

RAPPORT DE POSITIONNEMENT SUR L'UTILISATION DE LA SPECTROMETRIE DE MASSE HAUTE RESOLUTION POUR LE CRIBLAGE ENVIRONNEMENTAL

Thème F

Métrologie pour l'anticipation de la future surveillance et
l'amélioration des connaissances sur les substances émergentes

Togola, S. Lardy-Fontan, F. Lestremau, C. Soulier

Avec la collaboration de C Margoum et P. Bados

Décembre 2015

Programme scientifique et technique
Année 2015

Document final

En partenariat avec



Avec le soutien de



et de



Contexte de programmation et de réalisation

Ce rapport a été réalisé dans le cadre du programme d'activité AQUAREF pour l'année 2015 ...

Auteur (s) :

Anne Togola
BRGM
a.togola@brgm.fr

Sophie Lardy-Fontan
LNE
sophie.lardy-fontan@lne.fr

François Lestremau
INERIS
Francois.LESTREMAU@ineris.fr

Coralie Soulier
BRGM
c.soulier@brgm.fr

Christelle Margoum
IRSTEA
Christelle.margoum@irstea.fr

Philippe Bados
IRSTEA
Philippe.bados@irstea.fr

Vérification du document :

Valéria Dulio
INERIS
Valeria.DULIO@ineris.fr

Les correspondants

Onema : Pierre Francois Staub, pierre-francois.staub@onema.fr

BRGM : Jean Philippe Ghestem , jp.ghestem@brgm.fr

Référence du document : **Togola A, Lardy-Fontan S., Lestremau F., Soulier C. (2015)**
Rapport de positionnement sur l'utilisation de la spectrométrie de masse haute
résolution pour le criblage environnemental Rapport final. BRGM/RP-65420-FR 75,
65 p., 14 ill, 3 ann

Droits d'usage :	<i>Accès libre</i>
Couverture géographique :	<i>International</i>
Niveau géographique :	<i>National</i>
Niveau de lecture :	<i>Professionnels, experts</i>
Nature de la ressource :	<i>Document</i>

1	INTRODUCTION	13
2	LES APPROCHES SMHR MENANT A L'IDENTIFICATION	15
2.1	Approche quantitative.....	15
2.2	Approche qualitative.....	16
2.2.1	Criblage ciblé/ suspect	
2.2.2	Criblage non-ciblé	
3	INTERET DE L'INTEGRATION DE LA HRMS DANS LA SURVEILLANCE	21
3.1	acquisition de données quantitatives sur l'état chimique.....	21
3.2	acquisition de données qualitatives	21
3.2.1	Mise à jour des listes de substances d'intérêt : campagne prospective, liste de vigilance	
3.3	lien etat ecologique/pression chimique	21
3.3.1	Contrôle d'enquête : recherche/ identification de sources de pollutions, alerte, ...	
3.4	Intérêt de la bancarisation/ de la possibilité de traitement a posteriori..	25
3.5	les expériences européennes	25
3.5.1	Cas de la Norvège (Schlabach, Haglund et al. 2013)	
3.5.2	Joint Danube Survey (2015)	
3.5.3	Le programme d'analyse spéciale du CIPR (Commission internationale pour la protection du Rhin)	
3.6	conclusion sur l'applicabilité des approche SMHR a la surveillance	30
4	VERROUS ASSOCIES A CES NOUVEAUX OUTILS	31
4.1	Verrous analytiques	31
4.2	Verrous lies au traitement des donnees brutes	33
4.2.1	Prise en compte des blancs/contrôles qualités	
4.2.2	Paramètres à prendre en compte dans le retraitement des chromatogrammes	
4.2.3	Bases de données	
4.3	Verrous liés à l'échantillothèque	37
4.3.1	Bancarisation des métadonnées associées à l'analyse	
4.3.2	Bancarisation des métadonnées associées au retraitement	
4.3.3	Besoin que « n'importe qui » puisse retraiter les données acquises et bancarisées	
4.4	Poursuite des actions européenne et française en cours	38
5	CONCLUSIONS	39
6	Bilan sur le criblage de molécules en fonction des objectifs de la surveillance	40

Rapport de positionnement sur l'utilisation de la spectrométrie de masse haute résolution pour le criblage environnemental
A. Togola, S. Lardy-Fontan, F. Lestremau, C. Soulier

RESUME

La spectrométrie de masse pour le criblage non ciblé a connu un essor considérable ces dernières années dans de nombreux domaines dont celui de l'environnement. Plusieurs travaux sont désormais focalisés sur l'amélioration des connaissances pour la mise en œuvre de cette technique afin d'identifier de nouvelles molécules d'intérêts. Cependant il est important de disposer de textes de recommandations pour cette technique sur lesquelles s'appuyer afin d'éviter que des pratiques trop libérales soient mises en œuvre par les laboratoires privés au risque d'aboutir à des résultats ne correspondant pas la nature et composition de l'échantillon. L'intégration de ces nouvelles approches pour la surveillance réglementaire et de ce fait la mise en œuvre par des laboratoires prestataires nécessitera l'existence de cadre(s) et recommandations.

Ce rapport introduit les différentes approches pouvant mener à l'identification de composés d'intérêts ainsi que leur faisabilité et leur intérêt pour la surveillance. En effet cette technique apporterait des avantages pour la mise à jour des substances d'intérêt en simplifiant le travail conséquent de priorisation des molécules à rechercher. De plus l'information acquise non sélective enregistrée permet d'aller chercher une « nouvelle » molécule identifiée dans des échantillons *a posteriori*. Ces approches sont aussi importantes pour documenter l'état écologique des milieux en les couplant à tests biologiques dans le but de mettre en exergue les composés responsables d'une toxicité. De plus, sans aller jusqu'à l'identification, la comparaison d'empreinte moléculaire peut permettre de caractériser les sources de pollution ou mettre en évidence la présence de produits de transformation. Ces approches sont d'ores et déjà utilisées dans le cadre réglementaire dans certains pays européens. Les verrous associés à ces techniques sont aussi répertoriés dans ce document. Les futurs travaux de groupes de travail national et européen visent à apporter des éléments de réponse à ce sujet.

Mots clés (thématique et géographique) :

Spectrométrie de Masse Haute Résolution (SMHR), surveillance réglementaire, molécules organiques, criblage

ABSTRACTS

Mass spectrometry for non-targeted screening has expanded considerably in recent years in many areas, including the environment. Several studies are now focused on improving knowledge for implementation of this technique to identify new molecules of interest. However it is important to have recommendations texts for this technique on which to build in order to avoid overly liberal practices which are implemented by private laboratories may lead to results that do not match the nature and composition of the sample. The integration of these new approaches to regulatory oversight and therefore the implementation by provider's laboratories require the existence of framework and recommendations.

This report introduces the different approaches that may lead to the identification of compound of interest as well as their feasibility and their importance in monitoring. Indeed this technique would bring benefits to updating substance of interest by simplifying the prioritization's work of molecules to look for. Also recorded nonselective acquired information allows checking retrospectively if a "new" identified molecule is present in samples. These approaches are also important to document the ecological status of environments when coupled with biological tests in order to highlight the compounds responsible of toxicity. Also without going to the identification, comparison of molecular fingerprint can be used to characterize the sources of pollution or highlight the presence of transformation products. These approaches are already used in the regulatory framework in some European countries. Obstacles associated with these techniques are also listed in this document. Future works of national and European working groups aim to bring answers to this.

Key words (thematic and geographical area):

High Resolution Mass Spectrometry (HRMS), regulatory monitoring, organic compounds, screening



Rapport de positionnement sur l'utilisation de la spectrométrie de masse haute résolution pour le criblage environnemental

Rapport final,
BRGM/RP 65420-FR –
Décembre 2015

Rapport de positionnement sur l'utilisation de la spectrométrie de masse haute résolution pour le criblage environnemental

Rapport final

BRGM/RP-65420-FR

Décembre 2015

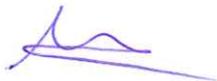
Étude réalisée dans le cadre de la convention ONEMA-BRGM 2013-2015

Togola A, Lardy-Fontan S., Lestremau F., Soulier C.

Avec la collaboration de
C.Margoum et P. Bados

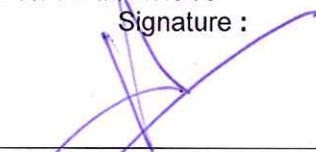
Vérificateur :

Nom : Nicole Baran
Fonction : Correspondant DAPP
Date : 04/02/2016
Signature :



Approbateur :

Nom : Hervé Gaboriau
Fonction Directeur :
Date : 24/02/2016
Signature :



Le système de management de la qualité et de l'environnement
est certifié par AFNOR selon les normes ISO 9001 et ISO 14001.

Mots-clés :

CORRESPONDANTS ONEMA : Pierre Francois Staub

En bibliographie, ce rapport sera cité de la façon suivante :

Togola A, Lardy-Fontan S., Lestremau F., Soulier C. (2015) Rapport de positionnement sur l'utilisation de la spectrométrie de masse haute résolution pour le criblage environnemental Rapport final. BRGM/RP-65420-FR 75, 65 p., 14 ill, 3 ann.

Synthèse

La spectrométrie de masse pour le criblage non ciblé a connu un essor considérable ces dernières années dans de nombreux domaines dont celui de l'environnement. Plusieurs travaux sont désormais focalisés sur l'amélioration des connaissances pour la mise en œuvre de cette technique afin d'identifier de nouvelles molécules d'intérêts. Cependant il est important de disposer de textes de recommandations pour cette technique sur lesquelles s'appuyer afin d'encadrer les pratiques potentiellement mises en œuvre par les laboratoires privés pour éviter d'aboutir à des résultats ne correspondant pas la nature et composition de l'échantillon. L'intégration de ces nouvelles approches pour la surveillance réglementaire et de ce fait la mise en œuvre par des laboratoires prestataires nécessitera l'existence de cadre(s) et recommandations.

Ce rapport introduit les différentes approches pouvant mener à l'identification de composés d'intérêts ainsi que leur faisabilité et leur intérêt pour la surveillance. En effet cette technique apporterait des avantages pour la mise à jour des substances d'intérêt en simplifiant le travail conséquent de priorisation des molécules à rechercher. De plus l'information acquise non sélective enregistrée permet d'aller chercher une « nouvelle » molécule identifiée dans des échantillons *a posteriori*. Ces approches sont aussi importantes pour documenter l'état écologique des milieux en les couplant à tests biologiques dans le but de mettre en exergue les composés responsables d'une toxicité. De plus, sans aller jusqu'à l'identification, la comparaison d'empreinte moléculaire peut permettre de caractériser les sources de pollution ou mettre en évidence la présence de produits de transformation. Ces approches sont d'ores et déjà utilisées dans le cadre réglementaire dans certains pays européens. Les verrous associés à ces techniques sont aussi répertoriés dans ce document. Les futurs travaux de groupes de travail national et européen visent à apporter des éléments de réponse à ce sujet.

Sommaire

1	Introduction.....	13
2	Les approches SMHR menant à l'identification.....	15
2.1	APPROCHE QUANTITATIVE	15
2.2	APPROCHE QUALITATIVE.....	16
2.2.1	Criblage ciblé/ suspect.....	17
2.2.2	Criblage non-ciblé.....	19
3	Intérêt de l'intégration de la HRMS dans la surveillance	21
3.1	ACQUISITION DE DONNEES QUANTITATIVES SUR L'ETAT CHIMIQUE ...	21
3.2	ACQUISITION DE DONNEES QUALITATIVES.....	21
3.2.1	Mise à jour des listes de substances d'intérêt : campagne prospective, liste de vigilance.....	21
3.3	lien etat ecologique/pression chimique	22
3.3.1	Contrôle d'enquête : recherche/ identification de sources de pollutions, alerte, ...	24
3.4	INTERET DE LA BANCARISATION/ DE LA POSSIBILITE DE TRAITEMENT A POSTERIORI.....	25
3.5	LES EXPERIENCES EUROPEENNES.....	25
3.5.1	Cas de la Norvège (Schlabach, Haglund et al. 2013).....	25
3.5.2	Joint Danube Survey (2015)	26
3.5.3	Le programme d'analyse spéciale du CIPR (Commission internationale pour la protection du Rhin).....	28
3.6	CONCLUSION SUR L'APPLICABILITE DES APPROCHE SMHR A LA SURVEILLANCE	30
4	Verrous associés à ces nouveaux outils	31
4.1	VERROUS ANALYTIQUES	31
4.2	VERROUS LIES AU TRAITEMENT DES DONNEES BRUTES.....	33
4.2.1	Prise en compte des blancs/contrôles qualités.....	33
4.2.2	Paramètres à prendre en compte dans le retraitement des chromatogrammes	34
4.2.3	Bases de données	35
4.3	VERROUS LIES A L'ECHANTILLOTHEQUE.....	37
4.3.1	Bancarisation des métadonnées associées à l'analyse.....	37

4.3.2	Bancarisation des métadonnées associées au retraitement.....	37
4.3.3	Besoin que « n'importe qui » puisse retraiter les données acquises et bancarisées 37	
4.4	POURSUITE DES ACTIONS EUROPEENNE ET FRANÇAISE EN COURS ...	38
5	Conclusions	39
6	Bilan sur le criblage de molécules en fonction des objectifs de la surveillance	40

Liste des illustrations

Illustration 1 – Enjeux présents et futurs pour la surveillance des milieux aquatiques	13
Illustration 2 – Les différentes applications possibles de l’analyse par spectrométrie de masse haute résolution	15
Illustration 3 – Substances classifiées en catégories selon les niveaux de confiance avec lesquelles leur identité peut être confirmée (selon (Schymanski, Singer et al. 2014))	17
Illustration 4 – Chromatogramme regroupant les signaux détectés dans un échantillon (ciblé, suspect et non-ciblé).....	19
Illustration 5 – Difficultés pour l’identification de substances en lien avec les niveaux de confiance (Letzel Thomas, Lucke Thomas et al. 2014)	20
Illustration 6 – Développement de méthodes in vitro et de biomarqueurs pour la surveillance des substances chimiques et le diagnostic de pollutions multiples (Sanchez and Aït-Aïssa 10-12 mars 2010).	24
Illustration 7 – Fréquence de détection des composés mis en évidence par l’approche non ciblée. La fréquence d’identification est illustrée par un code couleur : vert pour faible, jaune pour moyenne et rouge pour haute fréquence de détection.	25
Illustration 8 – Distribution des 7767 marqueurs retenus, classés par pays. L’intensité du signal est figurée par le code couleur , de bleu (faible), à rouge (fort), chaque trait horizontal correspondant à un marqueur (IKSR 2015).	26
Illustration 9 – Similarités des profils de pollution selon les pays le long du Danube obtenues par Analyses des Composantes Principales (IKSR 2015).....	27
Illustration 10 – Fréquence de détection des 110 composés suspects détectés dans les échantillons, sur les 315 recherchés (IKSR 2015).....	28
Illustration 11 – De l’échantillonnage au résultat final : Etapes et contrôles qualité	31
Illustration 12 – Procédures et méthodes d’analyse harmonisées pour la SMHR	33
Illustration 13 – Influence du dimensionnement de la fenêtre du temps de rétention sur le nombre de signaux « aténolol » identifiés.....	35

Liste des annexes

Annexe 1 Principe de la spectrométrie de masse haute résolution	49
Annexe 2 Bases de données	57
Annexe 3 Prédiction des temps de rétention.....	61

GLOSSAIRE

BDD : Base de données

CEN : Comité Européen de Normalisation

CG : Chromatographie en phase gazeuse

CIPR : Commission internationale pour la protection du Rhin

CL : Chromatographie en phase liquide

Contrôles d'enquête : contrôles qui sont destinés à identifier l'origine d'une dégradation de l'état des eaux.

Contrôles opérationnels : contrôles qui sont destinés à évaluer l'état et l'évolution des masses d'eau présentant un risque de ne pas atteindre les objectifs environnementaux

Contrôles de surveillance : contrôles qui sont destinés à évaluer les incidences de l'activité humaine et les évolutions à long terme de l'état des masses d'eau.

Criblage ciblé : analyse d'une liste finie et connue de substances a priori généralement à des fins quantitatives.

Criblage suspect : analyse orientée dont l'objectif est d'amener une preuve de la présence de substances par rapport à des informations préexistantes.

Criblage non ciblé : analyse dont l'objectif est d'identifier toutes molécules présentes dans l'échantillon sans aucun a priori.

DCE : Directive loi-cadre sur l'eau

EDA : Effect Directed Analysis

EPA (Environmental Protection Agency) : Agence de la protection de l'environnement

Ion fragment : Ion formé après fragmentation de l'ion moléculaire par perte d'un groupement. Les fragments donnent une information clé sur la structure de la molécule.

Ion moléculaire (ou ion parent) : cation radicalaire ayant la même masse que la molécule neutre, mais avec un électron en plus ou en moins dû au bombardement d'électrons hautement énergétiques dans la source d'ionisation du spectromètre de masse.

ISO (International Organization for Standardization) : Organisation internationale de normalisation

Massif isotopique : Massif regroupant des pics de masses proches à M+1 ou M+2 correspondant à la combinaison des différents isotopes qui composent la molécule.

NF : Norme Française

NQE : Normes de Qualité Environnementale fixées par la DCE

SPAS : Substances pertinentes à surveiller

PSEE : Polluants spécifiques de l'état écologique

SMHR : Spectrométrie de masse haute résolution.

Spectrométrie de masse (SM) : technique d'analyse physico-chimique permettant de détecter, d'identifier et de quantifier des molécules d'intérêt par mesure exacte de leur masse et permettant de caractériser la structure chimique des molécules en les fragmentant.

TOF (Time Of Flight) : Temps de vol

TR : Temps de rétention

1 Introduction

Depuis une vingtaine d'années, les besoins liés à la détermination de molécules organiques à l'état de traces et/ou ultra traces dans des matrices complexes sont croissants. Les grandes avancées techniques et instrumentales qui ont été acquises au cours de la dernière décennie ont ouvert le champ des "possibles" et sont désormais à la base d'un grand nombre de réglementations nationales, européennes ou internationales dans des domaines très variés.

Dans le domaine de la surveillance de l'environnement, les enjeux relatifs à la fiabilité des déterminations sont tels qu'ils nécessitent désormais de disposer d'instruments de mesure présentant des caractéristiques de performances aussi bien en terme de sensibilité (capacité à détecter de faibles quantités) et de spécificité (capacité à dissocier une molécule cible d'une autre présentant des propriétés moléculaires et structurales proches). En outre, face aux enjeux de la surveillance environnementale, le besoin de connaissance quant à la présence de potentiels composés toxiques, de polluants en mélange et notamment de polluants inconnus dits émergents tels que les produits de transformation est devenu crucial.

Le criblage ciblé, permettant de confirmer ce qui est connu dans un échantillon (Illustration 1), nécessite un pré travail de priorisation des molécules d'intérêt pour établir la liste des composés à rechercher. Au-delà de cette approche, de nouvelles sont désormais envisageables, du fait des évolutions technologiques des appareils de mesures mais surtout des améliorations des outils permettant le traitement de la donnée. Ces nouvelles approches de criblage, permettant d'identifier ce qui n'est pas connu dans un échantillon (Illustration 1), ne sont pas encore opérationnelles, mais les avancées rapides et les actions européennes et françaises concertées en cours laissent présager une accélération de la levée des verrous. Les exercices d'identification de nouveaux composés d'intérêt ont montré leur richesse en termes d'informations (détermination des nouvelles listes des substances pertinentes à surveiller - SPAS et des polluants spécifiques de l'état écologique - PSEE) mais aussi leurs limites (lourdeur de gestion, manque de réactivité...).

Pour la surveillance des milieux aquatiques, les enjeux présents et futurs sont doubles :

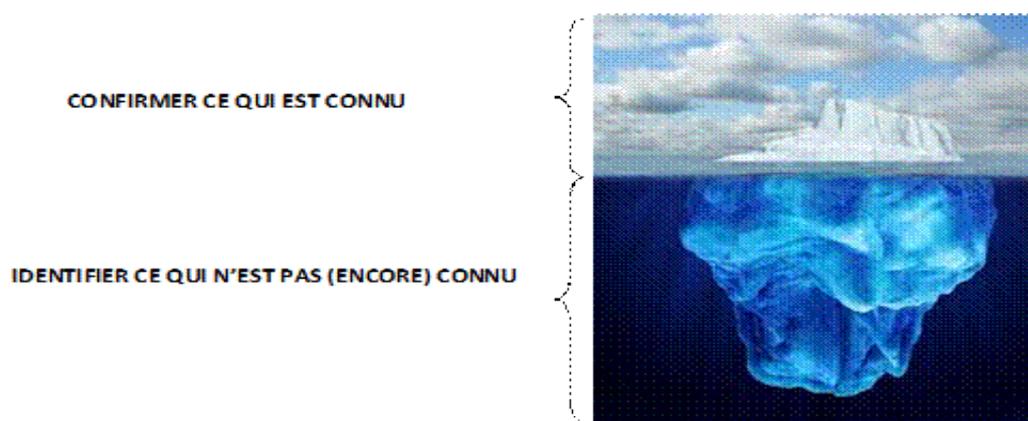


Illustration 1 – Enjeux présents et futurs pour la surveillance des milieux aquatiques

Ce document présente les approches de la spectrométrie de masse haute résolution (SMHR) pour l'identification de molécules organiques dans les échantillons environnementaux en montrant les avantages et les verrous de cette technique. Il a pour but de montrer l'intérêt et la faisabilité d'intégrer ces nouveaux outils dans la surveillance.

2 Les approches SMHR menant à l'identification

Le principe des analyseurs haute résolution repose sur l'acquisition de l'intégralité de l'information présente dans l'extrait ou dans l'échantillon : tout composé verra sa présence enregistrée de manière non sélective par le détecteur. Il est cependant important de préciser que ce type d'outils n'est pas universel et que la méthode d'extraction joue un rôle important sur le type de composés présents dans l'extrait.

Tout composé présent dans l'échantillon ne sera pas "analysable" par un outil unique. L'exhaustivité ne pourra être atteinte qu'en utilisant différents types d'instruments (chromatographie en phase liquide pour les molécules polaires, chromatographie en phase gazeuse pour les molécules apolaires...). Certaines molécules très complexes à mesurer ne pourront être observées car nécessitant des méthodes d'analyse très spécifiques. Le principe de l'approche repose sur l'utilisation d'un nombre réduit de méthodes génériques, couvrant un large spectre de molécules et de ce fait non spécifiques. Du fait de l'enregistrement non-sélectif des données, il est possible de rechercher *a posteriori* des molécules qui n'étaient pas ciblées au départ.

Cette information acquise sera ensuite retraitée de différentes façons selon la finalité de l'étude.

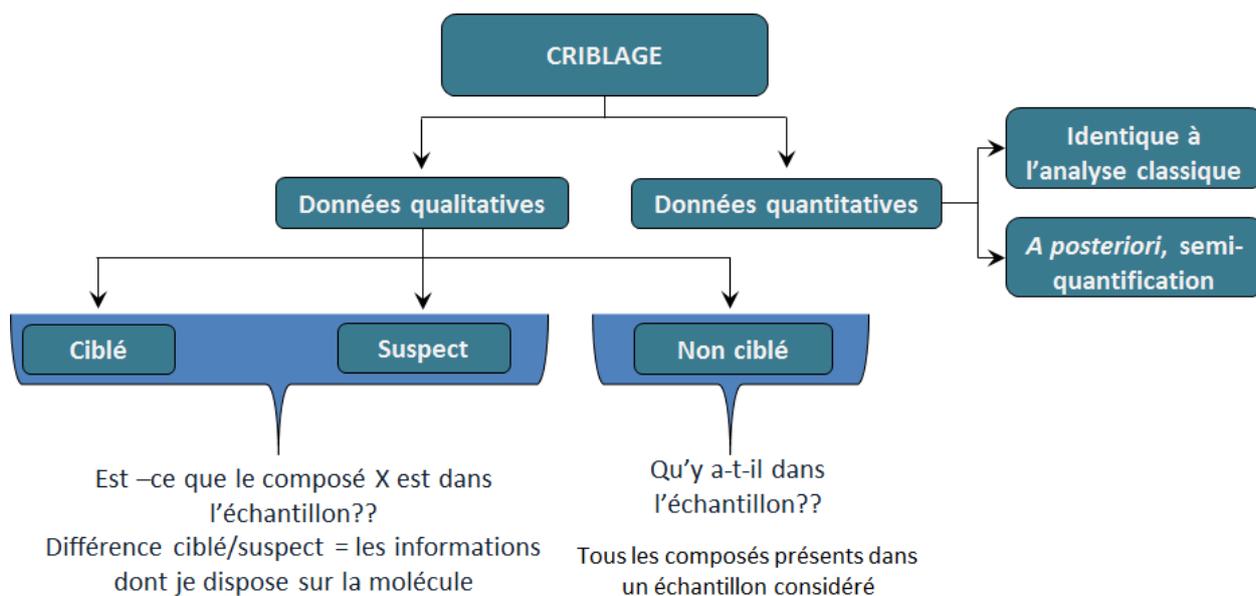


Illustration 2 – Les différentes applications possibles de l'analyse par spectrométrie de masse haute résolution

2.1 APPROCHE QUANTITATIVE

La mise en œuvre d'instruments de mesure reposant sur le couplage de la chromatographie en phase gazeuse ou phase liquide et de la SMHR est largement documentée dans un certain nombre de normes EPA, ISO, CEN, NF pour le dosage de molécules organiques dans des matrices complexes. Citons à titre d'exemple la norme EPA 1698: Steroids and Hormones in Water, Soil, Sediment, and Biosolids by CGHR/SMHR, la norme ISO 13914:2013 Qualité du sol - Détermination des dioxines et furanes comme biphényles polychlorés par chromatographie en phase gazeuse avec spectrométrie de masse à haute résolution (CG/SMHR) ou bien encore la norme ISO 18073:2004 Qualité de l'eau - Dosage des dioxines et furanes tétra- à octachlorés -

Méthode par dilution d'isotopes CGHR/SMHR. Ainsi jusqu'en 2014, la CGHR/SMHR était reconnue comme l'unique méthode de référence permettant de doser les dioxines, à des fins réglementaires, à l'état ultra-traces dans les matrices alimentaires.

Ces couplages mettant en œuvre la SMHR dans un but quantitatif peut présenter deux avantages principaux:

➤ Leur spécificité peut permettre de diminuer le nombre de faux positifs liés à des interférents ou bien encore certains composés isobariques. Le couplage chromatographique reste nécessaire pour la différenciation des composés isomériques.

➤ Leur spécificité peut permettre d'atteindre dans certains cas des limites de quantification inférieures à celles des couplages reposant sur la spectrométrie de masse basse résolution. Ainsi, dans le rapport publié en 2012 «Analytical Methods for the new proposed Priority Substances of the European Water Framework Directive (WFD)», Loos met en évidence que le recours à ces couplages, pour atteindre les limites de quantification compatibles avec les exigences de la DCE, pourrait être souhaitable voire même indispensable (concept de meilleures techniques disponibles) pour les substances les plus critiques ayant une Norme de Qualité Environnementale (NQE) faible ou difficile à atteindre: cyperméthrine, dichlorvos, dioxines, heptachlorepoxyde, HCB, PBDE, hormones notamment....

Cependant dans ce contexte, ses principaux désavantages sont le coût de sa mise en œuvre, sa plus faible présence dans les laboratoires. Sa prise en main s'avère fastidieuse (sensibilité, masse exacte, calibration, etc.) par rapport aux spectromètres de masse basse résolution (triple quadripôle), mais l'appropriation en cours par les laboratoires prestataires avec une forte augmentation du nombre d'équipements tend à fortement diminuer cet inconvénient.

Lorsqu'ils sont utilisés à des fins quantitatives, leur mise en œuvre est encadrée par les référentiels déjà existants et ne sera donc pas abordé au travers du présent travail.

2.2 APPROCHE QUALITATIVE

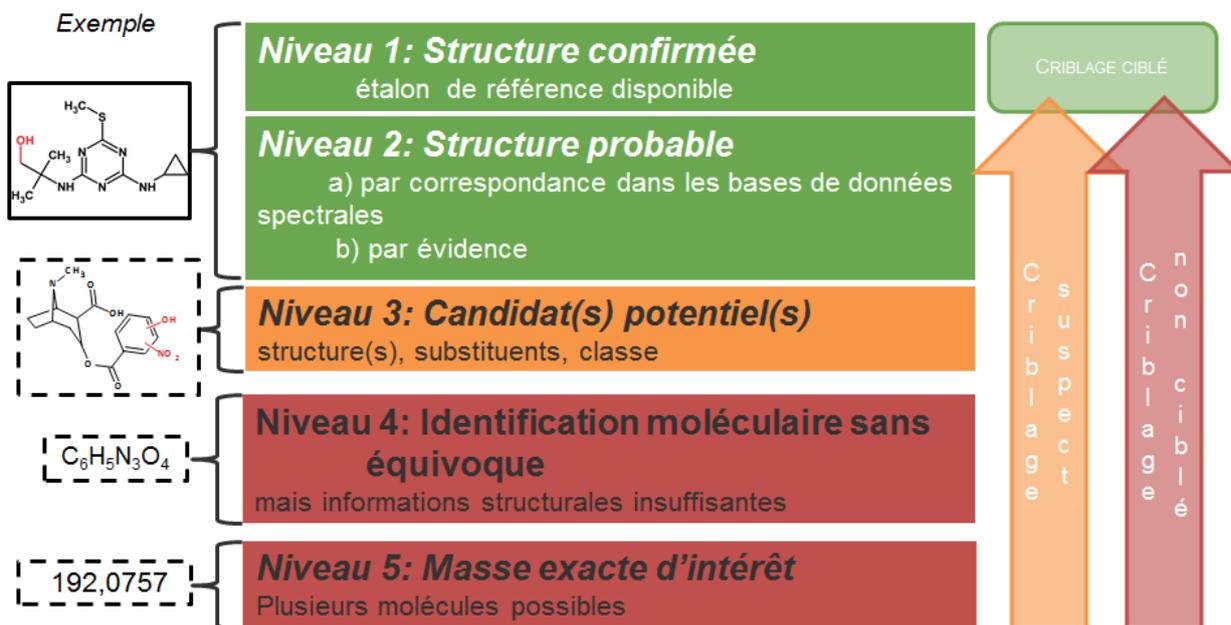
Toutes les molécules présentes dans l'extrait ou l'échantillon et analysables par la méthode analytique utilisée sont détectées et enregistrées. Le traitement qualitatif de ces données est classé selon deux approches (Illustration 2) :

➤ Le criblage ciblé et suspect pour l'identification d'une molécule connue dans l'échantillon

➤ Le criblage non ciblé pour l'identification de molécules présentes dans l'échantillon et non préalablement ciblées/identifiées/pressenties/recherchées

Le principe d'identification d'une molécule en SMHR est détaillé en annexe 1. A partir des paramètres d'identification d'une molécule (masse exacte, temps de rétention, données de fragmentation) « disponibles » dans un échantillon et des bases de données spectrales et de la littérature, différents niveaux de confiance sur l'identification sont attribués. Dans le cadre des travaux NORMAN, cinq niveaux de confiance accordés à l'identification d'une molécule ont été mis en place (Schymanski et al. 2014b) (Illustration 3). Les molécules identifiées par criblage ciblé correspondent au niveau de confiance le plus élevé (niveau 1); celles par criblage suspect et non ciblé peuvent correspondre aux niveaux de confiance 2, 3, 4 et 5.

Niveaux de confiance pour l'identification d'une molécule



Schymanski, Jeon, Gulde, Fenner, Ruff, Singer & Hollender (2014) *ES&T*, DOI: 10.1021/es5002105

Illustration 3 – Substances classifiées en catégories selon les niveaux de confiance avec lesquelles leur identité peut être confirmée (selon (Schymanski et al. 2014b))

2.2.1 Criblage ciblé/ suspect

Dans le cadre de cette approche, la question posée est : “est-ce que la molécule X est présente dans l'échantillon?”. La recherche de la présence d'une molécule ou non est basée sur les informations disponibles au laboratoire. Chaque molécule possède sa propre “empreinte chimique” dont certains paramètres sont inhérents à la méthode analytique comme le temps de rétention (TR) et la présence de fragments et leur intensité (Annexe 1). Pour cela des bases de données (BDD) sont créées à partir de standard de référence et peuvent contenir le TR, la masse exacte de la molécule, les masses des fragments, les adduits potentiels, les rapports d'intensité des fragments, etc. Les critères employés pour des analyses multi-résidus comme édictés dans la future NF T 90-214 peuvent être utilisés pour l'identification des molécules ciblées (niveau 1).

2.2.1.1 Criblage ciblé

Dans le cas du criblage ciblé il s'agit de BDD internes, élaborées à partir de l'injection de standards de référence. Les informations présentes sont spécifiques à l'instrument analytique. La présence ou l'absence des molécules incluses dans cette BDD interne sont recherchées dans l'échantillon. Si certaines s'avèrent être présentes, elles seront identifiées avec un accord de confiance de niveau 1 (Illustration 3). La confiance accordée à l'identification de la molécule est alors maximale sans toutefois discriminer des énantiomères, à moins de mettre en œuvre une méthode chromatographique adaptée.

2.2.1.2 Criblage suspect

Le criblage suspect peut être assimilé à du criblage ciblé, mais le laboratoire ne possède pas les étalons analytiques pour les substances recherchées. Les molécules d'intérêt sont connues et leurs propriétés physico-chimiques (volatilité, polarité,...) peuvent être évaluées à travers une recherche bibliographique. Elles peuvent provenir d'un exercice de priorisation, de la découverte (identification) de polluants émergents par des instituts partenaires (NormaNews¹ par exemple), ou d'autres sources (bibliographiques, etc.). Les BDD utilisées sont soit bâties avec des données externes (bibliographie) soit empiriques ou théoriques. Ces BDD peuvent être de différentes formes et ne contiennent aucune information sur les TR car ils sont spécifiques de la méthode analytique employée et celle-ci est différente dans chaque laboratoire. Il existe cependant des outils de modélisation et/ou prédiction des TR (Annexe 3) qui les estiment aux vues des propriétés physico-chimiques des molécules et de leur TR issu de la littérature. La confiance accordée aux molécules identifiées par le criblage suspect peut être de niveaux 2, 3, 4 et parfois 5 ; cela dépend des paramètres « disponibles » pour chaque molécule et du degré d'information contenue dans les BDD utilisées (TR, masses exactes des fragments, etc.).

Comme détaillé en Annexe 2, les BDD peuvent être constituées de différentes informations. Dans le cas le plus simple, elles contiennent le nom de la molécule et la formule chimique ce qui amène à la disposition de la masse exacte. Ces informations ne sont cependant pas suffisantes pour pouvoir postuler avec certitude à une structure moléculaire car malgré la connaissance de la masse exacte, le résultat obtenu peut correspondre à une dizaine voire beaucoup plus de composés. La confiance d'identification accordée à ces molécules est de niveau 4 car seules la masse exacte et la formule brute sont disponibles (Illustration 3).

Ainsi, il est également nécessaire de disposer de bases de données spectrales comprenant des informations additionnelles comme des données de fragmentation (parfois en considérant les adduits). Ces données supplémentaires permettent d'identifier des molécules avec une confiance plus élevée (niveau 3 voire 2 ; Illustration 3). La connaissance de la fragmentation de certaines molécules permet parfois à elle seule de déterminer la structure chimique de la molécule.

Certaines BDD contiennent les temps de rétention des molécules les constituants, spécifiques de la méthode analytique employée. Mais dans le cas du criblage suspect, cette information même provenant d'un autre laboratoire peut aider à l'identification d'une structure probable de molécule (niveau de confiance 2, Illustration 3) en utilisant la modélisation (Annexe 3). Le temps de rétention chromatographique reflète les interactions entre les analytes, la nature de la phase stationnaire et la phase mobile employée. Ces temps de rétention sont normalement reproductibles et peuvent donc être utilisés afin de caractériser la nature des analytes élués (Annexe 1).

Pour confirmer la réelle présence d'une molécule identifiée par le criblage suspect, atteignant donc les critères de confiance de niveau 2, le standard de référence sera acheté puis injecté. Par la suite cette molécule sera inscrite dans la liste des molécules ciblées (basculément en niveau 1) et pourra de ce fait être identifiée avec une confiance maximale dans les échantillons.

¹ NormaNews est une activité collaborative destinée aux membres actifs du groupe « non-target screening » au sein du réseau NORMAN. Il s'agit d'une plateforme d'échanges sur laquelle sont partagées des informations concernant de « nouvelles » molécules identifiées afin que chaque membre du groupe fasse une analyse rétrospective dans ces échantillons. Cela permettra d'obtenir des données d'occurrences de ces « nouvelles » molécules en Europe et au-delà.

2.2.2 Criblage non-ciblé

Le criblage non ciblé concerne tous les signaux non attribués par le criblage ciblé ou suspect. Plusieurs milliers de signaux peuvent être concernés (Illustration 4). Chaque signal est caractérisé au minimum par un temps de rétention, une masse exacte (et son massif isotopique) et une intensité. L'ensemble de ces trois paramètres est appelé « marqueur » et se positionne en niveau 5 sur l'échelle d'identification.

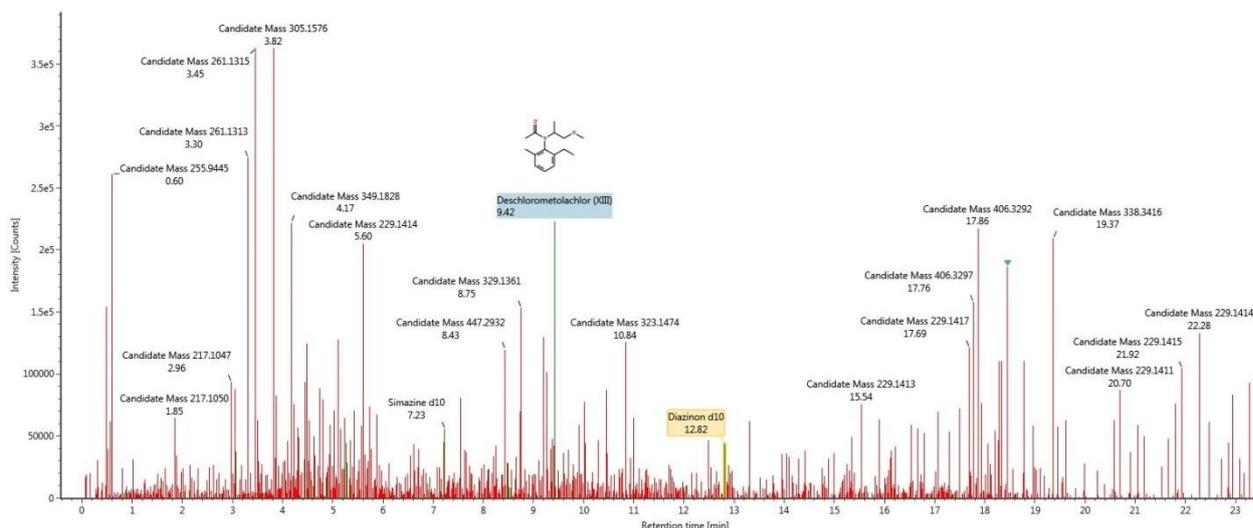
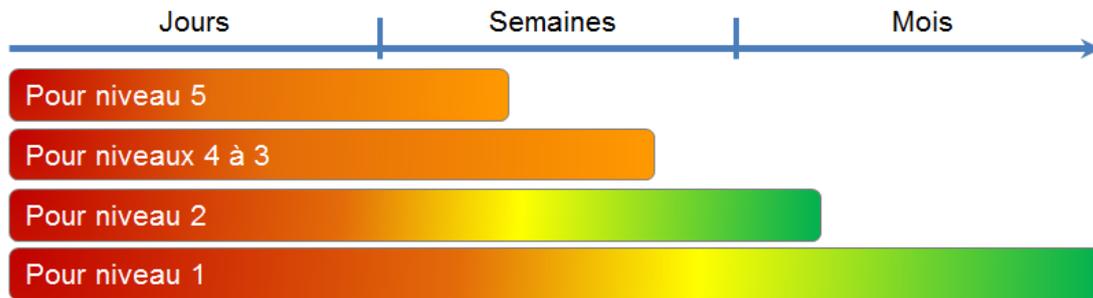


Illustration 4 – Chromatogramme regroupant les signaux détectés dans un échantillon (ciblé, suspect et non-ciblé).

Le travail nécessaire pour améliorer le niveau d'identification des composés, du niveau 5 (marqueurs) vers le niveau 1 (criblage ciblé, validation par standard) est à la fois long, incertain et doit être mené par des « experts » (Illustration 5). Compte-tenu des différentes possibilités d'identification, des analogies de structure et des verrous insolubles lors du processus d'identification, le nombre de marqueurs pouvant être finalement identifiés jusqu'au niveau 1 est faible.

Durée et effort analytique



Nombre de substances à traiter par analyse

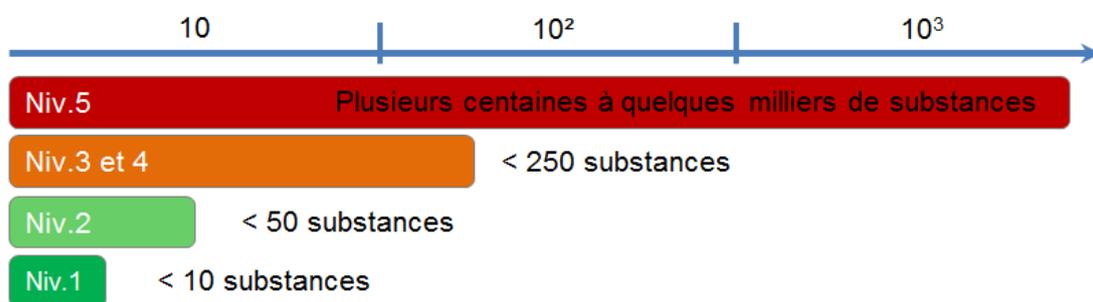


Illustration 5 – Difficultés pour l'identification de substances en lien avec les niveaux de confiance (Letzel Thomas et al. 2014)

Pour faciliter ce travail de traitement des données, il est intéressant d'orienter les efforts d'identification sur les marqueurs présentant un plus fort intérêt pour l'étude concernée. Pour cela différentes stratégies sont utilisées comme les approches guidées par les effets (approche Effect Directed Analysis, EDA) ou l'utilisation des comparaisons d'empreintes moléculaires. Ces approches sont détaillées dans la section suivante.

3 Intérêt de l'intégration de la HRMS dans la surveillance

La surveillance des milieux aquatiques telle qu'entreprise dans le cadre de la DCE a pour objectif d'évaluer la qualité des masses d'eau quant à leur état chimique et écologique. Dans ce contexte, les besoins sont extrêmement diversifiés. Les paragraphes ci-dessous essaieront de démontrer en quoi et comment les approches SMHR permettent de répondre positivement à ces différents enjeux.

L'évolution des modalités de la surveillance (pour une mise en œuvre lors du prochain cycle) étant d'actualité, sont présentés ici les deux volets : d'une part les possibles réponses à la surveillance actuelle et d'autre part l'amélioration de la surveillance par l'intégration de nouveaux outils.

3.1 ACQUISITION DE DONNEES QUANTITATIVES SUR L'ETAT CHIMIQUE

L'utilisation de la SMHR pour l'obtention de données quantitatives sur le bon état chimique est envisageable. L'approche ciblée permet d'atteindre au moins les niveaux de performances exigées par les objectifs réglementaires et obtenues par les approches actuelles. Rappelons toutefois que cette technique ne permet pas de couvrir l'ensemble des molécules. Elle est tout d'abord limitée par l'extraction si elle a lieu et le type de chromatographie en amont (composés plutôt polaires si chromatographie liquide et apolaires si chromatographie gazeuse). De plus malgré une méthode chromatographique la plus générale possible, certains composés spécifiques ne pourront pas être détectés (comme par exemple le glyphosate...).

En complément aux approches actuelles, les données exploitées a posteriori ne permettront pas d'obtenir une valeur quantitative robuste (cf 4.1) mais donneront accès à une approximation de la concentration (donnée semi-quantitative).

3.2 ACQUISITION DE DONNEES QUALITATIVES

La plus-value de ces nouvelles technologies est essentiellement l'acquisition de données qualitative (criblage : présence/absence).

3.2.1 Mise à jour des listes de substances d'intérêt : campagne prospective, listes de vigilances (au niveau national et européen)

L'objectif des campagnes prospectives est d'améliorer le diagnostic (acquisition de données d'occurrence sur les substances d'intérêt) pour mieux cibler les substances prioritaires pour les futures mesures de gestion.

Dans les dernières campagnes prospectives conduites au niveau national (recherche ciblée), un travail conséquent de priorisation des molécules à rechercher a été nécessaire. Ce travail a été effectué sur une liste de substances candidates sur la base des connaissances disponibles à la date de l'exercice. Les capacités des laboratoires à analyser les molécules identifiées ont servi pour la sélection finale des molécules à rechercher. Au final ce travail fastidieux mais riche en information a servi de base à l'établissement de la liste de substances à surveiller (SPAS). L'expérience de ces récentes campagnes prospectives montre que : 1) du fait de l'analyse ciblée, la liste des molécules à rechercher doit être définie très en amont de la campagne

d'analyse ; 2) certaines molécules d'intérêt n'ont pas fait l'objet de cette surveillance (soit parce que identifiées après l'établissement des listes, soit parce que non analysables par les laboratoires).

De ce fait pour compléter l'établissement des listes de surveillances, la SMHR permettrait :

- Une recherche par l'approche « criblage ciblé / suspect » de toute liste de molécules établie avant ou après l'acquisition de l'analyse, l'étape de recherche dans l'extrait ou l'échantillon n'étant qu'un processus de traitement de l'information. Différentes listes de molécules à rechercher peuvent ainsi être établies, utilisant le même jeu de données mais avec des objectifs cibles différents c'est-à-dire des listes de molécules différentes.
- Une recherche dans des échantillons préalablement bancarisés. Une molécule d'intérêt « récent » (c'est-à-dire mise en lumière par de nouvelles analyses ou de nouvelles connaissances) peut ainsi être recherchée dans des échantillons/extraits analysés préalablement et traités pour répondre à une liste initiale ne comprenant pas cette molécule.
- Une recherche par l'approche non ciblée, pour identification structurale des molécules présentes dans l'échantillon. Selon la réussite du processus d'identification, il est possible de « remonter » les différents niveaux d'identification (5 à 1), chacun d'entre eux permettant néanmoins d'obtenir des fréquences de détection (cf 3.4.1).

Un premier traitement de la donnée par l'approche suspecte pourra être mis en œuvre puis les résultats pourront être confirmés par l'injection des standards pour les composés d'intérêt (approche ciblée).

Ce traitement d'échantillons bancarisés pourra permettre une première estimation de la fréquence de détection de nouvelles molécules d'intérêt (pouvant intégrer le processus des campagnes prospectives). Les méthodes SMHR pouvant cependant s'avérer moins sensibles que les méthodes actuelles, car par définition moins spécifiques, des molécules nécessitant des limites de détection très faibles ne pourront peut-être pas être observées (ex des hormones). De même, l'effet de la préparation de l'extrait doit être considéré car cette étape peut avoir éliminé la molécule pourtant présente dans l'échantillon initial.

Liste de vigilance européenne:

L'objectif de cette liste de vigilance (EU Watch List) est d'acquérir à l'échelle européenne, des données d'occurrence sur des molécules d'intérêt et nourrir ainsi l'exercice de priorisation pour la mise à jour des Substances Prioritaires de la DCE. L'approche semi-quantitative apportée par le criblage ciblé pourrait répondre à ces objectifs, sous réserve que les méthodes d'analyses génériques et potentiellement d'extraction soient compatibles avec les molécules d'intérêt. Ce point est facilement vérifiable, il suffira a posteriori d'évaluer le protocole utilisé pour les échantillons « bancarisés » avec les nouveaux composés étalons d'intérêt.

3.2.2 Etat écologique/pression chimique

Un certain nombre de substances sont considérés/presentis comme les Polluants Spécifiques de l'Etat Ecologique (PSEE). De nouvelles approches sont en cours d'investigation pour mieux documenter l'état écologique des milieux. Ces outils font appel aux travaux sur la toxicité

notamment par le développement de bioessais, permettant d'améliorer le diagnostic risque environnemental, et la surveillance de l'état écologique

L'utilisation de la SMHR en parallèle des études écotoxicologiques permettrait d'élucider au moins partiellement quelles sont les molécules responsables de l'effet (réponse biologique) observé.

L'approche EDA (Effect Directed Analysis) permet de relier plus spécifiquement composés chimiques et effets (Illustration 6). L'extrait global est fractionné et pour chaque fraction la toxicité et la composition chimique sont comparées. A ce jour, la composition chimique est principalement abordée par l'approche ciblée classique (on recherche par une analyse spécifique la présence de composés préciblés). Mais les études récentes montrent qu'avec cette approche il n'est possible d'élucider qu'une fraction souvent faible et très variable selon les échantillons de la toxicité. Une étude de 2010 sur des sédiments montre une élucidation de 10 à 96% selon le type de tests mis en œuvre et le type de sédiment (Kinani et al. 2010).

L'utilisation des approches « criblage suspect » voire « criblage non ciblé » permettrait d'augmenter le nombre de molécules potentiellement responsables des effets observés.

L'approche suspecte peut être utilisée, en fonction des résultats des tests de toxicité pour rechercher dans l'échantillon des molécules connues pour être responsables des effets mis en évidence (perturbations endocriniennes, activité dioxin-like...).

L'approche non ciblée permettrait d'identifier de manière plus large, sans *a priori*, la/les molécules présentes dans la fraction toxique.

Cette approche est fastidieuse et complexe à mettre en œuvre, et nécessite un traitement particulier de l'échantillon (cf 6) mais pourrait permettre de grandes avancées pour mettre en lumière des polluants à enjeux forts nécessitant des mesures de gestion spécifiques si le critère « toxicité » est considéré prépondérant.

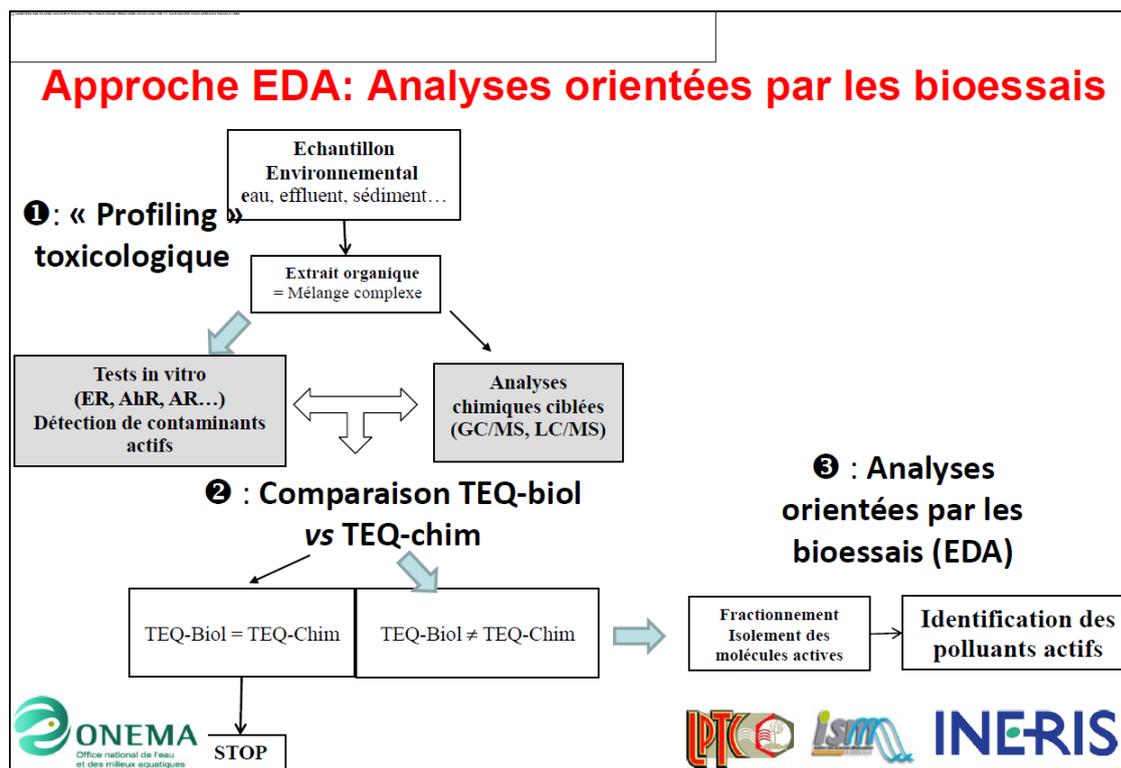


Illustration 6 – Développement de méthodes in vitro et de biomarqueurs pour la surveillance des substances chimiques et le diagnostic de pollutions multiples (Sanchez & Aït-Aïssa 10-12 mars 2010).

3.2.3 Contrôle d'enquête : recherche/ identification de sources de pollutions, alerte, ...

Recherche de sources, empreinte moléculaire d'une pollution

Il est possible d'utiliser le criblage non-ciblé pour caractériser les sources de pollution (stations d'épuration, diffuses, urbaines, etc.). Chaque site d'échantillonnage possède une empreinte moléculaire spécifique, qui peut toutefois évoluer en fonction des saisons, des sources de pollution, des intempéries, etc. Comparées à celles issues de sources de pollutions, il est possible de différencier ces sites. La comparaison d'empreintes moléculaires devrait permettre à terme de distinguer les différentes sources de pollutions sur les sites impactés en tenant compte des potentiels phénomènes de dégradation et d'adsorption. Cependant certains composés restent stables dans l'environnement.

Dans le même ordre d'idée, le suivi du profil moléculaire d'un site à enjeu peut permettre de rapidement alerter en cas de modification. La notion d'empreinte moléculaire ne nécessite pas dans un premier temps l'identification des marqueurs mais fait intervenir de lourds traitements statistiques pour lesquels la majorité des laboratoires ne sont pas encore compétents.

Efficacité des programmes de mesure

Le suivi des programmes de mesure afin de juger de leur efficacité est toujours complexe à envisager. En effet la « simple » diminution de la présence d'un composé n'est pas forcément significative, car elle peut être associée à l'apparition d'autres composés (produits de dégradation, etc.) qu'il peut s'avérer difficile d'anticiper. Néanmoins, l'acquisition d'une donnée « plus exhaustive » et indépendante des paramètres préciblés permet une plus grande souplesse dans les recherches de composés d'intérêt.

3.3 INTERET DE LA BANCARISATION/ DE LA POSSIBILITE DE TRAITEMENT A POSTERIORI

La grande force de ces nouvelles approches est la possibilité de pouvoir mettre en place une démarche évolutive et/ou itérative.

La possibilité de pouvoir rechercher dans un jeu de données (spectres de masse) bancarisées la présence/absence d'un composé nouvellement identifié permettra une meilleure réactivité à moindre coût (pour une première approche qualitative, pas de nécessité d'une nouvelle campagne d'échantillonnage/ d'analyse). Le jeu global d'échantillons pourra être partiellement exploité en fonction des spécificités locales.

Le fait de pouvoir répondre à différentes questions à partir d'un jeu de données bancarisées permet de prédire (temporellement et spatialement) le nombre de stations suivies qui pourront être exploitées différemment selon la problématique posée.

3.4 LES EXPERIENCES EUROPEENNES

3.4.1 Cas de la Norvège (Schlabach et al. 2013)

Une initiative norvégienne avait pour but de démontrer l'intérêt de l'approche non-ciblée dans la mise en évidence de nouveaux composés d'intérêt.

Un set d'échantillons variés (effluents de stations d'épuration (entrée/ sortie) ; sédiments, boues ; biote, air) a été analysé à la fois pour les composés polaires (CL/SMHR) et pour les apolaires (CG/SMHR et CGxCG/HRMS). Pour chacun des échantillons, l'approche non ciblée a été mise en œuvre.

En raison du temps et des ressources limitées de l'étude, le niveau maximal d'élucidation de l'identité des substances a été celui correspondant au niveau 2 (structure probable) de l'illustration 3, car l'étape de confirmation par injection de standards n'a pas été menée. Le niveau 5 n'a pas été considéré, car ils n'ont conservé que les composés pour lesquels une formule chimique est confirmée. Pour les 3 catégories considérées (niveaux 2, 3 et 4), la fréquence de détection dans le set d'échantillons a été calculée. Cet indicateur a ainsi permis de prioriser les molécules (structures « probables ») pour lesquelles la confirmation de l'identité (passage en niveau 1) et le développement de méthodes quantitatives semblent prioritaires.

	air urbain		Stations d'épuration					sédiment	biote
	périphérie	urbain	boues traitées	boue brute	eau entrée	MES entrée	eau sortie		
categories 3 +4	252	212	265	82	271	291	209	237	63-305
catégorie 2	56	62	141	205	105	48	330	178	25-103
dont									
pharma	26	18	39	25	33	15	242	25	4-54
Composés aromatiques polycycliques	1	9	50	94	9	3	4	96	45
additifs	12	15	27	42	34	15	28	15	3-9
phtalates	9	8	8	8	7	8	8	17	0-1
Organophosphorés	4	7	4	7	7	4	9	5	2-7
Pesticides	4	5	5	4	2	2	33	1	1-12
PFCs	0	0	0	0	3	0	0	0	0-1
autres	0	0	8	25	10	1	6	19	0

Illustration 7 – Fréquence de détection des composés mis en évidence par l'approche non ciblée. La fréquence d'identification est illustrée par un code couleur : vert pour faible, jaune pour moyenne et rouge pour haute fréquence de détection.

L'étude norvégienne conclut sur l'apport important de ces technologies, la plus grande facilité d'identification pour les méthodes CG/SM, compte-tenu de l'existence de bases de données universelles, la nécessité pour les approches CL/SM de confirmer par l'injection des standards analytiques, la plus-value d'inclure ces techniques analytiques dans les campagnes de screening et la surveillance réglementaire.

3.4.2 Joint Danube Survey (IKSR 2015)

Dans le cadre du réseau de surveillance sur le Danube, une campagne de démonstration a été menée sur 68 échantillons sur le Danube et ses principaux affluents. Les techniques de CL/SMHR et CG/SM ont été appliquées sur l'intégralité des échantillons. L'approche ciblée a été conduite pour 168 composés, dont 154 ont été détectés dans au moins un échantillon.

Concernant l'approche non ciblée :

Pour les techniques de CL/SMHR, sur 16214 marqueurs identifiés, 7767 ont été retenus (critères d'intensité, fréquence de détection, etc.) pour l'approche d'élucidation structurale. Parmi ces 7767 marqueurs, 955 n'ont pas pu franchir le niveau 5, 3442 ont été retrouvés dans les bibliothèques utilisées (équivalent niveau 2), dont 58 avec un degré de confirmation supérieure (correspondance des fragments théoriques et mesurés) et 3370 ont pu être classés en niveau 3 ou 4.

Une première visualisation des résultats permet de mettre en évidence les disparités entre les pays (Illustration 8), tant en termes de nombres que d'intensité des marqueurs retenus. Le même laboratoire a analysé l'intégralité des échantillons de tous les pays, permettant de s'affranchir de potentiels différences méthodologiques.

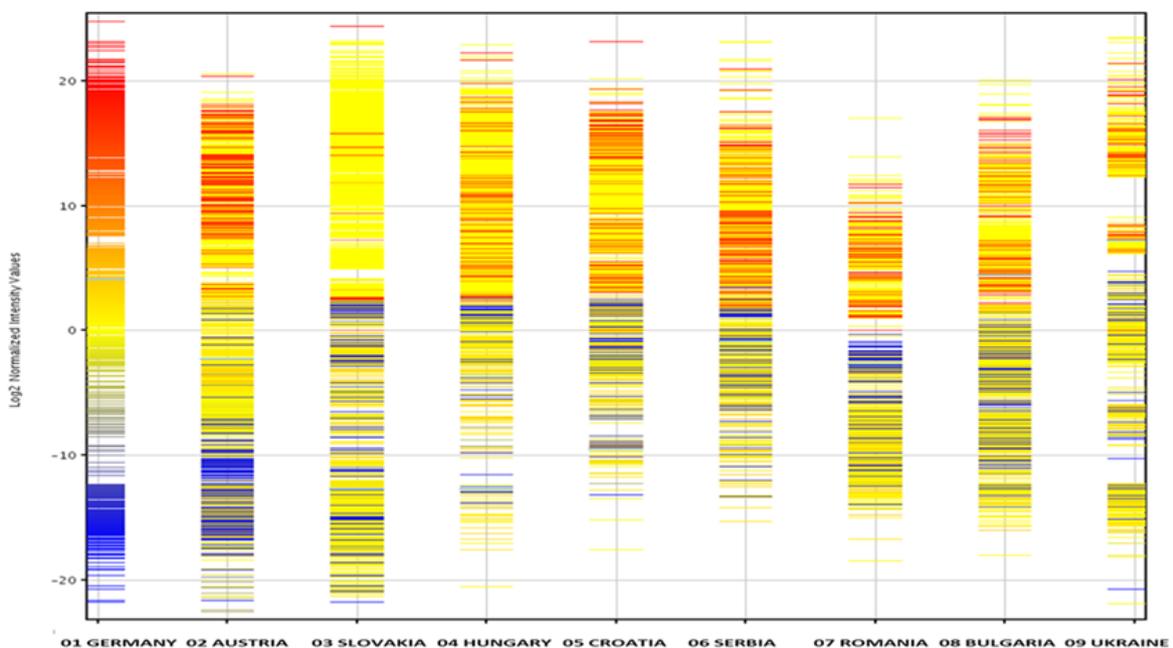


Illustration 8 – Distribution des 7767 marqueurs retenus, classés par pays. L'intensité du signal est figurée par le code couleur, de bleu (faible), à rouge (fort), chaque trait horizontal correspondant à un marqueur (IKSR 2015).

L'utilisation des statistiques (analyses en composantes principales) sur les 7767 marqueurs sélectionnés (Illustration 9), permet de mettre en évidence des similitudes entre certains pays d'une part Serbie, Roumanie, Bulgarie et Ukraine et entre Croatie et Hongrie. Au contraire, certains pays (Allemagne, Autriche et Slovaquie) ont des empreintes plus spécifiques.

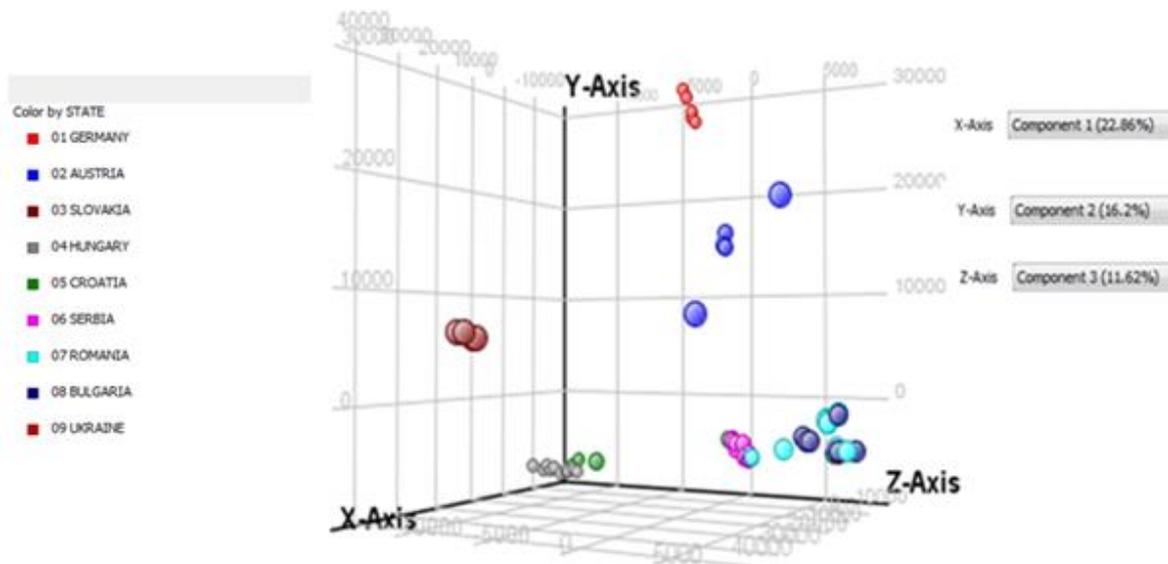


Illustration 9 – Similarités des profils de pollution selon les pays le long du Danube obtenues par Analyses des Composantes Principales (IKSR 2015).

Pour les techniques de CG/SM (2 méthodes), appliquées sur 22 échantillons, une proposition d'identification a pu être faite pour un total de 586 composés. Selon les méthodes, entre 29% et 38% des marqueurs n'ont pu être identifiés.

Comparativement, le précédent exercice du Joint Danube Survey, utilisant une technique d'extraction différente, avait permis l'identification de 158 composés sur 124 stations échantillonnées. Ce résultat montre l'intérêt de l'étape de préparation de l'échantillon, indépendamment de l'évolution des techniques de mesures.

Approche criblage suspect :

Une liste de 315 composés a été recherchée dans les échantillons. Parmi ceux-ci, 110 substances ont été détectées dans au moins 1 échantillon. Les fréquences de détection ont été calculées (Illustration 10), pour chacune d'entre elles, permettant de prioriser celles dont l'identification doit être confirmée (passage niveau 1) et le développement de méthodes de quantification spécifiques potentiellement envisagé.

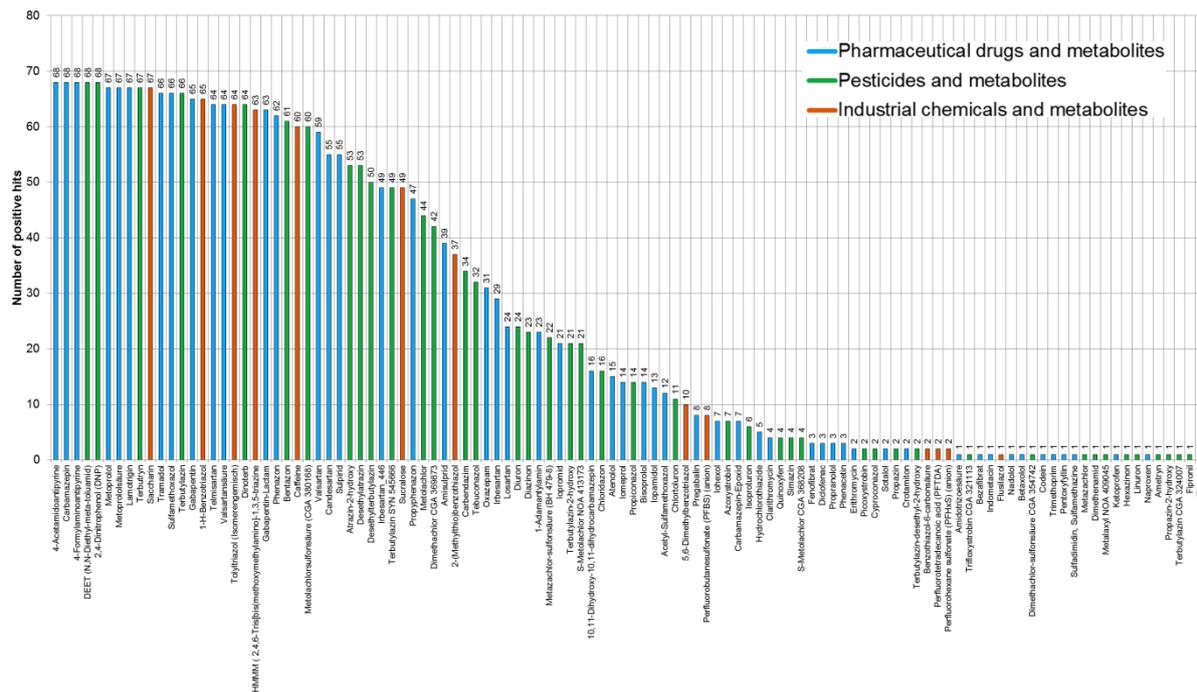


Illustration 10 – Fréquence de détection des 110 composés suspects détectés dans les échantillons, sur les 315 recherchés (IKSR 2015).

Analyse rétrospective :

Les analyses obtenues par CG/SM ont été bancarisées pour une analyse rétrospective. Des molécules identifiées par l’approche non ciblée lors des analyses CL/SMHR ont été recherchées dans ces données, permettant la confirmation de la détection de 13 composés (analysables par les deux méthodologies) et montrant la faisabilité de ce traitement a posteriori.

Conclusion de l’étude

L’intérêt de cette campagne est manifeste, avec la mise en évidence de nouveaux composés, des profils de contamination tout au long du Danube, une identification des affluents vecteurs de contamination (même sans identification des composés incriminés). Les équipes signalent l’intérêt d’améliorer cette surveillance en réalisant de nouveaux échantillonnages saisonniers (4 fois par an) pour obtenir une meilleure image des pollutions plus aléatoires (cycliques ?) comme par les pesticides, les antibiotiques. Ils mettent de plus en évidence la nécessité d’intégrer ces approches dans la priorisation européenne (action NORMAN sur les polluants émergents)

3.4.3 Le programme « d’analyse spéciale » du CIPR (Commission internationale pour la protection du Rhin)

En 2011 et 2013, la CIPR a organisé des exercices pour identifier à l’aide de la SMHR de nouvelles substances dans l’eau du Rhin (Rhin 2014).

Pour la campagne de 2011, 10 stations de surveillance sur le Rhin et une par principal affluent (Meuse, Main, Neckar, Aare, Glatt, Thur, lac de Constance) ont été retenues. Un seul prélèvement par station a été effectué, correspondant selon les cas à des échantillons composites de 1, 7 ou 14 jours. 300 substances ont été recherchées en approche ciblée.

Concernant le volet non - ciblé, en 2011, 8 000 substances inconnues ont été retenues (attribution d'une formule brute « vraisemblable » (niveau 4) après filtre des marqueurs attribués au bruit de fond. Les molécules ont été priorisées pour l'étape d'identification selon deux approches :

- Une première se focalisant sur les composés comportant un des groupements chlorés (souvent des composés persistants, avec un fort impact potentiel sur le Rhin, mis en évidence par le profil isotopique très spécifique du chlore). Pour cette première sélection, le processus d'identification complet a été mis en œuvre et l'acide 2,4-Dichlorobenzoïque, l'acide 3-(2-chloroanilino)-propanoïque et la tizanidine (myorelaxant) ont été identifiés (niveau 1).
- Une deuxième approche s'est focalisée sur les 150 marqueurs les plus intenses. Pour ces composés, une comparaison du jeu d'échantillon (correspondant au suivi longitudinal du Rhin de l'amont vers l'aval) a été effectuée. Un point de référence théorique a été fixé à Bâle. Les marqueurs présents en amont de la 1ère station de Bâle ont été éliminés (considérés comme du bruit de fond) puis n'ont été retenus que les marqueurs présents dans tous les échantillons en aval de ce point (126). Un examen attentif de des 126 marqueurs a été mené, permettant d'obtenir 57 marqueurs exploitables (spectres trop bruyants, interférences, blancs pollués...).

Pour les substances les plus fréquemment détectées, le processus d'identification complet a été mis en œuvre. Ainsi 4 nouveaux composés ont pu être identifiés et confirmés (niveau 1) : 1,3-Diméthyl-2-imidazolidone , la mélamine, le triéthylphosphate et le tétraglyme (Ruff et al. 2015).

Sur la campagne de 2013, la méthadone a été identifiée et une recherche de l'émetteur principal a été menée en parallèle avec succès.

Les résultats des travaux de 2011 et 2013 ont permis l'inclusion de nouvelles substances dans le suivi régulier du Rhin. Suite à la campagne 2013, un atelier s'est déroulé à Bonn (mars 2015) intitulé « Perfectionner la surveillance des substances dans le Rhin » rassemblant un large auditoire de plus de 80 participants. Les apports de ces nouvelles approches ont été relevés mais aussi les verrous à lever, en vue de la prochaine campagne spéciale prévue en 2017.

La complexité des analyses est soulignée. A l'heure actuelle, les résultats des analyses ne sont donc pas toujours absolument comparables :

- Les « profils moléculaires » qui sont obtenus dans la fenêtre analytique de l'appareil / de la méthode, ne coïncident pas systématiquement, entre les différents appareils utilisés.
- Les critères de sélection de pics significatifs à évaluer entraînent des différences, selon les différents opérateurs, appareils ou méthodes utilisées.

Le plus haut degré de comparabilité est atteint quand tous les échantillons sont analysés par le même appareil et évalués par les mêmes spécialistes. A défaut, un niveau suffisant de comparabilité est atteint pour la communication au sein de la CIPR si on s'accorde sur des standards de diffusion des résultats sur la base des enseignements tirés de comparaisons inter laboratoires.

Il est considéré comme « indispensable que les laboratoires associés à la surveillance du Rhin ajustent leurs positions sur les standards de communication des données obtenues par analyse non ciblée afin d'assurer la comparabilité des résultats de ce type d'analyse, également au niveau de leur mode de présentation au grand public ».

La nécessité d'une comparaison inter-laboratoires est mise en évidence. Cet exercice est d'ores et déjà prévu pour 2016 et en cours d'élaboration.

Les remarques de ce groupe d'experts recourent l'intégralité des verrous identifiés par AQUAREF dans le chapitre 4.

3.5 CONCLUSION SUR L'APPLICABILITE DES APPROCHE SMHR A LA SURVEILLANCE

Les différentes applications décrites dans les paragraphes précédents montrent comment les technologies de type SMHR pourront s'inscrire pleinement dans les objectifs de la Directive Cadre sur l'Eau, que ce soit en amont de la surveillance, pour déterminer les molécules d'intérêt à considérer, que ce soit dans le contrôle opérationnel, la recherche de cause ou le suivi de l'efficacité de programme de mesures.

L'exhaustivité relative de l'analyse permet d'élargir l'information acquise, en ayant accès à de nouvelles informations sans être contraint par la définition de composés d'intérêt.

La bancarisation de l'analyse et les possibilités de traitements a posteriori de la donnée permettent un large panel d'utilisation de ces données pour répondre à différents objectifs à partir de la même acquisition analytique.

La même donnée peut être retraitée à plusieurs périodes, selon l'avancée des connaissances (ex : nouvelles molécules à rechercher), les nouveaux questionnements (dégradation d'un site : quel était son état initial ?).

Les apports de ces nouvelles technologies pour la surveillance vont néanmoins de pair avec un accroissement de la complexité analytique par rapport aux techniques actuelles. La nouveauté de ces approches entraîne un manque de validation des méthodologies, une absence de guide méthodologique...Le chapitre suivant a pour but d'identifier les différents verrous méthodologiques et organisationnel qui semble nécessaire de lever avant d'envisager un déploiement dans le cadre de la surveillance réglementaire.

La partie 6 fait le récapitulatif des différentes applications envisageables pour ces approches, en y associant le degré de maturité de l'approche (verrous, transférabilité...).

4 Verrous associés aux nouveaux outils SMHR

Afin d'aider le lecteur à bien appréhender ces verrous il est nécessaire de repréciser les principes du processus de mesure qui conduisent sur la base d'un prélèvement en la production d'une donnée. Un parallèle entre l'approche conventionnelle et l'approche SMHR développée dans le présent document va être discuté.

4.1 VEROUS ANALYTIQUES

Pour rappel, le processus de mesure repose sur une succession d'étapes et de contrôles qualité depuis l'échantillonnage jusqu'à l'expression finale d'un résultat qui doivent être clairement définis, maîtrisés et si possible harmonisés (Illustration 11).

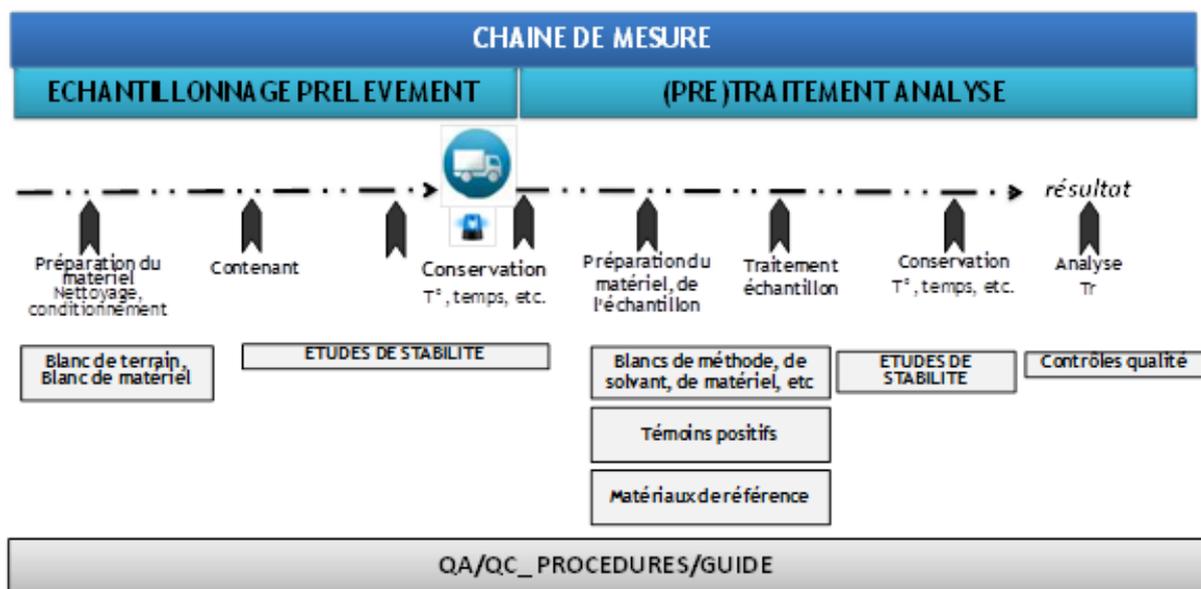


Illustration 11 – De l'échantillonnage au résultat final : Etapes et contrôles qualité

Dans le cadre de la surveillance conventionnelle c'est-à-dire orientée sur une liste de molécules connues, l'objectif est d'obtenir une information quantitative quant à l'occurrence de celles-ci dans les milieux aquatiques. Pour ce faire, l'ensemble des opérateurs impliqués dans la surveillance : préleveurs et laboratoires d'analyse mettent en œuvre des procédures harmonisées et des méthodes dont les performances sont connues pour chacune des molécules concernées. Dans cette approche, l'ensemble de ces éléments participe à la sélection de la meilleure approche disponible pour garantir la qualité de la donnée.

- Approche conventionnelle "molécules cibles" :
 - les procédures et méthodes sont optimisées pour chaque molécule ciblée.
 - leurs performances connues a priori participent à l'évaluation de leur adéquation aux objectifs

Au contraire, dans le cadre d'une surveillance basée sur une approche non ciblée dont l'objectif est d'identifier des molécules "nouvelles" pertinentes pour une future surveillance, l'ensemble des opérateurs impliqués dans la surveillance: préleveurs et laboratoires d'analyse mettent en œuvre des procédures et des méthodes de compromis. L'évaluation de l'adéquation des

méthodes (performances) quant à la molécule nouvellement connue ne pourra être déterminée qu'a posteriori.

- Approche “molécules non cibles” :
 - repose sur des approches de compromis ;
 - les méthodes et procédures d'identification d'une molécule sont dépendantes des méthodes mises en œuvre.
 - l'adéquation des procédures et méthodes pour une molécule sont connues a posteriori et impliquent une moins grande confiance quant à la qualité de la donnée aux objectifs.

Ceci implique deux considérations :

- ce n'est pas parce qu'une molécule est identifiée qu'elle était réellement présente dans le milieu. C'est le concept de faux positif ;
- ce n'est pas parce qu'une molécule n'est pas mise en évidence qu'elle n'était pas réellement présente dans le milieu. C'est le concept du faux négatif.

Ainsi on comprendra que dans une optique de mise en œuvre aux objectifs de la surveillance discutées précédemment, il est indispensable de:

- mettre en place des procédures et méthodes d'analyse harmonisées accompagnées de contrôle qualité;
- définir des critères d'acceptabilité des données brutes ;
- définir des workflow de traitements “chimométriques” et d'interprétation des données brutes ;
- réfléchir à un mode de bancarisation,

afin de permettre l'exploitabilité des informations recueillies.

En effet, il n'existe aucune idée préconçue sur la nature des molécules à rechercher par l'approche SMHR. Ainsi, pour chaque étape analytique (échantillonnage, préparation d'échantillon, séparation chromatographique, détection), la méthode peut ne pas être compatible avec les molécules potentiellement présentes. Afin de pouvoir maximiser la détection de molécules, il faut mettre en place pour chaque étape analytique des méthodes complémentaires qui permettront de couvrir l'ensemble des propriétés physico-chimiques (de volatils à non-volatils, de non-polaires à très polaires) et à terme d'acquérir une réponse spectrométrique pour toutes les molécules présentes dans le milieu (Illustration 12).

La mise en place d'une méthode générique permettra de couvrir autour de 70% des molécules présentes (extraction générique mais analyse par CL et par CG/HRMS) à des niveaux de sensibilités suffisants par rapport aux concentrations environnementales. Cependant, pour le reste des molécules, des méthodes spécifiques à chaque famille (volatils, très polaires, s'ionisant avec une phase mobile particulière,...) devront toujours être mises en place.

L'utilisation de traceurs (composé isotopique) représentant les principales familles ou propriétés physico-chimique d'intérêt devra être employée afin de pouvoir refléter l'efficacité de la détection selon la matrice/outil (eau/sédiment/biote/échantillonneur passif) utilisé et sa complexité en termes de composition matricielle.

Le groupe de travail national mis en place sur le criblage non ciblé (hors AQUAREF) vise à échanger et établir une tendance de pratiques appliquées pour cet exercice. Sur cette base et sur les travaux AQUAREF envisagés en 2016 et 2017 (essais collaboratif entre les laboratoires), des recommandations seront effectuées par AQUAREF.

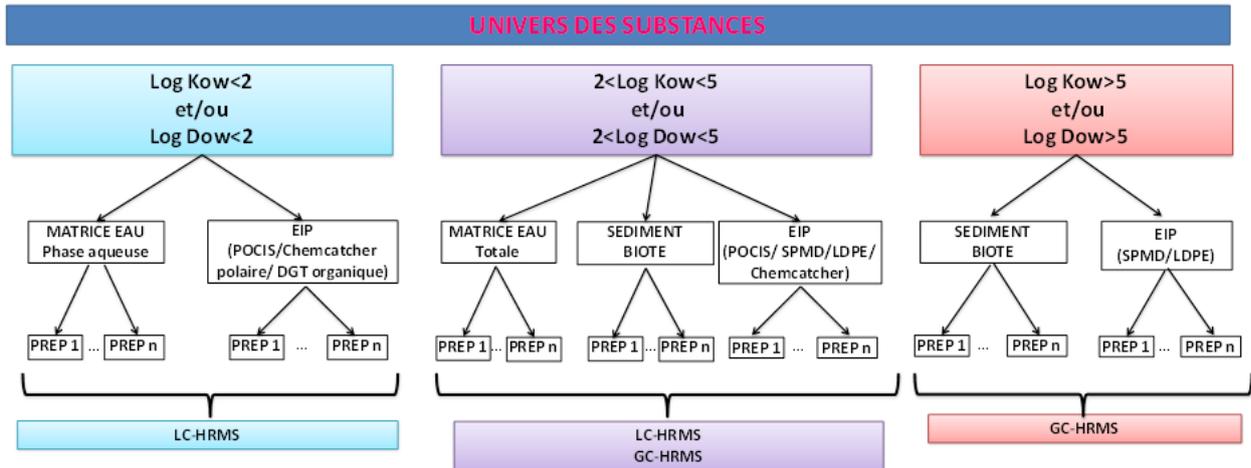


Illustration 12 – Procédures et méthodes d’analyse harmonisées pour la SMHR

Dans le cas de l’approche des molécules « non ciblées », pour chaque échantillon il est nécessaire de :

- Connaître les méthodologies utilisées, conditions d’extraction et d’analyse, afin de faciliter le traitement des données *a posteriori* (bancairisation).
- Fixer des conditions pour la mise en œuvre de la surveillance pour l’homogénéité des données.

4.2 VERROUS LIÉS AU TRAITEMENT DES DONNÉES BRUTES

Un certain nombre de verrous liés au traitement des données brutes ont été identifiés. Ils sont de manière générale communs aux criblages ciblés, suspect et non ciblé, mais d’importance différente. Ils sont décrits dans les paragraphes suivants et repris dans un tableau bilan en fonction de leur contribution à chaque type de criblage.

4.2.1 Prise en compte des blancs/contrôles qualités

Une des difficultés essentielles du traitement des données par SMHR est que tout ce qui atteint le détecteur devient ionisé dans la source du spectromètre et est enregistré. Aucun filtre ou sélection n’est effectué. Les contaminations sont pratiquement impossibles à éviter dans ce type d’analyse. Ainsi, si les composés d’intérêt sont, dans le meilleur des cas, enregistrés, il en est également de même pour de nombreuses substances parasites qui ont été collectées entre l’échantillonnage et l’analyse. Cela a pour conséquence d’entraîner une pollution inévitable des données obtenues et de complexifier le traitement de celles-ci.

En SMHR il y a un fort enjeu sur les blancs « propres ». La vérification du bruit de fond, blanc système (phase mobile) et blanc d’injection est indispensable avant le lancement d’une analyse. Les sources de contamination pouvant se trouver à différents niveaux : pollution de l’appareil analytique (colonne, source, etc.), des phases mobiles et/ou du blanc d’injection. De plus il est nécessaire de faire un « blanc d’extraction », suivant le protocole d’extraction en parallèle des échantillons. Celui-ci permettra de mettre en évidence des contaminations potentielles des échantillons au cours de l’extraction.

Avant de procéder à une analyse des données pour la recherche des molécules d’intérêt présentes dans l’échantillon, il est donc, au préalable, indispensable d’effectuer une comparaison des blancs du processus analytique. AQUAREF devra ainsi parvenir à établir une

procédure adaptée à la gestion des blancs et à recommander des pratiques d'usage dans ce cadre.

4.2.2 Paramètres à prendre en compte dans le retraitement des chromatogrammes

Une fois le signal enregistré, les logiciels de traitement de données doivent être capables d'identifier les différents marqueurs individuellement par rapport à la masse d'informations qui a été générée. En effet, l'enregistreur (le spectromètre de masse) collecte continuellement des spectres de masse. Cependant, un même pic chromatographique peut correspondre à plusieurs composés (coélution) et donc contenir plusieurs spectres de masse. Il est nécessaire de déconvoluer les signaux.

Les chromatogrammes doivent être « propres » en éliminant les artefacts et le bruit. A l'heure actuelle la majorité des logiciels de constructeur le font automatiquement, il est possible de sélectionner des filtres sur l'intensité.

Les extractions des masses exactes pour des molécules ciblées, suspectes ou non ciblées peuvent se faire sur la totalité du chromatogramme, ce qui va entraîner de multiples réponses. En effet comme expliqué en annexe 1, plusieurs molécules possèdent la même masse exacte mais pas la même structure (et donc des temps de rétention différents). Si l'information TR est connue il est nécessaire d'ajuster la fenêtre de temps de rétention. Si on prend le cas de l'aténolol (Illustration 13), plus la tolérance sur le temps de rétention est réduite (fenêtre de recherche), moins de signaux chromatographiques sont pris en compte. Dans le cas de l'analyse « suspecte » (connaissance imparfaite du temps de rétention), le choix du dimensionnement de cette fenêtre peut entraîner l'identification d'un faux positif ou négatif. C'est à ce niveau que l'utilisation des modèles de prédiction des temps de rétention simplifie le travail de l'opérateur (Annexe 2).

D'autres filtres (tolérances) peuvent être utilisés sur l'erreur de masse, au niveau du composé ciblé, des fragments (théoriques ou expérimentaux) à détecter... Le choix des valeurs appliquées comme filtre doit être totalement réfléchi et expliqué. De mauvais filtres peuvent donner lieu à l'identification de faux positifs (molécule identifiée mais pas réellement présente dans le milieu) et faux négatifs (molécule non identifiée alors que présente dans l'échantillon).

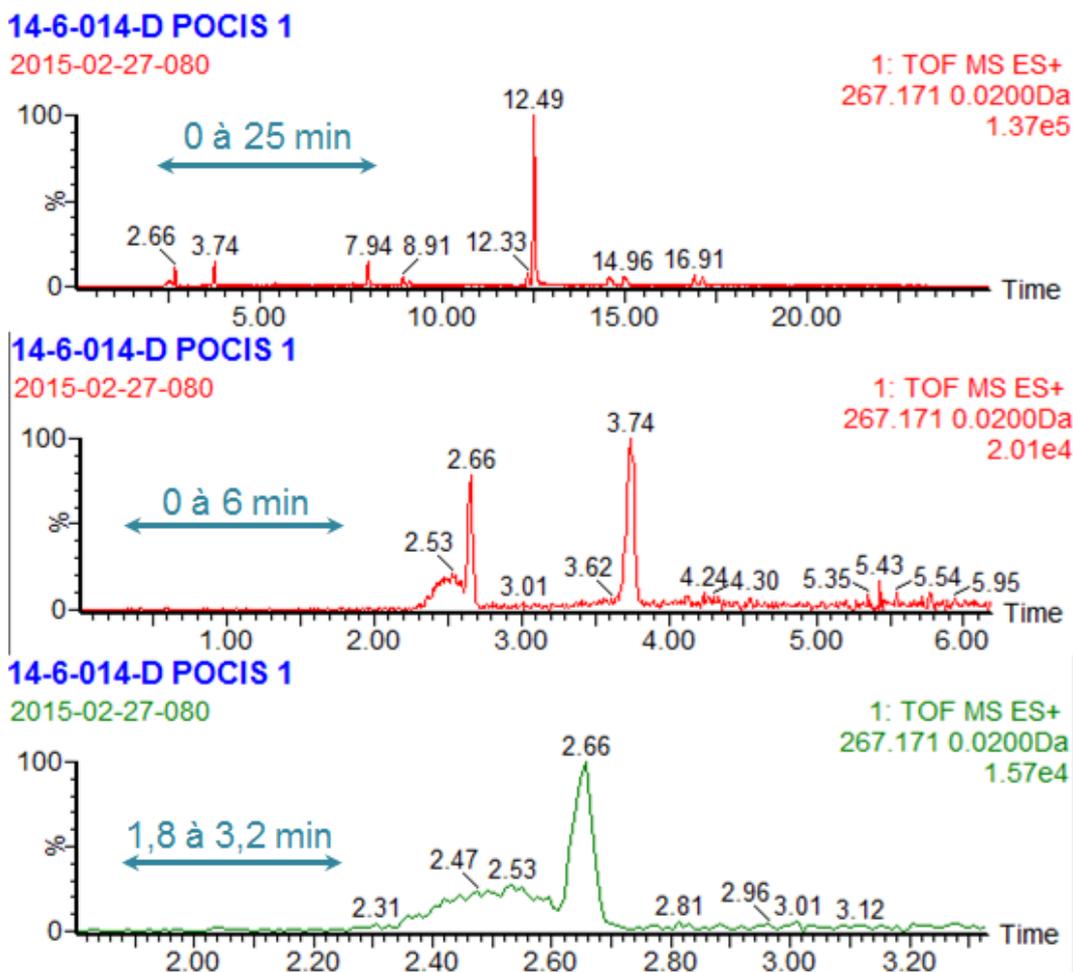


Illustration 13 – Influence du dimensionnement de la fenêtre du temps de rétention sur le nombre de signaux « atenolol » identifiés.

4.2.2.1 Multi-constructeurs /logiciels

Chaque constructeur a développé ses propres algorithmes que ce soit pour la prise en compte du blanc ou pour le traitement du signal. Il existe également des logiciels libres indépendants du constructeur permettant de traiter les échantillons. Or les résultats du traitement d'un même signal par différents logiciels pourraient ne pas être pas forcément identiques. Ce point sera particulièrement évalué dans l'étude technique AQUAREF 2016 où si cela est possible, un même signal sera traité avec des logiciels de constructeurs différents afin d'étudier la variabilité de résultats.

4.2.3 Bases de données

Depuis de nombreuses années les BDD pour la chromatographie gazeuse couplée à la spectrométrie de masse sont utilisées pour l'identification de molécules. Elles peuvent être utilisées pour tout appareil, les données sont reproductibles et robustes du fait du mode d'ionisation par impact électronique à énergie d'ionisation constante (70 eV). Cependant les contraintes ne sont pas les mêmes lorsqu'il s'agit de BDD dont les techniques d'ionisation ne sont pas reproductibles (ESI, APCI...). Plusieurs verrous sont mis en évidence et discutés dans les paragraphes ci-dessous.

4.2.3.1 Qualité des bases de données

L'identification des molécules reposant sur la comparaison avec des BDD spectrales dépend ainsi fortement de la qualité de celles-ci. En effet, si leur établissement n'est pas effectué dans des conditions représentatives, les données utilisées ne permettront pas d'identifier les molécules d'intérêt et peuvent même aboutir à l'identification de faux positifs ou négatifs. L'alimentation des spectres des BDD doit ainsi être effectuée avec des données fiables et dûment vérifiées. Ainsi, les bases de référence devraient contenir des spectres produits uniquement avec des étalons analytiques. Si ce n'est pas le cas, cela devrait être clairement mentionné afin de fournir aux utilisateurs une indication sur la fiabilité de la source. Les BDD des fabricants peuvent être considérées comme fiables car elles ont été établies avec les spectres des étalons. Pour la base de données MassBank, des recommandations ont également été établies dans ce sens et les données sont généralement validées avant intégration.

Globalement, pour chaque identification à partir de bases de données spectrales, il est indispensable de mentionner quelle base de données a été utilisée. Un critère de confiance pourrait être attribué suivant l'origine des bases de données utilisées (ou éventuellement au niveau du spectre).

4.2.3.2 Confirmation de l'identité par rapport aux bases de données

Les spectres des BDD peuvent être extrêmement différents en termes de qualité d'information qui les compose. Indépendamment du matériel utilisé, certaines molécules ne se fragmentent pas ou peu en fonction de l'énergie de collision associée ; ce qui induit que les spectres de ces molécules contiennent des informations différentes sur les fragments. Il devrait ainsi être défini quels sont les critères d'identification associés à l'utilisation des informations contenues dans les bases de données spectrales. Comme pour la confirmation d'identité à base d'étalon, les critères d'identification par point et par abondances isotopiques édictés dans la NF XP T 90-214 pourraient également être utilisés dans ce cadre.

Dans certains cas et pour certaines molécules, plusieurs bases de données indépendantes pourraient être utilisées conjointement afin de pouvoir augmenter le degré de certitude sur l'identité de la molécule suspectée.

4.2.3.3 Compatibilité des spectres des bases de données

Un des verrous majeurs de ce type de démarche est qu'il n'y a pas une harmonisation des pratiques en termes de constitution des bases de données. Chaque constructeur ayant sa technologie propre, les spectres produits peuvent être très différents pour une même molécule. De plus, les stratégies en termes de production de données spectrales sont différentes. Ainsi, certaines bases de données sont mises en œuvre en faisant varier l'énergie de collision par palier fixe (par exemple : 0, 10, 20, 30 ... eV) alors que d'autres le sont en effectuant des gradients. De même, les modes d'acquisition utilisés peuvent varier : acquisition de données dépendantes (typiquement en SM/SM les ions fragments ne proviennent que de l'ion « isolé » par le 1^{er} quadrupole) ou acquisition de données indépendantes (mode All ion SM/SM chez Agilent, MS^E chez Waters). Ainsi, cela implique que pour une même molécule analysée, les fragments formés ne vont pas être nécessairement identiques ou dans les mêmes proportions. Pour la préparation d'une réunion consacrée au criblage non-ciblé, il avait été demandé à chaque participant d'injecter des composés connus afin de pouvoir appréhender les différences de fragmentation. Dans le cas du diuron, si certaines similitudes ont été remarquées, des différences notables ont également été constatées.

Ces verrous s'appliquent particulièrement pour les composés analysés par chromatographie en phase liquide couplée à la SMHR. Pour la chromatographie en phase gazeuse couplée à la spectrométrie de masse, les énergies d'ionisation généralement fixées (à 70 eV) ce qui permet de disposer de bases utilisables quel que soit le fabricant.

En CL/SMHR, la question de l'interopérabilité des BDD entre différents constructeurs voire même entre différentes générations d'un même équipement représente, à l'heure actuelle, un réel frein dans l'efficacité des recherches sur la comparaison de spectres. Ce point devra être approfondi afin de pouvoir statuer sur les pratiques d'utilisation des bases de données dans cette optique.

4.2.3.4 Compatibilité des formats des bases de données

Les bases de données spectrales sont établies selon des formats propres aux fabricants ou aux développeurs. Ainsi, elles ne sont pas compatibles entre elles pour effectuer une recherche automatisée. Une conversion des formats ou de de l'extension des fichiers de spectres devrait être opérée pour les rendre compatibles ce qui actuellement n'est pas couramment faisable.

4.3 VERROUS LIES A L'ECHANTILLOTHEQUE

4.3.1 Bancarisation des métadonnées associées à l'analyse

Plusieurs paramètres analytiques peuvent jouer sur l'exhaustivité de la méthode utilisée pour le criblage de molécules (paragraphe 4.1). Il est important de lier aux données brutes les paramètres d'échantillonnage, d'extraction et d'analyse (méthode, appareil, ionisation, énergie de collision, etc.). En effet si une nouvelle molécule est identifiée, il sera possible de vérifier que celle-ci est extractible/analysable avec la méthode appliquée.

La bancarisation de ces paramètres pourra aussi servir pour faire de la semi-quantification *a posteriori*. En effet, avec le standard de référence d'une nouvelle molécule identifiée il sera possible d'estimer la réponse de celle-ci par rapport aux traceurs.

4.3.2 Bancarisation des métadonnées associées au retraitement

Les traitements des données brutes évoluent au fur et à mesure des connaissances. Par exemple des molécules identifiées comme « suspectes » dans un échantillon pourront devenir ciblées dans un autre à partir du moment où leurs étalons de référence auront été acquis et injectés. Et des molécules identifiées comme « non-ciblées » dans un échantillon pourront être ensuite ajoutées aux BDD de molécules suspectes, etc. Un échantillon traité deux fois (la première au temps t et la seconde $t+6$ mois par exemple) aura des résultats différents car les informations contenues dans les BDD auront évoluées. Il est donc important d'associer aux données brutes les informations concernant les BDD ou les traitements utilisés.

4.3.3 Besoin que « n'importe qui » puisse retraiter les données acquises et bancarisées

La bancarisation de toutes ces données entraîne un problème de stockage en lien avec la taille des fichiers d'analyse produits. A l'heure actuelle ce problème se pose à tous les laboratoires utilisant cette technique d'analyse. Les données brutes issues des échantillons, accompagnées des blancs, contrôle qualité, etc. représentent rapidement un volume de données très conséquent et à conserver de manière fiable et pérenne.

Il est aussi important de préciser que toutes les données brutes sont volumineuses et doivent être conservées dans le cas de traitement a posteriori. De plus elles ont un format propre aux constructeurs, ce qui implique un problème de compatibilité. En effet il est important de pouvoir s'échanger des données entre laboratoires dans le but de faire par exemple des exercices d'intercalibration.

4.4 POURSUITE DES ACTIONS EUROPEENNE ET FRANÇAISE EN COURS

Les travaux NORMAN et AQUAREF viseront à apporter des éléments de réponse à ces trois grands types de verrous : préparation de l'échantillon, traitement du signal et base de données.

L'action AQUAREF 2016, devrait permettre de répondre aux verrous concernant le traitement du signal. Un extrait d'un échantillon d'eau et un extrait de POCIS seront envoyés aux laboratoires partenaires dans le but de mieux appréhender les différences de résultats obtenus entre un même extrait (excluant ainsi la variabilité de l'échantillonnage/préparation d'échantillon) traités par des opérateurs et du matériel différents. Elle contribuera ainsi à une meilleure compréhension des variations provoquées particulièrement dans l'utilisation de filtres de traitement de données et de pratiques dans la prise en compte des blancs.

L'action AQUAREF 2017, si mise en œuvre, devrait permettre de répondre aux verrous concernant la préparation d'échantillon. Elle aura pour but de comprendre l'intégralité de la chaîne de mesure, en incluant les étapes d'extraction et d'analyse en amont du retraitement. La réalisation de cet essai comparatif s'effectuera sur 2 ou 3 sites. La comparaison permettra d'apporter des éléments sur l'importance de la variabilité des résultats en prenant en compte la préparation d'échantillon. La pertinence d'utiliser des approches différentes mais potentiellement complémentaires pourra aussi être évaluée.

Les verrous/questions concernant les bases de données sont abordés dans le cadre des travaux NORMAN. Les membres d'AQUAREF vont continuer à suivre et participer à ces activités (action F1b- Stratégie d'interprétation et exploitation des données de criblage non-ciblé en lien avec le réseau NORMAN pour préparer l'application de ces techniques dans la future surveillance 2016-2018). Une interaction plus active est prévue avec notamment, le cas échéant, la participation aux actions NormaNews, à Massbank et Envilight. Une deuxième action plus technique vise à renforcer les connaissances dans la pratique des analyses non ciblées afin de pouvoir alimenter des recommandations techniques. Ainsi, l'action « Application approche criblage non-ciblé sur site pilote - objectif: comparaison des pratiques des laboratoires et recommandations pour application dans la future surveillance » (G2c) est planifiée pour le cycle triennuel 2016-2018.

L'action AQUAREF de 2018, si mise en œuvre, sera dédiée à effectuer le bilan des travaux accomplis en 2016 et 2017. Elle visera dans un premier temps à établir un bilan des essais comparatifs. Les différents retours d'expérience seront mis à profit afin de pouvoir constituer un guide technique. Il intégrera des préconisations sur les verrous identifiés dans ce travail pour l'application des techniques de criblage pour la future surveillance. Ces recommandations pourront également être utilisées pour des travaux de normalisation consacrés à ce domaine.

Une journée technique est également prévue en 2018 afin d'informer les laboratoires prestataires des différentes conclusions obtenues à la suite de ces études et de leur apporter des recommandations de pratiques afin de fiabiliser les données produites dans ce type d'étude.

5 Conclusions

Les travaux sur l'analyse non ciblée ont pris un essor important ces dernières années dans de nombreux domaines liés à la mesure de micropolluants. Ainsi, ce type d'activité s'est notamment largement développé dans le champ des analyses alimentaires.

Dans le domaine environnemental, de nombreux travaux sont désormais focalisés sur l'amélioration des connaissances pour la mise en œuvre de cette technique. Au niveau Européen, certains des principaux acteurs de ce domaine font partie du réseau NORMAN. Ainsi, la majeure partie des activités de NORMAN dans le cadre du volet analyses physico-chimiques est dédiée à cette thématique. Les groupes de recherche les plus actifs dans ce domaine sont les Suisses de l'EAWAG et les allemands d'UFZ. Ces activités se sont traduites par l'organisation d'un essai interlaboratoire en 2013 pour lequel 18 instituts de 12 pays différents ont participé (Schymanski et al. 2015). Des workshops ont été également régulièrement organisés sur cette thématique (Septembre 2014-Zurich (Suisse), Septembre 2015- Rhodes (Grèce)) afin de présenter les résultats et conclusions de ces études et établir des perspectives. Ainsi, lors de la dernière réunion à Rhodes, les sujets ont particulièrement porté sur la création d'une liste collaborative de polluants émergents (NormaNews), sur des progrès récents dans la prédiction des temps de rétention, et la bancarisation de données brutes SMHR (EnviLight). Les membres d'AQUAREF ont participé régulièrement à ces différentes manifestations afin de pouvoir suivre les évolutions dans ce domaine.

Ces différentes réunions ont également confirmé l'engouement au niveau des principaux instituts de recherche et de surveillance de l'environnement européens, signe qu'il est indispensable qu'en France nos laboratoires s'emparent du sujet. Cette constatation a débouché sur la création d'un groupe d'utilisateurs pratiquant ou ayant un intérêt pour la pratique d'analyses de criblage suspecté ou non ciblé. Deux réunions regroupant une dizaine d'établissement comprenant des organismes de recherche publics et privés, dont des membres d'AQUAREF ont eu lieu le 22 Janvier 2015 et le 23 Juin 2015. Une troisième est prévue début Janvier 2016. Ces réunions ont permis de recenser les capacités en termes instrumentales et le niveau d'expertise de certains laboratoires nationaux. Elles ont également permis de mettre en évidence les principaux points d'intérêt communs à tous les participants. Ainsi, il apparaît nécessaire de dégager des pratiques qui pourraient être recommandées dans tous les compartiments du processus analytiques, du prélèvement jusqu'à l'analyse des données. Lors de la réunion de normalisation de la commission T91M en Octobre 2015, des discussions ont également été entamées à ce sujet avec les laboratoires prestataires afin de réfléchir à la création d'un groupe de travail qui pourra établir un texte de référence visant à encadrer et délimiter les pratiques dans ce domaine.

Globalement, des demandes fortes ont été exprimées par différents laboratoires afin de pouvoir disposer de textes comportant des recommandations sur lesquelles s'appuyer. L'intérêt d'un tel texte sera également de contribuer à éviter que des pratiques trop libérales soient mises en œuvre par les laboratoires privés au risque d'aboutir à des résultats ne correspondant pas la nature et composition de l'échantillon. L'intégration de ces nouvelles approches pour la surveillance réglementaire et de ce fait la mise en œuvre par des laboratoires prestataires nécessitera l'existence de cadre(s) et recommandations.

6 Bilan sur le criblage de molécules en fonction des objectifs de la surveillance

Les fiches suivantes visent à synthétiser les différents éléments du rapport en associant à chaque objectif de la surveillance pour lequel l'approche criblage nous paraît pertinente :

- Les bénéfices envisagés
- les verrous associés
- le degré de maturité de l'application
- la faisabilité et facilité du transfert (si possible) vers les laboratoires opérateurs.

Légende :



Facile, Faible, Rapide

Terme



Moyen



Complexe, Long

Cette classification est effectuée dans l'état actuel de nos connaissances et est amenée à évoluer au cours des prochaines années, en fonction de la levée de certains verrous mais aussi de la mise en évidence de certains autres.

SURVEILLANCE REGLEMENTAIRE ETAT CHIMIQUE	
DESCRIPTION DE L'OBJECTIF	Analyse quantitative de composés ciblés (agréments) dans une matrice
INTERET	Abaissement des limites de quantification Diminution des faux positifs
BENEFICES	Compatibilité avec les exigences de l'agrément Amélioration de la qualité des mesures
TYPE(S) DE TRAITEMENT	criblage ciblé
COMPLEXITE	
ANALYSE A POSTERIORI	Non applicable
VERROUS METHODOLOGIQUES	
APPLICABILITE LARGE ECHELLE	
TEMPS D'ACQUISITION	
TRANSFERABILITE TRANSFERT	

APPROCHES BIOANALYTIQUES	
DESCRIPTION DE L'OBJECTIF	Compréhension des causes de déclassement (mauvais état chimique des masses d'eau) par approche bioanalytique (ex EDA)
INTERET:	Amélioration des listes de surveillance Amélioration de la gestion des milieux
BENEFICES	Identification des molécules responsables des effets toxiques dans les milieux
TYPE(S) DE TRAITEMENT	criblage ciblé, suspect Analyse non-ciblée/identification de composés
COMPLEXITE	
ANALYSE A POSTERIORI	Applicable
VERROUS METHODOLOGIQUES	
APPLICABILITE LARGE ECHELLE	
TEMPS D'ACQUISITION	
TRANSFERABILITE TRANSFERT	

CONTROLE D'ENQUETE	
DESCRIPTION DE L'OBJECTIF	Identification de l'origine de la dégradation de l'état des masses d'eaux.
INTERET:	Amélioration de la gestion des milieux Gestion des pollutions à la source
BENEFICES	Application du principe pollueur /payeur
TYPE(S) DE TRAITEMENT	criblage suspect analyse non-ciblée :empreintes moléculaires/ identification
COMPLEXITE	
ANALYSE A POSTERIORI	Applicable
VERROUS METHODOLOGIQUES	
APPLICABILITE LARGE ECHELLE	
TEMPS D'ACQUISITION	
TRANSFERABILITE TRANSFERT	

CONTROLE D'INVESTIGATION	
DESCRIPTION DE L'OBJECTIF	Evaluation de l'état et de l'évolution des masses d'eau présentant un risque de ne pas atteindre les objectifs environnementaux
INTERET	Cartographie les masses d'eau par des empreintes chimiques
BENEFICES	Comparaison des masses d'eau dans le temps et l'espace : mise en évidence de ce qui leur est commun et de ce qui diffère.
TYPE(S) DE TRAITEMENT	criblage ciblé, suspect Analyse non-ciblée/ empreintes moléculaires
COMPLEXITE	
ANALYSE A POSTERIORI	Applicable
VERROUS METHODOLOGIQUES	
APPLICABILITE LARGE ECHELLE	
TEMPS D'ACQUISITION	
TRANSFERABILITE TRANSFERT	

ETUDE PROSPECTIVE	
DESCRIPTION DE L'OBJECTIF	Identification de nouvelles molécules d'intérêt sur la base de leur occurrence dans l'environnement
INTERET	Recherche orientée large de composés d'intérêt dans un échantillon
BENEFICES	Information en termes de fréquence de détection Information semi quantitative à posteriori: ordre de grandeur de présence dans l'environnement
TYPE DE TRAITEMENT	Criblage ciblé et suspect
COMPLEXITE	
ANALYSE A POSTERIORI	Applicable
VERROUS METHODOLOGIQUES	
APPLICABILITE LARGE ECHELLE	
TEMPS D'ACQUISITION	
TRANSFERABILITE TRANSFERT	

7 Bibliographie

Bibliographie

- Bade R, Bijlsma L, Sancho JV, Hernandez F (2015): Critical evaluation of a simple retention time predictor based on LogKow as a complementary tool in the identification of emerging contaminants in water. *Talanta* 139, 143-149
- Bellosta V (Octobre 2005): Identification des composés organiques par spectrométrie de masse. *ESPCI*, pp. 44
- Heberger K (2007): Quantitative structure-(chromatographic) retention relationships. *Journal of Chromatography A* 1158, 273-305
- IKSR 2015: Joint Danube Survey 3 - A comprehensive analysis of Danube water quality
- Kinani S, Bouchonnet S, Creusot N, Bourcier S, Balaguer P, Porcher JM, Ait-Aïssa S (2010): Bioanalytical characterisation of multiple endocrine- and dioxin-like activities in sediments from reference and impacted small rivers. *Environ Pollut* 158, 74-83
- Letzel Thomas, Lucke Thomas, Schulz Wolfgang, Sengl Manfred, Letzel Marion (2014): OMI (Organic Molecule Identification) in water using LC-MS(/MS): Steps from "unknown" to "identified": a contribution to the discussion. *Lab & More International* 1, 14-18
- Put R, Heyden YV (2007): Review on modelling aspects on reversed-phase liquid chromatographic quantitative structure-retention relationships. *Analytica Chimica Acta* 602, 164-172
- Rhin ClpIPd 2014: Programme d'analyse chimique "Rhin" - Analyse spéciale CLHP-SM/SM 2013
- Rouessac F, Rouessac A (2004): *Analyse Chimique _ Méthodes et techniques instrumentales modernes*, 462 pages pp
- Ruff M, Mueller M, Loos M, Singer H (2015): Quantitative target and systematic non-target analysis of polar organic micro-pollutants along the river Rhine using high-resolution mass-spectrometry - Identification of unknown sources and compounds. *Water Research* 87, 145-154
- Sanchez W, Aït-Aïssa S (10-12 mars 2010): Développement de méthodes in vitro et de biomarqueurs pour la surveillance des substances chimiques et le diagnostic de pollutions multiples, Journée Micropolluants Aquatiques
- Schlabach M, Haglund P, Rostkowski P, Dye C 2013: Non-target screening - A powerful tool for selecting environmental pollutants

- Schymanski EL, Jeon J, Gulde R, Fenner K, Ruff M, Singer HP, Hollender J (2014a): Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence. *Environmental Science & Technology* 48, 2097-2098
- Schymanski EL, Singer HP, Longree P, Loos M, Ruff M, Stravs MA, Ripolles Vidal C, Hollender J (2014b): Strategies to Characterize Polar Organic Contamination in Wastewater: Exploring the Capability of High Resolution Mass Spectrometry. *Environmental Science & Technology* 48, 1811-1818
- Schymanski EL et al. (2015): Non-target screening with high-resolution mass spectrometry: critical review using a collaborative trial on water analysis. *Analytical and Bioanalytical Chemistry* 407, 6237-6255

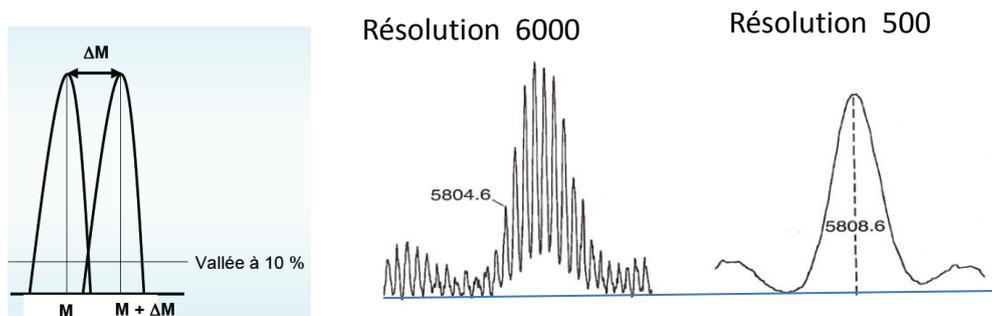
Annexe 1

Principe de la spectrométrie de masse haute résolution

La spectrométrie de masse est basée sur la détermination des masses des molécules présentes dans un échantillon. Pour cela, le composé à analyser est transformé en ion par bombardement avec des électrons ou des atomes ou des photons,... Puis ils sont soumis à l'action d'un champ électrique et / ou magnétique sous vide. Les forces qui s'exercent sur ces ions permettent alors de déterminer leur rapport *masse/charge* (m/z) (Rouessac & Rouessac 2004).

Notion de résolution

L'objectif principal d'un spectromètre de masse est de trier les molécules en fonction du rapport m/z . De ce fait la résolution, aptitude à séparer des ions de masses très proches, est une caractéristique importante de l'instrument. On peut définir la résolution comme le rapport $m/\Delta m$ (m =masse) la différence de masse minimale entre le pic considéré et son voisin le plus proche dont il peut être distingué (Annexe 1 - Illustration 1). Selon cette définition un instrument qui est en mesure de distinguer les masses 100 et 100,1 possède une résolution de $100/(100,1-100)=1000$.



Annexe 1 - Illustration 1 – Principe de la résolution en spectrométrie de masse

Il existe différents types de spectromètre de masse, issus de technologies différentes et présentant des résolutions différentes (Annexe 1 - Illustration 2) adaptés à différents objectifs de mesure et à des coûts évidemment très différents selon le niveau de technologie.

Analyseurs	Résolution $M/\Delta M$	Gamme m/z
Cyclotron	1 000 000	4 000
Trappe orbitale	100 000	4 000
Secteur Magnétique	20 000	20 000
Temps de vol (TOF)	20 000	500 000
Trappe ionique	5 000	6 000
Quadripôle	2 000	6 000

Annexe 1 - Illustration 2 – Caractéristiques des analyseurs (basse résolution $R < 2\,000$, haute résolution $R > 10\,000$ selon (Bellosta Octobre 2005))

L'utilisation d'un spectromètre de masse haute résolution permet de mesurer la masse d'un ion avec suffisamment de précision pour déterminer sa composition chimique (Bellosta Octobre 2005). Pour une molécule donnée, une analyse en basse résolution indiquera une masse de

320,1 uma (unité de masse atomique) quand une analyse en haute résolution indiquera une masse de 320,1252 uma. L'augmentation de la résolution permet ainsi de mieux définir les massifs isotopiques. Il est donc possible d'accéder aux modifications de ratio isotopiques sur des molécules (sans atteindre toutefois le niveau de capacité des systèmes dédiés aux analyses isotopiques).

Cette performance permet l'identification d'un composé, car il n'existe qu'un nombre restreint de composés organiques ayant exactement la même masse atomique à 0,0001 uma près, ceci étant dû aux masses atomiques qui ne sont pas des entiers mais aussi des abondances naturelles des éléments qui ne sont pas négligeables (Annexe 1 _ Tableau 1)

Masse exacte des isotopes				
Elément	Masse atomique	Nucléide	Masse	Abondance relative
Hydrogène	1,00794	¹ H	1,00783	100
		D (² H)	2,0141	0,016
Carbone	12,01115	¹² C	12.00000 (std)	100
		¹³ C	13,00336	1,11
Azote	14,0067	¹⁴ N	14,0031	100
		¹⁵ N	15,0001	0,38
Oxygène	15,9994	¹⁶ O	15,9949	100
		¹⁷ O	16,9991	0,04
		¹⁸ O	17,9992	0,20
Fluor	18,9984	¹⁹ F	18,9984	100
Silice	28,0855	²⁸ Si	27,9769	100
		²⁹ Si	28,9765	5,10
		³⁰ Si	29,9738	3,35
Phosphore	30,9738	³¹ P	30,9738	100
Souffre	32,066	³² S	31,9721	100
		³³ S	32,9715	0,78
		³⁴ S	33,9679	4,40
Chlore	35,4527	³⁵ Cl	34,9689	100
		³⁷ Cl	36,9659	32,5
Brome	79,9094	⁷⁹ Br	78,9183	100
		⁸¹ Br	80,9163	98
Iode	126,9045	¹²⁷ I	126,9045	100

Annexe 1 _ Tableau 2 : Masse exacte et abondance relative des isotopes (Bellosta Octobre 2005)

Par exemple une molécule de masse 60 uma peut correspondre à différentes formules brutes dont les masses exactes sont différentes (Annexe 1 - Illustration 3).

Si l'on analyse en basse résolution (technologie de type quadripolaire), on obtient une masse de 60,0 uma et il est impossible de choisir entre les formules possibles ; on ne peut donc pas identifier quelle est la molécule présente dans l'échantillon.

En haute résolution, la mesure fournie par l'appareil sera 60,0575 uma ce qui au vu de la précision de la mesure permet d'identifier sans ambiguïté la formule C_3H_8O (précision $\pm 0,001$).

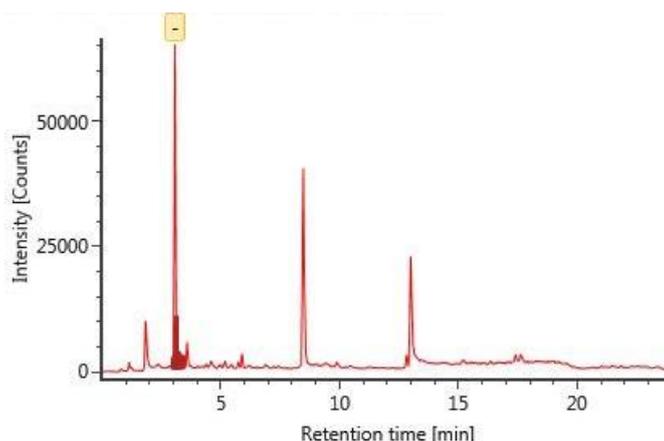


Annexe 1 - Illustration 3 – différentes structure moléculaires associées à une masse de 60,0 uma (Bellosta Octobre 2005)

Principe général d'identification des molécules

De manière générale, la SMHR est précédée d'une technique de séparation comme la chromatographie. Les composés présents dans un extrait vont être séparés plus ou moins rapidement en fonction de leur affinité physico-chimique avec la phase stationnaire de la colonne chromatographique et les phases mobiles (pour la chromatographie liquide) ou en fonction de leurs propriétés volatiles par les rampes de température (chromatographie gazeuse). La chromatographie permet en amont de la SMHR d'analyser les composés au fur et à mesure de leur élution en sortie de colonne. Le temps de rétention de chaque composé est un paramètre important pour l'identification. En effet en utilisant les mêmes paramètres d'analyse, ce temps de rétention sera toujours identique pour un composé donné, avec dans certains cas des légers décalages dus à la matrice.

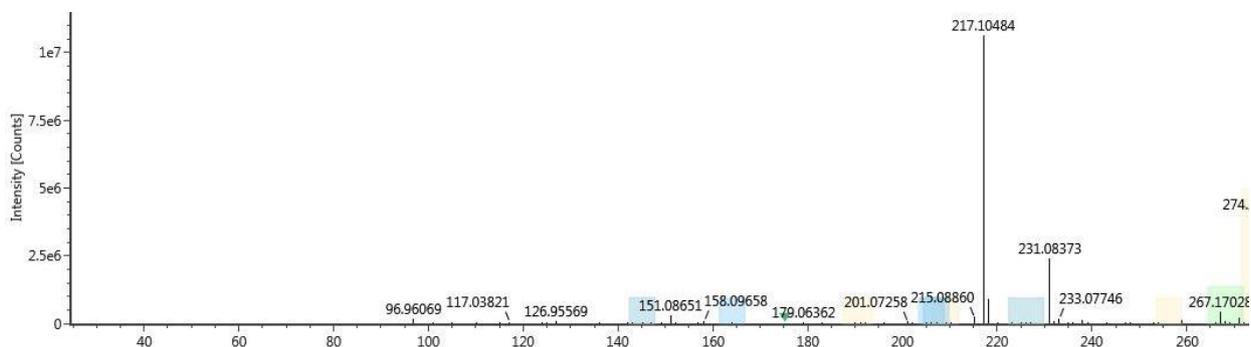
Le premier résultat obtenu après analyse d'un échantillon est le chromatogramme où est exprimée l'intensité du signal en fonction du temps de rétention dans la colonne (Annexe 1 - Illustration 4). Un signal chromatographique peut correspondre à une ou plusieurs molécules. Il est important de bien noter que chaque molécule a sa propre empreinte chimique (masse exacte, massif isotopique, fragments, ratio ion moléculaire/fragment). L'objectif est d'identifier toutes les empreintes chimiques afin d'identifier les molécules présentes dans un échantillon. Prenons l'exemple du signal à 3,08 minutes dans l'Annexe 1 - Illustration 4.



Annexe 1 - Illustration 4 – Chromatogramme

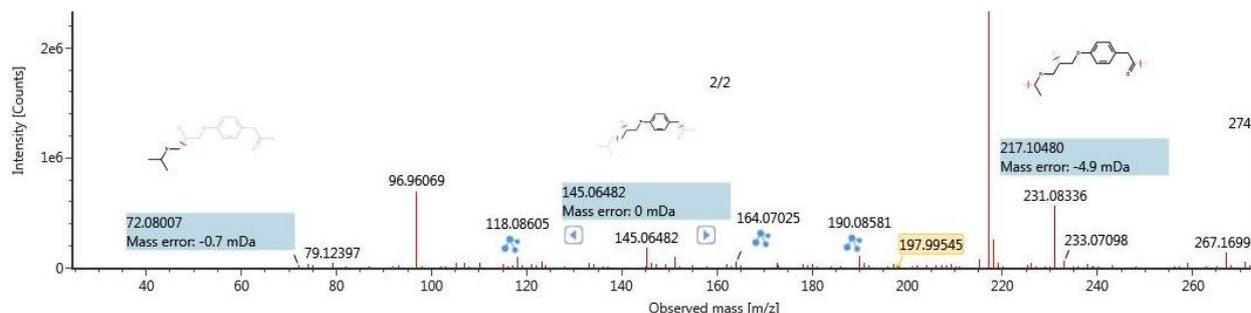
Pour identifier une molécule dont le signal chromatographique a été détecté, il est nécessaire de voir le spectre de masse, qui représente l'empreinte chimique, associé (Annexe 1 - Illustration 5), exemple pour le signal chromatographique à 3,08 minutes). Le paramètre important est d'avoir accès à la masse exacte de la molécule, plus précisément de l'ion moléculaire, ainsi qu'au massif isotopique. Ceci permet d'attribuer une ou plusieurs formule(s)

brute(s) à cette molécule. Chaque signal obtenu correspond à une masse détectée au temps de rétention 3,08 min. La masse détectée à 267,1703 uma pourrait correspondre à la formule brute C₁₄H₂₂N₂O₃.



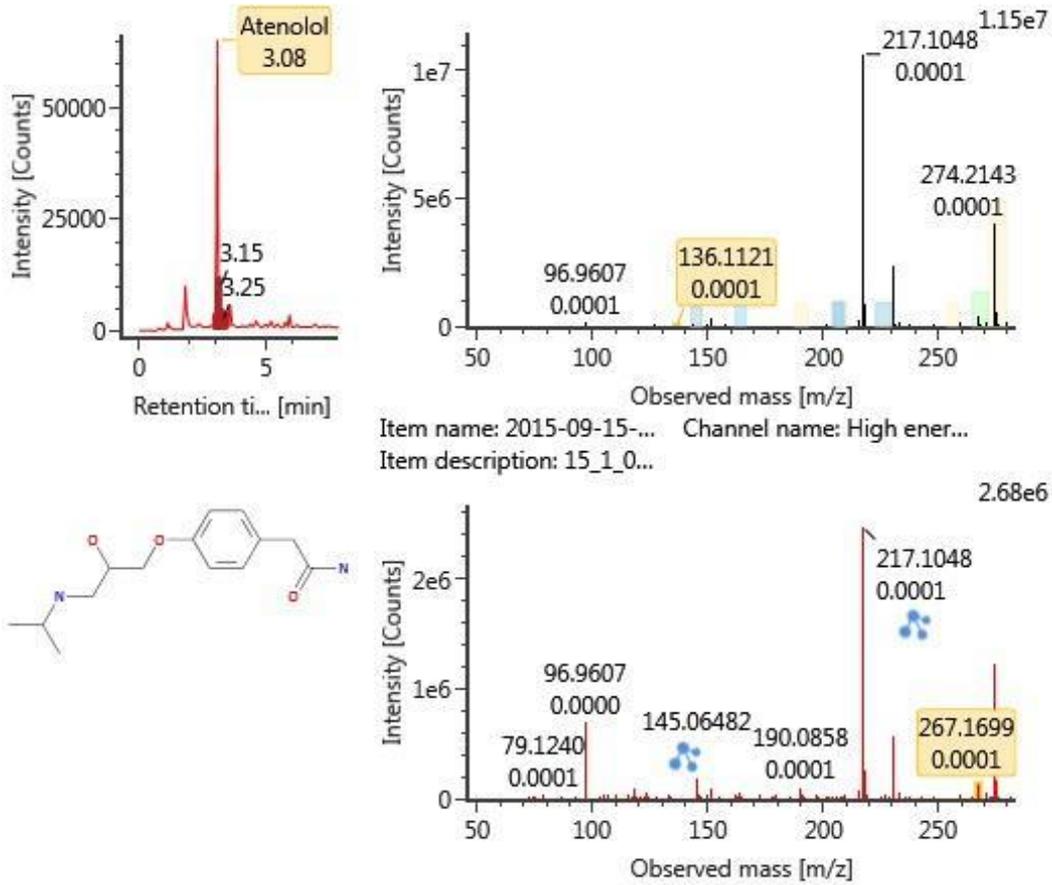
Annexe 1 - Illustration 5 – Spectre de masse de la molécule non fragmentée correspondant au signal chromatographique à 3,08 minutes

Chaque composé analysé par spectrométrie de masse peut être fragmenté par application d'un potentiel électrique et par le choc avec les molécules d'un gaz inerte dans une cellule de collision. Cette énergie permet de casser certaines liaisons spécifiques de la molécule afin d'obtenir des fragments caractéristiques du composé de départ. Cela pourrait s'apparenter à une empreinte digitale propre à chaque molécule. Ces fragments peuvent dans certains cas discriminer deux composés ayant la même masse mais de structure différente entraînant des fragments différents. Ces informations permettent d'attribuer une ou plusieurs structures à la molécule. En faisant varier le potentiel appliqué dans la cellule de collision, une fragmentation plus ou moins intense va être obtenue ce qui permet d'enrichir les données par rapport aux caractéristiques de la molécule étudiée. De plus les rapports d'intensité entre l'ion moléculaire et les ions fragments sont les mêmes d'un échantillon à l'autre pour un même appareil et une unique méthode d'analyse. La molécule correspondant au signal chromatographique à 3,08 minutes a été fragmentée. La structure des fragments a pu être déterminée pour les masses 217,1048, 190,0858, 164,0702, 145,0648, 118,0860 et 72,0801 (Annexe 1 - Illustration 6). L'ensemble de ces données permet d'identifier le signal comme étant l'aténolol.



Annexe 1 - Illustration 6 – Spectre de masse de la molécule fragmentée correspondant au signal chromatographique à 3,08 minutes

L'analyse du standard de référence de l'aténolol permet de confirmer l'identification de cette molécule en comparant le temps de rétention, la masse exacte et les données de fragmentations (Annexe 1 - Illustration 7).



Annexe 1 - Illustration 7 – Signal chromatographique et spectre de masse de l'aténolol

Annexe 2

Bases de données

Depuis la généralisation des spectromètres de masse à haute résolution, de nombreuses bases de données (BDD) ont été constituées et/ ou sont en cours de constitution. Ainsi tous les constructeurs en ont développé pour leur instrument. Elles sont généralement réparties en catégories comprenant dans la majorité des cas, pour les composés d'intérêt dans la pollution des milieux aquatiques, les pesticides et les composés pharmaceutiques. Celles-ci incluent généralement les données de fragmentation et dans certains cas, les temps de rétention chromatographiques.

En complément de ces BDD spécifiques, d'autres existent, soit recensant uniquement les formules développées, noms et masses exactes des composés (Chemspider, PubChem,...), soit comprenant des spectres (ainsi que parfois certaines métadonnées associées). A ce titre, il peut être cité MzCloud qui concerne essentiellement la technologie Orbitrap et la base de données MassBank qui est reliée aux activités de Norman.

Massbank est une BDD collaborative et librement accessible qui regroupe des spectres de composés émergents établis avec des instruments de différents constructeurs. Elle est régulièrement alimentée par les participants à ce projet, principalement les groupes de recherche japonais, suisse, allemand, grec ou norvégien. Ainsi, lorsque des polluants émergents sont identifiés par des partenaires européens, ils sont intégrés à la base de données d'où ils peuvent être recherchés par les autres partenaires. En retour, l'identification de composés inconnus (par analyse non-ciblée) au niveau français permettra également d'aider à pouvoir déterminer la présence de la substance identifiée à un niveau européen/global.

Les bases de données spectrométriques ont été principalement développées dans le cadre de l'alimentaire et contiennent ainsi majoritairement des pesticides et des produits médicamenteux (et vétérinaires) (Annexe 2). Massbank a fourni et continue une action complémentaire orientée à compléter ces listes pour les différentes classes de polluants retrouvés dans l'environnement, notamment les produits de transformation. Le contenu des bases de données est cependant actuellement limité en comparaison aux bases de données spectrales en CG/SM (200 000 composés) mais ce nombre est prévu en augmentation au fil de l'amélioration des connaissances.

L'intégration de la France et d'AQUAREF dans ce type de travaux est stratégiquement primordiale afin de pouvoir disposer d'une base de données qui permettra une recherche beaucoup plus rapide et flexible dans les campagnes de surveillance des polluants émergents.

Le Tableau 1 ci-dessous indique pour Mars 2015 l'état des différentes bases de données disponibles.

De plus, une plateforme d'échange de liste de composés identifiés ou suspectés, « NORMAN Suspect List Exchange », a été créée afin de faciliter les échanges sur les connaissances des polluants émergents (site internet NORMAN). Ces listes intègrent ainsi certains travaux de groupes de recherche de l'EAWAG (Suisse) et l'UFZ ainsi que les données recueillies lors d'un essai inter-laboratoire menés en 2013 sur le criblage non-ciblé (Schymanski et al. 2014a).

Le principal avantage de disposer de nombreuses BDD est leur complémentarité. Celles des constructeurs ne sont pas identiques en termes de composés. Ainsi, si un même échantillon est analysé par deux appareils de constructeur (et de base de données) différentes, des résultats complémentaires pourront être obtenus. Ce point sera ainsi vérifié lors d'un test prévu dans le cadre du programme AQUAREF 2016.

Tableau 1 – Récapitulatif sur les bases de données existantes (extrait de [2])

Nom base de données ou librairies	2013		Mars 2015
	Nombre total de composés	Nombre de composés avec spectres	Nombre total de composés
ChemSpider	32 millions	-	32 millions
DAIOS	1 404	> 1 000	1 404
PubChem	63 105 228	-	68 479 719
STOFF-IDENT	7 864	-	7 864
MassBank SM/SM	-	3 350	3 350
mzCloud	-	1 956	2 510
NIST IE-SM	-	212 961	242 477
NIST SM/SM	-	4 628	8 171
Wiley Registry of Mass Spectral Data (IE-SM)	-	289 000	638 000
Agilent Broecker, Herre & Pragst Toxicology/Forensics	8 998	3 497	8 998
Librairie Agilent Pesticide CL/Q-TOF SM/SM	1 664	~700	1 664
Librairie Agilent Pesticide CG/Q-TOF IE-SM	750	750	750
Librairie Agilent METLIN Synthetic Substance	64 092	10 000	64 092
Base de données en ligne Agilent METLIN Scripps	83 135	12 171	240 566
Librairie Agilent Veterinary Drug	1 684	770	1 684
Bruker ToxScreener (incl. Pesticide Screener)	-	704	1 753
Librairie Sciex / AB Sciex CL/SM/SM Meta	-	2 381	2 381
Thermo Environmental Food Safety (EFS) avec/sans temps de rétention (TR)	-	entre 90 et 454	732
Thermo toxicology	-	entre 36 et 618	654
Base de données Waters sans TR	-	730	730

Annexe 3

Prédiction des temps de rétention

Des modèles QSRR (Quantitative Structure-Retention relationship) peuvent être développés liant la nature des analytes et leur temps de rétention chromatographique. Par conséquent, le temps de rétention expérimentalement observé peut être exploité pour vérifier avec le modèle obtenu la pertinence de propositions de structure moléculaire pour un composé inconnu détecté. Le degré de précision atteint par ces modèles ne permet ainsi pas d'affirmer de l'identité d'une molécule mais il peut contribuer à éliminer certaines hypothèses si l'écart entre le temps prédit et observé est trop important.

Pour établir un modèle, le principe est de construire une base de données en injectant un nombre conséquent (au moins une centaine) de composés étalons ce qui permet d'obtenir des temps de rétention de composés de structures connues. Ces substances sont ensuite définies par des descripteurs afin d'aboutir à des modèles permettant de lier leur structure moléculaire aux temps de rétention.

Concernant la chromatographie en phase gazeuse, des modèles robustes sont déjà disponibles car des modèles simplifiés peuvent être utilisés. En effet, la phase mobile est généralement constituée d'hélium et, ainsi, seuls les analytes et la nature de la phase stationnaire constituent des variables. Un modèle appelé indice de Kovats est ainsi couramment employé. Il est particulièrement précis pour les hydrocarbures mais pour d'autres types de composés avec des groupements chimiques plus variés, d'autres modèles basés sur les ANN (réseau de neurones artificiels) sont préférés. Tous ces modèles permettent d'obtenir une précision de l'ordre de 1-3% (Heberger 2007).

La prédiction des temps de rétention en chromatographie en phase liquide est beaucoup plus complexe car de nombreuses interactions interviennent. Ainsi, des modèles généralement plus élaborés doivent être développés. Le modèle le plus simple et accessible est basé sur la corrélation des temps de rétention avec les log Kow (ou Log D) des composés étudiés (Bade et al. 2015). Il présente cependant des limites majeures car le log D ne permet pas de retranscrire toutes les interactions intervenant dans ce type de chromatographie. De plus, les valeurs déterminées pour le log kow (ou Log D) peuvent présenter de fortes imprécisions car elles découlent souvent elles-mêmes de prédiction à base de modèles.

Des modèles plus complexes peuvent également être mis en œuvre. Le plus employé de ces modèles est le LSER (Linear Solvation Energy Relationships) (Put & Heyden 2007). Ils comportent 5 descripteurs (la réfractivité molaire, le volume moléculaire, la dipolarité/polarisabilité du soluté, hydrogènes accepteur et donneur) qui offrent une meilleure définition des caractéristiques structurales des molécules étudiées. Bien qu'améliorant la précision des prédictions, ils présentent des limites particulièrement pour les composés polaires ou comportant des groupements ionisables.

Comme présenté lors du Workshop Norman en Septembre 2015, d'autres types de modèles sur des réseaux de neurones et/ou des algorithmes génétiques sont actuellement explorés par des groupes de recherche afin d'améliorer la précision des prédictions. Ces différents types de modèle ont été recensés dans une revue (Put & Heyden 2007).



Centre scientifique et technique
Direction des Laboratoires
3, avenue Claude-Guillemin
BP 36009 – 45060 Orléans Cedex 2 – France – Tél. : 02 38 64 34 34
www.brgm.fr