**4. TITLE AND SUBTITLE:**
Advanced Chemical Analysis Capability for Alternative Water Source Research

**6. AUTHOR(S):**
- Dr. Tanner Schaub, NMSU Chemical Analysis & Instrumentation Lab
- Dr. Jacqueline Jarvis, NMSU Chemical Analysis & Instrumentation Lab
- Dr. Pei Xu, NMSU Department of Civil Engineering
- Dr. Nirmala Khandan, NMSU Department of Civil Engineering

**14. ABSTRACT:**
With this project, we have adapted CAIL’s state-of-the-art instrumentation capability to serve regional water research needs, specifically through the establishment of novel characterization approaches for organic molecules in water based on high resolution mass spectrometry, including contaminants of emerging concern and their conversion products.

**15. SUBJECT TERMS:**
- Wastewater analysis, contaminants of emerging concern, CECs, aqueous organics, wastewater treatment, mass spectrometry, chemical analysis

**16. SECURITY CLASSIFICATION OF:**
- **a. REPORT** U
- **b. ABSTRACT** U
- **THIS PAGE** U

---

**Abstract:**

**Online at [https://www.usbr.gov/research/dwpr/DWPR_Reports.html](https://www.usbr.gov/research/dwpr/DWPR_Reports.html)**

**1. REPORT DATE (DD-MM-YYYY)**
September 2018

**2. REPORT TYPE**
Final

**3. DATES COVERED**
5/1/2017-4/30/2018

**5a. CONTRACT NUMBER**
Agreement No. R16AC00002

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**6. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
New Mexico State University
New Mexico Water Resources Research Institute
PO Box 30001, MSC 3167
Las Cruces, NM 88003

**7. PERFORMING ORGANIZATION REPORT NUMBER**

**8. SPONSOR/MONITOR’S ACRONYM(S)**
Reclamation

**10. SPONSOR/MONITOR’S REPORT NUMBER(S)**
DWPR Report No. NMSU002

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Available from the National Technical Information Service, Operations Division, 5285 Port Royal Road, Springfield VA 22161

**13. SUPPLEMENTARY NOTES**

**19b. TELEPHONE NUMBER (Include area code)**
303-445-2013

---

**Standard Form 298**
Prescribed by ANSI Std. Z39.18
Advanced Chemical Analysis Capability for Alternative Water Source Research

Prepared for the Bureau of Reclamation Under Agreement No. R16AC0002

by

Dr. Tanner Schaub, NMSU Chemical Analysis and Instrumentation Lab

Dr. Jacqueline Jarvis, NMSU Chemical Analysis and Instrumentation Lab

Dr. Pei Xu, NMSU Department of Civil Engineering

Dr. Nirmala Khandan, NMSU Department of Civil Engineering
Mission Statements

The U.S. Department of the Interior protects America’s natural resources and heritage, honors our cultures and tribal communities, and supplies the energy to power our future.

The mission of the Bureau of Reclamation is to manage, develop, and protect water and related resources in an environmentally and economically sound manner in the interest of the American public.

Disclaimer

The views, analysis, recommendations, and conclusions in this report are those of the authors and do not represent official or unofficial policies or opinions of the United States Government, and the United States takes no position with regard to any findings, conclusions, or recommendations made. As such, mention of trade names or commercial products does not constitute their endorsement by the United States Government.

Acknowledgments

Work on this project was made possible by funding from a cooperative agreement between the U.S. Bureau of Reclamation (BOR) and New Mexico State University (#R16AC00002, Center for the Development and Use of Alternative Water Supplies) and additional support from the National Science Foundation Major Research and Instrumentation Program (NSF MRI #1626468) and the Center for Animal Health and Food Safety at New Mexico State University.
Acronyms and Abbreviations

Reclamation       Bureau of Reclamation
NMSU              New Mexico State University
Chemical Analysis and Instrumentation Laboratory  CAIL
Fourier transform ion cyclotron resonance mass spectrometry FT-ICR MS
Contaminants of emerging concern      CECs
Tandem mass spectrometry             MS/MS
HRMS  high resolution mass spectrometry (FT-ICR and Orbitrap Fusion)
LC/MS  liquid chromatography mass spectrometry
SPE   solid-phase extraction

Measurements

μg/L        microgram per liter
m/z         mass-to-charge ratio
ppm         parts-per-million
ppb         parts-per-billion
Da          Dalton
DBE  molecular double bond equivalent, the number of molecular rings plus double bonds in a molecule.
Advanced Chemical Analysis Capability for Alternative Water Source Research

Contents

Page
Acronyms and Abbreviations ................................................................. iii
Measurements ........................................................................................ iv
Executive Summary ................................................................................ 1
1. Introduction ........................................................................................ 1
   1.1. Background ........................................................................... 1
   1.2. Objective .............................................................................. 4
   1.3. Personnel ............................................................................. 5
2. Approach ........................................................................................... 6
   2.1. Project Facility and Equipment .............................................. 6
   2.2. Methods ............................................................................... 8
      2.2.1. Sample Preparation ..................................................... 8
      2.2.2. Instrument Configuration and Modes of Operation ......... 9
      2.2.3. Mass Spectral Database/Library Considerations .......... 9
      2.2.4. Peak Detection and Integration ................................... 10
3. Results ............................................................................................ 10
   3.1. Instrumentation Development .............................................. 10
   3.2. Database selection and implementation................................. 11
   3.3. Nanoflow Liquid Chromatography/Mass Spectrometry ........... 13
   3.4. Analysis of Primary Settled Municipal Wastewater and Treated Wastewater ........................................................................ 15
      3.4.1. Broadband Compositional Analysis .............................. 15
   3.5. Nanoflow LC and MS/MS Identification of Contaminants ........ 19
4. Conclusions ..................................................................................... 21
   4.1. Established Resource .......................................................... 21
   4.2. Near-term Priorities .............................................................. 21
References ............................................................................................ 22

Figures

Figure 1. Broadband FT-ICR mass spectrum (left) of residual organics in a Permian Basin produced water sample........................................ 4

Figure 2. Improvement in mass spectral bandwidth with the new stand-alone data station..................................................... 11

Figure 3. Database curation.................................................................. 13

Figure 4. High resolution liquid chromatography/mass spectrometry (LC/MS) analysis of treated wastewater................................. 14

Figure 5. Schematic of the algal wastewater treatment process........... 16
Figure 6. High resolution Orbitrap Fusion mass spectra for Las Cruces municipal wastewater and process related treatment samples.............16

Figure 7: Heteroatom class distributions derived from the mass spectra of the raw wastewater.................................................................17

Figure 8. Compositional space coverage for organic species present within wastewater and treated wastewater samples derived from positive-ion ESI Orbitrap Fusion mass spectra.........................................................18

Tables

Table 1. Spiked compound identification in primary wastewater.........14

Table 2. Algal wastewater treatment remediation of contaminants.......20
Executive Summary

Research projects that address non-traditional water source utilization and treatment in southern New Mexico require innovative analytical chemistry support to determine water source quality, to evaluate the efficacy of treatment technologies, and to monitor associated systems such as environmental discharge and food safety considerations. With the availability of two modern, ultrahigh resolution mass spectrometers and all ancillary sample preparation equipment, the NMSU Chemical Analysis and Instrumentation Laboratory (CAIL) is a regionally unique resource for advanced chemical analysis. With this project, we have adapted CAIL’s state-of-the-art instrumentation capability to serve regional water research needs, specifically through the establishment of novel and robust characterization approaches for organic molecules in water based on high resolution mass spectrometry, including contaminants of emerging concern and their conversion products. Specific objectives are to improve instrument performance for wastewater applications (accomplished), identify and implement an appropriate mass spectral library search tool (accomplished), parametric selection (accomplished/in adaptation for various sample types and extractions), and implement appropriate extraction techniques (accomplished). This project established a new resource for water characterization at New Mexico State University, the utilization and optimization of which is now underway for specific applications.

1. Introduction

1.1. Background

The development of non-traditional water sources for agricultural and other uses is of ever-increasing importance in the western United States. Each impaired or non-traditional water source presents unique challenges, and successful formulation of treatment strategies requires input from interdisciplinary teams comprised of engineers, scientists and industrial collaborators. The ability to define the composition of a non-traditional water source is of critical importance to establish potential uses as well as to monitor the efficacy of treatment and transformation processes.

Complete chemical characterization of water is a multi-faceted problem. The breadth of potential contamination for a given water supply can include extremely diverse constituents that in turn require diverse
analytical techniques. Primary metrics such as pH, total dissolved and suspended solids, metals quantitation and the determination of cations/anions, are well-addressed by basic equipment that is readily available in research laboratories. Conversely, the analysis of dissolved organic contaminants is a more challenging problem and current approaches involving liquid chromatography mass spectrometry (LC/MS) generally present only a limited view of organic pollutant/dissolved organic composition.

Simply stated, most current LC/MS analytical approaches for organic molecule/pollutant characterization of impaired water sources do not achieve the performance metrics required for the task. With regard to municipal wastewater, targeted analyses monitor only handfuls of selected compounds, whereas successful treatment and use of these water sources requires global monitoring of contaminants and their transformation products within mixtures of organic molecules that can number in the multiple-thousands. Characterization of organics in produced water requires even greater analytical performance to address mixtures that contain compounds with mass differences of as little as 0.002 Da and require resolution, mass measurement accuracy and dynamic range that is only available with the highest performance mass spectrometers.

In 2016, the EPA released the fourth version of its contaminant candidate list (CCL 4). The CCL 4 consists of 97 chemicals (or groups) and 12 microbial agents that include pesticides, biological toxins, disinfection byproducts, and pharmaceuticals. These compounds are currently not regulated by the safe drinking water act (SDWA) or the EPA, but are believed to be present in public water systems and are of “emerging concern”. The primary methods used for the detection and quantitation of contaminants of emerging concern (CECs) from aqueous media involve the use of liquid chromatography coupled to mass spectrometry.

Current LC/MS water characterization methods commonly employ low resolution mass spectrometers (quadrupole time-of-flight, triple-quadrupoles and ion traps) coupled with standard liquid chromatography and spectral database matching for the targeted analysis of organics. With these platforms, there is a growing concern about the true identity (i.e., correct molecular assignment of a mass spectral component signal including structural isomers) of compounds that are claimed to be positively identified, especially in multi-disciplinary research settings where mass spectrometry may be used without thorough understanding of its scope and limitations. Even for targeted analyses, these platforms perform only modestly as illustrated, for example, in a recent inter-laboratory comparison between analytical methods for quantitation of
target CECs in drinking and surface water that showed false positive and false negative rates of >10% for those relatively simple (i.e. not spectrally complex) samples.\(^1\)

Although desirable, targeted quantitative analysis has multiple drawbacks. First, targeted analysis requires references standards for all target compounds, which are not always available and/or may be expensive.\(^3\) Ultimately, targeted analysis is generally limited to a few hundred compounds.\(^3\) Furthermore, targeted analysis requires the concentration of target compounds to be high enough to facilitate tandem mass spectrometry.\(^3\) Schymanski et al. determined that for wastewater samples, only 1.2% of all the detected mass spectral signals (and only 4 of the 30 most intense signals) belonged to 376 target compounds.\(^3\) Similarly, inter-laboratory comparisons have shown inconsistent results between laboratories resulting in both false positive and false negative results, which demonstrates the need for more robust methods of CEC identification and monitoring.

Therefore, the global analysis of CECs in aqueous environments requires high resolution mass spectrometry (i.e., FT-ICR MS or high resolution Orbitrap MS). Krauss et al. described three main workflows for the use of LC-HRMS in environmental analysis: (1) quantitative target analysis with reference standards, (2) suspect screening without reference standards, and (3) non-target screening of unknowns.\(^3\) High resolution MS has begun to find increased usage for the identification of CECs, and identification criteria of suspect and unknown compounds are in development.\(^3\) These criteria include unambiguous elemental formulae assignment, isotopic pattern validation, predicted chromatographic behavior, and matching of MS/MS spectra to spectral databases.\(^3\) Low resolution mass spectrometers do not have an adequate dynamic range to detect low concentrations of contaminants, lack mass measurement accuracy required for molecular elemental composition determination, or necessary resolution to isolate signals from nominally isobaric compounds.\(^3\) Additionally, high resolution platforms allow targeted analyses to be performed simultaneously with non-target analysis to streamline CEC analyses, making it one of the most promising avenues for identification of unknown CECs in aqueous environments.\(^3\)

Because both the FT-ICR and Orbitrap Fusion platforms identify compounds unambiguously at the level of elemental composition, subsequent tandem mass spectrometry measurements (i.e., fragmentation of individual components and measurement of their fragments using the Orbitrap Fusion) combined with nanoflow liquid chromatography and modern, cloud-based database matching algorithms enable identification and tracking of CECs, their transformation products and other non-database matched organics. Through prior Department of
Energy collaborations, we have had a narrow opportunity to analyze several produced water samples by FT-ICR MS. For some of those materials, significant diversity exists within the distribution of dissolved organic compounds as shown in Figure 1. The observed level of complexity requires ultrahigh resolution to observe all components, and sub-part per million mass measurement accuracy facilitates elemental composition assignment directly from measured m/z values.

**Figure 1.** Broadband FT-ICR mass spectrum (left) of residual organics in a New Mexican/Permian Basin-produced water sample where >7,700 compounds are identified at the level of elemental composition. The zoom inset (right) illustrates spectral complexity within a 0.5 Da mass window and gives a comparison of FT-ICR MS resolution to that of a lower resolution time-of-flight mass spectrometer.

### 1.2. Objective

Our blanket project objective is to tailor CAIL HRMS analytical instrumentation capability for the characterization of organics in impaired water sources. This development will allow CAIL to provide innovative feedback to multidisciplinary teams of water researchers in New Mexico and beyond at an unprecedented level of specificity.

With this project, we establish methodology to address organic compound determination in water. Specifically, we utilize two powerful, ultrahigh resolution mass spectrometers: an FT-ICR mass spectrometer that delivers unmatched mass resolving power and mass measurement accuracy (each of those metrics being at least an order of magnitude better for this instrument than other mass spectrometers), and an Orbitrap Fusion Tribrid mass spectrometer (commissioned January 2017) that provides high resolution, accurate mass measurement, and multiple fragmentation mechanisms for molecular structural elucidation. Both instruments can be coupled to our nanoflow liquid chromatography system...
(also commissioned in January 2017) for compositional analysis of complex mixtures with excellent specificity and dynamic range. The FT-ICR MS technology has been widely applied to the analysis of petroleum to characterize of tens of thousands of compounds simultaneously at the level of elemental composition.2–10 Within the CAIL laboratory, we have utilized this technology for the characterization of multiple complex mixture types in biofuel research,11–28 environmental applications,29,30 and biological systems.31 High resolution/accurate mass measurement by FT-ICR MS with a liquid chromatography interface has been shown for the identification of CECs32,33 and the observation of molecular-level transformations in aqueous organic matter associated with photochemical degradation.34 We have utilized FT-ICR MS for broadband compositional analysis (non-targeted). This is a proven and central component of the total platform. Primary development efforts of this project relate to Orbitrap Fusion, LC/MS and mass spectral database implementation outcomes and are discussed below.

1.3. Personnel

PI Dr. Tanner Schaub Dr. Schaub is the director of the NMSU Chemical Analysis and Instrumentation Laboratory and the director of the NMSU Center for Animal Health and Food Safety. Currently, Dr. Schaub’s lab is funded by a continuing collaboration with the Chemical and Biological Processes Development Group at Pacific Northwest National Laboratory, the NSF EPSCoR program “Energize New Mexico”, the NSF Major Research Instrumentation Program, the NSF CMI program, and the DPWR program of the U.S. Bureau of Reclamation.

Co-PI Dr. Nagamany Nirmalakhandan is a professor and the Ed Harold Foreman endowed chair holder in the NMSU Department of Civil Engineering. Dr. Khandan is the campus PI of the NSF ERC for Reinventing the Nation’s Urban Water Infrastructure (ReNUWIt). His ReNUWIt project has resulted in an algal testbed at the Las Cruces WW Treatment Plant for energy-efficient wastewater treatment.

Co-PI Dr. Pei Xu is an associate professor in the NMSU Department of Civil Engineering. Dr. Xu’s multiple funded research areas include water and wastewater engineering; membrane processes; desalination; potable and non-potable water reuse; produced water treatment; oxidation and photocatalysis; biological and bioelectrochemical processes; removal of emerging contaminants; and membrane fouling.

Co-PI Dr. Jacqueline Jarvis is a research assistant professor with the NMSU Chemical Analysis and Instrumentation Laboratory. Her research focuses on the application of advanced mass spectrometry techniques for energy research, environmental samples and biomedical applications.
2. Approach

Identification of the distribution (and removal/transformation) of contaminants from urban wastewater requires three approaches. First, identification of known contaminants (e.g., those represented in current LC/MS libraries) is achieved through rigorous HRMS with tandem mass spectrometry. Second, quantitation of known compounds may be achieved for compounds for which analytical standards are available and mass spectral database entries exist. Finally, high performance mass spectrometry facilitates the observation of broadband organic compound distributions (including natural organic matter), and allows for tracking of reaction products and dissolved organic material for which reference spectra and/or analytical standards are not available (i.e., true analytical unknowns).

For our project, we incorporate solid-phase extraction (SPE) to concentrate the full complement of organic compounds (i.e., CECs and degradation products, dissolved organic matter) from aqueous media. The use of ultrahigh resolution mass spectrometry allows detection of thousands of compounds from SPE extracts simultaneously within full-scan mode. Identification of CECs and degradation products is accomplished through unambiguous elemental formulae assignment provided by the accurate mass measurement inherent to HRMS. Additionally, the ultrahigh resolution, high mass accuracy, and high dynamic range of FT-MS allows for detection of ions with mass differences of as little as 0.002 Da, which ensures that low concentration compounds will be detected and identified.

Separation and identification of CECs and degradation products by tandem MS incorporates nano-liquid chromatography for targeted and non-target analyses. Retention time and fragment ion spectral matching with reference standards is used to identify and quantitate target compounds, whereas high resolution product ion spectra identify suspect and non-target compounds. Additionally, nano-liquid chromatography provides higher sensitivity than standard LC and separates isomeric compounds for identification by tandem MS. For wastewater, we utilize Las Cruces wastewater and treated wastewater from NMSU College of Engineering research projects for analytical platform optimization. Our efforts focus on three aspects:

1. **HRMS of municipal wastewater; nanoFlow LC and direct infusion.** We evaluated multiple varieties of mass spectral data based on high resolution/accurate mass measurement to determine complementarity and identify ideal instrument workflows that eliminate false positive/negative conclusions.
2. **Tandem Mass Spectrometry with multiple fragmentation mechanisms:** Building from accurate mass measurement of compounds (isolated by nano-flow liquid chromatography and also within the mass.
spectrometer), we employ both low- (CID) and higher-energy (HCD) collisionally-induced dissociation to generate product ion spectra for compound identification. With CAIL’s Orbitrap Fusion, product ion spectra may be generated at both high resolution/mass accuracy or in rapid, low-resolution mode.

3. Data Analysis and Spectral Matching: We have implemented a mass spectral database processing tool with the NIST 2017 library for compound identification using accurate mass precursor measurement and MS/MS.

2.1. Project Facility and Equipment

The NMSU Chemical Analysis and Instrumentation Laboratory is an integrated chemical instrumentation facility that specializes in the application of advanced mass spectrometry for complex mixture analysis in alternative fuels research, environmental applications, petroleum, and biological systems. The laboratory is equipped with a variety of mass spectrometers, chromatography systems, and analytical equipment. Since its inception in 2008, the CAIL lab has delivered advanced mass spectrometry and chemical analyses for >60 on-campus interdisciplinary collaborations that span four colleges and eleven departments. CAIL supports student research and sustains active collaboration with three national Laboratories (PNNL, NREL, LANL) and multiple universities both nationally and internationally. The laboratory is supported by research grants and contracts and by the NMSU Agricultural Experiment Station. CAIL prioritizes analyses in support of current and funded collaborations, followed by analyses performed to collect pre-data for collaborative proposals. CAIL only participates in fee-for-service activity on a limited basis. Projects that proceed via fee-for-service arrangements are screened carefully to avoid compromising lab productivity and goals.

Our ultrahigh resolution mass spectrometry systems enable the analysis of complex materials such as petroleum, dissolved aqueous organic matter, and environmental samples where multiple-thousands of compounds are observed simultaneously. We develop modes of instrument operation and advanced data analysis tools that are specific for each application area.

Instrumentation

- **Orbitrap Fusion Tribrid Mass Spectrometer**: Three orthogonal tandem mass spectrometry functionalities with part per million mass measurement accuracy and high resolution.
- **NanoFlow Liquid Chromatograph** – Thermo UltimateNano, state of the art nanoLC system.
- **FT-ICR Mass Spectrometer**: The FT-ICR mass spectrometer provides the highest resolution and mass measurement accuracy available to any mass spectrometer.
- **NanoMate Chip-based Electrospray Robot**: The NanoMate enables true nanoflow liquid chromatography and direct infusion electrospray ionization. Each sample flows through its own sample path to avoid carry over and on-the-fly fraction collection is available.
- **Liquid Chromatography/Mass Spectrometry (LC-MS)**: A second LC/MS system is based on a linear quadrupole ion trap (Thermo LTQ) and allows routine quantitative analysis for a variety of analytes ranging from pesticides to peptides. The LTQ is a highly sensitive mass spectrometer with tandem mass measurement capability.
- **Gas Chromatography/Mass Spectrometry (GC-MS)**: Two Agilent 6890N GC/MS systems are available for analysis of targeted volatile and semi-volatile compounds. The lab maintains both NIST and AMDIS searchable databases.
- **Inductively Coupled Plasma/Mass Spectrometry (ICP-MS)**: An Agilent 7500ce provides parts-per-billion quantitation of trace metals.

### 2.2. Methods

#### 2.2.1. Sample Preparation

We investigated two commercially available solid phase extraction cartridges (Waters HLB and Bond Elut Plexa), between which there was little difference in extraction efficiency or performance. For LC and direct infusion mass spectrometry, Las Cruces municipal primary-settled wastewater samples were adjusted to pH 2, vacuum filtered (0.45-0.7 µm) prior to extraction with Waters HLB (3 mL, 60 mg) solid phase extraction (SPE) cartridges (the same procedure was used for the Plexa cartridges). Cartridges were preconditioned with 2 mL of methanol followed by 2 mL of HPLC grade water. The sample water (50-200 mL) was passed through the cartridge followed by 2 mL of HPLC grade water to ensure all the sample water had eluted from the cartridge. Flow rates were not measured. The stationary phase was allowed to dry to remove any residual water. The analytes of interest were eluted with 2 mL of methanol and dried under N₂ gas. The analytes were resuspended in 100 µL of methanol and brought to a total volume of 1 mL in HPLC grade water prior to injection into the nano LC Orbitrap Fusion mass spectrometer.
2.2.2. Instrument Configuration and Modes of Operation

High resolution mass spectrometers are operated in one of two modes: 1) direct infusion, or 2) coupled to a nanoflow liquid chromatograph. We discuss only Orbitrap Fusion operation in this and subsequent sections. For direct infusion, SPE elute samples were mixed (50:50) with methanol:water solution that contains 0.1% formic acid to facilitate ionization. Electrospray ionization (both polarities) is performed with an Advion Triversa NanoMate sample handling/electrospray ionization robot. This apparatus can deliver as little as 500 nanoliters of water for analysis times of ~5-10 minutes for that volume. During that time, the mass spectrometer is configured to take 50 broadband mass spectra at high mass resolving power \( m/\Delta m_{50\%} = 400,000 \) at \( m/z \ 200 \), where \( m/\Delta m_{50\%} \) is the spectral peak width at half peak height) that are signal averaged to improve dynamic range. That composite broadband spectrum is then analyzed on-the-fly by the control software to formulate a decision tree for selecting and fragmenting individual compounds present in the solution (i.e., “data dependent MS/MS”). Compounds are isolated one at a time by an inline linear quadrupole ion trap at a spectral width of 0.5 m/z units and subjected to either collisionally induced dissociation (CID) or high energy collisional dissociation (HCD).

Product/fragment ion spectra are generated at a resolving power of 30,000 (m/z 200). For nanoflow LC/MS operation, the mass spectrometer is operated as described above including data-dependent MS/MS. The chromatographic flow rate is 300 nanoliters/minute for a 70 minute linear gradient from 98:2% to 5:95% (water:acetonitrile) through a homemade 50 micrometer diameter capillary LC column with a C18 (contains 18 carbon atoms) stationary phase. NanoLC eluent is delivered to the NanoMate robot for electrospray ionization and introduction to the mass spectrometer via a vendor-supplied infusion mandrel.

2.2.3. Mass Spectral Database/Library Considerations

Our data analysis and processing platform required some development, and the final implementation is described here. Namely, raw data files were converted to mzXML format using MSConvert (ProteoWizard) with a threshold of the 100 most intense ions and only MS2 spectra. The mzXML files were automatically searched through the NIST 2017 tandem MS library subject to the following parameters: importation of accurate mass spectrum type with four decimal places for the precursor and product ions, precursor mass accuracy tolerance of 1 ppm, product ion tolerance of 0.5 m/z, and constrained to only ion trap, FT and HCD spectra. The top hit for each MS2 spectrum was output into the NIST generated text result file. We implemented in-house written code to convert the NIST text file into an Excel file that can easily be filtered and sorted for the best library matches. Work is continuing to automate library
searching and output of user ready Excel files for multiple samples at one time.

2.2.4. Peak Detection and Integration

Masses for all observed compounds are recorded at better than 1 part-per-million in our measurements, which provides elemental composition directly from each measured mass-to-charge ratio. That mass measurement accuracy is further utilized as a constraint when searching/matching fragment ion spectra to the database. The extracted ion chromatographic peak areas for precursor ions that resulted in matches with the NIST library (reverse match factor $> 700$) were integrated. The peak areas of identified compounds were normalized to the peak area of reserpine (indole alkaloid, an internal standard substance) and fold change differences between samples are reported.

3. Results

3.1. Instrumentation Development

During the project period, we significantly improved the performance of the Orbitrap Fusion mass spectrometer, which is now the central component of our water analysis platform. CAIL hosted three instrumentation development researchers from the Swiss institute École Polytechnique Fédérale de Lausanne (EPFL) for a week-long collaboration kick-off meeting and hardware installation in September. The NMSU/EPFL collaboration focuses on the implementation of stand-alone signal processing hardware interfaced to the lab’s Orbitrap Fusion mass spectrometer. These capabilities increase instrument performance and transform the Orbitrap Fusion into the most powerful of its kind in the world. Specifically, we purchased $57K worth of equipment from EPFL (purchased with CAIL funds). The new capability enables multiple improvements in performance:

• Resolution: absorption mode Fourier-Transform, longer (full) transient ion detection for ultrahigh mass resolution
• Sensitivity: longer transients and ion accumulation, amplified transients
• Mass range: detection of ions at both higher and lower mass ranges compared to commercial data station, see Figure 2.
• Speed: parallel ion detection and ion accumulation/fragmentation for increased duty cycle
Figure 2. Improvement in mass spectral bandwidth with the new stand-alone data station (red) compared to the standard Thermo data station (black) illustrated with a complex sample of digested proteins. Notice areas of previously unobserved masses at both high and low mass. This is one improvement among several supplied by the new hardware that ensures comprehensive coverage of wastewater organic compounds for a variety of sample types.

3.2. Database selection and implementation

A central component of our data analysis platform is the 2017 version of the NIST mass spectral database. The delivery of this product by the vendor was delayed by several months. Nonetheless, we received the NIST mass spectral library in the sixth month of our project and interfaced it with our data processing server. We ran tests of high resolution data from our instruments collected for treated Las Cruces municipal wastewater. We curated the NIST database to fit our instrumentation specificity.

The 2017 NIST database is essentially a digital library of mass spectra that correspond to fragmented ions. The important improvement of this database as compared to previous versions is that the values that represent the mass of the starting (parent) molecule/ion are reported accurately to the 5th decimal place. That feature aligns with the performance of our instrumentation and allows us to constrain the search space during the data-to-library matching process. That capability improves speed, accuracy, and reduces false positive/negative rates.
The new NIST database contains an enormous number of fragment mass spectra as delivered by the vendor (i.e., over 570,000 spectra). Our instrument data files for water analysis typically contain ~45,000 scans, and given these parameters the computational requirements for that matching procedure exceed our server capability. However, the database includes MS/MS spectra for an array of instrument parameters, instrument types, and ion adducts that are not relevant to our analysis. In fact, the 575k spectra included in the library only correspond to ~17k original unique compounds. Therefore, we implemented a process for culling/curating the database to remove duplicates and unnecessary/irrelevant spectra in order to improve our analysis speed and throughput (see Figure 3).
3.3. Nanoflow Liquid Chromatography/Mass Spectrometry

Co-PIs Xu and Khandan have provided wastewater samples to CAIL for preliminary analysis by ultrahigh resolution Orbitrap mass spectrometry. The samples are primary wastewater from the Las Cruces wastewater treatment plant and treated samples that consist of wastewater from Dr. Xu’s membrane treatment process and Dr. Khandan’s algal water treatment system. CAIL staff have analyzed multiple sample types from each source by both liquid chromatography and direct infusion (no chromatography) mass spectrometry.

We have generated nanoflow LC/mass spectrometry data for samples of wastewater samples mentioned above to which a select group of standard compounds have been added. Those analyses allow us to evaluate the efficacy of our process with respect to sensitivity as well as identification accuracy and robustness of our data processing tools. A portion of the chromatogram and corresponding mass spectrum is shown below for a wastewater sample that has been exposed to a reverse-osmosis treatment process in the Xu lab (Figure 4) and the confirmation of detection of the spiked samples (100 ppb) is shown in Table 1.
Figure 4. High resolution liquid chromatography/Orbitrap Fusion mass spectrometry (LC/MS) analysis of treated wastewater. The data show good chromatographic peak shape and high dynamic range for a primary wastewater sample spiked with reference pesticides and pharmaceuticals.

Table 1. Library matching score and mass measurement error for 17 pesticides and pharmaceuticals spiked into raw primary wastewater and analyzed by the Orbitrap Fusion platform. Excellent library scores and a mass measurement accuracy of 225 parts-per-billion (RMS) indicate robust identification confidence. We are in the process of formalizing score cut-off values and correlating those to the performance metrics of our instrument.

<table>
<thead>
<tr>
<th>Match Result Name</th>
<th>Retention Time (minutes)</th>
<th>Score</th>
<th>Mass Measurement Error (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetaminophen</td>
<td>13.3</td>
<td>98.2</td>
<td>0.017</td>
</tr>
<tr>
<td>Trimethoprim</td>
<td>15.6</td>
<td>99.0</td>
<td>0.195</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>16.1</td>
<td>100.0</td>
<td>0.037</td>
</tr>
<tr>
<td>Dimethoate</td>
<td>18.1</td>
<td>99.0</td>
<td>0.200</td>
</tr>
<tr>
<td>Bromacil</td>
<td>21.0</td>
<td>100.0</td>
<td>0.797</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>21.3</td>
<td>100.0</td>
<td>0.353</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>22.4</td>
<td>96.1</td>
<td>0.168</td>
</tr>
<tr>
<td>Atrazine</td>
<td>24.6</td>
<td>93.6</td>
<td>-0.069</td>
</tr>
<tr>
<td>Prometryne</td>
<td>24.6</td>
<td>99.0</td>
<td>0.268</td>
</tr>
<tr>
<td>Norflurazon</td>
<td>25.8</td>
<td>100.0</td>
<td>0.368</td>
</tr>
<tr>
<td>Naproxen</td>
<td>26.5</td>
<td>95.3</td>
<td>0.256</td>
</tr>
<tr>
<td>Vinclozolin M2</td>
<td>30.0</td>
<td>99.0</td>
<td>0.174</td>
</tr>
<tr>
<td>Terbufos sulfone</td>
<td>31.1</td>
<td>100.0</td>
<td>0.568</td>
</tr>
<tr>
<td>Malathion</td>
<td>32.3</td>
<td>100.0</td>
<td>0.057</td>
</tr>
<tr>
<td>Parathion</td>
<td>36.0</td>
<td>100.0</td>
<td>0.379</td>
</tr>
<tr>
<td>S-Bioallethrin</td>
<td>42.3</td>
<td>98.4</td>
<td>0.038</td>
</tr>
<tr>
<td>Chlorpyrifos-ethyl</td>
<td>43.7</td>
<td>100.0</td>
<td>0.005</td>
</tr>
</tbody>
</table>
3.4. Analysis of Primary Settled Municipal Wastewater and Treated Wastewater

Following the development of the integrated high-resolution mass spectrometry analysis platform, we began analyses of primary settled municipal wastewater and treated wastewater samples from the NMSU engineering team. We are in the process of finalizing those observations into several publications (see section 4.1 Near-term priorities). We omit the engineering aspects of this project, as that is not the charge of this project. Rather, we demonstrate the operational output of the capability developed by this project.

To illustrate a portion of our platform, we include data from a collaboration with Dr. Khandan’s group at the NMSU College of Engineering (Figures 4-6, below). A schematic of the treatment process is shown in Figure 6.

![Figure 6. Schematic of the sampling points (circles) for a comparison between standard treatment and an algal-based treatment process executed by Dr. Khandan's group, using City of Las Cruces municipal wastewater.](image)

3.4.1 Broadband Compositional Analysis

For each sampling point, Figure 7 shows broadband mass spectra collected by the Orbitrap Fusion mass spectrometer in direct infusion mode, where ~4000 – 7,000 individual mass spectral signals are observed, per sample. The broadband mass spectra are distilled in various ways to observe compositional change throughout the treatment process.
Figure 6. High resolution Orbitrap Fusion mass spectra for Las Cruces municipal wastewater and process related treatment samples using an algal reactor (collaboration with N. Khandan, NMSU Engineering). The evenly spaced spectral signals observed in raw wastewater, roughing filter and algal reactor day 0 samples are from polymers that are eventually removed by the process.

Heteroatom class distributions (Figure 8) and compositional space coverage (Figure 9) of organic species present within wastewater samples derived from positive-ion ESI Orbitrap Fusion mass spectra (i.e., non-targeted analysis) are presented. For all water samples, we observe C$_n$H$_{2n}$O$_x$ species at the highest abundance; however, their composition is different within the raw wastewater versus the treated products. The double bond equivalent (DBE) versus carbon number plots for the raw wastewater and the first stages of treatment (i.e., after roughing filter and algal reactor day 0) show distributions of compounds characteristic of ethylene oxide and propylene oxide polymers. These distributions are mainly present at DBE values 0, 1, 4, 5, and 9, consisting of high degrees of alkylation (7$\leq$C#$\leq$60) and higher relative abundances of even carbon number species versus odd-carbon species.
Figure 7: Heteroatom class distributions derived from the (+) ESI mass spectra of the raw wastewater (gray), wastewater after the roughing filter (blue), wastewater after activated sludge treatment (red), wastewater at day 0 in algal reactor (green), and wastewater at day 6 in algal reactor (purple). The relative abundance of all species within heteroatom class groups were summed and only heteroatom class groups above 1% relative abundance are shown. The gray stripes within the Ox heteroatom class group represent the summed relative abundance of species with the double bond equivalents associated with high concentrations of ethylene oxide (C2H4O) and propylene oxide (C3H6O) polymers.

Variable polymer distributions affect the observed relative abundance of all Ox species, and the relative abundance of all Ox species with DBE values 0, 1, 2, 4, 5, and 9 have been plotted as gray diagonal stripes over the total Ox abundance (Figure 8). The raw wastewater, wastewater after the roughing filter, and the wastewater at day 0 within the algal reactor show that >88% of the relative abundance of the Ox species belongs to species that contain DBE values of the main polymer distributions – these compound types are omnipresent in municipal wastewater and are components of soaps, shampoos, lubricants, and coolants, for example. However, only 60% and 42% of the relative abundance (i.e. proportion of total monoisotopic signal magnitude) of the Ox species within the wastewater after the activated sludge treatment and the wastewater at day 6 within the algal reactor belong to DBE values associated with the main polymer distributions.
The compositional space coverage (Figure 9) exhibited by the O\textsubscript{x} species within the wastewater after the activated sludge treatment and the wastewater after six days within the algal reactor are more consistent with dissolved organic matter, where the relative abundance of the even- and odd-carbon number species are similar and show a gradual decrease in abundance with increasing alkylation and aromaticity (increasing DBE value) through the treatment progression. The wastewater at Day 6 within the algal reactor has the most diverse composition due to organic input from the algae (e.g., cellular exudates). However, both the wastewater after the activated sludge treatment and the wastewater after six days in the algal reactor still show compounds characteristic of ethylene oxide polymers, mostly at DBE 4 and 5.

These results illustrate successful broadband characterization of organic components at the level of elemental composition, which is one objective of our approach (the other being tandem MS/MS identification of specific contaminants).
3.5. Nanoflow LC and MS/MS Identification of Contaminants

Additionally, we have performed nanoflow liquid chromatography with Orbitrap Fusion mass spectrometry on the raw wastewater and algal treatment system samples. We identified 61 database-present compounds with excellent specificity. Namely, all parent compounds were matched with sub-part-per-million mass measurement accuracy and high fragment ion spectral matching scores (e.g. >700 reverse match score). Table 2 lists 23 contaminants that were observed to be either partially or completely removed (i.e., below detection limit) by the algal wastewater treatment system. Those compounds include prescription drugs (antidepressants, beta blockers, antiarrhythmic agents, antibiotics, blood pressure medication, etc.), NSAIDS (non-steroidal anti-inflammatory drugs), opioids, insecticide and industrial chemicals (corrosion inhibitor/anti scaling agent). We detected thirty-one database matched compounds in the algal treatment “day 0” sample and six in the “Day 6” sample that were not observed in the raw wastewater, most of which are biologically derived materials such as enzymes, metabolites, natural alkaloids, and vitamins.
Table 2. Algal wastewater treatment remediation of contaminants. Fold-change differences (i.e., the measure describing how much a quantity changes between an original and a subsequent measurement) are specified for samples collected on the first day of treatment and six days later. “ND” indicates that a compound was not observed in the treated sample. Detection limits for each compound have not yet been established.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Fold-change vs. raw wastewater</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Algal Treatment,</td>
</tr>
<tr>
<td></td>
<td>Day 0</td>
</tr>
<tr>
<td>Hydroxymorphinan</td>
<td>0.54</td>
</tr>
<tr>
<td>Hydroxy eicosatetraenoic acid</td>
<td>ND</td>
</tr>
<tr>
<td>Naltrexol</td>
<td>ND</td>
</tr>
<tr>
<td>Atenolol</td>
<td>ND</td>
</tr>
<tr>
<td>Benzoylecgonine</td>
<td>0.73</td>
</tr>
<tr>
<td>Benzyldimethyltetradecy lammor</td>
<td>ND</td>
</tr>
<tr>
<td>Caffeine</td>
<td>0.90</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>0.61</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>ND</td>
</tr>
<tr>
<td>Desacetyldiltiazem</td>
<td>ND</td>
</tr>
<tr>
<td>Desmethylverapamil</td>
<td>ND</td>
</tr>
<tr>
<td>Diethyltoluamide</td>
<td>0.74</td>
</tr>
<tr>
<td>Flecainide</td>
<td>ND</td>
</tr>
<tr>
<td>Levofoxacin</td>
<td>ND</td>
</tr>
<tr>
<td>Lidocaine</td>
<td>ND</td>
</tr>
<tr>
<td>Lumichrome</td>
<td>0.53</td>
</tr>
<tr>
<td>Metoprolol</td>
<td>0.90</td>
</tr>
<tr>
<td>Naproxen</td>
<td>ND</td>
</tr>
<tr>
<td>Oxycodone</td>
<td>ND</td>
</tr>
<tr>
<td>Desmethylvenlafaxine</td>
<td>0.54</td>
</tr>
<tr>
<td>Sitagliptin</td>
<td>ND</td>
</tr>
<tr>
<td>Trimethoprim</td>
<td>0.80</td>
</tr>
<tr>
<td>Valsartan</td>
<td>0.68</td>
</tr>
<tr>
<td>Venlafaxine</td>
<td>ND</td>
</tr>
</tbody>
</table>
4. Conclusions

4.1. Established Resource

With this project we have developed a powerful and dynamic platform for the analysis of organic compounds in water samples. Novel analytical performance in the form of high resolution, accurate mass measurement and a robust database matching approach will provide excellent validation and monitoring capability for projects that seek the development of treatment and usage methods for impaired water sources. The ability to describe broadband chemical composition for dissolved organic matter as well as to monitor specific contaminants, is a dual-level approach that is both unique and necessary. This approach avoids the pitfalls associated with low-resolution mass spectrometry and targeted approaches.

4.2. Near-term Priorities

We are in the process of finalizing several key datasets that will demonstrate the utility of this platform. First, we are summarizing our observations for RO/membrane-treated wastewater samples provided by Dr. Xu for the following manuscript:


Second, we will combine some of the data illustrated above for a paper on the algal wastewater treatment system:


Finally, we have the first data collected for spiked wastewater samples, and several additional calibrations will be performed in the coming weeks. These data will establish all performance metrics for our method including detection limits, linear dynamic range, and false positive/negative rates for a variety of pesticides, herbicides, and contaminants of emerging concern. Thereafter, we will draft a manuscript that describes the sample preparation, instrumentation parameters, and data analysis aspects that are novel to our approach:

References


Advanced Chemical Analysis Capability for Alternative Water Source Research


Advanced Chemical Analysis Capability for Alternative Water Source Research


