UNIVERSITEIT VAN AMSTERDAM

MSc in Chemistry
Analytical Sciences

Literature Thesis

Analytical methodologies for the characterisation of aquatic natural organic matter

---

by

Dennis Vughs
UvA: 10771506

June 2018

12 ECTS
April 2018 – July 2018

Supervisor/Examiner:
Prof. Dr. W.P. de Voogt
Supervisor:
Dr. A. Kolkman

Examiner:
Dr. J.R. Parsons

KWR Watercycle Research Institute

Abstract

Natural organic matter (NOM) is a complex heterogeneous mixture of naturally occurring organic compounds which exist ubiquitously in natural waters worldwide. The presence of NOM in water has a negative impact on drinking water quality and treatment processes, due to colour, taste and odour, fouling of membranes, promoting biological growth. NOM can act as a precursor for the formation of disinfection by-products. In order to minimise these unwanted effects, it is of great importance to significantly reduce the concentration of NOM in treated water. Therefore proper characterisation of NOM in raw and drinking water is needed, for which many different analytical methodologies are currently used.

In this literature study, a review of methods used for NOM analysis and characterisation is provided, including methods for isolation/concentration (e.g. reverse osmosis, freeze-drying and passive sampling) and fractionation (e.g. resin fractionation, size exclusion chromatography and solid phase extraction), followed by general analysis methods (e.g. dissolved organic carbon, specific UV absorbance and fluorescence analysis) and in-depth characterisation methods (e.g. Fourier transform infrared spectroscopy, nuclear magnetic resonance and Fourier transform ion cyclotron resonance mass spectrometry). And finally applications of hyphenated analytical techniques are discussed (e.g. size exclusion chromatography - organic carbon detector, liquid chromatography - high resolution mass spectrometry and pyrolysis gas chromatography - mass spectrometry). In addition, recommendations are made for the most fitting analysis techniques for NOM characterisation to be used or implemented at KWR and an analytical approach is provided using LC-HR-MS for the comprehensive characterisation of NOM.

Abbreviations

AF4	: Asymmetrical Flow Field-Flow Fractionation
DBP	: Disinfection By-Product
DI	: Direct Infusion
DOC	: Dissolved Organic Carbon
DOM	: Dissolved Organic Matter
DON	: Dissolved Organic Nitrogen
EEM	: Excitation Emission Matrix
ELSD	: Evaporative Light Scattering Detector
ESI	: Electrospray Ionisation
FA	: Fulvic Acid
FFF	: Field-flow fractionation
FTICR-MS	: Fourier Transform Ion Cyclotron Resonance Mass Spectrometry
FTIR	: Fourier Transform Infrared Spectroscopy
GC	: Gas Chromatography
HA	: Humic Acids
HILIC	: Hydrophilic Interaction Chromatography
HPLC	: High-Performance Liquid Chromatography
HPSEC	: High Pressure Size Exclusion Chromatography
HR-MS	: High Resolution Mass spectrometry
HS	: Humic Substances
ICP-MS	: Inductively Coupled Plasma Mass Spectrometry
LC-MS	: Liquid Chromatography – Mass Spectroscopy
MF	: Microfiltration
MS	: Mass Spectrometry
MSSVPy	: Microscale Sealed Vessel Pyrolysis
NOM	: Natural Organic Matter
NF	: Nanofiltration
NHS	: Non-humic Substances
NMR	: Nuclear magnetic resonance
OCD	: Organic Carbon Detector
OND	: Organic Nitrogen Detector
PARAFAC	: Parallel Factor Analysis
PDA	: Photodiode Array
POC	: Particulate Organic Carbon
POM	: Particulate Organic Matter
PRAM	: Polarity Rapid Assessment Method
Pyr-GC-MS	: Pyrolysis Gas Chromatography - Mass spectrometry
RO	: Reverse Osmosis
RP	: Reversed Phase
SEC	: Size Exclusion Chromatography
SPE	: Solid Phase Extraction
SUVA	: Specific Ultraviolet Absorbance
TDN	: Total Dissolved Nitrogen
TOC	: Total Organic Carbon
UF	: Ultrafiltration
UV	: Ultraviolet
UV-Vis	: Ultraviolet-visible Spectrophotometry

Contents

1	**Introduction**	**5**
1.1	Introduction	5
1.2	Natural organic matter	6
1.3	Dissolved black carbon	8
2	**Isolation and concentration methods**	**9**
2.1	Isolation and concentration	9
2.2	Fractionation methods	10
3	**General analysis methods and in-depth characterisation methods for NOM**	**15**
3.1	Analytical methods for determining general parameters of NOM	15
3.2	Analytical methods for elemental and structural identification of NOM	17
4	**Chromatographic and hyphenated characterisation methods and applications**	**25**
4.1	Chromatography characterisation methods	25
4.2	Liquid chromatography	25
4.3	Gas chromatography	30
4.4	Field-flow fractionation	33
5	**Recommended isolation and analytical methodologies for the characterisation of NOM at KWR**	**35**
5.1	Isolation, concentration and fractionation methods	35
5.2	General analysis methods	35
5.3	Hyphenated and in-depth characterisation methods	36
5.4	Overview recommended analytical methodologies for NOM characterisation at KWR	37
5.5	Analytical approach for NOM characterisation using the LC-Orbitrap Fusion	37
6	**Conclusion**	**39**
7	**References**	**40**

1 Introduction

1.1 Introduction

Natural organic matter (NOM) is a complex heterogeneous mixture of naturally occurring organic compounds which exist ubiquitously in natural waters. NOM originates from organisms (dead or living) such as plants, animals, microorganism, and transformation products of these sources [1]. The composition and amount of NOM in natural water depends on its surrounding biogeochemical environment (e.g. source of organic matter, biological processes, temperature, pH and ionic strength), geology, topography and climate [2-5]. The presence of NOM in water has a negative impact on drinking water quality and treatment processes due to its colour, taste and odour, fouling of membranes, promoting biological growth and corrosion in water distribution networks. NOM can act as a precursor for the formation of disinfection by-products (DBPs) during advanced oxidation water treatment [6, 7]. In particular the formation of these DBPs are of great concern, which have shown to possess mutagenic and genotoxic properties and cause adverse health effects [8-10]. Over 600 DBPs have previously been identified (e.g. trihalomethanes and haloacetic acids) and new emerging DBPs are still found due to improving detection levels and methods [9, 11, 12].

In order to minimise these unwanted effects, it is of great importance to significantly reduce the concentration of NOM in treated water. Due to global warming the concentration of NOM is increasing worldwide [13] and also the seasonal variability of NOM increases which presents major challenges for drinking water treatment. Therefore proper characterisation of NOM in raw and drinking water is essential for monitoring and optimisation of water treatment. Other applications in which NOM characterisation is important are: determining the quality of reclaimed water in waste water treatment in which NOM plays a central role for reuse [14] and for reclaimed NOM from recycled waste water compost [15] or reclaimed NOM from drinking water treatment which can be used as a soil improver [16].

Because NOM consists of thousands of different organic compounds it is impractical to characterise each individual compound. For that reason traditionally NOM characterisation was performed using general parameters such as dissolved organic carbon (DOC), dissolved organic nitrogen (DON), fluorescence and specific ultraviolet absorbance (SUVA). Significant advances in instrumentation have led to more compositional information on NOM such as Fourier transform infrared spectroscopy (FTIR), nuclear magnetic resonances spectroscopy (NMR), liquid chromatography-mass spectrometry (LC-MS), pyrolysis gas chromatography-mass spectrometry (pyr-GC-MS) and Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS).

Due to increasing water scarcity [17] (due to climate change and increasing populations) and water quality problems, more efficient water treatment methods are needed for the removal of NOM, which requires more information on the organic matter present in water, and increased necessity for extensive NOM characterisation. Due to these reasons KWR Watercycle Research Institute (KWR), an institute for water research located in the Netherlands, has expressed the ambition to perform NOM characterisation. Currently KWR has limited experience with extensive NOM characterisation, and requested an overview of the available analytical methodologies for NOM characterisation and recommendations about which techniques are interesting for implementation at KWR.

This literature thesis aims to summarise the isolation and fractionation methods and analytical techniques used for the characterisation of NOM in water, and to provide recommendations for the most fitting analysis techniques for NOM characterisation at KWR.

1.2 Natural organic matter

Natural organic matter (NOM) is a term used to designate all carbon-based compounds found within natural and engineered ecosystems, other than living organism and compounds of anthropogenic origin. Aquatic NOM originates from the decomposition of terrestrial soil and plant materials and from by-products of algae, bacteria and aquatic plants. NOM is present in all natural waters, is composed of a very complex mixture of organic compounds and possesses a large variety of properties. Due to the interactions between the hydrologic cycle, geosphere and biosphere, NOM is found in most sources of drinking water [18]. NOM in aquatic systems can be divided into two fractions; a dissolved organic matter (DOM) fraction and a particulate organic matter (POM) fraction. DOM is the truly dissolved fraction of NOM which passes through a 0.45 µm pore-size filter. POM consist of particles which cannot pass a 0.45 µm filter.

Humic and non-humic substances

Aquatic NOM consist of humic substances (HS) and non-humic substances (NHS) (see Figure 1-1). Humic substances are a complex heterogeneous mixture of mainly unidentified organic material, which originates from the decomposition of terrestrial soil and plant materials. The proportion of humic substances in aquatic NOM is in the range of 50-70% [19]. HS are generally less biodegradable than NHS, which can serve as substrate for microorganisms [20]. NHS consists primarily of identifiable organic compounds such as amino acids, nucleic acids, fatty acids, carbohydrates, hydro carbons and phenolic compounds [21],but also include larger structures such as polysaccharides and peptides/proteins. HS are large complex compounds which vary widely in charge (highly charged to uncharged), molecular size and molecular weight (few hundred to million Da), structure and combinations of various functional groups (e.g. phenolic, carboxylic, amino, esteric, nitroso and quinine) which are generally negatively charged at neutral pH [3]. The HS fraction of aquatic NOM is relatively hydrophobic, while the NHS fraction is seen as the hydrophilic fraction.

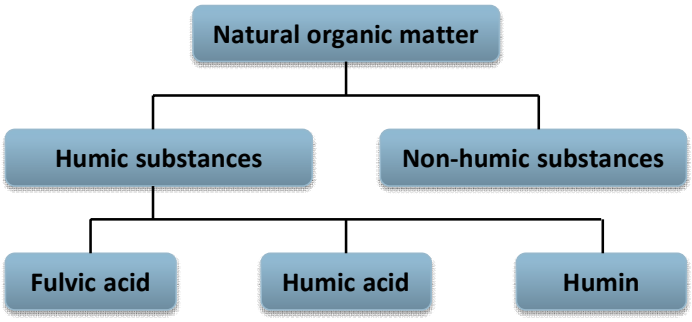


Figure 1-1: Schematic overview of natural organic matter fractions

Humic substances

Humic substances are often classified into the following three fractions based on the solubility under acidic or alkaline conditions:

- Humic acid (HA): Is the fraction which is soluble in in alkali, but is insoluble when the pH decreases below 2.
- Fulvic acid (FA): is the fraction which is soluble in water at any pH.
- Humin: is the fraction that is not soluble at any pH value.

FA is the major fraction of HS in dissolved organic matter, which is generally distributed in 3:1 ratio between FA and HA [22]. The common properties of HA and FA are presented in Table 1-1.

TABLE 1-1 COMMON PROPERTIES OF HUMIC ACID AND FULVIC ACID

Property	Humic acid	Fulvic acid
Elemental composition (% by weight)		
Carbon	50-60	40-50
Hydrogen	4-6	4-6
Oxygen	30-35	44-50
Nitrogen	2-4	<1-3
Sulfur	1-2	0-2
Solubility in strong acid	Not soluble	Soluble
Apparent molecular weight range (Da)	Few 100 to several million	180-10,000
Functional group distribution (% of oxygen is indicated in functional groups)		
Carbonyl (-COOH)	14-45	58-65
Phenol (-Ph)	10-38	9-19
Alcohol (-R-OH)	13-15	11-16
Carbonyl (-C=H)	4-23	4-11
Methoxyl (-O-CH_3)	1-5	1-2

Adapted from Snoeyink and Jenking [23], Xing [24], Sillanpää [18].

Aquatic FA has a molecular weight range of 180 – 10,000 Da, whereas aquatic HA has a molecular weight range of ~200 – to millions Da. Aquatic HAs are larger than FAs and are more colloidal in nature. FA has more oxygen containing functional groups (i.e. carboxyl, phenol, alcohol, carbonyl and methoxy) and less carbon (percentage wise) than HA, whereas HA has more aromaticity and is therefore more hydrophobic than FA. Due to the larger percentage of polar functional groups FA is more soluble in water. The hypothetical molecular structures of HA and FA are presented in Figure 1-2 and 1-3, respectively.

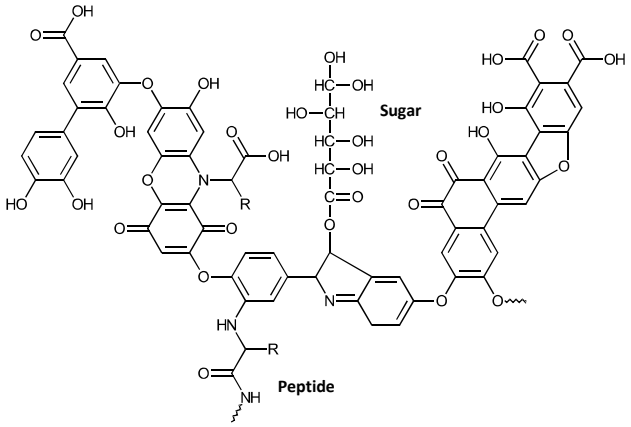


Figure 1-2: Hypothetical structure of humic acid. Adapted from Stevenson [25].

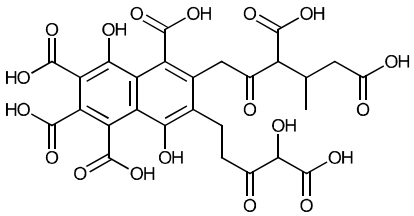


Figure 1-3: Hypothetical structure of fulvic acid. Adapted from Buffle [26].

1.3 Dissolved black carbon

Black carbon is the product of incomplete combustion of organic carbon which abounds in the environment [27]. Black carbon is formed by the incomplete combustion of organic matter from natural sources (e.g. wildfire and volcanic eruptions) or from anthropogenic sources (e.g. fossil fuel combustion and biomass burning). These sources produce black carbon residues in and on soils as well as airborne combustion products. Over time black carbon degrades in soil by microbial activity and oxygen atoms can be incorporated into the condensed aromatic structures of charcoal [28]. This causes the black carbon to dissolve partially in water and migrate through the soil as dissolved black carbon. It is suspected that DOM could be an important transport medium of dissolved black carbon, although limited data on this subject is available [27]. Dissolved black carbon was found in substantial amounts in DOM pools of different sources (oceans, rivers and estuaries) [29] and a few characterisation studies were already performed using FTICR-MS [30, 31]. Currently, little is known about the presence and behaviour of dissolved black carbon in aquatic environments. However, the characterisation of dissolved black carbon falls beyond the scope of this literature study.

2 Isolation and concentration methods

2.1 Isolation and concentration

For the characterisation of aquatic NOM, isolation and concentration of NOM are usually necessary in order to achieve sufficient enrichment for analysis and/or to separate NOM into molecular groups having the same chemical or physical properties. Many analytical methods for isolation and concentration of NOM are available, of which the most widely used are reverse osmosis (RO), freeze-drying, and pressure reducing evaporation [6, 32].

2.1.1 Reverse osmosis

RO is a membrane process which is used in industry (e.g. water filtration and desalting) that leads to the nonspecific isolation of NOM and is used in many NOM isolation studies [33-36]. RO isolation is widely employed because it has been shown that it is superior to other isolation techniques due to relatively small alteration effect on the physical and chemical properties of NOM [36, 37]. The disadvantage of RO isolation is that it also results in concentration of inorganic ions (e.g. salts) which may lead to membrane fouling and alteration of NOM characteristics. Therefore RO isolation is often used in combination with electrodialysis, cation exchange membranes, nano filtration and XAD resins [38-41].

2.1.2 Freeze-drying

Freeze-drying is often used for the preparation of NOM samples in order to concentrate the sample and keeping the sample stable for characterisation purposes [42]. The major advantage of freeze-drying is that it has a low operating temperature which leads to minimal damage of heat sensitive NOM components, such as peptide/proteins, amino sugars and polysaccharides. A limitation of freeze-drying is that it is time-consuming and can require additionally desalting when high salt concentrations are present in the water. Furthermore, freeze-drying is a process which is different for each analyte, resulting in that each class of compounds has its own optimal conditions, which can lead to inconsistent dryness in the sample, reduced stability or rehydration [43].

2.1.3 Pressure reduced evaporation

Pressure reduced evaporation (vacuum distillation) is a technique in which evaporation of solvents is accelerated by working under reduced pressure with or without additional heating.

2.1.4 Electrodialysis

Electrodialyisis is used as an NOM purification step by removing the salts which are present in water by cation and anion exchange membranes. For this reason electrodialysis is often coupled with other isolation or fractionation techniques [33, 38].

2.1.5 Passive sampler

In 2006, passive sampling was first used by Lam and Simpson [44] as an alternative extraction method for the isolation of DOM in freshwater. Passive sampling is a sampling technique based on the transfer of analyte molecules from the sample medium (i.e. natural water) to a collecting medium (solvent or porous sorbent), which is caused by the difference in chemical potentials of the analyte between the sample medium and collecting medium. The transfer of analyte molecules continues between the two media until equilibrium is reached in the system or the sampling is stopped. Lam and Simpson [44] developed a passive sampler (Figure 2-1) which is made of a porous HDPE casing to house the sampler

unit which consist of polyvinylidene fluoride (PVF) membrane and DEAE-cellulose resin. The PVF membrane ensures that only analytes with a molecular weight lower than 1000 kDa can enter the membrane, the DEAE-cellulose anion exchange resin concentrates the negatively charged analytes.

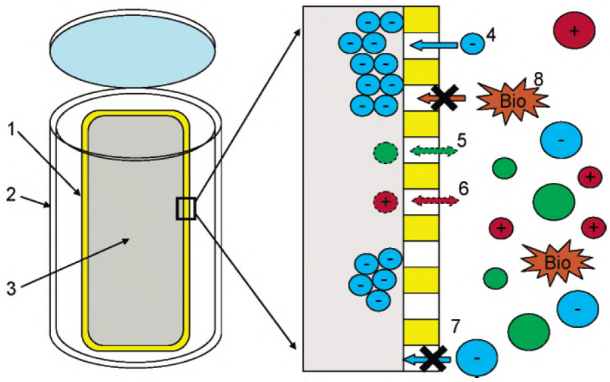


Figure 2-1: Schematic showing of the components of the passive sampler. (1) The 1000 kDa polyvinylidene fluoride membrane. (2) Porous HDPE casing. (3) DEAE-cellulose resin. (4) Negatively charged DOM enters the membrane and is retained and concentrated. (5+6) Dissolved neutral or positively charge analytes can enter the membrane but are not retained. (7+8) Large species cannot enter the membrane. Adapted from Lam and Simpson [44].

Passive sampling has some clear advantages over UF and SPE methods, such as on site sampling, less contamination due to sampling, sample handling and storage. Furthermore passive sampling provides an integrative DOM sample over a defined time period in comparison with regular sampling (i.e. grab sample) which provides a snapshot of DOM the composition of a single moment. A practical disadvantage is that the sampling time is longer (e.g. 1–4 weeks). However, it was reported by Warner et al. [45] that there is a statistically significant difference in NOM quality between grab samples and passive sampling extracts. The extracts obtained with passive samplers indicated a higher abundance of high molecular weight terrestrial DOM than there is actually present in the sampled stream water. Therefore caution is warranted, when passive samplers are used for the characterisation of low molecular weight DOM. In summary, passive samplers are a promising approach for DOM isolation, providing a cost effective, time integrated DOM extract for analysis.

2.2 Fractionation methods

NOM is a complex heterogeneous mixture of organic constituents, which are highly diverse in their structure, but they also share many common properties and can therefore be separated or fractionated into groups by their common properties. One of major challenges in characterising NOM, is to develop fractionation procedures which provide the pure compounds that could subsequently be characterised by standard techniques. The most widely used NOM fractionation methods are resin fractionation, SEC, membrane filtration, PRAM, reversed phase high performance liquid chromatography (HPLC) and field flow fractionation [2, 37, 46-49].

2.2.1 Resin fractionation

Resin fractionation is one of the most widely used techniques to isolate and fractionate aquatic NOM into different subgroups of compounds with similar properties. The classification of NOM into humic substances (hydrophobic) and non-humic substances (hydrophilic) is based on what is or what is not adsorbed onto XAD resins [50]. This classification is not universally accepted, since some of the hydrophilic (non-humic) fractions

of NOM exhibit the same properties as observed for "regular" humic fractions [51]. Currently, there are two well-known resin fractionation methods for NOM developed by Leenheer [52] and Aiken et al. [53]. The method of Aiken et al. involves XAD-4 and XAD-8 resins which fractionate NOM into three fractions, a hydrophobic (XAD-8), a transphilic (XAD-4) and a hydrophilic fraction (which is not retained on both resins). The analytical procedure developed by Leenheer [52] provides more information by separating NOM into six fractions based on their size and polarity. Therefore, three columns with resins, XAD-8, AG-MP-50 (cation) and DUOLITE A-7 (anion) were separately used. The XAD-8 column yields three fractions; hydrophobic base (HPOB) which is eluted with hydrochloric acid, hydrophobic acid (HPOA) eluted with sodium hydroxide and hydrophobic neutral (HPON) eluted with methanol. The second column containing the cation exchange resin (AG-MP-50) yields the hydrophilic base (HPIB) fraction. The third column containing the anionic exchange resin yields the hydrophilic acid (HPIA) fraction. Finally, the fraction which is not retained by the three resins is defined as the hydrophilic neutral fraction (HPIN). Figure 2-2 shows a NOM classification method, which classifies NOM based on size, polarity, acid/neutral/base properties and compound classes.

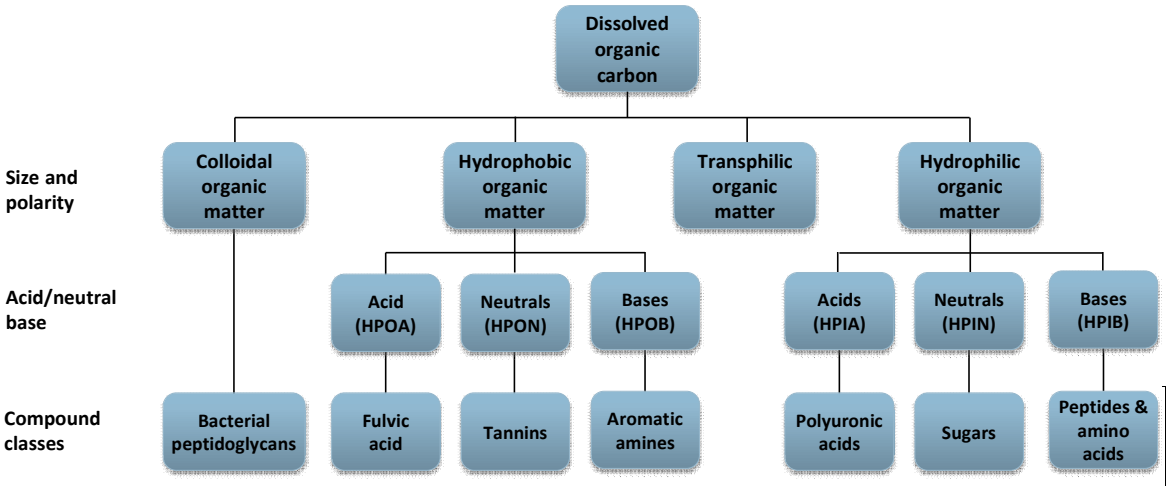


Figure 2-2: NOM fractionation diagram. Adapted from Leenheer and Croué [54].

Many researchers have adopted and modified the resin fractionation methods of Aiken et al. [53] and Leenheer [52] for their own research [55-57], and resin fractionation is still used in NOM characterisation studies today [58-60]. Resin fractionation was also adopted by the International Humic Substance Society (IHSS) as a standard method for the isolation of aquatic humic and fulvic acid [61]. Although resin fractionation techniques are widely used and accepted, they are time-consuming and require at least 24 hours to produce six fractions when 1 L of sample is used [52, 57]. Furthermore, possible chemical or physical alteration of NOM could occur, due to extreme pH levels and pH changes during fractionation, contamination through resin bleeding, size-exclusion effects and irreversible adsorption of NOM compounds to the applied resin [6].

2.2.2 Solid phase extraction

Pre-packed hydrophobic solid phase extraction (SPE) sorbent has only recently been studied for NOM isolation and fractionation [62-65]. Before the year 2000, XAD resin were the sorbents of choice for NOM extraction, but they are steadily replaced by C18 and polystyrene divinylbenzene (PS-DVB) based pre-packed SPE sorbents. Ditmar et al. [64] reported in a comparison study between different SPE sorbents that PPL PS-DVB based SPE sorbent showed the highest recovery, followed by slightly lower recoveries for C18 functionalised silica

sorbents. Furthermore it was shown by using NMR and HR-MS based characterisation that both sorbents generally extract the same analogue classes of compounds [64, 66]. The best performing PS-DVB sorbent is PPL bond ELUT which has a high surface area (600 m^2 g^{-1}) and also offers a substantial recovery of both polar and non-polar compounds, high recoveries for small compounds (< 3 kDa) and also provides enhanced selectivity for the whole bulk of NOM. PPL based PS-DVB sorbents are able to recover up to 62% of the NOM in river water [64], followed by 38-63% for C18 based sorbents [65, 67] and circa 30% for XAD-8 resin [68]. For the samples obtained of these relatively new PS-DVB sorbents it is regarded that they are acceptably representative of the true NOM composition [69]. Because of the aforementioned advantages PS-DVB resins have gained widespread acceptance [43].

There is still some discussion in literature about the use of disks or cartridges for SPE extraction of NOM. Kim et al. [66] reasons that SPE disks can be used at higher flow rates than cartridges and still yield a high recovery of ~60% using C18 sorbent instead of PS-DVB. However, Kim determined the recovery by UV-Vis instead of the conventional DOC analysis, which makes a direct comparison with other studies difficult.

The major advantage of using SPE for the isolation and extraction of NOM is that the extraction procedure is simple and fast. The disadvantage of SPE extraction is that most current SPE methods in literature can only isolate a hydrophobic and a hydrophilic fraction [70]. In order to obtain more fractions Ratpukdi et al. [70] developed a novel rapid fractionation technique that can separate NOM into six fractions in approximately 6 hours which is faster than traditional fractionation techniques. For this extraction procedure three Bond Elute ENV cartridges, one Strata-X-C cartridge and one Strata-X-AW cartridge were used, yielding the following six fractions; HPON, HPOB, HPOA, HPIB, HPIA and HPIN. An overview of the fractionation setup is shown in Figure 2-3. Satisfactory recoveries were obtained for Suwannee river NOM, with recoveries ranging from 63.5 ± 7.7% to 101.7 ± 11.2% for the different fractions, with the exception of HPOB fraction which yielded a substantially lower recovery (2.3%) A possible explanation for this low recovery is that the total amount of carbon in that fraction is low (1%) and that the compounds in that fraction have a strong affinity for the ENV sorbent. The recoveries for Red river NOM was for all fractions above 78.1%. This method is a good alternative for the method developed by Croue et al. [33] (fractionation by XAD resins), due to less time consumption (6 hours vs. 24 hours) and sample amount (1 L vs. > 100 L), while also yielding similar recoveries.

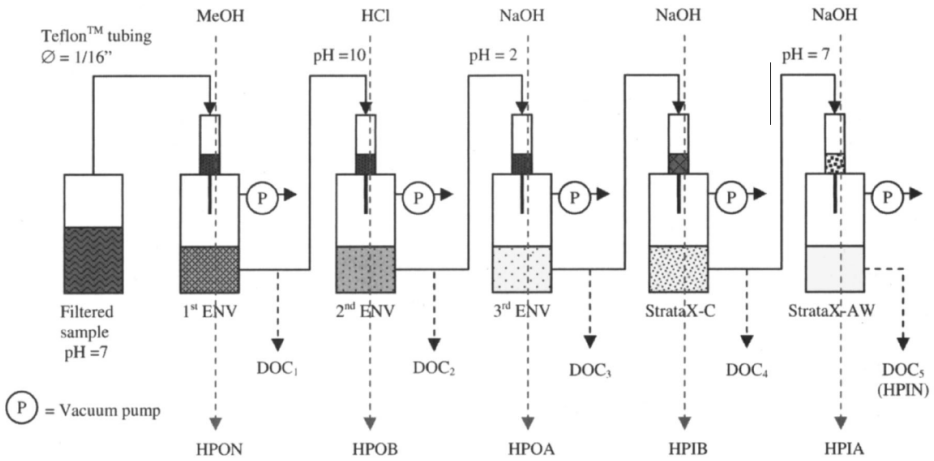


Figure 2-3: SPE fractionation setup and procedure for obtaining six fractions. Source Ratpukdi et al. [70].

2.2.3 Membrane filtration

Membrane filtration techniques are widely used in studies related to water research and water treatment and mainly include ultrafiltration (UF), microfiltration (MF) and nanofiltration (NF). In 1970 it was already demonstrated that aquatic NOM could be fractionated by using membranes with different molecular weight cut-off ranges [71]. Ultrafiltration is widely adopted since it was used by Benner et al. [72] for DOM extraction. The main difference between ultrafiltration and RO is the membrane permeability. With RO all compounds and ions (including salts) are retained by the membrane, while ultrafiltration only retains compounds above a certain molecular weight cut-off (e.g. 0.5, 1, 3, 10, 30 and 100 kDa). For NOM extraction/fractionation a typical molecular weight cut of 1kDa is used, meaning that small compounds and salts are passed through the membrane. The DOM recovery for ultrafiltration is often higher than that of SPE, and has the advantage that it does not require any chemical manipulations, which possibly can alter the chemical composition of DOM (e.g. SPE is acidified to pH 2 and RO rinsing is performed at pH 12). Another advantage of ultrafiltration is that it typically uses high flow rates, which makes it possible to extract large sample volumes relatively fast. One of the disadvantages of ultrafiltration is membrane fouling which leads to a reduction in productivity [73]. Another disadvantage is that significant differences were observed in organic matter recovery when different ultrafiltration systems (i.e. different manufacturers) or operation conditions were used [74]. Furthermore, it is questionable if the DOM that is extracted by ultrafiltration is representive of the DOM present in the original sample. Benner et al. [72] showed that the C/N ratios obtained from ultrafiltration extracts were similar to those of the initial water samples. However, it was shown that when the mass balance was determined by DOC and UV-Vis the recoveries were very different in comparison with the initial water samples, which could mean there is a bias in compounds retained by ultrafiltration [65]. It is important to consider that DOM/NOM extracts obtained from SPE are different from the extracts recovered by ultrafiltration [75].

2.2.4 Polarity rapid assessment method

Polarity rapid assessment (PRAM) is a novel polarity characterisation method for NOM developed by Rosario-Ortiz et al. [48], in which the polarity is characterised by quantifying the amount of material adsorbed onto different SPE sorbents. The analytical procedure is relatively simple; circa 200 mL of water sample is loaded onto SPE cartridges with different types of sorbent in parallel (e.g. C18, C8, C2, CN, silica and diol), for which the retention coefficient is calculated by measuring the initial concentration and the maximum breakthrough concentration by UV_{254} or TOC analysis (see Figure 2-4). In this way the polarity of NOM can be characterised multidimensionally.

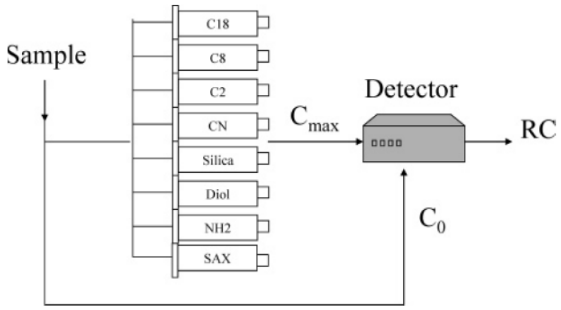


Figure 2-4: Experimental setup for PRAM (source: Rosario-Ortiz et al. [48]) The retention coefficient (RC) is calculated by means of the maximum breakthrough concentration an initial concentration. Other combinations of SPE cartridges can also be used in parallel.

An advantage of PRAM is that the analytical procedure is performed under ambient conditions with limited pre-treatment steps, resulting in maintaining the chemical

characteristics of NOM and yielding accurate representation of the polarity of NOM. Since the development of PRAM, the method has been increasingly applied for NOM polarity characterisation, but was also adopted and used for characterisation of nitrosoamines precursors [76, 77].

2.2.5 Size exclusion chromatography

Size exclusion chromatography (SEC) is an analytical technique which separates or fractionates compounds based on their size and sometimes molecular weight [78]. The analytes are injected onto an analytical column which contains a stationary phase made of a porous gel, in which small molecules can enter more pores than larger molecules, and therefore elute slower than the larger molecules. SEC was first applied for NOM fractionation in 1963 by Posner et al. [79], although the applications were hindered by the poor resolution of the soft gels that were used. This was later improved by the development of high pressure size exclusion chromatography (HPSEC) [80], which became quite popular for the fractionation of NOM, and was coupled to various detection techniques such as UV and photodiode array detection, organic carbon detection (OCD), fluorescence detection, and FTIR [81-84]. HPSEC hyphenated techniques are discussed further in detail in section 4.2.1. A limitation of SEC is that polystyrene sulfonate calibration standards are used for molecular weight determination, which are different compared to complex NOM structures and therefore the results obtained are approximations.

2.2.6 Other fractionation methods

Reversed phase high performance liquid chromatography (HPLC) is occasionally used as fractionation technique for NOM [85-87] especially in comparison with HPSEC. Reversed phase HPLC employs a non-polar stationary phase (e.g. C18 column) and polar mobile phase in which the NOM compounds are eluted from the column by polarity. It is also used as a separation technique in combination with different types of detectors (UV, Flu, MS, PDA) for the characterisation of NOM. HPLC hyphenated techniques are discussed in chapter 4.

Another liquid chromatography technique which occasionally is used for separation or fractionation of NOM is hydrophilic interaction chromatography (HILIC). HILIC is used in several studies as a fractionation method prior to NMR analysis [88, 89]. In section 4.2.3 HILIC chromatography and applications are discussed more in-depth.

Field-flow fractionation (FFF) is a chromatography-like technique which can be used for analysing molecular size distribution according to their diffusion coefficient. FFF provides a molecular size distribution of macromolecules which can be detected (online or off-line) by various types of detectors (UV, PDA, fluorescence). FFF is used as a fractionation technique for colloidal NOM but is also used for characterisation of NOM (i.e. molecular size distribution). FFF applications for the characterisation of NOM are discussed in section 4.4.

3 General analysis methods and in-depth characterisation methods for NOM

3.1 Analytical methods for determining general parameters of NOM

NOM characterisation can be performed by the determination of general parameters or by characterisation of elemental composition or structural identification. In many cases NOM is represented by general parameters such as DOC, dissolved organic nitrogen, UV-Vis absorbance, fluorescence, and SUVA analysis. The analysis of these parameters is relatively fast and does not require complex sample pre-treatment and instrumentation. In the following sections (3.11 – 3.15) the analytical methods for these general parameters are described in further detail. Analytical methods for the determination of elemental composition or structural properties of NOM are discussed in 3.2.

3.1.1 Total organic carbon and dissolved organic carbon

TOC and DOC are probably the most frequently used parameters to indicate the amount of carbon in a NOM sample. TOC is the sum of particulate organic carbon (POC) and DOC present in a sample, minus the inorganic carbon (i.e. carbon dioxide and carbonic acid salts). TOC is determined by a so called TOC analyser. DOC sometimes referred as DOM, is the amount of carbon dissolved within an aquatic system (see Figure 3-1). The widely accepted definition of DOC is the amount of organic carbon which can pass a 0.45 µm filter [90]. DOC analysis is performed by 0.45 µm filtration of the sample and subsequent analysis using a TOC analyser.

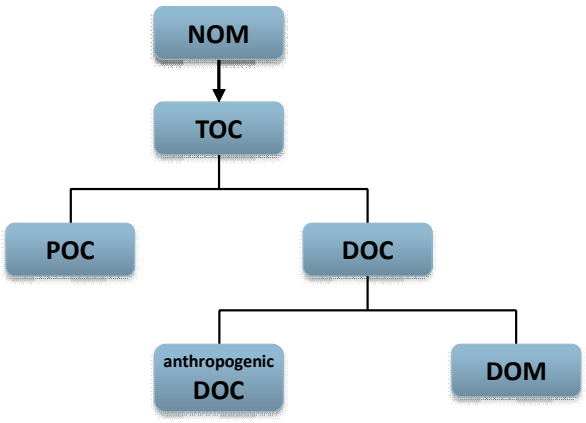


Figure 3-1: schematic overview of the different organic carbon/matter fractions in NOM.

3.1.2 Dissolved organic nitrogen

Dissolved organic nitrogen (DON) is ubiquitous in aquatic systems and is one of the most relevant parameters for drinking water treatment and thus NOM characterisation. DON is of growing interest due to the linkage with the formation of nitrogenous DBPs [91, 92] and membrane fouling. As with the determination of DOC, DON is the fraction which can pass a 0.45 µm filter. The concentration of DON can be calculated by determining the amount of total dissolved nitrogen (TDN) followed by subtracting the inorganic nitrogen fraction (i.e. nitrate, nitrite and ammonia). However, accurate DON calculation is rather difficult due to

analytical errors and uncertainty of three analyses (TDN, ammonia and nitrate/nitrite) which leads to high standard deviations and sometimes even to negative DON concentrations [93].

3.1.3 UV-Vis

Ultraviolet and visible (UV-Vis) absorption spectroscopy is one of the most popular techniques in analytical sciences. In this technique the concentration of a compound in a sample is determined by the absorption of ultraviolet or visible light. UV-Vis is based on the Beer-Lambert law, which states that the absorbance of a solution is directly proportional to the concentration of the absorbing analytes in the solution and the path length. It was shown that the UV absorbance spectra of NOM decreased monotonically with increasing wavelengths [94]. For NOM measurement any wavelength from 220 to 280 nm is applicable, but this depends on absorbance wavelength of the chromophores present in the NOM structure. Absorbance in the spectral region around 254 nm is determined by π - π* electron transitions, which is typically for substituted benzenes, aromatic rings or polyphenols and therefore UV_{254} is used as a surrogate measurement for the aromaticity of NOM. In general the UV-Vis absorption spectrum of NOM appears relatively featureless with no discrete peaks available, which makes NOM characterisation rather difficult. Therefore ratios of absorbance between specific wavelengths (e.g. A_{254}/A_{204}, A_{250}/A_{365} and A_{254}/A_{436}) are also used in order to provide more insight about the NOM composition. An example is the ratio between 250 and 365 nm, which has shown to be inversely related to the molecular weight and inversely proportional with the amount of aromatic material present in a sample [95]. The drawback of UV-Vis is its limited applicability for NOM characterisation and providing information about chromophoric NOM only, rather than about the total NOM pool.

3.1.4 Specific UV-absorbance

SUVA is defined as the absorbance of UV light in a water sample at a specified wavelength which is divided by the DOC concentration [96]. The wavelength used for this parameter can differ, but is usually 254 nm. $SUVA_{254}$ is an indicator of the aromaticity of a water sample, whereby a SUVA value > 4 indicates the presence of hydrophobic and mainly aromatic NOM, while a SUVA value < 3 indicates the presence of hydrophilic NOM [97]. SUVA correlates well with the specific UV ratio of $A_{254}/A_{204,}$ because both parameters also correlate with amount of carbon in NOM, but $SUVA_{254}$ is better suited to discriminate between different DOM compositions [94].

3.1.5 Fluorescence

Fluorescence is a spectrochemical method in which the analytes are excited by irradiation at specific wavelength and subsequently emit light at a different wavelength. Fluorescence spectroscopy has been frequently applied for the characterisation of NOM, because of its high sensitivity and good selectivity in comparison with UV spectroscopy [98] and because it is a relatively inexpensive tool that requires little or no sample pre-treatment. Fluorescence does not occur for every compound. Mainly compounds containing multiple aromatic rings or planar and cyclic rings containing double bonds, show fluorescence. The excitation and emission wavelengths are quite specific for a particular molecular conformation, which can be useful for characterising NOM [99, 100]. Excitation emission matrix 3D fluorescence spectroscopy (EEM) is a technique which is developed in the mid-1990s [101], whereby the excitation and fluorescence intensity can be scanned simultaneously in order to generate a map of optical space. See Figure 3-2 for example of a typical EEM of IHSS Suwannee River humic and fulvic acid. EEM spectra visualise a range of excitation and emission wavelengths (~200 nm to ~500 nm) and are therefore more informative in comparison with the traditional single scan technique.

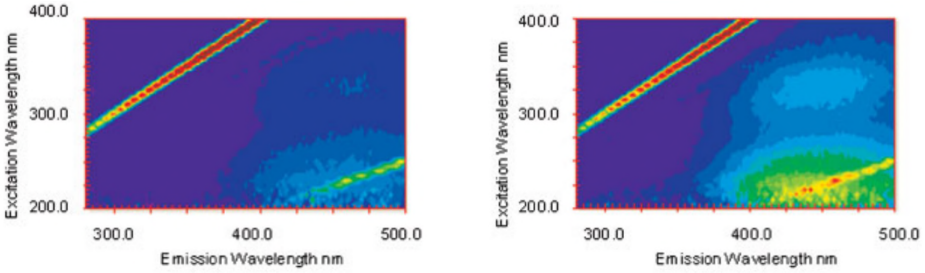


Figure 3-2: EEM of IHSS Suwannee River (left) humic and (right) fulvic acid. Source Hudson et al. [102].

EEM spectra of typical raw water contain two main fluorescence peaks referred to as protein and humic-like fluorophores [103] or referred to as fulvic, humic, and protein-like fluorophores [102, 104]. Other methods for NOM characterisation using EEM spectra are: fluorescence regional integration (FRI) [105], multivariate data analysis (e.g. partial least squares regression and principal component analysis) [106] and parallel factor analysis (PARAFAC) [107]. PARAFAC deconvolutes complex EEM spectra into individual components, some of which have been attributed to humic or protein-like NOM [107].

3.2 Analytical methods for elemental and structural identification of NOM

The general parameters discussed above provide some information on aquatic NOM, however are limited in determining the elemental composition and structural identification of NOM. In the last two decades, significant advances in instrumental techniques such as Fourier transform infrared spectroscopy (FTIR), nuclear magnetic resonances spectroscopy (NMR) and Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) have led to more information about NOM composition.

3.2.1 Elemental analysis

Determination of the elemental composition of NOM can be performed qualitatively by determining which elements are present or quantitatively by determining the exact amount of each element. The elemental composition of NOM is usually determined by an elemental analyser [108]. With an elemental analyser mainly the elements CHNS are determined by combustion analysis, in which a sample is burned in abundance of oxygen and then the different fractions are trapped on selective trap columns collecting the combustion products such as N_2, CO_2, H_2O and SO_2. The amount of these combustion products is used to calculate the composition of NOM. The determination of oxygen is conducted separately in an elemental analyser by the catalytic conversion of bound oxygen to CO. The elemental composition is useful for determining H/C, O/C or N/C ratios, which can be used for NOM characterisation [108]. For example the H/C ratio is a measure of the amount of saturation or aromaticity of NOM, in which a relatively high value means more saturation and less aromaticity. Also the polarity of NOM can be calculated by calculation the (O+N)/C ratio, whereby a low ratio means a high hydrophobicity.

3.2.2 Fourier transform infrared spectroscopy

Fourier transform infrared spectroscopy (FTIR) has been widely used for years for NOM characterisation [109-111]. The basic principle of infrared spectroscopy, is that infrared light is absorbed by an analyte at a specific wavelengths due to molecular vibrations such as stretching, bending and scissoring. Only vibrational modes in which the dipole moment changes in a molecule can be detected by infrared spectroscopy. Therefore specific functional groups absorb at specific wavelengths, which yield an absorption spectrum that can be regarded as a unique fingerprint of a molecule or sample. The interpretation of such an absorption spectrum can be quite difficult for a NOM sample, due to the overlapping of

spectral features, which is caused by the complexity of NOM. In Table 3-1 an overview is shown of the main FTIR bands detected in humic substances.

TABLE 3-1 MAIN BANDS OF FTIR SPECTRA OF HUMIC SUBSTANCES

Band (cm^{-1})	Assignment
3400	Associated O–H stretch (alcohols, phenols and carboxylic groups)
2850-2960	C–H stretch (CH_3 and CH_2)
2620	O–H stretch (hydrogen-bonded carboxylic groups)
1720	C=O stretch (carboxylic groups)
1630	C=C stretch (alkenes and aromatic rings)
1540	N–H bend (N–H structures)
1455	C–H bend (CH_3 and CH_2)
1410	O–H bend (carboxylic groups)
1375	C–H bend (CH_3)
1260 and 1220	C–O stretch (carboxylic groups, phenols, aromatic/unsaturated ethers)
1095 and 1030	C–O stretch (alcohols, aliphatic ethers)
805	C–H bend (tri- and tetra-substituted aromatic rings)

Adapted from Rodriques and Núñez [112]

In order to obtain good quality spectra for NOM characterisation, sample pre-treatment is often required. IR spectra of solids or liquids are usually recorded by transmitting infrared radiation directly through the sample. Therefore it is important to use an appropriate intermediate for the sample material and infrared beam. For NOM characterisation, the typical approach is to make a pellet of dried sample material and the optically transparent KBr. With this approach good resolution and transmittance is obtained. The disadvantage of using KBr for sample preparation is that it can be considered as sample destructive, because the sample is thoroughly mixed with KBr. Furthermore KBr is hygroscopic and precautions have to be made in order to keep the KBr sample mixture water free, because water yields a strong FTIR signal and causes substantial interference in a FTIR spectrum. In general a NOM FTIR spectrum looks relatively simple in comparison with a spectrum of a pure compound, due to the overlapping spectral bands caused by the heterogeneity of NOM. An approach for enhancing the resolution of an FTIR spectrum is by calculating a second derivative spectrum [113]. In a complex sample when multiple bands are overlapping a second derivative spectrum shows the original bands by examining the obtained valleys (see Figure 3-3).

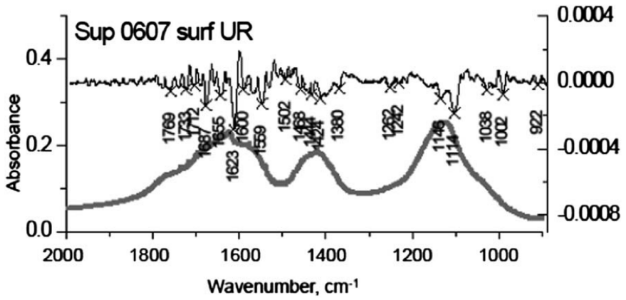


Figure 3-3: FTIR spectrum of DOM from Lake superior surface water and second derivative spectrum. The valleys (i.e. downwards spikes) show the original bands that overlap in the original spectrum. Source Minor et al. [114].

Attenuated total reflectance (ATR) is another infrared spectroscopy approach in which dried NOM samples are pressed to thin films against an ATR crystal. ATR spectroscopy uses a beam of IR light which penetrates through the crystal and creates an evanescent wave in the sample by reflecting at least once on the internal surface which is in contact with the sample. In order to form this evanescent wave the crystal should have a higher refractive index than the sample. When the sample interacts with this evanescent wave less light will be reflected

back to the detector. Since the depth of penetration of the evanescent wave into a sample is wavelength dependent, it is expected that ATR spectra show lower responses at higher wavelengths. One of the major advantages of ATR IR spectroscopy is that it is a non-destructive technique; it is relatively simple to recover the sample from the surface of the crystal. Another major advantage is that it minimizes sample preparation and thus the sample to sample reproducibility is improved. Due to these advantages ATR is becoming increasingly popular for NOM characterisation [110, 115, 116].

3.2.3 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) is one of the most powerful analytical techniques for characterising NOM [117]. It is extensively applied for studying NOM [88, 117-120], and it is often used together with FT-ICR-MS [121, 122] (see 3.2.4). Almost all elements are in principal observable using NMR, which means that NMR can be applied for almost every inorganic and organic matrix. However, NMR is mainly used for organic structures, because the organic nuclei (^{1}H, ^{13}C, ^{15}N, and ^{31}P) produce relatively narrow NMR signals. The basic principle of NMR works as follows; when a molecule is situated inside a magnetic field and a specific radiation frequency is applied, the nuclei of the atoms may experience a transition between spin states in which a nucleus flips from one orientation to another. The continuous transition between spin states is called resonance and is dependent of the applied frequency, magnetic field strength, type of nucleus and chemical environment of the atom. When a nuclei goes back to its original state, it emits a electromagnetic signal which is detected by the instrument. The specific frequency at which a nucleus resonates is known as the "chemical shift" which gives information about the different types of chemical groups present in a molecule and their bonding environment. When using one-dimensional (1D) NMR for obtaining a NMR spectrum, the types of nuclei are identified by chemical shift and the number of each type is determined by the peak intensity. With multidimensional NMR it is possible to determine how the nuclei are connected in order the resolve the structure. The theoretical resolution of NMR is substantial, whereby ^{1}H and ^{13}C NMR have a peak capacity of 5000 and 30,000, respectively. However, the peak capacity for multidimensional NMR increases considerably to, 2D ^{1}H-^{13}C NMR > 1,000,000 and for 3D ^{1}H-^{13}C-^{15}N NMR > 100,000,000 [117]. This shows that NMR has considerable potential to resolve overlapping signals in a complex heterogeneous matrix such as NOM. Although the peak capacity using NMR is considerable, the heterogeneity and complexity of NOM limits practical application of NMR data.

Figure 3-4 illustrates the complexity and spectral overlap of the solution-state two dimensional ^{1}H-^{13}C NMR correlation of Pacific Ocean DOM.

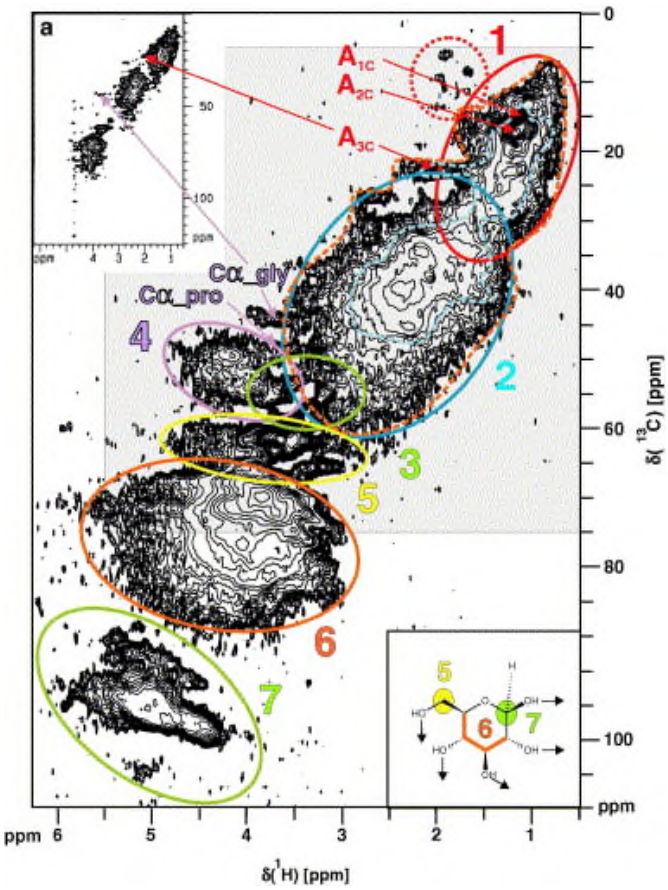


Figure 3-4: Two dimensional ^{1}H- ^{13}C NMR spectrum of Pacific Ocean DOM. Assignment of the major constituents is as follows: 1 = methyl bound to carbon and sulphur (dotted circle), 2 = methylene and methane cross-peaks without direct bonds to hetero atoms, 3 = low intensity cross-peaks from methoxyl, 4 = cross-peaks mainly of α-CH in proteins and vicinal dicarboxylic acids, 5 = carbohydrate methylene cross-peaks, 6 = carbohydrate methine cross-peaks, and 7 = anomeric units in carbohydrates, A_{1C} = (poly)alanine CH_3, A_{2C} = methylated carbohydrates, A_{3C} = N-acetyl carbohydrate. Source Hertkorn et al [123]. and Simpson et al [117].

NMR analysis of NOM can be performed using both solution and solid-state NMR. However, solid state NMR for ^{13}C NMR experiments is preferred over solution-state due to some advantages such as the possibility to use spectral editing of strongly overlapping bands and because of the higher concentrations in solid samples. Mostly ^{1}H and ^{13}C NMR techniques are used for NOM characterisation. However, Thorn and Cox [124] have showed that peptides and other nitrogen containing compounds can be identified with ^{15}N NMR.

3.2.3.1 Solution-state NMR and solid-state NMR

Currently, solution-state NMR delivers the highest resolution and is more comprehensive than solid-state NMR for soluble components. Solution-state NMR experiments are therefore preferred for naturally soluble samples such as NOM. In order to improve the applicability of NMR it is often advantageous to isolate components/fractions of NOM samples (e.g. fulvic acid, humic acid) in order to decrease the heterogeneity of the samples, for the removal of inorganic ions, and for the removal of paramagnetics, which improves both sensitivity and resolution. Fractionation techniques such as SEC [125], HILIC [88], and resin fractionation yielded satisfactory results. Furthermore, it is preferred to isolate and concentrate NOM samples in order to improve sensitivity. However, Lam and Simpson [126] showed that solution-state NMR can be performed directly on natural water, without any sample

concentration. The major advantage of this approach is that the sample composition is not modified due to isolation and concentration steps. On the other hand, the drawbacks of this approach are: long analysis time (i.e. >24 h) due to the low concentration, only 1D NMR can routinely be performed and challenging water suppression approaches are needed for obtaining undisturbed NMR spectra.

Solid-state NMR is traditionally performed on dried samples and is most used for ^{13}C NMR in environment samples rather than for ^{1}H NMR [117], since the ^{1}H – ^{1}H dipole interactions are very strong with solid-state NMR, which result in a wide spectral profile and limit chemical shift information. For environmental samples, solid-state ^{13}C NMR provides a great overview in the distribution of carbon within a sample and for which only limited sample pre-treatment is needed. Since NOM is a heterogeneous mixture, the chemical shift regions within in a NMR spectrum can be quite large. The ^{13}C NMR chemical shift range can be divided into the following regions: aliphatic, unsaturated, alkoxyl, and carbonyl region [127]. In Table 3-2 an overview is shown of the general assignment of structural features of NOM and their respective shift regions obtained using solid-state ^{13}C NMR. Some advantages of solid-state NMR over liquid NMR are, less interference of heterogeneous sample dissolution, no concentration limit, no solvent effects, and it can be used for insoluble NOM fractions [128].

TABLE 3-2: OVERVIEW OF CHEMICAL SHIFT REGION AND ASSIGNMENT OF STRUCTURAL FEATURES IN SOLID-STATE ^{13}C NMR FOR THE INTERPRETATION OF NOM SPECTRA

Chemical shift (ppm)	Assignment
0-50	Aliphatic or paraffinic carbon chain
50-75	Methoxyl groups
75-112	Carbohydrate RC–OH or RC–OR functional groups
112-145	Aromatic carbons
145-163	Phenolic groups
163-180	Carboxylic, carbonyl, amine, and ester carbons

Source Chen et al [127] .

NMR has shown to be a powerful discovery tool for molecular characterisation of NOM, because it provides unrivalled molecular information of the original sample. While 1D NMR provides some structural information, 2D NMR is really needed to obtain a comprehensive overview of a NOM sample. The disadvantage of 2D NMR is that interpretation of the enormous amount of data is challenging and still needs further study. Advances in the characterisation of NOM using NMR are expected to continue, but likely also require direct hyphenation of NMR, HR-MS and chromatography [117, 128]. The hyphenation of these techniques would provide chromatographic separation, while mass spectrometry would provide molecular formulas and NMR could link the nuclei to form and identify the chemical structure.

3.2.4 Mass spectrometry – FTICR-MS

Mass spectrometry is probably the most used technique for the comprehensive characterisation of NOM. In recent applications the following three mass spectrometry approaches were mainly used for the characterisation of NOM: (i) high resolution mass spectrometry (HR-MS) without prior fractionation [123, 129-133], (ii) HR-MS with (chromatographic) fractionation [134-136] and (iii) high performance liquid chromatography (HPLC) coupled on-line to HR-MS [137-139]. For all these aforementioned approaches the NOM/DOM samples were first isolated before mass spectrometry analysis (see chapter 2 for isolation methods). Currently the most used and most powerful mass spectrometry technique for comprehensive NOM characterisation is Fourier transform ion cyclotron

resonance mass spectrometry (FTICR-MS). The basic principle of FTICR-MS is as follows: Ions are generated using an external ion source, and are then injected into a panning trap which normally has a cubical or cylindrical shape. The ions are trapped by a magnetic field (super conducting magnet), where they are excited at their resonant cyclotron frequency by an electric field that Is perpendicular to the magnetic field. When the excitation field is removed, the ions rotate at their cyclotron frequency and induce a charge on a pair of electrodes and are then detected. The masses of the ions are resolved by their ion cyclotron resonance (rotational) frequency.

3.2.4.1 Mass accuracy and resolution

FTICR-MS is widely used for NOM characterisation due to its ultrahigh resolution (> 200,000) and mass accuracy (< 1 ppm) and capability for determining the elemental composition of the detected ions, without the need for chromatographic separation. With the ultrahigh resolution offered by FTICR-MS it is possible to detect several thousands of ions in one spectrum of a single NOM sample. An excellent mass accuracy is important for calculating the correct elemental composition of an unknown ion or compound. The mass accuracy shows how well the observed mass corresponds with the theoretical mass and is of importance when multiple elemental compositions could be assigned for an unknown mass or compound. Resolution is especially important for resolving closely spaced signals which is needed when a complex mixture such as NOM is measured. A typical NOM sample shows more than 20 peaks per nominal mass over a range of 300 mass units in a single spectrum [130]. The average resolution used in NOM characterisation studies is in the range of 300,000 – 750,000, although resolutions up to 20,000,000 are currently possible using FTICR-MS. See Figure 3-5 for an example of a FTICR-MS spectrum of Suwannee River NOM, recorded at a ultrahigh resolution of 500,000. The resolution achieved using FTICR mainly depends on the strength of the magnet field. A stronger magnetic field is also beneficial for the mass accuracy, dynamic range and speed of analysis. The major drawback of FTICR-MS is that generally a large number of spectra are needed to be accumulated (e.g. > 100) in order to improve mass accuracy and signal to noise ratio and to obtain excellent resolution. This long accumulation time makes it very challenging to hyphenate FTICR with a separation technique such as liquid chromatography. Therefore the online coupling between faster scanning ultrahigh resolution mass spectrometers such as the Orbitrap mass spectrometers and liquid chromatography are becoming increasingly popular [137-139]. In chapter 4, applications using a LC-Orbitrap for the characterisation of NOM are discussed.

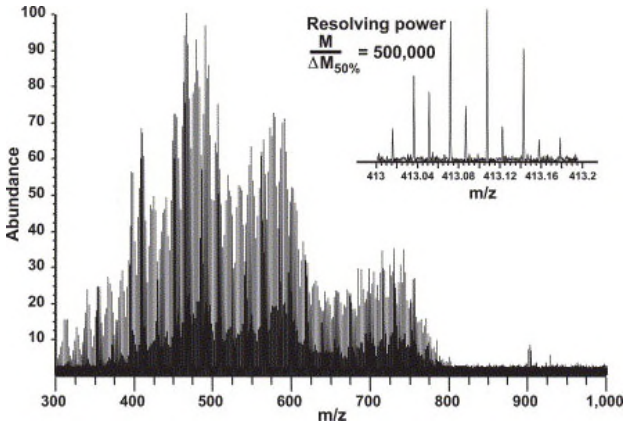


Figure 3-5: ESI FTICR mass spectrum of Suwannee river fulvic acids. The mass spectrum represents 200 co-added scans recorded using 9.4 T ESI FTICR-MS. The inset shows the expanded region near m/z 413, showing the ultrahigh resolution (500,000) of this technique. Source Kujawinski et al [140].

3.2.4.2 Ionisation techniques

There are many ionisation techniques available for performing mass spectrometry whereof electrospray ionisation (ESI) is the most widely used in general and also the most used for NOM analysis. This is because ESI ionises a wide range of molecules from slightly polar to hydrophilic and also ionises acidic and basic functional groups which are abundantly present in NOM. Furthermore, NOM samples generally consist of water and or organic solvent mixtures, which do not interfere with the ionisation and mass spectral information. ESI can be operated in the positive and negative mode. The most frequently used ionisation mode for NOM analysis using FTICR is negative mode [141], which is likely based on the assumption that NOM contains mainly acidic functional groups and which ionises (deprotonates) well in the negative mode. However, it is known that NOM compounds also contain basic functional groups (e.g. amines), which readily ionise (protonates) in the positive mode. Ohno et al. [141] performed a comparison study between negative and positive ESI-HRMS for extracted soil DOM, in which it was shown that most compounds were detected in the negative mode, but it was also shown that a 43% increase in total number of compounds detected was achieved when using both ionisation modes. Especially for aliphatic and carbohydrates classes of molecules, the combined modes showed improved detection. Therefore it is recommended to use both negative and positive ESI for a complete characterisation of NOM. An advantage of ESI is that it is a soft ionisation technique, therefore almost no in-source fragmentation will occur and only a small amount of fragment ions can be present in the obtained spectra. A disadvantages of ESI is that adduct formation can occur when salts are present in the sample (e.g. NaCl). The adducts can interfere with annotation of NOM compounds. Another disadvantage of ESI is that multiple charged species can be formed when analysing relatively large compounds (> 500 m/z). However, recent studies have shown that the majority of ions formed with ESI in NOM samples are singly charged [131, 133]. Also other ionisation techniques such as atmospheric pressure chemical ionisation (APCI) [142] and atmospheric pressure photo ionisation (APPI) [143, 144] have been used for NOM characterisation, although to a much smaller extent. These techniques are useful for detecting non-polar compounds, which are not detected properly using ESI.

3.2.4.3 Characterisation

Due to the complex nature of NOM, analysis of NOM samples by FTICR-MS yields large data sets with thousands of components per sample. The first step in data processing is to assign the molecular formulas for the components detected, which can be calculated using the accurate mass of a component. When the molecular formulas are determined a van Krevelen diagram [145] can be constructed for visualising the obtained data. The van Krevelen diagram plots the molar H/C ratio on y-axis and molar O/C ratio on the x-axis. The diagram is also employed for visualising the different compound classes in NOM. See Figure 3-6 for an example of a van Krevelen plot of swamp DOM. The clear trend lines in the van Krevelen diagram mark different chemical transformations such as methylation, hydration, carboxylation an hydrogenation.

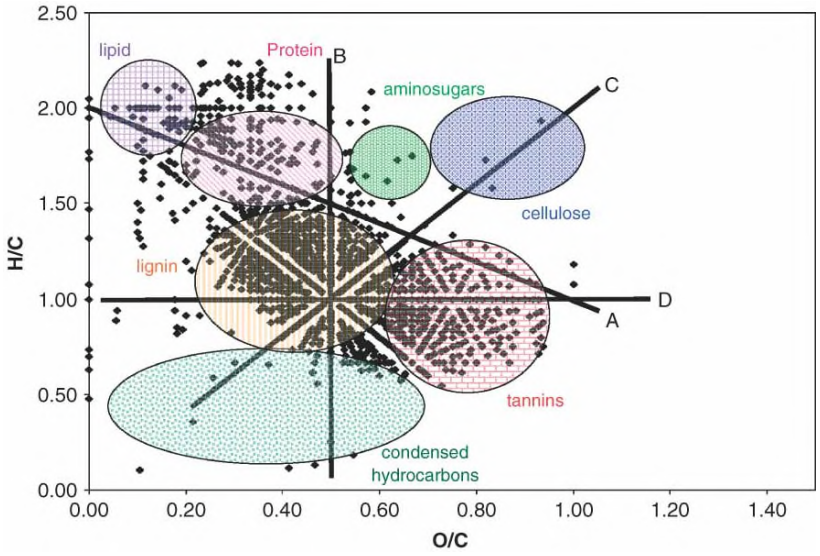


Figure 3-6: Van Krevelen diagram of swamp DOM. The compound classes are represented by the circles in the plot. The lines mark the following chemical reactions: (A) methylation/demethylation, (B) hydrogenation/dehydrogenation, (C) hydration/condensation, (D) oxidation and reduction. Source Sleighter et al. [130].

In addition to the two-dimensional van Krevelen diagram is it is also possible to visualise the information in three dimensions, by adding another molar ratio as z-axis (e.g. N/C or S/C). Currently, the van Krevelen diagram is the most widely used approach for visualising molecular formulas of large NOM data sets [146]. A drawback of the van Krevelen plot is that the dimensions of the diagram are not truly orthogonal, because the H/C and O/C ratios are not truly independent from each other. Moreover, when a molecule falls into one of the compound classes, it is not certain that it really belongs to that compound class, just because of the calculated O/C and H/C ratio. In the end van Krevelen diagram provides a broad overview of the properties of a NOM sample and is useful for characterising NOM, however, the diagnostic value remains limited. Another way for characterising NOM is by determining the different elemental ratios (H/C, O/C, N/C and S/C) and double bond equivalents (DBE) for the whole sample in order to obtain a certain profile. The average H/C ratio gives information about the aromaticity of a sample, while O/C and N/C give information about the polarity of a sample. In the end there are many ways to characterise a NOM sample using HR-MS data, but online or offline coupling using other techniques such as NMR or liquid chromatography will provide more comprehensive data.

In the literature many applications of FTICR-MS have been described for characterising NOM by direct infusion of the sample (after isolation) into the mass spectrometer. Until recently this approach was only applicable for FTICR-MS, due to the ultrahigh resolution. But with the introduction of the Orbitrap mass spectrometer [147] which provided resolutions up to 100,000, this approach was also applicable for the Orbitrap mass spectrometer. Since then, NOM characterisation studies are also performed using the Orbitrap mass spectrometer [148-150]. The first generation of Orbitrap mass spectrometers (2005) were outperformed by FTICR mass spectrometers [148]. Current high-end Orbitrap mass spectrometers such as the Orbitrap Fusion can achieve resolutions up to 500,000 – 1,000,000, which is more than sufficient for NOM characterisation. An advantage of the Orbitrap mass spectrometer is that it has relatively fast scan times (1 Hz @ 500,000 resolution) and can therefore be easily coupled to liquid chromatography. In Chapter 4 various LC-Orbitrap applications are discussed.

4 Chromatographic and hyphenated characterisation methods and applications

4.1 Chromatography characterisation methods

As discussed in the previous chapters, NOM is a very complex heterogeneous mixture of organic compounds for which it is challenging to provide a comprehensive characterisation. This holds especially when analytical detection techniques are applied to direct NOM characterisation without any prior fractionation or separation. Non-selective analytical methods (e.g. DOC, SUVA, UV, Flu) only give information about the bulk properties of a sample, or to limited fractions of sample (when fractionated). Only analytical techniques such as multidimensional NMR or HR-MS can provide enough selectivity for obtaining information about the molecular composition of NOM, without any prior fractionation. However, these unfractionated samples are quite complex and often lead to considerable spectral overlap. Thus, in order to improve NOM characterisation for both selective and non-selective detection techniques, hyphenated chromatography techniques are needed for further reducing the complexity of a NOM sample before detection. This chapter discusses the following chromatographic methods which have been applied for the online separation and fractionation of NOM; liquid chromatography, gas chromatography and field-flow fractionation. Due to nature of NOM (molecular size and polarity) liquid chromatography will be discussed in more detail.

4.2 Liquid chromatography

Liquid chromatography is the most used chromatographic separation technique for the online fractionation and separation of NOM. The following liquid chromatography methods are discussed: size exclusion chromatography, reversed phase liquid chromatography, hydrophilic interaction chromatography (HILIC) and 2D LC x LC chromatography.

4.2.1 High pressure size exclusion chromatography

High pressure size exclusion chromatography (HPSEC) is a chromatographic technique which separates analytes according to their size and occasionally to the molecular weight (see 2.2.5 for more information). Multiple detection techniques have been coupled to HPSEC for NOM characterisation such as UV or PDA detection [84], organic carbon detection (OCD) [82, 151] and fluorescence detection [152]. The most widely used detection techniques for online HSPEC fractionation of NOM are UV or PDA detectors, which provide quantitative and qualitative data about the molecular size distribution in a NOM sample [46, 81, 84, 153, 154]. The advantage of using UV or PDA detectors is that they are rapid and are relatively easy to use. A limitation of using UV or PDA detectors is that they only detect analytes at the selected wavelength. Furthermore, it only provides information about chromophoric NOM, rather than the total NOM pool. Another detection technique which is frequently used for NOM characterisation is fluorescence spectroscopy. However, HPSEC online specific Ex/Em fluorescence detection provides only limited information and less resolution is obtained in comparison with UV absorbance detection [152]. On the other hand, online 3D EEM fluorescence detection (see also 3.1.5) provides, in addition to molecular size, more information about the chemical and structural properties of NOM compared to UV absorbance [152], Although the interpretation of the resulting spectra is more difficult.

The organic carbon detector (OCD) is another online SEC detector which is frequently used for the characterisation of NOM. It offers more reliable quantitation over UV absorbance spectroscopy, because it directly detects the amount of carbon present in a fraction. Furthermore, when using an OCD detector it becomes possible to determine the amount of biopolymers and high molecular weight (HMW) polysaccharides, which are not detectable by UV absorbance detectors (see Figure 4-1). Since the introduction of SEC-OCD (also known as LC-OCD) in 1969 [155], substantial improvements have been made to its design, increasing the detection limits and making this analytical method applicable for most natural waters. Because of the excellent quantitation properties of the OCD detector and relative ease of use, SEC-OCD is very useful technique for characterising the NOM content in water treatment systems and determining the NOM removal efficiency.

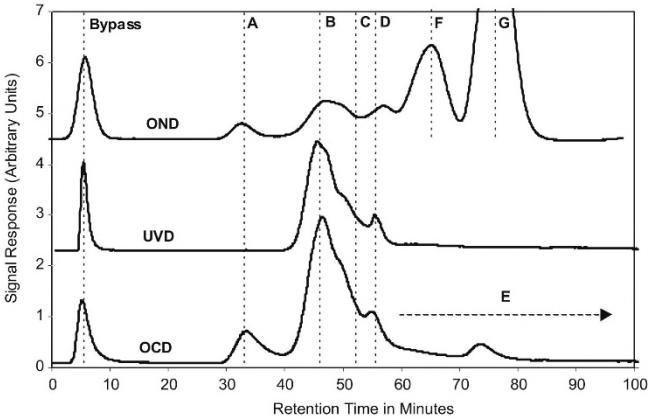


Figure 4-1: SEC-OCD-UV-OND chromatogram of the Pfinz river, showing the responses of the OCD detector, UV-detector (UVD) at 254 nm and organic nitrogen detector (OND). Fraction A: Biopolymer and HMW polysaccharides, fraction B: humic substances, fraction C: Building blocks (breakdown products of HS), fraction D: low molecular weight organic acids, fraction E: low molecular weight neutrals (alcohols, aldehydes, ketones, and amino acids), fraction F: nitrate, fraction G: ammonium. Source Huber et al. [82].

In the literature one study was found about the online coupling of SEC with HR-MS. Reemtsma et al. [136] demonstrated an online coupling between HPSEC and FTCIR-MS, in which three fractions were obtained using SEC; a low molecular weight fraction, a medium molecular weight fraction and high molecular weight fraction. The advantage of this method is that it enables to distinguish between fragments of high molecular and low molecular fraction and to find characteristic differences in the elemental compositions detected between the different fractions.

Recently, Huber et al. [82] demonstrated a novel organic nitrogen detector (OND) coupled to SEC for the detection of organic nitrogen (e.g. bound to NOM) and inorganic nitrogen. Using the OND detector, many of the same fractions could be detected as with the OCD detector, due to presence of nitrogen in biopolymer (amino acids) and humic substances. In addition, the OND can also detect nitrate and ammonium (see Figure 4-1). SEC is frequently coupled to multiple online detection techniques such as UV-OCD [151], and OCD-UV-OND [82] for obtaining better quantitative and qualitative data of NOM.

4.2.2 Reversed phase liquid chromatography

Reversed phase liquid chromatography (RP-LC) is the most popular mode for performing liquid chromatography in general. RP-LC employs a non-polar stationary phase (e.g. C18 column) and polar mobile phase in which analytes are eluted from the analytical column by polarity. RP-LC is not often used for the characterisation of NOM using online spectroscopy detection techniques such UV, PDA and fluorescence [156-158]. This is due to the low

selectivity of these techniques, causing that mostly an unresolved “hump” is observed in the chromatogram when using RP-LC that provides only limited information about the composition of NOM. Therefore RP-LC is more used as fractionation technique when low selectivity detectors are employed [85-87].

4.2.2.1 Reversed phase liquid chromatography – high resolution mass spectrometry

Until recently, few applications for NOM characterisation were described in the literature about the coupling of reversed phase liquid chromatography (RP-LC) and mass spectrometry. In 2008 Mawhinney et al. [159] described a method for characterising fulvic acids using RP-LC coupled to a quadrupole time-of-flight mass spectrometer (QTOF). Using the RP-LC-QTOF method fulvic acids could be separated and detected by the TOF mass analyser, and structural information was obtained by recording product ion spectra of fulvic acids. Due to the relatively low resolution of the employed QTOF mass spectrometer (9900 @ m/z 498) and the complexity of a fulvic acid sample, it was not possible to resolve all the detected species. As discussed before (3.2.4.1) the coupling of liquid chromatography with ultrahigh resolution FTICR-MS is challenging, due to the long scan times FTCIR-MS requires for obtaining satisfactory spectra. TOF mass spectrometers are easily hyphenated to liquid chromatography, but lack the resolution and thus resolving power which is needed for the comprehensive characterisation of NOM. With the introduction of the high resolution Orbitrap mass spectrometer [147], a fast scanning ultrahigh resolution technique became available for coupling to liquid chromatography. Although, the first studies using an Orbitrap mass spectrometer for the characterisation of NOM used a direct infusion (DI) approach, without any chromatographic separation [148, 150, 160]. It is only since recently that several studies demonstrated methodologies using RP-LC Orbitrap-MS for characterising NOM [137-139].

Petras et al. [138] was the first to describe a non-target workflow for the characterisation of DOM using RP-LC Orbitrap-MS. In this study a RP-LC system using a Kinetex C18 column with acetonitrile, ultrapure water and 0.1% formic acid as mobile phase was used, coupled to an Q Exactive Orbitrap mass spectrometer with a resolution of 140,000 (at m/z 200), which was used in the positive and negative mode. The samples (5 replicates per sample) were enriched by SPE pre-treatment using PPL cartridges as described by Dittmar et al. [64]. The data analysis was performed using MZmine2 [161], which was used for peak picking and data extraction. The molecular formulas were annotated using an in-house R script. Furthermore spectral networking using MS/MS data was used for improving the annotation of the molecular formulas. From the obtained data, Van Krevelen, O/C and H/C plots were derived. This study showed that using liquid chromatography, molecular formula annotation was enhanced by the alignment of chromatographic peak profiles and using both ionisation modes which reduces the redundancy of different ion adducts.

In the study of Patriarca et al. [137] an online HPLC-ESI-HR-MS method was developed for the analysis and comparison of different DOM samples. Furthermore, a comparison was made between analysing the samples using liquid chromatography and direct infusion. For this study an RP-LC chromatographic system was employed using an Agilent PLRP-S poly(styrene/divinylbenzene column) using a mobile phase consisting of acetonitrile, ultrapure water and 0.1% formic acid. For the data acquisition a LTQ-Velos-Pro Orbitrap mass spectrometer was used, equipped with an ESI source which was operated in the negative mode. The spectra were recorded at a resolution of 100,000 (at m/z 400). The samples were extracted by SPE using PPL extraction cartridges and were analysed in triplicate. Data analysis was performed by cutting the chromatograms in segments of 3 or 4 min. and by exporting the spectral data as a mass list. The formulas were assigned by an in-house developed Matlab routine. The data was visualised by using van Krevelen and Kendrick

diagrams which were used for comparing the results between chromatographic analysis and direct infusion for Suwannee river fulvic acid (see Figure 4-2). In Suwannee river fulvic acid, 4157 peaks were detected using the LC method compared to the 3416 peaks that were detected using the DI method, which is an increase of 22%. The van Krevelen plot also shows when superimposing the LC and DI profiles, that the same region is covered, which suggest that all the Fulvic acid NOM was successfully eluted from the analytical column. In the end 1116 peaks were uniquely detected with the LC method and 375 peaks with the DI method. In the Kendrick mass diagram, the Kendrick mass defect (KMD) is plotted as a function of the Kendrick mass, which visualizes homologues series, that differ by a methylene unit and have the same KMD. Most importantly, this Kendrick diagram shows that the majority of the low molecular weight fulvic acids (≈ 200 m/z) is only detected using the LC method.

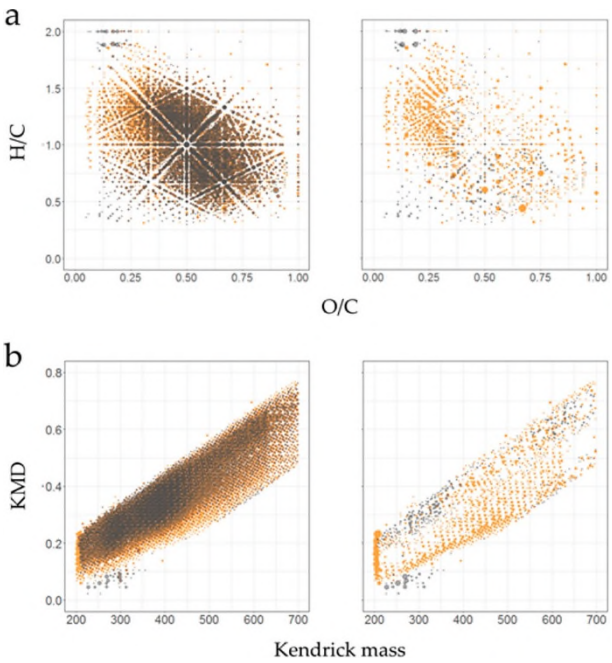


Figure 4-2: Van Krevelen and Kendrick diagrams of Suwannee river fulvic acid comparing the results obtained between liquid chromatography and direct infusion. A). Left van Krevelen diagram showing all the assigned peaks by the LC method (orange, 4157 formulas) and DI method (black 3416 formulas) ; Right: unique assigned peaks for the LC method (orange, 1116 formulas) and DI method (black, 375 formulas); B) Corresponding Kendrick diagrams of the data described by the van Krevelen diagrams in panel A. Source Patriarca et al. [137].

The study of Patriarca et al. showed that when chromatography is applied, less charge competition occurs at the ESI source, which enhances signal intensities and S/N ratios and substantially increases the number of detected species. Another benefit of using chromatography is that less contaminants (e.g. salts) are introduced into the ESI source, which reduces signal suppression. In the end it was demonstrated that substantially more information was obtained using the LC method in comparison with the DI method.

In another recent study by Verkh et al. [139], a non-target HPLC-HR-MS method was employed for the data analysis of DOM in waste water treatment. In this study the DOM profile of influent water was compared with the DOM profile of effluent water. The chromatographic separation was performed using a Waters Acquity UPLC BEH C18 column using a mobile phase consisting of acetonitrile, ultrapure water and 0.1% formic acid. No sample isolation or enrichment was performed on the samples. After filtration, 50 µL of sample was directly injected onto the analytical column for analysis. Mass spectrometry

analysis was conducted using a LTQ-Velos-Pro Orbitrap, equipped with an ESI source which was operated in the positive and negative mode at a resolution of 100,000 (at m/z 400). The data analysis was performed using MZmine2 which was used for peak picking, detection of adducts and calculating the molecular formula. Using an in-house developed R script, the molecular formula candidate with the smallest absolute mass deviation and highest isotopic pattern score was selected. From the obtained data, Van Krevelen and Kendrick diagrams were made, which were used for comparing influent DOM with effluent DOM. With this approach significant changes in DOM were identified, furthermore, it also showed that it has the potential to provide a chemical fingerprint of DOM transformation in waste water treatment.

These three approaches using LC-Orbitrap for NOM characterisation show that chromatographic separation provides another dimension of information (i.e. retention time) and is beneficial for distinguishing molecular features by reducing the complexity of NOM, in comparison with DI approaches. Furthermore, it was shown that less signal suppression was observed when using chromatographic separation and ESI, which increases S/N ratios and increases the change of detecting low intensity species.

4.2.3 HILIC chromatography

Hydrophilic interaction chromatography (HILIC) is still a relatively new technique developed for the separation of polar analytes [162]. HILIC is a variant of normal phase chromatography and partly overlaps with other separation techniques such as ion chromatography and RP-LC. Like normal phase chromatography, HILIC also uses polar stationary phases such as silica, amino or cyano. However, the mobile phase used is comparable to those employed by RP-LC. A typical HILIC application starts with a high percentage of organic phase such as acetonitrile which is replaced over time by an increasing percentage of water. The principle behind HILIC is that the mobile phase forms a water-rich layer on the surface of the polar stationary phase, which acts as the true stationary phase creating a liquid-liquid extraction system. With HILIC chromatography the analytes are eluted in the order of increasing hydrophilicity, although the elution is dependent of the interactions between the analyte and the stationary phase such as dipole-dipole interactions, electrostatic interactions, hydrogen bonding and hydrophobic interactions.

The first application of HILIC for the characterisation of NOM was reported by Woods et al [88]. HILIC was used as a fractionation technique for the separation of Suwannee river DOM into 80 fractions which were characterised (offline) by NMR. For the fractionation a diol functionalised silica HILIC column was used which was monitored by both a PDA and a fluorescence detector (see Figure 4-3).

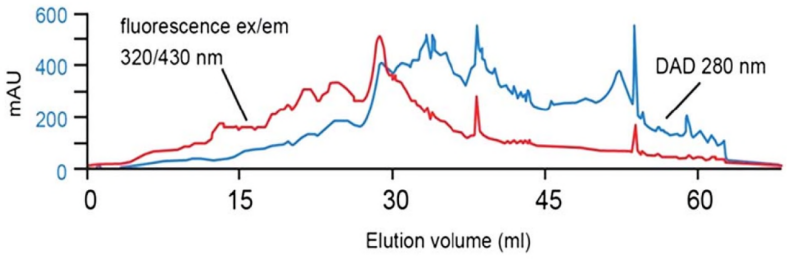


Figure 4-3: HILIC chromatogram of Suwannee river DOM recorded using DAD (blue) and fluorescence (red) detection. Source Sandron et al. [43] and Woods et al. [88].

It was shown that DOM components were eluted with increasingly polarity along the chromatogram, which demonstrates that the method is applicable for the separation of a wide variety of compounds that are present in a complex mixture such as DOM. Furthermore,

the HILIC fractionation reduced the chemical heterogeneity substantially, which resulted in that distinct signals were detected during ^{1}H NMR analysis that are indicative for discrete chemical compounds, which otherwise remained undetected when no fractionation was applied. In the end this application for DOM separation showed that molecular-level elucidation is possible using NMR when DOM is substantially simplified by HILIC fractionation.

4.2.4 LC x LC chromatography

In order to substantially reduce the complexity of NOM for comprehensive characterisation, several studies explored the possibilities of two dimensional LC x LC chromatography (2D-LC). Duarte et al. [163] investigated two different 2D-LC approaches for the characterisation of fulvic acids reference samples (IHSS). For the first approach a C18 column was used in isocratic mode (20% acetonitrile) as first dimension, for the second dimension a SEC column (polyhydroxymethacrylate copolymer) was employed in isocratic mode (11% acetonitrile in 20 mM NH_4HCO_3, pH 8). For the second approach, the first dimension consisted of a HILIC column (alkyl diol stationary phase) which was used in reversed phase mode (also known as per aqueous liquid chromatography) in isocratic mode (11% acetonitrile in 20 mM NH_4CO_3, pH 8). The second dimension was the same as the second dimension of the first approach. For both approaches three online detectors were used: UV absorbance (254 nm), fluorescence (Ex/Em, 240/450 nm) and evaporative light scattering detector (ELSD). The results showed that the strategy using comprehensive 2D-LC for NOM characterisation is promising, although the 2D chromatograms still showed some unresolved fractions. Both method yielded comparable results, and hold promise for further optimisation. Furthermore, It was shown that smaller molecular weight group fractions seem to be more hydrophobic of nature and the apparent molecular weight distribution of the fulvic acids samples determined by both approaches using UV detection was lower than reported in literature.

Another 2D-LC application was reported by Woods et al [89], who employed a 2D-LC approach (HILIC x HILIC) coupled (offline) with NMR for the characterisation of Suwannee river DOM. For the first dimension a diol functionalised silica HILIC column was used, for the second dimension a bare silica was employed. However, the columns used are not truly orthogonal to each other. In 2D-LC, high orthogonality leads to higher resolvability, but can also be troublesome due to incompatible mobile phases and column bleed, which can lead to problems with subsequent NMR analysis [89]. Despite the limited orthogonality, improvement in DOM fraction was achieved, which enabled extensive and in-depth NMR analyses. Furthermore, this study showed that highly oxidized sterols and hopanoid-type structures are also present in DOM and are potentially major components in DOM.

4.3 Gas chromatography

Gas chromatography (GC) is a common chromatographic technique for the separation of compounds in the gas phase by distributing the sample components between stationary and mobile phase. The mobile phase is a carrier gas (i.e. helium, hydrogen, nitrogen) which flows through the analytical column (stationary phase) and separates compounds according to their thermal volatility. Therefore, gas chromatography is mainly applicable for volatile compounds or compounds who can be made volatile by derivatisation. The separation of a complex samples containing NOM using GC is challenging, because compounds containing functional groups such as OH, COOH, NH, and SH, can form hydrogen bonds that can reduce volatility and can also interact negatively with the stationary phase (e.g. peak broadening). Furthermore, many NOM compounds are thermally labile meaning that temperature control during GC analysis is challenging. Besides, the larger NOM structures are often not volatile enough for GC analysis. In order to overcome these difficulties, degradation and derivatisation reactions such as pyrolysis, alkylation and silylation are often used. Using

pyrolysis, large complex molecules can be broken down into smaller and more volatile fragments by applying heat under a controlled atmosphere. The major disadvantage of pyrolysis is the unintended decomposition of thermally labile classes of compounds. Derivatisation reactions, such as alkylation, target active hydrogens on amines and acidic hydroxyl groups and replace these by an aliphatic or aliphatic-aromatic group by esterification. A common alkylation reagent for NOM analysis is tetramethylammonium hydroxide (TMAH) [164-166]. Silylation derivatisation reactions target active hydrogens on alcohols, acids, amines, amides, thiols, aldehydes and replace these with a silyl group (e.g. trimethylsilyl). Silylation is very effective for volatilising non-volatile compounds and reducing their polarity. A silylation reagent often used for NOM analysis is Bis(trimethylsilyl)trifluoroacetamide (BSTFA) [167, 168].

There are two modes for NOM analysis by GC; (i) a target approach by analysing classes of molecules such as fatty acids, lignin, lipids, phenols, and sugars [165, 167-169] or (ii) a non-target screening approach for characterising the total NOM pool in a sample [164, 166, 170]. For the characterisation of the total NOM pool, pyrolysis is often employed for the degradation of the large and complex molecules in NOM into smaller more amenable fragments. Different modes of pyrolysis and non-target screening applications are discussed in section 4.3.1.

4.3.1 Pyrolysis-gas chromatography/mass spectrometry

Pyrolysis-gas chromatography/mass spectrometry (Py-GC/MS) is considered a powerful technique for NOM characterisation, due to its convenience, speed of analysis and for providing structural information about NOM [171]. With Py-GC/MS it is possible to perform pyrolysis on large complex molecules and break them down into smaller fragments. After pyrolysis, the smaller fragments of the parent compound are transferred to the analytical column were the fragments (or analytes) are separated by volatility and structural interaction with the stationary phase. The analytes are then detected by a mass spectrometer, which yield mass spectra, which can be compared with the mass spectra obtained from analytical standards or mass spectral libraries for identification. Py-GC/MS can be employed for the quantitative and qualitative analysis of NOM since it produces a fingerprint of the organic material.

A variety of pyrolysis methodologies have been used for the analysis of NOM such as platinum filament, Curie point and microscale sealed vessel pyrolyser (MSSVpy). Platinum filament is probably the most commonly used pyrolyser. This type of pyrolyser uses quartz sample tubes for holding the sample, which has the disadvantage that the sample does not make direct contact with the filament, and which makes it difficult to determine the exact pyrolysis temperature. The Curie point pyrolyser is technique that uses a ferromagnetic wire sample holder that is heated to Curie point of this ferromagnetic sample holder. At the Curie point (range of 300 – 800 °C) the sample holder's permanent magnetism turns into induced magnetism. The drawback of this technique is that it is limited to the Curie point temperature of the sample holder, and there is almost no flexibility in using other temperatures. New pyrolysis techniques using controlled thermal treatment strategies such as MSSVPy, have shown to promote the reduction of polar NOM moieties and substantially improve the yields of GC amenable fragments [172, 173]. MSSVpy is performed using very small amounts of sample which are placed in a sealed tube and heated in the range of 250 – 350 °C for several days. This pyrolysis technique has shown to provide better results for polar biochemicals [172, 173].

Schulten et al. [174] used two approaches for determining the basic molecular structures of humic macromolecules in DOM, namely Curie point Py-GC/MS and pyrolysis-field ionisation

MS. For this study, samples of Hohloh lake were pre-treated using freeze-drying and XAD resin fractionation. The results obtained using Py-GC/MS showed fourteen intense peaks and many co-eluting peaks in the chromatogram. Using in-house and reference mass spectral libraries, many compounds classes were identified such as benzenes (42 identified structures), furans (35) and phenols (26). Furthermore, a substantial amount of aromatic structures were detected, which could not be elucidated. Therefore Schulten et al. suggested to use complementary analysis techniques such as isotope ratio measurements for improving identification.

Kracht et al. [170] was the first to combine Py-GC/MS with isotope ratio mass spectrometry (Py-GC/MS-IRMS), in order to improve identification and to obtain more comprehensive information about the origin of DOM. The approached was applied to freeze-dried peat samples and freeze-dried bog water. Only the most intense peaks in the IRMS chromatograms were identified using mass spectrometry (GC-MS), because only those were detected simultaneously by mass spectrometry and isotope ratio analysis. Over 60% of the major pyrolysis products that were studied, were derivatives of furan and phenol. Furthermore, several derivatives of pyran, benzene and cycloalkenones were identified using this approach. In the end, the isotopic and structural information of Py-GC/MS-IRMS provided extra information on peat formation. A limitation of this approach is that freeze-drying is used as sample pre-treatment, which is time consuming and requires additional sample desalting for samples with high salt content. Furthermore, freeze-drying is solute dependent, which means that every class of compounds has different requirements for freeze-drying conditions[43].

Berwick et al. [173] made a comparison between MSSVpy and traditional and flash pyrolysis techniques (e.g. platinum filament and Curie point), in order to determine the characterisation potential of MSSVpy for North Pine reservoir NOM. In this study, water samples were concentrated 5-fold using reverse osmosis and rotary evaporation, which was subsequently extracted by superimposed XAD-8/XAD-4 resin system. This yielded hydrophobic (HPO) and transphilic (TPI) fractions. The results of this study showed that a total of 131 products were detected using MSSVpy in the HPO and TPI fractions which were classified into ten major product groups. See Figure 4-4 for the ten major product groups and the TIC chromatogram of the HPO and TPI fraction. The MSSVpy TPI fraction consisted for more than 30% out of nitrogen containing products, which originate from humic substances but also from peptides, proteins and amino sugars. This shows that MSSVpy is really applicable for large complex structures such as humic acid and biomolecules. Furthermore, the intensity of the products detected during the MSSvp analysis was about half an order of magnitude higher than that of traditional flash pyrolysis, which demonstrates the higher sensitivity of MSSVpy, In conclusion, MSSVpy has shown that it is a promising new analytical approach for NOM characterisation.

Py-GC/MS has been used frequently for the study for natural organic matter, due to its structural elucidation capabilities and has made a substantial contribution to the understanding of natural organic matter. However, some limitations exist in using Py-GC-MS such as extensive fragmentation, secondary reactions of primary products and undesirable reactions of polar pyrolysis products, which all hamper the representation or elucidation of the parent material. Therefore, pyrolysis is often used as a complementary technique in combination with other techniques such as isotope ratio, SEC and NMR analysis in order to improve structural elucidation of NOM.

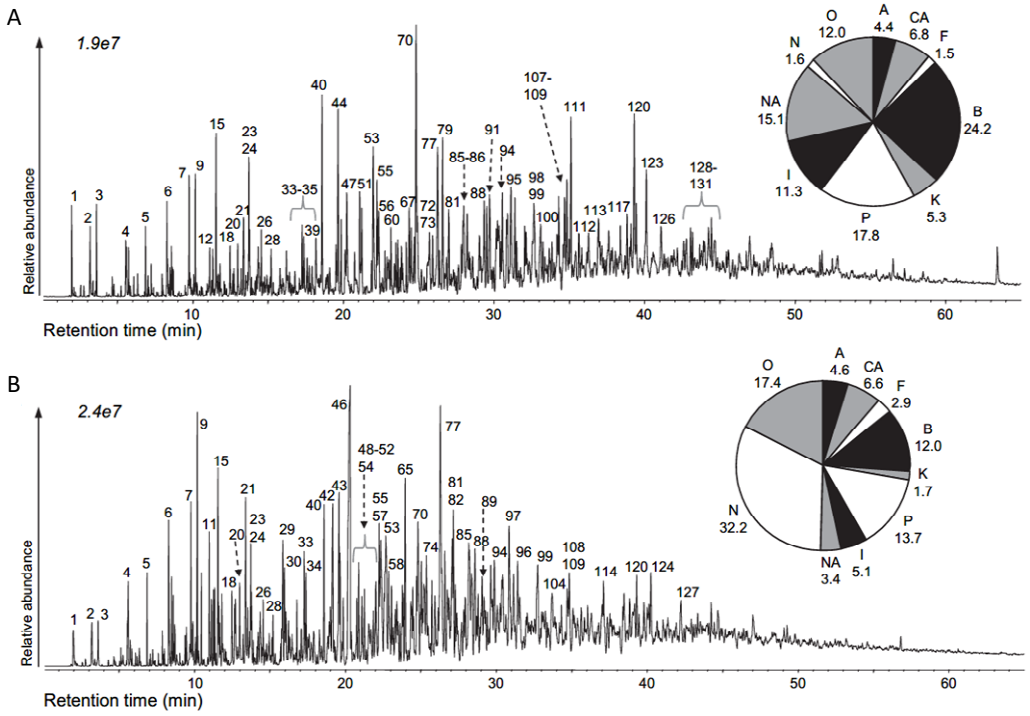


Figure 4-4: MSSVpy TIC GC-MS data of the HPO fraction (A) and HPI fraction (B). The pie chart reflects relative proportions of ten major groups. A = aliphatic, CA = cycloaliphatic, F = Furans, B = benzenes, K = ketones, P = phenols, I = indenes, NA = naphthalenes, N = nitrogen products and O = other. Adjusted from Berwick et al. [173].

4.4 Field-flow fractionation

Field-flow fractionation (FFF) is a chromatography-like technique for the separation/fractionation of macromolecules, colloids and particles. This technique was first introduced and reported by Giddings [175]. FFF is performed by applying an external field perpendicular to the direction of the sample flow in a small thin channel. The external field suppresses the particles near the accumulation wall of the channel. After a while the particles relax under the applied field, and are eluted by the applied laminar channel flow. The smaller particles elutes faster than the larger particles. FFF is used for the fractionation of colloidal organic matter but also for characterisation and determination of the molecular size distribution of DOM [176-178]. Most FFF applications for DOM analysis are conducted using an UV absorbance detector [176-178], however, more recent applications also used other detectors such as fluorescence, inductively coupled plasma mass spectrometry (ICP-MS), and solid state NMR [178-180]. An example of a FFF application using multiple detections techniques was demonstrated by Stolpe et al. [178]. They used FFF coupled on-line to UV absorbance and fluorescence detectors and ICP-MS in order to investigate the colloidal size of chromophoric DOM, fluorescent organic matter and colloidal size distributions of trace elements in various rivers in the region of the northern Gulf of Mexico. Combining the results obtained from these techniques Stolpe et al. showed that two major populations of colloids were found as carriers for elements in the investigated waters; (i) small chromophoric DOM (0.5-4 nm) that binds most elements and (ii) larger (5-40 nm) Fe-rich colloids binding Mn, P, and Pb.

Currently, one of the major issues with FFF is the availability of appropriate calibration standards for molecular size distribution. In most cases polystyrene sulfonate standards are used, which share little similarity with complex NOM structures, and lead to erroneous molecular weight estimations [43]. Another disadvantage is the presence of surfactants in

the carrier solution, which can cause denaturation of NOM components in the sample and therefore leads to changes in the tertiary structure. Currently, FFF lacks the maturity of other separation techniques and published methods [181]. Therefore, FFF should be viewed as a complementary technique to other NOM characterisation techniques.

5 Recommended isolation and analytical methodologies for the characterisation of NOM at KWR

This chapter discusses the most promising isolation/fractionation methods and analytical techniques used for the characterisation of NOM in water, and provides recommendations for the most fitting analysis techniques for NOM characterisation to be used or implemented at KWR. Additionally, a LC-HR-MS NOM characterisation approach is provided.

5.1 Isolation, concentration and fractionation methods

The literature study has shown that SPE currently is the most promising approach for NOM isolation and concentration. The SPE method developed by Ditmar et al. [64] using PPL PS-DVB cartridges showed that the extraction procedure is simple, fast and also provides high extraction yields (> 60%). Furthermore, the SPE method can be easily implemented for NOM characterisation at KWR. A disadvantage of this SPE procedure is that only limited amount fractions can be obtained. A solution for this could be the SPE fractionation method developed by Ratpukdi et al. [70], who used four different SPE columns (total of five) in order to separate NOM into six fractions in less approximately 6 hours, which is substantially faster than traditional fractionation techniques. The drawback of this approach is the low recovery of the hydrophobic base fraction which is only 2.3%. Alternatively, XAD resin extraction still remains a viable option for the isolation and fractionation of NOM. Although, sample pre-treatment is very labour intensive and time consuming (> 24h), good extraction yields are obtained. Using the extraction method developed by Leenheer and Croué [54] six different fractions can be obtained while the extraction method of IHSS [61] provides two fractions; a humic and fulvic acid fraction. A XAD isolation and fractionation setup is available at KWR and therefore remains an option, especially when humic and fulvic acid fractions are required.

When size separated NOM fractions are required, there are basically two options; size exclusion chromatography or field-flow fractionation. Both techniques are perfectly capable of providing size distributed fractions. A limitation of both techniques is that polystyrene sulfonate standards are used, which are different than complex NOM structures, which results in an approximation of the molecular weight. Currently, size exclusion chromatography is not used at KWR, however, asymmetrical flow field-flow fractionation (AF4) a sub-technique of FFF is available at KWR and is perfectly suited for obtaining size separated NOM fractions when required.

5.2 General analysis methods

For NOM analysis there are only a few general parameters that are of importance (see section 3.11-3.15). By far the most important general parameter for NOM analysis is DOC, because the amount of DOC measured relates directly to amount of DOM that is present in a sample. A TOC analyser is available at KWR, and was already employed in previous NOM studies [12, 182]. The determination of the DON content can be of importance, although, accurate DON analysis is rather difficult. Other parameters and analysis methods such as UV, fluorescence and SUVA can be valuable for NOM analysis in general, but are limited by the lack of

selectivity compared to other analytical techniques and are therefore not an viable option for NOM characterisation at KWR.

5.3 Hyphenated and in-depth characterisation methods

Multiple analytical techniques have been discussed in this literature study for the characterisation of NOM. Hyphenated analytical techniques such as SEC-OCD-UV-OND (or LC-OCD) have shown to provide valuable qualitative and quantitative data for NOM characterisation despite the limited selectivity of the applied detectors. SEC-OCD has shown to be an important analytical tool for creating NOM profiles and for water treatment applications. For water treatment applications SEC-OCD can provide the amount of NOM removal for each fraction, which is valuable information for water treatment research. The SEC-OCD technique is not available at KWR, however, NOM SEC-OCD analysis is performed by multiple contract laboratories and is therefore easily outsourced.

5.3.1 NMR and HR-MS

In order to perform NOM characterisation at a molecular level, only two analytical techniques are eligible for this purpose; NMR and HR-MS analysis. Literature has shown that both techniques are extensively used for providing a comprehensive overview of the amount and type of NOM structures present in a sample.

NMR has shown to be a powerful discovery tool for the molecular characterisation of NOM. While 1D NMR provides limited structural information, 2D NMR provides a comprehensive overview of the type of structures present in a sample. However, when identification of individual compounds is required, extensive fractionation before NMR analysis has to be conducted, in order to substantially reduce the complexity of NOM. At this time, no NMR instrumentation is available at KWR, and it is not likely that NMR instrumentation is acquired in the near future. Therefore, when NMR analysis is required for comprehensive characterisation of NOM, collaborations with NMR research groups are needed.

Ultrahigh resolution mass spectrometry has shown to play a key role in the understanding of NOM by providing molecular-level characterisation. The elemental compositions of thousands of NOM compounds have already been determined by HR-MS. Literature has shown that NOM characterisation was primarily performed by using FT-ICR-MS and by means of direct infusion. With the introduction of the Orbitrap mass spectrometer, the coupling between ultrahigh resolution mass spectrometry and liquid chromatography became possible and showed to be a valuable improvement for NOM characterisation (see 4.2.2.1). Since KWR has an ultrahigh resolution (500.000) Orbitrap Fusion mass spectrometer at their disposal, the LC-HR-MS approach is the recommended approach for NOM characterisation at KWR. In section 5.5 a LC-HR-MS workflow is given in order to perform NOM characterisation at KWR, based on recent LC-HR-MS studies.

In the end, when a comprehensive characterisation of NOM at molecular level is required, the best results are obtained by combining multiple analytical techniques such as multidimensional NMR and LC-HR-MS

5.4 Overview recommended analytical methodologies for NOM characterisation at KWR

Below, an overview is presented with recommended analytical methodologies for the characterisation of NOM at KWR:

General parameters

- DOC analysis (TOC analyser) – for determining the amount of NOM present in a sample.

Concentration and isolation methods

- PPL SDB SPE extraction – Fast extraction and concentration method, developed by Ditmar et al. [64].
- XAD resin extraction – labour intensive extraction and concentration method for obtaining humic and fulvic acid fractions, according to IHSS isolation method [61].

Fractionation

- SPE extraction – for obtaining six fractions with different polarities, using three different types of SPE cartridges according to method developed by Ratpukdi et al. [70].
- XAD resin extraction (optional) – labour intensive extraction and concentration method for obtaining six fractions with different polarities using the method developed by Leenheer and Croué [54].
- AF4 analysis – fractionation of NOM by molecular size.

(Comprehensive) characterisation methods

- LC-HR-MS – comprehensive NOM characterisation using LC-Orbitrap Fusion, based on the analytical approach provided in section 5.5.
- One or multi-dimensional NMR – for comprehensive NOM characterisation (only in collaboration with NMR research groups).
- SEC-OCD-UV-OND analysis – for obtaining a NOM profile of the different size separated fractions (i.e. biopolymer, humic substances, building blocks, etc.). Analysis is performed by a contract laboratory.

5.5 Analytical approach for NOM characterisation using the LC-Orbitrap Fusion

In this section an analytical approach is proposed for the characterisation of NOM using the LC-Orbitrap Fusion at KWR. This NOM characterisation approach is derived from recent studies [137-139] using LC-HR-MS and based on the instrumentation, methods and software tools available at KWR, and is intended as an initial starting point for the characterisation of NOM. The analytical approach is shown below.

Sample preparation

The sample preparation for LC-Orbitrap analysis is dependent on the applied sample pre-treatment procedure. For the LC separation it is important to use as little as possible amount of solvent in the sample (< 10%) in order to obtain a satisfactory separation. Before analysis, dilute the sample when needed until a maximum of 10% organic solvent is reached and filter the sample using a 0.45 µm filter.

LC analysis

For the LC separation of NOM, it is recommended to use reversed phase chromatography employing a C18 analytical column. Literature has shown that primarily acetonitrile is used as organic modifier for the mobile phase [137-139], although methanol is also a viable option.

The recommended mobile phase is:

- Mobile phase A: ultrapure water + 0.05 - 0.1% formic acid
- Mobile phase B: acetonitrile + 0.05 - 0.1% formic acid

The gradient should start with a low percentage of mobile phase B (i.e. 5 - 10%) and increasing (linear) slowly to 100% B, for obtaining a satisfactory separation.

HR-MS analysis

HR-MS analysis at KWR is performed using an Orbitrap Fusion mass spectrometer. It is recommended to use a non-target screening approach for NOM characterisation. This requires a full scan data acquisition of a large m/z range in order to detect as many NOM compounds as possible. Furthermore data dependent acquisition should be used, in order to obtain MS^2 spectra of the individual NOM compounds. The resolution should be set at the highest resolution possible while ensuring that enough data points per peak are recorded. The mass accuracy should be as low as possible in order for determining the correct elemental composition of the individual NOM compounds. By using a lock mass, the mass accuracy can be lowered to < 1 ppm. It is also recommended to perform data acquisition in both positive and negative mode, in order to obtain a comprehensive overview of NOM compounds. Additionally, direct infusion experiments can also be conducted in order to compare the results obtained between LC and DI.

Overview Orbitrap Fusion analysis requirements:

- Full scan data acquisition: range 100 – 1500 m/z
- Data dependent MS^2 acquisition
- Resolution: 240,000 – 500,000
- Mass accuracy calibration: < 2 ppm
- Apply lock mass: improving mass accuracy < 1 ppm
- Positive and negative mode analysis

Data processing

There are multiple software tools available for processing the acquired data. The recommended software tools are MZmine2 or Compound Discoverer. Both tools are used for peak picking, grouping of fragments and adducts, and determining elemental composition of a compound. Currently, MZmine2 also supports the Seven Golden Rules [183] for restricting the amount of elemental compositions generated, which helps by determining the correct formula. The molecular formula is calculated by means of the mass error of the proposed formula and the isotopic pattern score. Additionally, an R or Matlab routine can be used to improve the molecular formula selection. When the molecular formulas are known, van Krevelen and Kendrick diagrams can be made. Van Krevelen diagrams can be used for grouping the different compound classes in NOM, and Kendrick diagrams for determining the homologues series that are present. Additionally, H/C, O/C, N/C, S/C ratios and double bond equivalents can be calculated for obtaining a specific sample profile.

6 Conclusion

This literature thesis provides a comprehensive overview of the analytical methodologies used for the characterisation of aquatic NOM. Many analytical methodologies are being used for the isolation, concentration, fractionation and characterisation of NOM. Characterisation of NOM has progressed substantially in recent years, due to the development of advanced instrumentation and analytical methods, such as multidimensional NMR and HR-MS. Despite the progress made in recent years, NOM characterisation on a molecular level still remains very challenging, due to its inherent complexity and heterogeneity. As a result limited information about NOM is currently known.

For the characterisation of NOM at KWR, multiple recommendations were made. The most promising extraction technique for NOM to be implemented at KWR is PPL SDB SPE extraction, which is relatively fast and provides satisfactory extraction yields (~ 60%). When fractions of different polarities are required, the method developed by Ratpukdi et al [70]. using different types of SPE extraction cartridges can be applied. For comprehensive NOM characterisation it is highly recommended to employ LC-HR-MS instead of DI-HR-MS, due to the substantial increase in detected compounds it provides. LC-HR-MS can be performed by using the LC-Orbitrap Fusion at KWR, for which an analytical approach is provided in this study.

One of the major issues with NOM analysis that still persists, is the collection of a representative sample. Many isolation, concentration or fractionation methods can alter the composition of NOM, due to extreme conditions that are often used for extracting or collecting NOM. Furthermore, extraction yields of NOM are in general substantially below 100%, which means that only a part of the total NOM pool is extracted. Therefore it is of importance to always consider the strengths, flaws and biases of the extraction methods used when the acquired data are interpreted.

The future of NOM characterisation lies in the development of unbiased and non-altering methods for NOM extraction and in combining multiple structural characterisation techniques, providing a more comprehensive view of the NOM composition, using complementary techniques such as multidimensional NMR and HR-MS. However, substantial progress needs to be made, in order to effectively process and interpret the enormous and complicated datasets generated by these techniques.

7 References

1. Chow, C.W.K., et al., *The impact of the character of natural organic matter in conventional treatment with alum.* Water Science and Technology, 1999. **40**(9): p. 97-104.
2. Leenheer, J.A., et al., *Comprehensive isolation of natural organic matter from water for spectral characterizations and reactivity testing*, in *ACS Symposium Series*. 2000. p. 68-83.
3. Gjessing, E.T., *Physical and chemical characteristics of aquatic humus*. Ann Arbor Science. 1976: Ann Arbor.
4. Fabris, R., et al., *Comparison of NOM character in selected Australian and Norwegian drinking waters.* Water Research, 2008. **42**(15): p. 4188-4196.
5. Aiken, G. and E. Cotsaris, *Soil and hydrology: Their effect on NOM.* Journal / American Water Works Association, 1995. **87**(1): p. 36-45.
6. Matilainen, A., et al., *An overview of the methods used in the characterisation of natural organic matter (NOM) in relation to drinking water treatment.* Chemosphere, 2011. **83**(11): p. 1431-42.
7. Jacangelo, J.G., et al., *Selected processes for removing NOM: on overview.* Journal (American Water Works Association), 1995. **87**(1): p. 64-77.
8. Singer, P.C., *DBPs in drinking water: Additional scientific and policy considerations for public health protection.* Journal (American Water Works Association), 2006. **98**(10): p. 73-80.
9. Richardson, S.D., et al., *Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research.* Mutation Research/Reviews in Mutation Research, 2007. **636**(1): p. 178-242.
10. Richardson, S.D. and S.Y. Kimura, *Water Analysis: Emerging Contaminants and Current Issues.* Analytical Chemistry, 2016. **88**(1): p. 546-582.
11. Krasner, S.W., et al., *Occurrence of a New Generation of Disinfection Byproducts.* Environmental Science & Technology, 2006. **40**(23): p. 7175-7185.
12. Kolkman, A., et al., *Tracing Nitrogenous Disinfection Byproducts after Medium Pressure UV Water Treatment by Stable Isotope Labeling and High Resolution Mass Spectrometry.* Environ Sci Technol, 2015. **49**(7): p. 4458-4465.
13. Delpla, I., et al., *Impacts of climate change on surface water quality in relation to drinking water production.* Environment International, 2009. **35**(8): p. 1225-1233.
14. Tran, N.H., et al., *A critical review on characterization strategies of organic matter for wastewater and water treatment processes.* Bioresour Technol, 2015. **193**: p. 523-33.
15. Zbytniewski, R. and B. Buszewski, *Characterization of natural organic matter (NOM) derived from sewage sludge compost. Part 1: chemical and spectroscopic properties.* Bioresource Technology, 2005. **96**(4): p. 471-478.
16. Vughs, D., A. Kolkman, and M. Maessen, *Chemical screening of anion exchange (AIX) regenerant.* 2016, KWR: Nieuwegein.
17. Gosling, S.N. and N.W. Arnell, *A global assessment of the impact of climate change on water scarcity.* Climatic Change, 2016. **134**(3): p. 371-385.
18. Sillanpää, M., *Natural Organic Matter in Water* IWA Publishing, 2015.
19. Thurman, E.M., *Organic Geochemistry of Natural Waters.* Martinus Nijhoff/DR. W. Junk publishers, 1985.
20. Hesse, S., G. Kleiser, and F.H. Frimmel, *Characterization of refractory organic substances (ROS) in water treatment.* Water Science and Technology, 1999. **40**(9): p. 1-7.
21. Thomas, J.D., *The role of dissolved organic matter, particularly free amino acids and humic substances, in freshwater ecosystems.* Freshwater Biology, 1997. **38**(1): p. 1-36.
22. Perdue, E.M. and J.D. Ritchie, *Dissolved Organic Matter in Freshwaters*, in *Treatise on Geochemistry*, K.K. Turekian, Editor. 2003, Pergamon: Oxford. p. 273-318.
23. Snoeyink, V.L. and D. Jenkins, *Water chemistry* 1981.
24. Xing, Y., *Characterization of dissolved organic carbon in prairie surface water using FTIR spectroscopy (M.Sc. thesis)* University of Saskatchewan, Saskatoon, Saskatchewan, Canada, 2010.
25. F.J., S., *Humus Chemistry: Genesis, Composition, Reactions.* Wiley interscience, New York, 1982.

26. Buffle, J., F.L. Greter, and W. Haerdi, *Measurement of complexation properties of humic and fulvic acids in natural waters with lead and copper ion-selective electrodes.* Analytical Chemistry, 1977. **49**(2): p. 216-222.
27. Dittmar, T., *The molecular level determination of black carbon in marine dissolved organic matter.* Organic Geochemistry, 2008. **39**(4): p. 396-407.
28. Marques, J.S.J., et al., *Dissolved Black Carbon in the Headwaters-to-Ocean Continuum of Paraíba Do Sul River, Brazil.* Frontiers in Earth Science, 2017. **5**: p. 11.
29. Wagner, S., R. Jaffé, and A. Stubbins, *Dissolved black carbon in aquatic ecosystems.* Limnology and Oceanography Letters, 2018. **3**(3): p. 168-185.
30. Hockaday, W.C., et al., *Direct molecular evidence for the degradation and mobility of black carbon in soils from ultrahigh-resolution mass spectral analysis of dissolved organic matter from a fire-impacted forest soil.* Organic Geochemistry, 2006. **37**(4): p. 501-510.
31. Dittmar, T. and B.P. Koch, *Thermogenic organic matter dissolved in the abyssal ocean.* Marine Chemistry, 2006. **102**(3): p. 208-217.
32. Pan, Y., et al., *Characterization of natural organic matter in drinking water: Sample preparation and analytical approaches.* Trends in Environmental Analytical Chemistry, 2016. **12**: p. 23-30.
33. Croué, J.-P., *Isolation of Humic and Non-Humic NOM Fractions: Structural Characterization.* Environmental Monitoring and Assessment, 2004. **92**(1): p. 193-207.
34. Green, N.W., et al., *Suwannee River Natural Organic Matter: Isolation of the 2R101N Reference Sample by Reverse Osmosis.* Environmental Engineering Science, 2014. **32**(1): p. 38-44.
35. Pressman, J.G., et al., *Disinfection byproduct formation in reverse-osmosis concentrated and lyophilized natural organic matter from a drinking water source.* Water Research, 2012. **46**(16): p. 5343-5354.
36. Song, H., et al., *Isolation and fractionation of natural organic matter: evaluation of reverse osmosis performance and impact of fractionation parameters.* Environmental Monitoring and Assessment, 2008. **153**(1): p. 307.
37. Alasonati, E., et al., *Characterization of the colloidal organic matter from the Amazonian basin by asymmetrical flow field-flow fractionation and size exclusion chromatography.* Water Research, 2010. **44**(1): p. 223-231.
38. Koprivnjak, J.F., et al., *Chemical and spectroscopic characterization of marine dissolved organic matter isolated using coupled reverse osmosis–electrodialysis.* Geochimica et Cosmochimica Acta, 2009. **73**(14): p. 4215-4231.
39. Serkiz, S.M. and E.M. Perdue, *Isolation of dissolved organic matter from the suwannee river using reverse osmosis.* Water Research, 1990. **24**(7): p. 911-916.
40. Sharp, M.M. and I.C. Escobar, *Effects of dynamic or secondary-layer coagulation on ultrafiltration.* Desalination, 2006. **188**(1): p. 239-249.
41. Vetter, T.A., et al., *Combining reverse osmosis and electrodialysis for more complete recovery of dissolved organic matter from seawater.* Separation and Purification Technology, 2007. **56**(3): p. 383-387.
42. McCurry Daniel, L., F. Speth Thomas, and G. Pressman Jonathan, *Lyophilization and Reconstitution of Reverse-Osmosis Concentrated Natural Organic Matter from a Drinking Water Source.* Journal of Environmental Engineering, 2012. **138**(4): p. 402-410.
43. Sandron, S., et al., *Chromatographic methods for the isolation, separation and characterisation of dissolved organic matter.* Environmental Sciences: Processes and Impacts, 2015. **17**(9): p. 1531-1567.
44. Lam, B. and A.J. Simpson, *Passive Sampler for Dissolved Organic Matter in Freshwater Environments.* Analytical Chemistry, 2006. **78**(24): p. 8194-8199.
45. Warner, D.L., D. Oviedo-Vargas, and T.V. Royer, *Evaluation of passive samplers for the collection of dissolved organic matter in streams.* Environmental Monitoring and Assessment, 2014. **187**(1): p. 4208.
46. Lamsal, R., et al., *Characterization and removal of natural organic matter by an integrated membrane system.* Desalination, 2012. **303**: p. 12-16.
47. Pelekani, C., et al., *Characterization of Natural Organic Matter Using High Performance Size Exclusion Chromatography.* Environmental Science & Technology, 1999. **33**(16): p. 2807-2813.
48. Rosario-Ortiz, F.L., S. Snyder, and I.H. Suffet, *Characterization of the Polarity of Natural Organic Matter under Ambient Conditions by the Polarity Rapid Assessment Method (PRAM).* Environmental Science & Technology, 2007. **41**(14): p. 4895-4900.
49. Zhou, Z. and L. Guo, *A critical evaluation of an asymmetrical flow field-flow fractionation system for colloidal size characterization of natural organic matter.* Journal of Chromatography A, 2015. **1399**: p. 53-64.

50. Malcolm, R.L. and P. MacCarthy, *Quantitative evaluation of XAD-8 and XAD-4 resins used in tandem for removing organic solutes from water.* Environment International, 1992. **18**(6): p. 597-607.
51. Barrett, S.E., S.W. Krasner, and G.L. Amy, *Natural Organic Matter and Disinfection By-Products*. ACS Symposium Series. Vol. 761. 2000: American Chemical Society. 448.
52. Leenheer, J.A., *Comprehensive approach to preparative isolation and fractionation of dissolved organic carbon from natural waters and wastewaters.* Environmental Science & Technology, 1981. **15**(5): p. 578-587.
53. Aiken, G.R., et al., *Isolation of hydrophilic organic acids from water using nonionic macroporous resins.* Organic Geochemistry, 1992. **18**(4): p. 567-573.
54. Leenheer, J.A. and J.-P. Croué, *Peer Reviewed: Characterizing Aquatic Dissolved Organic Matter.* Environmental Science & Technology, 2003. **37**(1): p. 18A-26A.
55. Chow, C.W.K., R. Fabris, and M. Drikas, *A rapid fractionation technique to characterise natural organic matter for the optimisation of water treatment processes.* Journal of Water Supply: Research and Technology - Aqua, 2004. **53**(2): p. 85.
56. Croue, J.-P., G. Korshin, and M. Benjamin, *Characterization of Natural Organic Matter in Drinking Water*. American Water Works Association, 2000.
57. Marhaba, T.F., Y. Pu, and K. Bengraine, *Modified dissolved organic matter fractionation technique for natural water.* Journal of Hazardous Materials, 2003. **101**(1): p. 43-53.
58. Avsar, E., et al., *Chemical characterization of natural organic matter and determination of disinfection by-product formation potentials in surface waters of Istanbul (Omerli and Buyukcekmece water dam), Turkey.* Fresenius Environmental Bulletin, 2014. **23**(2 A): p. 494-501.
59. Postigo, C., et al., *Characterization of iodinated disinfection by-products in chlorinated and chloraminated waters using Orbitrap based gas chromatography-mass spectrometry.* Analytical and Bioanalytical Chemistry, 2016. **408**(13): p. 3401-3411.
60. Penru, Y., et al., *Characterization of natural organic matter from Mediterranean coastal seawater.* Journal of Water Supply: Research and Technology - AQUA, 2013. **62**(1): p. 42-51.
61. *International Humic Substances Society*. 2018; Available from: http://humic-substances.org/isolation-of-ihss-aquatic-humic-and-fulvic-acids/.
62. Kaiser, E., et al., *Solid-State and Multidimensional Solution-State NMR of Solid Phase Extracted and Ultrafiltered Riverine Dissolved Organic Matter.* Environmental Science & Technology, 2003. **37**(13): p. 2929-2935.
63. Schwede-Thomas, S.B., et al., *Characterizing the properties of dissolved organic matter isolated by XAD and C-18 solid phase extraction and ultrafiltration.* Aquatic Sciences, 2005. **67**(1): p. 61-71.
64. Dittmar, T., et al., *A simple and efficient method for the solid - phase extraction of dissolved organic matter (SPE - DOM) from seawater.* Limnology and Oceanography: Methods, 2008. **6**(6): p. 230-235.
65. Kruger, B.R., B.J. Dalzell, and E.C. Minor, *Effect of organic matter source and salinity on dissolved organic matter isolation via ultrafiltration and solid phase extraction.* Aquatic Sciences, 2011. **73**(3): p. 405-417.
66. Kim, S., et al., *High resolution electrospray ionization mass spectrometry and 2D solution NMR for the analysis of DOM extracted by C18 solid phase disk.* Organic Geochemistry, 2003. **34**(9): p. 1325-1335.
67. Dalzell, B.J., E.C. Minor, and K.M. Mopper, *Photodegradation of estuarine dissolved organic matter: a multi-method assessment of DOM transformation.* Organic Geochemistry, 2009. **40**(2): p. 243-257.
68. Hedges, J.I., *Global biogeochemical cycles: progress and problems.* Marine Chemistry, 1992. **39**(1): p. 67-93.
69. Perminova, I.V., et al., *Molecular Mapping of Sorbent Selectivities with Respect to Isolation of Arctic Dissolved Organic Matter as Measured by Fourier Transform Mass Spectrometry.* Environmental Science & Technology, 2014. **48**(13): p. 7461-7468.
70. Ratpukdi, T., et al., *Rapid Fractionation of Natural Organic Matter in Water Using a Novel Solid-Phase Extraction Technique.* Water Environment Research, 2009. **81**(11): p. 2299-2308.
71. Gjessing, E.T., *Communication. Ultrafiltration of aquatic humus.* Environmental Science & Technology, 1970. **4**(5): p. 437-438.
72. Benner, R., et al., *Bulk chemical characteristics of dissolved organic matter in the ocean.* Science, 1992. **255**(5051): p. 1561-1564.
73. Shan, L., et al., *Natural organic matter fouling behaviors on superwetting nanofiltration membranes.* Water Research, 2016. **93**: p. 121-132.
74. Buesseler, K.O., et al., *An intercomparison of cross-flow filtration techniques used for sampling marine colloids: Overview and organic carbon results.* Marine Chemistry, 1996. **55**(1): p. 1-31.

75. Simjouw, J.P., E.C. Minor, and K. Mopper, *Isolation and characterization of estuarine dissolved organic matter: Comparison of ultrafiltration and C18 solid-phase extraction techniques.* Marine Chemistry, 2005. **96**(3-4): p. 219-235.
76. Chen, C., et al., *Applying polarity rapid assessment method and ultrafiltration to characterize NDMA precursors in wastewater effluents.* Water Research, 2014. **57**: p. 115-126.
77. Liao, X., et al., *Applying the polarity rapid assessment method to characterize nitrosamine precursors and to understand their removal by drinking water treatment processes.* Water Research, 2015. **87**: p. 292-298.
78. Paul-Dauphin, S., et al., *Probing Size Exclusion Mechanisms of Complex Hydrocarbon Mixtures: The Effect of Altering Eluent Compositions.* Energy & Fuels, 2007. **21**(6): p. 3484-3489.
79. Posner, A.M., *Importance of Electrolyte in the Determination of Molecular Weights by 'Sephadex' Gel Filtration, with Especial Reference to Humic Acid.* Nature, 1963. **198**: p. 1161.
80. Fukano, K., et al., *Evaluation of new supports for high-pressure aqueous gel permeation chromatography: TSK-GEL SW type columns.* Journal of Chromatography A, 1978. **166**(1): p. 47-54.
81. Her, N., et al., *Characterizing dissolved organic matter and evaluating associated nanofiltration membrane fouling.* Chemosphere, 2008. **70**(3): p. 495-502.
82. Huber, S.A., et al., *Characterisation of aquatic humic and non-humic matter with size-exclusion chromatography – organic carbon detection – organic nitrogen detection (LC-OCD-OND).* Water Research, 2011. **45**(2): p. 879-885.
83. Wagner, M., et al., *Characterization and quantification of humic substances 2D-Fluorescence by usage of extended size exclusion chromatography.* Water Research, 2016. **93**: p. 98-109.
84. Szabo, H.M. and T. Tuhkanen, *The application of HPLC–SEC for the simultaneous characterization of NOM and nitrate in well waters.* Chemosphere, 2010. **80**(7): p. 779-786.
85. Morelli, E., F. Puntoni, and A. Seritti, *Fluorescence characteristics of reverse phase HPLC fractionated marine organic matter.* Environmental Technology, 1993. **14**(10): p. 941-948.
86. Mills, G.L. and J.G. Quinn, *Isolation of dissolved organic matter and copper-organic complexes from estuarine waters using reverse-phase liquid chromatography.* Marine Chemistry, 1981. **10**(2): p. 93-102.
87. Lombardi, A.T., et al., *Reverse phase high performance liquid chromatographic fractionation of marine organic matter.* Environmental Technology, 1992. **13**(11): p. 1013-1021.
88. Woods, G.C., et al., *HILIC-NMR: Toward the Identification of Individual Molecular Components in Dissolved Organic Matter.* Environmental Science & Technology, 2011. **45**(9): p. 3880-3886.
89. Woods, G.C., M.J. Simpson, and A.J. Simpson, *Oxidized sterols as a significant component of dissolved organic matter: Evidence from 2D HPLC in combination with 2D and 3D NMR spectroscopy.* Water Research, 2012. **46**(10): p. 3398-3408.
90. Danielsson, L.G., *On the use of filters for distinguishing between dissolved and particulate fractions in natural waters.* Water Research, 1982. **16**(2): p. 179-182.
91. Lee, W., P. Westerhoff, and J.-P. Croué, *Dissolved Organic Nitrogen as a Precursor for Chloroform, Dichloroacetonitrile, N-Nitrosodimethylamine, and Trichloronitromethane.* Environmental Science & Technology, 2007. **41**(15): p. 5485-5490.
92. Chuang, Y.-H., et al., *The contribution of dissolved organic nitrogen and chloramines to nitrogenous disinfection byproduct formation from natural organic matter.* Water Research, 2013. **47**(3): p. 1308-1316.
93. Hansell, D.A., *Results and observations from the measurement of DOC and DON in seawater using a high-temperature catalytic oxidation technique.* Marine Chemistry, 1993. **41**(1-3): p. 195-202.
94. Hur, J., M.A. Williams, and M.A. Schlautman, *Evaluating spectroscopic and chromatographic techniques to resolve dissolved organic matter via end member mixing analysis.* Chemosphere, 2006. **63**(3): p. 387-402.
95. De Haan, H. and T. De Boer, *Applicability of light absorbance and fluorescence as measures of concentration and molecular size of dissolved organic carbon in humic Lake Tjeukemeer.* Water Research, 1987. **21**(6): p. 731-734.
96. Weishaar, J.L., et al., *Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon.* Environmental Science & Technology, 2003. **37**(20): p. 4702-4708.
97. Edzwald, J.K. and J.E. Tobiason, *Enhanced Coagulation: US Requirements and a Broader View.* Water Science and Technology, 1999. **40**(9): p. 63.
98. Ishii, S.K.L. and T.H. Boyer, *Behavior of Reoccurring PARAFAC Components in Fluorescent Dissolved Organic Matter in Natural and Engineered Systems: A Critical Review.* Environmental Science & Technology, 2012. **46**(4): p. 2006-2017.

99. Miano, T., G. Sposito, and J.P. Martin, *Fluorescence spectroscopy of model humic acid-type polymers.* Geoderma, 1990. **47**(3): p. 349-359.
100. Bieroza, M., A. Baker, and J. Bridgeman, *Relating freshwater organic matter fluorescence to organic carbon removal efficiency in drinking water treatment.* Science of The Total Environment, 2009. **407**(5): p. 1765-1774.
101. Coble, P.G., *Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy.* Marine Chemistry, 1996. **51**(4): p. 325-346.
102. Hudson, N., A. Baker, and D. Reynolds, *Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review.* River Research and Applications, 2007. **23**(6): p. 631-649.
103. Baghoth, S.A., et al., *Characterization of NOM in a drinking water treatment process train with no disinfectant residual.* Water Science and Technology: Water Supply, 2009. **9**(4): p. 379-386.
104. Baker, A., et al., *Relating dissolved organic matter fluorescence and functional properties.* Chemosphere, 2008. **73**(11): p. 1765-1772.
105. Chen, W., et al., *Fluorescence Excitation−Emission Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter.* Environmental Science & Technology, 2003. **37**(24): p. 5701-5710.
106. Persson, T. and M. Wedborg, *Multivariate evaluation of the fluorescence of aquatic organic matter.* Analytica Chimica Acta, 2001. **434**(2): p. 179-192.
107. Stedmon, C.A. and S. Markager, *Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis.* Limnology and Oceanography, 2005. **50**(2): p. 686-697.
108. Li, A., et al., *Polarity based fractionation of fulvic acids.* Chemosphere, 2009. **77**(10): p. 1419-1426.
109. Stevenson, F.J. and K.M. Goh, *Infrared spectra of humic acids and related substances.* Geochimica et Cosmochimica Acta, 1971. **35**(5): p. 471-483.
110. Hay, M.B. and S.C.B. Myneni, *Structural environments of carboxyl groups in natural organic molecules from terrestrial systems. Part 1: Infrared spectroscopy.* Geochimica et Cosmochimica Acta, 2007. **71**(14): p. 3518-3532.
111. Park, H., et al., *Characterization of natural organic matter treated by iron oxide nanoparticle incorporated ceramic membrane-ozonation process.* Water Research, 2012. **46**(18): p. 5861-5870.
112. Rodríguez Francisco, J. and A. Núñez Luis, *Characterization of aquatic humic substances.* Water and Environment Journal, 2011. **25**(2): p. 163-170.
113. Abdulla, H.A.N., et al., *Changes in the compound classes of dissolved organic matter along an estuarine transect: A study using FTIR and 13C NMR.* Geochimica et Cosmochimica Acta, 2010. **74**(13): p. 3815-3838.
114. Minor, E.C., et al., *Structural characterization of dissolved organic matter: A review of current techniques for isolation and analysis.* Environmental Sciences: Processes and Impacts, 2014. **16**(9): p. 2064-2079.
115. Cho, J., et al., *Characterization of clean and natural organic matter (NOM) fouled NF and UF membranes, and foulants characterization.* Desalination, 1998. **118**(1): p. 101-108.
116. McGuire Molly, M. and K. Herman Ellen, *A novel ATR - FTIR technique for identifying colloid composition in natural waters.* Hydrological Processes, 2014. **29**(6): p. 1314-1323.
117. Simpson, A.J., M.J. Simpson, and R. Soong, *Nuclear Magnetic Resonance Spectroscopy and Its Key Role in Environmental Research.* Environmental Science & Technology, 2012. **46**(21): p. 11488-11496.
118. Zheng, G. and W.S. Price, *Direct Hydrodynamic Radius Measurement on Dissolved Organic Matter in Natural Waters Using Diffusion NMR.* Environmental Science & Technology, 2012. **46**(3): p. 1675-1680.
119. Mao, J., et al., *Advanced Solid-State NMR Characterization of Marine Dissolved Organic Matter Isolated Using the Coupled Reverse Osmosis/Electrodialysis Method.* Environmental Science & Technology, 2012. **46**(11): p. 5806-5814.
120. Mopper, K., et al., *Advanced Instrumental Approaches for Characterization of Marine Dissolved Organic Matter: Extraction Techniques, Mass Spectrometry, and Nuclear Magnetic Resonance Spectroscopy.* Chemical Reviews, 2007. **107**(2): p. 419-442.
121. Hertkorn, N., et al., *High-field NMR spectroscopy and FTICR mass spectrometry: Powerful discovery tools for the molecular level characterization of marine dissolved organic matter.* Biogeosciences, 2013. **10**(3): p. 1583-1624.
122. Zhong, J., et al., *Combining advanced NMR techniques with ultrahigh resolution mass spectrometry: A new strategy for molecular scale characterization of macromolecular components of soil and sedimentary organic matter.* Organic Geochemistry, 2011. **42**(8): p. 903-916.

123. Hertkorn, N., et al., *Characterization of a major refractory component of marine dissolved organic matter.* Geochimica et Cosmochimica Acta, 2006. **70**(12): p. 2990-3010.
124. Thorn, K.A. and L.G. Cox, *N-15 NMR spectra of naturally abundant nitrogen in soil and aquatic natural organic matter samples of the International Humic Substances Society.* Organic Geochemistry, 2009. **40**(4): p. 484-499.
125. Woods, G.C., et al., *Online High-Performance Size Exclusion Chromatography−Nuclear Magnetic Resonance for the Characterization of Dissolved Organic Matter.* Environmental Science & Technology, 2010. **44**(2): p. 624-630.
126. Lam, B. and A.J. Simpson, *Direct 1H NMR spectroscopy of dissolved organic matter in natural waters.* Analyst, 2008. **133**(2): p. 263-269.
127. Chen, J., et al., *Spectroscopic characterization of the structural and functional properties of natural organic matter fractions.* Chemosphere, 2002. **48**(1): p. 59-68.
128. Cook, R.L., *Coupling NMR to NOM.* Analytical and Bioanalytical Chemistry, 2004. **378**(6): p. 1484-1503.
129. Sleighter, R.L., et al., *Multivariate Statistical Approaches for the Characterization of Dissolved Organic Matter Analyzed by Ultrahigh Resolution Mass Spectrometry.* Environmental Science & Technology, 2010. **44**(19): p. 7576-7582.
130. Sleighter Rachel, L. and G. Hatcher Patrick, *The application of electrospray ionization coupled to ultrahigh resolution mass spectrometry for the molecular characterization of natural organic matter.* Journal of Mass Spectrometry, 2007. **42**(5): p. 559-574.
131. Kim, S., R.W. Kramer, and P.G. Hatcher, *Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram.* Analytical Chemistry, 2003. **75**(20): p. 5336-5344.
132. Koch, B.P., et al., *Fundamentals of Molecular Formula Assignment to Ultrahigh Resolution Mass Data of Natural Organic Matter.* Analytical Chemistry, 2007. **79**(4): p. 1758-1763.
133. Kujawinski, E.B., et al., *The application of electrospray ionization mass spectrometry (ESI MS) to the structural characterization of natural organic matter.* Organic Geochemistry, 2002. **33**(3): p. 171-180.
134. Liu, Z., et al., *The chemical changes of DOM from black waters to coastal marine waters by HPLC combined with ultrahigh resolution mass spectrometry.* Estuarine, Coastal and Shelf Science, 2011. **92**(2): p. 205-216.
135. Koch, B.P., et al., *Advanced characterization of marine dissolved organic matter by combining reversed-phase liquid chromatography and FT-ICR-MS.* Marine Chemistry, 2008. **111**(3): p. 233-241.
136. Reemtsma, T., et al., *Differences in the molecular composition of fulvic acid size fractions detected by size-exclusion chromatography–on line Fourier transform ion cyclotron resonance (FTICR–) mass spectrometry.* Water Research, 2008. **42**(1): p. 63-72.
137. Patriarca, C., et al., *Online HPLC-ESI-HRMS Method for the Analysis and Comparison of Different Dissolved Organic Matter Samples.* Environmental Science & Technology, 2018. **52**(4): p. 2091-2099.
138. Petras, D., et al., *High-Resolution Liquid Chromatography Tandem Mass Spectrometry Enables Large Scale Molecular Characterization of Dissolved Organic Matter.* Frontiers in Marine Science, 2017. **4**: p. 405.
139. Verkh, Y., M. Rozman, and M. Petrovic, *A non-targeted high-resolution mass spectrometry data analysis of dissolved organic matter in wastewater treatment.* Chemosphere, 2018. **200**: p. 397-404.
140. Kujawinski, E.B., et al., *Probing molecular-level transformations of dissolved organic matter: insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry.* Marine Chemistry, 2004. **92**(1): p. 23-37.
141. Ohno, T., R.L. Sleighter, and P.G. Hatcher, *Comparative study of organic matter chemical characterization using negative and positive mode electrospray ionization ultrahigh-resolution mass spectrometry.* Analytical and Bioanalytical Chemistry, 2016. **408**(10): p. 2497-2504.
142. Hertkorn, N., et al., *Natural organic matter and the event horizon of mass spectrometry.* Analytical Chemistry, 2008. **80**(23): p. 8908-8919.
143. Hockaday William, C., et al., *Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment.* Limnology and Oceanography: Methods, 2009. **7**(1): p. 81-95.
144. D'Andrilli, J., et al., *Comprehensive characterization of marine dissolved organic matter by Fourier transform ion cyclotron resonance mass spectrometry with electrospray and atmospheric pressure photoionization.* Rapid Communications in Mass Spectrometry, 2010. **24**(5): p. 643-650.
145. Van Krevelen, D.W., *Graphical-Statistical Method for the Study of Structure and Reaction Processes of Coal.* Fuel, 1950. **29**: p. 269-284.

146. Reemtsma, T., *Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry: Status and needs*. Journal of Chromatography A, 2009. **1216**(18): p. 3687-3701.
147. Makarov, A., et al., *Performance Evaluation of a Hybrid Linear Ion Trap/Orbitrap Mass Spectrometer*. Analytical Chemistry, 2006. **78**(7): p. 2113-2120.
148. Hawkes, J.A., et al., *Evaluation of the Orbitrap Mass Spectrometer for the Molecular Fingerprinting Analysis of Natural Dissolved Organic Matter*. Analytical Chemistry, 2016. **88**(15): p. 7698-7704.
149. Urai, M., et al., *Molecular characterization of dissolved organic matter in various urban water resources using Orbitrap Fourier transform mass spectrometry*. Water Science and Technology: Water Supply, 2014. **14**(4): p. 547-553.
150. Phungsai, P., et al., *Molecular characterization of low molecular weight dissolved organic matter in water reclamation processes using Orbitrap mass spectrometry*. Water Research, 2016. **100**: p. 526-536.
151. Allpike, B.P., et al., *A new organic carbon detector for size exclusion chromatography*. Journal of Chromatography A, 2007. **1157**(1): p. 472-476.
152. Wu, F.C., et al., *Rapid quantification of humic and fulvic acids by HPLC in natural waters*. Applied Geochemistry, 2007. **22**(8): p. 1598-1605.
153. Awad, J., et al., *Characterization of dissolved organic matter for prediction of trihalomethane formation potential in surface and sub-surface waters*. Journal of Hazardous Materials, 2016. **308**: p. 430-439.
154. Chow, C.W.K., et al., *Assessing Natural Organic Matter Treatability Using High Performance Size Exclusion Chromatography*. Environmental Science & Technology, 2008. **42**(17): p. 6683-6689.
155. Axt, G., *Ein Gerat zur kontinuierlichen Messung des organischen Kohlenstoffes in Trink- und Abwasserbereichen*. Vom Wasser, 1969. **36**: p. 328-339.
156. Khundzhua, D.A., et al., *An analysis of dissolved organic matter from freshwater Karelian Lakes using reversed-phase high-performance liquid chromatography with online absorbance and fluorescence analysis.* Moscow University Physics Bulletin, 2017. **72**(1): p. 68-75.
157. Trubetskaya, O.E., et al., *RP-HPLC and spectroscopic characterization of Suwannee River water NOM after concentrated urea treatment and dialysis.* Desalination and Water Treatment, 2016. **57**(12): p. 5358-5364.
158. Wu, F.C., R.D. Evans, and P.J. Dillon, *Separation and Characterization of NOM by High-Performance Liquid Chromatography and On-Line Three-Dimensional Excitation Emission Matrix Fluorescence Detection*. Environmental Science & Technology, 2003. **37**(16): p. 3687-3693.
159. Mawhinney, D.B., et al., *Characterization of fulvic acids by liquid chromatography-quadrupole time-of-flight mass spectrometry*. Journal of Chromatography A, 2009. **1216**(9): p. 1319-1324.
160. Cortés-Francisco, N. and J. Caixach, *Molecular Characterization of Dissolved Organic Matter through a Desalination Process by High Resolution Mass Spectrometry*. Environmental Science & Technology, 2013. **47**(17): p. 9619-9627.
161. Pluskal, T., et al., *MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data*. BMC Bioinformatics, 2010. **11**(1): p. 395.
162. Alpert, A.J., *Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds*. Journal of Chromatography A, 1990. **499**: p. 177-196.
163. Duarte, R.M.B.O., A.C. Barros, and A.C. Duarte, *Resolving the chemical heterogeneity of natural organic matter: New insights from comprehensive two-dimensional liquid chromatography*. Journal of Chromatography A, 2012. **1249**: p. 138-146.
164. Peuravuori, J. and K. Pihlaja, *Advanced TMAH and TMAAc thermochemolysis–pyrolysis techniques for molecular characterization of size-separated fractions from aquatic dissolved organic matter*. Analytical and Bioanalytical Chemistry, 2007. **389**(2): p. 475-491.
165. van Heemst, J.D.H., et al., *Characterization of Estuarine and Fluvial Dissolved Organic Matter by Thermochemolysis using Tetramethylammonium Hydroxide*. Acta hydrochimica et hydrobiologica, 2000. **28**(2): p. 69-76.
166. Frazier, S.W., et al., *Characterization of organic matter from natural waters using tetramethylammonium hydroxide thermochemolysis GC-MS*. Journal of Analytical and Applied Pyrolysis, 2003. **70**(1): p. 99-128.
167. Aluwihare, L.I., D.J. Repeta, and R.F. Chen, *Chemical composition and cycling of dissolved organic matter in the Mid-Atlantic Bight*. Deep Sea Research Part II: Topical Studies in Oceanography, 2002. **49**(20): p. 4421-4437.
168. Maie, N., et al., *Chemical characteristics of dissolved organic matter in an oligotrophic subtropical wetland/estuarine*. Limnology and Oceanography, 2005. **50**(1): p. 23-35.

169. Ohlenbusch, G., M.U. Kumke, and F.H. Frimmel, *Sorption of phenols to dissolved organic matter investigated by solid phase microextraction.* Science of The Total Environment, 2000. **253**(1): p. 63-74.
170. Kracht, O. and G. Gleixner, *Isotope analysis of pyrolysis products from Sphagnum peat and dissolved organic matter from bog water.* Organic Geochemistry, 2000. **31**(7): p. 645-654.
171. Parsi, Z., et al., *Analytical pyrolysis as a tool for the characterization of natural organic matter—A comparison of different approaches.* Journal of Analytical and Applied Pyrolysis, 2007. **79**(1): p. 9-15.
172. Greenwood, P.F., et al., *Bacterial biomarkers thermally released from dissolved organic matter.* Organic Geochemistry, 2006. **37**(5): p. 597-609.
173. Berwick, L., P.F. Greenwood, and R.J. Smernik, *The use of MSSV pyrolysis to assist the molecular characterisation of aquatic natural organic matter.* Water Research, 2010. **44**(10): p. 3039-3054.
174. Schulten, H.R., *Analytical pyrolysis and computational chemistry of aquatic humic substances and dissolved organic matter.* Journal of Analytical and Applied Pyrolysis, 1999. **49**(1): p. 385-415.
175. Giddings, J.C., F.J.F. Yang, and M.N. Myers, *Flow field-flow fractionation: A versatile new separation method.* Science, 1976. **193**(4259): p. 1244-1245.
176. Beckett, R., Z. Jue, and J.C. Giddings, *Determination of molecular weight distributions of fulvic and humic acids using flow field-flow fractionation.* Environmental Science & Technology, 1987. **21**(3): p. 289-295.
177. Lyvén, B., et al., *Optimisation of on-channel preconcentration in flow field-flow fractionation for the determination of size distributions of low molecular weight colloidal material in natural waters.* Analytica Chimica Acta, 1997. **357**(3): p. 187-196.
178. Stolpe, B., et al., *Size and composition of colloidal organic matter and trace elements in the Mississippi River, Pearl River and the northern Gulf of Mexico, as characterized by flow field-flow fractionation.* Marine Chemistry, 2010. **118**(3): p. 119-128.
179. Zanardi-Lamardo, E., C.A. Moore, and R.G. Zika, *Seasonal variation in molecular mass and optical properties of chromophoric dissolved organic material in coastal waters of southwest Florida.* Marine Chemistry, 2004. **89**(1): p. 37-54.
180. Assemi, S., et al., *Characterization of natural organic matter fractions separated by ultrafiltration using flow field-flow fractionation.* Water Research, 2004. **38**(6): p. 1467-1476.
181. Messaud, F.A., et al., *An overview on field-flow fractionation techniques and their applications in the separation and characterization of polymers.* Progress in Polymer Science, 2009. **34**(4): p. 351-368.
182. Vughs, D., et al., *Application of effect-directed analysis to identify mutagenic nitrogenous disinfection by-products of advanced oxidation drinking water treatment.* Environmental Science and Pollution Research, 2016.
183. Kind, T. and O. Fiehn, *Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry.* BMC Bioinformatics, 2007. **8**.

UNIVERSITEIT VAN AMSTERDAM

MSc in Chemistry
Analytical Sciences

Master Thesis

Aromatic amino acids as a source for nitrogen containing by-products formed by advanced oxidation water treatment

---

by

Dennis Vughs
UvA: 10771506

March 2018

42 ECTS
April 2016 – March 2018

Supervisor/Examiner:
Prof. Dr. W.P. de Voogt
Supervisor:
Dr. A. Kolkman

Examiner:
Dr. J.R. Parsons

KWR Watercycle Research Institute

Acknowledgements

First of all I would like to express my sincere gratitude to my supervisors, Dr. Annemieke Kolkman and Prof. Dr. Pim de Voogt for their guidance during this thesis. I also would like to thank Dr. Bram Martijn for his input and expertise on disinfection by-products and assistance involving the aromatic amino acids labeling experiments. Furthermore, I would like to thank Dr. Kirsten Baken for performing the toxicological evaluation. And finally I would like to thank all my colleagues from the LMC and CWG teams of KWR for their assistance and for creating such a nice environment to work in.

Abstract

For the preparation of drinking water, advanced oxidation processes such as, UV/H_2O_2 and ozone are becoming increasingly important for disinfection and for the removal of organic micro pollutants which are present in source water. In earlier research it was shown that medium pressure UV treatment of artificial water containing natural organic matter and nitrate may lead to the formation of nitrogen containing by-products. A total of 84 N-DBPs were detected in artificial water of which 22 were detected in a full-scale drinking water facility using MP UV/H_2O_2. It was shown that the chemical response detected by a suspect screening was comparable with the response obtained with Ames fluctuation assay using strains TA98 and TA100. This implies that the 22 N-DBPs are possibly responsible for the positive response in the Ames fluctuation test, but the genotoxic potential of the identified by-products does not explain the observed Ames response. Of the 84 detected N-DBPs only 14 were identified, and without the identity and any information about the toxic potency of these N-DBPs, it is not possible to perform substance-specific health risk assessment.

In the present study a different strategy for the detection of N-DBPs was pursued by using model compounds and applying MP UV treatment. It is known that water containing aromatic amino acids (i.e. tyrosine, phenylalanine and tryptophan) and nitrate or nitrite becomes mutagenic after UV irradiation, which can be a potential generator of genotoxic N-DBPs. Labeling experiments were conducted involving stable isotopically labeled nitrated and aromatic amino acids and MP UV treatment. Many by-products were detected using this strategy, of which only a few were detected in the artificial water sample of the prior study. For tryptophan one N-DBP was linked to the N-DBPs in artificial water (i.e. containing NOM and nitrate) and was identified as 3-nitroindole. However, the results suggest that 3-nitroindole detected in artificial water does not originate from tryptophan, but from another source. For phenylalanine four unidentified N-DBPs (also isomers) were linked to the N-DBPs in artificial water. The results indicate that these four N-DBPs really originate from phenylalanine in artificial water and thus demonstrate that aromatic amino acids are a source for some of the N-DBPs formed by MP UV treatment of artificial water.

Five N-DBPs were unambiguously identified during this study, bringing the total of identified N-DBPs in artificial water to 19. The toxicity evaluation of the newly identified N-DBPs based on prediction of structural characteristics indicated potential mutagenicity, although it remains uncertain if these N-DBPS will give a positive response in the Ames test.

Furthermore a LC-QTOF target method was developed for determination of the 19 identified N-DBPs in drinking- and surface water resulting in satisfactory LODs and LOQs. The method developed was applied in a seven months monitoring study of a full-scale drinking water facility using MP UV treatment. Fifteen of the 19 N-DBPs were detected in this monitoring study, with concentrations ranging from 1.0 to 44 ng/L. Five N-DBPs were detected at a relatively high concentration in MP UV treated water. The total concentration of detected N-DBPs in a full-scale treatment facility demonstrates the relevance of the N-DBPs identified and the N-DBP target method developed.

Contents

1	**Introduction**	**5**
1.1	Disinfection by-products	5
1.2	Objective MSc thesis	6
1.3	Structure of thesis	7
2	**Development and validation of a LC-QToF method for the determination of aromatic amino acids in water**	**8**
2.1	Amino acids analysis	8
2.2	Material & methods	9
2.3	Results method development and optimisation	10
2.4	Method validation	11
2.5	Source and artificial water analysis	12
2.6	Summary/conclusion	13
3	**Labeling experiments**	**14**
3.1	Stable isotope labeling strategy	14
3.2	Tryptophan labeling experiments	14
3.3	Tyrosine and phenylalanine labeling experiments	23
3.4	SPE experiments of tryptophan and phenylalanine by-products	29
3.5	Conclusion labeling experiments	31
4	**Identification, toxicological evaluation and quantification of N-DBPs**	**32**
4.1	Identification of N-DBPs	32
4.2	Toxicological evaluation of the identified N-DBPs	33
4.3	Development of a LC-QToF target method for N-DBPS in water	34
4.4	Method validation in ultrapure water	35
4.5	Monitoring study full-scale drinking water treatment facility	37
4.6	Results monitoring study	37
4.7	N-DBP concentration in extracts of previous study	40
5	**Conclusion & recommendations**	**42**
6	**References**	**44**

1 Introduction

1.1 Disinfection by-products

For the production of drinking water, surface water is gaining importance due to increasing populations and limited availability of groundwater. However, surface waters may contain a large variety of organic micropollutants, such as pharmaceuticals, pesticides and industrial compounds, for which the traditional water treatment systems (e.g. rapid sand filtration, coagulation, granular activated carbon filtration) do not constitute a robust barrier [1]. Advanced oxidation processes (AOP), such as UV/H_2O_2 and ozone are becoming increasingly important for the preparation of drinking water, for effectively removing these micropollutants [2-4]. However, it is known that advanced oxidation processes can produce potentially harmful disinfection by-products (DBPs) [5-8]. Many studies were performed to investigate the formation, identities and occurrence of these DBPs [9]. Such studies are challenging, due to the vast amount of DBPs that can be formed. Many factors can contribute to the formation of DBPs such as, disinfection method, type of source water and process conditions. Information on the potential human health risk of these DBPs is often unknown because of their unknown identity and/or the lack of toxicity data. A relatively new subgroup of DBPs are the nitrogenous DBPs (N-DBPs) [10-12], which have a higher human toxicological potential than the well-known chlorinated DBPs [9, 10, 12].

Earlier research has shown that medium pressure (MP) UV/H_2O_2 treatment in drinking water production may lead to the formation of N-DBPs [7, 13, 14]. These N-DBPs are formed through a complex mechanism of nitrate photolysis by UV in which nitrate is converted into the stable nitrite [15, 16]. During this nitrate to nitrite reduction, various nitrate intermediate radicals are formed [16], which have the ability to react with natural organic matter (NOM) that is present in source water. This ultimately results in the incorporation of the nitrogen-atom of nitrate into aquatic NOM [14].

In previous research, an innovative approach was developed in order to trace N-DBPs, combining stable isotope labeled nitrate with high-resolution mass spectrometry (HRMS) [13]. It was shown that multiple N-DBPs were formed after MP UV treatment of artificial water containing nitrate and NOM. Using this approach a total of 84 N-DBPs were detected in artificial water. A suspect screening for these 84 N-DBPs in water samples from a full-scale drinking water facility using MP UV/H_2O_2, resulted in the detection of 22 N-DBPs. The Ames mutagenicity test, a way to determine genotoxicity of (treated) water [5, 6, 17], was also performed. It was shown that chemical response detected by the suspect screening was comparable with the response obtained with Ames fluctuation assay using Salmonella strains TA98 and TA100. This implies that some of the 22 N-DBPs are possibly responsible for the positive response in the Ames fluctuation test. The genotoxic effect was shown to be effectively removed from treated drinking water with granular activated carbon (GAC) filtration and/or dune infiltration [5, 7].

Without the identity and any information about the toxic potency of N-DBPs, it is not possible to perform substance-specific health risk assessment. Therefore, it is important to identify these N-DBPs and to investigate what their mutagenic response is in the Ames test. In our follow-up study this was investigated by applying a fractionation method to MP UV treated water containing nitrate and NOM. Next, the different fractions were analysed by mutagenicity testing and chemical suspect screening [18]. This showed that the presence of

N-DBPs and mutagenicity in the Ames fluctuation test were correlated. Five potentially genotoxic by-products, with relatively high concentrations, were linked to fractions in which mutagenicity was observed. Of the 84 known N-DBPs formed by MP UV treatment, 14 by-products were unambiguously identified [13, 18]. However, the genotoxic potential of the identified by-products does not explain the observed Ames response, and the subject for the present research project is to further identify products that may explain the observed genotoxicity.

1.2 Objective MSc thesis

A different strategy for detecting by-products and assessment of their mutagenic response, is to use model compounds and apply MP UV treatment. Literature has shown that a possible source for the by-products, besides NOM, could be aromatic amino acids, which are present in NOM and surface water [19, 20]. Suzuki et al. [21] showed that aromatic amino acids (i.e. tryptophan, phenylalanine and tyrosine) become mutagenic after UV irradiation in water containing nitrate and nitrite. For the three aromatic amino acids the highest Ames response was observed for tryptophan, using the salmonella strain TA98 with and without S9 mix. Furthermore Aljammaz [22] showed that in water containing aromatic amino acids the concentration of inorganic nitrogen (i.e. nitrate, nitrite and ammonia) is substantially decreased after MP UV treatment, which indicates that at least a part of inorganic nitrogen is converted to organic nitrogen by nitration of the aromatic amino acids. Once again water containing tryptophan showed the largest nitrogen gap, indicating that tryptophan is most susceptible for nitration and thus by-product formation. Because aromatic amino acids are expected to be present in source water (i.e. surface water) [19, 20] for MP UV drinking water treatment, the role of aromatic amino acids in by-products formation needs to be further investigated.

In order to explain the observed genotoxic response in MP UV treated water and to perform substance-specific health risk assessment, the identities of these N-DBPs needs to be known. Based on evidence by Suzuki and Aljammaz, the following hypothesis was made: aromatic amino acids are a source for the formation of some of the genotoxic N-DBPs formed by MP UV water treatment.

The first part of the present MSc thesis therefore aims to: (i) investigate the role of aromatic amino acids in N-DBPs formation, by irradiation of aromatic amino acids under MP UV conditions, in combination with stable isotope labeling and high resolution mass spectrometry; and (ii) identification of N-DBPs formed from aromatic amino acids and further identification of relevant N-DBPs formed during full-scale MP UV water treatment. To address the first goal, the labeling strategy developed by Kolkman et al. [13] using ^{14}N and ^{15}N nitrate was used, and was expanded by also using labeled and unlabeled aromatic amino acids, for the detection of N-DBPs and for obtaining structural information.

The second part of the MSc thesis aims to: (i) perform a toxicological evaluation of newly identified N-DBPs in the present study; and (ii) conduct and evaluate a seven month monitoring study for identified by-products in order to determine the relevance of N-DBPs in a full-scale drinking water treatment facility using MP UV treatment.

1.3 Structure of thesis

In the first part of the present thesis, the role of aromatic amino acids in N-DBPs formation is investigated. Therefore the presence of aromatic amino acids in in source water (Lake IJssel) and artificial water (i.e. NOM and nitrate dissolved in ultrapure water) has to be demonstrated first. Therefore an analytical method is developed in chapter 2, using liquid chromatography (LC) coupled to a high resolution quadrupole time of flight mass spectrometer (QToF), for the quantitative determination of aromatic amino acids in source and artificial water. In chapter 3 the developed LC-QToF method is optimised for stable isotope labeling experiments involving aromatic amino acids. Experiments are conducted using tyrosine, phenylalanine, tryptophan, labeled and unlabeled nitrate, labeled tryptophan and MP UV treatment. Subsequently, the artificial water and full-scale water treatment facility sample of the prior study [13] in which 84 and 22 N-DBPs were detected, is screened for aromatic amino acids N-DBPs (formed with labeling experiments) in order to determine if some of unidentified N-DBPs found in the prior study could originate from aromatic amino acids and be identified. In chapter 4, more of the 84 previously detected N-DBPs are identified and a toxicological evaluation is performed on the identified N-DBPs. Finally, an analytical target method (LC-QToF) is developed and validated for all identified N-DBPs in the current and prior study, in order to perform and evaluate a seven month monitoring study for N-DBPs in a full-scale drinking water treatment facility using MP UV treatment. Finally, in chapter 5 the conclusions of this thesis are presented and recommendations are discussed.

2 Development and validation of a LC-QToF method for the determination of aromatic amino acids in water

Only if aromatic amino acids are truly present in source water (Lake IJssel) and artificial water (i.e. NOM and nitrate dissolved in ultrapure water), N-DBPs originating from aromatic amino acids can be formed and labeling experiments are meaningful. Therefore, first an analytical method has to be developed for the determination of free dissolved aromatic amino acids (i.e. tyrosine, phenylalanine and tryptophan) in water.

This chapter describes the analytical method, method development, validation and sample analysis of free dissolved aromatic amino acids in water.

2.1 Amino acids analysis

Amino acids present in surface water play an important role in the biogeochemistry of nitrogen and carbon [23], and are therefore studied extensively. They can be analysed with a large variety of analytical techniques (e.g. HPLC, GC-MS, CE, IC) in many different types of matrices. Due to their hydrophilicity and zwitterionic nature, analysis of amino acids can be challenging. Derivatization techniques are therefore widely used to improve detection and chromatographic separation in biological and environmental matrices. But derivatization techniques have some major drawback such as, instable derivatives, low derivatives yield and being labour intensive. Therefore direct analysis techniques without derivatization are becoming more popular using analytical methodologies such as, CE-MS [24], HPAEC-PAD [25], ion-pair chromatography (LC) coupled to MS [26], and HILIC-MS [27]. For the analysis of the total amount of amino acids in a sample (i.e. bound species and biopolymer), often an acid or alkaline hydrolysis is employed before chromatographic separation. Since it was shown that free dissolved amino acids are a potential source for N-DBP formation [11, 21, 22], hydrolysis is not needed.

Because this study focuses on the moderately polar aromatic amino acids and their by-products, sufficient retention and detection is expected using a regular reversed phase C18 method. As a starting point for the analytical method development of free dissolved tyrosine (Tyr), phenylalanine (Phe) and tryptophan (Trp) (see Figure 1 for structures) in water, the non-target high resolution screening method employed in the previous studies was used [13, 18]. A high resolution mass spectrometer was used because of its capabilities for detecting unknown aromatic amino acids by-products, which is needed for the labeling experiments in chapter 3.

TYROSINE PHENYLALANINE TRYPTOPHAN

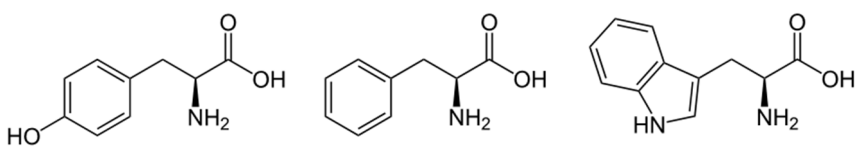


Figure 1: Structures of tyrosine, phenylalanine and tryptophan

2.2 Material & methods

In this section, the final optimized analytical method and sample pre-treatment is described for the analysis of tyrosine, phenylalanine and tryptophan in water.

2.2.1 Chemicals

All solvents used were of analytical grade quality. Acetonitrile and methanol (ultra gradient HPLC grade) were obtained from Avantor Performance Materials B.V. (Deventer, the Netherlands). Formic acid (HPLC quality) was purchased from Sigma-Aldrich (Steinheim, Germany). The aromatic amino acids reference standards; L-tyrosine, L-phenylalanine and L-tryptophan were acquired from Sigma-Aldrich. The isotopically labeled internal standards; L-tyrosine-d4, L-phenylalanine-d5 and L-tryptophan-d5 were purchased from Toronto Research Chemicals (Toronto, Canada). In table 1, the CAS number, formula, accurate mass of the protonated molecule ($[M+H]^+$) and Log D are shown for the aromatic amino acids and their corresponding internal standards. Ultrapure water was obtained by purifying demineralized water in an Elga Purelab Chorus ultrapure water system. (High Wycombe, United Kingdom). Stock solutions of the reference and internal standards were prepared in methanol and ultrapure water (20/80% v/v) at a concentration of 100 and 50 mg/L, respectively. Stock solutions were stored at -25° C. Standards were prepared from the stock solutions by dilution in ultrapure water, and were prepared shortly before analysis.

TABLE 1: CAS NUMBER, FORMULA, ACCURATE MASS PROTONATED MOLECULE AND LOG D OF AROMATIC AMINO ACIDS AND INTERNAL STANDARDS

Name	CAS number	Formula	Accurate mass $[M+H]^+$	Log D* (pH 4)
L-tyrosine	60-18-4	$C_9H_{11}NO_3$	182.0812	-1.49
L-phenylalanine	63-91-2	$C_9H_{11}NO_2$	166.0863	-1.20
L-tryptophan	73-22-3	$C_{11}H_{12}N_2O_2$	205.0972	-1.10
Internal standards				
L-tyrosine-d4	62595-14-6	$C_9H_7D_4NO_3$	186.1063	n.d.
L-phenylalanine-d5	56253-90-8	$C_9H_6D_5NO_2$	171.1176	n.d.
L-tryptophan-d5	62595-11-3	$C_{11}H_7D_5N_2O_2$	210.1285	n.d.

n.d. not determined

** calculated using ChemAxon*

2.2.2 Sample pre-treatment

Fifty mL of water sample was transferred into a 50 mL flask, to which the internal standards were added (2.0 µg/L). After homogenization the samples were filtered using a 0.2 µm Phenomenex Phenex regenerated cellulose filter (Utrecht, Netherlands) and were transferred to an autosampler vial for LC-QToF analysis.

2.2.3 LC-QToF analysis

The LC system consisted of a LC-30AD binary gradient pump, SIL-30AC auto sampler and a CTO-20AC column oven (Shimadzu Corporation, Kyoto, Japan). The chromatographic separation was achieved using a Xbridge BEH C18 XP (2.1 x 100 mm, 2.5 µm, Waters, Milford, MA, USA) preceded by a Phenomenex SecurityGuard Ultra column (Phenomenex, Torrance, USA) at a temperature of 25° C. The mobile phase consisted out of solvent A; ultrapure water with 0.05% formic acid (v/v) and solvent B; acetonitrile with 0.05% formic acid (v/v). The gradient elution started with 4% B and was held constant for 1 minute, and was then followed by a linear gradient to 100% B in 7 min, and was held constant at 100% B for 4 min. Then the mobile phase was returned to initial gradient conditions in 0.5 min and was held

for 4.5 min. The mobile phase flow rate was 0.3 mL/min and the injection volume was set to 50 µL.

Detection was performed using an AB Sciex TripleTOF 5600+ high resolution QToF mass spectrometer (AB Sciex, Concord, Canada) operated in positive electrospray (ESI) mode with a DuoSpray ion source. External mass calibration was automatically performed after thirty consecutive samples by a calibration delivery system (AB Sciex) using the APCI probe of the DuoSpray ion source. The source conditions were as follows: ion spray voltage, 5.0 kV; ion source gas 1 and 2 at 40 and 50 psi, respectively; curtain gas, 25 psi; temperature, 500 °C and declustering potential, 70 V. Full scan accurate MS and MS/MS mass spectra were acquired from 100 to 800 Da with a resolving power of 30,000 FWHM (at m/z 400). In order to unambiguously confirm the identities of tyrosine, phenylalanine and tryptophan, MS/MS spectra were recorded with a collision energy of 35 eV and collision energy spread (CES) of 15 eV. The recording of MS/MS spectra of the analytes, specified in the mass list, was continuously acquired (no threshold) from 40 to 300 Da. Data acquisition and processing were performed using Analyst TF 1.6 and Multiquant 3.0 software (AB Sciex).

2.3 Results method development and optimisation

For the method development of free dissolved aromatic amino acids in water, the LC-QToF non-target screening method employed in the previous study [18] was used as starting point . Due to the hydrophilicity of the aromatic amino acids (see table 1 for Log D values), sample pre-treatment using the solid phase extraction method described in the previous study is not possible. The recovery for these aromatic amino acids would just be too low. Therefore the decision was made, to use a direct injection approach in which water samples are directly injected onto the column, in order to minimize the loss of aromatic amino acids. The same approach is used for the labeling experiments described in chapter 3, also to minimize the loss of N-DBPs during sample pre-treatment.

A reversed phase Xbridge BEH C18 XP analytical column was used for the method development. And for the mobile phase a combination of ultrapure water (A) and acetonitrile (B) with formic acid as modifier (0.05% v/v) was used. The initial gradient of the non-target screening method started with 5% B. First the injection volume was optimised. In order to obtain the most sensitivity, a large injection volume (for a 2.1 mm column) of 100 µL was tested first. This resulted in a broad peak for tyrosine (most polar) and also a moderately broad peak for phenylalanine. Satisfactory peak shapes were obtained for all aromatic amino acids using a 50 µL injection. The gradient was then further optimised by lowering the initial gradient to 4% B, and by holding the gradient for 1 min, improving retention for tyrosine. Furthermore the linear gradient was shortened from 40 min (100% B) to 7 min (100% B), improving total analysis time from 52 min to 17 min.

Since mass spectrometric analysis of amino acids can be performed in the positive or negative mode using electrospray ionisation, first a comparison was made between both ionisation modes. It was determined that the sensitivity was improved substantially (> 2x) in the positive mode. Therefore the mass spectrometric detection was performed by the detection of the protonated molecular ion ($[M+H]^+$) using an extracted ion chromatogram window of 10 ppm. In order to improve selectivity and sensitivity, continuous MS2 spectra recording of the aromatic amino acids was added to the QToF acquisition method. The most intense fragments per compound were selected for quantification purposes. This resulted in improved selectivity and thus lower detection limits for all aromatic amino acids. See figure 2 for a comparison between the extracted ion chromatogram (EIC) of the protonated molecular ion ($[M+H]^+$) and EIC of the most intense fragment per compound.

Fragment ions were acquired with a collision energy of 20, 35 and 50 eV, which was automatically averaged (CES) to obtain MS2 spectra with many fragment ions (see attachment I for MS2 spectra of tyrosine, phenylalanine and tryptophan).

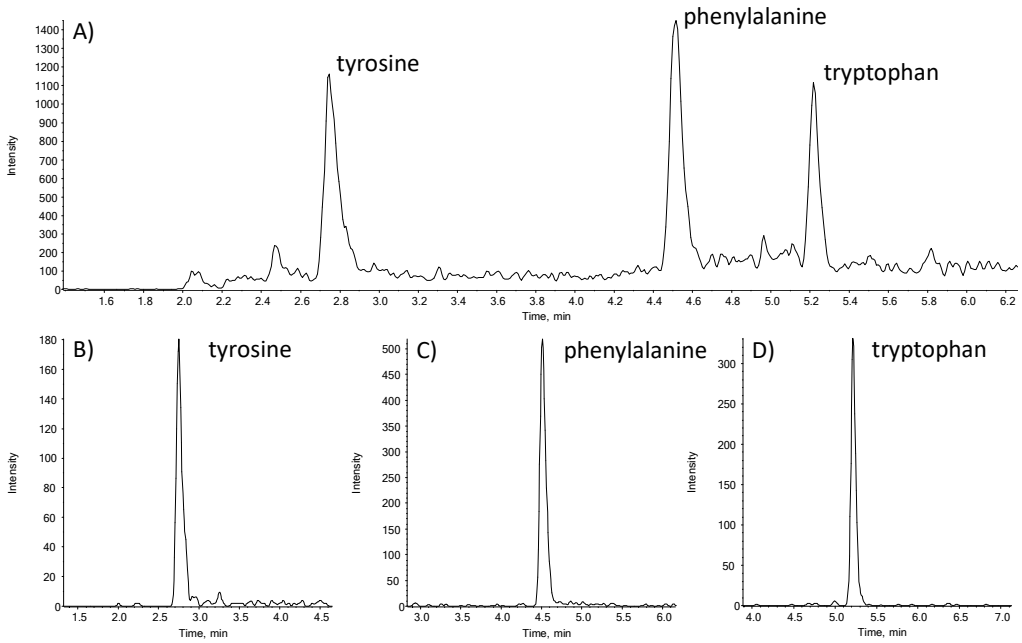


Figure 2: Chromatograms of drinking water spiked with 0.25 µg/L aromatic amino acids. A) Summed EIC $[M+H]^+$ trace (10 ppm) of aromatic amino acids. B) Tyrosine MS2 EIC of m/z 91.0542. C) Phenylalanine MS2 EIC 120.0808. D) Tryptophan MS2 EIC of m/z 118.0651.

Sample pre-treatment consisted of adding isotopically labeled internal standards of the aromatic amino acids to the sample followed by filtration using a 0.20 µm filter prior to LC-QToF analysis. No loss of aromatic amino acids was observed during the filtration step.

2.4 Method validation

The developed analytical method was validated for drinking and surface water. The limit of detection (LOD), limit of quantification (LOQ), repeatability and recovery were determined for all aromatic amino acids in both matrices. The validation results are shown in table 2 for drinking water and table 3 for surface water.

TABLE 2: VALIDATION RESULTS AROMATIC AMINO ACIDS IN DRINKING WATER (N=8)

Compounds	LOD (µg/L)	LOQ (µg/L)	Repeatability 1 µg/L (%)	Recovery 1 µg/L (%)
L-tyrosine	0.033	0.10	3.5	98.5
L-phenylalanine	0.008	0.10	2.6	98.5
L-tryptophan	0.016	0.10	2.2	101.9

TABLE 3: VALIDATION RESULTS AROMATIC AMINO ACIDS IN SURFACE WATER (N=8)

Compounds	LOD (µg/L)	LOQ (µg/L)	Repeatability 1 µg/L (%)	Recovery 1 µg/L (%)
L-tyrosine	*	0.10	2.5	93.1
L-phenylalanine	*	0.10	0.8	94.3
L-tryptophan	*	0.10	2.4	101.6

** Due to the presence of significant amounts of aromatic amino acids in surface water, it was not possible to determine all validation characteristics. The validation results of drinking water can be used as reference.*

Satisfactory LOD and LOQ results were obtained for the developed analytical method in drinking water. For surface water the LOD could not be determined due to the presence of significant amounts of aromatic amino acids and therefore the LOD of drinking water was used as a reference. The LOQ (i.e. ≥ 3x LOD) was determined for all aromatic amino acids at 0.10 µg/L. Recoveries in drinking- and surface water are between 90.0 and 105 % and are satisfactory. The reproducibility for all compounds is lower than 4 % (at 1 µg/L). The validation results for aromatic amino acids in drinking- and surface water show that the analytical method developed can successfully be applied for the determination of aromatic amino acids in water.

2.5 Source and artificial water analysis

After validation, the method developed was applied to the analysis of aromatic amino acids in source and artificial water. A sample was taken from Lake IJssel which is used as source water for Heemskerk drinking water treatment facility, which uses UV/H_2O_2 for disinfection. 22 N-DBPs were detected after MP UV treatment in this drinking water treatment facility. Two artificial water samples were also prepared containing Pony Lake or Suwannee river NOM. Artificial water containing Pony Lake NOM was treated with MP UV in the previous study [13], resulting in the detection of 84 N-DBPS. Suwannee River NOM is probably the best characterized NOM [28] and was therefore used as a reference. Results of aromatic amino acid analysis in source and artificial water are shown in table 4.

TABLE 4: RESULTS AROMATIC AMINO ANALYSIS IN SOURCE AND ARTIFICAL WATER

Compounds	Lake IJssel water intake (µg/L)	Pony Lake NOM 5.4 mg/L (µg/L)	Suwannee river NOM 5.5 mg/L (µg/L)
L-tyrosine	0.15	2.7	0.80
L-phenylalanine	0.16	1.5	0.44
L-tryptophan	< 0.1 (0.08)*	0.61	0.14

** Detected concentration was lower than LOQ but higher than LOD.*

Tyrosine, phenylalanine and tryptophan were detected in all samples with concentrations ranging from 0.08 to 2.7µg/L. Tryptophan was detected in Lake IJssel below the LOQ but higher than the LOD, therefore the reported concentration is semi-quantitative. It was shown that aromatic amino acids are present in moderate concentrations in Lake IJssel, meaning, meaning that aromatic amino acids can be a potential source for N-DBPs.

2.6 Summary/conclusion

A LC-QToF method was developed for the determination of aromatic amino acids in drinking- and surface water. Satisfactory LOD and LOQ results were obtained for both drinking- and surface water. Aromatic amino acids were detected in source and artificial water, showing that they can be a potential source for the formation of N-DBPS.

The developed analytical method demonstrated that it is well applicable for analysis of aromatic amino acids, and will therefore be used for the labeling experiments in chapter 3. However some adjustment will be made to the method in order to perform non-target screening for unknown aromatic N-DBPs (e.g. longer gradient and information dependent MS2 acquisition).

3 Labeling experiments

In chapter 2 it was shown that aromatic amino acids are present in source and artificial water, meaning that the possibility exist that some of the N-DBPs could originate from aromatic amino acids after MP UV treatment. In this chapter stable isotope labeling experiments will be conducted to really find out whether this is the case.

3.1 Stable isotope labeling strategy

In the previous study, an innovative stable isotope labeling strategy was developed for tracing N-DBPs in artificial water [13], based on incorporation of the nitrate atom originating from nitrate into an newly formed N-DBP after MP UV treatment. This strategy will also be used for labeling experiments with aromatic amino acids.

The labeling strategy works as follows: when stable isotope nitrate ($^{Fp}O_3^-$) is added to artificial water and normal nitrate ($^{14}NO_3^-$) is added to another artificial water sample from the same source and both are MP UV treated, ^{15}N will be incorporated into a newly formed N-DBP in the first artificial water sample, and ^{14}N will be incorporated into the same N-DBP in the second artificial water. This will result in a mass difference of 0.99704 Da between the N-DBP formed with normal nitrate ($^{14}NO_3^-$) and labeled nitrate ($^{15}NO_3^-$). This mass difference can be detected by high resolution mass spectrometry in combination with a non-target screening approach (see figure 3). Only N-DBPs will have this characteristic mass difference, and can therefore be distinguished from regular DBPs and background ions.

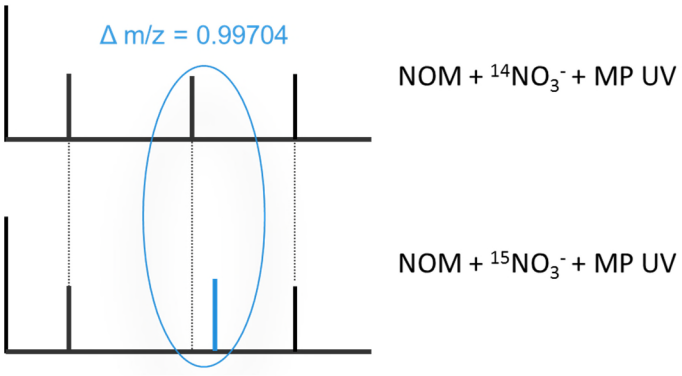


Figure 3: Stable isotope labeling strategy for the detection of N-DBPs. A mass spectrum is shown with a Δm/z of 0.99704 between the stable isotope labeled N-DBP compared to the normal N-DBP.

3.2 Tryptophan labeling experiments

Since tryptophan showed the highest Ames response and nitrogen gap in other studies [21, 22] after MP UV treatment, tryptophan is the obvious choice for starting the labeling experiments. Two stable isotope-labeled substances were used for labeling experiments, i.e. nitrate ($^{15}NO_3^-$) and tryptophan-$^{13}C_{11}$ ($^{13}C_{11}H_{12}N_2O_2$).The stable isotope-labeled tryptophan was used to provide additional certainty that the formed by-product really originates from tryptophan itself, and not from an unwelcome contamination. Another advantage of using labeled tryptophan is that the number of carbon atoms present in the formed by-products can easily be determined. A sample scheme was made for the tryptophan labeling experiments, in which different sample compositions are tested (table 5). All the samples consisted out of ultrapure water to which a combination of; unlabeled tryptophan, labeled

tryptophan, unlabeled nitrate, labeled nitrate and Pony Lake NOM was added. For the labeling experiments a relatively high concentration (5 mg/L) of labeled and unlabeled tryptophan was used, in order to ensure that high concentration of by-products were formed in, which easily should be detected using mass spectrometry. Nitrate and NOM concentration were at the same level as in the previous studies [13, 18], in order to obtain comparable results and to ensure by-product formation. NOM was added to some samples to find out if NOM or NOM intermediate products can react with tryptophan to form other by-products.

All samples were prepared in fourfold (with exception of the untreated reference) which then were subdivided into two duplicate sample sets, of which one duplicate sample set was MP UV treated while the other remained untreated. Of each sample 100 mL was prepared and was transferred to a glass sample bottle and was stored at 1-5 °C until MP UV treatment.

TABLE 5: OVERVIEW TRYPTOPHAN LABELING SAMPLE SCHEME

Samples	MP UV treatment	Unlabeled Tryptophan 5 mg/L	Labeled Tryptophan 5 mg/L	Unlabeled nitrate 10 mg/L	Labeled nitrate 10 mg/L	Pony Lake NOM 5 mg/L
Untreated reference (ultrapure water)						
Untreated + Trp, without nitrate		x				
UV treated + Trp, without nitrate	x	x				
Untreated + Trp + $^{14}NO_3^-$		x		x		
UV treated + Trp + $^{14}NO_3^-$	x	x		x		
Untreated + Trp + $^{15}NO_3^-$		x			x	
UV treated + Trp + $^{15}NO_3^-$	x	x			x	
Untreated + Trp + $^{14}NO_3^-$ + NOM		x		x		x
UV treated + Trp + $^{14}NO_3^-$ + NOM	x	x		x		x
Untreated + Trp + $^{15}NO_3^-$ + NOM		x			x	x
UV treated + Trp + $^{15}NO_3^-$ + NOM	x	x			x	x
Untreated + Trp-^{13}C, without nitrate			x			
UV treated + Trp-^{13}C, without nitrate	x		x			
Untreated + Trp-^{13}C + $^{14}NO_3^-$			x	x		
UV treated + Trp-^{13}C + $^{14}NO_3^-$	x		x	x		
Untreated + Trp-^{13}C + $^{15}NO_3^-$			x		x	
UV treated + Trp-^{13}C + $^{15}NO_3^-$	x		x		x	
Untreated + Trp-^{13}C + $^{14}NO_3^-$ + NOM			x	x		x
UV treated + Trp-^{13}C + $^{14}NO_3^-$ + NOM	x		x	x		x
Untreated + Trp-^{13}C + $^{15}NO_3^-$ + NOM			x		x	x
UV treated + Trp-^{13}C + $^{15}NO_3^-$ + NOM	x		x		x	x
UV treated + Trp-^{13}C + $^{14}NO_3^-$ + $^{15}NO_3^-$	x	x		x*	x*	
UV treated + Trp-^{13}C $^{14}NO_3^-$ + $^{15}NO_3^-$ + NOM	x	x		x*	x*	x

All samples were prepared in duplicate
** 5 mg/L*

3.2.1 MP UV treatment

The samples were sent to PWN technologies for MP UV treatment using a collimated beam set-up. Fifty-five mL of sample was transferred into a 60 × 35 mm crystallizing dish and was MP UV treated in open air at room temperature. The MP UV dose was delivered by a Trojan collimated beam apparatus using a 3 kW medium pressure Hg lamp. UV dose calculations were performed according to Bolton and Linden [29]. UV intensity was measured using a radiometer (International Light IL1700). A MP UV dose of 600 mJ/cm^2 was applied to each

sample. After irradiation the samples were returned and stored at 1-5 °C awaiting sample pre-treatment and LC-QToF analysis.

3.2.2 Sample pre-treatment

Twenty-five mL of water sample was transferred into a 25 mL flask, to which the internal standard tryptophan-d5 was added (100 µg/L). After homogenization, the samples were filtered using a 0.20 µm filter and were transferred to an autosampler vial for LC-QToF analysis.

3.2.3 LC-QToF analysis

For the analysis of tryptophan labeling samples, the aromatic amino acid method used in chapter 2 was partially adjusted. A longer linear gradient was applied, increased from 7 minutes to 100 %B, into 40 minutes to 100% B, in order to achieve a better separation for by-products and also to detect less-polar by-products. Data acquisition was performed in positive and negative ionisation mode. And instead of using a mass list for triggering MS/MS spectra, information dependent acquisition (IDA) was used for triggering MS/MS spectra. Eight IDA MS/MS spectra were triggered per full scan cycle, only for signals higher than 100 counts in combination with background subtraction and dynamic exclusion. The remainder of the LC-QToF settings can be found in attachment II (materials & methods labeling experiments).

3.2.4 Mass spectrometric data analysis

After LC-QToF analysis, the raw mass spectrometric data was processed using MasterView (Sciex) and differential analysis was performed in order to detect differences between $^{14}NO_3^-$, $^{15}NO_3^-$, Trp$^{12}C_{11}$, Trp-$^{13}C_{11}$, TRP-$^{12}C_{11}D_5$ and NOM MP UV treated samples, and the control samples. The intensity threshold for MasterView was set at 2000 counts for the positive and negative mode. The chromatographic data was compared from 1.5 to 35 min, with a mass range of 65-800 Da and an EIC width of 0.02 Da and retention window of 1min.

The nitrate labeling strategy was used for the detection of all N-DBPs formed by the MP UV treatment of nitrate. For this the UV treated Trp + $^{14}NO_3^-$ sample was compared with UV treated Trp + $^{15}NO_3^-$ sample, wherein all detected compounds with a mass difference of 0.99704 Da between the $^{14}NO_3^-$ and $^{15}NO_3^-$ sample, which were not present in the control samples, were detected as N-DBPs. Also the UV treated Trp + $^{14}NO_3^-$ sample was compared with UV treated Trp-$^{13}C_{11}$ + $^{14}NO_3^-$ sample, in order to detect all by-products originating from tryptophan, including by-products that were formed only by UV photolysis without interactions of nitro radicals. An overview of the amount of detected by-products is shown in table 6.

TABLE 6: OVERVIEW OF THE NUMBER OF DETECTED TRYPTOPHAN BY-PRODUCTS

Type of compounds	Number of accurate masses		Summed concentration Trp-d5 equivalents (µg/L)	
	Positive	Negative	Positive	Negative
DBPs (all)	957	1127	4049	3292
N-DBPs	157	278	680	741

Many by-products were formed by MP UV treatment of water containing tryptophan. In total 957 and 1127 accurate masses were detected in positive and negative mode, respectively. Of these detected accurate masses only a relative small number were N-DBPs, 157 for positive mode and 278 for negative mode. The number of accurate masses of DBPs and N-DBPs stated in the table are not all unique. Some of the detected accurate masses are from

fragments or adducts of by-products, therefore the actual number of by-products will be lower.

In order to obtain a good overview and to find out if most of the by-products are detected with the applied approach, a mass balance was made up. For this the concentration of tryptophan before UV treatment and after UV treatment was calculated using Trp-d5. All the detected by-products were also semi-quantified as Trp-d5 internal standard equivalent. The mass balance is shown in table 7.

TABLE 7: MASS BALANCE OF TRYPTOPHAN AND BY-PRODUCTS AFTER MP UV (TRP-D5 EQUIVALENTS)

Type of compounds	Positive mode		Negative mode	
	Concentration (µg/L Trp-d5 eq.)	(%)	Concentration (µg/L Trp-d5 eq.)	(%)
Tryptophan before UV	5188	-	5188	-
Tryptophan after UV	1117	21.5	1560	30.1
DBPs after UV	3369	64.9	2551	49.2
N-DBPs after UV	680	13.1	741	14.3
Sum Trp + DBPs	5166	99.6	4852	93.6

The concentration of tryptophan after MP UV is substantially decreased to 21.5% and 30.1% of its initial amount, for positive and negative mode respectively. This means that a considerable amount of tryptophan is converted by UV photolysis and nitro radicals into DBPs and N-DBPs, which is confirmed by the detected amounts DBPs and N-DBPs. In the end, 99.6% and 93.6% (positive and negative mode) of the mass balance is accounted for, demonstrating that most of by-products are detected using this approach. However, there are some remarks for calculating the mass balance like this. First, the concentration of the by-products cannot be determined exactly, because the ionisation efficiency is different for each by-product. Therefore the calculated concentration is an indication. Second, there is no correction made for the amount of nitrate/nitro groups reacting with tryptophan to form N-DBPs. Nevertheless, the mass balance still gives a good overview of the performed experiment and also shows that most of the by-products are probably detected.

3.2.5 Tryptophan labeling results

For the detection of N-DBPs formed by MP UV treatment of tryptophan, the UV treated Trp + $^{14}NO_3^-$ sample was compared with UV treated Trp + $^{15}NO_3^-$ sample as described in 3.2.4. An example of an N-DBP detected using the labeling strategy is shown figure 4. In this figure an extracted ion chromatogram is shown of an N-DBP with the elemental composition $C_8H_6N_2O_2$. In the $^{14}NO_3^-$ MP UV treated sample, a chromatographic peak is detected for m/z 161.0360, but is not present in the $^{14}NO_3^-$ sample without MP UV treatment. In the $^{15}NO_3^-$ MP UV treated sample, a chromatographic peak is detected for m/z 162.0331 and no peak is visible for m/z 161.0360. In the sample with an equal amount (1:1) of $^{14}NO_3^-$ and $^{15}NO_3^-$, both peaks are detected in the same ratio and at the same retention time. This confirms that detected compound at m/z 161.0360 is really a N-DBP. The 1:1 mixture sample can also be used to search for peak pairs with a 1:1 ratio and mass difference of m/z 0.99704 (for a single N incorporation), making data analysis easier.

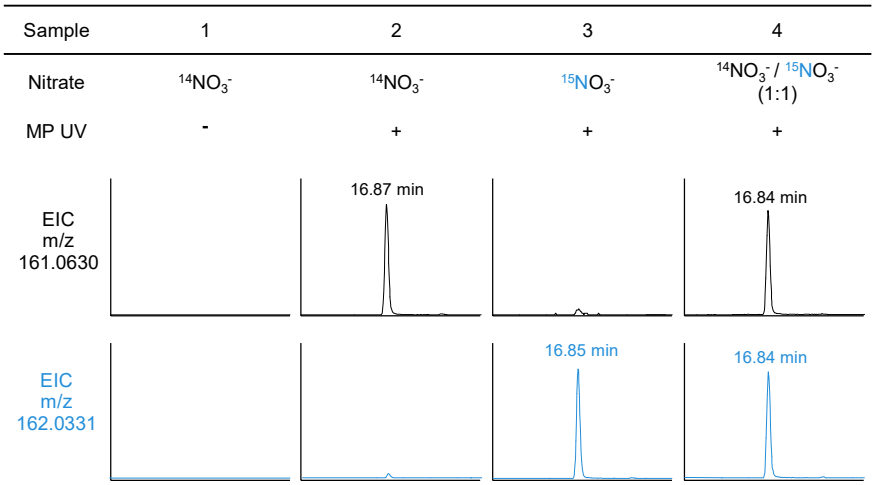


Figure 4: Example of a detected N-DBP ($C_8H_6N_2O_2$) in negative mode that is formed after MP UV treatment. The EICs of m/z 161.0360 and m/z 162.0331 are shown.

In total 157 and 278 accurate masses of N-DBPs were detected in the positive and negative mode, respectively The N-DBPs were considered to be "real" when they were present in the same ratio in the $^{14}NO_3^-$ and $^{15}NO_3^-$ sample after MP UV treatment and were present in the 1:1 mixture sample in the same ratio and were also detected in the ^{13}C-Trp sample after MP UV treatment. The ^{13}C-Trp sample was used for confirmation and determination of the amount of ^{13}C atoms in the formed by-products, which give extra information about the elemental compositions of the by-products. The top 10 and top 15 of the highest detected N-DBPs in the positive and negative mode are shown respectively in tables 8 and 9. The elemental composition and mass error were determined for the listed N-DBPs, and for some of the detected N-DBPs the identity was also determined. For the identification, a diagram containing common tryptophan modifications by nitrating agents [30] was used, or were determined using the elemental composition and MS/MS spectrum. All identifications are tentative.(i.e. level 2/3 according to Schymanski [31])

TABLE 8: MOST ABUNDANT TRYPTOPHAN N-DBPS IN POSITIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	Accurate mass ^{13}C-Trp	RT (min)	Intensity	Concn. Trp-d5 equiv (µg/L)	13C atoms	Formula	Δppm	Identity
282.0729	283.0702	293.1097	6.36	344064	61	11	$C_{11}H_{11}N_3O_6$	3.7	dihydroxy-nitrotryptophan
220.0723	221.0691	230.1051	6.97	329039	58	10	$C_{10}H_9N_3O_3$	2.4	
220.0721	221.0691	230.1051	5.59	300943	53	10	$C_{10}H_9N_3O_3$	1.5	
266.0780	267.0749	277.1146	3.90	174436	31	11	$C_{11}H_{11}N_3O_5$	3.2	hydroxy-nitrotryptophan
248.0673	249.0642	259.1038	3.90	139152	25	11	$C_{11}H_9N_3O_4$	2.9	
266.0779	267.0749	277.1145	5.22	137670	24	11	$C_{11}H_{11}N_3O_5$	3.2	hydroxy-nitrotryptophan
264.0625	265.0593	275.0989	4.02	118528	21	11	$C_{11}H_9N_3O_5$	2.7	
250.0828	251.0801	261.1196	8.91	108081	19	11	$C_{11}H_{11}N_3O_4$	2.3	
254.0779	255.0748	264.1112	7.03	104660	19	10	$C_{10}H_{11}N_3O_5$	2.6	
203.0453	204.0423	213.0784	6.96	97688	17	10	$C_{10}H_6N_2O_3$	0.9	

TABLE 9: MOST ABUNDANT TRYPTOPHAN N-DBPS IN NEGATIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	Accurate mass ^{13}C-Trp	RT (min)	Intensity	Concn. Trp-d5 equiv (µg/L)	13C atoms	Formula	Δppm	Identity
189.0311	190.0281	198.0611	14.53	253492	42	9	$C_9H_6N_2O_3$	2.3	nitroindole-carbaldehyde
280.0579	281.0550	291.0944	6.36	240716	40	11	$C_{11}H_{11}N_3O_6$	1.4	dihydroxy-nitrotryptophan
217.0260	218.0229	227.0595	10.85	237069	40	10	$C_{10}H_8N_2O_5$	1.5	
236.0681	237.0649	246.1014	5.59	230689	39	10	$C_{10}H_{11}N_3O_4$	1.4	
264.0632	265.0600	275.0998	3.90	214684	36	11	$C_{11}H_{11}N_3O_5$	1.9	hydroxy-nitrotryptophan
264.0631	265.0600	275.0996	5.21	162315	27	11	$C_{11}H_{11}N_3O_5$	1.5	hydroxy-nitrotryptophan
262.0476	263.0443	273.0839	4.03	136529	23	11	$C_{11}H_9N_3O_5$	1.4	
248.0677	249.0647	259.1052	8.58	104598	18	11	$C_{11}H_{11}N_3O_4$	-0.3	nitrotryptophan
233.0205	234.0174	243.0544	8.89	99350	17	10	$C_{10}H_6N_2O_5$	0.0	
190.0260	191.023	198.0527	10.66	88417	15	8	$C_8H_5N_3O_3$	0.4	
280.0574	281.0543	291.0942	11.9	83280	14	11	$C_{10}H_9N_3O_4$	-0.4	
161.0360	162.0331	169.0629	16.87	70977	12	8	$C_8H_6N_2O_2$	2.2	
189.0308	190.0278	198.0609	15.48	64659	11	9	$C_9H_6N_2O_3$	0.2	
235.0359	236.0330	245.0692	13.67	56744	9	10	$C_{10}H_8N_2O_5$	-1.0	
280.0573	281.0542	291.0939	10.16	50687	8	11	$C_{11}H_{11}N_3O_6$	-1.1	

Dihydroxy-nitrotryptophan is detected as highest in the positive mode and as second highest in the negative mode. The identity of the most intense N-DBP with m/z 189.0311 in the negative mode is uncertain. A possible candidate is nitroindole-carbaldehyde, but many structural isomers are possible. Another frequently detected N-DBP (multiple isomers) is nitrotryptophan, which is detected using both ionisation modes. As expected there is a substantial overlap between the N-DBPs detected in the positive and negative mode, due to the presence of functional groups that are ionisable in the positive (e.g. nitrogen) and negative (e.g. carboxyl and hydoxy) ionisation mode.

3.2.6 NOM samples

For the tryptophan labeling experiments also some samples were prepared containing NOM (see 3.2) in order to find out if NOM or NOM intermediate products can react with tryptophan to form other by-products, or have any effect on by-product formation. For this the UV treated Trp + $^{14}NO_3^-$ + NOM sample was compared with UV treated Trp + $^{15}NO_3^-$ + NOM sample. The top 25 highest detected by-products were then compared with the top 25 by-products formed without NOM to check if there was any difference. See table 10 for an overview of the total amount of N-DBPs detected in the NOM sample, compared with the amount of N-DBPs detected without NOM.

TABLE 10: COMPARSION OF THE AMOUNT OF N-DBPS DETECTED WITH AND WITHOUT NOM

Sample	Detected accurate masses		Summed concentration Trp-d5 equivalents (µg/L)	
	Positive	Negative	Positive	Negative
Trp + $^{14}NO_3^-$ without NOM	157	278	680	741
Trp + $^{14}NO_3^-$ with NOM	165	189	595	459

The same N-DBPs were found with and without NOM, although the concentrations detected in the NOM samples are on average lower. The total amount and concentration of by-products detected in the NOM samples is lower. This is probably due to the available amount of nitrate present in the sample, resulting in competition between the formation of

tryptophan N-DBPs and NOM N-DBPs. So in the end more by-products are probably formed in the presence of NOM, but remain undetected because they fall below the threshold of detection. This experiment shows that presence of NOM has a relatively small effect on tryptophan N-DBP formation, and will therefore not be used for the stable isotope labeling experiments with tyrosine and phenylalanine.

3.2.7 Suspect screening of 84 N-DBPs

The goal of the tryptophan labeling experiments was to find out if some of unidentified N-DBPs detected in artificial water and/or full-scale water treatment facility (both MP UV treated) in the prior study could originate from tryptophan. So a suspect screening was performed for the 84 N-DBPs detected in previous study (see attachment III for the list). For the suspect screening the MP UV treated Trp + $^{14}NO_3^-$ sample was used. In order to confirm a possible match, the artificial water samples and samples of the full-scale water treatment facility (both SPE extracts) of the prior study were analysed again using the analytical method for the tryptophan labeling experiments. The sample extracts were pre-treated in the prior study using the AMES SPE protocol, in order to achieve sufficient sensitivity for the Ames fluctuation assay and non-target HR-MS screening. Because the aromatic amino labeling experiments were conducted at relatively high concentrations, SPE treatment was not needed, and sufficient sensitivity was achieved using direct injection.

With the suspect screening, only one N-DBPs at m/z 161.0360 (in negative mode) was detected in the original $^{14}NO_3^-$ artificial water sample of the prior study, was also detected in the MP UV treated Trp + $^{14}NO_3^-$ sample of the current study. See table 11 for the results. The N-DBP with m/z 161.0360 was not detected in the full-scale water treatment sample.

TABLE 11: RESULTS SUSPECT SCREENING OF THE MOST ABUNDANT TRYPTOPHAN N-DBPS IN THE ARTIFICAL WATER SAMPLE OF THE PRIOR STUDY

Accurate mass $^{14}NO_3^-$	RT (min)	Intensity	Concn. Trp-d5 equiv (µg/L)	Formula	RT Trp N-DBPs Trp + $^{14}NO_3^-$ sample (min)	RT Original $^{14}NO_3^-$ sample (min)	RT, MS1 and MS2 Confirmed
161.0360	16.87	70977	12	$C_8H_6N_2O_2$	16.87	16.88	yes

With the suspect screening there were no N-DBPs detected in the positive mode in the original $^{14}NO_3^-$ artificial water sample. In the negative mode one by-product with m/z 161.0360 was detected and confirmed in the original $^{14}NO_3^-$ artificial water sample by matching retention time and MS/MS spectrum (see figure 5 and table 11). This demonstrates that experimental design of this study has worked, and that one by-product detected in the prior study could potentially originate of tryptophan.

In the prior study only 16 N-DBPs (of which 6 uniquely) of the 84 N-DBPs were detected in the positive mode, therefore it was taken into account that N-DBPs originating from tryptophan most likely would be detected in the negative mode. Of the top 15 by-products formed by the MP UV tryptophan labeling experiments, only number 12 (m/z 161.0360), a relatively low intensity N-DBPs was detected in artificial water. This could mean that the other by-products are not formed during the MP UV irradiation of artificial water, or that these other formed by-products are not sufficiently extracted from the water using AMES pre-treatment protocol, which was used as sample pre-treatment in the prior study. So in order to investigate if the detected N-DBP, really originate of tryptophan, SPE experiment will be conducted to determine the extraction recovery of the top 15 N-DBPs in negative mode. See 3.4 for the SPE experiments.

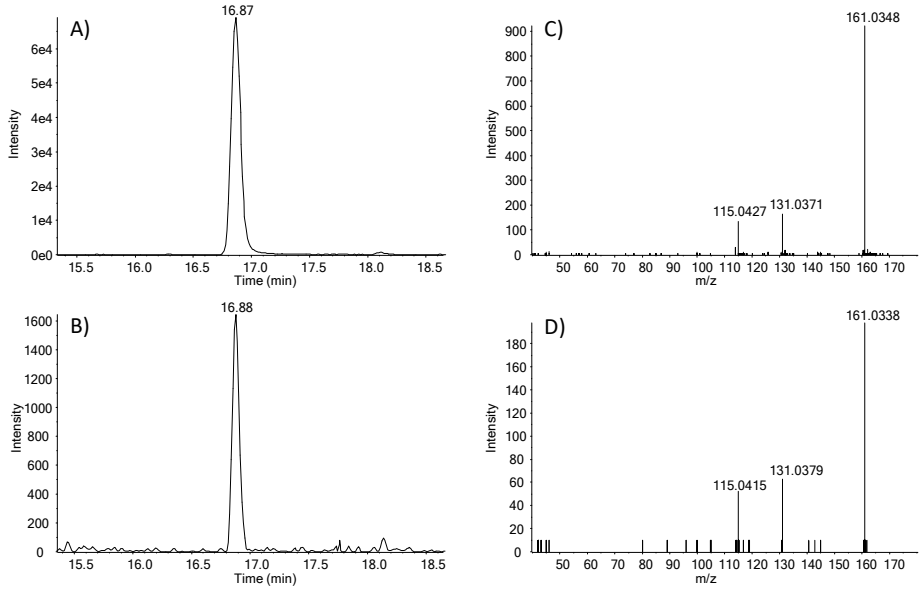


Figure 5: confirmation of N-DBP m/z 161.0360 in artificial water. A) EIC of m/z 161.0360 in Trp + $^{14}NO_3^-$ after MP UV sample. B) EIC of m/z 161.0360 in artificial water after MP UV. C) MS2 spectrum of m/z 161.0360 in Trp + $^{14}NO_3^-$ after MP UV sample. D) MS2 spectrum of m/z 161.0360 in artificial water after MP UV.

3.2.8 Identification of m/z 161.0360

Before a toxicological evaluation can be performed for the N-DBPs with m/z 161.0360, the identity must be known. The identification process for m/z 161.0360 ($C_8H_6N_2O_2$) started with investigating the loss of elements during N-DBP formation. The molecular formula of the by-product ($C_8H_6N_2O_2$) was subtracted from that of tryptophan ($C_{11}H_{12}N_2O_2$). So during N-DBP formation C_3H_6 was lost, and the by-product was very likely nitrated, and gained also a nitro (NO_2) group. This means that of the original tryptophan structure $C_3H_6NO_2$ was lost. This loss shows that the acid functional group and amine functional group of the basic amino acid structure were probably lost during N-DBP formation, whereby the nitrated indole structure remains. See figure 6 for the structure of indole.

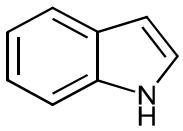


Figure 6: structure of indole (CAS nr. 120-72-9)

The theory that the formed by-product is nitroindole, is supported by the observed retention for this by-product (16.87 min) compared with that of tryptophan (6.86). Nitroindole is less polar due to the missing amine and acid functional group, and should therefore have a substantially higher retention, which is the case. Next a reference spectrum was sought and found of 6-nitroindole in the NIST EI-GC-MS library. This EI-GC-MS spectrum (figure 7) containing nominal masses, was compared with that of the MS2 spectrum of the by-product (figure 5C) and showed similarities. There is a mass difference of one Da between the EI-GC-MS and LC-MS MS2 $[M-H]^-$ spectrum, which can be explained by the loss of a proton (negative mode ionisation).

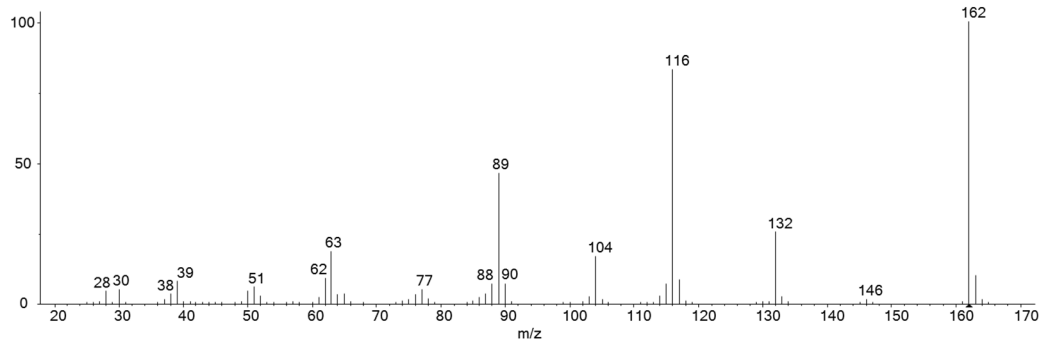


Figure 7: EI-GC-MS reference spectrum of 6-nitroindole

Because nitroindole has many structural isomers, the most common structural isomers were ordered as reference standard for the confirmation of the identity of the by-product. The results of the reference standards and the Trp + $^{14}NO_3^-$ MP UV sample analysis is shown in table 12.

TABLE 12: RESULTS REFERENCE STANDARD ANALYSIS OF NITROINDOLE ISOMERS

Compound	CAS nr	RT sample (min)	RT reference standard (min)	RT, MS1 and MS2 Confirmed
7-Nitroindole	6960-42-5	16.19	18.70	no
6-Nitroindole	4769-96-4	16.19	18.28	no
5-Nitroindole	6146-52-7	16.19	17.49	no
4-Nitroindole	4769-97-5	16.19	17.06	no
3-Nitroindole	4770-03-0	16.19	16.18	yes

The identity of 3-nitroindole was unambiguously confirmed as one of the N-DBPs being formed after tryptophan MP UV irradiation. The retention time and MS2 spectrum of the Trp + $^{14}NO_3^-$ MP UV sample matches exactly with that of the 3-nitroindole reference standard (see figure 8).

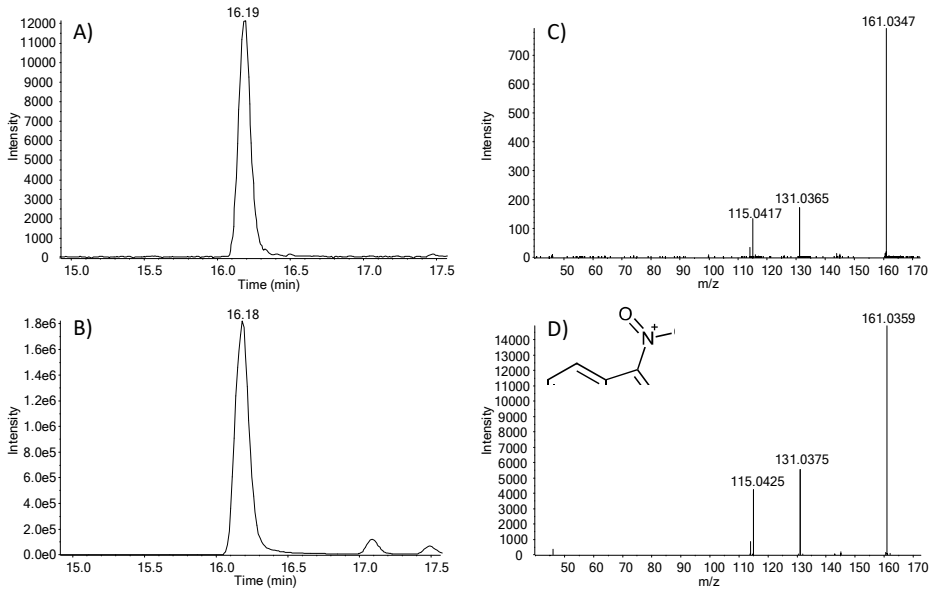


Figure 8: Confirmation of 3-nitroindole. A) EIC of 3-nitroindole (m/z 161.0360) in Trp + $^{14}NO_3^-$ MP UV sample. B) EIC of 3-nitroindole reference standard (m/z 161.0360). C) MS2 spectrum of 3-nitroindole in Trp + $^{14}NO_3^-$ MP UV sample. D) MS2 spectrum of 3-nitroindole reference standard.

3.3 Tyrosine and phenylalanine labeling experiments

After completing the tryptophan labeling experiments, the labeling experiments were continued with the two remaining aromatic amino acids, tyrosine and phenylalanine. The same approach was used as for the tryptophan labeling experiments, with some adjustments. For the tyrosine and phenylalanine labeling experiments no isotope labeled ^{13}C carbon reference standard was used, because the tryptophan experiments showed that the added value of such a standard is limited and therefore too expensive. Also no experiments were conducted with added NOM to the samples, because it had a limited effect on N-DBP formation during the tryptophan labeling experiments. An overview of final sample scheme for the tyrosine and phenylalanine labeling experiments is shown in table 13.

TABLE 13: OVERVIEW OF THE TYROSINE AND PHENYLALANINE LABELING SAMPLE SCHEME

Samples	MP UV treatment	Tyrosine 5 mg/L	Phenylalanine 5 mg/L	Unlabeled nitrate ($^{14}NO_3^-$) 10 mg/L	Labeled nitrate ($^{15}NO_3^-$) 10 mg/L
Untreated reference (ultrapure water)					
Untreated + Tyr, without nitrate		x			
UV treated + Tyr, without nitrate	x	x			
Untreated + Tyr + $^{14}NO_3^-$		x		x	
UV treated + Tyr + $^{14}NO_3^-$	x	x		x	
Untreated + Tyr + $^{15}NO_3^-$		x			x
UV treated + Tyr + $^{15}NO_3^-$	x	x			x
Untreated + Phe, without nitrate			x		
UV treated + Phe, without nitrate	x		x		
Untreated + Phe + $^{14}NO_3^-$			x	x	
UV treated + Phe + $^{14}NO_3^-$	x		x	x	
Untreated + Phe + $^{15}NO_3^-$			x		x
UV treated + Phe + $^{15}NO_3^-$	x		x		x
UV treated + Tyr + $^{14}NO_3^-$ + $^{15}NO_3^-$	x	x		x*	x*
UV treated + Phe + $^{14}NO_3^-$ + $^{15}NO_3^-$	x		x	x*	x*

All samples were prepared in duplicate
** 5 mg/L*

The samples were MP UV treated, pre-treated, and analysed according to the same conditions as described in paragraphs 3.2.1- 3.2.3.

3.3.1 Mass spectrometric data analysis

The mass spectrometric data analysis was performed under same conditions as described in paragraph 3.2.4. The nitrate labeling strategy was used again for the detection of N-DBPs formed by MP UV treatment of nitrate. The MP UV treated Tyr or Phe + $^{14}NO_3^-$ sample was compared with MP UV treated Tyr or Phe + $^{15}NO_3^-$ sample for the detection of N-DBPs. In order to detect all by-products, including by-products that were formed by UV photolysis and without interactions of nitro radicals, the MP UV untreated Tyr or Phe + $^{14}NO_3^-$ sample was compared with the MP UV treated Tyr or Phe + $^{14}NO_3^-$ sample. An overview of the number of detected by-products for tyrosine and phenylalanine is shown in table 14.

TABLE 14: OVERVIEW DETECTED TYROSINE AND PHENYLALANINE BY-PRODUCTS

Type of compounds	Detected accurate masses		Summed concentration (µg/L)	
	Positive	Negative	Positive	Negative
Tyrosine			(Tyr-d4 equiv.)	(Tyr-d4 equiv.)
DBPs (all)	142	128	1270	2274
N-DBPs	17	24	643	1037
Phenylalanine			(Phe-d5 equiv.)	(Phe-d5 equiv.)
DBPs (all)	89	126	488	1642
N-DBPs	30	70	246	1138

There are fewer accurate masses detected with the tyrosine (142 and 128; pos/neg) and phenylalanine (89 and 126; pos/neg) labeling experiments in comparison with tryptophan (957 and 1127; pos/neg). An even lower number of N-DBPs was detected for tyrosine (17 and 24; pos/neg) and phenylalanine (28 and 70; pos/neg). Although relatively high concentrations of N-DBPS were detected for tyrosine (1037 µg/L) and phenylalanine (1138 µg/L) compared with tryptophan (741 µg/L). This shows that the average concentration of the detected tyrosine and phenylalanine accurate masses is higher than that of tryptophan and that tyrosine and phenylalanine are also susceptible for nitration.

A mass balance was made up for tyrosine and phenylalanine, in order to find out if most of the by-products are really detected. The mass balance for tyrosine and phenylalanine is shown in table 15.

TABLE 15: MASS BALANCE OF TYROSINE AND PHEYNYLALANINE BY-PRODUCTS AFTER MP UV (TYR-D4 OR PHE-D5 EQUIVALENTS)

Type of compounds	Positive mode		Negative mode	
	Concentration µg/L	(%)	Concentration µg/L	(%)
Tyrosine				
Tyrosine before UV	5178	-	5178	-
Tyrosine after UV	3169	61.2	2936	56.7
DBPs after UV	627	12.1	1237	23.9
N-DBPs after UV	643	12.4	1037	20.0
Sum Tyr + DBPs	4439	85.7	5210	100.6
Phenylalanine				
Phenylalanine before UV	5188	-	5188	-
Phenylalanine after UV	3874	74.7	3771	72.7
DBPs after UV	242	4.7	504	9.7
N-DBPs after UV	246	4.7	1138	21.9
Sum Tyr + DBPs	4362	84.1	5413	104.3

After MP UV treatment a relatively large amount of tyrosine (61.2 and 56.7%) and phenylalanine (74.7 and 72.7%) remains intact in comparison with that of tryptophan (21.5 and 30.1%) for the positive and negative mode. This shows that only a small amount of tyrosine and phenylalanine is converted by UV photolysis and nitro radicals into DBPs and N-DBPs at MP UV conditions. The overall result of the mass balance shows that for tyrosine 85.7% and 100.6% (pos and neg) of the mass balance is accounted for, and for 84.1 and

104.3% (pos and neg) for phenylalanine. This means that most of the by-products are detected. The mass balance results in the negative mode are greater than 100%, which can be explained by that some fragments and adducts are also included in the detection of the by-products. And because the ionisation efficiency is different for each by-product, resulting in less reliable concentrations.

3.3.2 Tyrosine labeling results

For the detection of tyrosine N-DBPs formed by MP UV treatment, the UV treated Tyr + $^{14}NO_3^-$ sample was compared with UV treated Tyr +$^{15}NO_3^-$ sample as described in 3.2.4. Only 17 and 24 accurate masses of N-DBPs were detected in the positive and negative mode. The top 2 and top 5 of the highest detected tyrosine N-DBPs in the positive and negative mode are shown in table 16 and 17.

TABLE 16: MOST ABUNDANT TYROSINE N-DBPS IN POSITIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	RT (min)	Intensity	Concn. Tyr-d4 equiv (µg/L)	Formula	Δppm	Identity
227.0664	228.0637	5.60	1646000	599	$C_9H_{10}N_2O_5$	0.7	Nitrotyrosine
227.0663	228.0632	2.78	71667	26	$C_9H_{10}N_2O_5$	-0.2	Nitrotyrosine

TABLE 17: MOST ABUNDANT TYROSINE N-DBPS IN NEGATIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	RT (min)	Intensity	Concn. Tyr-d4 equiv (µg/L)	Formula	Δppm	Identity
225.0520	226.0492	5.59	978589	617	$C_9H_{10}N_2O_5$	1.8	Nitrotyrosine
225.0517	226.0487	2.80	194186	122	$C_9H_{10}N_2O_55$	0.5	Nitrotyrosine
146.9667	148.9608	2.19	85259	54	?		2x N15 label
439.0374	440.0343	2.78	17116	11	?		
177.0305	178.0282	17.85	9892	6	$C_8H_6N_2O_3$	1.3	

Two isomers of nitrotyrosine were detected in high concentrations in the positive (625 µg/L) and negative mode (739 µg/L). In the positive mode nitrotyrosine accounted for more than 97% of total amount of the formed N-DBPs and in negative mode for more than 71%. The remaining N-DBPS in the negative mode with exception of m/z 146.9667 were too low in concentration for identification. The N-DBP with m/z 146.9667 showed a double nitrogen label, meaning that it had reacted twice with nitrate, but could not be identified.

3.3.3 Phenylalanine labeling results

Phenylalanine N-DBPs formed by MP UV treatment were detected by comparing the UV treated Phe + $^{14}NO_3^-$ sample with the UV treated Tyr +$^{15}NO_3^-$ sample as described in 3.2.4. In total 30 and 70 accurate masses of N-DBPs were detected in the positive and negative mode. The top 5 and top 10 of the highest detected phenylalanine N-DBPs in the positive and negative mode are shown in table 18 and 19, respectively.

TABLE 18: MOST ABUNDANT PHENYLALANINE N-DBPS IN POSITIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	RT (min)	Intensity	Concn. Phe-d5 equiv (µg/L)	Formula	Δppm	Identity
227.0666	228.0635	5.60	479466	67	$C_9H_{10}N_2O_5$	1.1	Nitrotyrosine
227.0665	228.0635	6.09	438853	61	$C_9H_{10}N_2O_5$	0.8	nitro-hydroxyphenylalanine
227.0664	228.0635	6.39	417768	58	$C_9H_{10}N_2O_5$	0.7	nitro-hydroxyphenylalanine
227.0664	228.0635	5.90	141834	20	$C_9H_{10}N_2O_5$	0.7	nitro-hydroxyphenylalanine
227.0664	228.0633	3.93	53856	8	$C_9H_{10}N_2O_5$	0.7	nitro-hydroxyphenylalanine

TABLE 19: MOST ABUNDANT PHENYLALANINE N-DBPS IN NEGATIVE MODE

Accurate mass $^{14}NO_3^-$	Accurate mass $^{15}NO_3^-$	RT (min)	Intensity	Concn. Phe-d5 equiv (µg/L)	Formula	Δppm	Identity
225.0519	226.0489	6.09	390075	148	$C_9H_{10}N_2O_5$	0.5	nitro-hydroxyphenylalanine
225.0517	226.0490	5.60	353717	134	$C_9H_{10}N_2O_5$	0.0	Nitrotyrosine
222.0156	224.0097	18.45	322402	122	$C_8H_5O_5N_3$	0.7	2x N15 label
225.0519	226.0488	6.39	280288	106	$C_9H_{10}N_2O_5$	0.0	nitro-hydroxyphenylalanine
222.0156	224.0097	17.79	279937	106	$C_8H_5O_5N_3$	-0.2	2x N15 label
267.0005	506.0145	17.31	128372	49	$C_8H_4N_4O_7$	-0.8	3x N15 label
177.0311	178.0279	17.23	99836	38	$C_8H_6N_2O_3$	2.5	nitrooxindole?
504.0208	506.0145	6.08	95040	36	?		2x N15 label
436.0008	437.9951	17.79	89085	34	?		2x N15 label
177.0311	178.0279	14.72	86245	33	$C_8H_6N_2O_3$	3.0	
391.0165	392.0135	14.72	78018	30	?		
243.8992	244.8901	2.29	73738	28	?		
257.0410	258.0384	5.47	71785	27	$C_9H_{10}N_2O_7$	-1.3	
168.0305	169.0274	10.31	52166	20	$C_7H_7NO_4$	1.0	
208.0250	209.0223	5.91	48094	18	$C_9H_7NO_5$	0.5	

Nitrotyrosine and the structural isomer nitro-hydroxyphenylalanine were detected in high concentrations in the positive and negative mode. In the positive mode, nitrotyrosine and nitro-hydroxyphenylalanines together account for more than 86% of the concentration of detected by-products. In the negative mode this accounts only for less than 36% and many other by-products are also detected. For the negative detected phenylalanine by-products, it is striking that four by-products are detected with a double nitrogen label and even one with a triple nitrogen label (see figure 9 for the mass spectrum of m/z 267.0005, containing a triple nitrogen label). This shows that phenylalanine is very susceptible for double or even triple nitration.

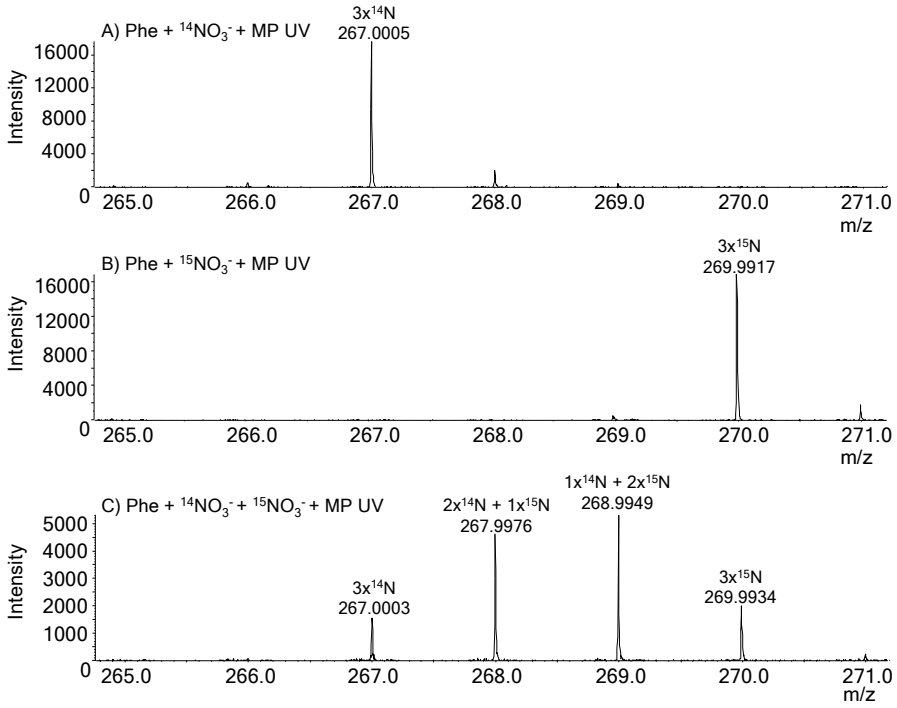


Figure 9: Mass spectra of N-DBP m/z 267.005 in various samples showing different ^{14}N and ^{15}N labels. A) Mass spectrum of Phe + $^{14}NO_3^-$ + MP UV sample, showing a triple ^{14}N label. B) Mass spectrum of Phe + $^{15}NO_3^-$ + MP UV sample, showing a triple ^{15}N label. C) Mass spectrum of Phe +$^{14}NO_3^-$ + $^{15}NO_3^-$ + MP UV sample, showing a triple ^{14}N, double ^{14}N + single ^{15}N, single ^{14}N + double ^{15}N and triple ^{15}N label.

3.3.4 Suspect screening of 84 N-DBPs

After the tyrosine and phenylalanine experiments a suspect screening was performed using the 84 N-DBPs detected in previous study (see attachment III for the list). The suspect screening was performed on the MP UV treated Phe + $^{14}NO_3^-$ and MP UV treated Phe + $^{15}NO_3^-$ sample. The suspect screening was performed as described in 3.2.7. See table 20 for the results of the suspect screening.

TABLE 20: RESULTS SUSPECT SCREENING OF THE MOST ABUNDANT PHENYLALANINE N-DBPS IN THE ARTIFICAL WATER SAMPLE OF THE PRIOR STUDY

Accurate mass $^{14}NO_3^-$	RT (min)	Intensity	Concn. Phe-d5 equiv (µg/L)	Formula	RT Original $^{14}NO_3^-$ sample (min)	RT, MS1 and MS2 Confirmed
222.0156	18.45	322402	122	$C_8H_5O_5N_3$	18.42	yes
222.0156	17.79	279937	106	$C_8H_5O_5N_3$	17.76	yes

With the suspect screening there were no tyrosine N-DBPs detected in the original $^{14}NO_3^-$ artificial water sample in the positive and negative mode and for phenylalanine no N-DBPs were detected in the positive mode. For phenylalanine in the negative mode, two by-products with m/z 222.0156 (isomers) were detected and confirmed in the original $^{14}NO_3^-$ artificial water sample by matching retention time and MS/MS spectrum (see figure 10), but these were not detected in the full-scale water treatment sample. This shows once again that aromatic amino acids could potentially be the source of some of the by-products detected in the prior study.

Of the most abundant 15 by-products formed by the MP UV phenylalanine labeling experiments, only number 3 and 5 (both m/z 222.0156) corresponded with the list of suspects. As before with the tryptophan labeling experiments, this could mean that the other by-products are not formed during the MP UV irradiation of artificial water or that these by-products are not sufficiently extracted from the water. This will be further investigated in 3.4.

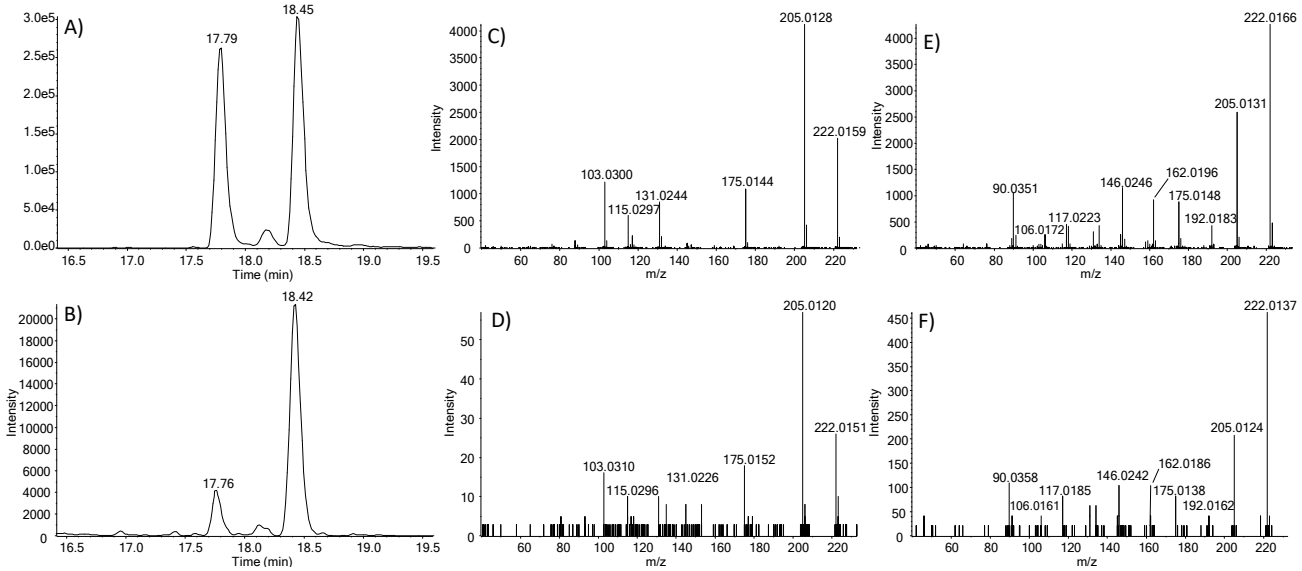


Figure 10: confirmation of N-DBP m/z 222.0156 (2 isomers) in artificial water. A) EIC of m/z 222.0156 (2 isomers) in Phe + $^{14}NO_3^-$ after MP UV sample. B) EIC of m/z 222.0156 in artificial water after MP UV. C) MS2 spectrum at RT 17.79 of m/z 222.0156 in Phe + $^{14}NO_3^-$ after MP UV sample. D) MS2 spectrum at RT 17.76 of m/z 222.0156 in artificial water after MP UV. E) MS2 spectrum at RT 18.45 of m/z 222.0156 in Phe + $^{14}NO_3^-$ after MP UV sample. F) MS2 spectrum at RT 18.42 of m/z 222.0156 in artificial water after MP UV.

3.3.5 Identification of m/z 222.0156

The identification of the two isomers with m/z 222.0156 ($C_8H_5N_3O_5$) containing both a double nitrogen label started by annotating the MS/MS spectrum and searching for known losses (See attachment IV for the annotated MS/MS spectra). This showed mainly losses of NO and OH, meaning that there are probably nitro and hydroxyl functional groups present in the by-products. Next the formula of the by-products ($C_8H_5N_3O_5$) was compared with that of phenylalanine ($C_9H_{11}NO_2$). This showed there was a loss of one carbon and at least one oxygen (assuming that the by-product was double nitrated, and also alcohol oxidation had occurred), indicating that the polar acidic functional group of the amino acid was probably lost during formation of the by-product. This is also supported by the observed retention time (17.79 and 18.45 min) of both by-products in comparison with that of phenylalanine (4.65 min), which means that the by-products are less polar than the parent compound. The by-products are expected to consist of a benzene ring containing two nitro and one hydroxy functional group and also containing the remainder of the amine part of the amino acid (C_2H_xN). Then a ChemSpider and PubChem database search was performed using the formula of the by-products which resulted in 37 and 65 hits. The structures of these possible candidates were thoroughly investigated and finally one good candidate was found: 3,5-Dinitro-4-hydroxy-benzylcyanid (see figure 11). There were no other structural isomers in the ChemSpider and Pubchem database, although there are many more structural isomers possible. The only candidate was checked by computational (in silico) ms/ms fragmentation using MetFrag [32], in which the computational MS/MS spectrum was compared with the measured MS/MS spectrum and showed a decent match (8 of 13 fragments explained) with the by-product detected at 18.45min. Unfortunately this candidate was not available for

purchase, and remains therefore unconfirmed. There are many structural isomers possible for this structure, so synthesis is not option. The only option that remains for identification is using nuclear magnetic resonance (NMR), but that falls beyond the scope of this study.

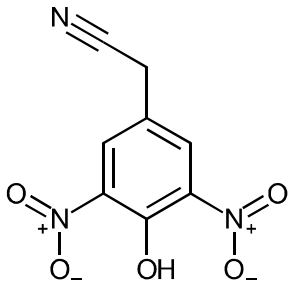


Figure 11: structure of 3,5-dinitro-4-hydroxy-benzylcyanid (CAS nr. 55770-69-9)

The basic structure of this by-product is benzyl cyanide, which is a known toxic compound for which an oral LD50 in mouse was derived at 45.5 mg/kg [33]. Furthermore, Ma et al [34] demonstrated that benzyl cyanide is the major by-product of phenylalanine chlorination in drinking water. This means that it is a possibility that the detected by-product with m/z 222.0156 really could be a derivative from benzyl cyanide and makes this by-product interesting for further research.

3.4 SPE experiments of tryptophan and phenylalanine by-products

A total of three N-DBPs were detected with the tryptophan and phenylalanine MP UV labeling experiments that were also present in the artificial water sample of the prior study, meaning that these by-products in the artificial water sample potentially could originate from aromatic amino acids. However of the top 15 by-products detected with the tryptophan and phenylalanine labeling experiments, only these three by-products (i.e. 3-nitroindole and two isomers with m/z 222.0156) were detected in the artificial water sample, of which 3-nitroindole is a relatively low intensity N-DBP (nr 12). This would imply that other by-products of the top 15 are not formed by MP UV treatment of artificial water, which makes it more likely that the three detected by-products could originate from other sources than from aromatic amino acids. Alternatively, it might be that these other by-products are formed, but are not sufficiently extracted from the water using the AMES pre-treatment protocol, which was used as sample pre-treatment in the prior study for the artificial water sample. In order to investigate if the detected N-DBPs really originate of tryptophan or phenylalanine, SPE experiments were conducted according to the AMES pre-treatment protocol to determine the extraction recovery of the top 15 detected N-DBPs.

For the SPE experiments the same samples were used as for the aromatic amino acids labeling experiments, using the AMES SPE pre-treatment protocol for determining the recovery of the by-products. All samples for the labeling experiments were stored in the freezer at -25°C to reduce the risk of degradation. The following samples were used for the SPE experiment in duplo: UV treated + Trp + $^{14}NO_3^-$ sample, UV treated + Phe + $^{14}NO_3^-$ sample and an ultrapure water sample (blank). In order to compare the by-products in the samples after the AMES SPE pre-treatment with the samples that were not pre-treated, the final SPE extracts were diluted again with ultrapure water so that effectively no sample enrichment had occurred. In this way, a fair comparison could be made between the samples with and without sample pre-treatment in order to determine the recovery of the by-products. See attachment V for the applied conditions for the tryptophan and phenylalanine by-products SPE experiment.

3.4.1 Results of SPE experiments

Because quantification of by-products is not possible without knowing their identity, the recovery was calculated using the intensity of the by-products. The intensity of the by-products in the samples that were not pre-treated was compared with intensity of the by-products after SPE treatment, making it possible to calculate the recovery. The recoveries of the top 15 by-products of tryptophan and phenylalanine are shown in figure 12 and 13.

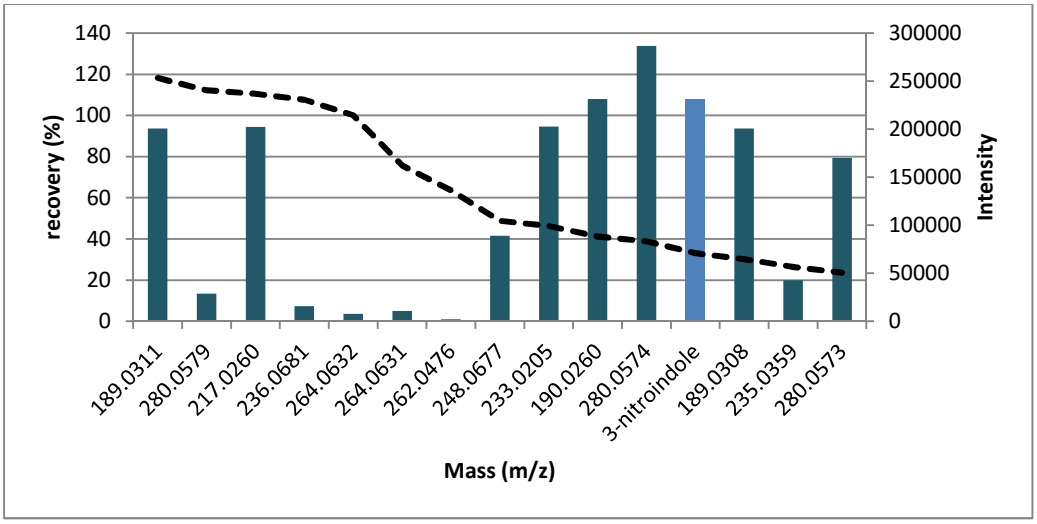


Figure 12: Recoveries of tryptophan by-products in SPE treated samples. Recoveries were calculated as the ratios of intensities of each N-DBP after and before pre-treatment. By-product 3-nitroindole indicated in blue. The dashed line represents the intensity of the by-products when no sample pre-treatment is applied (direct injection)

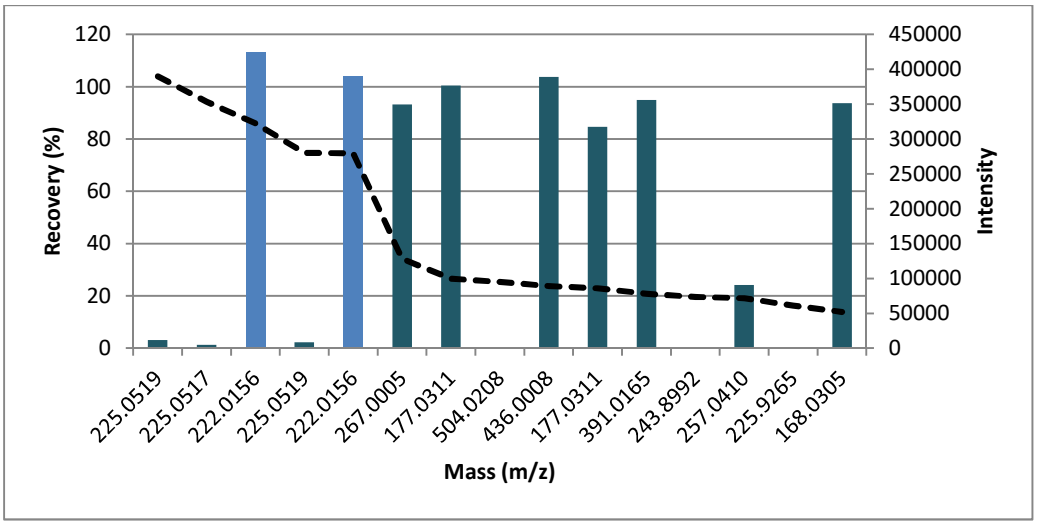


Figure 13: Recoveries of phenylalanine by-products in SPE treated samples. Recoveries were calculated as the ratios of intensities of each N-DBP after and before pre-treatment. The two isomers with m/z 222.0516 are indicated in blue. The dashed line represents the intensity of the by-products when no sample pre-treatment is applied (direct injection).

The results of the SPE recovery experiment for tryptophan by-products show that there are six by-products with poor recoveries (< 20%) while the other nine by-products have a moderate to good recoveries (41 – 134%), including 3-nitroindole (108%). Of those nine by-

products only the low intensity by-product 3-nitroindole was detected in the artificial water of the prior study. This implies that it is quite unlikely that 3-nitroindole present in the artificial water sample is a MP UV by-product of tryptophan. It is more probable that the 3-nitroindole in the artificial water sample originates from another source. Using a retrospective suspect screening it was determined that unnitrated indole is not the source.

The SPE recovery results for the phenylalanine by-products show that there are seven by-products with poor recoveries (< 25%) and eight others with good recoveries (84 – 114), including the two unidentified isomers with m/z 222.0156. Of those eight by-products, initially only the two isomers with m/z 222.0156 were detected in the artificial water sample of the prior study. However, a suspect screening conducted at a very low intensity also revealed the presence of two isomers with m/z 177.0311(of the top 15) in the artificial water samples, just barely above the detection limit (s/n > 3). This means that four of the eight by-products with a good SPE recovery are present in the artificial water sample. This result makes it quite probable that the two unidentified isomers with m/z 222.0156 present in the artificial water sample are really MP UV by-products of phenylalanine and shows that aromatic amino acids can be responsible for the formation for some of the N-DBPs formed by MP UV treatment of artificial water.

3.5 Conclusion labeling experiments

It was shown that the labeling strategy involving nitrate, aromatic amino acids and MP UV worked. Many by-products were detected using this strategy. For tryptophan, one by-product was detected in the artificial water sample of the prior study and was identified as 3-nitroindole. However, other high intensity tryptophan by-products were not detected in the artificial water, meaning that it is quite likely that the 3-nitroindole in the artificial water sample does not originate from tryptophan, but from another source.

For phenylalanine, two unidentified by-products with m/z 222.0156 (isomers) were initially detected in the artificial water sample of the prior study. Two other phenylalanine by-products with m/z 177.0311 (isomers) were detected at a very low intensity in the artificial water, meaning that four of the eight by-products with a good recovery are detected in the artificial water sample. This makes it quite probable that the two unidentified isomers with m/z 222.0156 present in the artificial water sample are really MP UV by-products of phenylalanine, and shows that aromatic amino acids are probably responsible for some of the N-DBPs formed by MP UV treatment of artificial water.

None of the tryptophan and phenylalanine by-products were detected in the full-scale MP UV water treatment sample of the prior study. This could mean that the concentration of aromatic amino acids in Lake IJssel is just too low in order to detect their by-products. This is supported in chapter 2, in which the detected the concentrations of tryptophan and phenylalanine were 6 and 9 times lower in Lake IJssel in comparison with artificial water. This makes the detected by-products of tryptophan and phenylalanine less relevant for further research.

4 Identification, toxicological evaluation and quantification of N-DBPs

In this chapter, structure elucidation of a number of the 84 previously detected but unidentified N-DBPs was performed and the results of a toxicological evaluation which was performed on the identified N-DBPs are discussed. Furthermore an analytical target method (LC-QToF) was developed and validated for the identified N-DBPs in the current and prior study, in order to conduct and evaluate a seven months monitoring study for N-DBPs in a full-scale drinking water treatment facility using MP UV treatment.

4.1 Identification of N-DBPs

In total 14 N-DBPs of the 84 detected by-products detected in artificial water were identified in the two prior studies [13, 18]. The identification of the remaining unidentified N-DBPs remains important, because the response of the identified N-DBPs does not explain the overall AMES response detected in artificial water and full-scale MP UV treated water [18]. The identification was mainly focussed on N-DBPs that are also formed by full-scale MP UV water treatment. The N-DBPs were identified by using an in silico fragmentation tool (MetFrag), and by expert knowledge. It was assumed that all the N-DBPs consist of a benzene ring to which a nitro group is attached and could contain the following functional groups: hydroxyl, methyl, methoxyl, carboxylic acid, sulfonic acid and amine. Possible candidates were purchased and were analysed using the LC-QToF method as described in 3.2.3 and compared with the LC-QToF data from the N-DBPs in the sample. The results of the QToF analysis for the identification of N-DBPs is shown in table 21. The chromatograms and mass spectra of the identified N-DBPs are shown in attachment VI.

TABLE 21: RESULTS Q-TOF ANALYSIS OF REFERENCE STANDARDS FOR THE UNAMBIGUOUS IDENTIFICATION OF N-DBPS

Compound	CAS nr	Formula	Accurate mass [M-H]	RT sample (min)	RT reference standard (min)	RT*, MS1** and MS2 Confirmed
2-Methyl-4-nitrophenol	99-53-6	$C_7H_7O_3N$	152.0353	16.25	16.25	yes
5-Methyl-2-nitrophenol	700-38-9	$C_7H_7O_3N$	152.0353	16.23	18.88	no
4-Methyl-2-nitrophenol	119-33-5	$C_7H_7O_3N$	152.0353	16.23	18.93	no
3-Methyl-4-nitrophenol	2581-34-2	$C_7H_7O_3N$	152.0353	16.23	15.30	no
2-Methyl-3-nitrophenol	5460-31-1	$C_7H_7O_3N$	152.0353	16.23	16.19	no
4-Methyl-3-nitrophenol	2042-14-0	$C_7H_7O_3N$	152.0353	16.23	15.95	no
3-Methyl-2-nitrophenol	4920-77-8	$C_7H_7O_3N$	152.0353	16.23	16.43	no
2-Methyl-5-nitrophenol	5428-54-6	$C_7H_7O_3N$	152.0353	16.23	16.83	no
2-Methoxy-4-nitrophenol	3251-56-7	$C_7H_7O_4N$	168.0302	13.40	13.43	yes
2-Methoxy-5-nitrophenol	636-93-1	$C_7H_7O_4N$	168.0302	13.40	13.04	no
2-Amino-3-nitrobenzoic acid	606-18-8	$C_7H_6N_2O_4$	181.0255	14.66	14.64	yes
6-Nitroanthranilic acid	50573-74-5	$C_7H_6N_2O_4$	181.0255	14.69	9.17	no
5-Nitroanthranilic acid	616-79-5	$C_7H_6N_2O_4$	181.0255	10.38	12.40	no
4-Nitroanthranilic acid	619-17-0	$C_7H_6N_2O_4$	181.0255	10.38	14.00	no
4-Hydroxy-3-nitrobenzenesulfonic acid	6313-34-4	$C_6H_5O_6NS$	217.9765	4.86	4.87	yes

** Retention time differs less than 0.10 min relative to the reference standard*
*** Measured accurate mass falls within a 5 ppm mass range of the theoretical mass*

Four N-DBPs were unambiguously identified by analysing the reference standards and matching the accurate mass, retention time, and MS/MS fragmentation spectra of the reference standard with the N-DBPs detected in the artificial water sample. The identities of 2-metyl-4-nitrophenol, 2-methoxy-4-nitrophenol, 2-amino-3-nitrobenzoic acid and 4-hydroxy-3-nitrobenzenesulfonic acid were thus confirmed.

4.2 Toxicological evaluation of the identified N-DBPs

Five N-DBPs were newly identified (one in chapter 3 and four in chapter 4) for which it is now possible to conduct a toxicity evaluation. The toxicity evaluation was performed by Dr. Kirsten Baken of KWR Watercycle Research Institute, see attachment VII t/m IX. Table 22 summarizes all the findings.

TABLE 22: SUMMARY OF TOXICOLOGICAL EVALUATION OF N-DBPS OF WHICH THE IDENTITY WAS CONFIRMED

Substance	Toxicity data	Structural alerts	Read across
2-Amino-3-nitrobenzoic acid	-	mutagenicity carcinogenicity reproductive and developmental toxicity	positive for *in vitro* mutagenicity negative for *in vivo* genotoxicity negative for carcinogenicity
3-Nitroindole	-	mutagenicity skin sensitization	positive for *in vitro* mutagenicity negative for *in vivo* genotoxicity negative for carcinogenicity positive for reproductive toxicity
2-Methyl-4-nitrophenol	anti-androgenic activity vasodilation	mutagenicity carcinogenicity hemolytic anemia & hepatotoxicity *or* energy metabolism dysfunction	positive/negative for *in vitro* mutagenicity positive for carcinogenicity positive for developmental toxicity
2-Methoxy-4-nitrophenol (4-Nitroguaiacol)	-	mutagenicity carcinogenicity energy metabolism dysfunction	negative for *in vitro* mutagenicity positive for carcinogenicity positive for developmental toxicity
4-Hydroxy-3-nitrobenzenesulfonic acid	-	mutagenicity estrogen receptor binding energy metabolism dysfunction	negative for *in vitro* mutagenicity

4.2.1 Conclusion toxicological evaluation

Because limited toxicity data was available for the toxicological evaluation of the by-products identified, the prediction of DNA binding and/or genotoxic potential was based on structural characteristics. For all five N-DBPs the OECD QSAR Toolbox indicated a potential for DNA binding and mutagenicity, and potential carcinogenicity was indicated for 2-amino-3-nitrobenzoic acid, 2-methyl-4-nitrophenol and 2-methoxy-4-nitrophenol. Although these predictions may indicate mutagenic potency, it is not certain that these N-DBPS show a positive response in the Ames test. Therefore it remains uncertain if the identified N-DBPS have contributed to the overall Ames response detected in artificial water and full-scale MP UV treated water in the prior study [13]. In the end, this can be determined by applying the identified N-DBPs for Ames testing.

4.3 Development of a LC-QToF target method for N-DBPS in water

For the analytical method development and validation of the 19 identified N-DBPs in water (table 23), the LC-QTOF method which was applied for the labeling experiments (chapter 3) was used again and was modified for target method analysis. A high resolution LC-QToF method was preferred over a regular triple quadrupole LC-MS system, due to the screening and structure elucidation capabilities of the QToF mass spectrometer for unidentified N-DBPs. The data dependent acquisition of the QToF instrumental method was adapted so that the MS/MS spectra of the 19 N-DBPs were continuously acquired within a specified time window (set in the mass list), while the remaining available scan time (i.e. cycle time) was used for triggering MS/MS spectra of the highest detected peaks (maximum of eight) detected in the full scan MS mode.

The sample pre-treatment consisted of solid phase extraction according to the AMES II protocol and was used for concentrating the samples and thus increasing sensitivity for detecting N-DBPs. The AMES II protocol was also used in the two prior studies and therefore only small changes were made to the pre-treatment procedure, so that it was possible to compare results between different studies. The following modifications were made to the pre-treatment procedure: (i) sample volume was reduced from 1000 mL to 500 mL, due to sample availability and (ii) two additional internal standards were added for quantification purposes, i.e. nitrophenol-d4 and neburon. See attachment X (materials en methods) for the final sample pre-treatment procedure and LC-QToF method.

TABLE 23: CAS NUMBERS, FORMULAS, AND ACCURATE MASSES OF THE IDENTIFIED N-DBPS

Compound	CAS nr	Formula	Accurate mass [M-H$^-$]
4-Nitrophenol	100-02-7	$C_6H_5NO_3$	138.0197
2-Methyl-4-nitrophenol	99-53-6	$C_7H_7O_3N$	152.0353
4-Nitrocatechol	3316-09-4	$C_6H_5NO_4$	154.0146
4-Nitro-1,3-benzenediol	3163-07-3	$C_6H_5O_4N$	154.0146
2-Nitrohydroquinone	16090-33-8	$C_6H_5O_4N$	154.0146
3-Nitroindole	4770-03-0	$C_8H_6N_2O_2$	161.0357
2-Methoxy-4-nitrophenol	3251-56-7	$C_7H_7O_4N$	168.0302
2-Amino-3-nitrobenzoic acid	606-18-8	$C_7H_6N_2O_4$	181.0255
2-Hydroxy-5-nitrobenzoic acid	96-97-9	$C_7H_5NO_5$	182.0095
4-Hydroxy-3-nitrobenzoic acid	616-82-0	$C_7H_5NO_5$	182.0095
2-Hydroxy-3-nitrobenzoic acid	85-38-1	$C_7H_5NO_5$	182.0095
2,4-Dinitrophenol	51-28-5	$C_6H_4N_2O_5$	183.0047
5-Nitrovanillin	6635-20-7	$C_8H_7NO_5$	196.0252
4-Nitrobenzenesulfonic acid	138-42-1	$C_6H_5NO_5S$	201.9816
4-Nitrophthalic acid	610-27-5	$C_8H_5NO_6$	210.0044
2-Methoxy-4,6-dinitrophenol	4097-63-6	$C_7H_6N_2O_6$	213.0153
4-Hydroxy-3-nitrobenzenesulfonic acid	6313-34-4	$C_6H_5O_6NS$	217.9765
3,5-Dinitrosalicylic acid	609-99-4	$C_7H_4N_2O_7$	226.9946
Dinoterb	1420-07-1	$C_{10}H_{12}O_5N_2$	239.0673

4.4 Method validation in ultrapure water

The developed LC-QToF target method for the determination of 19 N-DBPs in water was validated in ultrapure water first. The validation was performed including SPE sample pre-treatment (n=5). The limit of detection (LOD), limit of quantification (LOQ), repeatability, recovery and instrumental repeatability were determined for all N-DBPs. The instrumental repeatability was determined by analysing a calibration standard in eightfold (no SPE). The validation results are shown in table 24.

TABLE 24: SPE VALIDATION RESULTS (N=5) FOR N-DBPS IN ULTRAPURE WATER AND INSTRUMENTAL REPEATABILITY RESULTS

Compound	LOD (ng/L)	LOQ (ng/L)	Recovery 25 ng/L (%)	Repeatability 25 ng/L (%)	Instrumental repeatability 25 ng/L (%) (n=8)
2-Methoxy-4,6-dinitrophenol	0.26	1.0	105.2	2.4	1.6
4-Nitrophenol	1.7	5.0	91.9	3.0	2.0
4-Nitrocatechol	0.89	2.0	82.6*	6.2*	2.5
2-Hydroxy-5-nitrobenzoic acid	0.65	2.0	92.3	0.9	1.4
5-Nitrovanillin	0.36	1.0	100.9	6.5	2.8
4-Nitrophthalic acid	1.01	10	119.5	7.9	1.7
2,4-Dinitrophenol	0.39	1.0	95.5	3.1	2.5
4-Nitro-1,3-benzenediol	0.61	2.0	94.1	2.7	1.7
2-Nitrohydroquinone	1.95	5.0	71.7*	14.5*	3.5
4-Nitrobenzenesulfonic acid	0.69	3.0	24.2	14.1	1.5
4-Hydroxy-3-nitrobenzoic acid	0.42	2.0	109.5	1.9	1.3
2-Hydroxy-3-nitrobenzoic acid	0.66	2.0	88.8	0.9	1.7
Dinoterb	0.64	2.0	94.1	4.4	2.0
3,5-Dinitrosalicylic acid	0.39	3.0	45.2	13.6	2.9
2-Methyl-4-nitrophenol	0.06	1.0	92.1	1.9	1.6
2-Methoxy-4-nitrophenol	1.33	3.0	95.2	2.7	2.5
4-Hydroxy-3-nitrobenzenesulfonic acid	0.43	2.0	26.2	6.2	1.6
2-Amino-3-nitrobenzoic acid	0.84	2.0	92.1	2.2	2.2
3-Nitroindole	0.33	2.0	90.2	1.7	1.2

** n=4*

Satisfactory LOD and LOQ results were obtained for all N-DBPs in ultrapure water, resulting in a LOQ between 1.0 – 10 ng/L in ultrapure water. The recoveries for most N-DBPs are between 71.1 – 119.5 % and are within acceptable range, except for 4-nitrobenzenesulfonic acid, 4-hydroxy-3-nitrobenzenesulfonic acid and 3,5-dinitrosalicylic acid which have a recovery of 24.2, 26.2 and 45.2%, respectively. The repeatability for all compounds is lower than 15% and is satisfactory. The instrumental repeatability is good and for all N-DBPs lower than 3.6%. The low recovery for 4-nitrobenzenesulfonic acid and 4-hydroxy-3-nitrobenzenesulfonic acid was expected in advance, due to hydrophilic nature (logP = -1.35 and -1.25) of both compounds and therefore limited retention was expected with the applied SPE cartridge. In the end, good validation results are obtained for ultrapure water which can be used to compare to the validation results in drinking- and surface water (4.4.1).

4.4.1 Method validation for drinking- and surface water

For the monitoring study of Heemskerk full-scale drinking water treatment facility the LC-QToF method was also validated for drinking- and surface water. For the preparation of validation samples, drinking water was obtained from the tap at KWR and surface water was obtained from the Lekkanaal. The limit of detection (LOD), limit of quantification (LOQ), repeatability and recovery were determined for all N-DBPs in both drinking- and surface water. The validation results are shown in table 25 and 26.

TABLE 25: SPE VALIDATION RESULTS (N=5) FOR N-DBPS IN DRINKING WATER

Compound	LOD	LOQ	Recovery 25 ng/L	Repeatability 25 ng/L
	(ng/L)	(ng/L)	(%)	(%)
2-Methoxy-4,6-dinitrophenol	0.29	1.0	74.4	1.7
4-Nitrophenol	0.79	5.0	91.1	7.4
4-Nitrocatechol	0.18	2.0	87.2	6.6
2-Hydroxy-5-nitrobenzoic acid	0.50	2.0	93.1	0.7
5-Nitrovanillin	0.38	1.0	112.2	3.8
4-Nitrophthalic acid	5.9	10	275.0	3.0
2,4-Dinitrophenol	0.27	1.0	132.9	6.9
4-Nitro-1,3-benzenediol	0.45	2.0	86.6	5.0
2-Nitrohydroquinone	2.6	5.0	78.0	3.8
4-Nitrobenzenesulfonic acid	1.2	3.0	80.9	19.2
4-Hydroxy-3-nitrobenzoic acid	0.64	2.0	43.4	2.5
2-Hydroxy-3-nitrobenzoic acid	0.19	2.0	80.2	2.3
Dinoterb	0.44	2.0	109.3	2.8
3,5-Dinitrosalicylic acid	1.33	3.0	63.8*	18*
2-Methyl-4-nitrophenol	0.49	1.0	137.7	6.0
2-Methoxy-4-nitrophenol	0.40	3.0	107.0	3.2
4-Hydroxy-3-nitrobenzenesulfonic acid	0.31	2.0	58.5	13.8
2-Amino-3-nitrobenzoic acid	0.47	2.0	48.3	3.1
3-Nitroindole	0.48	2.0	130.0	4.5

** n=4*

TABLE 26: SPE VALIDATION RESULTS (N=5) FOR N-DBPS IN SURFACE WATER

Compound	LOD	LOQ	Recovery 25 ng/L	Repeatability 25 ng/L
	(ng/L)	(ng/L)	(%)	(%)
2-Methoxy-4,6-dinitrophenol	*	1.0	87.4	1.2
4-Nitrophenol	*	5.0	111.4	5.3
4-Nitrocatechol	*	2.0	91.9	2.8
2-Hydroxy-5-nitrobenzoic acid	*	2.0	114.8	2.3
5-Nitrovanillin	*	1.0	116.0	2.0
4-Nitrophthalic acid	*	10	245.1	2.5
2,4-Dinitrophenol	*	1.0	137.6	2.0
4-Nitro-1,3-benzenediol	0.66	2.0	99.2	4.1
2-Nitrohydroquinone	*	5.0	87.9	6.4
4-Nitrobenzenesulfonic acid	*	3.0	80.2	5.4
4-Hydroxy-3-nitrobenzoic acid	*	2.0	57.1	2.4

2-Hydroxy-3-nitrobenzoic acid	*	2.0	95.0	0.7
Dinoterb	0.99	2.0	124.7	3.2
3,5-Dinitrosalicylic acid	*	3.0	71.8**	26.4**
2-Methyl-4-nitrophenol	*	1.0	154.8	3.7
2-Methoxy-4-nitrophenol	*	3.0	118.7	4.0
4-Hydroxy-3-nitrobenzenesulfonic acid	0.69	2.0	43.2	12.4
2-Amino-3-nitrobenzoic acid	0.63	2.0	62.5	2.0
3-Nitroindole	0.49	2.0	134.0	3.7

** Due to the presence of the N-DBPs in surface water, it is not possible to determine all validation characteristics. The validation results of drinking water can then be used as reference.*
*** n=4*

Satisfactory LOD and LOQ results were obtained for the N-DBPs in drinking- and surface water, which results in LOQs between 1.0 – 10 ng/L. The recoveries in drinking and surface water for almost all N-DBPs are between 43.2 – 154.8 % and are within acceptable range, excluding 4-nitrophthalic acid which has a recovery of 275.0 and 245.1 for surface and drinking water. The repeatability for almost all compounds is lower than 20% and is satisfactory. Only the repeatability for 3,5-dinitrosalicylic acid is relatively high, and was determined at 26.4%. The high recovery for 4-nitrophthalic acid is was not observed with the validation in ultrapure water and might therefore be caused by severe ion enhancement in the mass spectrometer. This can possibly solved by using an isotopically labeled internal standard, although it is not commercially available for 4-nitrophthalic acid.

The results of the validation of N-DBPs in drinking- and surface water is satisfactory and show that the LC-QToF target method can be applied for the monitoring study of a full-scale drinking water treatment facility.

4.5 Monitoring study full-scale drinking water treatment facility

For the monitoring of N-DBPs formed by MP UV treatment, the Heemskerk full-scale treatment facility was selected (PWN, Heemskerk, The Netherlands). In the prior study, 22 N-DBPs were detected after MP UV/H_2O_2 treatment in this facility. At this facility surface water from Lake IJssel is treated by applying advanced oxidation and adds H_2O_2 before MP UV treatment. After MP UV/H_2O_2 treatment, granulated active carbon (GAC) filtration is applied for quenching the excess of H_2O_2 and is followed by infiltration into the dunes. Artificial dune recharge water is reclaimed after at least > 21 days, which is first aerated and filtrated (i.e. rapid sand filtration) before drinking water distribution.

For the monitoring study three types of samples were selected:

- Influent MP UV/H_2O_2 treatment
- Effluent MP UV/H_2O_2 treatment (before GAC filtration)
- Aerated reclaimed dune infiltration water

The sampling campaign was conducted from July 2016 until January 2017 with a sampling frequency of once a month. In total seven samples were collected per sampling point. The samples were collected and stored (1-5 °C) in 1L glass sample bottles upon arrival at KWR. Then the samples were transferred to 1L HDPE bottles and stored in a freezer (-25 °C) until LC-QToF analysis.

4.6 Results monitoring study

The results of the monitoring study for the influent MP UV/H_2O_2 and effluent MP UV/H_2O_2 samples are presented in table 27 and 28. No N-DBPs were detected in the aerated reclaimed dune infiltration water sample and the results for this sample can be found in attachment XI.

TABLE 27: RESULTS MONITORING STUDY - INFLUENT MP UV/H_2O_2 TREATMENT SAMPLE

Compound	Influent MP UV/H_2O_2 treatment						
	27/07/16 ng/L	05/09/16 ng/L	21/09/16 ng/L	20/10/16 ng/L	21/11/16 ng/L	14/12/16 ng/L	11/01/17 ng/L
2-Methoxy-4,6-dinitrophenol	< 1.0	5.9	6.8	< 1.0	< 1.0	1.0	1.2
4-Nitrophenol	< 5.0	< 5.0	< 5.0	< 5.0	5.8	5.5	8.8
4-Nitrocatechol	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Hydroxy-5-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
5-Nitrovanillin	< 1.0	7.5	5.6	< 1.0	< 1.0	< 1.0	< 1.0
4-Nitrophthalic acid	< 10	< 10	< 10	< 10	< 10	< 10	< 10
2,4-Dinitrophenol	4.2	5.0	8.8	8.0	7.3	7.7	8.7
4-Nitro-1,3-benzenediol	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Nitrohydroquinone	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0
4-Nitrobenzenesulfonic acid	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	3.0
4-Hydroxy-3-nitrobenzoic acid	< 2.0	2.2	< 2.0	< 2.0	< 2.0	< 2.0	2.0
2-Hydroxy-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	4.3	2.2
Dinoterb	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
3,5-Dinitrosalicylic acid	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
2-Methyl-4-nitrophenol	1.4	1.6	2.4	2.5	6.5	7.1	10.2
2-Methoxy-4-nitrophenol	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
4-Hydroxy-3-nitrobenzenesulfonic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Amino-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
3-Nitroindole	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
Summed concentration	5.6	22	24	11	20	26	36

TABLE 28: RESULTS MONITORING STUDY - EFFLUENT MP UV/H_2O_2 TREATMENT SAMPLE

Compound	Effluent MP UV/H_2O_2 treatment						
	27/07/16 ng/L	05/09/16 ng/L	21/09/16 ng/L	20/10/16 ng/L	21/11/16 ng/L	14/12/16 ng/L	11/01/17 ng/L
2-Methoxy-4,6-dinitrophenol	< 1.0	2.6	6.9	1.4	22	21	41
4-Nitrophenol	< 5.0	7.2	9.2	6.6	25	29	31
4-Nitrocatechol	< 2.0	< 2.0	< 2.0	< 2.0	2.1	3.1	3.7
2-Hydroxy-5-nitrobenzoic acid	< 2.0	15	16	11	28	44	44
5-Nitrovanillin	< 1.0	< 1.0	1.4	< 1.0	< 1.0	< 1.0	< 1.0
4-Nitrophthalic acid	< 10	35	28	19	26	42	39
2,4-Dinitrophenol	4.5	1.0	1.2	< 1.0	4.7	3.1	5.6
4-Nitro-1,3-benzenediol	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Nitrohydroquinone	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0
4-Nitrobenzenesulfonic acid	< 3.0	9.7	8.1	12	28	29	40
4-Hydroxy-3-nitrobenzoic acid	< 2.0	3.5	5.8	3.6	11	16	16
2-Hydroxy-3-nitrobenzoic acid	< 2.0	2.4	< 2.0	2.3	5.6	8.8	8.6
Dinoterb	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
3,5-Dinitrosalicylic acid	< 3.0	< 3.0	< 3.0	< 3.0	3.0	5.6	4.6
2-Methyl-4-nitrophenol	1.3	< 1.0	1.2	< 1.0	1.3	2.3	2.3
2-Methoxy-4-nitrophenol	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
4-Hydroxy-3-nitrobenzenesulfonic acid	< 2.0	< 2.0	< 2.0	< 2.0	16	7.3	23
2-Amino-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	2.4	2.9
3-Nitroindole	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	2.0	2.3
Summed concentration	5.8	76	78	55	174	216	263

Fifteen of the 19 N-DBPs were detected during the monitoring study in MP UV/H_2O_2 treated water, with concentrations ranging from 1.0 to 44 ng/L. Five N-DBPs were detected at a relatively high concentrations in MP UV treated water; 2-methoxy-4,6-dinitrophenol (41 ng/L), 4-nitrophenol (31 ng/L), 2-hydroxy-5-nitrobenzoic acid (44 ng/L), 4-nitrophthalic acid (42 ng/L) and 4-nitrobenzenesulfonic acid (40 ng/L). As expected, the highest concentrations of N-DBPs were detected in the MP UV treated water, while much lower concentrations were detected in the influent sample. Three N-DBPs (5-nitrovanillin, 2,4-dintrophenol, and 2-methyl-4-nitrophenol) were detected in higher concentrations in the influent sample than in the MP UV treated sample, which suggest that these N-DBPs are probably not formed by MP UV in the full-scale treatment facility. Furthermore, there were no N-DBPs detected in the sample after dune infiltration, meaning that there are no N-DBPs present in the final drinking water.

3-Nitroindole, a by-product detected in the tryptophan labeling experiments, was also detected in the MP UV/H_2O_2 treated water, although at low concentrations of 2.0 - 2.3 ng/L in two of the seven samples. A suspect screening was performed for detection of other by-products of tryptophan, but none were found. A suspect screening was also performed for by-products of phenylalanine, but yielded no results. These results show that the N-DBPs of aromatic amino acids play a limited role in in the Heemskerk full-scale treatment facility.

In order to obtain a good overview of the detected N-DBPs in the monitoring study, the concentrations were summed per sample (see figure 14).

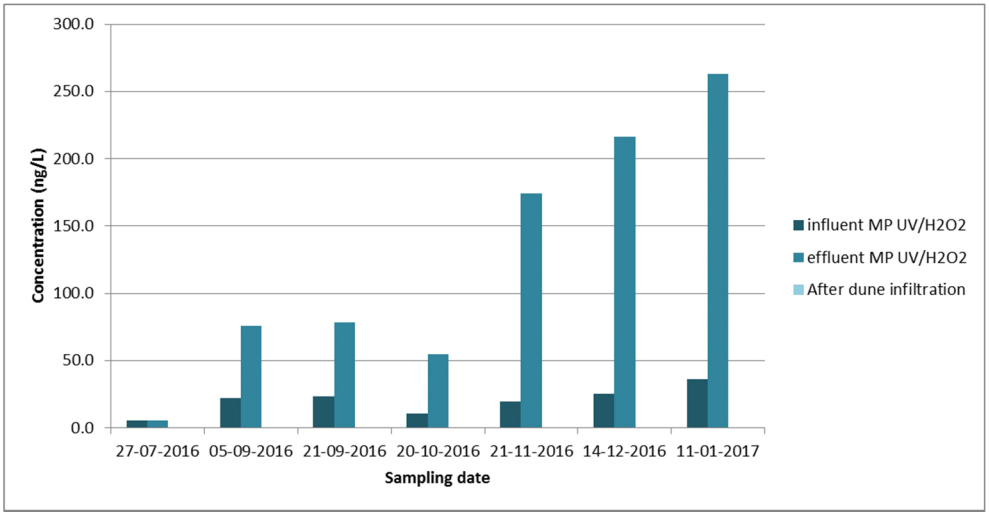


Figure 14: Summed concentrations of the detected by-products in the Heemskerk full-scale water treatment facility in influent MP UV/H_2O_2, effluent MP UV/H_2O_2 and after dune infiltration water, analysed using the QToF N-DBP target method.

A clear trend is visible when looking at the total amount of detected by-products per month. Low summed concentrations of N-DBPs were detected in July (sum is 5.8 ng/L) but increases considerably during the year, with a maximum of 263 ng/L detected in January. This trend cannot be explained by the concentration of dissolved organic carbon (DOC) or nitrate present in the water. The concentration of DOC does not fluctuate much during the year (5-7 mg/L), and the concentration of nitrate was low from August to December (< 2 mg/L) [35]. The increase in N-DBP concentration can currently not be explained and is subject for further research.

Furthermore the results show that the best period for detecting by-products is from November until January, although the period from February until June was not monitored and could possibly yield even higher concentrations.

4.7 N-DBP concentration in extracts of previous study

In the previous study [13] the Heemskerk full-scale drinking water treatment facility was also monitored and in March 2013 the same sampling points were also sampled. In the previous study the concentration of the mainly unidentified N-DBPs was estimated by using bentazon-d6 equivalents. Now a large portion of by-products is identified which are formed in a full-scale MP UV facility, it is possible to determine the actual concentrations of these by-products in these "old" extracts. The extracts were pre-treated in 2013 according to the Ames II protocol and were stored in the freezer at -25°C. The extracts were analysed with the LC-QToF target method as described in 4.3. The results of the target N-DBP analysis is shown in table 29. In the previous study a total amount of 82 ng/L bentazon-d6 equivalents were detected in influent MP UV/H_2O_2 water, whereof 47 ng/L were of identified N-DBPs.

TABLE 29: RESULTS N-DBP ANALYSIS USING THE QTOF TARGET METHOD IN EXTRACTS OF 2013

Compound	influent MP UV/H_2O_2 15-03-2013 (ng/L)	effluent MP UV/H_2O_2 15-03-2013 (ng/L)	After dune infiltration 15-03-2013 (ng/L)
2-Methoxy-4,6-dinitrophenol	< 0.5	1.4	< 0.5
4-Nitrophenol	< 3.0	4.7	< 3.0
4-Nitrocatechol	< 1.0	1.7	< 1.0
2-Hydroxy-5-nitrobenzoic acid	< 1.0	6.1	< 1.0
5-Nitrovanillin	< 0.5	0.9	< 0.5
4-Nitrophthalic acid	< 5.0	6.3	< 5.0
2,4-Dinitrophenol	0.6	< 0.5	< 0.5
4-Nitro-1,3-benzenediol	< 1.0	< 1.0	< 1.0
2-Nitrohydroquinone	< 3.0	< 3.0	< 3.0
4-nitrobenzenesulfonic acid	< 2.0	4.0	< 2.0
4-Hydroxy-3-nitrobenzoic acid	< 1.0	2.4	< 1.0
2-Hydroxy-3-nitrobenzoic acid	< 1.0	< 1.0	< 1.0
Dinoterb	< 1.0	< 1.0	< 1.0
3,5-Dinitrosalicylic acid	< 2.0	< 2.0	< 2.0
2-Methyl-4-nitrophenol	1.5	< 0.5	< 0.5
2-Methoxy-4-nitrophenol	< 2.0	< 2.0	< 2.0
4-Hydroxy-3-nitrobenzenesulfonic acid	< 1.0	< 1.0	< 1.0
2-Amino-3-nitrobenzoic acid	< 1.0	< 1.0	< 1.0
3-Nitroindole	< 1.0	< 1.0	< 1.0
Summed concentration	2.1	28	0.0

The total amount of N-DBPs detected in the MP UV/H_2O_2 treated sample (28 ng/L) is relatively low in comparison with the results of the current monitoring study and is also lower than 47 ng/L bentazon-d6 equivalents of the previous study. A possible explanation for the low concentrations in comparison with previous study could be that the concentration of some of N-DBPs declined due to break down over time, although such effects were previously not observed. Another explanation is due to the quantification with bentazon-d6, which yield a different ionisation response [36] in comparison with the individual N-DBP.

4.7.1 Conclusion

In the end it is good to finally confirm the presence of several N-DBPs in the samples of the previous study and that it is finally possible to quantify them. Because the identity and concentration is now known for many N-DBPs detected in full-scale drinking water treatment facility using MP UV, it is now possible to test the identified N-DBPs (or mixture of) in the Amest test and to gain more insight in the relation between the observed mutagenicity and the identified N-DBPs.

5 Conclusion & recommendations

In this study a LC-QToF method was developed for the determination of aromatic amino acids for which satisfactory LOD and LOQ results were obtained for both drinking- and surface water. Aromatic amino acids were detected in Lake IJssel source water and artificial water, indicating that they can be a potential source for the formation of N-DBPs.

Labeling experiments were conducted using a non-target screening LC-QToF method and it was demonstrated that the labeling strategy involving nitrate, aromatic amino acids and MP UV treatment worked and many by-products were detected using this strategy. For tryptophan, one by-product was detected in the artificial water sample and was identified as 3-nitroindole. However, other high intensity tryptophan by-products were not detected in artificial water, suggesting that 3-nitroindole in the artificial water sample does not originate from tryptophan, but originates from another source. For phenylalanine, four unidentified phenylalanine N-DBPs were detected in artificial water, for which was shown that these N-DBPs are truly MP UV by-products of phenylalanine. This demonstrates that aromatic amino acids are probably responsible for some of the N-DBPs formed by MP UV treatment of artificial water. However, none of the tryptophan and phenylalanine by-products were detected in the full-scale MP UV water treatment sample of the prior study, making the detected by-products of tryptophan and phenylalanine less relevant for further research.

Of the 84 by-products detected in artificial water (prior study), five N-DBPs were unambiguously identified during this study, bringing the total of identified N-DBPs in artificial water to 19. For all of the newly identified N-DBPs an indication for DNA binding and mutagenicity was obtained and an indication for potential carcinogenicity for 2-amino-3-nitrobenzoic acid, 2-methyl-4-nitrophenol and 2-methoxy-4-nitrophenol.

A LC-QToF target method was developed for determination of the 19 identified N-DBPs in drinking- and surface water and satisfactory LOD and LOQ results were obtained. The method was applied for a seven months monitoring study of a full-scale drinking water facility using MP UV treatment. Fifteen of the 19 known N-DBPs were detected during the monitoring study in MP UV/H_2O_2 treated water, with concentrations ranging from 1.0 to 44 ng/L. Five N-DBPs were detected at relatively high concentrations in MP UV treated water; 2-methoxy-4,6-dinitrophenol (41 ng/L), 4-nitrophenol (31 ng/L), 2-hydroxy-5-nitrobenzoic acid (44 ng/L), 4-nitrophthalic acid (42 ng/L) and 4-nitrobenzenesulfonic acid (40 ng/L). The monitory study showed a large seasonal variation in the summed concentration of detected N-DBPs (5.8 ng/L in July and 263 ng/L in January) and also showed there were no N-DBPs detected in the samples after dune infiltration. The amount of detected N-DBPs in a full-scale treatment facility demonstrates the relevance of the identified N-DBPs and the developed N-DBP target method.

Because the identity and concentration is now known for many N-DBPs detected in a full-scale drinking water treatment facility using MP UV; it is recommended to conduct N-DBP target analysis of MP UV treated water of a drinking water facility, and then to perform Ames testing with and without (mixtures of) identified N-DBPs at their detected concentration, in order to gain more insight between the relation of identified N-DBPs and observed mutagenicity in MP UV treated water.

Other recommendations:

- Conducting a monitoring study of a year for a full-scale drinking water treatment facility using MP UV in order to obtain a good overview of the concentrations of the N-DBPs per month and to find an explanation between the concentration differences per month
- Performing labeling experiments in MP UV influent water, in order to detect other types of N-DBPs, such as N-DBPs formed from organic micro pollutants.

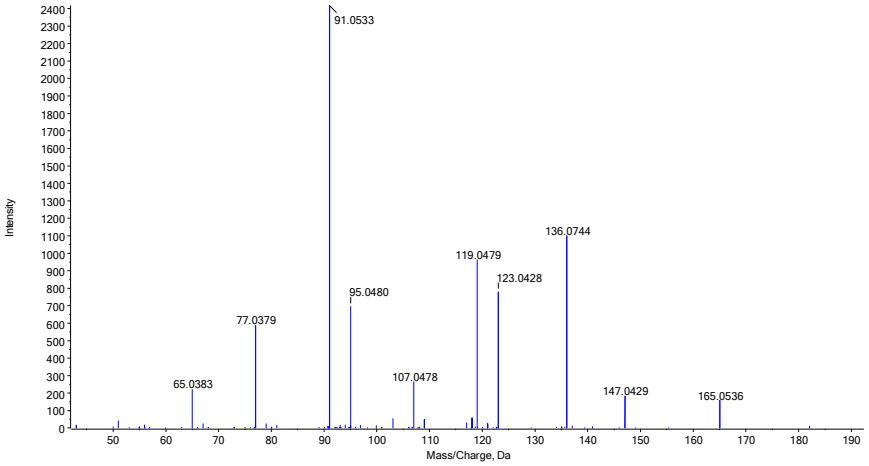
6 References

1. Kruithof, J.C. and J.C. Schippers, *Drinking water treatment in the 1990s.* Water Supply, 1994. **12**(1-2): p. SS51-SS54.
2. Homem, V. and L. Santos, *Degradation and removal methods of antibiotics from aqueous matrices - A review.* Journal of Environmental Management, 2011. **92**(10): p. 2304-2347.
3. Lekkerkerker-Teunissen, K., et al., *Transformation of atrazine, carbamazepine, diclofenac and sulfamethoxazole by low and medium pressure UV and UV/H2O2 treatment.* Separation and Purification Technology, 2012. **96**: p. 33-43.
4. Yang, W., H. Zhou, and N. Cicek, *Treatment of organic micropollutants in water and wastewater by UV-based processes: A literature review.* Critical Reviews in Environmental Science and Technology, 2014. **44**(13): p. 1443-1476.
5. Heringa, M.B., et al., *Formation and removal of genotoxic activity during UV/H2O2–GAC treatment of drinking water.* Water Res, 2011. **45**(1): p. 366-374.
6. Hofman-Caris, R.C.H.M., et al., *Influence of process conditions and water quality on the formation of mutagenic byproducts in UV/H2O2 processes.* Water Research, 2015. **74**: p. 191-202.
7. Martijn, A.J. and J.C. Kruithof, *UV and UV/H2O2 Treatment: The Silver Bullet for By-product and Genotoxicity Formation in Water Production.* Ozone: Science & Engineering, 2012. **34**(2): p. 92-100.
8. Richardson, S.D. and S.Y. Kimura, *Water Analysis: Emerging Contaminants and Current Issues.* Analytical Chemistry, 2016. **88**(1): p. 546-582.
9. Richardson, S.D., et al., *Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: A review and roadmap for research.* Mutation Research - Reviews in Mutation Research, 2007. **636**(1-3): p. 178-242.
10. Bond, T., et al., *Occurrence and control of nitrogenous disinfection by-products in drinking water - A review.* Water Research, 2011. **45**(15): p. 4341-4354.
11. Bond, T., M.R. Templeton, and N. Graham, *Precursors of nitrogenous disinfection by-products in drinking water--A critical review and analysis.* Journal of Hazardous Materials, 2012. **235-236**: p. 1-16.
12. Shah, A.D. and W.A. Mitch, *Halonitroalkanes, halonitriles, haloamides, and N-nitrosamines: A critical review of nitrogenous disinfection byproduct formation pathways.* Environmental Science and Technology, 2012. **46**(1): p. 119-131.
13. Kolkman, A., et al., *Tracing Nitrogenous Disinfection Byproducts after Medium Pressure UV Water Treatment by Stable Isotope Labeling and High Resolution Mass Spectrometry.* Environ Sci Technol, 2015. **49**(7): p. 4458-4465.
14. Thorn, K.A. and L.G. Cox, *Ultraviolet irradiation effects incorporation of nitrate and nitrite nitrogen into aquatic natural organic matter.* Journal of Environmental Quality, 2012. **41**(3): p. 865-881.
15. Goldstein, S. and J. Rabani, *Mechanism of nitrite formation by nitrate photolysis in aqueous solutions: The role of peroxynitrite, nitrogen dioxide, and hydroxyl radical.* Journal of the American Chemical Society, 2007. **129**(34): p. 10597-10601.
16. Mack, J. and J.R. Bolton, *Photochemistry of nitrite and nitrate in aqueous solution: A review.* Journal of Photochemistry and Photobiology A: Chemistry, 1999. **128**(1-3): p. 1-13.
17. Monarca, S., et al., *The influence of different disinfectants on mutagenicity and toxicity of urban wastewater.* Water Research, 2000. **34**(17): p. 4261-4269.
18. Vughs, D., et al., *Application of effect-directed analysis to identify mutagenic nitrogenous disinfection by-products of advanced oxidation drinking water treatment.* Environmental Science and Pollution Research, 2016.

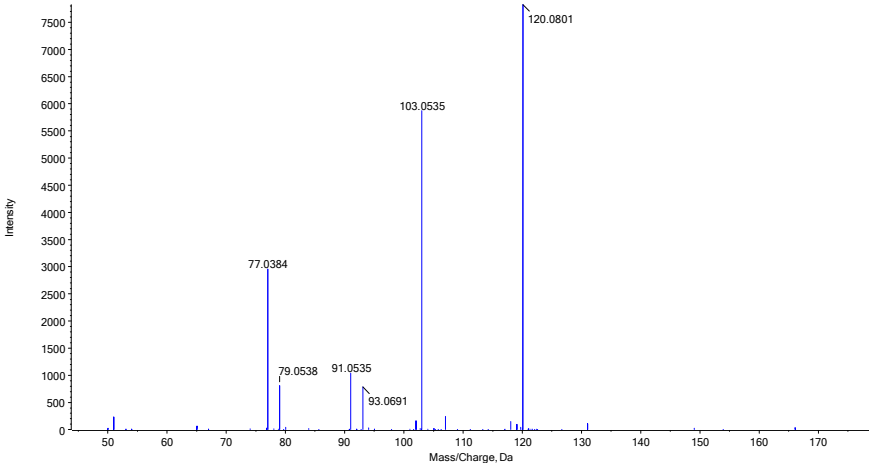
19. Dotson, A. and P. Westerhoff, *Occurrence and removal of amino acids during drinking water treatment.* Journal (American Water Works Association), 2009. **101**(9): p. 101-115.
20. How, Z.T., et al., *Analysis of free amino acids in natural waters by liquid chromatography-tandem mass spectrometry.* Journal of Chromatography A, 2014. **1370**: p. 135-146.
21. Suzuki, J., et al., *Formation of mutagens by photolysis of amino acids in neutral aqueous solution containing nitrite or nitrate ion.* Chemosphere, 1985. **14**(5): p. 493-500.
22. Aljammaz, M., *Identification of Genotoxic Reaction Products Formed Upon Drinking Water Treatment by Medium Pressure Ultra Violet Light with Hydrogen Peroxide (MP UV/H2O2).* PWN, 2014.
23. Wu, F.C. and E. Tanoue, *Tryptophan in the sediments of lakes from Southwestern China Plateau.* Chemical Geology, 2002. **184**(1): p. 139-149.
24. Desiderio, C., et al., *Capillary electrophoresis-mass spectrometry for the analysis of amino acids.* Journal of Separation Science, 2010. **33**(16): p. 2385-2393.
25. Martens, D.A. and K.L. Loeffelmann, *Soil Amino Acid Composition Quantified by Acid Hydrolysis and Anion Chromatography−Pulsed Amperometry.* Journal of Agricultural and Food Chemistry, 2003. **51**(22): p. 6521-6529.
26. Piraud, M., et al., *Ion-pairing reversed-phase liquid chromatography/electrospray ionization mass spectrometric analysis of 76 underivatized amino acids of biological interest: A new tool for the diagnosis of inherited disorders of amino acid metabolism.* Rapid Communications in Mass Spectrometry, 2005. **19**(12): p. 1587-1602.
27. Gao, J., et al., *Robust analysis of underivatized free amino acids in soil by hydrophilic interaction liquid chromatography coupled with electrospray tandem mass spectrometry.* Journal of Chromatography A, 2016. **1449**: p. 78-88.
28. Harris, B.D., et al., *Characterization of Disinfection By-Products from Chromatographically Isolated NOM through High-Resolution Mass Spectrometry.* Environmental Science & Technology, 2015. **49**(24): p. 14239-14248.
29. Bolton, J.R. and K.G. Linden, *Standardization of methods for fluence (UV Dose) determination in bench-scale UV experiments.* Journal of Environmental Engineering, 2003. **129**(3): p. 209-215.
30. Nuriel, T., A. Hansler, and S.S. Gross, *Protein nitrotryptophan: Formation, significance and identification.* Journal of Proteomics, 2011. **74**(11): p. 2300-2312.
31. Schymanski, E.L., et al., *Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence.* Environ Sci Technol, 2014.
32. Ruttkies, C., et al., *MetFrag relaunched: incorporating strategies beyond in silico fragmentation.* Journal of Cheminformatics, 2016. **8**(1): p. 3.
33. Tanii, H. and K. Hashimoto, *Studies on the mechanism of acute toxicity of nitriles in mice.* Archives of Toxicology, 1984. **55**(1): p. 47-54.
34. Ma, X., et al., *Identification and characterization of phenylacetonitrile as a nitrogenous disinfection byproduct derived from chlorination of phenylalanine in drinking water.* Water Research, 2016. **102**(Supplement C): p. 202-210.
35. *RIWA-Rijn annual report* 2016.
36. van Leerdam, J.A., et al., *Exploring the boundaries of non-target screening with Liquid Chromatography coupled to ESI-MS* 2017, KWR: Nieuwegein. p. 89.
37. Taneda, S., et al., *Estrogenic and Anti-androgenic Activity of Nitrophenols in Diesel Exhaust Particles (DEP).* Biological and Pharmaceutical Bulletin, 2004. **27**(6): p. 835-837.
38. Mori, Y., et al., *Isolation of Nitrophenols from Diesel Exhaust Particles (DEP) as Vasodilatation Compounds.* Biological and Pharmaceutical Bulletin, 2003. **26**(3): p. 394-395.

Attachment I MS/MS spectra of aromatic amino acids

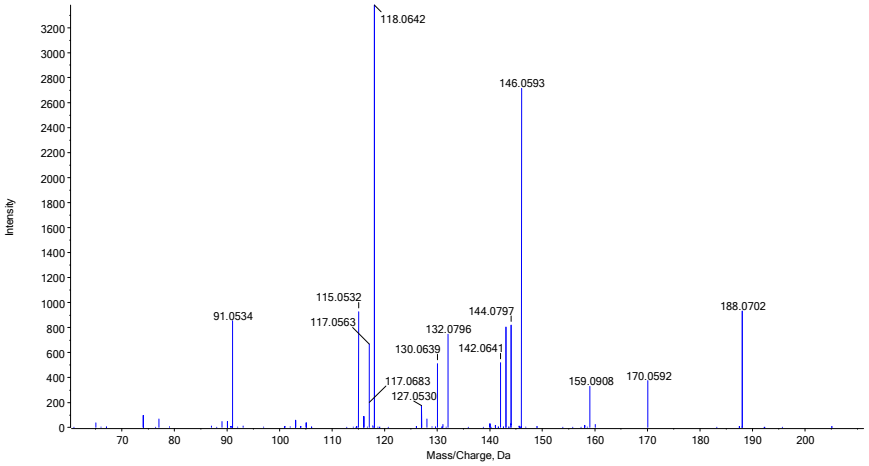
TYROSINE MS/MS SPECTRUM (20, 35 AND 50 EV, AVERAGED)



PHENYLALANINE MS/MS SPECTRUM (20, 35 AND 50 EV, AVERAGED)



TRYPTOPHAN MS/MS SPECTRUM (20, 35 AND 50 EV, AVERAGED)



Attachment II Materials & Method labeling experiments

II.1 Chemicals

All solvents used were of analytical grade quality. Acetonitrile and methanol (ultra gradient HPLC grade) was obtained from Avantor Performance Materials B.V. (Deventer, the Netherlands). Formic acid (HPLC quality), potassium nitrate ($K^{14}NO_3$) and $K^{15}NO_3$ (98% enrichment) were purchased from Sigma-Aldrich (Steinheim, Germany). The aromatic amino acids reference standards; L-tyrosine, L-phenylalanine and L-tryptophan were acquired from Sigma-Aldrich. The isotopically labeled internal standards; L-Tyrosine-d4, L-Phenylalanine-d5 and L-Tryptophan-d5 were purchased from Toronto Research Chemicals (Toronto, Canada). L-tryptophan-$^{13}C_{11}$ was obtained from Cambridge Isotope Laboratories. Ultrapure water was obtained by purifying demineralized water in an Elga Purelab Chorus ultrapure water system. (High Wycombe, United Kingdom). Pony Lake NOM was obtained from International Humic Substances Society. Stock solutions of the reference and internal standards were prepared in methanol and ultrapure water (20/80% v/v) at a concentration of 100 and 50 mg/L, respectively. Stock solutions were stored at -25 °C.

II.2 LC-QTOF analysis

HPLC settings (Shimadzu Nexera; LC-30AD, SIL-30AC, CTO-20AC):

- Column: Waters Xbridge C18 XP, 2.1 x 150 mm, 2.5 µm
- Mobile phase A: ultrapure water + 0.05% formic acid
  Mobile phase B: acetonitrile + 0.05% formic acid
- Gradient: 4% B held for 1 min. Then linear from 4% to 100% B in 40 min. Held at 100% B for 5 min. Then return to initial conditions in 1 min and held for 6 min.
- Flow: 300 µL/min
- Injection volume: 50 µL
- Column oven temperature: 25°C

QToF-MS settings (AB SCIEX TripleTOF 5600+):

- Resolution: > 30.000 @ m/z 400 (MS and MS/MS mode)
- Mass accuracy < 5 ppm
- Mass range Full scan: 65-800 Da
- Mass range MS2 scan: 40-800 Da
- Ionisation: positive and negative mode
- Source: electrospray (ESI)
- TurboIonSpray heater: 500°C
- Ion Spray Voltage: 5000 and 3000 volt for positive and negative mode
- Curtain gas: 25 psi
- Gas 1: 40 psi
- Gas 2: 50 psi
- Divert valve: 3.0 min
- Collision energy: 20, 35, 50 eV (averaged)
- Data dependant MS/MS scans: 8 per cycle (50ms), threshold 100 counts and dynamic background subtraction.

Attachment III Screening list of nitrogenous disinfection by-products

Accurate mass	Retention time* (min)	Positive/ Negative mode	Most probable formula	Detected in fractionated samples	Detected in unfractionated samples	Compound
138.0199	11.79	Negative	$C_6H_5O_3N$	yes	yes	4-nitrophenol
152.0359 (1)	10.10	negative	$C_7H_7O_3N$	yes	yes	
152.0361 (2)	15.18	negative	$C_7H_7O_3N$	yes	yes	
153.0073	10.54	negative		yes	yes	
154.0148 (1)	9.23	negative	$C_6H_5O_4N$	yes	yes	4-nitrocatechol
154.0148 (2)	10.06	negative	$C_6H_5O_4N$	yes	yes	2-nitrohydroquinone
154.0148 (3)	**	negative	$C_6H_5O_4N$	yes?	yes	4-nitro-1,3-benzenediol
161.0364	15.06	negative	$C_8H_6O_2N_2$	yes	no	
168.0306	12.23	negative	$C_7H_7O_4N$	yes	yes	
179.0101	10.62	negative	$C_7H_4O_4N_2$	yes	yes	
181.0263	13.59	negative	$C_7H_7O_4N_2$	yes	yes	
182.0098 (2)	10.21	negative	$C_7H_5O_5N$	yes	yes	4-hydroxy-3-nitrobenzoic acid
182.0098 (3)	12.81	negative	$C_7H_5O_5N$	yes	yes	2-hydroxy-5-nitrobenzoic acid
182.0098 (4)	**	negative	$C_7H_5O_5N$	yes	yes	2-hydroxy-3-nitrobenzoic acid
183.0055	14.03	negative	$C_6H_4O_5N_2$	yes	yes	2,4-dinitrophenol
195.0055	10.97	negative	$C_7H_4O_5N_2$	yes	yes	
196.0254	7.68	negative	$C_8H_7O_5N$	yes	yes	
196.0258 (1)	10.10	negative		yes	yes	
196.0259 (2)	11.51	negative		yes	yes	5-nitrovanillin
196.0260 (3)	12.89	negative		yes	yes	
198.0047	7.88	negative	$C_7H_5O_6N$	yes	yes	
201.9818	5.00	negative	$C_6H_5NO_5S$	yes	yes	4-nitrobenzenesulfonic acid
208.0255	11.77	negative	$C_9H_7O_5N$	yes	yes	
210.0048 (1)	7.48	negative	$C_8H_5O_6N$	yes	yes	4-nitrophthalic acid
210.0048 (2)	8.90	negative	$C_8H_5O_6N$	yes	yes	
211.0004	12.90	negative	$C_7H_4O_6N_2$	yes	yes	
212.0204	10.55	negative	$C_8H_7O_6N$	yes	yes	Structural isomer of 5-Hydroxy-4-methoxy-2-nitrobenzoic acid
213.0154	14.58	negative	$C_7H_6O_6N_2$	yes	yes	2-methoxy-4,6-dinitrophenol
216.0268	11.53	positive		yes	yes	
216.0268	6.81	positive		yes	yes	
217.9767	5.73	negative		yes	yes	
222.0165	16.75	negative	$C_8H_5O_5N_3$	yes	yes	
223.0005	10.63	negative	$C_8H_4O_6N_2$	yes	yes	
223.9957	13.52	negative	$C_7H_3O_6N_3$	yes	yes	
225.9994 (1)	6.37	negative	$C_8H_5O_7N$	yes	yes	
225.9996 (2)	7.39	negative	$C_8H_5O_7N$	yes	yes	
226.9948	13.28	negative	$C_7H_4O_7N_2$	yes	yes	3,5-dinitrosalicylic acid
238.0726	17.84	negative	$C_{11}H_{13}O_5N$	yes	yes	
239.0677	25.87	negative	$C_{10}H_{12}O_5N_2$	yes	yes	dinoterb
240.0151	7.67	negative	$C_9H_7O_7N$	yes	yes	
243.9923	5.20	negative		no	no	
252.0153	8.74	negative	$C_{10}H_7O_7N$	yes	yes	
254.0314	10.64	negative	$C_{10}H_9O_7N$	yes	yes	
266.1037	22.94	negative	$C_{13}H_{17}O_5N$	yes	yes	
270.0755 (1)	11.32	negative		yes	yes	
270.0755 (2)	11.61	negative		yes	yes	
272.0891	11.53	positive		yes	yes	

274.0935	13.89	negative	$C_{11}H_{17}O_7N$	yes	yes	
298.0940	13.10	negative	$C_{13}H_{17}O_7N$	yes	yes	
316.1413 (1)	18.35	negative	$C_{14}H_{23}O_7N$	yes	yes	
316.1417 (2)	19.00	negative	$C_{14}H_{23}O_7N$	yes	yes	
316.1417 (3)	19.74	negative	$C_{14}H_{23}O_7N$	yes	yes	
318.1550 (1)	19.80	positive	$C_{14}H_{23}O_7N$	yes	yes	
318.1550 (2)	18.42	positive	$C_{14}H_{23}O_7N$	yes	yes	
319.0209	16.16	negative		yes	yes	
331.0554 (1)	12.39	negative		yes	yes	
331.0554 (2)	12.57	negative		yes	yes	
331.0554 (3)	13.08	negative		yes	yes	
335.0534	18.95	negative		yes	yes	
340.1388	27.36	positive	$C_{16}H_{21}O_7N$	yes	yes	
340.1388	26.99	positive	$C_{16}H_{21}O_7N$	yes	yes	
340.1388	28.08	positive	$C_{16}H_{21}O_7N$	yes	yes	
372.1491	24.10	negative		yes	yes	
386.1096 (1)	10.94	negative		yes	yes	
386.1096 (2)	11.84	negative		yes	yes	
386.1653	26.10	negative		yes	yes	
387.1091	20.87	negative		yes	no	
400.1262 (1)	11.24	negative		yes	yes	
400.1262 (2)	12.19	negative		yes	yes	
400.1802	28.34	negative		yes	yes	
404.2758 (1)	17.34	positive	$C_{19}H_{37}O_8N_3$	yes	yes	
404.2758 (2)	17.54	positive	$C_{19}H_{37}O_8N_3$	yes	no	
404.2758 (3)	17.72	positive	$C_{19}H_{37}O_8N_3$	yes	no	
404.2758 (4)	18.02	positive	$C_{19}H_{37}O_8N_3$	yes	yes	
404.2758 (5)	18.36	positive	$C_{19}H_{37}O_8N_3$	yes	yes	
404.2758 (6)	18.86	positive	$C_{19}H_{37}O_8N_3$	no	no	
404.2758 (7)	20.40	positive	$C_{19}H_{37}O_8N_3$	yes	no	
408.1308 (1)	11.70	negative		yes	yes	
408.1310 (2)	13.21	negative		yes	yes	
410.1468	14.46	negative		yes	yes	
414.1418 (1)	11.48	negative		yes	yes	
414.1421 (2)	12.57	negative		yes	yes	
428.1213	10.26	negative		yes	yes	
442.1365 (1)	9.49	negative		yes	yes	
442.1365 (2)	10.19	negative		yes	yes	
447.1455	24.00	negative		no	no	
474.1501	11.66	positive		yes	yes	
448.2675 (1)	17.32	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (2)	17.52	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (3)	17.70	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (4)	17.99	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (5)	18.35	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (6)	18.86	negative	$C_{20}H_{39}O_8N_3$	no	no	
448.2675 (7)	20.40	negative	$C_{20}H_{39}O_8N_3$	no	no	
452.1203 (1)	11.12	negative		yes	yes	
452.1210 (2)	12.44	negative		yes	yes	

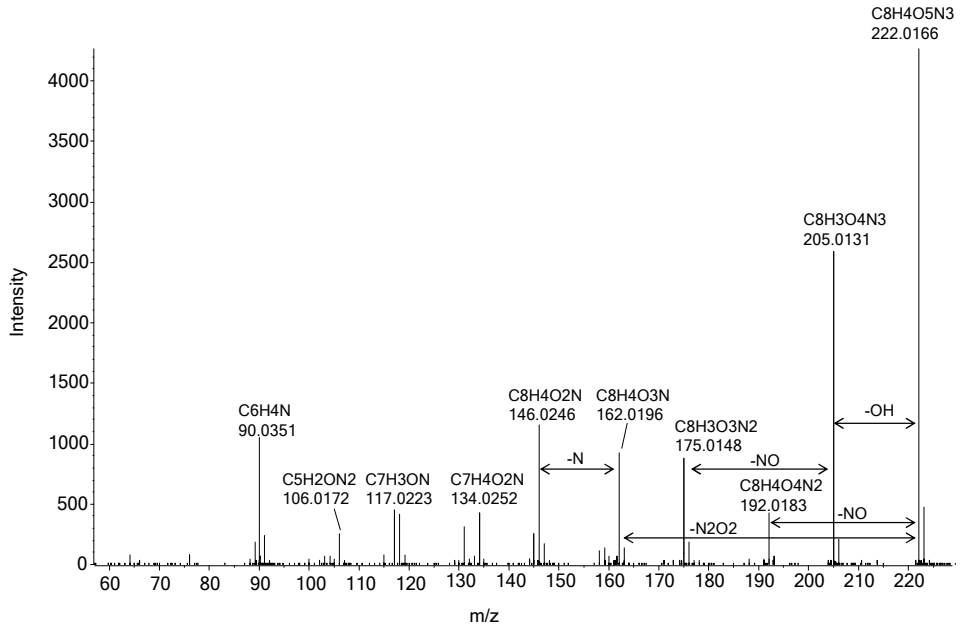
** Retention time of the by-products as determined with stable isotope labeling using the Orbitrap mass spectrometer*

*** These N-DBPs were not detected during the initial isotope labeling experiment, but in the follow-up study by Vughs et al.*

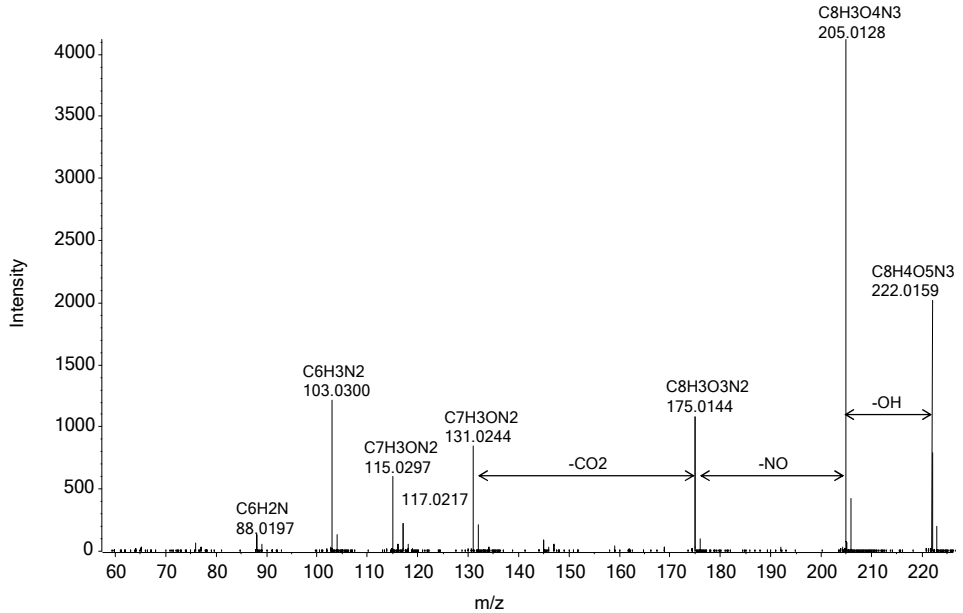
For more information about the detection and identification of the by-products, see publication: Kolkman, A.; Martijn, B. J.; Vughs, D.; Baken, K. A.; van Wezel, A. P., Tracing Nitrogenous Disinfection Byproducts after Medium Pressure UV Water Treatment by Stable Isotope Labeling and High Resolution Mass Spectrometry. Environmental Science & Technology 2015, 49, (7), 4458-4465.

Attachment IV Annotated MS/MS spectra of phenylalanine N-DBPs

ANNOTATED MS/MS SPECTRUM (20, 35 AND 50 EV, AVERAGED) OF BY-PRODUCT MZ 222.0156 (RT 18.45)



ANNOTATED MS/MS SPECTRUM (20, 35 AND 50 EV, AVERAGED) OF BY-PRODUCT MZ 222.0156 (RT 17.79)



Attachment V Tryptophan and phenylalanine N-DBPs SPE experiment conditions

V.1 Samples tryptophan and phenylalanine N-DBPs SPE experiment

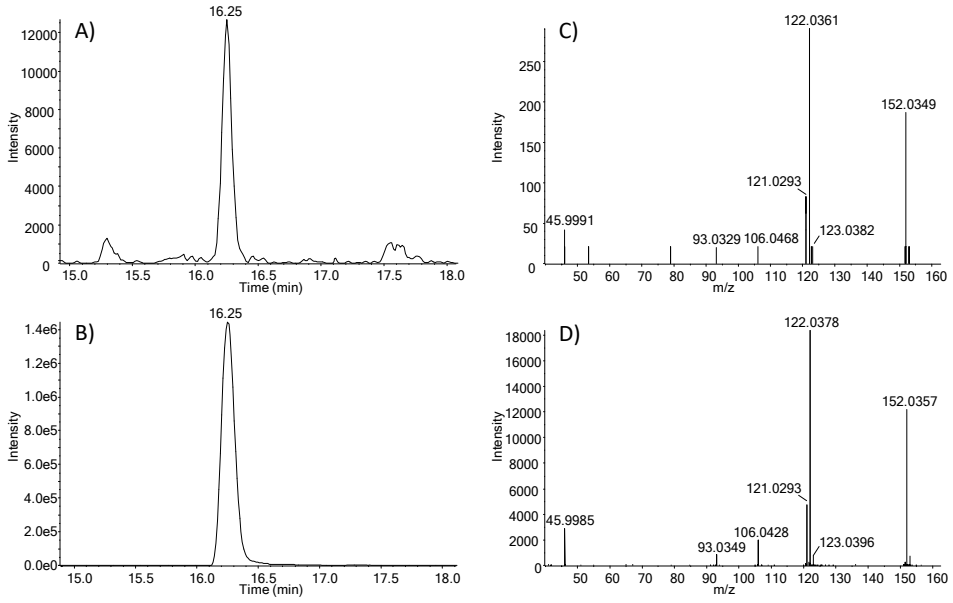
The following samples were used for the SPE experiment in duplo: UV treated + Trp + $^{14}NO_3^-$ sample, UV treated + Phe + $^{14}NO_3^-$ sample and an ultrapure water sample (blank). Ten mL of sample was diluted with 990 ml of ultrapure water prior the SPE sample pre-treatment.

V.2 AMES SPE pre-treatment protocol

For the AMES sample pre-treatment, 1L of sample was acidified to pH 2.3 using hydrochloric acid and loaded on a 200 mg Waters OASIS HLB glass SPE cartridge (Etten-Leur, Netherlands). Then the SPE column was dried for 1 hour by air and elution was performed with 7.5 mL of 8:2 (v/v) acetonitrile/methanol. The eluate was evaporated using a Barkey optocontrol (Leopoldshöhe, Germany) with a gentle nitrogen stream at circa 75 °C (block temperature at 300 °C) until a volume of 250 µL was reached. In order to compare the by-products in the samples after the AMES SPE pre-treatment with the samples that were not pre-treated, 9750 µL of ultrapure water added to the extract, so that effectively no sample enrichment had occurred (10 mL before SPE to 10 mL after SPE). In this way, a fair comparison could be made between the samples with and without sample pre-treatment in order to determine the recovery of the by-products. Finally, the samples were analysed using the LC-QToF as described in paragraph 3.2.3.

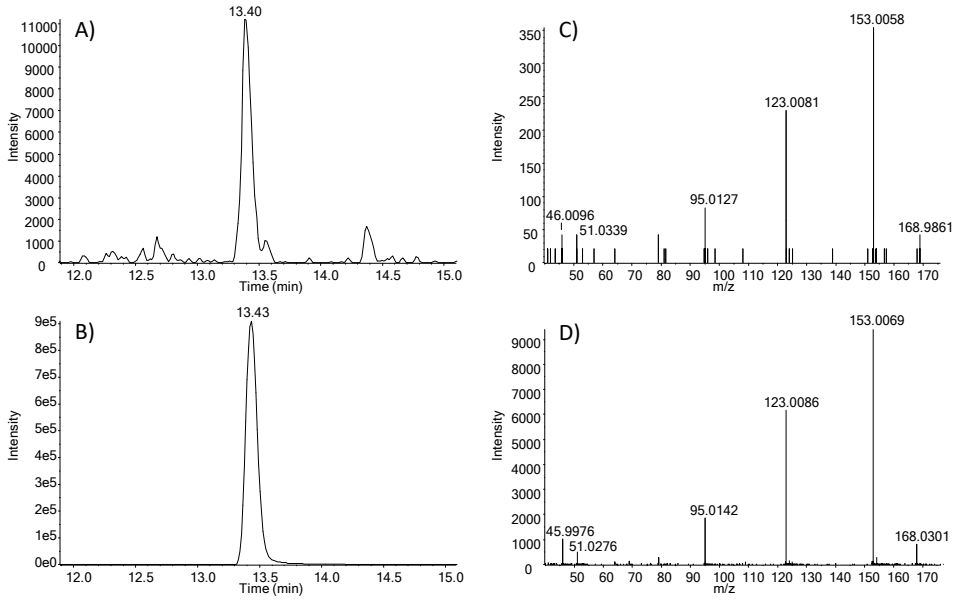
Attachment VI Confirmation of identified N-DBPs

CONFIRMATION OF 2-METHYL-4-NITROPHENOL IN ARTIFICIAL WATER



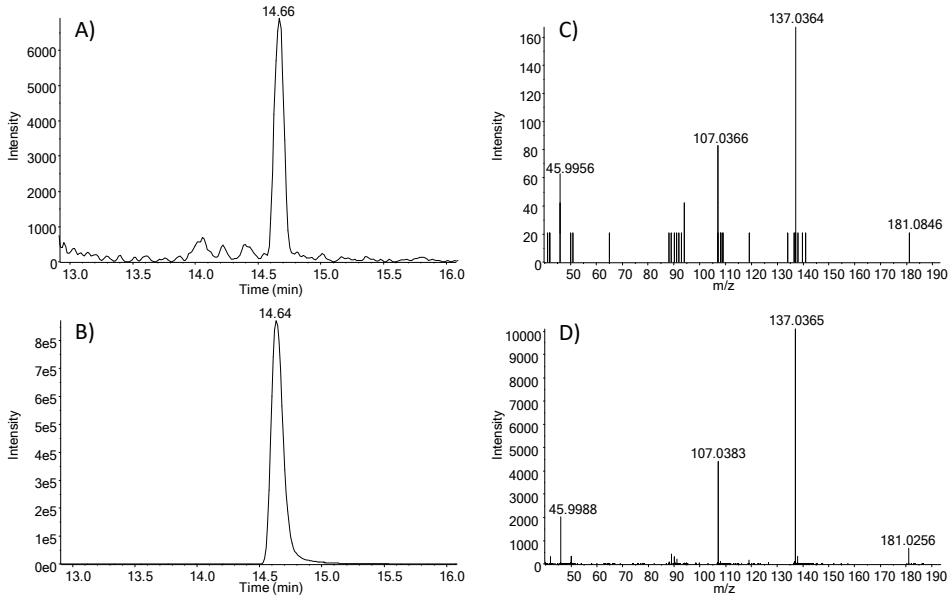
A) EIC of m/z 152.0353 in artificial water after MP UV. B) EIC of 2-methyl-4-nitrophenol reference standard (m/z 152.0353). C) MS2 spectrum of m/z 152.0353 in artificial water after MP UV. D) MS2 spectrum of 2-methyl-4-nitrophenol reference standard.

CONFIRMATION OF 2-METHOXY-4-NITROPHENOL IN ARTIFICIAL WATER



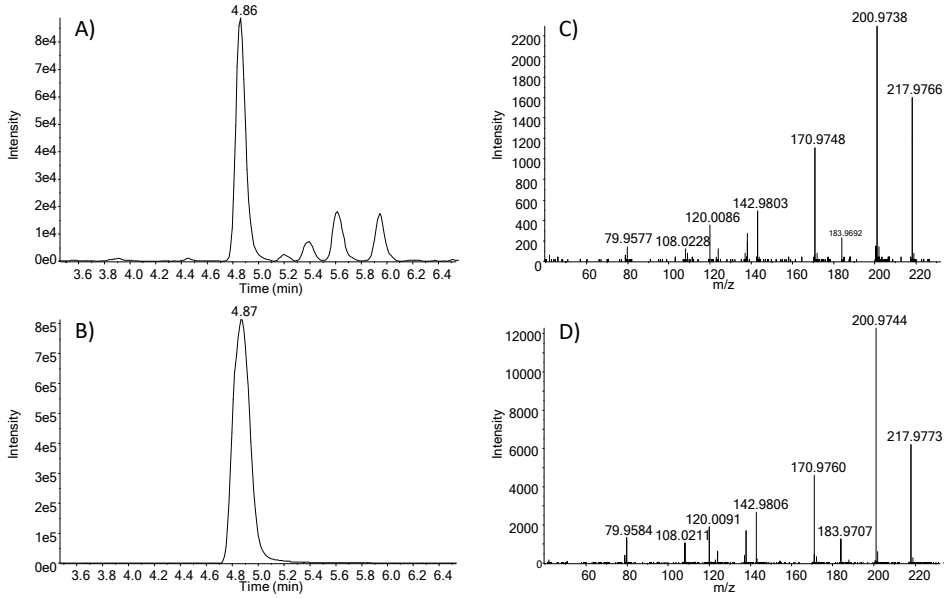
A) EIC of m/z 168.0302 in artificial water after MP UV. B) EIC of 2-methoxy-4-nitrophenol reference standard (m/z 168.0302). C) MS2 spectrum of m/z 168.0302 in artificial water after MP UV. D) MS2 spectrum of 2-methoxy-4-nitrophenol reference standard.

CONFIRMATION OF 2-AMINO-3-NITROBENZOIC ACID IN ARTIFICIAL WATER



A) EIC of m/z 181.0255 in artificial water after MP UV. B) EIC of 2-Amino-3-nitrobenzoic acid reference standard (m/z 181.0255). C) MS2 spectrum of m/z 181.0255 in artificial water after MP UV. D) MS2 spectrum of 2-Amino-3-nitrobenzoic acid reference standard.

CONFIRMATION OF 4-HYDROXY-3-NITROBENZENESULFONIC ACID IN ARTIFICIAL WATER



A) EIC of m/z 217.9765 in artificial water after MP UV. B) EIC of 4-hydroxy-3-nitrobenzenesulfonic acid reference standard (m/z 217.9765). C) MS2 spectrum of m/z 217.9765 in artificial water after MP UV. D) MS2 spectrum of 4-hydroxy-3-nitrobenzenesulfonic acid reference standard.

Attachment VII KWR toxicological evaluation

The following toxicity evaluation was performed by Dr. Kirsten Baken of KWR Watercycle Research Institute.

TOXICOLOGICAL EVALUATION

The following information sources were consulted:

- Organizations: European Food Safety Authority (EFSA), European Chemicals Agency (ECHA), US Environmental Protection Agency (EPA), Food and Drug Administration, and Dutch National Institute for Environment and Health (RIVM)
- Toxicological meta-databases: International Toxicity Estimates for Risk (ITER) comprising data from the WHO International Programme on Chemical Safety (IPCS), U.S. EPA, Agency for Toxic Substances and Disease Registry (ATSDR), Health Canada, International Agency for Research on Cancer (IARC), en RIVM; TOXNET; and the OECD eChemPortal linking to e.g. ACToR, ECHA, HSDB, INCHEM, UNEP and IRIS
- OECD QSAR Toolbox v3.4.0.17 software.

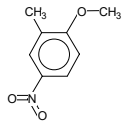
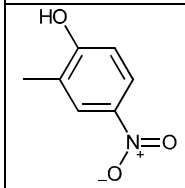
For none of the substances, toxicological study results or human health risk assessments were retrieved from any of the information sources. Additional literature search by name and CAS number of all chemicals using PubMed and Scopus resulted in only two publications, reporting anti-androgenic activity [37] and vasodilating properties [38] of 2-methyl-4-nitrophenol.

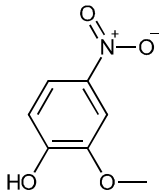
Two chemicals, 2-methyl-4-nitrophenol and 2-methoxy-4-nitrophenol (which is in use as a hair dye ingredient), are present in the ECHA database. ECHA labeled these chemicals as eye, skin and respiratory irritants. Both substances have not been REACH registered, but are listed in the REACH Annex III inventory of indications for hazardous toxicological or ecotoxicological properties. The REACH inventory and chemical profiling using the OECD QSAR Toolbox show potential mutagenic and carcinogenic activity for these substances based on their chemical structure (see attachment VIII). The other substances were not included in the OECD QSAR Toolbox. Their InChI identifiers were imported for chemical profiling; structural alerts for mutagenicity and/or carcinogenicity were identified in these chemicals as well (see attachment IX). In addition, the OECD QSAR Toolbox reports structural alerts for additional endpoints for all substances (attachment VIII and IX). Read across was performed to predict mutagenicity, carcinogenicity, and chronic toxicity based on measured data for identified structural analogues included in the OECD QSAR Toolbox .

SUMMARY OF TOXICOLOGICAL EVALUATION

Substance	**Toxicity data**	**Structural alerts**	**Read across**
2-Amino-3-nitrobenzoic acid	-	mutagenicity carcinogenicity reproductive and developmental toxicity	positive for *in vitro* mutagenicity negative for *in vivo* genotoxicity negative for carcinogenicity
3-Nitroindole	-	mutagenicity skin sensitization	positive for *in vitro* mutagenicity negative for *in vivo* genotoxicity negative for carcinogenicity positive for reproductive toxicity
2-Methyl-4-nitrophenol	anti-androgenic activity vasodilation	mutagenicity carcinogenicity hemolytic anemia & hepatotoxicity *or* energy metabolism dysfunction	positive/negative for *in vitro* mutagenicity positive for carcinogenicity positive for developmental toxicity
2-Methoxy-4-nitrophenol (4-Nitroguaiacol)	-	mutagenicity carcinogenicity energy metabolism dysfunction	negative for *in vitro* mutagenicity positive for carcinogenicity positive for developmental toxicity
4-Hydroxy-3-nitrobenzenesulfonic acid	-	mutagenicity estrogen receptor binding energy metabolism dysfunction	negative for *in vitro* mutagenicity

Attachment VIII Structural alerts reported by ECHA and the OECD QSAR Toolbox

CAS	Substance	ECHA database	OECD QSAR Toolbox
99-53-6[1]	2-Methyl-4-nitrophenol		
		**Suspected mutagen:** The Toolbox profiler 'DNA alerts for AMES, MN and CA by OASIS v.1.3' gives an alert for mutagenicity; The Toolbox profiler 'in vitro mutagenicity (Ames test) alerts by ISS' gives an alert for mutagenicity; ISS Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) **Suspected carcinogen:** The Toolbox profiler 'Carcinogenicity (genotox and nongenotox) alerts by ISS' gives an alert for carcinogenicity; ISS Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (moderate reliability) **Suspected persistent in the environment:** Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is Possible NON Readily Biodegradable (good reliability);The Danish QSAR database contains information indicating that the substance is predicted as non readily biodegradable[2]	**DNA Binding & DNA alerts for AMES, CA and MNT by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids **DNA Binding by OECD:** SN1 >> Nitrenium Ion formation >> Aromatic nitro **Carcinogenicity (genotox and nongenotox) alerts by ISS:** Nitro-aromatic (Genotox) Structural alert for genotoxic carcinogenicity **in vitro mutagenicity (Ames test) alerts by ISS** Nitro-aromatic **in vitro mutagenicity (Micronucleus) alerts by ISS** Nitro-aromatic **Oncologic Primary Classification** Aromatic Amine Type Compounds **Repeated dose (HESS)** Nitrobenzenes (Hemolytic anemia with methemoglobinemia) Rank A Nitrobenzenes (Hepatotoxicity) Rank C
			**DNA Binding by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids **DNA Binding by OECD:** Michael addition >> P450 Mediated Activation to Quinones and Quinone-type Chemicals >> Alkyl phenols SN1 >> Nitrenium Ion formation >> Aromatic nitro **Carcinogenicity (genotox and nongenotox) alerts by ISS:** Nitro-aromatic (Genotox) Structural alert for genotoxic carcinogenicity **in vitro mutagenicity (Ames test) alerts by ISS**

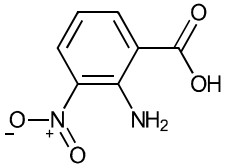
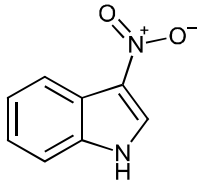
			Nitro-aromatic **in vitro mutagenicity (Micronucleus) alerts by ISS** Nitro-aromatic **Oncologic Primary Classification** Aromatic Amine Type Compounds Phenol Type Compounds **Repeated dose (HESS)** Nitrophenols/ Halophenols (Energy metabolism dysfunction) Rank B
3251-56-7	2-Methoxy-4-nitrophenol 	**Suspected mutagen:** The Toolbox profiler 'DNA alerts for AMES, MN and CA by OASIS v.1.3' gives an alert for mutagenicity; The Toolbox profiler 'in vitro mutagenicity (Ames test) alerts by ISS' gives an alert for mutagenicity; CAESAR Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Suspect Mutagen (moderate reliability); ISS Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (good reliability); KNN Mutagenicity model in VEGA (Q)SAR platform predicts that the chemical is Mutagen (moderate reliability) **Suspected carcinogen:** The Toolbox profiler 'Carcinogenicity (genotox and nongenotox) alerts by ISS' gives an alert for carcinogenicity; CAESAR Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (good reliability); ISS Carcinogenicity model in VEGA (Q)SAR platform predicts that the chemical is Carcinogen (good reliability) **Suspected persistent in the environment**: Ready biodegradability model (IRFMN) in VEGA (Q)SAR platform predicts that the chemical is Possible NON Readily Biodegradable (good reliability);The Danish QSAR database contains information indicating that the substance is predicted as non-readily biodegradable[3]	**DNA Binding & DNA alerts for AMES, CA and MNT by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids **DNA Binding by OECD:** Michael addition >> P450 Mediated Activation to Quinones and Quinone-type Chemicals >> Hydroquinones SN1 >> Nitrenium Ion formation >> Aromatic nitro **in vitro mutagenicity (Ames test) alerts by ISS** Nitro-aromatic **in vitro mutagenicity (Micronucleus) alerts by ISS** H-acceptor-path3-H-acceptor Nitro-aromatic **Carcinogenicity (genotox and nongenotox) alerts by ISS:** Nitro-aromatic (Genotox) Structural alert for genotoxic carcinogenicity **Oncologic Primary Classification** Aromatic Amine Type Compounds Phenol Type Compounds **Repeated dose (HESS)** Nitrophenols/ Halophenols (Energy metabolism dysfunction) Rank B

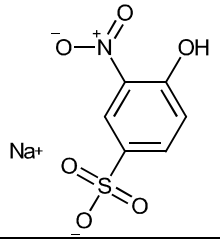
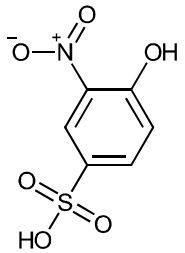
1 OECD QSAR Toolbox reports two structures for the same CAS number

2 https://echa.europa.eu/nl/information-on-chemicals/annex-iii-inventory/-/dislist/details/AIII-100.002.512

3 https://echa.europa.eu/nl/information-on-chemicals/annex-iii-inventory/-/dislist/details/AIII-100.019.854

Attachment IX Structural alerts reported by the OECD QSAR Toolbox

CAS	InChI	Substance	OECD QSAR Toolbox
606-18-8	InChI=1S/C7H6N2O4/c8-6-4(7(10)11)2-1-3-5(6)9(12)13/h1-3H,8H2,(H,10,11)	2-Amino-3-nitrobenzoic acid 	**DNA Binding by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids **DNA Binding by OECD:** SN1 >> Nitrenium Ion formation >> Aromatic nitro **in vitro mutagenicity (Ames test) alerts by ISS** Nitro-aromatic **in vitro mutagenicity (Micronucleus) alerts by ISS** H-acceptor-path3-H-acceptor Nitro-aromatic **Carcinogenicity (genotox and nongenotox) alerts by ISS:** Nitro-aromatic (Genotox) Structural alert for genotoxic carcinogenicity **Oncologic Primary Classification** Aromatic Amine Type Compounds **Protein binding by OASIS v1.4** AN2 >> Michael-type addition to quinoid structures >> Substituted Anilines **DART scheme v.1.0** Known precedent reproductive and developmental toxic potential NO2-alkyl/NO2-benzene derivatives (8b)
4770-03-0	InChI=1S/C8H6N2O2/c11-10(12)8-5-9-7-4-2-1-3-6(7)8/h1-5,9H	3-nitroindole 	**DNA Binding & DNA alerts for AMES, CA and MNT by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Conjugated Nitro Compounds SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Conjugated Nitro Compounds **DNA Binding by OECD:** SN1 >> Nitrenium Ion formation >> Unsaturated heterocyclic nitro **in vivo mutagenicity (Micronucleus) alerts by ISS** H-acceptor-path3-H-acceptor **Oncologic Primary Classification** Aromatic Amine Type Compounds **Protein binding alerts for skin sensitization by OASIS v1.4** Michael Addition >> Michael addition on conjugated systems with electron withdrawing group >> Nitroalkenes

6313-34-4	InChI=1S/C6H5NO6S.Na/c8-6-2-1-4(14(11,12)13)3-5(6)7(9)10;/h1-3,8H,(H,11,12,13);/q;+1	sodium;4-hydroxy-3-nitrobenzenesulfonic acid 	**DNA Binding by OASIS v.1.4:** Radical >> Radical mechanism via ROS formation (indirect) >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids SN1 >> Nucleophilic attack after reduction and nitrenium ion formation >> Nitrophenols, Nitrophenyl Ethers and Nitrobenzoic Acids **in vivo mutagenicity (Micronucleus) alerts by ISS** H-acceptor-path3-H-acceptor **Oncologic Primary Classification** Aromatic Amine Type Compounds Phenol Type Compounds **Estrogen Receptor Binding** Strong binder, OH group **Repeated dose (HESS)** Benzene/ Naphthalene sulfonic acids (Less susceptible) Rank C Nitrophenols/ Halophenols (Energy metabolism dysfuntion) Rank B
616-85-3	*In OECD QSAR Toolbox database*	*4-hydroxy-3-nitrobenzenesulphonic acid* 	

Attachment X Materials & Methods LC-QToF N-DBPs target method

X.1 Chemicals

All solvents used were of analytical grade quality. Acetonitrile and methanol (ultra gradient HPLC grade) was obtained from Avantor Performance Materials B.V. (Deventer, the Netherlands). Formic acid (HPLC quality) and hydrochloric acid 30% suprapur were purchased from Sigma-Aldrich (Steinheim, Germany) and Merck (Darmstadt, Germany), respectively. The internal standards 4-nitrophenol and neburon were obtained from Sigma-Aldrich and bentazon-d6 was obtained from LGC Standards GmbH (Wesel, Germany). The following N-DBPs were purchased from Sigma-Aldrich, 4-nitrophenol, 4-nitrocatechol, 2-hydroxy-5-nitrobenzoic acid, 5-nitrovanillin, 4-nitrophthalic acid, 2,4-dinitrophenol, 4-hydroxy-3-nitrobenzoic acid, 2-hydroxy-3-nitrobenzoic acid, dinoterb, 3,5-dinitrosalicylic acid, 2-methyl-4-nitrophenol, 2-methoxy-4-nitrophenol and 2-amino-3-nitrobenzoic acid. 2-methoxy-4,6-dinitrophenol and 4-hydroxy-3-nitrobenzenesulfonic acid were obtained from Vitas-M laboratory (Moscow, Russia). 4-nitro-1,3-benzenediol was purchased from Santa Cruz Biotechnology. 2-nitrohydroquinone and 3-nitroindole were obtained from Chemos GmbH (Regenstauf, Germany) and Oxchem (Wood Dale, IL, USA), respectively. 4-nitrobenzenesulfonic acid was obtained from TCI Europe (Zwijndrecht, Belgium).Ultrapure water was obtained by purifying demineralized water in an Elga Purelab Chorus ultrapure water system. (High Wycombe, United Kingdom). Stock solutions of the reference and internal standards were prepared in methanol and ultrapure water (20/80% v/v) at a concentration of 100 and 50 mg/L, respectively. Stock solutions were stored at -25 °C.

X.2 Sample pre-treatment

For the sample pre-treatment, 500 mL of sample was acidified to pH 2.3 using hydrochloric acid and was loaded onto a SPE cartridge (OASIS HLB, 200 mg, glass, 6 cc) obtained from Waters (Etten-Leur, Netherlands). Then the SPE column was dried for 1 hour by air and elution was performed with 7.5 mL of 8:2 (v/v) acetonitrile/methanol. The eluate was evaporated using a Barkey optocontrol (Leopoldshöhe, Germany) with a gentle nitrogen stream at circa 75 °C (block temperature at 300 °C) until a volume of 250 µL was reached. Then 750 µL ultrapure water was added to extract, containing nitrophenol-d4, bentazon-d6 and neburon internal standards at a concentration of 13.33 µg/L. Then the extracted was filtered using a 0.2 µm Phenomenex Phenex regenerated cellulose filter (Utrecht, Netherlands) and was transferred to a 1.8 mL autosampler vial for LC-QToF analysis.

X.3 LC-QTOF analysis

HPLC settings (Shimadzu Nexera; LC-30AD, SIL-30AC, CTO-20AC):

- Column: Waters Xbridge C18 XP, 2.1 x 150 mm, 2.5 µm
- Mobile phase A: ultrapure water + 0.05% formic acid
  Mobile phase B: acetonitrile + 0.05% formic acid
- Gradient: linear from 5% to 100% B in 40 min. Held at 100% B for 5 min. Then return to initial conditions in 1 min and held for 6 min.
- Flow: 300 µL/min
- Injection volume: 10 µL
- Column oven temperature: 25°C

QToF-MS settings (AB SCIEX TripleTOF 5600+):

- Resolution: > 30.000 @ m/z 400 (MS and MS/MS mode)
- Mass accuracy < 5 ppm
- Mass range Full scan: 120-500 Da
- Mass range MS2 scan: 40-460 Da
- Ionisation: negative mode
- Source: electrospray (ESI)
- TurboIonSpray heater: 500°C
- Ion Spray Voltage: 3000 volt
- Curtain gas: 25 psi
- Gas 1: 40 psi
- Gas 2: 50 psi
- Divert valve:0 - 3.0 min to waste
- Collision energy: 20, 35, 50 eV (averaged)
- Data dependant MS/MS scans: 8 per cycle (50ms), threshold 100 counts and dynamic background subtraction.
- MS/MS inclusion list, see table below

MS/MS INCLUSIONLIST (2 MINUTE WINDOW)

Mass (Da)	Retention time (min)
213.0153	15.56
138.0197	12.95
154.0146	10.46
182.0100	13.56
196.0252	12.68
210.0040	8.46
183.0047	15.10
154.0146	13.04
154.0146	11.25
201.9816	6.31
182.0095	11.32
192.0095	10.97
239.0673	26.68
226.9946	14.69
152.0353	16.26
168.0302	13.40
217.9765	4.83
181.0255	14.64
161.0357	16.18

Attachment XI Result monitoring study

RESULTS MONITORING STUDY - AERATED RECLAIMED DUNE INFILTRATION WATER SAMPLE

Compound	Aerated reclaimed dune infiltration water						
	27/07/16 ng/L	05/09/16 ng/L	21/09/16 ng/L	20/10/16 ng/L	21/11/16 ng/L	14/12/16 ng/L	11/01/17 ng/L
2-Methoxy-4,6-dinitrophenol	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
4-Nitrophenol	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0
4-Nitrocatechol	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Hydroxy-5-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
5-Nitrovanillin	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
4-Nitrophthalic acid	< 10	< 10	< 10	< 10	< 10	< 10	< 10
2,4-Dinitrophenol	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
4-Nitro-1,3-benzenediol	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Nitrohydroquinone	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0	< 5.0
4-Nitrobenzenesulfonic acid	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
4-Hydroxy-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Hydroxy-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
Dinoterb	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
3,5-Dinitrosalicylic acid	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
2-Methyl-4-nitrophenol	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0	< 1.0
2-Methoxy-4-nitrophenol	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0	< 3.0
4-Hydroxy-3-nitrobenzenesulfonic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
2-Amino-3-nitrobenzoic acid	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
3-Nitroindole	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0	< 2.0
Summed concentration	0.0	0.0	0.0	0.0	0.0	0.0	0.0