An integrated metabolomics and proteogenomics approach reveals molecular alterations following carbamazepine exposure in the male mussel Mytilus galloprovincialis

Thibaut Dumas, Frédérique Courant, Christine Almunia, Julien Boccard, David Rosain, Geoffroy Duporte, Jean Armengaud, Hélène Fenet, Elena Gomez

An integrated metabolomics and proteogenomics approach reveals molecular alterations following carbamazepine exposure in the male mussel *Mytilus galloprovincialis*

Thibaut Dumas [1], Frédérique Courant [1*], Christine Almunia [2], Julien Boccard [3,4], David Rosain [1], Geoffroy Duporté [1], Jean Armengaud [2], Hélène Fenet [1] and Elena Gomez [1]

[1] HydroSciences Montpellier, IRD, CNRS, University of Montpellier, Montpellier, France;
[2] Université Paris-Saclay, CEA, INRAE, Département Médicaments et Technologies pour la Santé (DMTS), SPI, 30200 Bagnols-sur-Cèze, France;
[3] School of Pharmaceutical Sciences, University of Geneva, Geneva 1211, Switzerland;
[4] Institute of Pharmaceutical Sciences of Western Switzerland, University of Geneva, Geneva 1211, Switzerland

* Corresponding author at: HydroSciences Montpellier, IRD, CNRS, University of Montpellier, 15 avenue Charles Flahault, 34093 Montpellier, France. E-mail address: frederique.courant@umontpellier.fr

Abstract

Carbamazepine is one of the most abundant pharmaceutical active compounds detected in aquatic systems. Based on laboratory exposures, carbamazepine has been proven to adversely affect aquatic organisms. However, the underlying molecular events remain poorly understood. This study aims to investigate the molecular mechanisms potentially associated with toxicological effects of carbamazepine on the mussel *Mytilus galloprovincialis* exposed for 3 days at realistic concentrations encountered in coastal environments (80 ng/L and 8 µg/L). An integrated metabolomics and proteogenomics approach, including data fusion strategy, was applied to gain more insight in molecular events and cellular processes triggered by carbamazepine exposure. Consistent metabolic and protein signatures revealed a metabolic rewiring and cellular stress at both concentrations (e.g. intensification of protein synthesis, transport and catabolism processes, disruption of lipid and amino acid metabolisms). These highlighted molecular signatures point to the induction of autophagy, closely related with carbamazepine mechanism of action, as well as a destabilization of the lysosomal membranes and an enzymatic overactivity of the peroxisomes. Induction of programmed cell death was highlighted by the modulation of apoptotic cognate proteins. The proposed integrative omics data analysis was shown to be highly relevant to identify the modulations of the two molecular levels, i.e. metabolites and proteins. Multi-omics approach is able to explain the resulting complex

biological system, and document stronger toxicological pieces of evidence on pharmaceutical active compounds at environmental concentrations in sentinel organisms.

Keywords: Multi-omics, data fusion, pharmaceutical active compounds, mechanism of action, Mediterranean mussel, marine organisms

1. Introduction

The antiepileptic carbamazepine (CBZ), known for its low removal in wastewater treatment plants (WWTPs) [1] and its low (bio)degradability in the environment [2,3], has been proposed as a possible anthropogenic marker of urban pollution in the aquatic systems [4,5]. CBZ is among the most frequent pharmaceutical active compounds (PhACs) detected in the environment and can be a threat for aquatic organisms [6–9]. Marine ecosystems constitute the terminal receptacle of contaminated continental water and are also submitted to direct discharges from WWTPs. Therefore, those ecosystems are not spared by the occurrence of PhACs, including CBZ. The latter is not easily adsorbed by sediments or solid particles and consequently remains mainly in the aqueous phase [10,11], at concentrations ranging from < 1 to 1410 ng/L in marine ecosystems [12,13]. CBZ is thus bioavailable to numerous marine organisms, especially filter feeders such as bivalve mollusks.It has already been detected in tissues of marine bivalves despite its low bioaccumulation potential [13–17].

The exposure of marine bivalves to CBZ may lead to adverse effects. Several experimental studies revealed individual and sub-individual effects at different life stages (adult, larval). The most common effect reported was the induction of oxidative stress at environmental concentrations [18,19]. CBZ also exerts various effects on marine bivalves such as cytotoxicity and immunotoxicity [20–24], genotoxicity [22,25] and embryotoxicity [26]. Nevertheless, little is known about its mechanisms of action on invertebrates. Some mechanisms known in humans were confirmed in marine bivalves, such as its interaction with the adenylyl cyclase system leading to the reduction of intracellular cAMP levels in

Mytilus galloprovincialis [23]. In addition, its potential agonist role of gamma-aminobutyric acid (GABA) receptors in the clam *Ruditapes philippinarum* [21], and its interaction with voltage-dependent Na^+ and K^+ channels in *Scrobicularia plana* [27] were demonstrated. However, these different types of molecular alterations are not sufficient to fully explain the adverse effects reported in the literature. Since differences in CBZ pharmacokinetics and pharmacodynamics are expected between humans and bivalves, it is necessary to apply non hypothesis-driven approaches to further investigate mechanisms of action and toxicological effects of CBZ on marine bivalves.

The combination of different omics approaches (i.e., multi-omics) has been shown as a relevant way to examine the complex interactions and regulation events between different biological layers (e.g. DNA, RNA, proteins, metabolites) [28,29]. Indeed, the major motivation for integrating data from different levels (e.g. transcriptome, proteome, metabolome) is to improve the understanding of the underlying biological system, and to gain further insight into mechanistic knowledge that is hardly accessible by measuring a single level of the biological system (e.g. metabolomics alone) [28]. According to previous ecotoxicological studies, an integrated metabolomics and proteomics approach is suitable to unravel mechanisms of action and toxicological effects of contaminants, without preconceived assumptions [30–32]. However, due to a certain degree of heterogeneity between these large-scale approaches and the intrinsic complexity of the generated datasets, the integrative analysis of different omics remains challenging [33]. For this purpose, different fusion strategies can be considered, depending on the structure and the nature of the associated data [34]. When observations constitute the common mode, i.e. the same samples are characterized by several blocks of variables (e.g. metabolites and proteins), high/low/mid/ level approaches can be implemented [35]. High-level data fusion involves the individual investigation of each data block, and outputs are then collected to offer a global picture of all data. One potential advantage is to provide increased prediction accuracy, e.g. using a majority vote, but it does not allow links between signals from the different sources to be directly considered. For that purpose, low-level data fusion constitutes an attractive alternative, because the combination of data is carried out at the signal level. As relationships between multiple

blocks can be objectively evaluated, multiblock modelling provides a very efficient approach to this aim. However, it may suffer from different sources of biological and technical noise, especially in the case of multi-omics integration. In that context, mid-level strategies offer an interesting compromise between these two extremes, by starting with an individual evaluation of the different blocks to extract meaningful information that will be combined in a second phase. The first step of mid-level data fusion can be carried out using either dimensionality reduction based on latent variable models (e.g. Principal Component Analysis) or variable selection methods. The latter aims to select the most informative subset of variables from each block, therefore potentially increasing prediction accuracy. Another way of looking at variable selection is to remove signals that are not relevant to describe the biological situation of interest. For this purpose, it is possible to implement uninformative variable elimination (UVE) strategies. Processed data blocks can then be used as input of a multiblock model. A mid-level data integration strategy combining Monte Carlo UVE based on partial least squares regression (MCUVE-PLS) [36] and consensus orthogonal partial least squares-discriminant analysis (consensus OPLS-DA) [37] was implemented in this work. By focusing on a relevant subset of variables in each block, it offers the advantage of improved prediction performance compared to the whole dataset, while allowing potential links between blocks to be efficiently considered. Moreover, an objective evaluation of the most relevant data block(s) can be gained using multiblock modelling [35].

Hence, an integrated metabolomics and proteogenomics approach, based on mid-level data fusion and multiblock modelling, can be a suitable strategy to gain understanding of biological systems by reinforcing the biochemical connectivity between the two omics levels (*i.e.* proteins and metabolites). Such a strategy was applied in this study to fill the information gap about the mechanisms of action and potential toxicological effects of CBZ on marine bivalve mollusks. For this purpose, a 3-day exposure at a low (80 ng/L) and a high (8 µg/L) dose on the Mediterranean mussel *M. galloprovincialis* was conducted in controlled laboratory conditions.

2. Materials & Methods

2.1. Chemicals

Pesticide analytical-grade solvents (methanol, dichloromethane and ethanol) and LC/MS grade solvents (water, acetonitrile, formic acid 99%) were obtained from Carlo Erba (Val de Reuil, France). Ultrapure water was generated by a Simplicity UV system from Millipore (Bedford, MA, USA) with a specific resistance of 18.2 MΩ.cm at 25°C. Analytical pure standards used for identification of endogenous metabolites were purchased from the following suppliers: Sigma-Aldrich (now part of Merck), Santa Cruz Biotechnology, Toronto Research Chemicals and LGC Standards. Analytical pure standard used for xenobiotic quantification concern carbamazepine (CBZ, Sigma-Aldrich) and the deuterated carbamazepine-d8 (CBZ-d8, Toronto Research Chemicals). All chemicals used in this study were analytical grade (purity > 95 %).

2.2. Animals, experimental design and sample collection

Mussels *Mytilus galloprovincialis* were purchased from Les Compagnons de Maguelone mussel culture (Villeneuve-les-Maguelone, France) in March 2019 and were brought back to the laboratory in a mesh bag at room temperature (14°C) within an hour. On arrival at the laboratory, mussels were immersed in seawater. Then, 90 mussels with a homogenous shell size (7.0 cm ± 0.4 cm) were rinsed with seawater while algae and barnacles were removed from the shell. Mussels were then randomly distributed in 18 glass aquaria (n = 5 mussels per aquarium), each containing 2 L of filtered seawater (provided by IFREMER, Palavas, France; filter GF/F Ø 100 µm).

During all the experiment (acclimatization and exposure periods), seawater was continuously aerated and renewed daily, and mussels were fed once daily with the alga *Tetraselmis suecica* (Greensea, Mèze, France) at constant density (10 000 cells/mL). Physicochemical parameters were checked on a daily basis such as temperature (14.4 °C ± 0.5), pH (7.5 ± 0.06), oxygen (10.3 mg/L ± 0.1) and salinity (36.8 ‰ ± 0.2), and were maintained constant throughout the experiment.

After a 7-day acclimatization period, three exposure conditions were constituted (n = 30 mussels *per* condition) with an environmentally relevant CBZ concentration (80 ng/L; low dose), a higher CBZ concentration (8 µg/L; high dose) and a solvent control exposure (10 µl/L of methanol, solvent used for the preparation of CBZ solution, corresponding to a final dilution of 0.01 ‰). Each exposure solution was prepared extemporaneously every day. Aquaria seawater was spiked one hour after feeding to prevent any adsorption of CBZ on the algae. Seawater (500 mL) from aquaria was sampled in triplicate daily before each seawater renewal (static renewal) to quantify carbamazepine concentration.

After three days of exposure, no dead mussels were recorded. Mussels were dissected on ice to collect the digestive gland (used for both metabolomics and proteomics analysis), the gills (for further analyses not in the scope of this study) and the remaining soft tissues (for CBZ quantifications) in different cryotubes. In the same time, sex microscopy determination allowed the identification of 17 females, 10 males and 3 undifferentiated in the solvent control exposure condition, 13 females, 12 males and 5 undifferentiated in the low dose exposure condition, and 14 females, 12 males and 4 undifferentiated in the high dose exposure condition. The samples were directly frozen in liquid nitrogen before being stored at -80°C. Metabolomics and proteomics analyses were performed on the same samples in order to reinforce the biochemical connectivity between the two omics levels (*i.e.* proteins and metabolites).

In the present study, only samples of male mussels were considered for the following analysis in order to avoid any bias due to the gender confounding factor. Beside, two samples from the solvent control exposure condition and two samples from the high dose exposure condition were excluded from analysis due to insufficient digestive gland tissue to perform both metabolomics and proteomics analysis (n = 8 males for the solvent control, n = 12 males for the low dose and n = 10 males for the high dose). Mussels individuals were considered independent (as the unit of replication) since no clustering related to glass aquarium was observed using Principal Analysis

Components (PCA) or hierarchical clustering analysis for both metabolomics and proteomics analyses (results not shown).

2.3. Extraction and analysis of carbamazepine

Materials and methods describing extraction and analysis of CBZ in seawater exposure and mussel soft tissues were given in Supplementary Data. Briefly, extraction of CBZ in seawater exposure was carried out with a solid phase extraction (SPE) method using SPE cartridges Oasis HLB 500 mg 6 cc (Waters, Wexford, Ireland). The extraction protocol of CBZ in mussel soft tissues was based on a *QuEChERS* (Quick, Easy, Cheap, Efficient, Rugged and Safe) method adapted from Martínez Bueno et al. [15]. The targeted analysis of CBZ was performed on a Vanquish HPLC (Thermo Fisher Scientific, Bremen, Germany) coupled to a Q Exactive Orbitrap high resolution mass spectrometer (Thermo Fisher Scientific, Bremen, Germany) equipped with a heated electrospray ionization source (HESI).

2.4. Metabolomics analysis

2.4.1. Sample preparation

Digestive glands of male mussels were lyophilized and grinded until a fine powder was obtained. As described in previous studies [38,39], the biphasic solvent system methanol/dichloromethane/water (16/16/13; *v/v/v*) was used to extract metabolites from tissues. Briefly, 30 mg of tissue (± 0.40 mg) were first homogenized and extracted in a glass tube with 240 µL of methanol and 75 µL of ultrapure water and vortexed for 1 min. In the next step, 240 µL of dichloromethane and 120 µL of ultrapure water were added before vortexing the whole 1 min. The samples were left on ice for 15 min and then centrifuged at 2000× *g* for 15 min at 4°C. Volumes of 50 µL of the supernatant (polar phase) were collected and evaporated to dryness under a nitrogen stream and finally reconstituted in 200 µL of water/acetonitrile (95/5; *v/v*). The extracts were filtered using a 500 µL centrifugal filter with a 10k modified polyethersulfone membrane (VWR, Fontenay-sous-Bois, France), and then transferred into vials prior to liquid chromatography—high resolution mass spectrometry (LC-HRMS) analysis.

2.4.2. *Data acquisition and quality control*

Non-targeted analysis was carried out using a Vanquish HPLC (ThermoFisher Scientific, Bremen, Germany) coupled to a Q Exactive Orbitrap HRMS (ThermoFisher Scientific, Bremen, Germany) equipped with a heated electrospray ionization source (HESI).

A reversed phase pentafluorophenylpropyl (PFPP) analytical column (100 × 2.1 mm; 3 µm particle size; Sigma Aldrich, PA, Bellefonte, USA) was used for HPLC separation. Each sample (10 µL) was loaded onto the column with a full loop injection. The mobile phase was constituted of water and acetonitrile both modified with 0.1 % formic acid. The flow rate was set at 250 µL/min and the gradient system was as follows (water/acetonitrile): 95/5 until 3 min, 60/40 at 8 min, 50/50 at 9 min, 30/70 at 13 min, 5/95 from 15 to 18 min, going back to the initial condition at 21 min (95/5) and equilibrating the column until the 28th min.

The Q Exactive HRMS was tuned to a mass resolution of 35,000 (FWHM, m/z 200) with a mass spectrum range of 50–750 m/z. Data were acquired simultaneously in both positive (ESI^+) and negative (ESI^-) ionization modes with the following settings: spray voltage at 3.35 |kV|, sheath gas flow rate of 55, aux gas flow rate at 10, S-Lens RF level of 50, capillary temperature at 300°C and heater temperature at 250°C. Sample order was randomly distributed in the injection sequence to avoid any bias and to reduce the influence of potential confounding factors.

In order to assess the analytical repeatability and sensitivity of acquisitions, a quality control (QC) sample was injected at regular intervals throughout the sequence (every six samples). The QC sample was prepared by pooling 40 µL of each injected sample extract. Then, the relative standard deviation (RSD) for each features detected in QC samples was calculated. A controlled analytical variability is set such that 70 % of features must have an RSD < 30 % [40].

2.4.3. *Data pre-processing*

The strategy for data pre-processing was already presented in Dumas et al. [39,41]. Briefly, the raw data were converted into mzXML files with MSConvert freeware (ProteoWizard 3.0) [42]. A multi-step strategy was applied for processing ESI^+ and ESI^- acquisitions separately using the XCMS

package [43] in the R environment. Optimized XCMS parameters were implemented: *m/z* interval for peak picking was set at 0.0025, the signal-to-noise ratio threshold was set at 10, the group bandwidth was set at 8, and the minimum fraction was set at 0.5. XCMS returned results as a peak table containing variable identity (i.e., *m/z* and retention time) and feature abundances (i.e. peak area). After visual inspection of each extracted ion chromatogram, all features corresponding to baseline drift or background noise were discarded from the peak table. In addition, the Bioconductor package CAMERA [44] was used to remove isotopes, adducts, and fragments from the peak table, thus avoiding information redundancy. Based on QC sample injections, features with a RSD > 30 % were excluded for statistical analysis.

2.4.4. Metabolites annotation and identification

The public databases Human Metabolome Database (HMDB; http://www.hmdb.ca/; [45]) and LipidMaps (https://lipidmaps.org/; [46]) were used for feature annotation. A mass precision was fixed at 0.002 Da. The levels of confidence for the annotation were defined according to the recommendation of the Compound Identification work group of the Metabolomics Society [47]: (i) level 1 was characterized by unambiguous identification based on the accurate mass and retention time of the corresponding analytical standard injected under the same analytical conditions (in-house database), (ii) level 2 corresponded to putative annotation based upon physicochemical properties and/or spectral similarity with public databases (e.g. HMDB and LipidMaps), and (iii) level 3 was putatively characterized compound classes (e.g. based upon characteristic physicochemical properties of a chemical class of compounds). Compound classes were attributed thanks to the LipidMaps Database of Computationally-generated Bulk Lipid Species [48].

2.4.5. Differential metabolite abundance analysis

Abundance of metabolites (peak area of MS spectra) was compared between the solvent control condition and the CBZ treated conditions. Differential abundance was assessed with a Welch *t*-test and the False Discovery Rate (FDR) was controlled using the Benjamini–Hochberg FDR

correction. Metabolites with an absolute significant modulation higher than 30 % (adjusted p-value < 0.05) were considered statistically relevant.

2.5. Proteogenomics analysis

2.5.1. Protein extraction

Each digestive gland, previously lyophilized, was analyzed individually by shotgun proteomics in standard conditions. For this, 7 mg from each lyophilisate was mechanically homogenized by bead-beating in 70 µl of LDS sample buffer (Invitrogen, Fisher Scientific, Illkirch, France) with a Precellys Evolution homogenizer (Bertin Technologies, Montigny-le-Bretonneux, France). The homogenates were centrifuged at 10,000× *g* for 3 min in order to pellet cellular debris, and the resulting supernatant collected to a new tube. Samples were homogenized with a mix of 0.1 mm and 0.5 mm glass beads and 0.1 mm silica beads as previously recommended [49]. The homogenization was carried out until the dissolution of the whole material was achieved with 5 homogenization cycles at 6800 rpm with a pause of 30 s between each. Then, samples were centrifuged at 16 000× *g* for 2 min. Samples were then incubated for 5 min at 99 °C and briefly centrifuged. A 20 µL aliquot of each supernatant was then subjected to SDS-PAGE for a short electrophoretic migration, as described previously [50]. The whole-protein content from each well was extracted as a single polyacrylamide band, processed as described [51], and submitted to proteolysis with trypsin (Roche) using 0.01 % ProteaseMAX surfactant (Promega).

2.5.2. Liquid Chromatography - Tandem Mass Spectrometry method

The peptide mixtures were analyzed in data-dependent mode with a Q Exactive HF mass spectrometer (ThermoFisher Scientific) coupled with an UltiMate 3000 LC system (Dionex-LC Packings). This analytical system was operated essentially as described [52]. Peptides were resolved onto a nanoscale C18 PepMapTM 100 capillary column (LC Packings) with a 90-min gradient of acetonitrile, 0.1% formic acid, at a flow rate of 0.2 µL/min. Following a Top20 method, peptides were analyzed with scan cycles initiated by a full scan of peptide ions in the Orbitrap analyzer, followed by

high-energy collisional dissociation and MS/MS scans on the 20 most abundant precursor ions. Full scan mass spectra were acquired from *m/z* 350 to 1800 at a resolution of 60,000. Ion selection for MS/MS fragmentation and measurement was performed applying a dynamic exclusion window of 10 s.

2.5.3. Peptide assignation and proteomics data analysis

MS/MS spectra were assigned to peptide sequences by the MASCOT Daemon 2.3.2 search engine (Matrix Science, London, UK) searching against the digestive gland specific RNAseq-derived database obtained by *de novo* assembly with Trinity version 2.4 [53] and followed by ORF search and transcript translation with Transdecoder as described in Cogne et al. [54]. The mRNA sequencing used the Illumina method with paired-end 150bp sequencing strategy from which 40 millions of reads were used to assemble the transcriptome. The final protein database contains 86,022 putative protein sequences totaling 29,583,007 amino acids. For MS/MS spectra assignment, the parameters were: full-trypsin specificity, maximum of one missed cleavage, mass tolerances of 5 ppm on the parent ion and 0.02 Da on the MS/MS, carboxyamidomethylated cysteine (+57.0215) as a fixed modification, and oxidized methionine (+15.9949) and deamidation of asparagine and glutamine (+0.9848) as variable modifications. All peptide matches presenting a MASCOT ion score with a p-value of less than 0.05, corresponding to a FDR of 1% as evaluated with the DecoyPyrat procedure [55], were filtered and assigned to a protein following a parsimony rule. Functional annotation of identified proteins was perform by sequence similarity search using Swissprot and NCBInr databases with the blast tool alignment DIAMOND as previously described [56]. Peptide-to-spectrum matches were counted for each protein leader of the protein group defined on the basis of their sequence similarities.

2.5.4. Differential protein abundance analysis

The abundances of all MS/MS-detected proteins were based on cumulated spectral counting of peptides used to identify proteins. For the differential comparison of protein abundances, spectral counts were normalized as described in Liu et al. [57]. Statistical evaluation of differential detection was

carried out using the Tfold method i.e. log2 (fold change), combined with a Student *t*-test for the p-value calculation. The fold change was calculated according to the following ratio $\frac{Spectral\ count\ proteins\ from\ treated\ samples}{Spectral\ count\ proteins\ from\ non-treated\ samples\ (CTRL)}$. Benjamini-Hochberg correction was applied on p-values as following: $adjusted\ p-value = \min\left(\frac{p\ X\ nbp}{j}, 1\right)$ using the number of proteins detected with at least two different peptides for each sample (nbp), the *t*-test p-value (p) and the protein rank according to the classification of increase p-value (j). A Tfold absolute value higher than 1 (corresponding to an up-modulation of +100 % or a down-modulation of -50 %) was considered as differentially expressed between the solvent control condition and the CBZ treated conditions.

2.5.5. Mass spectrometry and proteomics data

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [58] partner repository with the dataset identifier PXD024201 and project DOI 10.6019/PXD024201. [The reviewers may access this private dataset using reviewer_pxd024201@ebi.ac.uk as Username and mCyQvaJY as Password].

2.6. Metabolomics and proteogenomics data fusion

Unit variance was used as a pre-treatment to scale the variables of the different blocks. MCUVE-PLS [36] and consensus OPLS-DA [37] models were computed with combinations of toolboxes and in-house functions under the MATLAB® 8 environment (The MathWorks, Natick, USA). MCUVE-PLS was carried out using the libPLS 1.98 package [59] using an ensemble of 10^4 models with a ratio of calibration samples to the total samples of 0.7. A reliability index threshold of 2 was applied to remove variables considered as uninformative. Then, consensus OPLS-DA was implemented using the K-OPLS package [60] and leave-one-out cross-validation was performed to assess the optimal model size and predictive ability. More information on MCUVE-PLS and consensus OPLS-DA are available in Supplementary Data.

2.7. Bioinformatics analysis

Bioinformatics analysis was carried out to translate metabolic or proteomic signatures, i.e. lists of markers of interest, into meaningful functional information, and get deeper insights into the biological events modulated by CBZ exposure. When a biological pathway is modulated in a given condition, levels of the corresponding proteins and metabolites are expected to change accordingly. In that context, Gene Ontology (GO) terms classification, i.e. metadata stored in databases linking biological entities to a biological context, can help to highlight meaningful alterations. GO classification was done for modulated proteins as previously described [56]. Due to the lack of an annotated genome, we used the top five homologs in the Swissprot database to associate GO terms to each protein. In addition, to decipher metabolism pathways disruption by the CBZ, proteins were also annotated and classified according to the KEGG pathway and KEEG BRITE databases (https://www.kegg.jp/).

3. Results

In this study, we performed untargeted metabolomics and shotgun proteomics analyses on the digestive gland of male mussels *M. galloprovincialis* exposed 3 days at two CBZ environmental concentrations 80 ng/L (low dose, LD; n = 12) and 8 µg/L (high dose, HD; n = 10) and a solvent control (SC; n = 8). Measured concentrations of CBZ in seawater exposure did not differ from the nominal concentrations by more than 10 % (Supplementary Data). Uptake of CBZ by mussel was confirmed according the CBZ concentration quantified in their soft tissues (Supplementary Data). The metabolic fingerprints were constituted of 3203 signals in ESI^- and 5216 signals in ESI^+, 80 % of the signals having a good analytical repeatability in both ESI modes, across the QC replicates (RSD < 30 %). The shotgun mass spectrometry analysis provided 864,275 Peptides Spectra Matches with a p-value equal to 0.05 or lower, from the 1,891,263 spectra recorded, leading to a mean of 46 % of the interpreted spectra. A number of 6,880 proteins were identified, using a parsimony method classified into 5,954 groups including proteins with at least 50 % of homology.

3.1. Metabolomics and proteomics integrated analysis

Data integration was implemented using a mid-level data fusion strategy involving MCUVE-PLS and consensus OPLS-DA. Figure 1 provides a comprehensive description of the statistical strategy used with the remaining number of variables after each step. The metabolomic datasets acquired in ESI^+ (MetaboPOS) and ESI^- (MetaboNEG) as well as the proteomic dataset (Proteo) were defined as three separate data blocks, accounting for 1178, 1146 and 2422 variables, respectively. These variables were considered as acceptable according to analytical criteria (e.g. QC RSD < 30 %, exclusion of isotopes, exclusion of signals with a retention time out of the analytical acquisition time). First, the three data blocks were processed individually and MCUVE-PLS was carried out to remove uninformative variables. A reliability index cut-off of 2 was found appropriate to remove unwanted signals, while preserving biological information. Subsets of 365 variables, representing 31 % of the total from the corresponding block, 399 (35 %) and 937 (39 %) variables were obtained for the MetaboNEG, MetaboPOS and Proteo blocks, respectively.

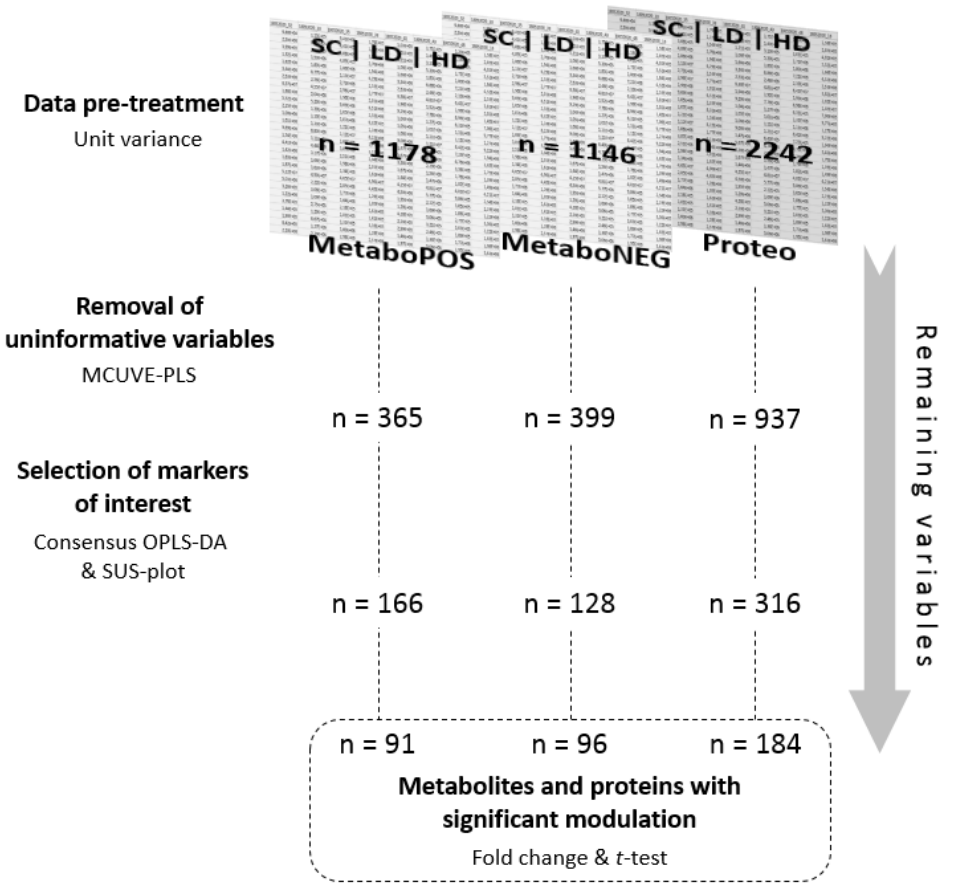


Figure 1: Strategy for selection of statistically relevant proteins and metabolites, and the remaining variables after each step for the respective blocks MetaboPOS (metabolomics dataset acquired in ESI+), MetaboNEG (metabolomics dataset acquired in ESI-) and Proteo (proteomics dataset).

Two consensus OPLS-DA models comparing the solvent control (SC) to each exposure condition separately, i.e. low dose 80 ng/L (SC*vs*LD) and high dose 8 µg/L (SC*vs*HD), were then evaluated. Both models were found optimal with two components (one predictive and one orthogonal) using leave-one-out cross-validation. The SC*vs*LD model was characterized by R2Y=0.990 and Q2Y=0.851 (Figure 2A), while the SC*vs*HD was associated with R2Y=0.985 and Q2Y=0.811 (Figure 2B), indicating that each model explained more than 98 % of the data and has a predictive performance higher than 80 %.

A Shared and Unique Structure (SUS) plot was then used to distinguish signals with common patterns of variations associated with both doses, from specific modulations related to a given exposure condition (Figure 2C). The coordinates of the signals on the SUS-plot were defined according to their loading provided by each consensus OPLS-DA model (abscissa axis: loading from the SC*vs*LD model; ordinate axis: loading from the SC*vs*HD model). Interestingly, the distribution of

the data points suggested an overall common modulation resulting from the two doses. Moreover, the lower density at the center of the plot was the consequence of the elimination of variables considered as non-informative by MCUVE-PLS. Thresholds were set to select the extreme data points of the SUS-plot associated with markers of interest. These variables were expected to be the most informative according to the models to explain the effects of CBZ at the low and high dose of exposure on the modulation of proteins and metabolites.

Hence, the selected signals considered as relevant included subsets of 128, 166 and 316 signals from the blocks MetaboNEG, MetaboPOS and Proteo, respectively. Among these selected signals, 45 were annotated from the block MetaboNEG, 37 from the MetaboPOS and 316 from the Proteo. The annotations are provided in Tables S1 and S2 for metabolites and proteins, respectively.

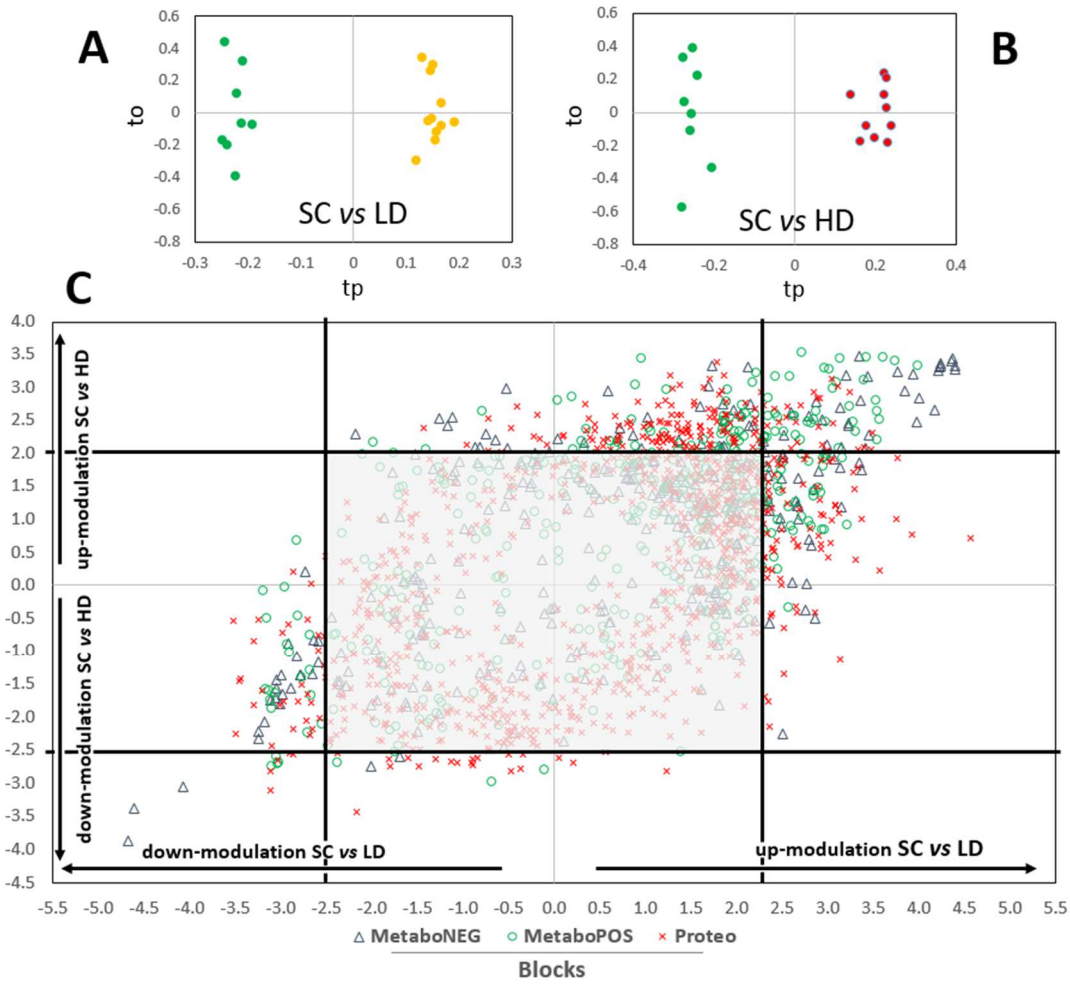


Figure 2: Score plots of the two consensus OPLS models with one orthogonal component (to) and one predictive component (tp): (**A**) solvent control exposure versus low dose exposure (SCvsLD) and (**B**) solvent control versus high dose (SCvsHD). (**C**) Shared and Unique Structure (SUS) plot integrating both comparisons SCvsLD and SCvsHD with common and specific modulations. Variables with a blue triangle belong to the MetaboNEG block (metabolomics dataset acquired in ESI-

), those with a green circle to the MetaboPOS block (metabolomics dataset acquired in ESI-) and those with a red cross to the Proteo block (proteomics dataset). The thresholds are represented with black lines and variables with a variation considered less relevant are shaded.

In order to highlight the significantly modulated proteins and metabolites, their modulation amplitude and p-value for both comparisons SC*vs*LD and SC*vs*HD were indicated in Tables S1 and S2. As expected, all the selected signals had not necessarily a significant modulation amplitude, especially for proteins. As no pre-selections of the metabolites and proteins was done on the basis of the modulation of their abundance, the selection of proteins and metabolites from the SUS-plot was not only based on a significant modulation, but mainly depended on correlation patterns between the variables. Compounds with a non-significant modulation amplitude, but localized at the extremes of the SUS-plot data points, may have a high correlation with other compounds, and the latter could have an important biological meaning. Therefore, selected (and annotated) signals with a modulation trend considered as relevant on the SUS-plot were selected for subsequent bioinformatics analyses.

Venn diagrams were built to highlight the number of metabolites or proteins with a common or specific modulation between the exposure conditions (Figure 3). Among the selected metabolites or proteins, less than half had a common modulation between the low and high dose of exposure. Hence, CBZ seemed to have partially similar effect at both doses but also some dose specificities. When looking at annotated signals, the common and specific patterns of modulation between the two doses were globally conserved, although less than one third of the metabolites could be annotated (Figure 3).

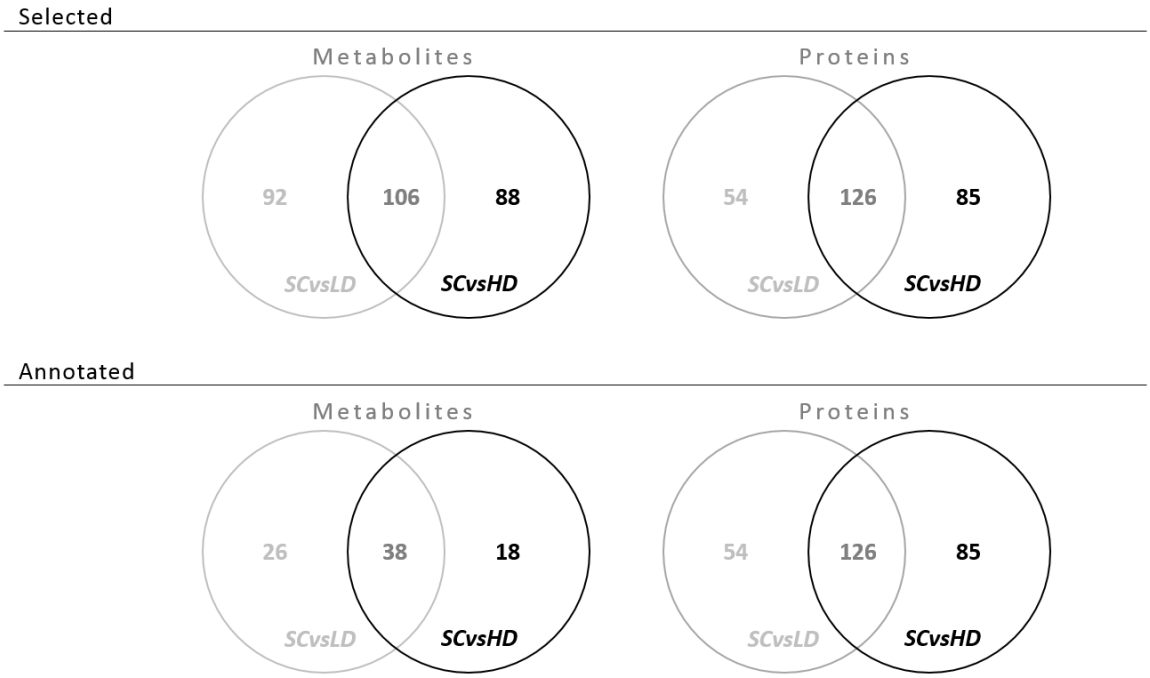


Figure 3: Venn diagrams highlighting specific and common variable modulations (metabolites and proteins) between both comparisons solvent control exposure versus low dose exposure (SCvsLD), and solvent control versus high dose (SCvsHD). Results are given for the selected signals according to the SUS-plot and then for annotated ones.

Bioinformatics tools were used to highlight biological processes and functions potentially altered by CBZ exposure from the selected subsets of proteins and metabolites. Finally, both proteins and metabolites were employed for a global analysis of metabolic pathways, in order to describe the relationships between both omics levels.

3.2. CBZ effects on proteome

Modulated proteins were queried against KEGG and assigned to pathways, themselves classified in three categories: genetic information processing, metabolism and cellular process (Figure 4). Not all the proteins queried against the KEGG database could be assigned to a KEGG ID, and the number of modulated proteins successfully associated to KEGG pathways was 97 for the low dose (73 % of the selected proteins) and 127 for the high dose (74 % of the selected proteins).

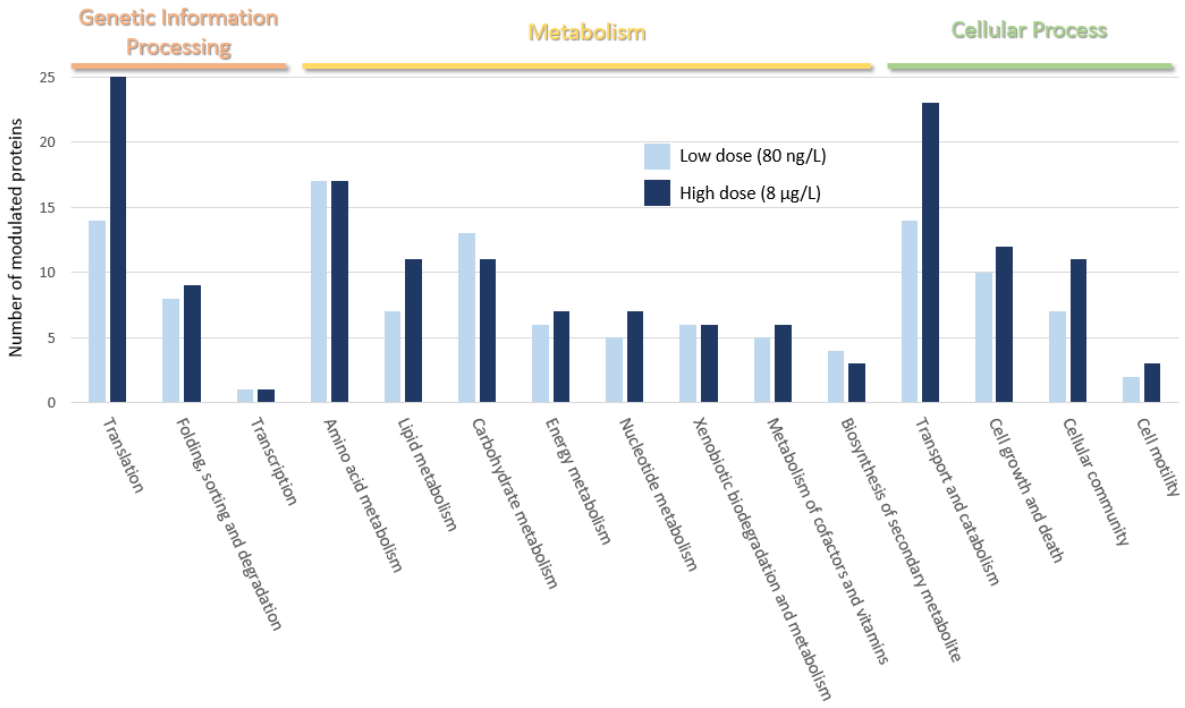


Figure 4: KEGG pathways in which the modulated proteins are involved at the low and high dose of carbamazepine.

The classification of proteins according to KEGG databases (Figure 4) highlighted cellular and metabolic processes largely impacted by proteins modulated at low or high doses. These processes were mainly related to translation process (twice the amount of protein modulated at the high dose), transport and catabolism (more pronounced at the high dose), metabolisms of amino acids and carbohydrates (equally between the two doses), degradation of lipids (more pronounced at the high dose) and cell growth and death (equally between the two doses).

An intensification of the translation process can be highlighted through the increasing number of modulated ribosomal proteins at the low dose (6 ribosomal proteins up-modulated and 6 down-modulated) and, to a greater extent at the high dose (19 ribosomal proteins mainly up-modulated) (Table S2). Increased numbers of ribosomal proteins lead to increases in the amount of RNA transcription and protein synthesis. The steps following translation, involving the folding and maturation of newly synthetized proteins, the sorting and the degradation of misfolded proteins, were also disturbed through the modulation of proteins within the endoplasmic reticulum (ER) at both doses (Figure S1). For example, the following proteins from ER were significantly up-modulated at the low and high dose, respectively: mannosyl-oligosaccharide 1,2-alpha-mannosidase IA (ERManI)

(+1340 % and +900 %), dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B-like (STT3) (+140 % and + 160 %), protein ERGIC-53 (+120 % and +120 %), while UDP-glucose:glycoprotein glucosyltransferase 1-like isoform X1 (HUGT) was significantly up-modulated only at the high dose (+70 %). Through an intensification of translation and ER activity, CBZ exposure induces protein synthesis and degradation of misfolded proteins. Accumulation of misfolded proteins can cause stress of ER. Significant up-modulations of two markers of ER stress were observed: bax inhibitor-1 protein (+1245 % at the low dose and +885 % at the high dose) and sarco/endoplasmic reticulum calcium ATPase (+200 % at the low dose and +160 % at the high dose).

Transport and catabolism processes appeared to intensify at the low dose and, to a greater extent, at the high dose (Figure 4). As described in Table S3, several proteins, mainly up-modulated, were involved in the internalization of compounds into the cell by endocytosis/phagocytosis. This also includes the degradation of compounds and cellular components mainly via the autophagic-lysosomal system, but also by phagolysosomes (fusion of lysosome with phagosome), and peroxisomes. Regarding the high number of down- and up-modulated proteins involved in the autophagic-lysosomal system (Table S3), CBZ triggered certain mechanisms inducing autophagy at the both doses. The most modulated protein was Bax inhibitor-1 protein, involved in the regulation of autophagy after ER stress. Another protein that may be related to autophagy induction is inositol monophosphatase 1, significantly up-modulated at the low and high doses (+328 % and +245 %, respectively).

Cell growth and death was the second most enriched cellular process at both doses (Figure 4). More specifically, modulated proteins belonged mainly to the cell death process that includes apoptotic processes. Induction of apoptosis was supported by several proteins that were significantly, or slightly, modulated at the high dose, such as TNF ligand-like 1 (+160 %; p-value = 0.042), TNF ligand superfamily member 10 (+40 %; p-value = 0.068), caspase 2 (-30 %; p-value = 0.067), cathepsin B (+80 %; p-value = 0.015) and mitogen-activated protein kinase kinase 1 (+470 %; p-value = 0.0015). At the low dose, the modulation of the quinone oxidoreductase PIG3-like isoform

X2 (-60 %; p-value = 0.016) from the p53 signaling pathway could be involved in the induction of apoptosis by generating oxidative stress. At the same dose, we also observed a tendency towards down-modulation of caspase 2 (-30 %; p-value = 0.059) and up-modulation of cathepsin B (+50 %; 0.081) as well as significant up-modulation of mitogen-activated protein kinase kinase (+400 %; p-value = 0.024), supporting the induction of programmed cell death even at an environmental concentration of CBZ.

Regarding metabolic pathways, most of the modulated proteins matched with the amino acid metabolism (17 proteins) and carbohydrate metabolism (13 proteins), while no marked difference was observed between the both doses (Figure 4). However, no particular amino acid metabolism or carbohydrate metabolism was significantly affected by CBZ, as modulated proteins were scattered in different metabolic pathways. Lipid metabolism was affected by CBZ exposure mainly at the high dose and to a lesser extent at the low dose (Figure 4). Several peroxisomal enzymes involved in fatty acid β-oxidation were up-modulated as illustrated in Table S3. Acyl-CoA oxidase, one of the key enzymes involved in the β-oxidation process, was significantly up-modulated (+230 %). Exposure to CBZ increased the fatty acid degradation that takes place in peroxisome and affected the lipid homeostasis.

3.3. CBZ effects on metabolome related to proteogenomics data

Among the 294 signals selected from the blocks MetaboNEG and MetaboPOS, 82 were annotated and reported in Table S1. Many signals were annotated at level 2, some of them with several potential annotations, while 7 metabolites were identified at level 1. A total of 41 annotated metabolites were significantly modulated (abundance modulation > 30 % and p-value < 0.05) at the low dose (SC*vs*LD) (29 up-modulated and 12 down-modulated), while 33 annotated metabolites were significantly modulated at the high dose (SC*vs*HD) (28 up-modulated and 5 down-modulated). The specific metabolic pathway or class compound assigned to each metabolite is provided in Table S1. Most of these metabolites belong to the amino acid metabolism, lipid metabolism and purine

metabolism. What is interesting here, is to investigate the consistent modulation of metabolites in accordance with protein modulations. As detailed in Table S4, several modulated proteins and metabolites shared common metabolic pathways. When direct connections were investigated between modulated proteins and metabolites, consistently affected pathways were observed for purine metabolism, phenylalanine metabolism in relation to tyrosine, and lipid degradation. Detailed metabolic pathways for purines, phenylalanine and tyrosine are illustrated in Figure 5.

At the low dose, modulation of phenylalanine (+64 %) was consistent with the up-modulation of two enzymes, namely aspartate aminotransferase (+70 %) involved in the conversion of phenylpyruvate to phenylalanine, and phenylalanine-4-hydroxylase (+110 %) which converts phenylalanine to tyrosine which was not highlighted as modulated. Even if tyrosine, precursor of catecholamines neurotransmitters (e.g., dopamine, metanephrine, norepinephrine, etc.), was not modulated, the dopamine metabolites 3,4-dihydroxyphenylacetaldehyde (-61 %) and the phenylalanine modulation indicates that the low dose of CBZ has modified the catecholamine metabolism. At the high dose, phenylalanine up-modulation was maintained, as well as the up-modulation of aspartate aminotransferase directly linked to this amino acid. Furthermore, a modulation trend was observed for monoamine oxidase (+210 %; p-value = 0.198; high biological variability) which catalyses the oxidative deamination of amines (i.e., dopamine, norepinephrine, epinephrine, serotonine) although amine compounds were not detected or modulated at 3 days of exposure.

Disturbances of purine metabolism were contrasted between the two CBZ doses (Figure 5). While purine-nucleoside phosphorylase (+180 %) and xanthine dehydrogenase/oxidase (+100 %) were up-modulated at the low dose, no related metabolite was affected. However, adenosine (+57 %; p-value = 0.052) and xanthine (+63 %) were up-modulated at the high dose consistently with an up-modulation of purine-nucleoside phosphorylase (+190 %) and 5'-nucleotidase (+130 %). An up-modulation of xanthine could be related to an increased nucleotides catabolism since it is the end product just before its conversion to uric acid.

A large part of metabolites, modulated either at the low dose, the high dose or both, belonged to the lipid metabolism. As reported in Table S1, the main compound classes of modulated lipids were fatty acids, branched fatty acids, fatty esters, acyl carnitines and acyl glycines. A disturbance of lipid metabolism is consistent with the proteogenomics data. For example, long-chain specific acyl-CoA dehydrogenase, actively involved in fatty acid degradation in mitochondria, was up-modulated at the low dose (+60 %). At the high dose, three different enzymes that complete the fatty acid degradation in mitochondria and peroxisome were up-modulated (acyl-coenzyme A oxidase 1; long-chain-fatty-acid--CoA ligase 4-like; peroxisomal bifunctional enzyme-like). In addition, CBZ affected the arachidonic acid metabolism by the down-modulation of the two enzymes prostaglandin-H2 D-isomerase/glutathione transferase (-60 %, low dose) and carbonyl reductase 1 (-40 %, low dose; -70 % high dose) as well as a prostaglandin and a thromboxane metabolites (Table S4). Arachidonic acid being the precursor of the eicosanoid family (prostaglandins, thromboxanes, and leukotrienes), a disturbance of its metabolism could have physiological consequences regarding the important role of eiconaoid family in organisms.

Finally, several annotated metabolites corresponding to dipeptides (e.g. gamma-l-glutamyl-l-leucine) were up-modulated at the low and/or high doses (Table S1). This is consistent with the increase in proteins synthesis and degradation as dipeptides are catabolic products of proteins.

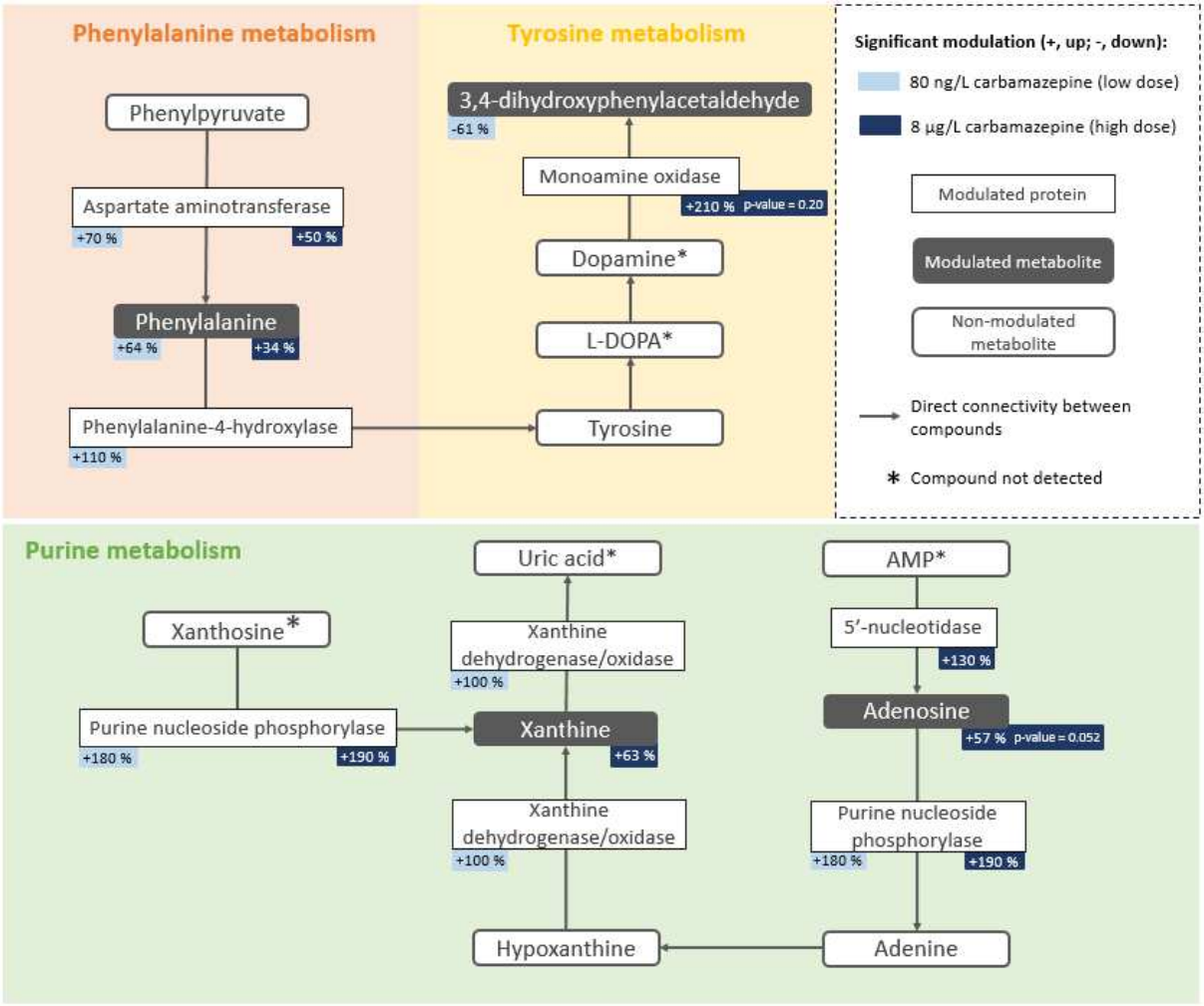


Figure 5: Metabolic pathways of phenylalanine, tyrosine and purine metabolisms represented by consistent changes in the levels of metabolites and proteins affected by the CBZ exposure at the low (80 ng/L) and high (8 µg/L) dose in the digestive gland of male mussel *Mytilus galloprovincialis*.

4. Discussion

The strategy of metabolomics and proteomics data fusion combining MCUVE-PLS and consensus OPLS-DA, aimed to highlight metabolic and protein correlated signatures, and then to study the possible relationships between these entities in response to the exposure of male *M. galloprovincialis* to a low and high dose of CBZ. The two consensus OPLS-DA models, each comparing the SC condition with either the low dose condition (SC*vs*LD) or the high dose condition (SC*vs*HD), were found to be robust based on cross-validation. The combination of these two models was carried out using a SUS-plot and allowed us to extract the most informative signals, related to common modulation patterns shared by the two doses or specific to each of them. The links between proteins and metabolites were then investigated using tools for visualizing metabolic pathways.

The interpretation of metabolomics data in combination with proteomics data is sometimes difficult when it comes to explain changes from one omics level to another [62]. The coordination between the modulation of a metabolite and the modulation of a protein can be constrained by feedback phenomena as well as by the time scale between them, or cannot be of the same amplitude and thus not be detected. Indeed, the concentration of a metabolite can modulate the expression of a gene, and changes in the expression of a gene can lead to an increase or decrease in the concentration of a regulatory enzyme or protein, subsequently affecting the concentration of the metabolite [62]. Based on the simultaneous measurement of metabolite and transcript or protein concentrations, different studies have reached contradictory conclusions on the coordination of the observed changes between the different levels [62]. Some studies have claimed that transcripts and metabolites are substantially co-regulated [63–65], providing examples of consistent correlations between biosynthetic enzymes and their products [63]. In contrast, other studies have shown that transcript and metabolite profiles tend to behave differently [66], and some have argued that correlational approaches are not specific enough to draw conclusions about functionally related genes and metabolites [67,68]. As our observations suggest, biochemical mechanisms probably lie between these two conclusions. On the one hand, proteins and metabolites were co-regulated (e.g.

555 peroxisomal and mitochondrial enzymes and fatty acids, phenylalanine and phenylalanine-4-
556 hydroxylase, adenosine and purine nucleoside phosphorylase-like isoform X1, etc.). On the other
557 hand, changes in other protein and metabolite abundances seemed uncoordinated or isolated. In
558 addition to the feedback phenomena and time scales that may explain the latter cases, it is also
559 highly likely that partial access to the metabolome and proteome, as well as the lack of knowledge
560 about protein-metabolite and protein-protein interactions may explain these observations. However,
561 despite the lack of data related to bivalve mollusks, our integrative analysis of metabolomics and
562 proteomics data revealed promising results and consistent relationships between the two datasets in
563 response of the male mussel *M. galloprovincialis* to CBZ exposure.

564 Globally, the presented results indicated a metabolic and cellular stress in the digestive gland
565 of male mussels induced by CBZ at both concentrations (80 ng/L and 8 µg/L). As discussed hereafter,
566 this stress was revealed by different marker compounds of the lysosomal-autophagic system,
567 lysosomal membrane destabilization, overactivity of peroxisomal and mitochondrial enzymes,
568 apoptosis, or compounds involved in proteins synthesis and energy metabolism.

569 Autophagy induction following CBZ exposure was demonstrated in this study by numerous
570 modulated proteins. Autophagy clears aggregated and misfolded proteins as well as damaged
571 organelles via cytosolic sequestration and subsequent lysosomal degradation. One of the known
572 mechanisms of action of CBZ in mammals is the induction of autophagy by inhibition of inositol
573 monophosphatase (IMPase), leading to free inositol depletion and reduced myo-inositol-1,4,5-
574 triphosphate (IP3) [69]. Although autophagy and its induction is not well understood in bivalve
575 mollusks [70], autophagy is a mechanism widely conserved in eukaryotes. Our findings showed a
576 significant modulation of IMPase at both exposure doses. It means that CBZ can act on the
577 intracellular inositol levels with consequences for the autophagy regulation, although inositol or IP3
578 were not detected by the metabolomics analysis. Another hypothesis can be put forward regarding
579 the induction of autophagy. Endoplasmic reticulum (ER) stress (e.g., accumulation of misfolded
580 proteins, change in Ca^{2+} level) is known as a particularly efficient stimulus for autophagy [71]. This study

581 demonstrated an ER stress through the up-modulation of protein synthesis and the potential
582 increase in misfolded proteins. ER stress was also highlighted by a high up-modulation of bax
583 inhibitor-1 and sarco/endoplasmic reticulum calcium ATPase at both CBZ doses. Sarco/endoplasmic
584 reticulum calcium ATPase is a Ca^{2+} transport protein of the ER membrane maintaining the low
585 cytosolic calcium level that enable a vast array of signaling pathways and physiological processes [72].
586 Bax inhibitor-1 is an ER-located protein, with anti-apoptotic functions, involved in the suppression of
587 intrinsic cell death mediated by ER stress or ER calcium release (see references in [73]). Evidences were
588 provided on the dual role of Bax inhibitor-1 to regulate both autophagy and apoptosis, two closely
589 related homeostatic processes [73]. Further research is then needed to elucidate the mechanism of
590 action of CBZ in bivalve mollusks, triggering autophagy either via depletion of inositol or by the
591 induction of ER stress.

592 In response to exposure in bivalves, lysosomes contribute to the degradation and
593 detoxification of contaminants through different enzymes [74]. Among these enzymes, cathepsins are
594 often involved in the detoxification of metals or organic contaminants [22]. In our study, the exposure
595 at the low and high doses induced an up-modulation of cathepsin B and a cathepsin-like protein-4,
596 and a down-modulation of cathepsin L. An over-expression of the gene coding for cathepsin L in the
597 digestive glands of *M. galloprovincialis* mussels, exposed for 3 days to 6.3 µg/L CBZ (close to the
598 present study exposure conditions) was reported in a recent study [22]. Exposure of the marine bivalve
599 *Ruditapes philippinarum* to 1 µg/L CBZ for 28 days also resulted in over-expression of the same
600 gene [21]. Changes in cathepsin abundance may reflect a general up-regulation of lysosomal proteolysis
601 due to either damage to cellular proteins or direct effects on protease activity due to CBZ
602 sequestration in lysosomes. However, an elevated lysosomal activity or sequestration of
603 contaminants in lysosomes can affect the lysosomal integrity (e.g. membrane permeability, size of
604 lysosome) [74,75]. In marine bivalves, β-glucuronidase is even used as lysosomal enzymatic markers of
605 changes in membrane permeability and lysosome size [74,76]. In our study, significant up-modulation of
606 β-glucuronidase was observed at the low (+210 %; p-value = 0.004) and high dose (+130 %; p-value =

607 0.034), although it was not selected by MCUVE-PLS. A modification of the membrane permeability of
608 the lysosomes or their morphology is possible at both exposure doses. At comparable doses of CBZ
609 (0.1 µg/L and 10 µg/L) but a longer exposure time (7 days), Martin-Diaz et al. [23] measured a dose-
610 dependent decrease in the stability of the lysosome membranes of haemocytes, collected from the
611 digestive gland of *M. galloprovincialis*. As a consequence, hydrolytic lysosomal enzymes can be
612 released in the cytosol causing severe cellular damages and even cell death [77].

613 Peroxisomes and mitochondria play a very important role in the phenomena of β-oxidation
614 of fatty acids to maintain lipid homeostasis in eukaryotic cells [78]. These enzymatic reactions produce
615 reactive oxygen species (ROS) such as hydrogen peroxide (H_2O_2), which is then degraded by
616 catalase [79]. In our study, an up-modulation of the peroxisomal enzymes responsible for the β-
617 oxidation of saturated and unsaturated fatty acids (e.g. acyl-CoA oxidase; long-chain acyl-CoA
618 synthetase; peroxisomal multifunctional enzyme type 2 isoform X2) was observed, mainly at the high
619 CBZ dose, as well as a slight up-modulation of catalase. In conjunction with this, the metabolomics
620 data also showed a disturbance in the metabolism of lipids and especially fatty acids. It also includes
621 a disturbance of arachidonic acid metabolism, precursor of eicosanoids, playing an important role in
622 bivalves reproduction and others physiological processes [80]. In response to a number of endogenous
623 compounds and xenobiotics, peroxisomes can undergo massive proliferation. This phenomenon is
624 usually accompanied by the induction of certain peroxisomal enzyme activities, particularly those of
625 the fatty acid β-oxidation system, while catalase may not be induced or weakly induced [79,81].
626 Consequently, peroxisome proliferation is considered as a potential source of oxidative stress for
627 cells, since ROS-producing enzymes are induced to a greater extent than the catalase capable of
628 detoxifying ROS [82,83], which could be the case in our study. Although the induction of peroxisomal
629 proliferation has already been observed in the cells of the digestive gland of *M. galloprovincialis* in
630 response to contaminants [84], to our knowledge this phenomenon has not been described in
631 invertebrates exposed to CBZ. More targeted measurements on the numerical and volume density of
632 peroxisomes and on the enzymatic activity of acyl-CoA oxidase would make it possible to verify this

hypothesis [84]. Peroxisome (or even peroxisome proliferation) dysfunction can cause an unbalanced ROS homeostasis leading to cellular damages and cell death [78].

In this study, organelles (e.g. lysosome, peroxysome and endoplasmic reticulum) stress and/or dysfunction triggered by CBZ exposure, as well as possible oxidative stress, can induce apoptosis [85]. This hypothesis is supported by several proteins involved in apoptotic processes, modulated at the low or high dose, such as TNF ligand-like 1, TNF ligand superfamily member 10, caspase 2, cathepsin B, apoptosis-inducing factor 3 isoform X1, mitogen-activated protein kinase kinase 1, and quinone oxidoreductase PIG3-like isoform X2 [86–91]. Although apoptosis was not directely measured in mollusks exposed to CBZ, Mezzelani et al. [92] observed several up-regulated genes with a role in apoptotic processes after *M. galloprovincialis* exposure to 1 µg/L CBZ (28 days). Furthermore, this programmed cell death was observed in the Chinese rare minnows *Gobiocypris rarus* exposed 28 days at 1, 10 and 100 µg/L CBZ [93]. After a 3-days exposure, we were probably at an early stage of the apoptosis signaling pathway. Prolonged exposure to CBZ would eventually lead to programmed cell death, even at environmentally relevant concentrations.

The general metabolic and cellular stress in digestive gland of male mussels triggered by CBZ exposure can also be demonstrated by an increase in protein turnover. This turnover was marked by an increase in both protein synthesis (e.g., increasing numbers of ribosomal proteins and proteins from endoplasmic reticulum) and catabolism (e.g., increasing numbers of dipeptides, catabolic products of proteins). Catabolic processes are thought to remove damaged proteins while the synthesis of new and protective proteins is vital for the preservation of cellular homeostasis during CBZ-induced stress. However, faced with environmental stressors or pollution, mussels generally adopt a strategy of depressing protein synthesis to protect themselves [94]. Organisms limit protein synthesis by storing inactive ribosomes that are rapidly reactivated when conditions improve. In the present study, it would be interesting to check whether ribosomal proteins are active in the translation process or wether they are stored for inactive ribosomes assembly. In any case, both strategies reveal a general stress induced by CBZ exposure at environmental concentrations.

659 The integrated analysis of proteomics and metabolomics data provided a detailed view of the
660 cellular and metabolic processes modulated in response to CBZ, indicating a general state of stress in
661 the cells of the digestive gland of male *M. galloprovincialis*. Our findings provide preliminary results
662 on CBZ mechanism of action related to autophagy induction in bivalve mollusks, either by inositol
663 depletion or ER stress. Furthermore, as the induction of oxidative stress is the most frequently
664 reported effect of CBZ in non-target organisms, this study could provide more details on the causes
665 of this stress. Indeed, oxidative stress can be triggered following a possible destabilization of the
666 lysosomal membranes, already demonstrated by other studies, but also through an enzymatic
667 overactivity of the peroxisomes (or even peroxisome proliferation), of which our results provide the
668 first evidence in *M. galloprovincialis*. As supported by others studies, ROS production and cellular
669 damages could lead to programmed cell death. Our results show several modulated proteins
670 involved in apoptotic processes. These cytotoxic effects generated by CBZ could also explain the
671 decrease in cell viability observed in studies conducted on mussel haemocytes. As example, decrease
672 in cell viability was measured in *M. galloprovincialis* haemocytes treated for 1 hour at different CBZ
673 concentrations (0.01; 0.1; 1; 10 and 100 µg/L) [24] or on haemocytes of the freshwater mussel
674 *Dreissena polymorpha* from 96h of treatment at the lowest CBZ dose tested 1 µg/L [97]. Induction of
675 cellular death or a decrease in cell viability would result in a weakening of the mussel immune system
676 [24,97]. The general stress triggered by CBZ was initiated as early as 3 days of exposure to an
677 environmental concentration (80 ng/L) but more intensively at the higher dose (8 µg/L). It would be
678 interesting to check whether prolonged exposure to the low dose would induce the level of stress
679 reached at the high dose, or whether organisms can adapt to prolonged exposure by developing
680 strategies to mitigate the toxicity. Further research is also needed to decipher the molecular intiating
681 events leading to the overall metabolic and cellular stresses observed in this study.

682

683 **Declaration of competing interest**

684 We declare that there is no conflict of interest for this paper.

Acknowledgement & Funding

This research was funded by the Agence Nationale de la Recherche (IMAP ANR-16-CE34-0006-01). The doctoral fellowship of Thibaut Dumas was financially supported by the Agence Nationale de la Recherche (IMAP ANR-16-CE34-0006-01). The authors thank the PONTEM (Platform Of Non-Target Environmental Metabolomics) platform from the consortium facilities MAMMA (Montpellier Alliance for Metabolomics and Metabolism Analysis). Technical assistance of Céline Roques and Bénédicte Marion (Institut des Biomolécules Max Mousseron, Montpellier, France) and Guylaine Miotello (Atomic Energy and Alternative Energies Commission, CEA, Bagnols-sur-Cèze, France) is gratefully acknowledged.

Appendix A. Supplementary Data

The following are available online at [**XXXX**], Figure S1: Folding, sorting and degradation processes disrupted by exposure to low and high doses of carbamazepine (KEGG pathway). Table S1: Annotation and identification of relevant metabolites selected according to the SUS-plot, modulated by CBZ exposure in male mussels either at the low dose, high dose, or both. Table S2: Annotation of relevant proteins selected according to the SUS-plot, modulated by CBZ exposure in male mussels either at the low dose, high dose, or both. Table S3: Relevant proteins selected by the models, involved in the transport and catabolism cellular process. Table S4: Modulated metabolites and proteins involved in common metabolic pathways.

References

(1) Margot, J.; Rossi, L.; Barry, D. A.; Holliger, C. A Review of the Fate of Micropollutants in Wastewater Treatment Plants. *WIREs Water* **2015**, *2* (5), 457–487. https://doi.org/10.1002/wat2.1090.

(2) Benotti, M. J.; Brownawell, B. J. Microbial Degradation of Pharmaceuticals in Estuarine and Coastal Seawater. *Environ. Pollut.* **2009**, *157* (3), 994–1002. https://doi.org/10.1016/j.envpol.2008.10.009.

(3) Calisto, V.; Domingues, M. R. M.; Erny, G. L.; Esteves, V. I. Direct Photodegradation of Carbamazepine Followed by Micellar Electrokinetic Chromatography and Mass Spectrometry. *Water Res.* **2011**, *45* (3), 1095–1104. https://doi.org/10.1016/j.watres.2010.10.037.
(4) Clara, M.; Strenn, B.; Kreuzinger, N. Carbamazepine as a Possible Anthropogenic Marker in the Aquatic Environment: Investigations on the Behaviour of Carbamazepine in Wastewater Treatment and during Groundwater Infiltration. *Water Res.* **2004**, *38* (4), 947–954. https://doi.org/10.1016/j.watres.2003.10.058.
(5) Hai, F. I.; Yang, S.; Asif, M. B.; Sencadas, V.; Shawkat, S.; Sanderson-Smith, M.; Gorman, J.; Xu, Z.-Q.; Yamamoto, K. Carbamazepine as a Possible Anthropogenic Marker in Water: Occurrences, Toxicological Effects, Regulations and Removal by Wastewater Treatment Technologies. *Water* **2018**, *10* (2), 107. https://doi.org/10.3390/w10020107.
(6) Fabbri, E.; Franzellitti, S. Human Pharmaceuticals in the Marine Environment: Focus on Exposure and Biological Effects in Animal Species. *Environ. Toxicol. Chem.* **2016**, *35* (4), 799–812. https://doi.org/10.1002/etc.3131.
(7) Mezzelani, M.; Gorbi, S.; Regoli, F. Pharmaceuticals in the Aquatic Environments: Evidence of Emerged Threat and Future Challenges for Marine Organisms. *Mar. Environ. Res.* **2018**, *140*, 41–60. https://doi.org/10.1016/j.marenvres.2018.05.001.
(8) Prichard, E.; Granek, E. F. Effects of Pharmaceuticals and Personal Care Products on Marine Organisms: From Single-Species Studies to an Ecosystem-Based Approach. *Environ. Sci. Pollut. Res. Int.* **2016**, *23* (22), 22365–22384. https://doi.org/10.1007/s11356-016-7282-0.
(9) Ebele, A. J.; Abou-Elwafa Abdallah, M.; Harrad, S. Pharmaceuticals and Personal Care Products (PPCPs) in the Freshwater Aquatic Environment. *Emerg. Contam.* **2017**, *3* (1), 1–16. https://doi.org/10.1016/j.emcon.2016.12.004.
(10) Ternes, T. A.; Herrmann, N.; Bonerz, M.; Knacker, T.; Siegrist, H.; Joss, A. A Rapid Method to Measure the Solid–Water Distribution Coefficient (Kd) for Pharmaceuticals and Musk Fragrances in Sewage Sludge. *Water Res.* **2004**, *38* (19), 4075–4084. https://doi.org/10.1016/j.watres.2004.07.015.
(11) Zhang, Y.; Geißen, S.-U.; Gal, C. Carbamazepine and Diclofenac: Removal in Wastewater Treatment Plants and Occurrence in Water Bodies. *Chemosphere* **2008**, *73* (8), 1151–1161. https://doi.org/10.1016/j.chemosphere.2008.07.086.
(12) Arpin-Pont, L.; Bueno, M. J. M.; Gomez, E.; Fenet, H. Occurrence of PPCPs in the Marine Environment: A Review. *Environ. Sci. Pollut. Res.* **2016**, *23* (6), 4978–4991. https://doi.org/10.1007/s11356-014-3617-x.
(13) McEneff, G.; Barron, L.; Kelleher, B.; Paull, B.; Quinn, B. A Year-Long Study of the Spatial Occurrence and Relative Distribution of Pharmaceutical Residues in Sewage Effluent, Receiving Marine Waters and Marine Bivalves. *Sci. Total Environ.* **2014**, *476*, 317–326. https://doi.org/10.1016/j.scitotenv.2013.12.123.
(14) Alvarez-Muñoz, D.; Huerta, B.; Fernandez-Tejedor, M.; Rodríguez-Mozaz, S.; Barceló, D. Multi-Residue Method for the Analysis of Pharmaceuticals and Some of Their Metabolites in Bivalves. *Talanta* **2015**, *136*, 174–182. https://doi.org/10.1016/j.talanta.2014.12.035.
(15) Martínez Bueno, M. J.; Boillot, C.; Fenet, H.; Chiron, S.; Casellas, C.; Gómez, E. Fast and Easy Extraction Combined with High Resolution-Mass Spectrometry for Residue Analysis of Two Anticonvulsants and Their Transformation Products in Marine Mussels. *J. Chromatogr. A* **2013**, *1305*, 27–34. https://doi.org/10.1016/j.chroma.2013.06.071.
(16) Wille, K.; Kiebooms, J. A. L.; Claessens, M.; Rappé, K.; Vanden Bussche, J.; Noppe, H.; Van Praet, N.; De Wulf, E.; Van Caeter, P.; Janssen, C. R.; De Brabander, H. F.; Vanhaecke, L. Development of Analytical Strategies Using U-HPLC-MS/MS and LC-ToF-MS for the Quantification of Micropollutants in Marine Organisms. *Anal. Bioanal. Chem.* **2011**, *400* (5), 1459–1472. https://doi.org/10.1007/s00216-011-4878-6.
(17) Boillot, C.; Martinez Bueno, M. J.; Munaron, D.; Le Dreau, M.; Mathieu, O.; David, A.; Fenet, H.; Casellas, C.; Gomez, E. In Vivo Exposure of Marine Mussels to Carbamazepine and 10-

Hydroxy-10,11-Dihydro-Carbamazepine: Bioconcentration and Metabolization. *Sci. Total Environ.* **2015**, *532*, 564–570. https://doi.org/10.1016/j.scitotenv.2015.05.067.

(18) Almeida, A.; Calisto, V.; Esteves, V. I.; Schneider, R. J.; Soares, A. M. V. M.; Figueira, E.; Freitas, R. Presence of the Pharmaceutical Drug Carbamazepine in Coastal Systems: Effects on Bivalves. *Aquat. Toxicol. Amst. Neth.* **2014**, *156*, 74–87. https://doi.org/10.1016/j.aquatox.2014.08.002.

(19) Almeida, Â.; Esteves, V. I.; Soares, A. M. V. M.; Freitas, R. Effects of Carbamazepine in Bivalves: A Review; Reviews of Environmental Contamination and Toxicology; Springer International Publishing: Cham, 2020; pp 1–19. https://doi.org/10.1007/398_2020_51.

(20) Aguirre-Martínez, G. V.; Buratti, S.; Fabbr, E.; DelValls, A. T.; Martín-Díaz, M. L. Using Lysosomal Membrane Stability of Haemocytes in Ruditapes Philippinarum as a Biomarker of Cellular Stress to Assess Contamination by Caffeine, Ibuprofen, Carbamazepine and Novobiocin. *J. Environ. Sci. China* **2013**, *25* (7), 1408–1418. https://doi.org/10.1016/s1001-0742(12)60207-1.

(21) Almeida, Â.; Freitas, R.; Calisto, V.; Esteves, V. I.; Schneider, R. J.; Soares, A. M. V. M.; Figueira, E.; Campos, B.; Barata, C. Effects of Carbamazepine and Cetirizine under an Ocean Acidification Scenario on the Biochemical and Transcriptome Responses of the Clam Ruditapes Philippinarum. *Environ. Pollut.* **2018**, *235*, 857–868. https://doi.org/10.1016/j.envpol.2017.12.121.

(22) Brandts, I.; Teles, M.; Gonçalves, A. P.; Barreto, A.; Franco-Martinez, L.; Tvarijonaviciute, A.; Martins, M. A.; Soares, A. M. V. M.; Tort, L.; Oliveira, M. Effects of Nanoplastics on Mytilus Galloprovincialis after Individual and Combined Exposure with Carbamazepine. *Sci. Total Environ.* **2018**, *643*, 775–784. https://doi.org/10.1016/j.scitotenv.2018.06.257.

(23) Martin-Diaz, L.; Franzellitti, S.; Buratti, S.; Valbonesi, P.; Capuzzo, A.; Fabbri, E. Effects of Environmental Concentrations of the Antiepilectic Drug Carbamazepine on Biomarkers and CAMP-Mediated Cell Signaling in the Mussel Mytilus Galloprovincialis. *Aquat. Toxicol. Amst. Neth.* **2009**, *94* (3), 177–185. https://doi.org/10.1016/j.aquatox.2009.06.015.

(24) Tsiaka, P.; Tsarpali, V.; Ntaikou, I.; Kostopoulou, M. N.; Lyberatos, G.; Dailianis, S. Carbamazepine-Mediated pro-Oxidant Effects on the Unicellular Marine Algal Species Dunaliella Tertiolecta and the Hemocytes of Mussel Mytilus Galloprovincialis. *Ecotoxicology* **2013**, *22* (8), 1208–1220. https://doi.org/10.1007/s10646-013-1108-3.

(25) Aguirre-Martínez, G. V.; DelValls, A. T.; Laura Martín-Díaz, M. Yes, Caffeine, Ibuprofen, Carbamazepine, Novobiocin and Tamoxifen Have an Effect on Corbicula Fluminea (Müller, 1774). *Ecotoxicol. Environ. Saf.* **2015**, *120*, 142–154. https://doi.org/10.1016/j.ecoenv.2015.05.036.

(26) Franzellitti, S.; Balbi, T.; Montagna, M.; Fabbri, R.; Valbonesi, P.; Fabbri, E.; Canesi, L. Phenotypical and Molecular Changes Induced by Carbamazepine and Propranolol on Larval Stages of Mytilus Galloprovincialis. *Chemosphere* **2019**, *234*, 962–970. https://doi.org/10.1016/j.chemosphere.2019.06.045.

(27) Freitas, R.; Almeida, Â.; Calisto, V.; Velez, C.; Moreira, A.; Schneider, R. J.; Esteves, V. I.; Wrona, F. J.; Figueira, E.; Soares, A. M. V. M. The Impacts of Pharmaceutical Drugs under Ocean Acidification: New Data on Single and Combined Long-Term Effects of Carbamazepine on Scrobicularia Plana. *Sci. Total Environ.* **2016**, *541*, 977–985. https://doi.org/10.1016/j.scitotenv.2015.09.138.

(28) Canzler, S.; Schor, J.; Busch, W.; Schubert, K.; Rolle-Kampczyk, U. E.; Seitz, H.; Kamp, H.; von Bergen, M.; Buesen, R.; Hackermüller, J. Prospects and Challenges of Multi-Omics Data Integration in Toxicology. *Arch. Toxicol.* **2020**, *94* (2), 371–388. https://doi.org/10.1007/s00204-020-02656-y.

(29) Pinu, F. R.; Beale, D. J.; Paten, A. M.; Kouremenos, K.; Swarup, S.; Schirra, H. J.; Wishart, D. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites* **2019**, *9* (4). https://doi.org/10.3390/metabo9040076.

(30) Chen, H.; Diao, X.; Wang, H.; Zhou, H. An Integrated Metabolomic and Proteomic Study of Toxic Effects of Benzo[a]Pyrene on Gills of the Pearl Oyster Pinctada Martensii. *Ecotoxicol. Environ. Saf.* **2018**, *156*, 330–336. https://doi.org/10.1016/j.ecoenv.2018.03.040.

(31) Ji, C.; Li, F.; Wang, Q.; Zhao, J.; Sun, Z.; Wu, H. An Integrated Proteomic and Metabolomic Study on the Gender-Specific Responses of Mussels Mytilus Galloprovincialis to Tetrabromobisphenol A (TBBPA). *Chemosphere* **2016**, *144*, 527–539. https://doi.org/10.1016/j.chemosphere.2015.08.052.

(32) Sun, J.; Zhou, Q.; Hu, X. Integrating Multi-Omics and Regular Analyses Identifies the Molecular Responses of Zebrafish Brains to Graphene Oxide: Perspectives in Environmental Criteria. *Ecotoxicol. Environ. Saf.* **2019**, *180*, 269–279. https://doi.org/10.1016/j.ecoenv.2019.05.011.

(33) Wu, C.; Zhou, F.; Ren, J.; Li, X.; Jiang, Y.; Ma, S. A Selective Review of Multi-Level Omics Data Integration Using Variable Selection. *High-Throughput* **2019**, *8* (1), 4. https://doi.org/10.3390/ht8010004.

(34) Steinmetz, V.; Sévila, F.; Bellon-Maurel, V. A Methodology for Sensor Fusion Design: Application to Fruit Quality Assessment. *J. Agric. Eng. Res.* **1999**, *74* (1), 21–31. https://doi.org/10.1006/jaer.1999.0428.

(35) Boccard, J.; Rudaz, S. Harnessing the Complexity of Metabolomic Data with Chemometrics. *J. Chemom.* **2014**, *28* (1), 1–9. https://doi.org/10.1002/cem.2567.

(36) Cai, W.; Li, Y.; Shao, X. A Variable Selection Method Based on Uninformative Variable Elimination for Multivariate Calibration of Near-Infrared Spectra. *Chemom. Intell. Lab. Syst.* **2008**, *90* (2), 188–194. https://doi.org/10.1016/j.chemolab.2007.10.001.

(37) Boccard, J.; Rutledge, D. N. A Consensus Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) Strategy for Multiblock Omics Data Fusion. *Anal. Chim. Acta* **2013**, *769*, 30–39. https://doi.org/10.1016/j.aca.2013.01.022.

(38) Bonnefille, B.; Gomez, E.; Alali, M.; Rosain, D.; Fenet, H.; Courant, F. Metabolomics Assessment of the Effects of Diclofenac Exposure on Mytilus Galloprovincialis: Potential Effects on Osmoregulation and Reproduction. *Sci. Total Environ.* **2018**, *613–614*, 611–618. https://doi.org/10.1016/j.scitotenv.2017.09.146.

(39) Dumas, T.; Bonnefille, B.; Gomez, E.; Boccard, J.; Castro, N. A.; Fenet, H.; Courant, F. Metabolomics Approach Reveals Disruption of Metabolic Pathways in the Marine Bivalve Mytilus Galloprovincialis Exposed to a WWTP Effluent Extract. *Sci. Total Environ.* **2020**, *712*, 136551. https://doi.org/10.1016/j.scitotenv.2020.136551.

(40) Want, E. J.; Masson, P.; Michopoulos, F.; Wilson, I. D.; Theodoridis, G.; Plumb, R. S.; Shockcor, J.; Loftus, N.; Holmes, E.; Nicholson, J. K. Global Metabolic Profiling of Animal and Human Tissues via UPLC-MS. *Nat. Protoc.* **2013**, *8* (1), 17–32. https://doi.org/10.1038/nprot.2012.135.

(41) Dumas, T.; Boccard, J.; Gomez, E.; Fenet, H.; Courant, F. Multifactorial Analysis of Environmental Metabolomic Data in Ecotoxicology: Wild Marine Mussel Exposed to WWTP Effluent as a Case Study. *Metabolites* **2020**, *10* (7), 269. https://doi.org/10.3390/metabo10070269.

(42) Kessner, D.; Chambers, M.; Burke, R.; Agus, D.; Mallick, P. ProteoWizard: Open Source Software for Rapid Proteomics Tools Development. *Bioinformatics* **2008**, *24* (21), 2534–2536. https://doi.org/10.1093/bioinformatics/btn323.

(43) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78* (3), 779–787. https://doi.org/10.1021/ac051437y.

(44) Kuhl, C.; Tautenhahn, R.; Böttcher, C.; Larson, T. R.; Neumann, S. CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets. *Anal. Chem.* **2012**, *84* (1), 283–289. https://doi.org/10.1021/ac202450g.

(45) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorndahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F.; Neveu, V.; Greiner, R.; Scalbert, A. HMDB 3.0—The

Human Metabolome Database in 2013. *Nucleic Acids Res.* **2013**, *41* (D1), D801–D807. https://doi.org/10.1093/nar/gks1065.

(46) Fahy, E.; Subramaniam, S.; Murphy, R. C.; Nishijima, M.; Raetz, C. R. H.; Shimizu, T.; Spener, F.; Meer, G. van; Wakelam, M. J. O.; Dennis, E. A. Update of the LIPID MAPS Comprehensive Classification System for Lipids. *J. Lipid Res.* **2009**, *50* (Supplement), S9–S14. https://doi.org/10.1194/jlr.R800095-JLR200.

(47) Blaženović, I.; Kind, T.; Ji, J.; Fiehn, O. Software Tools and Approaches for Compound Identification of LC-MS/MS Data in Metabolomics. *Metabolites* **2018**, *8* (2), 31. https://doi.org/10.3390/metabo8020031.

(48) Liebisch, G.; Fahy, E.; Aoki, J.; Dennis, E. A.; Durand, T.; Ejsing, C. S.; Fedorova, M.; Feussner, I.; Griffiths, W. J.; Köfeler, H.; Merrill, A. H.; Murphy, R. C.; O'Donnell, V. B.; Oskolkova, O.; Subramaniam, S.; Wakelam, M. J. O.; Spener, F. Update on LIPID MAPS Classification, Nomenclature, and Shorthand Notation for MS-Derived Lipid Structures. *J. Lipid Res.* **2020**, *61* (12), 1539–1555. https://doi.org/10.1194/jlr.S120001025.

(49) Hayoun, K.; Gouveia, D.; Grenga, L.; Pible, O.; Armengaud, J.; Alpha-Bazin, B. Evaluation of Sample Preparation Methods for Fast Proteotyping of Microorganisms by Tandem Mass Spectrometry. *Front. Microbiol.* **2019**, *10*, 1985. https://doi.org/10.3389/fmicb.2019.01985.

(50) Trapp, J.; Armengaud, J.; Pible, O.; Gaillard, J.-C.; Abbaci, K.; Habtoul, Y.; Chaumot, A.; Geffard, O. Proteomic Investigation of Male Gammarus Fossarum, a Freshwater Crustacean, in Response to Endocrine Disruptors. *J. Proteome Res.* **2015**, *14* (1), 292–303. https://doi.org/10.1021/pr500984z.

(51) Hartmann, E. M.; Allain, F.; Gaillard, J.-C.; Pible, O.; Armengaud, J. Taking the Shortcut for High-Throughput Shotgun Proteomic Analysis of Bacteria. In *Host-Bacteria Interactions: Methods and Protocols*; Vergunst, A. C., O'Callaghan, D., Eds.; Methods in Molecular Biology; Springer: New York, NY, 2014; pp 275–285. https://doi.org/10.1007/978-1-4939-1261-2_16.

(52) Klein, G.; Mathé, C.; Biola-Clier, M.; Devineau, S.; Drouineau, E.; Hatem, E.; Marichal, L.; Alonso, B.; Gaillard, J.-C.; Lagniel, G.; Armengaud, J.; Carrière, M.; Chédin, S.; Boulard, Y.; Pin, S.; Renault, J.-P.; Aude, J.-C.; Labarre, J. RNA-Binding Proteins Are a Major Target of Silica Nanoparticles in Cell Extracts. *Nanotoxicology* **2016**, *10* (10), 1555–1564. https://doi.org/10.1080/17435390.2016.1244299.

(53) Haas, B. J.; Papanicolaou, A.; Yassour, M.; Grabherr, M.; Blood, P. D.; Bowden, J.; Couger, M. B.; Eccles, D.; Li, B.; Lieber, M.; MacManes, M. D.; Ott, M.; Orvis, J.; Pochet, N.; Strozzi, F.; Weeks, N.; Westerman, R.; William, T.; Dewey, C. N.; Henschel, R.; LeDuc, R. D.; Friedman, N.; Regev, A. De Novo Transcript Sequence Reconstruction from RNA-Seq Using the Trinity Platform for Reference Generation and Analysis. *Nat. Protoc.* **2013**, *8* (8), 1494–1512. https://doi.org/10.1038/nprot.2013.084.

(54) Cogne, Y.; Degli-Esposti, D.; Pible, O.; Gouveia, D.; François, A.; Bouchez, O.; Eché, C.; Ford, A.; Geffard, O.; Armengaud, J.; Chaumot, A.; Almunia, C. De Novo Transcriptomes of 14 Gammarid Individuals for Proteogenomic Analysis of Seven Taxonomic Groups. *Sci. Data* **2019**, *6* (1), 184. https://doi.org/10.1038/s41597-019-0192-5.

(55) Wright, J. C.; Choudhary, J. S. DecoyPyrat: Fast Non-Redundant Hybrid Decoy Sequence Generation for Large Scale Proteomics. *J. Proteomics Bioinform.* **2016**, *9* (6), 176–180. https://doi.org/10.4172/jpb.1000404.

(56) Trapp, J.; Gouveia, D.; Almunia, C.; Pible, O.; Degli Esposti, D.; Gaillard, J.-C.; Chaumot, A.; Geffard, O.; Armengaud, J. Digging Deeper Into the Pyriproxyfen-Response of the Amphipod Gammarus Fossarum With a Next-Generation Ultra-High-Field Orbitrap Analyser: New Perspectives for Environmental Toxicoproteomics. *Front. Environ. Sci.* **2018**, *6*. https://doi.org/10.3389/fenvs.2018.00054.

(57) Liu, H.; Sadygov, R. G.; Yates, J. R. A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–4201. https://doi.org/10.1021/ac0498563.

(58) Perez-Riverol, Y.; Csordas, A.; Bai, J.; Bernal-Llinares, M.; Hewapathirana, S.; Kundu, D. J.; Inuganti, A.; Griss, J.; Mayer, G.; Eisenacher, M.; Pérez, E.; Uszkoreit, J.; Pfeuffer, J.; Sachsenberg, T.; Yılmaz, Ş.; Tiwary, S.; Cox, J.; Audain, E.; Walzer, M.; Jarnuczak, A. F.; Ternent, T.; Brazma, A.; Vizcaíno, J. A. The PRIDE Database and Related Tools and Resources in 2019: Improving Support for Quantification Data. *Nucleic Acids Res.* **2019**, *47* (D1), D442–D450. https://doi.org/10.1093/nar/gky1106.

(59) Li, H.-D.; Xu, Q.-S.; Liang, Y.-Z. LibPLS: An Integrated Library for Partial Least Squares Regression and Linear Discriminant Analysis. *Chemom. Intell. Lab. Syst.* **2018**, *176*, 34–43. https://doi.org/10.1016/j.chemolab.2018.03.003.

(60) Bylesjö, M.; Rantalainen, M.; Nicholson, J. K.; Holmes, E.; Trygg, J. K-OPLS Package: Kernel-Based Orthogonal Projections to Latent Structures for Prediction and Interpretation in Feature Space. *BMC Bioinformatics* **2008**, *9* (1), 106. https://doi.org/10.1186/1471-2105-9-106.

(61) Darzi, Y.; Letunic, I.; Bork, P.; Yamada, T. IPath3.0: Interactive Pathways Explorer V3. *Nucleic Acids Res.* **2018**, *46* (W1), W510–W513. https://doi.org/10.1093/nar/gky299.

(62) Bradley, P. H.; Brauer, M. J.; Rabinowitz, J. D.; Troyanskaya, O. G. Coordinated Concentration Changes of Transcripts and Metabolites in Saccharomyces Cerevisiae. *PLOS Comput. Biol.* **2009**, *5* (1), e1000270. https://doi.org/10.1371/journal.pcbi.1000270.

(63) Hirai, M. Y.; Klein, M.; Fujikawa, Y.; Yano, M.; Goodenowe, D. B.; Yamazaki, Y.; Kanaya, S.; Nakamura, Y.; Kitayama, M.; Suzuki, H.; Sakurai, N.; Shibata, D.; Tokuhisa, J.; Reichelt, M.; Gershenzon, J.; Papenbrock, J.; Saito, K. Elucidation of Gene-to-Gene and Metabolite-to-Gene Networks in Arabidopsis by Integration of Metabolomics and Transcriptomics. *J. Biol. Chem.* **2005**, *280* (27), 25590–25595. https://doi.org/10.1074/jbc.M502332200.

(64) Hirai, M. Y.; Yano, M.; Goodenowe, D. B.; Kanaya, S.; Kimura, T.; Awazuhara, M.; Arita, M.; Fujiwara, T.; Saito, K. Integration of Transcriptomics and Metabolomics for Understanding of Global Responses to Nutritional Stresses in Arabidopsis Thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (27), 10205–10210. https://doi.org/10.1073/pnas.0403218101.

(65) Nikiforova, V. J.; Daub, C. O.; Hesse, H.; Willmitzer, L.; Hoefgen, R. Integrative Gene-Metabolite Network with Implemented Causality Deciphers Informational Fluxes of Sulphur Stress Response. *J. Exp. Bot.* **2005**, *56* (417), 1887–1896. https://doi.org/10.1093/jxb/eri179.

(66) Gibon, Y.; Usadel, B.; Blaesing, O. E.; Kamlage, B.; Hoehne, M.; Trethewey, R.; Stitt, M. Integration of Metabolite with Transcript and Enzyme Activity Profiling during Diurnal Cycles in Arabidopsis Rosettes. *Genome Biol.* **2006**, *7* (8), R76. https://doi.org/10.1186/gb-2006-7-8-r76.

(67) Carrari, F.; Baxter, C.; Usadel, B.; Urbanczyk-Wochniak, E.; Zanor, M.-I.; Nunes-Nesi, A.; Nikiforova, V.; Centero, D.; Ratzka, A.; Pauly, M.; Sweetlove, L. J.; Fernie, A. R. Integrated Analysis of Metabolite and Transcript Levels Reveals the Metabolic Shifts That Underlie Tomato Fruit Development and Highlight Regulatory Aspects of Metabolic Network Behavior. *Plant Physiol.* **2006**, *142* (4), 1380–1396. https://doi.org/10.1104/pp.106.088534.

(68) Urbanczyk-Wochniak, E.; Baxter, C.; Kolbe, A.; Kopka, J.; Sweetlove, L. J.; Fernie, A. R. Profiling of Diurnal Patterns of Metabolite and Transcript Abundance in Potato (Solanum Tuberosum) Leaves. *Planta* **2005**, *221* (6), 891–903. https://doi.org/10.1007/s00425-005-1483-y.

(69) Sarkar, S.; Rubinsztein, D. C. Inositol and IP3 Levels Regulate Autophagy—Biology and Therapeutic Speculations. *Autophagy* **2006**, *2* (2), 132–134. https://doi.org/10.4161/auto.2387.

(70) Picot, S.; Faury, N.; Arzul, I.; Chollet, B.; Renault, T.; Morga, B. Identification of the Autophagy Pathway in a Mollusk Bivalve, Crassostrea Gigas. *Autophagy* **2020**, *16* (11), 2017–2035. https://doi.org/10.1080/15548627.2020.1713643.

(71) Høyer-Hansen, M.; Jäättelä, M. Connecting Endoplasmic Reticulum Stress to Autophagy by Unfolded Protein Response and Calcium. *Cell Death Differ.* **2007**, *14* (9), 1576–1582. https://doi.org/10.1038/sj.cdd.4402200.

(72) Primeau, J. O.; Armanious, G. P.; Fisher, M. E.; Young, H. S. The SarcoEndoplasmic Reticulum Calcium ATPase. In *Membrane Protein Complexes: Structure and Function*; Harris, J. R.,

Boekema, E. J., Eds.; Subcellular Biochemistry; Springer: Singapore, 2018; pp 229–258. https://doi.org/10.1007/978-981-10-7757-9_8.
(73) Castillo, K.; Rojas-Rivera, D.; Lisbona, F.; Caballero, B.; Nassif, M.; Court, F. A.; Schuck, S.; Ibar, C.; Walter, P.; Sierralta, J.; Glavic, A.; Hetz, C. BAX Inhibitor-1 Regulates Autophagy by Controlling the IRE1α Branch of the Unfolded Protein Response. *EMBO J.* **2011**, *30* (21), 4465–4478. https://doi.org/10.1038/emboj.2011.318.
(74) Martínez-Gómez, C.; Bignell, J.; Lowe, D. *Lysosomal Membrane Stability in Mussels.*; 56; International Council for the Exploration of the Sea (ICES), 2015; p 41. https://doi.org/10.25607/OBP-240.
(75) Boya, P.; Kroemer, G. Lysosomal Membrane Permeabilization in Cell Death. *Oncogene* **2008**, *27* (50), 6434–6451. https://doi.org/10.1038/onc.2008.310.
(76) Raftopoulou, E. K.; Dimitriadis, V. K. Aspects of the Digestive Gland Cells of the Mussel Mytilus Galloprovincialis, in Relation to Lysosomal Enzymes, Lipofuscin Presence and Shell Size: Contribution in the Assessment of Marine Pollution Biomarkers. *Mar. Pollut. Bull.* **2012**, *64* (2), 182–188. https://doi.org/10.1016/j.marpolbul.2011.12.017.
(77) Viarengo, A.; Lowe, D.; Bolognesi, C.; Fabbri, E.; Koehler, A. The Use of Biomarkers in Biomonitoring: A 2-Tier Approach Assessing the Level of Pollutant-Induced Stress Syndrome in Sentinel Organisms. *Comp. Biochem. Physiol. Toxicol. Pharmacol. CBP* **2007**, *146* (3), 281–300. https://doi.org/10.1016/j.cbpc.2007.04.011.
(78) Schrader, M.; Costello, J.; Godinho, L. F.; Islinger, M. Peroxisome-Mitochondria Interplay and Disease. *J. Inherit. Metab. Dis.* **2015**, *38* (4), 681–702. https://doi.org/10.1007/s10545-015-9819-7.
(79) Lawrence, A. J.; Hemingway, K. L. *Effects of Pollution on Fish: Molecular Effects and Population Responses*; John Wiley & Sons, 2008.
(80) STANLEY-SAMUELSON, D. W. The Biological Significance of Prostaglandins and Related Eicosanoids in Invertebrates12. *Am. Zool.* **1994**, *34* (6), 589–598. https://doi.org/10.1093/icb/34.6.589.
(81) Reddy, J. K.; Mannaerts, G. P. Peroxisomal Lipid Metabolism. *Annu. Rev. Nutr.* **1994**, *14* (1), 343–370. https://doi.org/10.1146/annurev.nu.14.070194.002015.
(82) Nemali, M. R.; Reddy, M. K.; Usuda, N.; Reddy, P. G.; Comeau, L. D.; Rao, M. S.; Reddy, J. K. Differential Induction and Regulation of Peroxisomal Enzymes: Predictive Value of Peroxisome Proliferation in Identifying Certain Nonmutagenic Carcinogens. *Toxicol. Appl. Pharmacol.* **1989**, *97* (1), 72–87. https://doi.org/10.1016/0041-008X(89)90056-2.
(83) Reddy, J. K.; Lalwai, N. D. Carcinogenesis by Hepatic Peroxisome Proliferators: Evaluation of the Risk of Hypolipidemic Drugs and Industrial Plasticizers to Humans. *Crit. Rev. Toxicol.* **1983**, *12* (1), 1–58. https://doi.org/10.3109/10408448309029317.
(84) Cajaraville, M. P.; Ortiz-Zarragoitia, M. Specificity of the Peroxisome Proliferation Response in Mussels Exposed to Environmental Pollutants. *Aquat. Toxicol.* **2006**, *78*, S117–S123. https://doi.org/10.1016/j.aquatox.2006.02.016.
(85) Johansson, A.-C.; Appelqvist, H.; Nilsson, C.; Kågedal, K.; Roberg, K.; Öllinger, K. Regulation of Apoptosis-Associated Lysosomal Membrane Permeabilization. *Apoptosis* **2010**, *15* (5), 527–540. https://doi.org/10.1007/s10495-009-0452-5.
(86) Abreu-Martin, M. T.; Chari, A.; Palladino, A. A.; Craft, N. A.; Sawyers, C. L. Mitogen-Activated Protein Kinase Kinase Kinase 1 Activates Androgen Receptor-Dependent Transcription and Apoptosis in Prostate Cancer. *Mol. Cell. Biol.* **1999**, *19* (7), 5143–5154. https://doi.org/10.1128/MCB.19.7.5143.
(87) Foghsgaard, L.; Wissing, D.; Mauch, D.; Lademann, U.; Bastholm, L.; Boes, M.; Elling, F.; Leist, M.; Jäättelä, M. Cathepsin B Acts as a Dominant Execution Protease in Tumor Cell Apoptosis Induced by Tumor Necrosis Factor. *J. Cell Biol.* **2001**, *153* (5), 999–1010. https://doi.org/10.1083/jcb.153.5.999.

(88) Guo, Y.; Srinivasula, S. M.; Druilhe, A.; Fernandes-Alnemri, T.; Alnemri, E. S. Caspase-2 Induces Apoptosis by Releasing Proapoptotic Proteins from Mitochondria. *J. Biol. Chem.* **2002**, *277* (16), 13430–13437. https://doi.org/10.1074/jbc.M108029200.

(89) Kumar, S. Caspase 2 in Apoptosis, the DNA Damage Response and Tumour Suppression: Enigma No More? *Nat. Rev. Cancer* **2009**, *9* (12), 897–903. https://doi.org/10.1038/nrc2745.

(90) Porté, S.; Valencia, E.; Yakovtseva, E. A.; Borràs, E.; Shafqat, N.; Debreczeny, J. E.; Pike, A. C. W.; Oppermann, U.; Farrés, J.; Fita, I.; Parés, X. Three-Dimensional Structure and Enzymatic Function of Proapoptotic Human P53-Inducible Quinone Oxidoreductase PIG3. *J. Biol. Chem.* **2009**, *284* (25), 17194–17205. https://doi.org/10.1074/jbc.M109.001800.

(91) Romero, A.; Novoa, B.; Figueras, A. The Complexity of Apoptotic Cell Death in Mollusks: An Update. *Fish Shellfish Immunol.* **2015**, *46* (1), 79–87. https://doi.org/10.1016/j.fsi.2015.03.038.

(92) Mezzelani, M.; Nardi, A.; Bernardini, I.; Milan, M.; Peruzza, L.; d'Errico, G.; Fattorini, D.; Gorbi, S.; Patarnello, T.; Regoli, F. Environmental Pharmaceuticals and Climate Change: The Case Study of Carbamazepine in M. Galloprovincialis under Ocean Acidification Scenario. *Environ. Int.* **2021**, *146*, 106269. https://doi.org/10.1016/j.envint.2020.106269.

(93) Yan, S.; Chen, R.; Wang, M.; Zha, J. Carbamazepine at Environmentally Relevant Concentrations Caused DNA Damage and Apoptosis in the Liver of Chinese Rare Minnows (Gobiocypris Rarus) by the Ras/Raf/ERK/P53 Signaling Pathway. *Environ. Pollut.* **2021**, *270*, 116245. https://doi.org/10.1016/j.envpol.2020.116245.

(94) Pytharopoulou, S.; Kouvela, E. C.; Sazakli, E.; Leotsinidis, M.; Kalpaxis, D. L. Evaluation of the Global Protein Synthesis in Mytilus Galloprovincialis in Marine Pollution Monitoring: Seasonal Variability and Correlations with Other Biomarkers. *Aquat. Toxicol.* **2006**, *80* (1), 33–41. https://doi.org/10.1016/j.aquatox.2006.07.010.

(95) Sokolova, I. M.; Frederich, M.; Bagwe, R.; Lannig, G.; Sukhotin, A. A. Energy Homeostasis as an Integrative Tool for Assessing Limits of Environmental Stress Tolerance in Aquatic Invertebrates. *Mar. Environ. Res.* **2012**, *79*, 1–15. https://doi.org/10.1016/j.marenvres.2012.04.003.

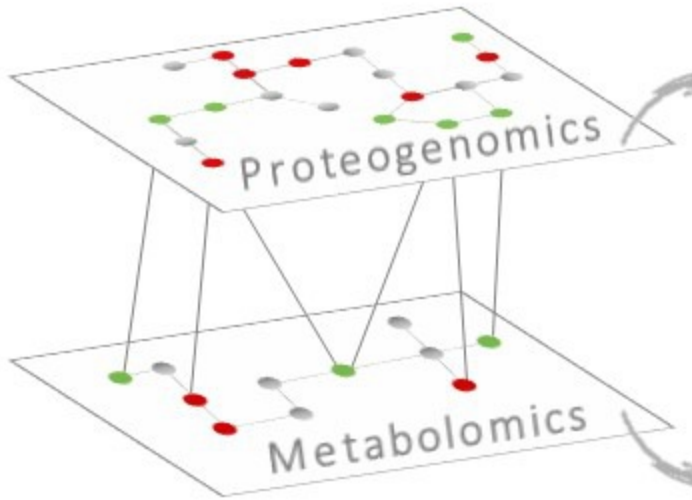
(96) Almeida, Â.; Freitas, R.; Calisto, V.; Esteves, V. I.; Schneider, R. J.; Soares, A. M. V. M.; Figueira, E. Chronic Toxicity of the Antiepileptic Carbamazepine on the Clam Ruditapes Philippinarum. *Comp. Biochem. Physiol. Toxicol. Pharmacol. CBP* **2015**, *172–173*, 26–35. https://doi.org/10.1016/j.cbpc.2015.04.004.

(97) Parolini, M.; Quinn, B.; Binelli, A.; Provini, A. Cytotoxicity Assessment of Four Pharmaceutical Compounds on the Zebra Mussel (Dreissena Polymorpha) Haemocytes, Gill and Digestive Gland Primary Cell Cultures. *Chemosphere* **2011**, *84* (1), 91–100. https://doi.org/10.1016/j.chemosphere.2011.02.049.

Solvent Control

Carbamazepine 80 ng/L

Carbamazepine 80 µg/L



Mid-level data fusion

MCUVE-PLS & consensus OPLS-DA

Metabolic & Cellular stress

- Amino acid, purine and lipid metabolisms
- Protein synthesis process
- Transport and catabolism processes

HYPOTHESES

- Autophagy induction
- Lysosomal membrane destabilization
- Peroxisomes and mitochondria dysfunction
- Apoptosis