of TDG. This method resulted in reduced complexity of the data and gave models with fewer components. The constructed models have a high predictive ability; this was also confirmed when test set samples with known production methods were used for model validation.
In papers I and II no new samples for validation of the models were available. Instead, a selection of samples was removed from the data set before modelling and was thereafter used as a test set. An alternative way of doing the studies (paper I and paper II) would have been to use the FOI samples to build the model and the LLNL samples for validation. In paper III, a set of separately produced HD samples was used as test set samples.

Two classification methods were used in paper III. RF and OPLS-DA models showed similar predictive performance in paper III, but there were still differences. OPLS-DA can classify samples as not belonging to any of the classes (Table 6), while RF will force the samples into one of the used classes. For unknown samples, there may be more misclassifications if RF models are used. One advantage with RF is that it can be used with many classes. Two examples of models showings classes according to chlorination models are shown in Figure 21 and Figure 22. Both crude samples and extracts from matrix spiked with sulfur mustard could be classified according to used chlorination methods.

In paper I, models both with, and without TMS-derivatised CASs were built. A reduction in the number of CASs used for model generation was evaluated. The predictive ability of the model was not reduced when 20 out of 67 variables were removed. Variable reduction was not attempted in paper II and paper III (it might have been interesting to do so). Most probably there is a number of potential CASs included in these models not contributing to route separation.

**Table 6.** Comparison of correct predictions from OPLS-DA and RF classification models using spiked matrix data and models for the chlorination method of HD. The chlorination methods are named A, B and C. Correct predictions are in bold. Data in this Table are from paper III.

<table>
<thead>
<tr>
<th>Correct method (nb of samples)</th>
<th>Predictions OPLS-DA</th>
<th>Predictions RF</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>A (10)</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>B (11)</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>C (12)</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
Figure 21. Model constructed samples (filled) and test set samples (non-filled) in a score plot of a PLS-DA model of sulfur mustard impurities, classes according to chlorination methods. Sample coloured according to synthesis route, filled circles are model constructing samples. Figure simplified from paper III.

Figure 22. Model constructed samples (filled) and test set samples (non-filled) in a score plot of a PLS-DA model of sulfur mustard impurities in spiked matrix samples, classes according to chlorination methods. Sample coloured according to synthesis route. Figure simplified from paper III.
From the impurity profiles from the DC study, paper IV, euclidian distances were calculated (Table 7). These distances can be used for separation of samples with same source or different source. In this case it is same production method or different production method. By construction a histogram the two classes of distances can be clearly seen, Figure 23. From these histogram can a kernel density estimation be made. This is then further used when calculating likelihood ratios of two samples where the question is if they have the came from the same source or different source. In this example, there is no overlap in the histograms and the profiles for the two DC production methods are clearly different. This data would need to be expands with DC impurity profiles from other methods and more replicates for the “same method” data as well. The use of likelihood ratio is tool used in forensic science for showing the strength of an evidence.

**Table 7.** Calculations of euclidian distances (ED).

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Mean ED</th>
<th>SD</th>
<th>Number of ED</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC1, within method</td>
<td>0.122</td>
<td>0.088</td>
<td>570</td>
</tr>
<tr>
<td>DC2, within method</td>
<td>0.100</td>
<td>0.059</td>
<td>529</td>
</tr>
<tr>
<td>DC1 vs DC2, between method</td>
<td>0.866</td>
<td>0.025</td>
<td>552</td>
</tr>
</tbody>
</table>

**Figure 23.** Euclidian distances of DC2 samples, all individual profiles compared. Blue bars indicate same source and orange bars indicate different source comparison.

Another way of visualization of same source/different source is by using PCA. By scaling the data one can decide if all variables should have similar impact or
if the largest one should have more impact in the models. Different scaling and pretreatments, and how it affects the score plots, are shown in Figure 24.

My view about the classification work is first of all that it works really well using multivariate classification models for this type of research. Classification models were built, had good predicative power and were successfully validated with test set samples. We could identify important CASs related to specific routes. The most important CASs were also manually seen in the chromatograms. One common issue for all three studies (paper I, paper II and paper III) is the usability of the models. They are all built on samples produced at one or two laboratories, with highly defined methods, where the aim was to produce samples as similar as possible. The models have then been validated with a sub-set of these samples, or samples produced later on but still using identical methods and starting materials. The variation studied within these models/samples is likely to be less than the real variation. There is thus a need for studies investigating the variations within each method and how the samples will be affected by storage conditions and length of time spent in storage. The models developed in paper I, paper II and paper III, can probably be used for authentic samples. However, it is the knowledge gained that is the great value of this work. Hopefully, the model will not be used for real cases, since this sample-is (fortunately) extremely rare, and the probability that samples would contain HD or VR is lower still. Instead, I believe that the knowledge developed when making the samples, analysing the CASs, identifying the CASs, interpreting and understanding why a certain CAS is formed in specific routes, is the true value. This knowledge is unique and will be used for a long time. An issue for the future is how to translate values such as are predicted into probabilities of belonging to a class. It will always be important to know the extent of variation within a class and that will have to be estimated or experimentally determined.
Figure 24. PCA models of DC1, DC2, and reference data. Samples coloured according to laboratory. A) UV-scaled data 2PCs, $R^2=0.81$, $Q^2=0.77$, B) Centered data 2PCs, $R^2=0.99$, $Q^2=0.98$, C) UV and $\log(x + 0.00001)$ scaled data 2PCs, $R^2=0.75$, $Q^2=0.68$. 
5.7 Recommendations for future studies

After worked with the attribution studies in paper I, II, III, and IV, I would like to make some recommendations for future studies in this area and also to point out how to use this knowledge in real cases. A key step is to have reference samples that have been analysed with known impurity profiles. If no reference samples exist the laboratory need to make a large sample set and the time line for answer a specific attribution question will be longer.

It is important to separate attribution studies from sample matching studies. Sample matching aims at investigating if two samples originate from the same source or from different source. While attribution studies are done to increase knowledge of impurities in threat samples related to for example synthesis method, precursor method, or batch differences. The goal is to be able to source new unknown samples to any of the investigated parameters. The choice of analytical techniques depends of the complexity of the study and how small differences in impurity profiles there will be. Chromatography in combination with mass spectrometry is well suited for analysing organic impurity profiles. The data processing of impurity profiles in attribution studies are well suited to be non-targeted approaches in discovery phases of a study. Non-targeted methods can be used for detection of important chemical attribution signatures. Then can target library of detected impurities be created. For sample matching cases, reference samples need to be analysed and relevant impurities selected and added to a target library. The impurity profiles (peak areas of the impurities) of the reference samples are further used. Distance measures such as euclidian distances can be used for comparison of impurity profiles. Multivariate data analysis models such as PLS-DA can also be used for classification of the reference samples. Sample matching reference data is also needed for the setup of kernel density estimation, these equation is needed for the calculation of likelihood ratio when two samples of unknown source are to be compared (Table 8). The value of the likelihood ratio shows a defined probability that the unknown samples have origin from the same source (or different source) when using the predefined reference data. A high value show a high probability of having same origin. For attribution studies, the reference samples are used for finding CASs and adding them to a library. Many reference samples are needed to cover the expected variance in the area. For the use of classification models for new samples it is important to ensure that the new samples are similar enough to be used with the developed model (Table 9).
<table>
<thead>
<tr>
<th>Same source or different source</th>
<th>Are reference samples available?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes,</td>
<td>-analyse new samples with same method as used for the reference samples.</td>
</tr>
<tr>
<td></td>
<td>-use target library for impurity profile detection.</td>
</tr>
<tr>
<td></td>
<td>-calculate ED for the two new samples.</td>
</tr>
<tr>
<td></td>
<td>-use KDE and get values for same source and different source</td>
</tr>
<tr>
<td></td>
<td>-calculate likelihood ratio for samples belong to same source and difference source using x number of reference samples. This gives a specific probability that the samples have same/different source.</td>
</tr>
<tr>
<td>No,</td>
<td>-make reference samples, if possible.</td>
</tr>
<tr>
<td></td>
<td>-analyse and determine relevant impurity profiles.</td>
</tr>
<tr>
<td></td>
<td>-calculated ED for reference samples</td>
</tr>
<tr>
<td></td>
<td>-make KDE for same source data and different source data.</td>
</tr>
</tbody>
</table>

Table 8. Suggested work flow for setting up same source/different source studies of threat samples. KDE – kernel density estimation.

<table>
<thead>
<tr>
<th>Attribution</th>
<th>Are reference samples available?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes,</td>
<td>-analyse new samples with same method as used for reference samples.</td>
</tr>
<tr>
<td></td>
<td>-use target library for impurity profile detection.</td>
</tr>
<tr>
<td></td>
<td>-create classification model.</td>
</tr>
<tr>
<td>No,</td>
<td>-make reference samples, if possible.</td>
</tr>
<tr>
<td></td>
<td>-analyse and determine relevant impurity profiles.</td>
</tr>
<tr>
<td></td>
<td>-create classification model.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Can classification model be used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes,</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>No,</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table 9. Work flow for future set up of attribution studies.
6. Conclusions

Analytical and multivariate methods for source attribution of HD, Bis(2-chloroethyl) sulfide and VR, Isobutyl S-2-diethylaminoethyl methylphosphonothiolate were studied. This thesis shows that the impurity profiles of CWAs, their “chemical fingerprints”, can be measured by GC/MS and used to classify samples according to synthetic route. The approach for selecting target compounds that will be used as variables in the multivariate modelling is critical. In this work, both targeted and a non-targeted approaches were used. Based on our current studies, the author’s conclusion is that a non-targeted approach is beneficial at an early stage of method development to ensure that all important variables are detected. Once a proper set of target compounds has been established, a targeted approach is then advantageous as it offers a more straightforward and efficient data evaluation. The use of GC/HRMS (paper III) resulted in more informative data sets and enabled a more detailed route sourcing than single quadrupole GC/MS. GC/HRMS offers better resolution and sensitivity, which enabled detection of lower abundant CASs resulting in data sets of improved quality. Hence, it was found necessary to apply GC/HRMS to differentiate between the applied synthetic route based upon the obtained impurity profiles of the starting material thiodiglycol (paper III). Source attribution for CWA samples spiked to matrix samples also required the use of GC/HRMS as the selected CASs were present at low concentrations. New CASs were also formed in the spiked matrix samples due to the reactivity of these compounds.

The GC/MS profiling method developed (paper IV) is robust and exhibits good between-laboratory reproducibility. The GC/MS methods that were applied in different laboratories were not strictly harmonised and allowed the use of in-house developed analytical methods and different instrumentation. This result is important as it allows for data exchange and international collaboration that could reveal the method used in the synthesis of a CWA and possibly identify who is responsible for a CWA attack.
7. Future perspectives and reflections

I believe chemical forensics and attribution studies on CWAs will continue to develop. This is a relatively new research area and much remains to be explored. More research studies are needed to improve our knowledge of CASs related to synthetic routes and batch differences, and the effects of storage and weathering of CWAs. The stability of a CAS is a key subject in these studies. Studies focusing on how CAS profiles can vary when doing the same synthesis with slightly different settings would also be of importance for the CWA research community.

The implementation of attribution studies into the OPCW’s work is another interesting issue. How will it be done? How can it be used and how will the designated laboratories contribute to this important area in the future? When investigating crimes scenes where a CWA has been deployed, the chemical analysis is only one part among many important things to consider. Methods for making sample matching need to be evaluated and published. How similar can we expect CAS profiles to be? Do all CASs need to be present? Are their relative intensities something that can be used when evaluating similarities in sample matching cases? If the OPCW would like to use the designated laboratories in any future investigations of crimes against the CWC where sample matching or origin of sample are to be investigated, then the designated laboratories need to start training now for these important tasks. Training should focus on how to perform the analysis, how to report data and how to perform sufficient quality control. One way forward would be to start on a small scale with a number of samples that can either be from the same source or different sources.

Another area for future development is analysis of CWA impurity profiles in biomedical samples. In biomedical CWA samples, it is not only the CWA that will react to form adducts or degradation products - the impurities can also react in a similar or totally different way within the body. CWA impurity profiles in biomedical samples is not a well-studied subject. There is one publication by John et al. [7], from an attack in Syria where sarin, degradation products and one impurity were detected in samples collected from a deceased victim. Sulfur mustard and related compounds in biomedical samples have also been studied [86, 87]. The development of instrumentation and data processing tools are promising for future studies of this kind that will require advanced analytical methods.

The major outcome from this thesis is probably the simple fact that the studies actually were published. Other defence research groups can now be
inspired and contribute route attributions studies using other CWAs or by investigating other parameters that will affect the CAS distribution. I also believe that it is the CASs identified in these studies that will be of most importance for us and others going forward. The classification models are the essential tools needed for finding the most relevant CASs for route separation.

Defence research in general can, of course, be sensitive. Not all research conducted at the FOI is suitable for publications since it may deal with gaps in defence capabilities or show the current state of knowledge in a particular area. However, it is important to share data processing methods with others and to discuss potential improvements. Our knowledge of sourcing CWA samples to specific batches, production methods or to link samples to each other can be one key among others in investigations seeking to identify the perpetrators after CWA attacks. We believe that our research will help the OPCW and UN and other organisations to attribute the use of CWAs to the source – be that nation states, specific batches or precursor chemicals. We also believe that this will, in the end, contribute to a world free of chemical weapons.
8. Populärvetenskaplig sammanfattning

Trots att kemiska stridsmedel (CWA) är förbjudna att använda så har ändå brott mot kemvapenkonventionen skett regelbundet, framförallt under det sista årtiondet med inbördeskriget i Syrien där sarin, senapsgas och klorgas använts. CWA har dessutom använts vid politiska mord och mordföröreningar. Vid sådana händelser är det viktigt för den juridiska processen att identifiera vilket CWA som använts och att hitta information om hur det tillverkades. Utvecklingen och användandet av avancerade analysermetoder i kombination med multivariata dataanalyser krävs för att producera forensiska bevis och underhållsbevis.

Denna avhandling beskriver forskning som genomförts med syfte att spåra syntesvägar för CWA. I tre studier har metoder utvecklats för detta och deras möjlighet att bestämma syntesväg bygger på användandet av föroreningar och biprodukter. Den relativa fördelningen av dessa ämnen ger upphov till en unik profil, ett kemiskt fingeravtryck, som kan användas för att retrospektivt se vilken syntesväg som används vid tillverkningen av ett visst CWA. Studien i den första publikationen är gjord på nervgasen rysk VX, S-[2-(dietylamino)ethyl] O-isobutyl metylfosfomentioat, och den andra och tredje studien gjordes på senapsgas, bis(2-kloroethyl)sulfid.

Avhandlingen beskriver försöksupplägg, val av kemiska analysermetoder, metoder för datautvärdering samt hur klassificeringsmodeller har använts. Studierna visade att klassificeringsmodeller tydligt kunde särskilja de sex produktionsmetoderna som användes i den första studien och de fem produktionsmetoderna som användes i den andra studien. I den tredje publikationen användes ett nytt förutsättningslöst utvärderingssätt i kombination en högupplöst masspektrometer vilket ledde till detektion av fler lågkonzentrationsämnen i senapsgasproverna. Metoden gav data med tillräcklig information för att klassificera prover utifrån syntesmetod för utgångsämnet tiodiglykol.

Alla klassificeringsmodeller har validerats med hjälp av testset med goda resultat. Alla testset-prover predikterades rätt i rysk VX-studien och första senapsgas-studien. Klassificeringen av syntesmetod för tiodiglykol, utifrån senapsgasanalyser, i den tredje publikationen var svårare men metoden kunde trots detta uppnå korrekt klassificerade av 56-89 % av testset-proverna. Förutom bildandet av klassificeringsmodeller har ämnen som är viktiga för separation av prover utifrån syntesväg, identifierats. Detta har lett till en ökad kunskap om vilka föroreningar som bildas vid dessa synteser. Stabiliteten på dessa föroreningar har undersökts och huvuddelen av dessa var stabila i rumstemperatur.
9. References


[19] OPCW, Statement by the director-general in response to the adoption of two decisions under article XV of the chemical weapons convention to amend the annex on chemicals, 2019, C-24/DG.20.


[70] C. Veenaas, P. Haglund, Methodology for non-target screening of sewage sludge using comprehensive two-dimensional gas chromatography coupled to high-resolution methvelica


10. Acknowledgements

Forskningen som presenterad i avhandlingen är finansierad av myndigheten för samhällsskydd och beredskap genom medel för transatlantiskt samarbete. Forskningen är även finansierad av försvarsdepartementet.

Jag vill tacka FOI som gett mig möjligheten att fördjupa kunskaperna inom kemiska profilering. FOI är en fantastisk arbetsplats där jag kunnat utvecklas och fått lära mig nytt hela tiden. Det doktorandperioden gett mig är en större insikt i värdet av att kommunicera forskning, hur viktigt det är att presentera forskningsresultat i publikationer och att presentera forskningsresultat på ett begripligt sätt på konferenser och workshops så att andra kan lära av oss och vi kan vara med i utvecklande samarbeten.

Stort tack till avdelningsledningen med Åsa, Niklas och Per som gett mig den här möjligheten. Och tack till Yvonne som var delaktig till att jag blev doktorand.

Tack till Johan vid Linköpings universitet, för att du gjorde det möjligt för mig att bli doktorand hos dig och all hjälp jag fått under jobben med publikationer och avhandlingen. Du har alltid haft tid för mig och det har inte känts som vi haft 80 mil mellan oss, utan väldigt nära och enkelt. Jag har särskilt uppskattnat ditt väntliga sätt och din positiva attityd.

Jag vill tacka Crister för att ha varit en utmärkt handledare som peppat mig när så behövts och däremellan kommit med bra synpunkter för att driva arbetet framåt och för givande diskussioner om forskningen. Och en bra reskompis!


Papers

The papers associated with this thesis have been removed for copyright reasons. For more details about these see:

https://doi.org/10.3384/9789179295844