NON-TARGET APPROACH FOR THE DETERMINATION OF NOVEL MICROPOLLUTANTS IN WASTEWATER USING LIQUID CHROMATOGRAPHY QUADRUPOLE-TIME OF FLIGHT MASS SPECTROMETRY (LC-QTOF-MS)

Pablo Gago-Ferrero[1], Anna A. Bletsou, Reza Aalizadeh, Nikolaos S. Thomaidis

National and Kapodistrian University of Athens. Laboratory of Analytical Chemistry, Department of Chemistry, Panepistimiopolis Zographou, 15771 Athens, Greece

INTRODUCTION

Wastewaters contain a very large list of micropollutants and transformation products of environmental concern. All these (mostly) synthetic organic chemicals enter the wastewater treatment plants (WWTP) with influents and due to incomplete or zero removal are released in the aquatic environment. Thus, the study of the fate of the emerging pollutants and their transformation products in WWTPs is of paramount environmental importance and can also provide valuable information related to consumption trends.

Target screening procedures are limited to a small fraction of these substances, due to the inability to obtain standards for all that substances and the ignorance of the existence of many of them. Recent advances in high resolution mass spectrometry (HRMS) have opened up new windows of opportunity in the field of complex samples analysis. Suspect screening, with suspected substances based on prior information but with no reference standard, is a powerful tool which allows a large increment in the number of compounds to be evaluated. However, in most cases many of the peaks showing greater intensity not correspond to substances included in the target and suspect screening lists. These substances are potentially relevant, due to their high concentration, and their identification is environmentally important. Nevertheless, full identification of unknown compounds is often difficult and there is no guarantee of a successful outcome. The aim of the present work is the development and application of a workflow for the tentative identification of relevant unknown substances (not detected in the previously applied target and suspect methods) using liquid chromatography quadrupole-time-of-flight mass spectrometry (LC–QToF-MS).

NON-TARGET SCREENING WORKFLOW

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

- **LC-QTOF-MS**

Blank subtraction

- Use of metabolomic tools: (*Metabolite detect, Bruker*).

Peak peacking and prioritization

- Algorithm based on molecular features.
- Prioritization of peaks criteria:
  - Intensity.
  - Distinctive isotopic pattern.

Determination of the elemental compositions of the unknowns

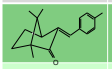
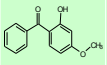
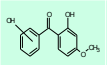
- Mass accuracy → threshold: 5 ppm.
- Agreement of the theoretical and measured isotopic pattern.
- Application of the Seven Golden Rules to asses the plausibility of the generated molecules.

Determination and evaluation of candidates (Tentative) Identification of TPs

- Databases (e.g. MassBank; Very limited data).
- Deep study of the MS/MS spectra.
- Use of *In-Silico* fragmentation software.
  - Smart formula 3D (Bruker).
  - Metfrag.
- Chromatographic retention time plausibility → application of an in-house developed model.

Figure 1. Flowchart of the non-target screening methodology.

IDENTIFICATION LEVELS

Example	Identification confidence level	Minimum data requirements
	**Confirmed Structure** *By reference standard*	MS, MS^2, RT, Reference standard
	**Probable Structure** *Using spectra database or by diagnostic evidence*	MS, MS^2 and Library MS^2 or Exp. data
	**Tentative candidate(s)** *Structure, substituent, class*	MS, MS^2, Exp. data
$C_8H_{22}O_3N_2$	**Unequivocal molecular formula**	MS isotope/adduct
195.1233	**Exact mass of interest**	MS

Confirmation

Retention time and MS/MS of chemical standards, when available.

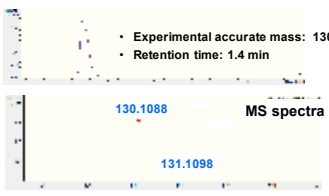
RESULTS

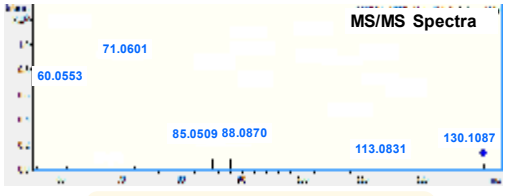
Table 1. Identification of unknown compounds corresponding to the most intense peaks.

Retention time (min)	Mass of ion [m/z] (peak of component)	Ion type	Intensity	Molecular formula	Proposed identification name	Level of confirmation of identification
1.28	164.1282	$[M+H]^+$	1508655	C7H17NO3		Unequivocal molecular formula
1.91	145.0977	$[M+H]^+$	2186079	C6H12N2O2	e.g. 4-(2-Hydroxyethyl)-2-piperazinone	Tentative candidates
2.27	96.0452	$[M+H]^+$	1145713	C5H5NO	2-Formyl-1H-pyrrole	Probable structure
4.19	195.1233	$[M+H]^+$	1405658	C8H18O5	tetraethyleneglycol	Tentative candidate
4.68	135.1018	$[M+H]^+$	1122821	C6H14O3		Unequivocal molecular formula
4.98	424.1857	$[M+H]^+$	1263654			Exact mass of interest
5.09	358.2078	$[M+NH4]^+$	1264684	C15H24N4O5		Unequivocal molecular formula
5.16	283.1753	$[M+H]^+$	1262520	C13H22N4O3		Unequivocal molecular formula
5.2	468.2108	$[M+H]^+$	1263126			Exact mass of interest
5.73	149.1176	$[M+H]^+$	1688072	C7H16O3		Unequivocal molecular formula
6.13	520.333	$[M+H]^+$	1262524			Exact mass of interest
6.44	608.3854	$[M+H]^+$	1262588			Exact mass of interest
9.1	232.1913	$[M+H]^+$	1160646	C12H25NO3	e.g. N,N-Bis(2-hydroxyethyl)octanamide	Tentative candidates
9.4	191.1647	$[M+H]^+$	1410087	C10H22O3		Unequivocal molecular formula
12.69	316.1955	$[M+H]^+$	1137576	C16H29NO3S	e.g. 1-{(2-Methoxyethyl)[(5-methyl-2-thienyl)methyl] amino}-3-[(2-methyl-2-propanyl)oxy]-2-propanol	Tentative candidates

The developed non-target approach was applied to a real influent wastewater sample from the WWTP of Athens. Fifteen peaks selected on the basis of intensity were evaluated. Results are summarized in Table 1. The developed workflow allowed the obtaining of unequivocal molecular formulas for most of the selected peaks and plausible candidates in some cases. Metformin was detected amongst the most intense peaks by using a target approach. Example 1 shows the application of the workflow to this peak, treating it as an unknown, to check the validity of the method. Example 2 describes the application of the workflow to a real unknown, providing four tentative candidates.

Example 1: Application of the workflow to metformin, treating the peak as unknown





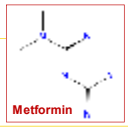
Number of possible formulas → 1
(Threshold of 5 ppm and 50 mSigma)

$C_4H_{12}N_5$

✓ **Chemspider Hits : 12**

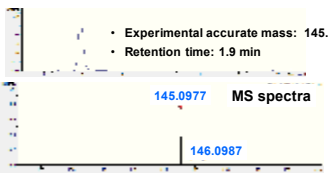
✓ **Compounds with score > 0.8 ➡ 4**

Candidates:



Metfrag Score:	1	0.86	0.81	0.81
Explained Fragments:	5	3	2	2
RT Pred. Model:	✓	✓	X	X
Chemspider data sources:	59	5	4	4
Chemspider References:	293	6	4	4

Example 2: Application of the workflow to a real unknown: tentative identification



Number of possible formulas → 1
(Threshold of 5 ppm and 50 mSigma)

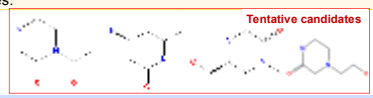
$C_6H_{12}N_2O_2$

✓ **Hits Chemspider: 336**

✓ **Compounds with score > 0.9➡ 28**

✓ **Only five with more than 3 fragment matches**

Candidates:



Metfrag Score:	1	0.97	0.95	0.95	0.95
Explained Fragments:	4	4	4	4	3
RT Pred. Model:	✓	✓	✓	✓	X
Chemspider data sources:	1	4	5	17	2
Chemspider References:	1	4	5	18	2

Acknowledgements

This project was implemented under the Operational Program «Education and Lifelong Learning» and funded by the European Union (European Social Fund) and National Resources – ARISTEIA 624 (TREMEPOL project).