Information of the Journal in which the present paper is published:

Analytical and Bioanalytical Chemistry, 2020

https://doi.org/10.1007/s00216-020-02789-0

MCR–ALS analysis of ^{1}H NMR by segments to study the zebra fish exposure to acrylamide

Yolanda Pérez,[§] Marta Casado,[†] Demetrio Raldúa,[†] Eva Prats,[†] Benjamín Piña,[†] Romà Tauler,[†] Ignacio Alfonso, and Francesc Puig-Castellví,[†,‡,*]

§NMR Facility, IQAC–CSIC, Jordi Girona 18–26, 08034 Barcelona

†Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona 18-26, 08034 Barcelona

Department of Biological Chemistry, IQAC–CSIC, Jordi Girona 18–26, 08034 Barcelona

KEYWORDS

MCR-ALS, NMR, chemometrics, metabolite discovery, 1H

ABSTRACT

A pipeline for the analysis of 1D ^{1}H NMR metabolomics data is proposed, consisting in the blind resolution and integration of all resonances of the spectral dataset by Multivariate Curve Resolution–Alternating Least Squares (MCR–ALS) and followed by the analysis of the resulting matrix of concentrations with chemometric methods such as Principal Component Analysis (PCA), ANOVA – Simultaneous Component Analysis (ASCA) and Partial Least Squares – Discriminant Analysis (PLS–DA). Since MCR–ALS does not require the use of spectral templates, concentrations for all resonances are obtained before being assigned without neglecting any resonance relevant for the metabolomics study in the analysis of these concentrations. In this work, the proposed pipeline performance was validated with 1D ^{1}H NMR spectra from a metabolomics study of zebra fish upon acrylamide (ACR) exposure. As a conclusion, this pipeline represents a framework for the high–throughput analysis of NMR metabolomics data that opens the path for truly untargeted NMR metabolomics analyses.

INTRODUCTION

Nuclear magnetic Resonance (NMR) spectroscopy is one of the most important analytical techniques used for metabolite discovery.[1–2] In the biological field, NMR is also employed in metabolomics studies to characterize the metabolites in cell extracts, tissues and living organisms for disease diagnosis and biomarker discovery.[3] The interpretation of NMR spectra from biological samples is not easy due to their complex composition. In a typical metabolomics NMR spectrum, hundreds of resonances from several dozens of metabolites are obtained.[4]

Since resonances cannot be directly assigned because of their strong overlap, the NMR spectra are usually investigated using multivariate data analysis approaches[5–10], such as Principal Component Analysis (PCA)[11], ANOVA – Simultaneous Component Analysis (ASCA)[12] or Partial Least Squares – Discriminant Analysis (PLS–DA)[13]. When these approaches are used for the analysis of NMR spectra, the chemical shifts of the resonances from the altered metabolites can be identified. However, the other spectroscopic parameters (*J*, multiplicity, proton number) are not directly retrieved from these analyses and further work needs to be done by the NMR spectroscopist to successfully characterize the metabolites of interest. Moreover, to obtain quantitative information, these chemometric analyses should be followed by the integration of the relevant resonances on the original NMR spectra using deconvolution approaches.

To improve the spectroscopic interpretation of these chemometric analyses of NMR data and to avoid the integration and deconvolution steps, we have designed a new metabolomic workflow that introduces as a preliminary step the application of the Multivariate Curve Resolution – Alternating Least Squares (MCR–ALS)[14] method. MCR–ALS allows resolving, for every metabolite, their resonances and their corresponding integrals (or relative concentrations) [15]. Thus, after obtaining the relative concentrations

from this step, chemometric (e.g., PCA, PLS–DA) and statistical (e.g., t–test) analysis can be performed on them.

MCR–ALS performance to resolve all the pure spectral resonances (and pure integrals) of the metabolites in the analyzed sample depends on the variance of its associated dataset.[14–17] As a matter of fact, the resolved components obtained in the MCR–ALS analysis of 1D ^{1}H NMR metabolomics datasets usually contain mixed contributions of several metabolites and they are descriptive of metabolic or physiologic processes.[18–20] Having said that, it is possible to achieve an improved performance and get the pure resonance profiles of the metabolites in the analyzed samples by using the DTC–MCR–ALS methodology, which constrains the search and resolution of resonances in the narrow spectral region where they are located.[18] The DTC–MCR–ALS approach works in two steps. In the first step, the NMR data is split into smaller datasets (or segments) comprising each one a spectral segment of the original dataset and then each one is analyzed by MCR–ALS. In a second step, the original full–length NMR data is analyzed with MCR–ALS in combination with a selectivity constraint built after examining the data from the first step using a decision tree approach.[18] Since the smaller segments analysed in the first step of the DTC–MCR–ALS contain resonances from a lower number of metabolites, the complexity of each data segment would be substantially lower than in the original raw data, and therefore a better MCR–ALS resolution of the pure resonances using a lower number of components is feasible.

In this work, we have investigated the use of the first step of the DTC–MCR–ALS as an independent data pre–processing method to retrieve the concentration and clean spectral profiles of the resonances present in the analysed ^{1}H NMR dataset. The principle of using MCR–ALS as a pre–processing tool for ^{1}H NMR data was first introduced by Ebrahimi *et al.* in 2016.[15] In this paper, MCR–ALS was applied on 7 segments from a 1D ^{1}H NMR dataset comprising 1,632 samples to obtain the resonance integrals from 7 known metabolites. This study can be regarded as targeted, although it showed it is possible to obtain

the concentration profiles of NMR resonances without needing to assume their peak shape as with deconvolution methods,[7,21–24] revealing a high potential for being applied in untargeted studies.

Contemporary to our work, Khakimov showed that MCR–ALS analysis on the NMR segments can be used to examine metabolomics data from urine samples.[25] By using this approach, they were able to recover the concentration profiles from 209 resonances. In addition, due to the complexity or urine samples, 94 NMR segments containing the most shifted resonances were analysed with binning instead of MCR–ALS. Their analyses revealed that the pre–processing based on MCR–ALS produced the lowest variability within group and the largest between groups if compared to other NMR pre–processing methods. However, due to the methodological purpose of the study, the differences between groups at a metabolic level were not examined.

In this regard, in the present study, we have investigated a ^{1}H NMR dataset of metabolic extracts by windowing and MCR–ALS analysis. Specifically, we have evaluated the metabolic alterations in zebra fish muscle and brain upon acrylamide (ACR) exposure. ACR is a neurotoxic, mutagenic, and reprotoxic compound that can be formed in cooking at elevated temperatures.[26–27] Due to its public health risk, ACR consumption levels are regulated in some countries.[28–29] Samples included in the dataset comprised an experimental design of 2 factors and 2 levels, which were investigated by PCA, ASCA and PLS–DA. Opposed to **Khakimov**'s study[25], we were able to analyze all the spectral segments with MCR–ALS since they do not contain strong peak shifts as in the urine samples. Moreover, the PLS–DA analyses revealed metabolomics differences between groups that could be linked to the system of study.

EXPERIMENTAL SECTION

Biological experiment and dissection. Wild–type zebrafish were obtained from Piscicultura Superior (Barcelona, Spain) and maintained in fish water (reverse–osmosis purified water containing 90 mg mL/

Instant Ocean® (Aquarium Systems, Sarrebourg, France), 0.58 mM $CaSO_4 \cdot 2H_2O$) at 28 ± 1°C under a 12L:12D photoperiod. Adult zebrafish (≈50:50 male:female ratio) were exposed for 72 h to 0.75 mM ACR (Sigma–Aldrich, St. Louis, MO) in fish water. Control fish were maintained in fish water under identical conditions. Experimental solutions were renewed after 48 h of exposure.

For sample collection fish were euthanized by inducing hypothermic shock in ice–chilled water (2° to 4°C). Brain and muscle tissue were immediately excised and pooled (4 brains or muscle/sample) and stored at –80°C for further analyses. All procedures were approved by the Institutional Animal Care and Use Committees at the CID–CSIC and conducted in accordance with the institutional guidelines under a license from the local government (agreement number 9027).

Metabolite extraction and sample preparation. Tissue samples were freeze–dried and homogenized in 1500 µl of a solution of cold $CHCl_3$:MeOH (2:1) using a TissueLyser (Qiagen, 2 stainless steel beads, 50 Hz, 2 min for brain, 4 min for muscle). Next, 350 µl of cold milliQ water were added and the biphasic system was mixed at 4ºC for 20 min by orbital shaking. After centrifugation (13,000 rpm, 4°C, 20 min), the aqueous phase was collected and transferred to a 2–ml eppendorf tube. The extraction process was repeated once. The two supernatants were combined and freeze–dried, and the resulting dried residue was stored at –80°C until analyzed.

The extracts were dissolved in 700 µl of deuterated phosphate buffer (Na_2DPO_4 25 mM, pH 7.0 in D_2O) with DSS 0.2 mM as internal standard. After centrifugation at 13,000 rpm for 5 min at 4ºC, 650 µl of the supernatant were collected and used for the NMR analysis. In total, 20 samples were prepared (5 per every individual condition).

NMR spectroscopy. 1D NOESY–presat NMR spectra (Bruker library – noesygppr1d) were recorded at 298 K on a Bruker Avance III console (Bruker Biospin, Germany) combined with a 9.7 T Oxford

Instruments magnet (^{1}H 500 MHz), equipped with a 5 mm inverse cryogenically cooled triple–resonance TCI (^{1}H, ^{13}C, ^{15}N, and ^{2}H lock) cryoprobe with a z–axis magnetic field–gradient capability.

For every tube, the sample was locked to the solvent, tuned and matched, shimmed and the optimal 90° pulse was calculated using the automated routine. The 90° length value and the saturation frequency and power (presaturation power corresponding to a 25 Hz field) for each sample and the rest of fixed acquisition parameters were used for the 1D NOESY acquisition. After spectral recording, for every sample, the optimal water saturation (the residual water peak should be small and to allow good phase correction) and the line width at half height of the DSS (which should be less than 0.7 Hz with no line broadening) were also checked.

In addition, f**or one of the samples of each group (untreated/exposed to ACR), an ensemble of 2D NMR experiments were acquired (^{1}H J-resolved, COSY and TOCSY and ^{1}H-^{13}C HSQC; acquisition parameters in Supplementary material).**

Pre-processing of NMR spectra. ^{1}H (1D NOESY–presat) NMR spectra were automatically referenced, phased and baseline corrected using TopSpin (Bruker BioSpin GmbH, Billerica, MA, USA) routines. Moreover, the 1D ^{1}H NMR spectra were apodized with a weighting exponential function of 0.3 Hz with MestreNova v.11.0 (Mestrelab Research, Spain), and imported to Matlab R2016a (The Mathworks Inc., Natick, MA, USA) as a spectral data matrix. Then, regions of 4.68 –5.14 ppm (water), 7.65 –7.68 ppm (chloroform), below 0.7 ppm (DSS), and above 9.5 ppm (empty) were removed. After the removal of these spectral regions, the number of data–points for each ^{1}H NMR spectra was 28,913. Minor resonance misalignments were corrected with *icoshift*[30] algorithm. Finally, the processed ^{1}H NMR spectra were normalized using the Probabilistic Quotient Normalization (PQN) [31] method.

Multivariate Curve Resolution – Alternating Least Squares (MCR-ALS). In the application of the MCR–ALS method, a data matrix, D, is decomposed by a bilinear model into the product of two factor

matrices which describe the set of concentrations, C, and spectral profiles, S^T, of the constituents of the analyzed samples (Figure 1).

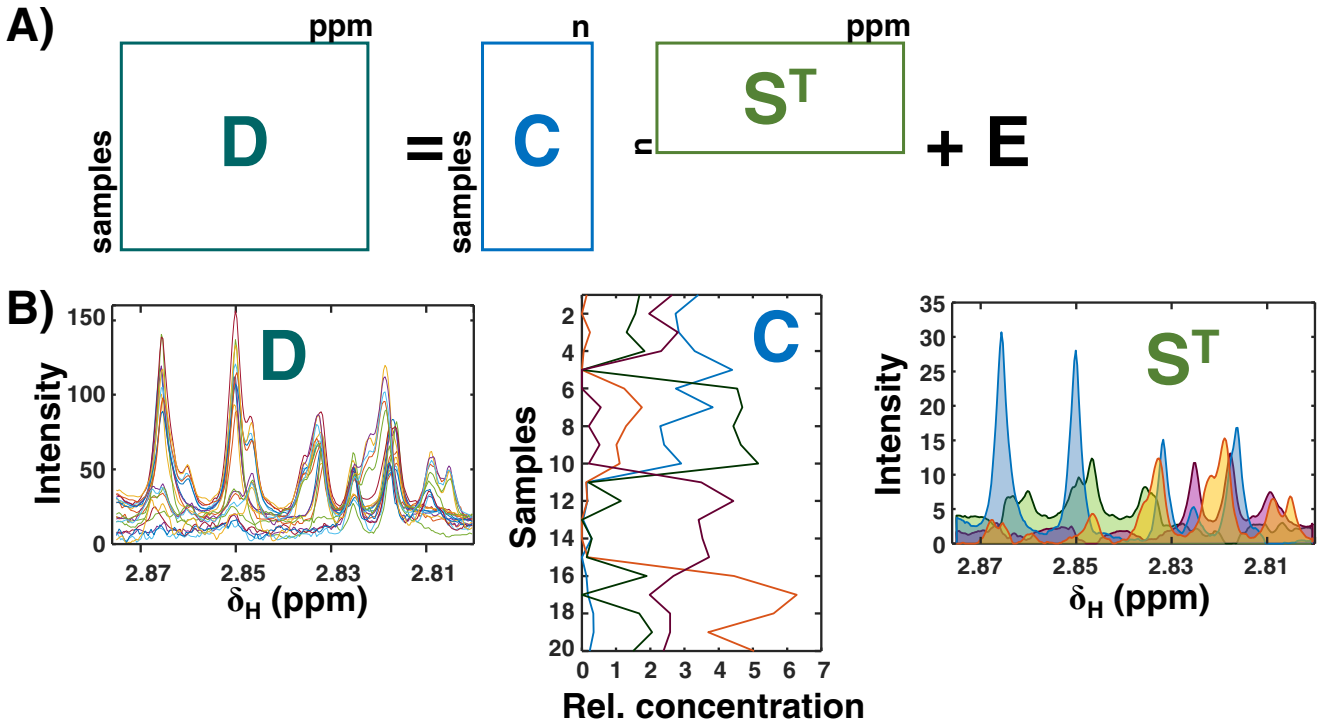


Figure 1. MCR–ALS. A) Matrix representation of the MCR–ALS bilinear decomposition. B) The 1D ^{1}H NMR dataset (D) can be decomposed into a set of concentrations (C) and spectral resonances (S^T) from *n* metabolites. The plot of C profiles shows the evolution of the resolved concentration profiles for the 3 metabolites (*n*=4) over the 20 analyzed samples. The plot of S^T spectra shows the resonances from the 4 metabolites.

For 1D ^{1}H NMR data, matrix C contains the set of relative concentrations corresponding to the S^T proton resonances in the 1D ^{1}H NMR spectra of the analyzed samples (Figure 1). The number of resolved profiles in C and S^T (or number of components) needs to be initially determined. In this case, the optimal number of components was assessed by Singular Value Decomposition analysis (see Supplementary material). [32] During the MCR–ALS bilinear decomposition, constraints should be

proposed to drive the resolution of C and S^T factor matrices to be in agreement with the previous physical knowledge about the physical nature of the profiles of the chemical constituents of the studied system. In this study, only non-negativity constraints were applied for the resolution of both C and S^T matrices.

The dataset consisted of 20 1D ^{1}H NMR spectra (in rows) with 28,913 δ_H values (in columns). This dataset was split into 98 spectral segments (step B1 in Figure 2). Every spectral window was set to contain at least one proton resonance, and limits (beginning and end) of these spectral segments were chosen at chemical shifts whose intensities were at the noise level with the aim not to split any resonance between two segments.

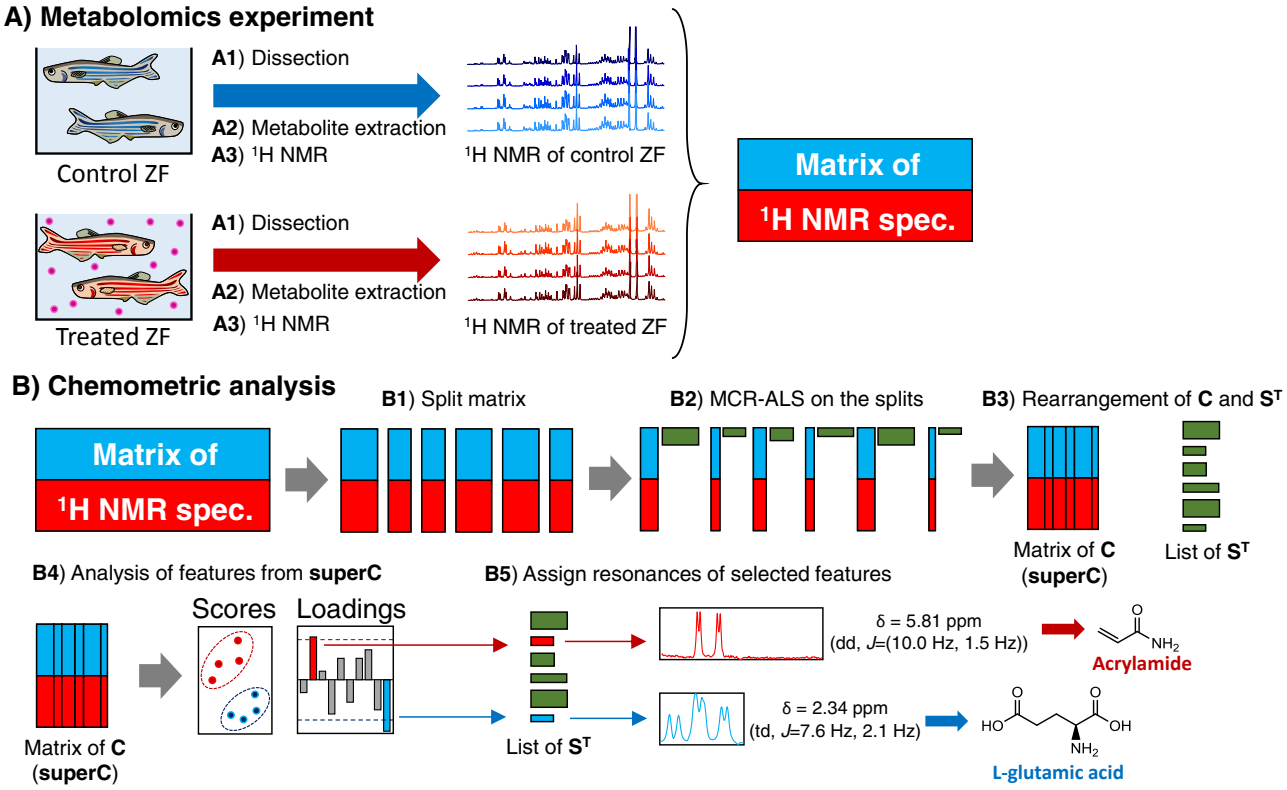


Figure 2. Proposed workflow for the NMR metabolomics experiment.

Every window was then examined separately with MCR-ALS (step B2 in Figure 2). MCR-ALS analyses were performed using the MCR-ALS GUI 2.0[33] toolbox under Matlab 2016b (The Mathworks Inc. Natick, MA, USA) environment.

A total number of 98 spectral segments were resolved by MCR–ALS analysis using between 1 and 4 components for each window. Thus, from these analyses, 98 C and 98 S^T matrices were obtained, and a total number of 222 components were obtained in the analysis of all the spectral segments of the 1D ^{1}H NMR dataset. The 98 C matrices were then combined row–wisely into a super–augmented C matrix (from now on, this is called the 'SuperC' matrix) with 20 rows (one per sample) and 222 columns (one per component) (step B3 in Figure 2).

The arrangement and indexing of these components in the superC matrix was performed using the same convention as when plotting NMR spectra: lower column indexes corresponded to MCR–ALS components associated with spectral features at lower chemical shifts, while components with spectral features at higher chemical shifts were given high column indexes. The superC matrix was then further investigated with other multivariate data analysis chemometric methods, such as PCA or PLS–DA (step B4 in Figure 2), and the relevant components were assigned by the inspection of their corresponding spectral features (step B5 in Figure 2).

Principal Component Analysis. PCA[11] was applied to the raw 1D ^{1}H NMR spectral data and to the superC data matrices. Prior to PCA analyses, all datasets were mean–centered. PCA results were validated using a venetian blinds cross–validation method. The analyses were carried out using the PLStoolbox 7.8.0 (Eigenvector Research Inc.).

ANOVA Simultaneous Component Analysis (ASCA). ASCA is a multivariate data analysis method that combines the power of ANOVA to separate variance sources with the advantages of Simultaneous Component Analysis (SCA) to the modeling of the individual separate effect matrices.[12] In this study, ASCA was employed to untangle the metabolic response linked to two effects of factors, the type of tissue and the ACR exposure (treatment), according to the model defined in eq. 1:

$$X = X_{Tissue} + X_{treatment} + X_{Treatment\_Tissue} + E, \qquad \text{(eq. 1)}$$

Where X corresponds to the superC matrix, X_{Tissue} and $X_{treatment}$ are descriptive of the separated metabolic fingerprints that are either determined by the two types of tissue or by the ACR exposure

treatment, and $X_{Treatment\_Tissue}$ is giving the metabolic fingerprint showing the possible specific interaction between these two factors (type of tissue and ACR exposure) for each set of samples. From the analysis of each of the three decomposed matrices, it was possible to investigate whether significant differences existed due to the type of tissue, to the ACR treatment only, or to the two factors at the same time (e.g., if ACR exposure causes different metabolic response depending on the type of tissue). The statistical significance of the effects of each factor and their interaction was evaluated using a permutation test (with 1,000 permutations).[34] Before the ASCA analysis, the superC matrix was mean–centered. ASCA analysis were performed using the PLStoolbox 7.8.0 (Eigenvector Research Inc.).

Partial Least Squares-Discriminant Analysis. In this study, 3 different PLS–DA analyses were performed. One first PLS–DA analysis (PLS–DA #1 in Figure 3) was used to discriminate the two analyzed tissues (brain and muscle) at control conditions and identify the metabolites responsible of the tissue discrimination. The two other PLS–DA analyses (PLS–DA #2 and #3 in Figure 3) were applied to discriminate between samples from control and exposed ZF for each of the two tissues, respectively, and to identify the metabolites associated to their separation. For each of the 3 PLS–DA analyses, the studied dataset had the concentration profiles from the superC matrix corresponding to the subgroup of samples to be compared. Before each PLS–DA analysis, the data were auto–scaled. PLS–DA analyses were validated with a venetian blinds type of cross–validation. The analyses were carried out using the PLStoolbox 7.8.0 (Eigenvector Research Inc.).

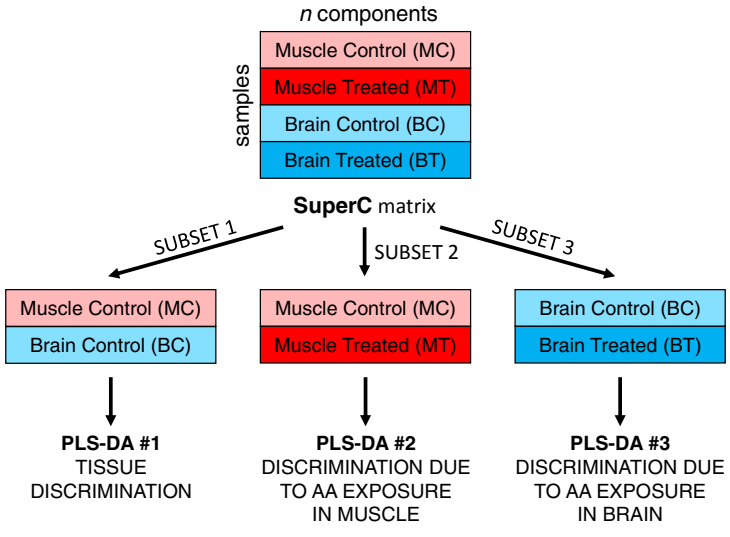


Figure 3. PLS–DA on subsets from SuperC matrix.

Spectral resonances integration with BATMAN. Results obtained in the PLS–DA analyses were verified by manual integration of the individual resonances using the BATMAN R–package[24] and by the posterior assessment of their statistical significance using a Student t–test.

RESULTS AND DISCUSSION

Exploratory analysis of the superC matrix by PCA. The PCA analysis of the superC matrix containing the relative concentration values of the 222 resolved features in the 20 zebrafish (muscle or brain) samples revealed an orthogonal distribution of the metabolism of this organism according to the type of tissue and to the ACR exposure (Figure 4B). Specifically, the samples were separated according to the two types of tissues (muscle and brain) in PC1 (63.43% of the explained variance), while PC2 (9.99%) allowed the separation according to the ACR exposure.

Whereas this orthogonal score discrimination was clearly detected in the PCA of the superC matrix, it could not be observed when the same analysis was performed on the original 1D ^{1}H NMR dataset(Figure 4A). In fact, results from the PCA analysis in Figure 4A erroneously suggested that the metabolic response of the ACR–treatment could be neglected when compared with the metabolic variance observed between the two tissues. In Figure 4A, the component that separates between the two tissues (PC1) accounted for the 97.12% of the explained variance of the dataset. In order to separate samples due to the ACR–treatment, it is needed to examine the dataset at the fifth PC (PC5, Figure 4C), which only accounted for the 0.18% of the explained variance of the dataset. However, the cluster separation in this component is not perfect either. Similarly, the effect of the ACR–treatment could be better observed with PCA when samples from each tissue were analyzed separately, but still the clusters were partially overlapped in the two cases (FigS4).

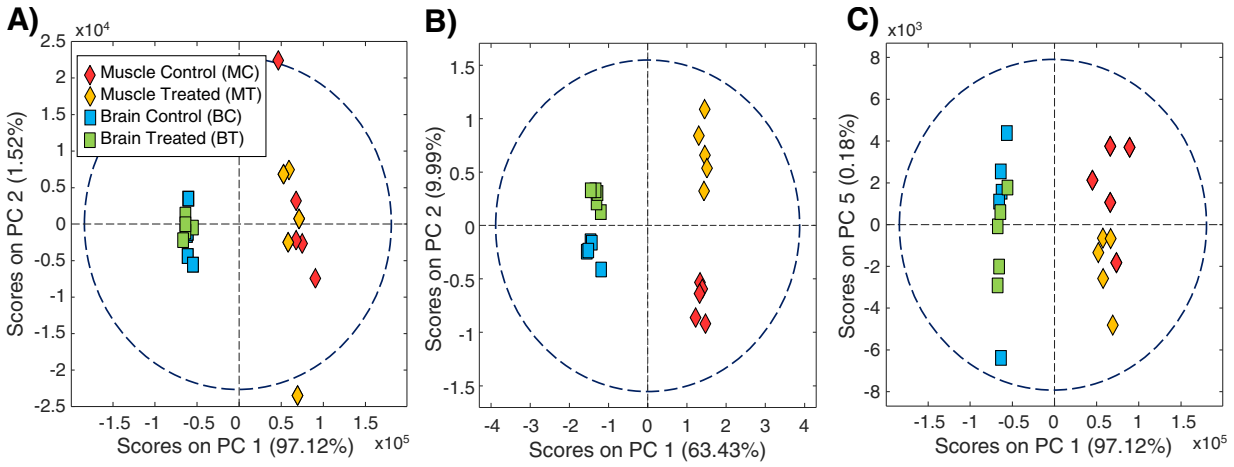


Figure 4. A) PCA scores (PC1 vs PC2) on the 1D ^{1}H NMR spectral matrix. B) PCA scores (PC1 vs PC2) on the superC matrix. C) PCA scores (PC1 vs PC5) on the 1D ^{1}H NMR spectral matrix.

We attribute the divergence in the results shown in Figure 4 to the different reliability of the examined data in the two cases. While each of the variables from the superC matrix are descriptive of pure resolved resonances, every single variable from the 1D ^{1}H NMR spectral matrix can be representative of multiple resonances at the same time (if signal overlap occurs) in addition to the instrumental noise. Due to the high signal overlap and instrumental noise, the contribution of the individual resonances within the corresponding variables is diluted. For this reason, sample differences would be less efficiently distinguished when resonances are severely overlapped, as in 1D ^{1}H NMR raw data, than when each variable is representing one single resolved resonance, as in the superC matrix.

On the other hand, the 1D ^{1}H NMR data contains a significant amount of variables only descriptive of instrumental noise. This is not the case for the superC matrix, since most of the instrumental noise is set apart in the residual matrix (E in Figure 1) during the MCR–ALS analysis. This is the reason why the PCA analysis of the superC matrix was much less affected by the noise contributions than the analysis of raw 1D ^{1}H NMR data.

Therefore, for a comprehensive untargeted NMR metabolomics analysis, we propose performing the chemometric study on the data set with the MCR–ALS resolved concentrations instead of on the corresponding 1D ^{1}H NMR spectral raw data.

Tissue specificity of ACR exposure in ZF. As shown above, the PCA analysis of the superC matrix detected two main sources of data variance, one related to the tissue (PC1 in Figure 4B) and another to the ACR exposure (PC2 in Figure 4B). In order to determine the significance of these two factors, and to evaluate whether there exists any interaction between the two factors, the ASCA method was applied and the results are shown in Figure 5.

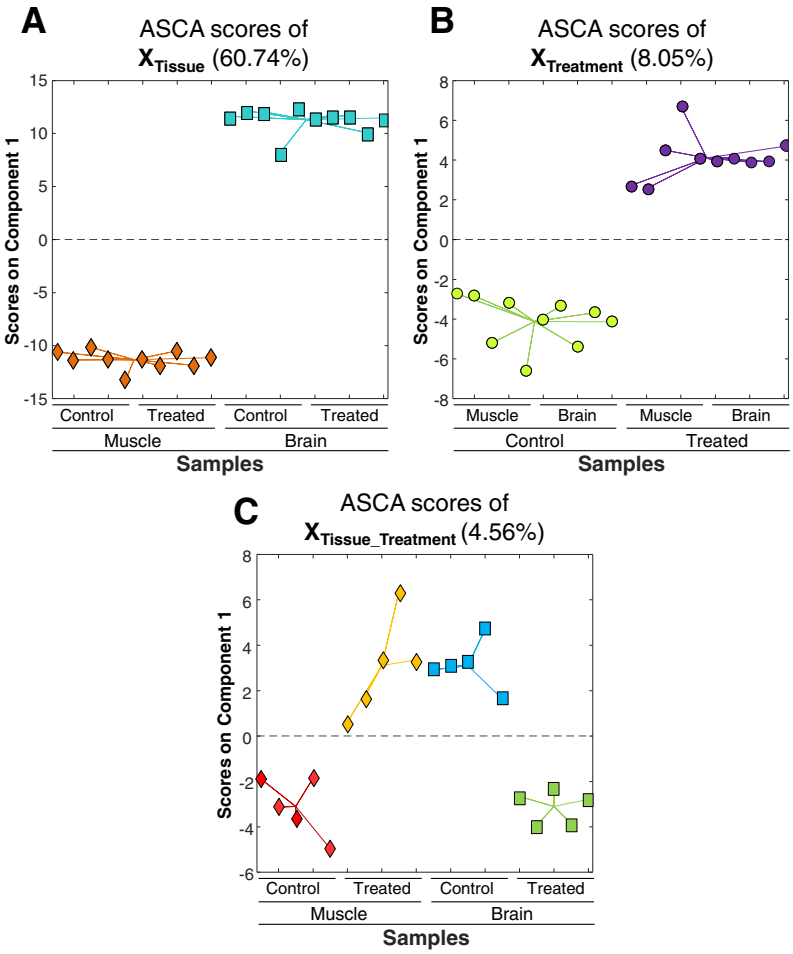


Figure 5. ASCA scores of (A) X_{Tissue}, (B) $X_{treatment}$, and (C) $X_{Treatment\_Tissue}$. In the legends, MC and MT are muscle extracts from control and ACR-treated ZF; BC and BT are the corresponding brain extract counterparts.

The effects of the two studied factors is clearly visualized in Figure 5. As in the PCA analysis, metabolic differences where more significant for the tissue type (60.74%, Figure 5A) rather than for the ACR exposure itself (8.05%, Figure 5B). In addition, the possible interaction between these two factors

(type of tissue and ACR exposure) can also be detected accounting for a 4.56% of the total data variance (Figure 5C).

These results were confirmed after the calculation of the statistical significances of the effects for the two factors and of their interaction using a permutation test.[34] In all the cases, the significance of the tests were high (p values = 0.001).

Metabolic discrimination of the tissues. In order to select the metabolites characteristic for the two tissues, a PLS–DA analysis was performed comparing the 222 features for the two type of tissues (muscle and brain) of control samples (PLS–DA #1 in Figure 3). As a result, 158 features were found to discriminate between muscle and brain. By looking at the spectral profiles associated with these features (S^T in Figure 1, and step B5 in Figure 2) studying the spectroscopic parameters (δ_H, J, multiplicity) that define these resonances, and after consulting these parameters on a NMR spectral library (such as HMDB[35] or BRMB[36]), it was possible to assign some of them to metabolites. It must be noted that this level of spectroscopic investigation cannot be performed by more conventional approaches. In these approaches, resonances are assigned according to their chemical shifts values (δ), whereas the other associated spectroscopic parameters cannot be easily determined due to the resonance overlapping. Moreover, since the data were auto–scaled, even the slightest variations among tissues can be detected. The resonances most distinctive for brain extracts were for γ–aminobutyric acid (GABA), acetic acid, N–acetyl–L–aspartic acid, L–glutamic acid, citric acid, carnosine, phosphothreonine, and some other resonances from lipidic compounds. On the other hand, creatine, L–histidine, taurine, glycerophosphorylcholine, L–alanine, fumaric acid, L–tyrosine, D–glucose, glycogen, UDP, and AXP were significantly more concentrated in muscle extracts than in brain extracts. In other words, brain samples characteristic compounds (compared to muscle samples) include neurotransmitters, neuronal osmolytes, and fats; in muscle, the characteristic metabolites (compared to brain samples) were those involved in its functioning, structure, redox, and energy metabolism.[37]

Metabolic fingerprint due to ACR treatment. The metabolic biomarkers of ACR exposure in muscle and in brain were determined by performing a PLS–DA analysis on each separated data subset of the superC matrix by comparing the control and treated samples related to the same tissue (PLS–DA #2 and #3 in Figure 3). Resulting from these two PLS–DA analyses, several features were detected to be significant in each of the two tissues: 93 features in the muscle subset and 79 features in the brain subset. After the spectral investigation of these significant features, we concluded that in the muscle tissue samples, ACR exposure caused an increase of the pools of the following metabolites: ACR, L–leucine, L–isoleucine, L–valine, L–lysine, GABA, phosphothreonine, methionine–sulfoxide, D–glucose, UDP, L–tyrosine and L–phenylalanine. In contrast, the pools of L–alanine, L–glutamic acid, glutathione, and glycerol were diminished due to ACR exposure. In the brain tissue samples, the ACR exposure increased significantly the pools of free ACR, L–valine, N–acetyl–S–(carbamoylethyl)–L–cysteine (AAMA), and methionine–sulfoxide. Conversely, the pools of carnosine, L–glutamic acid, glutathione, and NAD^+ were reduced due to this ACR treatment.

All these metabolic alterations reflect that ACR depleted the pools of intracellular glutathione by reacting with it. As a product of this reaction, in brain, ACR was accumulated as N–acetyl–S–(carbamoylethyl)–L–cysteine (AAMA). In addition, glutathione depletion caused a reduction of the organism oxidative stress defenses, which elicited the appearance of new oxidized species, such as methionine–sulfoxide pools. This enhanced oxidative stress may have caused the alteration of the amino acid pools in muscle as well, probably to be used for protein synthesis, in agreement with activation of protein synthesis of thioredoxins by ACR exposure.[38]

Biomarker discovery. Spectral features S^T associated with the same metabolite are strongly correlated, as they can be observed for instance in the heatmap plots shown in supplementary Figure S2. Thus, this fact can be used as a guide during the resonance assignment step. However, not all the selected features could be assigned to metabolites because their resonances are from compounds not included in metabolomics libraries.[39-40] In order to provide a more comprehensive description of the effects of ACR exposition on zebra fish, we have investigated unassigned features that increased due to

the ACR exposure. In particular, we focused on the features 78, 79, and 96 shown in Figure 6C. These 3 features were significantly important in muscle extracts of ACR treated zebrafish (Figure 6D-F). The reconstructed 1D ^{1}H NMR spectrum with these 3 features revealed two symmetric multiplets, centered at 2.60 and at 2.86 ppm, that resemble two double triplets (Figure 6C). This observed coupling is in agreement with the spin system of a $RSCH_2CH_2R'$ moiety. Since this spin system is also found in AAMA, we propose that this molecule is a reaction product from ACR. In humans, several ACR–derived compounds in addition to AAMA have been detected after oral exposition to ACR[41].

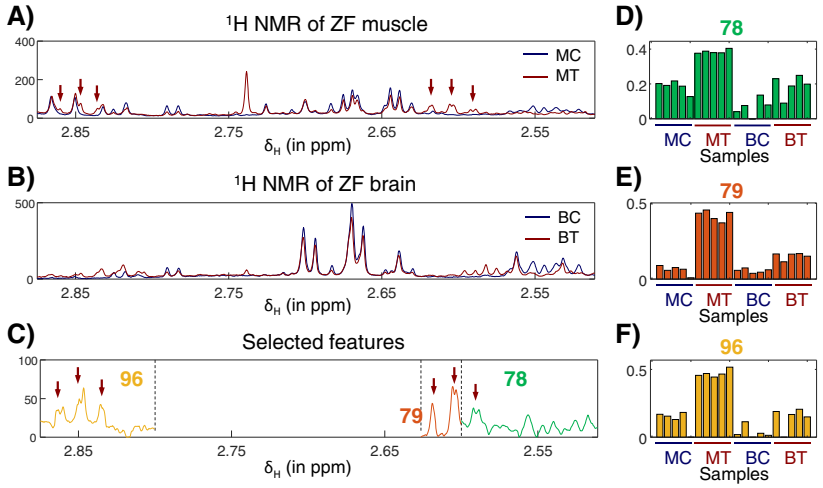


Figure 6. A-B) Representative 1D ^{1}H NMR spectra of muscle (A) and (B) brain extracts. 1D ^{1}H NMR spectra from control ZF are in blue, while spectra from treated ZF are coloured in red. C) Reconstructed 1D ^{1}H NMR with features 78, 79 and 96. D-F) Concentration profiles for features 78, 79 and 96. Arrows point the resonances from the two triple doublets centered at 2.60 and 2.86 ppm.

Hence, in this work, we detected two metabolites resulting from the ACR transformation inside the organism, AAMA and the 'ACR-derivative'. AAMA was found more prevalent in the brain of treated ZF, while the 'ACR-derivative' was more prominent in the corresponding muscle extracts. This tissue specificity may reflect the high reactivity of ACR in the two metabolically different tissues.

Method considerations. The proposed MCR-ALS-based chemometric methodological pipeline has been proven to be very useful for both metabolic fingerprint and biomarker discovery. Noteworthy, some considerations regarding this pipeline need to be made.

First, one of the most important steps in this pipeline is to set the limits of the NMR resonances in every window to be analyzed by MCR–ALS. In this study, we have chosen the window limits by visual inspection (see methods), relying on the user's spectroscopic knowledge avoiding to split resonances between two segments. However, more automatable alternative approaches could be designed. For instance, the 1D ^{1}H NMR dataset could have been examined with an intelligent bucketing procedure as in Sousa *et al.* (2013) [42], and MCR–ALS can be applied separately then to every bucket.

Second consideration is that MCR–ALS is a robust chemometric method which estimates the concentrations (C) associated with the same resonances (S^T) in all the investigated samples. However, this chemometric approach should not be used with an NMR dataset presenting inconsistencies between spectra (e.g., derived from incomplete shimming or phasing). Apart from that, the MCR–ALS can be used with 1D ^{1}H NMR data from any type of samples (e.g., extracts, serum, CSF) as long as the resonances are aligned across all the spectra. In metabolomics studies in the lab, where robust sample preparation methods are used[43], chemical shifting variations among the spectra can be usually corrected with simple resonance alignment methods, such as *icoshift*[30]. Also, for samples with complex background signals, such as serum samples, it may be recommended to add extra components in the MCR–ALS analyses to account for these background signals.

And as a third consideration, resonances presenting the same variance across samples and that are found within the same window will be resolved together, even if they are from different metabolites. Although this particularity can be seen as a limitation of the method, the resolved components will still be very useful and informative. The resolved spectral data will be much simpler than the original data (as all the other resonances presenting different variance will appear in other components) facilitating the resonance assignment. Moreover, the grouping of resonances in the same component suggests that the corresponding compounds are metabolically-related (since they present the same metabolic variance), resulting very useful for the biological interpretation.

CONCLUSIONS

The chemometric methodological pipeline presented here takes advantage of the rapid and powerful performance of the MCR–ALS chemometrics method to blindly resolve the concentrations and resonances from untangled metabolites, which allows the identification of the spectral features characteristic of the metabolomic system under study.

In this work we have proved that this methodology simplifies the metabolic analysis of 1D ^{1}H NMR spectra from complex biological extracts and that it can even be used for biomarker discovery. Moreover, the presented methodology is more sensitive to small metabolic variations than other chemometric methods, which do not take into account the prominent resonance overlapping in 1D ^{1}H NMR data.

In consequence, the proposed methodology can accelerate the discovery of unknown metabolites, and it has the potential to facilitate the study of large NMR datasets, such as those analyzed in cohort metabolomics studies[43].

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

NMR Experimental Parameters for 2D NMR spectra (PDF)

Supplementary results (PDF)

Table S1. NMR assignments (XLSX)

AUTHOR INFORMATION

Corresponding Author

*e-mail: francesc.puigcastellvi@agroparistech.fr

Tel: (+33) 01 44 08 18 43

Present Addresses

§NMR Facility, IQAC–CSIC, Jordi Girona 18–26, 08034 Barcelona

†Department of Environmental Chemistry, IDAEA–CSIC, Jordi Girona 18–26, 08034 Barcelona

‡Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 75005, Paris, France

Department of Biological Chemistry, IQAC–CSIC, Jordi Girona 18–26, 08034 Barcelona

Author Contributions

B.P. and D.R. designed the experiments. M.C. and Y.P. performed the experiments. F.P.-C. performed the chemometric analysis. All authors have contributed in the manuscript writing and have given approval to its final version.

Funding sources

The research leading to these results has received funding from the NATO SfP project MD.SFPP 984777.

Notes

The authors declare no competing interests.

ABBREVIATIONS

1D ^{1}H NMR, one-dimensional proton Nuclear Magnetic Resonance; AAMA, N-acetyl-S-(carbamoylethyl)-L-cysteine; ACR, acrylamide; ANOVA, ANalysis Of VAriance; ASCA; ANOVA–Simultaneous Component Analysis; AXP, adenosine nucleotides; BATMAN, **Bayesian automated metabolite analyzer for NMR**; C, concentrations matrix in the MCR-ALS analysis; D, input matrix in the MCR-ALS analysis; DSS, **2,2-Dimethyl-2-silapentane-5-sulfonate**; GABA, gamma-aminobutyric acid; GUI, Graphical User Interface; MCR-ALS, Multivariate Curve Resolution – Alternating Least

Squares; NMR, Nuclear Magnetic Resonance; NOESY, **Nuclear Overhauser Effect SpectroscopY**; PC1, first Principal Component; PC2, second Principal Component; PCA, Principal Component Analysis; PLS–DA, Partial Least Squares – Discriminant Analysis; PQN, Probabilistic Quotient Normalization; SCA, Simultaneous Component Analysis; S^T, spectrums matrix in the MCR–ALS analysis; SVD, Single Value Decomposition; UDP, uridine diphosphate.

REFERENCES

1. Boiteau, R. M.; Hoyt, D. W.; Nicora, C. D.; Kinmonth–Schultz, H. A.; Ward, J. K.; Bingol, K., Structure Elucidation of Unknown Metabolites in Metabolomics by Combined NMR and MS/MS Prediction. *Metabolites* 2018, *8* (1), 8.
2. Dona, A. C.; Kyriakides, M.; Scott, F.; Shephard, E. A.; Varshavi, D.; Veselkov, K.; Everett, J. R., A guide to the identification of metabolites in NMR–based metabonomics/metabolomics experiments. *Comput. Struct. Biotec.* 2016, *14*, 135–153.
3. Johnson, C. H.; Ivanisevic, J.; Siuzdak, G., Metabolomics: beyond biomarkers and towards mechanisms. *Nat. Rev. Mol. Cell Biol.* 2016, *17*, 451.
4. Bingol, K., Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods. *High-throughput* 2018, *7* (2), 9.
5. Cappello, T.; Maisano, M.; Mauceri, A.; Fasulo, S., ^{1}H NMR–based metabolomics investigation on the effects of petrochemical contamination in posterior adductor muscles of caged mussel Mytilus galloprovincialis. *Ecotox. Environ. Safe.* 2017, *142* (Supplement C), 417–422.
6. Nagato, E. G.; D'eon, J. C.; Lankadurai, B. P.; Poirier, D. G.; Reiner, E. J.; Simpson, A. J.; Simpson, M. J., ^{1}H NMR–based metabolomics investigation of *Daphnia magna* responses to sub–lethal exposure to arsenic, copper and lithium. *Chemosphere* 2013, *93* (2), 331–337.

7. Puig-Castellví, F.; Pérez, Y.; Piña, B.; Tauler, R.; Alfonso, I., Comparative analysis of ^{1}H NMR and ^{1}H–^{13}C HSQC NMR metabolomics to understand the effects of medium composition in yeast growth. *Anal. Chem.* 2018, *90* (21), 12422–12430.

8. Tomassini, A.; Vitalone, A.; Marini, F.; Praticò, G.; Sciubba, F.; Bevilacqua, M.; Delfini, M.; Di Sotto, A.; Di Giacomo, S.; Mariani, P.; Mammola, C. L.; Gaudio, E.; Miccheli, A.; Mazzanti, G., ^{1}H NMR-Based Urinary Metabolic Profiling Reveals Changes in Nicotinamide Pathway Intermediates Due to Postnatal Stress Model in Rat. *J. Proteome Res.* 2014, *13* (12), 5848–5859.

9. Ruan, L. Y.; Fan, J. T.; Hong, W.; Zhao, H.; Li, M. H.; Jiang, L.; Fu, Y. H.; Xing, Y. X.; Chen, C.; Wang, J. S., Isoniazid-induced hepatotoxicity and neurotoxicity in rats investigated by ^{1}H NMR based metabolomics approach. *Toxicol. Lett.* 2018, *295*, 256–269.

10. Fathi, F.; Oskouie, A. A.; Tafazzoli, M.; Naderi, N.; Sohrabzedeh, K.; Fathi, S.; Norouzinia, M.; Rostami Nejad, M., Metabonomics based NMR in Crohn's disease applying PLS-DA. *Gastroenterol. Hepatol. Bed Bench* 2013, *6* (Suppl 1), S82–S86.

11. Bro, R.; Smilde A. K., Principal Component Analysis. *Anal. Methods* 2014, *6* (9), 2812–2831.

12. Smilde, A. K.; Jansen, J. J.; Hoefsloot, H. C. J.; Lamers, R.-J. A. N.; van der Greef, J.; Timmerman, M. E., ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. *Bioinformatics* 2005, *21* (13), 3043–3048.

13. Barker, M., & Rayens, W., Partial least squares for discrimination. *J. Chemom.* 2003, *17* (3), 166–173.

14. Tauler, R.; Kowalski, B.; Fleming, S., Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial-Process. *Anal. Chem.* 1993, *65*.

15. Ebrahimi, P.; Larsen, F. H.; Jensen, H. M.; Vogensen, F. K.; Engelsen, S. B., Real-time metabolomic analysis of lactic acid bacteria as monitored by in vitro NMR and chemometrics. *Metabolomics* 2016, *12* (4), 77.

16. Winning, H.; Larsen, F. H.; Bro, R.; Engelsen, S. B., Quantitative analysis of NMR spectra with chemometrics. *J. Magn. Reson.* 2008*, 1,* 26–32.

17. Abdollahi, H.; Tauler, R., Uniqueness and rotation ambiguities in Multivariate Curve Resolution methods. *Chemometrics Intell. Lab. Syst.* 2011, *108* (2), 100–111.

18. Puig–Castellví, F.; Alfonso, I.; Tauler, R., Untargeted assignment and automatic integration of ^{1}H NMR metabolomic datasets using a multivariate curve resolution approach. *Anal. Chim. Acta* 2017, *964* (Supplement C), 55–66.

19. Karakach, T. K.; Knight, R.; Lenz, E. M.; Viant, M. R.; Walter, J. A., Analysis of time course ^{1}H NMR metabolomics data by multivariate curve resolution. *Magn. Reson. Chem.* 2009, *47* (S1), S105–S117.

20. Montoliu, I.; Martin, F. o.–P. J.; Collino, S.; Rezzi, S.; Kochhar, S., Multivariate Modeling Strategy for Intercompartmental Analysis of Tissue and Plasma ^{1}H NMR Spectrotypes. *J. Proteome Res.* 2009, *8* (5), 2397–2406.

21. Röhnisch, H. E.; Eriksson, J.; Müllner, E.; Agback, P.; Sandström, C.; Moazzami, A. A., AQuA: An Automated Quantification Algorithm for High–Throughput NMR–Based Metabolomics and Its Application in Human Plasma. *Anal. Chem.* 2018, *90* (3), 2095–2102.

22. Tardivel, P. J. C.; Canlet, C.; Lefort, G.; Tremblay–Franco, M.; Debrauwer, L.; Concordet, D.; Servien, R., ASICS: an automatic method for identification and quantification of metabolites in complex 1D ^{1}H NMR spectra. *Metabolomics* 2017, *13* (10), 109.

23. Cañueto, D.; Gómez, J.; Salek, R. M.; Correig, X.; Cañellas, N., rDolphin: a GUI R package for proficient automatic profiling of 1D ^{1}H–NMR spectra of study datasets. *Metabolomics* 2018, *14* (3), 24.

24. Hao, J.; Liebeke, M.; Astle, W.; De Iorio, M.; Bundy, J. G.; Ebbels, T. M. D., Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nat. Protoc.* 2014, *9* (6), 1416–1427.

25. Khakimov, B., Mobaraki, N., Trimigno, A., Aru, V., Engelsen, S. B., Signature Mapping (SigMa): An efficient approach for processing complex human urine 1H NMR metabolomics data. *Anal. Chim. Acta*, 2020, *In press*, 1–10, doi:10.1016/j.aca.2020.02.025.

2020,

26. Dearfield, K. L.; Abernathy, C. O.; Ottley, M. S.; Brantner, J. H.; Hayes, P. F., Acrylamide: its metabolism, developmental and reproductive effects, genotoxicity, and carcinogenicity. *Mutat. Res.-Rev. Genet.* 1988, *195* (1), 45–77.

27. Tareke, E.; Rydberg, P.; Karlsson, P.; Eriksson, S.; Törnqvist, M., Analysis of Acrylamide, a Carcinogen Formed in Heated Foodstuffs. *J. Agr. Food Chem.* 2002, *50* (17), 4998–5006.

28. Tepe, Y.; Çebi, A., Acrylamide in Environmental Water: A Review on Sources, Exposure, and Public Health Risks. *Expos. Health* 2017.

29. Duke, T. J.; Ruestow, P. S.; Marsh, G. M., The influence of demographic, physical, behavioral, and dietary factors on hemoglobin adduct levels of acrylamide and glycidamide in the general U.S. population. *Crit. Rev. Food Sci. Nutr.* 2018, *58* (5), 700–710.

30. Savorani, F.; Tomasi, G.; Engelsen, S. B., icoshift: A versatile tool for the rapid alignment of 1D NMR spectra. *J. Magn. Reson.* 2010, *202* (2), 190–202.

31. Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H., Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in ^{1}H NMR Metabonomics. *Anal. Chem.* 2006, *78* (13), 4281–4290.

32. Abdi, H., Singular Value Decomposition (SVD) and Generalized Singular Value Decomposition (GSVD). In *Encyclopedia of measurement and statistics*, Salkind, N. J., Ed. SAGE Publications: 2007; 907–912.

33. Jaumot, J.; de Juan, A.; Tauler, R., MCR–ALS GUI 2.0: New features and applications. *Chemometrics Intell. Lab. Syst.* 2015, *140*, 1–12.

34. Zwanenburg, G.; Hoefsloot, H. C. J.; Westerhuis J. A.; Jansen, J. J.; Smilde, A. K., ANOVA–principal component analysis and ANOVA–simultaneous component analysis: a comparison. *J. Chemometr.* 2011, *25*, 561–567.

35. 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.

36. Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z.; Nakatani, E.; Schulte, C. F.; Tolmie, D. E.; Kent Wenger, R.; Yao, H.; Markley, J. L., BioMagResBank. *Nucleic Acids Res.* 2008, *36* (suppl_1), D402–D408.

37. Shi, M.; Ellingsen, Ø.; Bathen, T. F.; Høydal, M. A.; Koch, L. G.; Britton, S. L.; Wisløff, U.; Stølen, T. O.; Esmaeili, M., Skeletal muscle metabolism in rats with low and high intrinsic aerobic capacity: Effect of aging and exercise training. *PLOS ONE* 2018, *13* (12), e0208703.38. Raldúa, D.; Casado, M.; Prats, E.; Faria, M.; Puig-Castellví, F.; Pérez, Y.; Alfonso, I.; Hsu, C.; Arick II, M. A.; Garcia-Reyero, N.; Ziv, T.; Ben-Lulu, S.; Admon, A.; Piña, B., Targeting redox metabolism: the perfect storm induced by acrylamide poisoning in the brain. *Sci. Rep.* 2020, *10*, 312.

39. da Silva, R. R.; Dorrestein, P. C.; Quinn, R. A., Illuminating the dark matter in metabolomics. *PNAS* 2015, *112* (41), 12549–12550.

40. Jones, O. A. H., Illuminating the dark metabolome to advance the molecular characterisation of biological systems. *Metabolomics* 2018, *14* (8), 101.

41. Kopp, E. K.; Dekant, W., Toxicokinetics of acrylamide in rats and humans following single oral administration of low doses. *Toxicol. Appl. Pharm.* 2009, *235* (2), 135–142.

42. Sousa, S. A. A.; Magalhães, A.; Ferreira, M. M. C., Optimized bucketing for NMR spectra: Three case studies. *Chemometrics Intell. Lab.Syst.* 2013, *122*, 93–102.

43. McHugh, C. E.; Flott, T. L.; Schooff, C. R.; Smiley, Z.; Puskarich, M. A.; Myers, D. D.; Younger, J. G.; Jones, A. E.; Stringer, K. A., Rapid, Reproducible, Quantifiable NMR Metabolomics: Methanol and Methanol: Chloroform Precipitation for Removal of Macromolecules in Serum and Whole Blood. *Metabolites* 2018, *8* (4), 93.

43. Trivedi, D. K.; Hollywood, K. A.; Goodacre, R., Metabolomics for the masses: The future of metabolomics in a personalized world. *New Horiz. Trans. Med.* 2017, *3* (6), 294–305.

.

MCR-ALS analysis of ^{1}H NMR spectra by segments to study the zebra fish exposure to acrylamide

Yolanda Pérez,[§] Marta Casado,[†] Demetrio Raldúa,[†] Eva Prats,[†] Benjamín Piña,[†] Romà Tauler,[†] Ignacio Alfonso, and Francesc Puig-Castellví,[†,‡,*]

§NMR Facility, IQAC-CSIC, Jordi Girona 18-26, 08034 Barcelona

†Department of Environmental Chemistry, IDAEA-CSIC, Jordi Girona 18-26, 08034 Barcelona

‡Department of Biological Chemistry, IQAC-CSIC, Jordi Girona 18-26, 08034 Barcelona

Corresponding author:

Francesc Puig-Castellví

E-mail address:

francesc.puigcastellvi@agroparistech.fr

Telephone:

(+33) 01 44 08 18 43

TABLE OF CONTENTS:

NMR Experimental Parameters for 2D NMR spectra

Pulse program noesygppr1d

Time domain 65536

Dummy scans 4

Scans 512

Sweep width 20 ppm

Acquisition time 3.277 s

Relaxation delay 4 s

Receiver gain 25.4

Dwell time 50.0 μs

Mixing time 0.01 s

Line broadening 0.3 Hz

Pulse program jresgpprqf

Time domain 8192 in F2 40 in F1

Dummy scans 16

Scans 64

Sweep width 16.62 ppm in F2 0.156 in F1

Acquisition time 0.493 s in F2 0.256 in F1

Relaxation delay 2.0 s

Receiver gain 25.4

Dwell time 60.2 μs

Line broadening 0.3 Hz in F2 0.3 Hz/SINE GB=0.1 in F1

Pulse program cosygpprqf

Time domain 8192 in F2 512 in F1

Dummy scans 16

Scans 16

Sweep width 15.96 ppm in F2 15.96 in F1

Acquisition time 0.513 s in F2 0.032 in F1

Relaxation delay 2 s

Receiver gain 25.4

Dwell time 62.67

Line broadening 0.3 Hz in F2 0.3 Hz in F1

Pulse program mlevgpphprzf

Time domain 8192 in F2 600 in F1

Dummy scans 8

Scans 32

Sweep width 15.96 ppm in F2 15.96 in F1

Acquisition time 0.513 s in F2 0.038 in F1

Relaxation delay 2 s

Receiver gain 203

Dwell time 62.67

Mixing time 0.06 s

Line broadening QSINE SSB=2 in F2 QSINE SSB=2 in F1

Pulse program hsqcetgpprsisp2.2

Time domain 3072 in F2 512 in F1

Dummy scans 64

Scans 48

Sweep width 15.96 ppm in F2 165.00 in F1

Acquisition time 0.193 s in F2 0.012 in F1

Relaxation delay 3 s

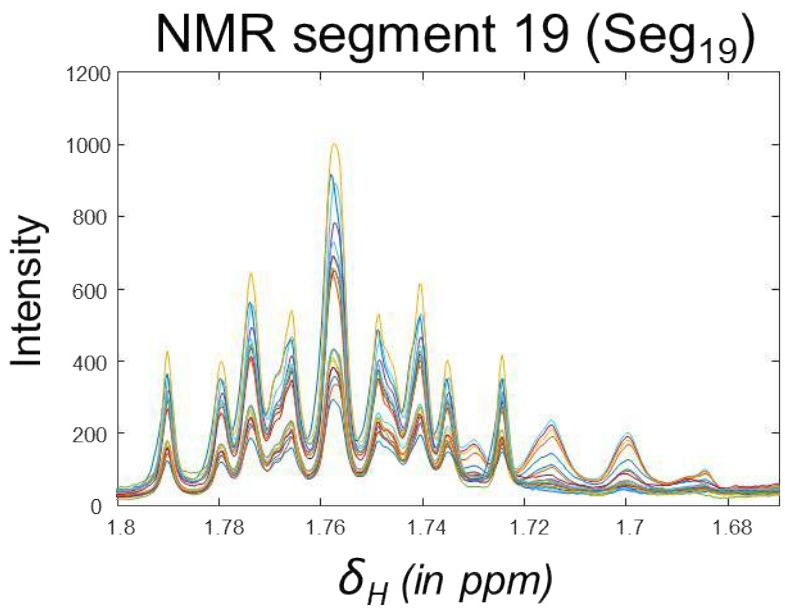
Receiver gain 203

Dwell time 62.67

Line broadening QSINE SSB=2 in F2 QSINE SSB=2 in F1

Supplementary Results

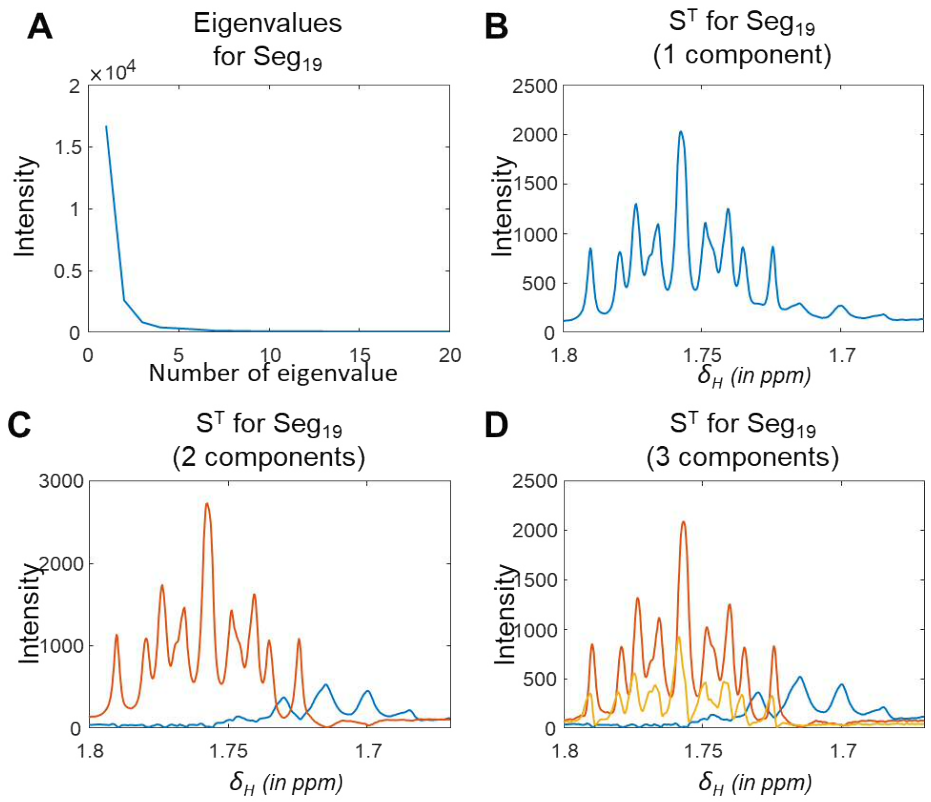
Selection of the optimal number of components in MCR-ALS



FigS1. Segment 19 (1.67–1.80 ppm).

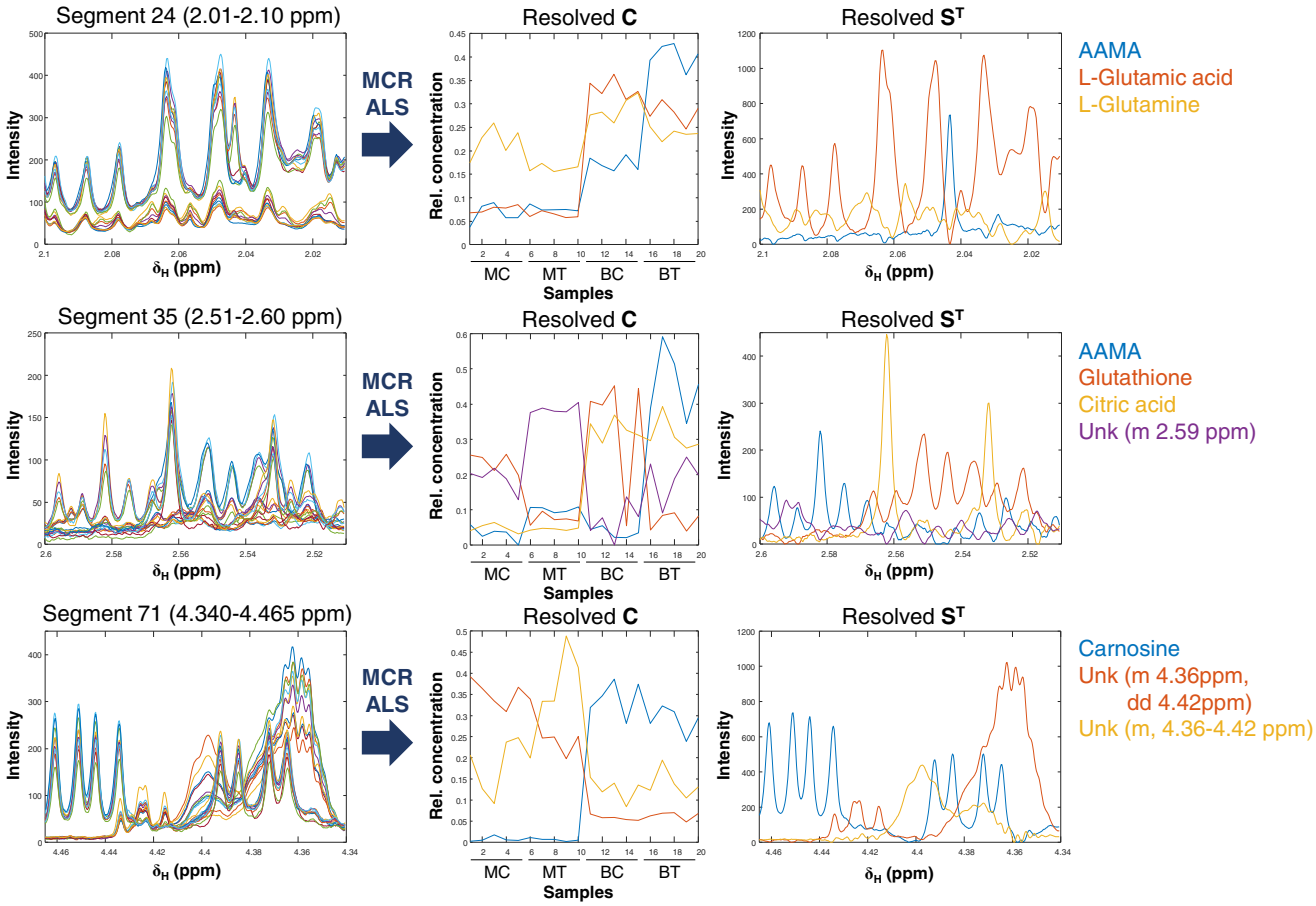
The optimal number of components to introduce in the MCR–ALS analysis can be found by using the Single Value Decomposition (SVD) method. By this method, the eigenvalues of the ^{1}H NMR segments are calculated. For instance, for segment 19 (Fig S1), the corresponding eigenvalues are given in FigS2A. In this figure, it can be observed that the first eigenvalues are larger than the later ones. It is assumed that the firsts eigenvalues are representative of meaningful data, while those close to zero are descriptive of noise data. So, it could be presumed that the number of eigenvalues representative of meaningful data is between 2 and 3, since the 4th eigenvalue is already too small.

If we analyze the NMR segment by using n=2 in the MCR–ALS analysis, two multiplets with good symmetry are resolved (FigureS2B). In case we use n–1 in the MCR–ALS analysis, then the two multiplets are resolved together (FigureS2C). Finally, by using n+1 as the number of components, one of the multiplets is split between two components (the red and the yellow in FigureS2D). Therefore, the optimal number of components to use for this segment is 2.



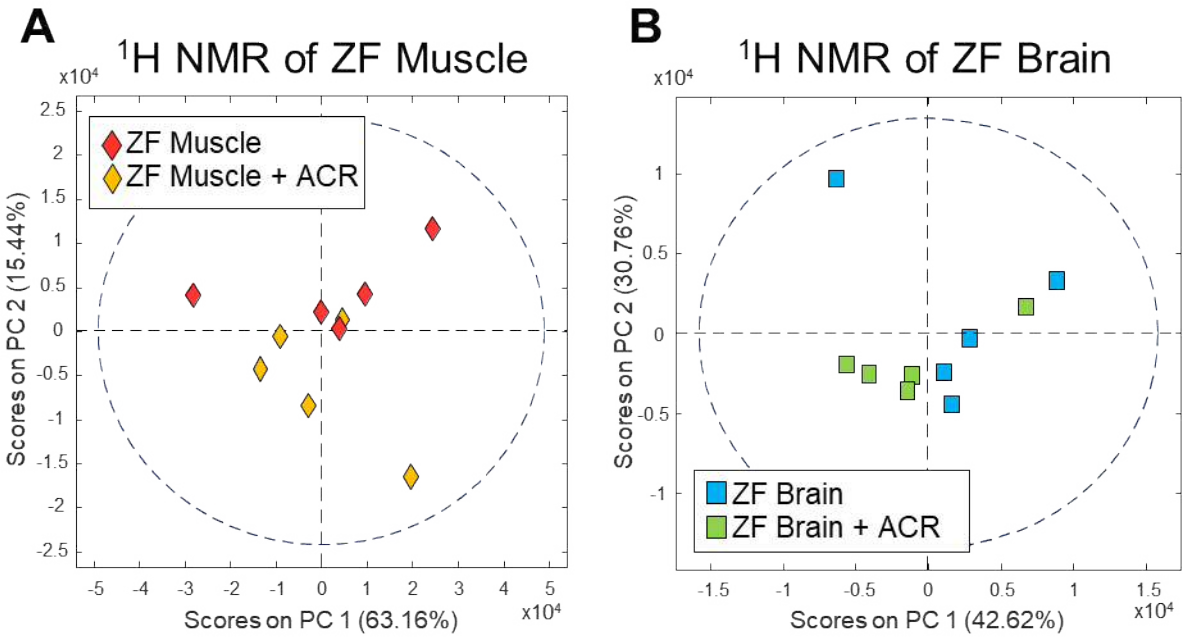
FigS2. Selecting the number of components by SVD.

Additional examples of the MCR-ALS resolution



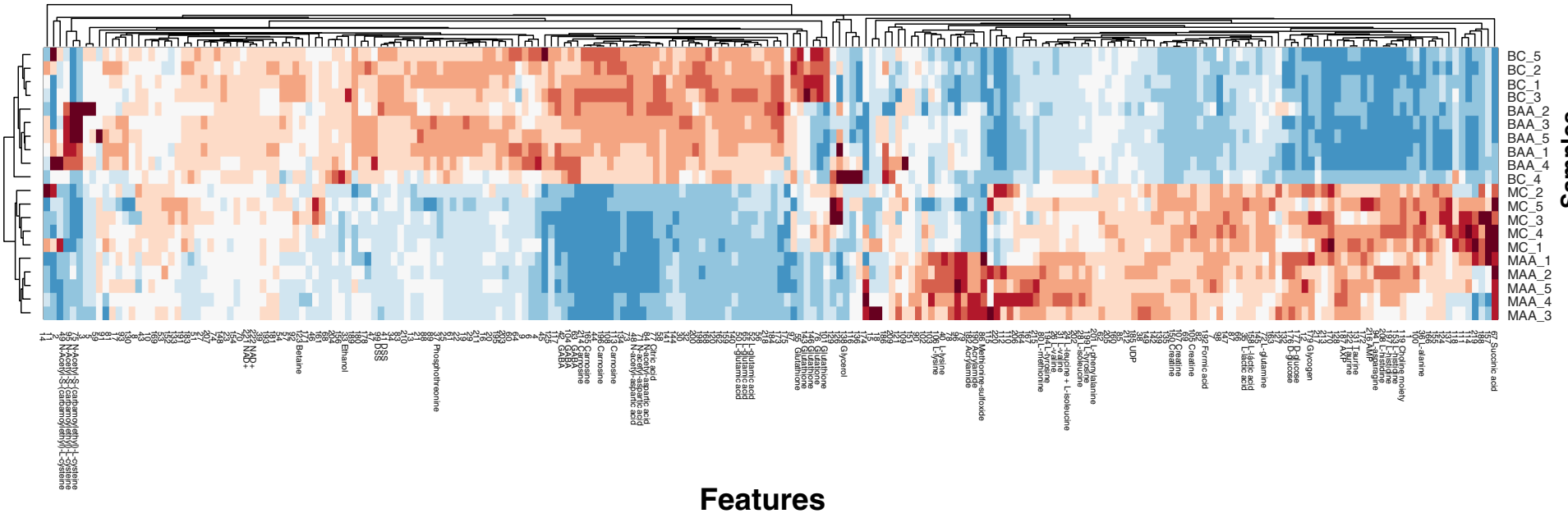
FigS3. MCR-ALS resolution for 3 ^{1}H NMR segments.

PCA of 1H NMR data (one tissue at a time)



FigS4. PCA on 1D ^{1}H NMR of extracts from ZF. A) Muscle extracts. B) Brain extracts. In the pre-processing, ^{1}H NMR spectra were mean-centered.

Heatmap analysis of the superC matrix

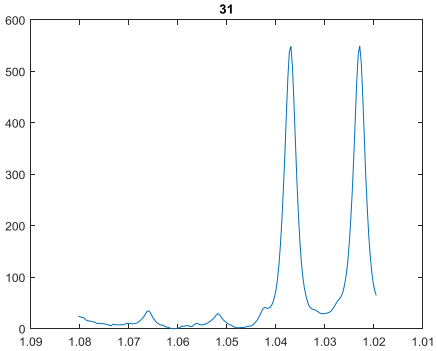


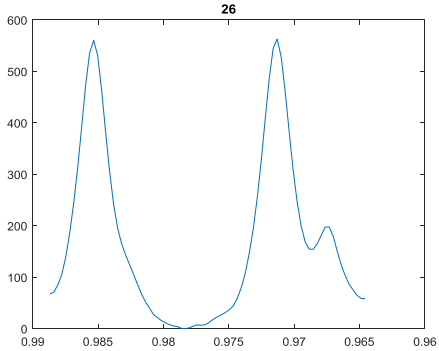
FigS5. Heatmap analysis of the superC matrix.

Assigned spectral features

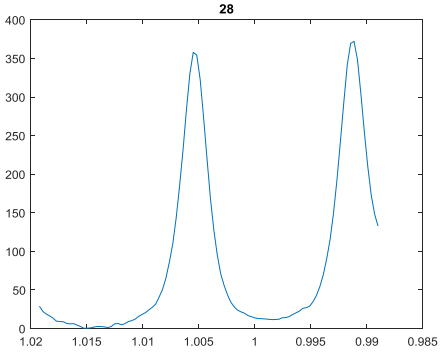
Numbers given in the figure titles refers to the features numbering (from 1 to 222). The concentration profiles associated to these spectral features are shown in the heatmap figure of FigS5. For spectral profiles containing contributions for more than one metabolite, the chosen name refers to the metabolite showing the maximal contribution.

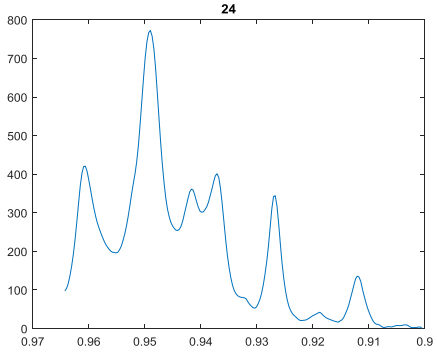
1) L–valine



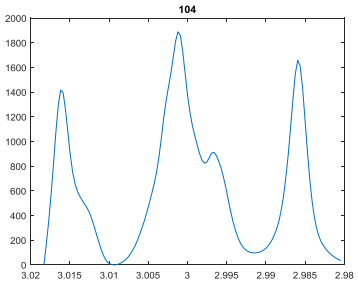


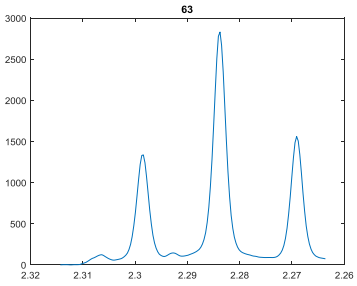
2) L–leucine + L–isoleucine

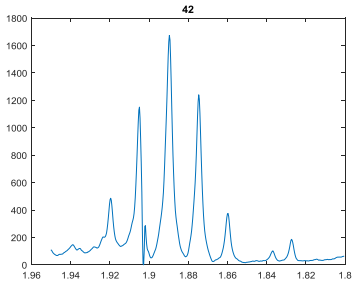




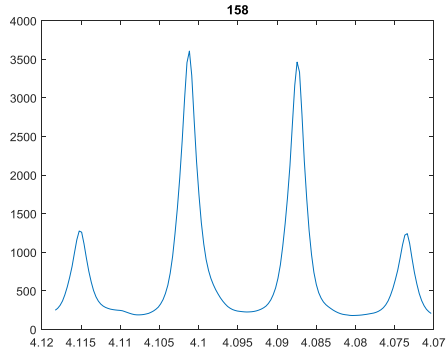
3) GABA

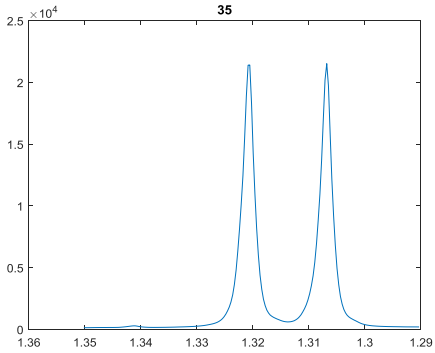




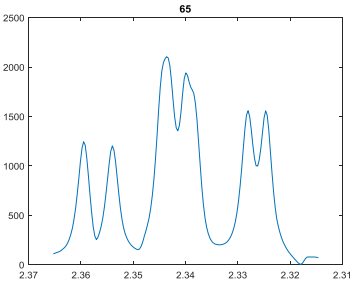


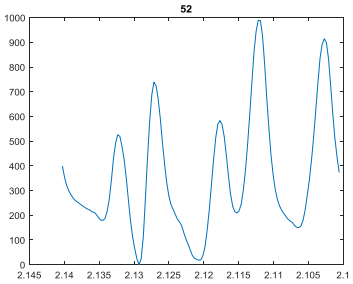
4) L-lactic acid

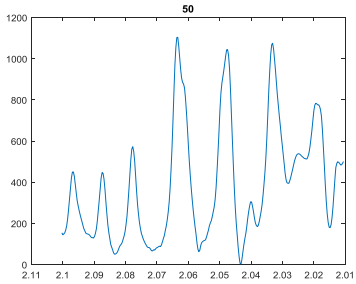




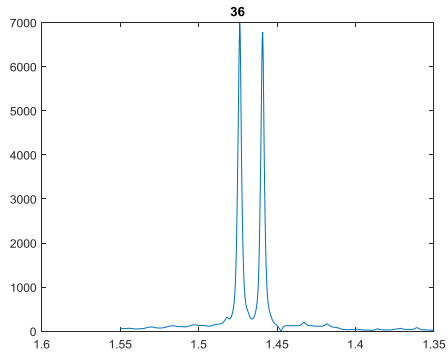
5) L-glutamic acid



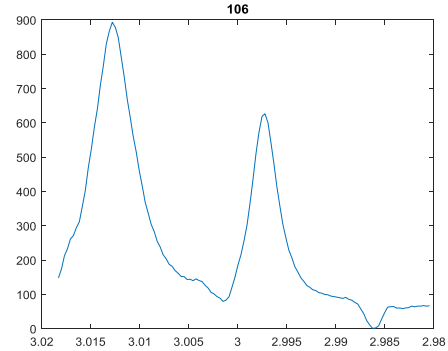


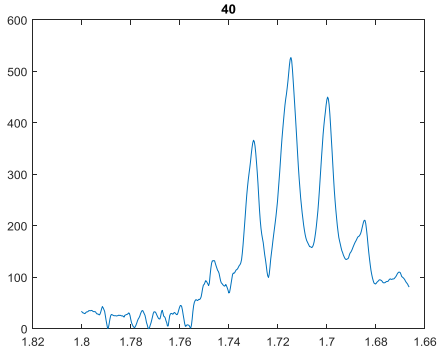


6) L-alanine

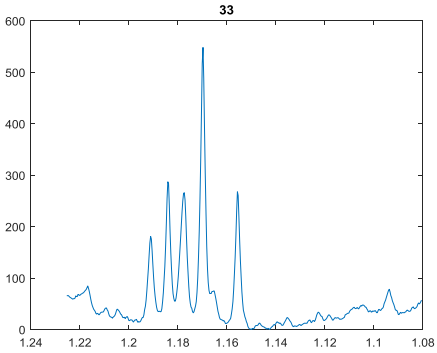


7) L-Lysine

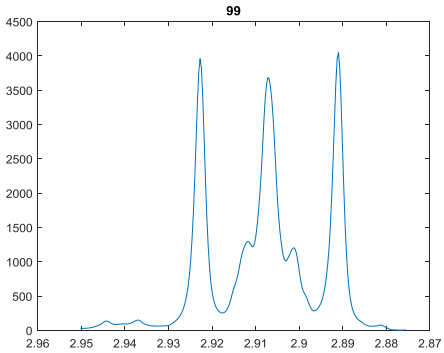


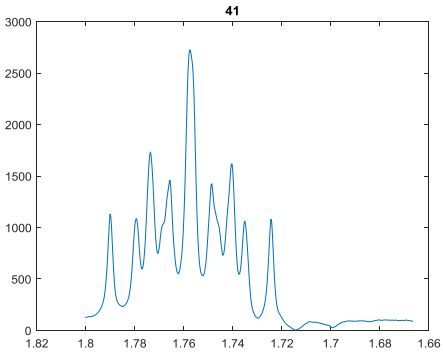


8) Ethanol

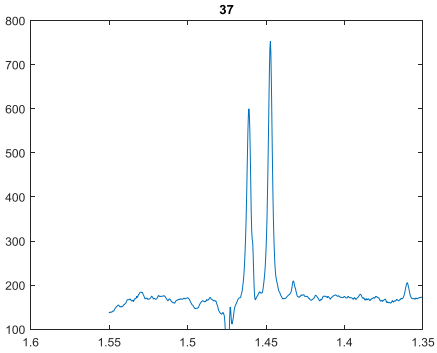


9) DSS

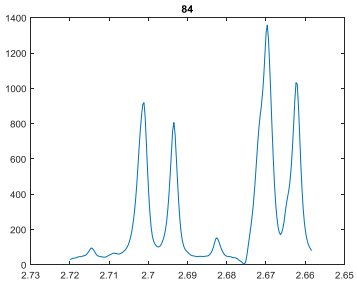


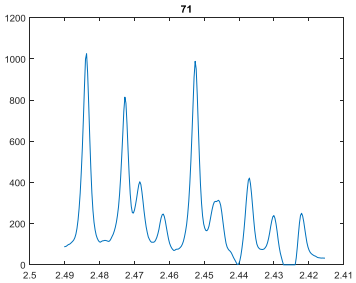


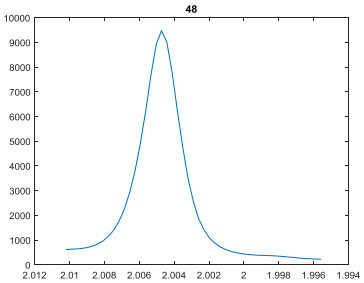
10) Phosphothreonine



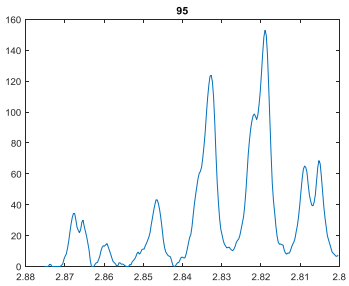
11) N-acetyl-aspartic acid

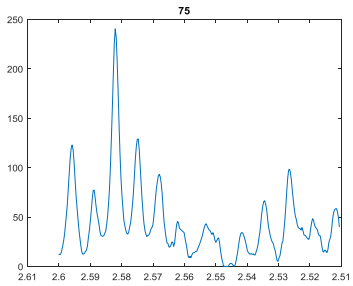


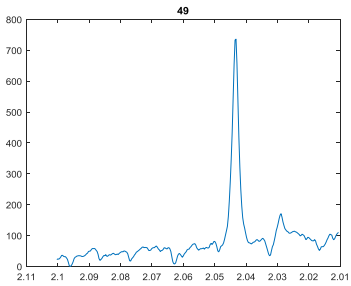




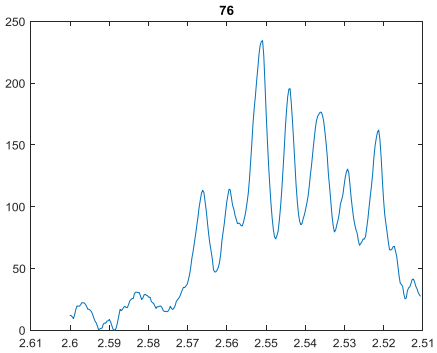
12) N-Acetyl-S-(carbamoylethyl)-L-cysteine

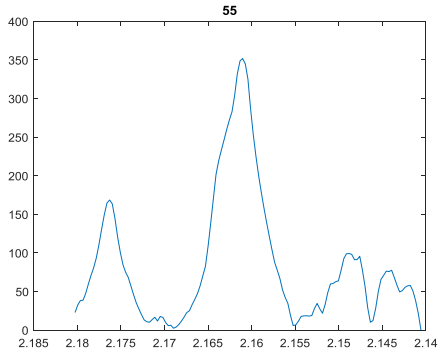


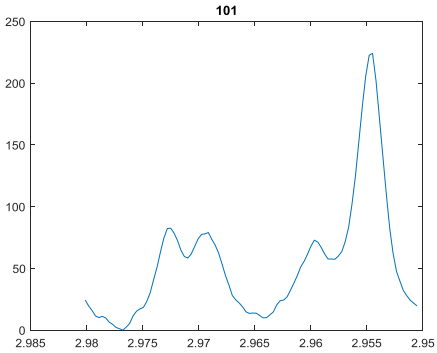


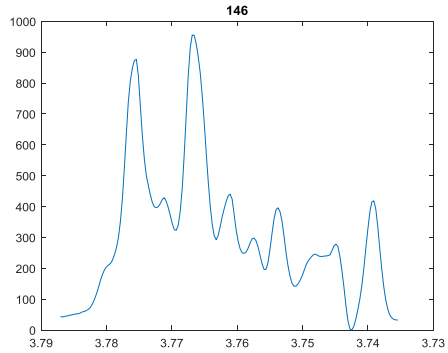


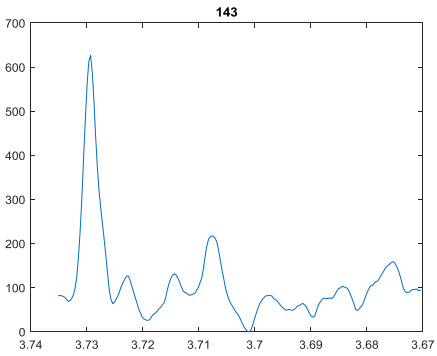
13) Glutathione



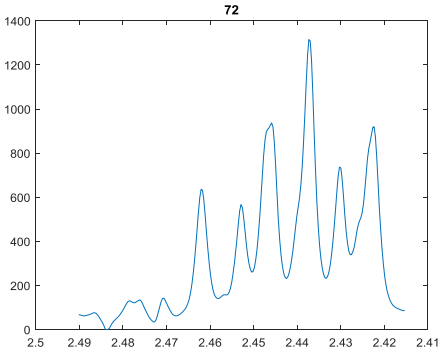




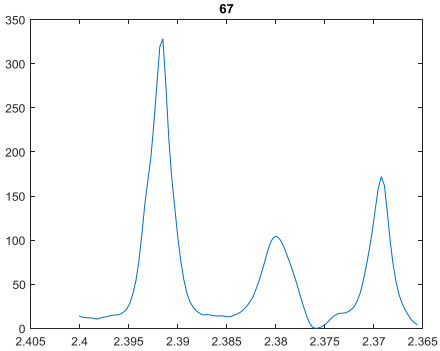




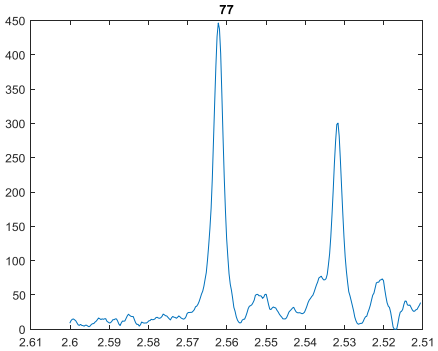
14) L-glutamine



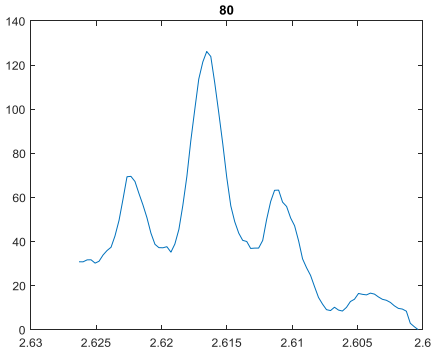
15) Succinic acid



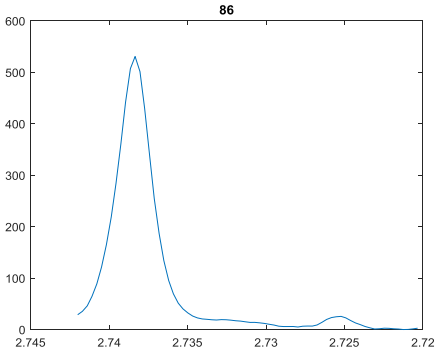
16) Citric acid



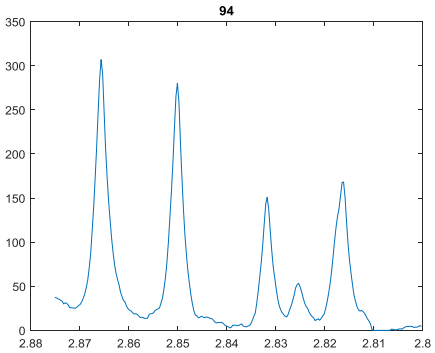
17) L-methionine



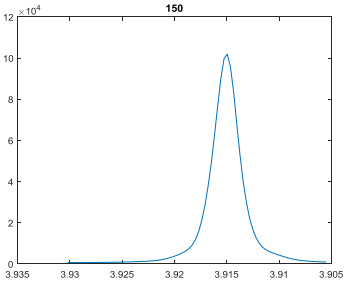
18) Methionine sulfoxide

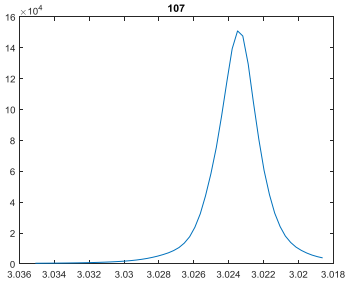


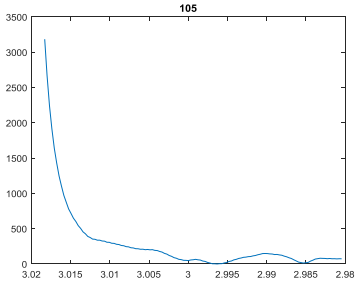
19) L-asparagine



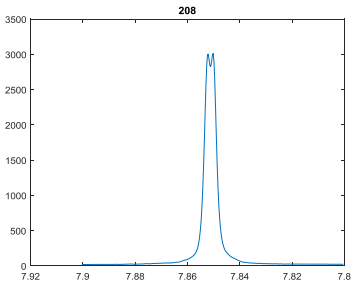
20) Creatine

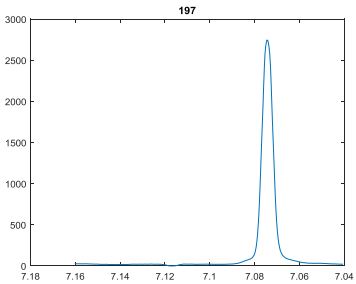


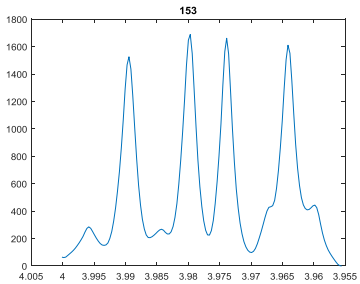




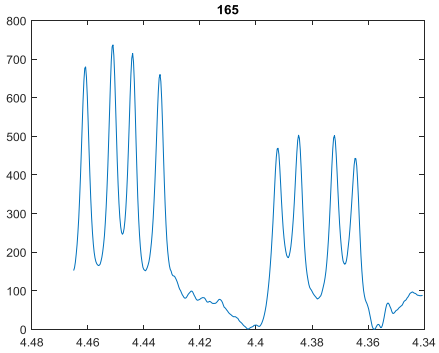
21) L-histidine

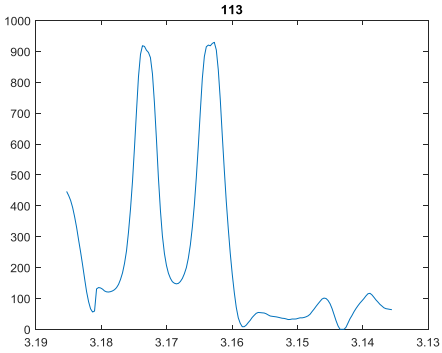


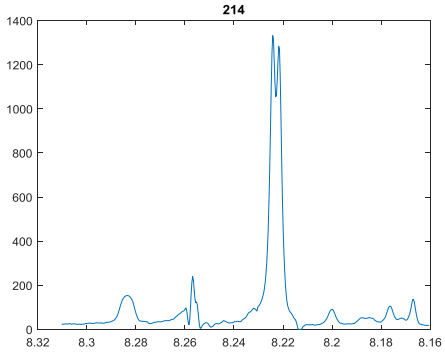


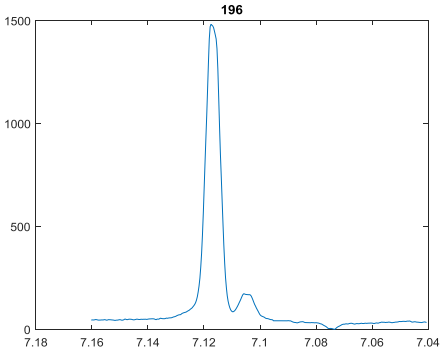


22) Carnosine

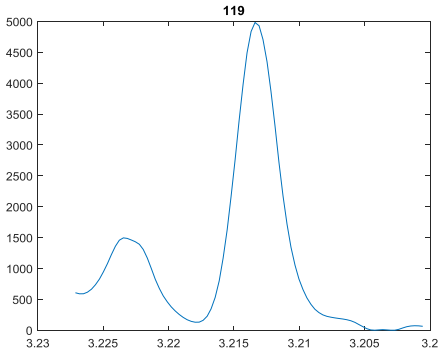




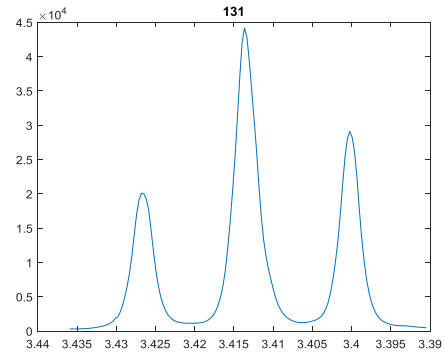


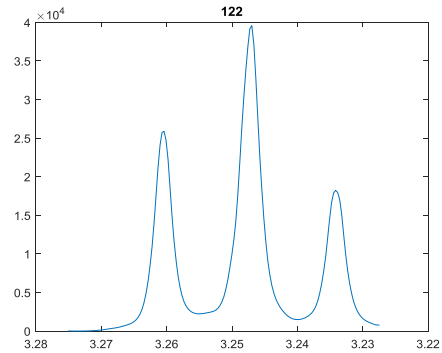


23) Choline moiety

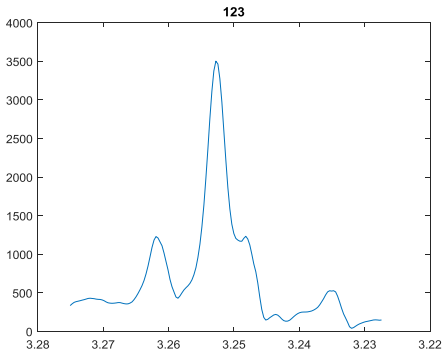


24) Taurine

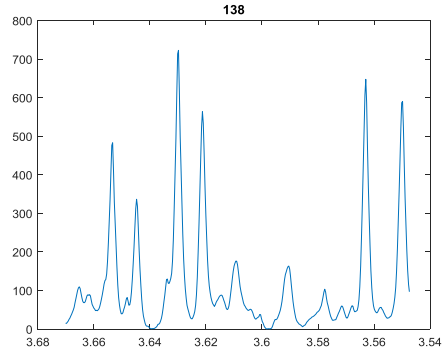




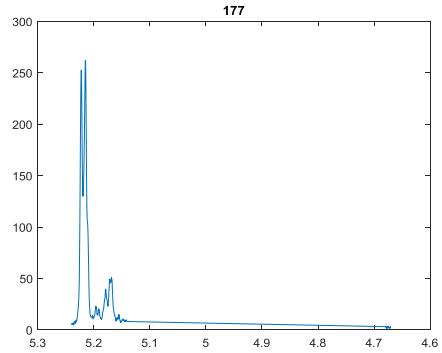
25) Betaine

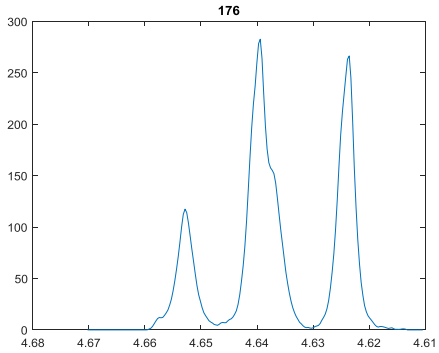


26) Glycerol

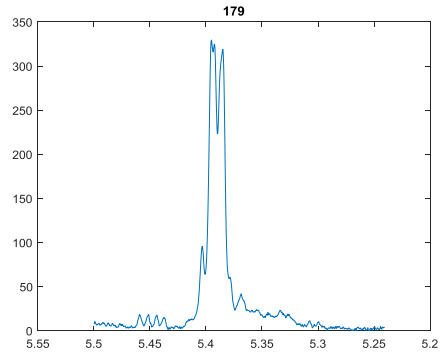


27) D-glucose

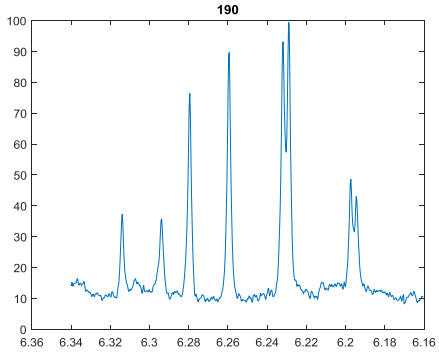


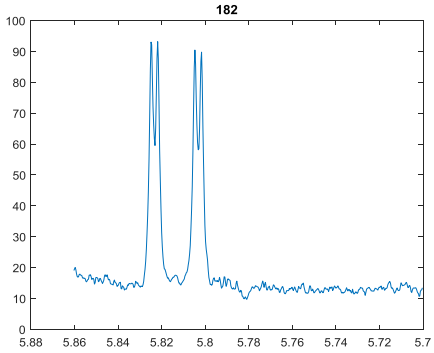


28) Glycogen

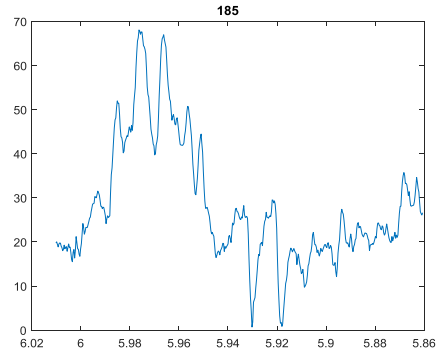


29) Acrylamide

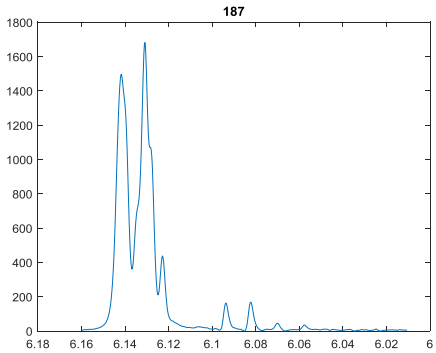




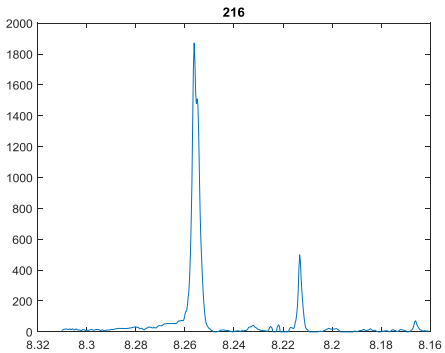
30) UDP



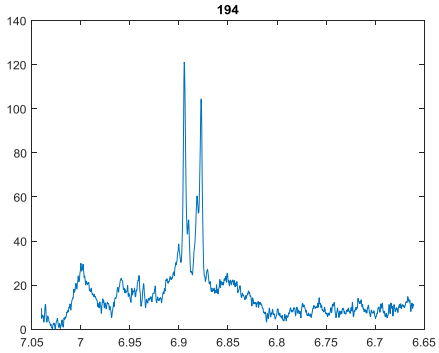
31) AXP

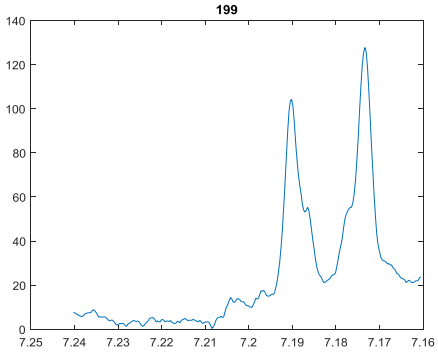


32) AMP

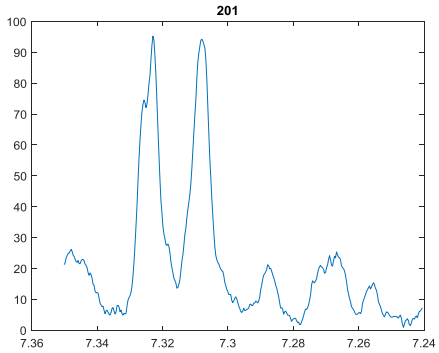


33) L-tyrosine





34) L-phenylalanine



35) NAD+