SUSPECT AND NON-TARGET SCREENING OF ORGANIC MICROPOLLUTANTS IN WASTEWATER THROUGH THE DEVELOPMENT OF A LC-HRMS BASED WORKFLOW

Pablo Gago-Ferrero

Laboratory of Analytical Chemistry
Department of Chemistry
University of Athens

Emerging pollutants (EPs)

- ✓ Pharmaceuticals
- ✓ Illicit drugs
- ✓ Personal care products
- ✓ Endocrine disruptive compounds (EDCs)
- ✓ Flame retardants
- ✓ Food additives
- ✓ Disinfection by-products
- ✓ Pesticides

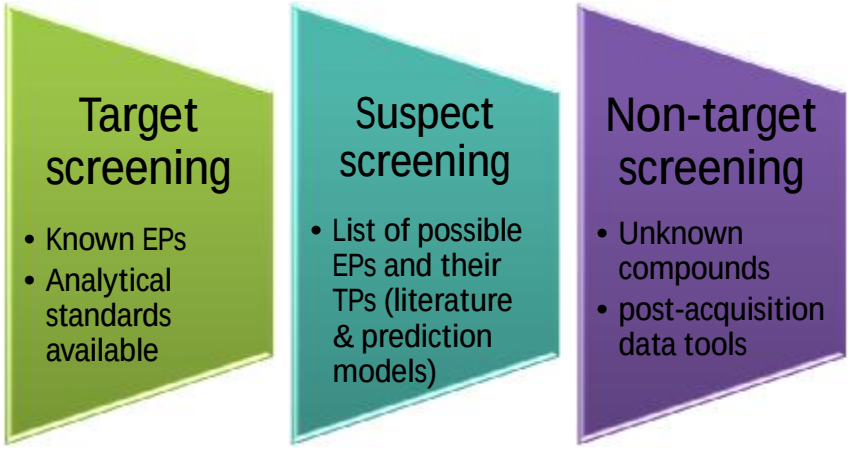
+

- ✓ Metabolites &
- ✓ Transformation Products (TPs)

aquatic environment

Wastewaters: Potentially tens of thousands of substances

Identification Approaches



Toxicity:

✓ Target / Non-target compounds

✓ Not identified compounds

Holistic view of risk: Target-based environmental monitoring should necessarily be accompanied by non-targeted analysis.

Objectives

Objectives

✓ Development and optimization of an integrated workflow to detect simultaneously target, suspect and unknown organic contaminants in wastewater samples, using liquid chromatography quadrupole-time-of-flight mass spectrometry (LC–QToF-MS).

✓Application of the developed methodology to the analysis of real wastewater.

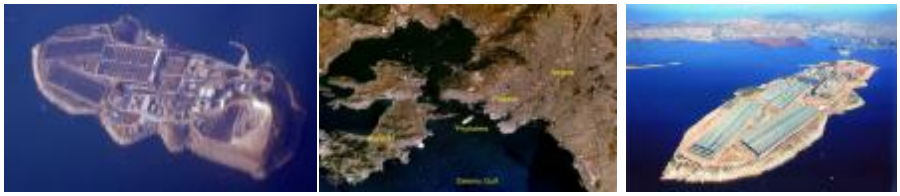
Sampling

Location: WWTP of Athens, Greece

Period: March 2014

Samples:

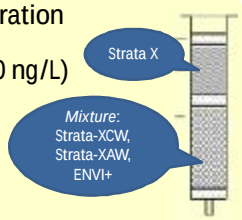
- 24-h composite flow-proportional samples of influent wastewaters & effluent wastewaters over a week (7 consecutive days)
- 2-h composite flow-proportional samples of influent wastewaters (Thursday & Saturday, 12 samples per day, from 02:00 to 00:00)



Sample Preparation – Instrumental Analysis

Sample preparation

- ✓ 200 mL filtered wastewater ⟶ 100 times preconcentration
- ✓ Isotopically labelled internal standards were spiked (100 ng/L)
- ✓ SPE ⟶ Mixed-bed cartridges
- ✓ Extraction: Neutral, Basic & Acidic Compounds



*Kern et al. Environmental Science and Technology (2009) 43(18):7039

Instrumental analysis:

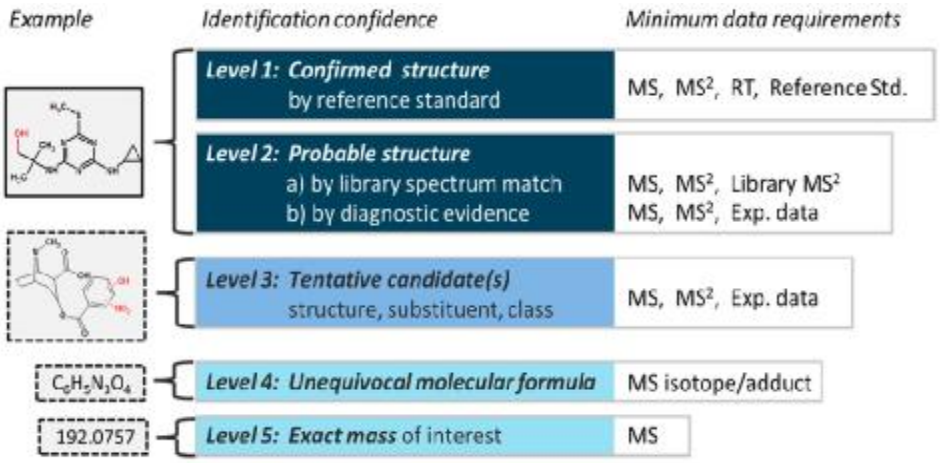
HPLC-HRMS-QToF-MS/MS

MS & MS/MS data in a single run

Target & Suspect Analysis: bbCID

Non-target Screening: AutoMS/MS

Levels of Identification Confidence



Schymanski et al. Environmental Science and Technology (2014) 48(4):2097

Suspect Screening

1. in-house database

- *more* than 10000 EPs and TPs

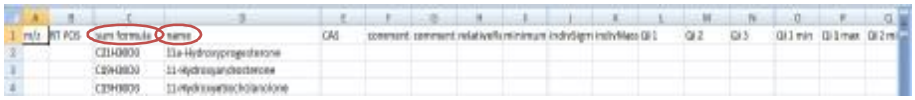
I. Human Metabolites *(Metabolite Predict, Bruker)*

II. Transformation Products *(UM-PPS, literature)*

III. Pharmaceuticals-Toxicological relevant compounds

IV. NORMAN association list of EPs of concern

...including information over:



2. Retention time prediction tool
KNN-GA-SVM

3. - High Resolution Mass Spectral Libraries
- In Silico fragmentation softwares (MassBank, MetFrag)

Suspect Screening – Reduction of features

I. Human Metabolites *(Metabolite Predict, Bruker, literature)*

1345 metabolites

III. Pharmaceuticals-Toxicological relevant compounds

5480 compounds

II. Transformation Products *(UM-PPS, literature)*

1835 TPs

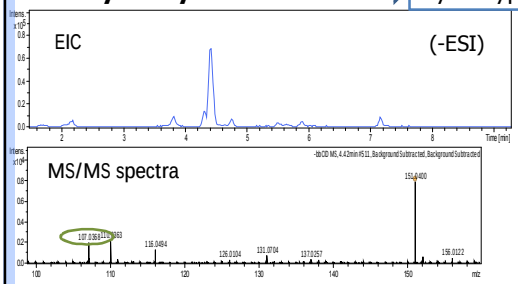
IV. NORMAN association list of EPs of concern

1200 compounds

Most intense and most relevant hits are prioritized for *further evaluation*

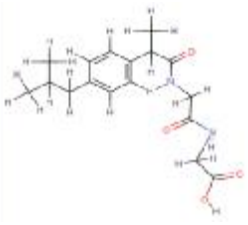
Suspect Screening – Human Metabolites

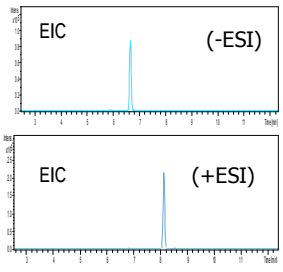
Acetylsalicylic acid Met12 → Hydroxyphenyl acetate



t_R = 4.43 min Predicted t_R = 4.39 min
Accuracy: 1.3 ppm
Isotopic fit: 0.6 mSigma

Ibuprofen Met13





t_R = 6.6 min Predicted t_R = 8.2 min
Accuracy: 1.3 ppm
Isotopic fit: 8.2 mSigma

t_R = 8.1 min Predicted t_R = 9.4 min
Accuracy: 1.7 ppm
Isotopic fit: 21.1 mSigma

Not available MS/MS data

Non-Target Screening – Introduction

WHY NON-TARGET?

TARGET SCREENING	✓ Known substance ✓ Reference standard available	✓ Unequivocal identification ✓ Possible quantification
SUSPECT SCREENING	✓ Suspect substance ✓ No reference standard available	✓ Qualitative detection possible

What proportion of substances present in the samples are actually detected with target and suspect screening?

Non-Target Screening - Introduction

- ✓ Many of the most intense peaks do not correspond to substances included in the target and suspect screening lists.
- ✓ These substances are potentially relevant, due to their high concentration.

✓ Identification of these substances is environmentally relevant

NON-TARGET SCREENING

- ✓ No former information on the analytes
- ✓ Molecular structures can be assigned on the basis of the exact mass, isotopic pattern and fragmentation information

✓ *Nevertheless, full identification of unknown compounds is often difficult & there is no guarantee of a successful outcome*

Non-Target Screening - Methodology

PROPOSED WORKFLOW

Full scan (MS) and Product ion spectra (MS/MS)
Accurate mass measurements

↓

Blank subtraction

↓

Automatic peak detection using Algorithms
(High number of peaks)

↓

Peak prioritization

↓

Determination of the Elemental compositions of the unknowns

↓

Determination and evaluation of candidates
(Tentative) Identification of TPs
- *Interpretation of the fragmentation pathway*
- *Chromatographic retention time plausibility*

↓

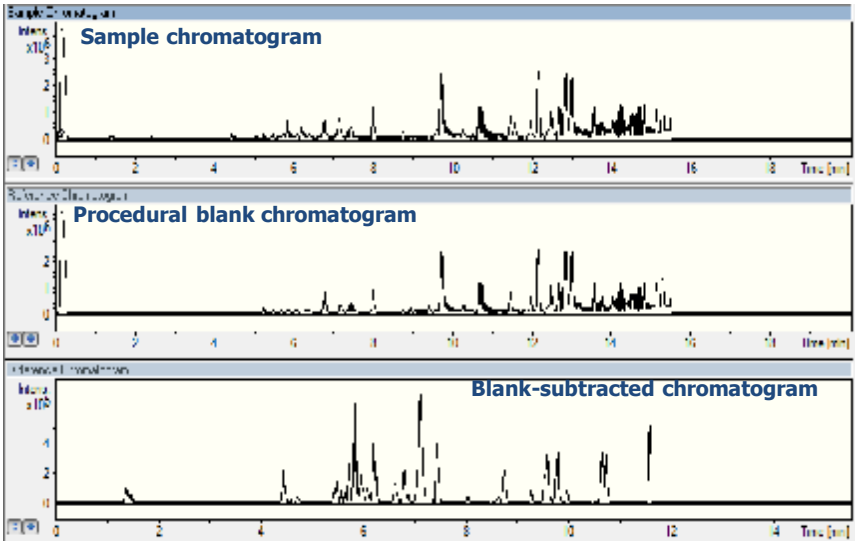
Confirmation: RT and MS/MS of chemical standards, when available

Non-Target Screening – Methodology

BLANK SUBTRACTION

✓Use of metabolomics tools

Metabolite Detect



PEAK PEAKING PROCEDURE

✓ Peak peaking: **Molecular features Algorithm**

- Using *Data analysis* and *Target analysis (Bruker)*
- Threshold: Signal/Noise >10

➡ A high number of peaks (> 3500) was obtained

#	RT [min]	Area	Int. Type	I	S/N	Max. m/z
1	1,1	12025,0	MolFeature	1108	12,1	151,0084
2	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
3	1,1	30852,8	MolFeature	2927	11,1	181,038
4	1,1	150714,1	MolFeature	14085	76,1	232,9988
5	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
6	1,1	74252,5	MolFeature	7411	44,9	166,9858
7	1,1	13137,9	MolFeature	1533	17	261,9721
8	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
9	1,1	9764,1	MolFeature	472	10,5	634,8444
10	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
11	1,1	9017,2	MolFeature	927	10,3	406,9612
12	1,1	8880,0	MolFeature	804	17,9	444,9128
13	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
14	1,1	73703,8	MolFeature	5180	57,6	378,9341
15	1,1	8340,0	MolFeature	628	14	334,9604
16	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
17	1,1	51524,7	MolFeature	4804	32,5	181,0011
18	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
19	1,1	9166,9	MolFeature	1016	11,3	378,8721
20	1,1	142217,7	MolFeature	10180	62,1	133,034
21	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]
22	1,1	109310,9	MolFeature	16292	36,8	216,9499
23	1,1	17145,8	MolFeature	1187	15,2	412,9716
24	1,1	[illegible]	MolFeature	[illegible]	[illegible]	[illegible]

PRIORITIZATION OF PEAKS FOR FURTHER EVALUATION

✓ Selection of the most relevant from the large peak list
(Not included either in the target or the suspect screening)

Criteria:

- Intensity
- Presence of a distinctive isotopic pattern

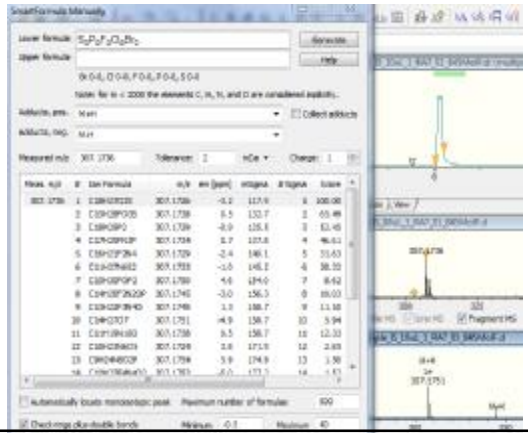
Non-target identification was performed on selected masses from the top most intense peaks

DETERMINATION OF ELEMENTAL COMPOSITION

1st step: **Generation of possible molecular formula(s)**

Criteria:

- Mass accuracy → threshold: 5 ppm / 2 mDa
- Agreement of the theoretical and measured isotopic pattern



DETERMINATION OF ELEMENTAL COMPOSITION: SEVEN GOLDEN RULES (SGR)

✓**Plausibility of the generated molecules** → Use of the **Seven Golden Rules** software

"Seven golden rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry"

i. Element number restrictions
ii. Lewis and Senior chemical rules check
iii. Isotopic pattern filter
iv. Hydrogen/carbon ratio check
v. Element ratio of nitrogen, oxygen, phosphorus and sulphur vs carbon check
vi. Element ratio probability check
vii. Check of the presence of trimethylsilylated compounds

30 million compounds database → Great reduction of the possibilities

✓ The correct molecular formula is assigned with a probability of 98%, if the formula exists in a compound database

Kind and Fiehn. BMC Bioinformatics 8:105 (2007)

Non-Target Screening – Methodology

EVALUATION OF POSSIBLE CANDIDATES

- ✓ Number of candidates to one molecular formula: **1** - **>2000**
  (*Chemspider, Pubmed databases*)

Approaches for tentative identification:

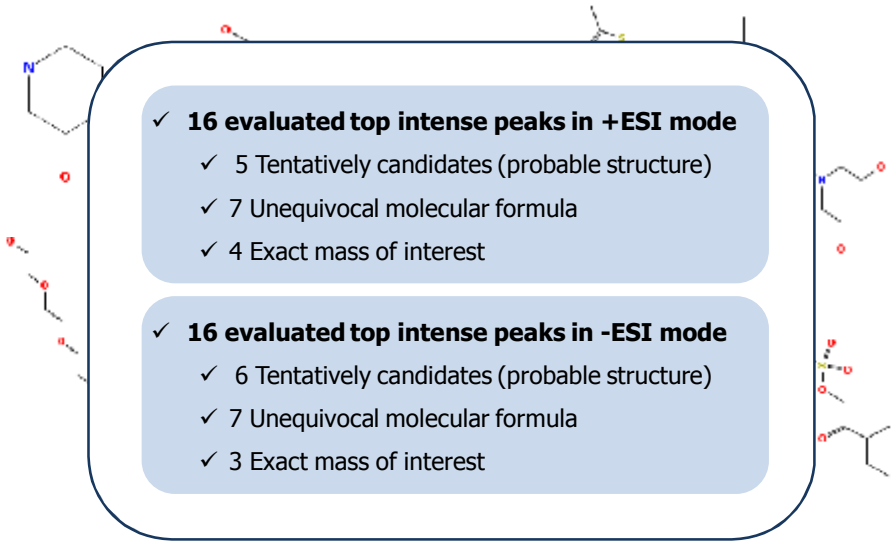
✓**Databases** (e.g. MassBank) →*Still very limited number of compounds (not very useful for non-target screening)*
✓Deep **study** of the **MS/MS spectra (AutoMSMS analysis)**
✓***In-silico*** **fragmentation software**
- Smart formula 3D *(Bruker)*
- Metfrag

✓**Chromatographic retention time** plausibility → Application of models
✓**Number of data sources and references** in different data bases (e.g. Chemspider)

- ✓ To confirm the identity of a substance, purchase of reference standard is required (if available)

Non-Target Screening - Results

RESULTS



- ✓ **16 evaluated top intense peaks in +ESI mode**
  - ✓ 5 Tentatively candidates (probable structure)
  - ✓ 7 Unequivocal molecular formula
  - ✓ 4 Exact mass of interest
- ✓ **16 evaluated top intense peaks in -ESI mode**
  - ✓ 6 Tentatively candidates (probable structure)
  - ✓ 7 Unequivocal molecular formula
  - ✓ 3 Exact mass of interest

Acknowledgments

Nikolaos Thomaidis
Katerina Psoma
Anna Bletsou
Reza Aalizadeh

This research has been co-financed by the European Union and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF)

Thank you for your attention!

National and Kapodistrian
UNIVERSITY OF ATHENS
Faculty of Chemistry

http://trams.chem.uoa.gr/