IDENTIFICATION OF UNKNOWNS IN REAL WASTEWATER THROUGH THE APPLICATION OF A LC-QTOF-MS BASED WORKFLOW

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

Wastewater contains a high number of organic micropollutants and transformation products of environmental concern. Recent approaches, combining methodologies based on target and suspect screening (for suspected substances based on prior information but with no reference standard) are important for the comprehensive characterization of environmental samples. Nevertheless, samples still contain many chromatographic peaks which do not correspond to substances included in target and suspect screening lists. These substances may be potentially relevant (e.g. due to their concentration or potential effects) and thus the identification of selected non-targets is important. However, full identification of unknown compounds is often difficult and there is no guarantee of a successful outcome. The aim of this work is to show some specific examples on the identification of unknown compounds in real wastewater (collected from the WWTP of Athens). Identifications were conducted using a developed integrated workflow based on LC–QToF-MS to detect formerly unknown organic contaminants in wastewater.

NON-TARGET SCREENING WORKFLOW

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

⬇

Blank subtraction

⬇

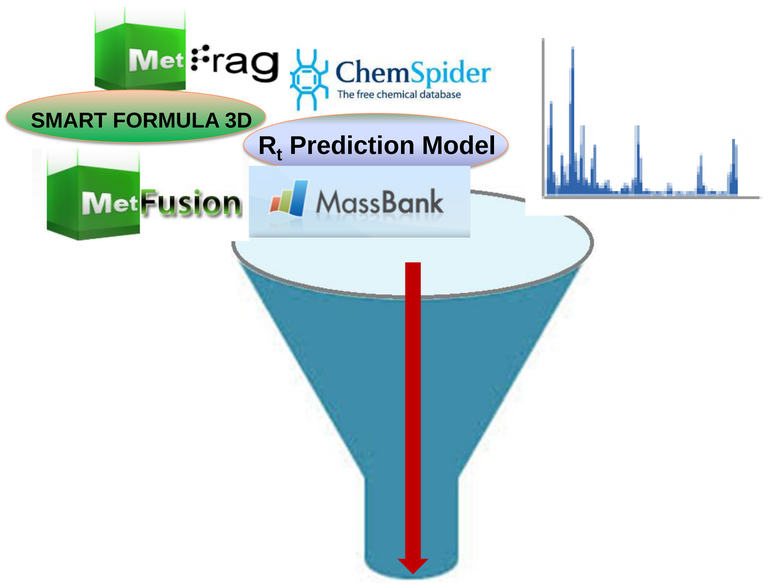
Peak peaking and prioritization

⬇

Determination of the elemental compositions of the unknowns

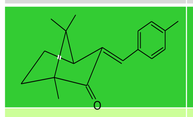
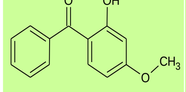
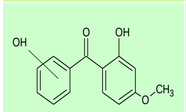
- **Mass accuracy**
- **Isotopic fit**
- **Seven Golden Rules (SGR)** [1] (to assess the plausibility of the generated molecules).

Determination and evaluation of candidates



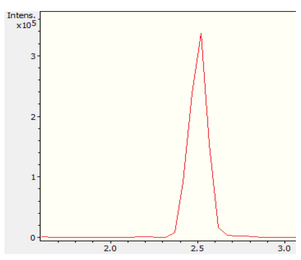
(Tentative) Identification

LEVELS OF IDENTIFICATION

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

NON-TARGET EXAMPLES ON THE IDENTIFICATION OF UNKNOWN COMPOUNDS

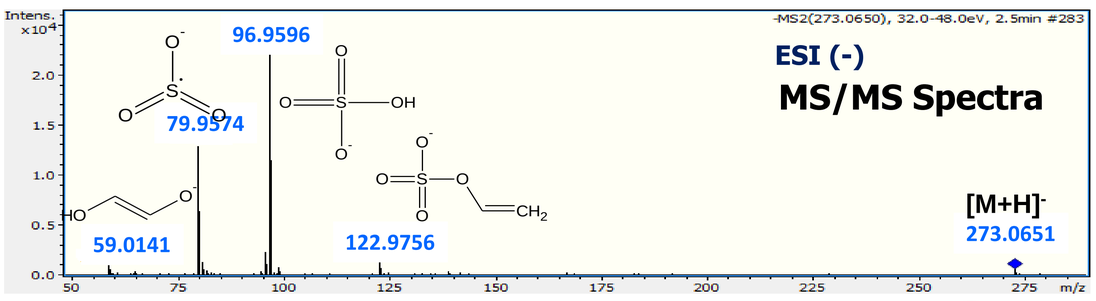
Example I: AGN-PC-03193Q



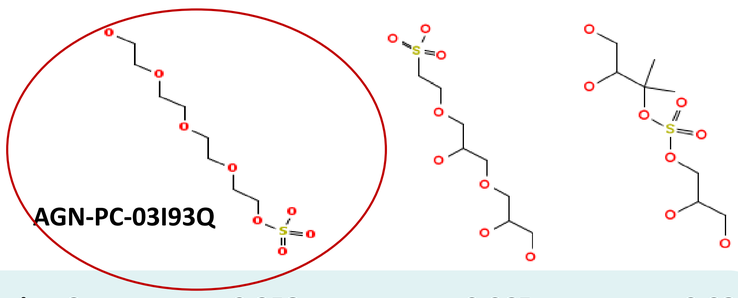
- Experimental accurate mass: 273.0651
- t_R= 2.4 min
- High intensity in ESI(-)

➡

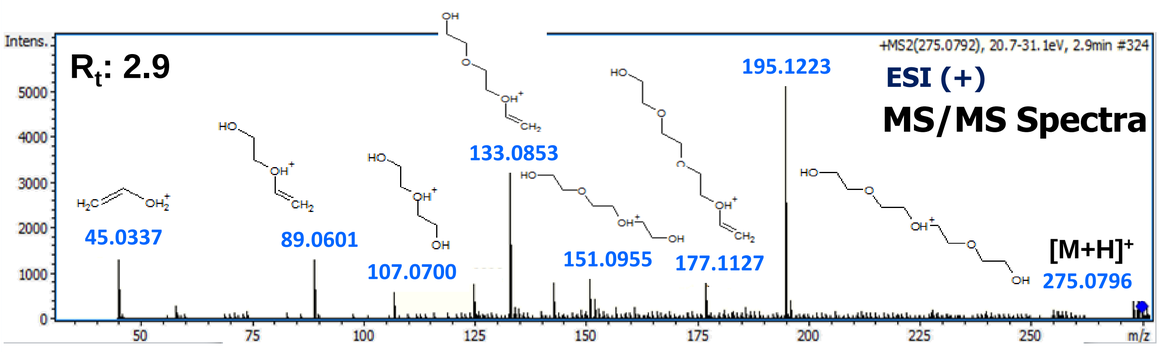
- **Number of possible formulas** (*2 mDa, 5 ppm, 50 mSigma)* → **11**
- **After Seven Golden Rules → 3**



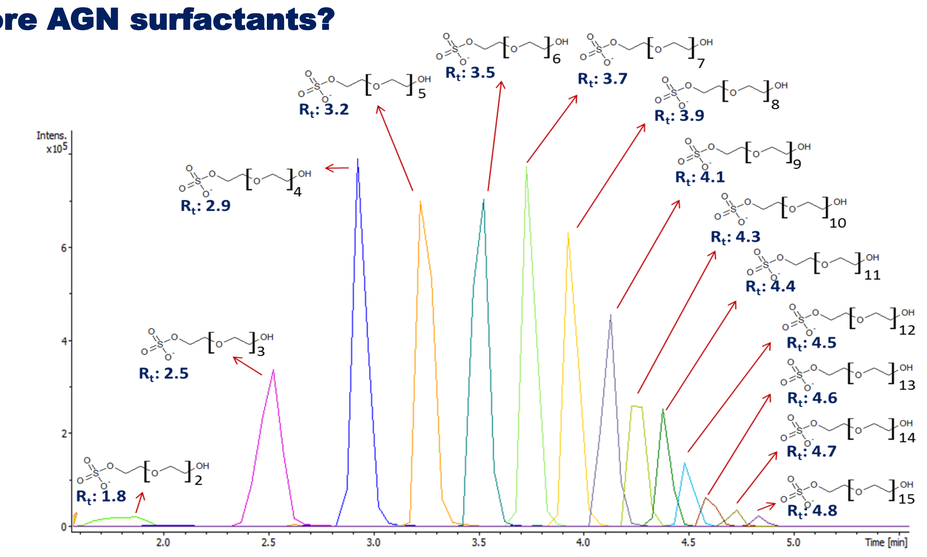
	$C_9H_{11}FN_4O_5$	$C_{12}H_{10}F_4N_2O$	$C_8H_{18}O_8S$
✓Hits Chemspider / Pubmed	5	23	3
✓Compounds with MetFusion matches	0	0	3



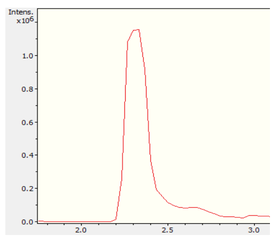
✓MetFusion Score:	0.258	0.065	0.035
✓Explained Fragments:	4	2	1
✓RT Pred. Model:	✓	✓	X
✓Number of patents:	104	2	1



More AGN surfactants?



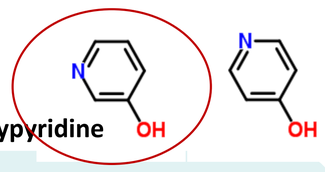
Example II: 3-Hydroxypyridine



- Experimental accurate mass: 96.0446
- t_R= 2.27 min
- High intensity in ESI(+)

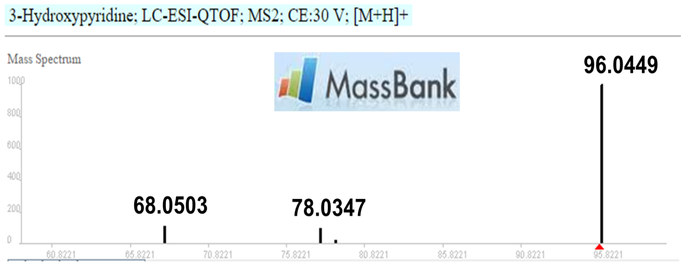
➡ **Number of possible formulas** (*2 mDa, 5 ppm, 50 mSigma)* → **1** ➡ C_5H_5NO

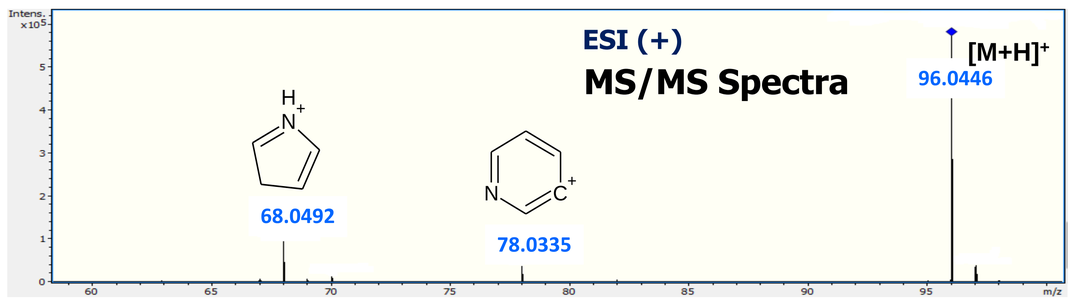
- ✓**Hits Chemspider: 38**
- ✓**Compounds with MetFusion score > 0.5 → 2**



- ✓ Identity confirmed with a commercial standard → **Level 1**

✓MetFusion score:	1.117	0.786
✓Explained Fragments:	2	2
✓RT Pred. Model:	✓	✓
✓Chemspider data sources:	116	105
✓Chemspider References:	287	243





The first example consists of the identification of the surfactant 2-[2-[2-(2-hydroxyethoxy)ethoxy]ethoxy]ethyl hydrogen sulfate (AGN-PC-03I93Q). After applying the mass accuracy and isotopic fit criteria 11 possible molecular formulas were found; after applying the SGR 3 plausible molecular formulas remained. Only the candidates with formula $C_8H_{18}O_8S$ could explain MS/MS fragments. The evaluation with *MetFusion* showed that AGN-PC-03I93Q was the only structure that explained all the fragments in the MS/MS spectrum. Moreover, this substance was the candidate with the highest commercial importance. Clear MS/MS spectra was also obtained in ESI(+) and again AGN-PC-03I93Q was the only compound that could explained all the fragments. With all these evidences a level of confidence 2b was assigned. At this point, a logical question was if more surfactants from the homologous series ($SO_4C_2H_4(OC_2H_4)_xOH$) were present. The presence of these compounds was screened and 14 were detected, most of them at high intensity. t_Rs were consistent with the model and increased constantly with the number of carbons. MS/MS spectra were also consistent in all cases. AGN surfactants were the compounds detected at the highest intensity in ESI(-), showing its high use in the evaluated area. Another example can be found in 3-hydroxypyridine. There were 38 compounds in *Chemspider* database corresponding to the determined molecular formula. Only two of them (3-hyroxypyridine and 4-hydroxypyruidine) had a good *MetFusion* score when evaluating the MS/MS spectra. Both compounds can explain the two fragments obtained in the MS/MS spectra, have plausible t_R according to the model and have high commercial importance. ESI(+) MS/MS spectra were available in *MassBank* for 3-hydroxypyridine and matched perfectly with the ones obtained experimentally. Standards of 3-hydroypyridine and 4-hydroxypyridine were purchased and through the comparison of the t_Rs the identity of the substance 3-hydroxypyridine was confirmed.

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). **References:** [1] Kind T and Fiehn O (2007) BMC Bioinformatics 8, 105.