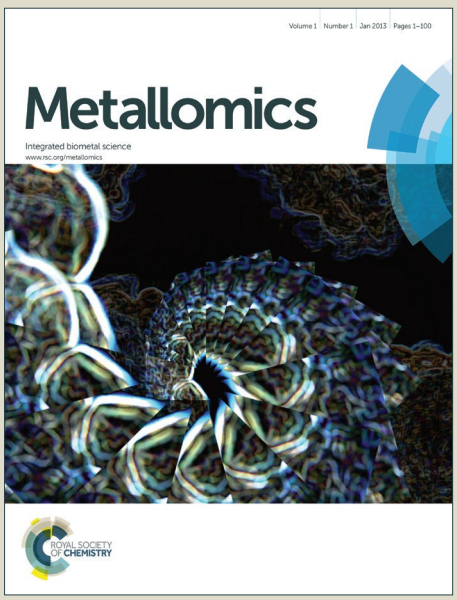
Metallomics

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/metallomics

SIGNIFICANCE STATEMENT

Copper containing fungicides are used to protect vineyards from fungal infections, but their residues on grapes can affect microorganisms living in vineyards, such as *Saccharomyces cerevisiae*. Here, a new LC-MS based metabolomics approach combined with chemometrics is used to explore Cu(II) toxicity in *S.cerevisiae* cultures. Control and Cu(II)-treated yeast samples are compared to identify what intracellular metabolites change due to a dose-response effect. The proposed untargeted metabolomics strategy revealed new Cu(II) targets for discovery and enabled biological interpretation of Cu(II) toxicological effects on *S.cerevisiae* metabolism.

LC-MS based metabolomics and chemometrics study of the toxic effects of copper on *Saccharomyces cerevisiae*

Mireia Farrés[a], Benjamí Piña[a], Romà Tauler[a]

[a] Institute of Environmental Assessment and Water Research (IDAEA), Spanish Council for Scientific Research (CSIC), Jordi Girona 18-26, 08034 Barcelona, Spain.

Abstract

Copper containing fungicides are used to protect vineyards from fungal infections. Higher residues of copper on grapes at toxic concentrations are a potentially toxic and affect the microorganisms living in vineyards, such as *Saccharomyces cerevisiae*. In this study, the response of metabolic profiles of *S.cerevisiae* at different concentrations of copper sulphate (control, 1 mM, 3 mM and 6 mM) were analysed by liquid chromatography coupled to mass spectrometry (LC-MS) and multivariate curve resolution-alternating least squares (MCR-ALS) using an untargeted metabolomics approach. Peak areas of the MCR-ALS resolved elution profiles in control and in Cu(II)-treated samples were compared using partial least squares regression (PLSR) and PLS-discriminant analysis (PLS-DA), and the intracellular metabolites best contributing to samples discrimination were selected and identified. Fourteen metabolites showed significant concentration changes upon Cu(II) exposure, following a dose-response effect. The observed changes were consistent with the expected effects of Cu(II) toxicity, including oxidative stress and DNA damage. This research confirmed that LC-MS based metabolomics coupled to chemometric methods are a powerful approach for discerning metabolomics changes in *S.cerevisiae* and for elucidating modes of toxicity of environmental stressors, including heavy metals like Cu(II).

1. Introduction

Copper (Cu(II)) containing fungicides have been used for more than one century in Europe on agricultural soils, such as vineyard soils. These fungicides have been used to protect crops from fungal infections such as downy mildew by *Plasmopara viticola* since the end of the 19th century. Copper is a potentially toxic metal, which, at low concentrations, can act as an essential micronutrient for microbial growth. At toxic concentrations, Cu(II) interacts with cellular nucleic acids and enzyme active sites, although a principal initial site of Cu(II) action is considered to be at plasma membrane. Thus, exposure of fungi and yeasts to elevated Cu(II) concentrations can lead to a rapid decline in membrane integrity [1]. Total Cu(II) concentrations in such soils can affect toxicologically to microorganisms living in vineyards. Cooper-based agrochemicals should not be toxic, under normal circumstances, but intensive and long term use of these fungicides has increased Cu(II) concentrations in soils significantly [2]. The use of copper formulates in biological vineyards has caused high levels in copper residues on grapes causing in some causes slow or stuck fermentations [3]. Several studies have been published about Cu (II) toxicity microorganisms [4-7] and specifically the effects of copper on *Saccharomyces cerevisiae* (yeast) have been the focus of many of these studies [1, 2, 8-11], although metabolomics studies have not been published yet.

The development of new omic methodologies and technologies for environmental risk assessment may represent great opprotunities for the identification of emerging risks in the application of fungicides on vineyards. The recent availability of a range of omic technologies provide researchers enormous opportunities to uncover the effects of xenobiotics on many parmameters simultaneously [12]. Omic technologies are valuable tools to measure biochemical changes associated with mode of action at the level of DNA/RNA (transcriptomics), proteins (proteomics) and at the metabolome (metabolomics). They provide the means to identify biomarkers for dose response modelling. These omics repertories might not be sufficiently informative *per se*, since these data are highly multivariate in nature, therefore advanced multivariate data analytical techniques able to cope with the challenges inherent with these very complex analytical data sets should be used. Chemometric methods offer multiple efficient and robust methods for modeling and analysis of complicated chemical/biological data tables, with the goal to produce more interpretable and reliable models capable of handling incomplete, noisy and collinear data structures [13].

Metabolomics enables the detection of possible alterations in the metabolome of organisms as a result of their exposure to bioactive compounds. However, the development of robust metabolomics

models for the study of the mode of action (MoA) of bioactive compounds is a complex procedure. The application of bioactive compounds at sublethal doses under experimental conditions is the preferable option for metabolomics thus enabling the detection of their primary effects on the metabolism of the biological system, while excluding undesirable secondary effects [14].

Untargeted analytical methods can detect hundreds of metabolites with no or with only a limited prior knowledge of the metabolite composition of samples. Liquid chromatography – mass spectrometry (LC-MS) - based approaches have been shown to be particularly useful for untargeted metabolomics [15]. Recently, several untargeted approaches have been proposed to analyse metabolomic profiles of *S.cerevisiae* using different analytical techniques coupled to chemometric evaluation [16-19]. In the present work, we propose and use an untargeted approach to study the metabolomic profiles of a laboratory *S.cerevisiae* strain (*BY4741*) exposed to different sublethal concentrations of copper sulfate ($CuSO_4$). The goal of this approach is to observe what are the main changes in the yeast metabolic profiles and to tentatively identify what metabolites had their concentrations changing more in relation to copper (Cu(II)) exposition. A similar chemometrics strategy to the one previously described in [19] is proposed in this work for the analysis of the LC-MS data and for the resolution of the metabolic profiles with the help of the multivariate curve resolution-alternating least squares (MCR-ALS) method [20-23]. Determination of biomarkes is performed using Student's *t*-test and by variable selection methods, like the variable importance in projection (VIP) [24] and selectivity ratio (SR) methods [25, 26], in partial least squares regression (PLSR) and PLS-discriminant analysis (PLS-DA) [24, 27, 28].

LC-MS analysis generates large data sets, due to the huge size of MS spectra and storage requirements, especially in full scan high *m/z* resolution LC-MS acquisition modes, different data analysis strategies have been proposed. One of them is data binning, which reduces storage and facilitates data analysis steps. However, in this case the resolution power of the raw measurements is lost and several steps are required for its recovery (see for instance strategies used in previous works [16, 19, 29, 30]. Another sounder strategy, is the one based on the use of the regions of interest (ROI), already proposed in some open source software packages for metabolomics such as the *XCMS* software [31]. These strategies take advantage of the sparse nature of the raw MS data and consider only intensity data values higher than a preselected threshold value and having peak elution profile features. They have been recently adapted to the MATLAB environment (see [32]) for chemometric data analysis.

2. Materials and Methods

2.1 *S.cerevisiae* culturing and exposure to sublethal concentrations of Cu(II)

Saccharomyces cerevisiae samples were exposed to copper (added as $CuSO_4$) at sublethal concentrations. *BY4741(MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*) yeast strain cells were pre-cultured in non-selective YPD (yeast peptone dextrose) medium on an orbital shaker (150 rpm) at 30 ºC overnight to obtain the initial culture (pre-culture). Four flasks of 250 ml of YPD medium were inoculated with the yeast pre-culture to an optical density at 600 nm (OD_{600}) of 0.1. These cultures were then grown for 6h to an OD_{600} of 0.8. Volumes of 75 ml of the yeast culture samples were exposed to increasing concentrations of $CuSO_4$ of 1 mM, 3 mM and 6 mM to for 3.5 hours to an OD_{600} of 2.5, with three replicates at each concentration. Three additional replicates were prepared as control samples without any addition of Cu(II).

2.2 Quenching and *S.cerevisiae* metabolites extraction

Metabolism of yeast culture was rapidly inactivated cooling down the samples on ice. Once cooled down, cells from the late exponential growth phase were harvested by centrifugation (4000 rpm for 15 min at 4 ºC) discarding the supernatant and washed twice with phosphate buffered saline (PBS) to adjust their pH to 7.4. Final cell pellets were kept cold until the extraction.

Intracellular metabolites were extracted from the *S.cerevisiae* culture using the boiling ethanol protocol as described previously [33]. Metabolites extraction was performed into 15 mL Falcon tubes, adding 5 mL of solvent (75 % ethanol) to the cell pellet and further incubation of the suspension for 3 min. at 80 ºC. After cooling down the mixture on ice, sample volume was concentrated and dried by evaporation using nitrogen gas (N_2). The residue was re-suspended to a final volume of 0.4 mL with the LC mobile phase (75 % acetonitrile). Prior to pouring the final volume to a vial, the solution was filtered through 0.2 µm GHP membranes (GHP, Acrodisc Syringe Filters) to further ensure removal of any residual protein/debris before LC analysis.

2.3 LC-MS analysis

A Waters Acquity UPLC system (Waters) and a TSKgel Amide-80 5-µm, (250 x 2.0 mm) HILIC (Hydrophilic Interaction LIquid Chromatography) column purchased from Tosoh Bioscience were

used. LC solvents were 0.5 mM ammonium acetate in 90 % acetonitrile at pH 5.5 (solvent A) and 2.5 mM ammonium acetate in 90 % acetonitrile in 60 % acetonitrile at pH 5.5 (solvent B). The gradient elution was as follows: t = 0, 25 % B; t = 8, 30 % B; t = 12, 60 % B; t = 17, 60 % B; t = 20, 25 % B; t = 27, 25 % B. Injection volume was 5 µL and flow rate was 0.15 mL/min.

A Waters LCT Premier orthogonal accelerated time of flight (TOF) mass spectrometer (Waters), operated in negative electrospray ionization (ESI) mode was used to acquire mass spectra profiles in full scan mode from 100 to 800 Da. The mass spectrometer was interfaced to computer workstation running MassLynx V 4.1 software for data acquisition and processing.

Since the aim was to separate highly polar molecules, such as the metabolites more abundant in yeast metabolome, the use of HILIC columns was specially adequate for their analysis. HILIC columns have grown popularity in recent years due to their high compatibility with mass spectrometry and to their detection capabilities improvement for polar metabolites [34]. In a previous work, this approach was already shown to be adequate for yeast growth metabolome analysis stressed by increasing temperatures [19]. In this work, the same analytical methodology has been extended to the analysis of yeast metabolome stressed by Cu(II) treatment.

2.4 LC-MS data pretreatment

Full scan MS spectra of the different chromatographic runs were saved in raw mode and were then converted to cdf format by MassLynx V4.1 software and imported to MATLAB R20012b (Mathworks Inc. Natick, MA, USA) computational environment using mzcdfread.m and mzcdf2peak.m functions from the Bioinformatics Toolbox.

Regions of interest (ROI) strategy has been used in this work to reduce size of MS spectra . The implementation of ROI approach requires the input of a signal-to-noise (SNR) threshold value, the mass accuracy of the mass spectrometer expressed in ppm, and the minimum number of retention times these signals were repeatedly obtained. In the present work, the values of these parameters were 250 (0.15% of maximum MS signal intensity), 0.05 (*m/z* resolution), and 10 retention times respectively. Each region of interest contained masses with significant intensity with no loss of the original spectral resolution. In this way, every MS spectra provided a data matrix with a number of rows equal to the number of measured retention times (ranging from 0 to 27 min.) in the chromatogram and a number of columns equal to the number of finally selected *m/z* ROI values considering all simultaneously analysed chromatographic runs.

2.5 LC-MS data analysis by MCR-ALS

As stated in previous work [19], due to the high number of highly overlapped peaks generated in LC-MS metabolomics analysis of yeast samples, the application of MCR-ALS is proposed and used to facilitate their resolution and identification. The goal of MCR-ALS analysis was to resolve directly the maximum number of individual elution profiles and pure mass spectral profiles of all possible metabolites extracted from the investigated yeast samples and to investigate the effect of Cu(II) exposure on them.

MCR-ALS is a powerful chemometrics method able to analyse multicomponent systems with strongly overlapping contributions from complex chemical systems, including chromatographic ones. The mathematical basis of the bilinear model used by MCR in the case of the particular case of the analysis of a single yeast sample (k) is shown in equation 1:

$$\mathbf{D}_k = \mathbf{C}_k\mathbf{S}^T + \mathbf{E}_k \quad \text{for} \quad k = 1, 2, \ldots, 12 \text{ samples} \qquad \textbf{Equation 1}$$

In this equation, the rows of the data matrices $\mathbf{D}_k$ (I x J) have the MS spectra at all retention times (i= 1,…I) in the chromatographic analysis of this yeast sample, and the columns have the corresponding elution profiles at all the measured mass spectra *m/z* channels (j=1,…J). $\mathbf{C}_k$ is the matrix of MCR-ALS resolved elution profiles in yeast sample k, and $\mathbf{S}^T$ is the matrix of their corresponding resolved pure mass spectra. $\mathbf{E}_k$ contains the unexplained variance related to background and noise contributions not modelled by $\mathbf{C}_k$ and $\mathbf{S}^T$.

MCR-ALS can be extended to the simultaneous analysis of multiple yeast samples analysed by LC-MS as it is shown in next equation 2:

$$\mathbf{D}_{aug} = \begin{bmatrix} \mathbf{D}_1 \\ \mathbf{D}_2 \\ \mathbf{D}_3 \\ \vdots \\ \mathbf{D}_K \end{bmatrix} = \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \mathbf{C}_3 \\ \vdots \\ \mathbf{C}_K \end{bmatrix} \mathbf{S}^T + \begin{bmatrix} \mathbf{E}_1 \\ \mathbf{E}_2 \\ \mathbf{E}_3 \\ \vdots \\ \mathbf{E}_K \end{bmatrix} \qquad \textbf{Equation 2}$$

or briefly:

$$\mathbf{D}_{aug} = \mathbf{C}_{aug}\mathbf{S}^T + \mathbf{E}_{aug} \qquad \textbf{Equation 3}$$

In the particular case under study, the augmented data matrix ($\mathbf{D}_{aug}$) has a number of 22662 rows, equal to the total number of recorded elution times in the simultaneous analysis of the different yeast samples and corresponding chromatographic runs (k = 1, 2, …, 12 samples, three control

samples, and three 1mM Cu(II), three 3mM Cu(II) and three 6mM Cu(II) exposed samples). The number of columns of $\mathbf{D}_{\text{aug}}$ is 1076, which is equal to the number of *m/z* ROI values finally considered. $\mathbf{C}_{\text{aug}}$ is the augmented matrix of the resolved elution profiles in the different chromatographic runs, $\mathbf{S}^{\text{T}}$ is the matrix of pure MS spectra of them and $\mathbf{E}_{\text{aug}}$ matrix is the noise and background signal not explained by the model [20]. One clear advantage of the MCR-ALS of column-wise augmented data matrices as shown in Equations 2 and 3 is that chromatographic peak alignment among chromatographic runs is not necessary because MCR-ALS allows for a complete freedom in the modelling of elution profiles [35] in the different runs.

Although the number of MCR-ALS components is usually estimated by singular value decomposition (SVD) [36], in this work an arbitrary sufficient large number of components was initially proposed. The criterion used to select this number is that it should explain a significant amount of variance and that it should give information about all possible detected chromatographic peaks in the system. In LC-MS data, due to the high selectivity of MS signals, chromatographic peak shape features are rather easily distinguished. Usually, apart from resolving chromatographic elution profiles of the extracted metabolites, other additional components describing solvent and background contributions were also considered.

Once the number of components and the initial estimates have been selected, the bilinear models described by Equations 2 and 3 were solved using a constrained ALS (Alternating Least Squares) iterative approach. In the present study, the applied constraints have been non-negativity for the elution profiles in, $\mathbf{C}_{\text{aug}}$, and for the MS spectra in $\mathbf{S}^{\text{T}}$ as well as their normalization. [20, 22].

MCR-ALS model quality was evaluated by the percent of explained variance (R^2) and the lack of fit (lof) values [37], see Equations 4 and 5 respectively:

$$\text{lof}\,\% = 100\sqrt{\frac{\sum_{i,j} e_{ij}^2}{\sum_{i,j} d_{ij}^2}} \quad e_{i,j} = d_{i,j} - \hat{d}_{i,j} \qquad \textbf{Equation 4}$$

$$R^2 = 100\sqrt{\frac{\sum_{i,j} d_{ij}^2 - \sum_{i,j} e_{ij}^2}{\sum_{i,j} d_{ij}^2}} \qquad \textbf{Equation 5}$$

where $d_{i,j}$ are the experimental values for variable j and sample i; $\hat{d}_{i,j}$ are the corresponding calculated MCR-ALS values (Equation 1).

2.6 Detecting metabolite concentration changes due to Cu(II) treatment

To discover what yeast intracellular metabolites concentrations changed because of the Cu(II) exposition, peak areas of the MCR-ALS resolved elution profiles ($\mathbf{C}_{\mathrm{aug}}$) were statistically evaluated. For this purpose, a Student's *t*-test ($p < 0.05$) was applied to compare the areas of MCR-ALS resolved compounds of control samples with those of the samples exposed to 6 mM of Cu(II).

Furthermore, differences between two groups of samples were also analysed by partial least squares – discriminant analysis (PLS-DA). PLS-DA [38] is a PLS regression (PLSR) method [27, 28, 39] where the set of predictor variables, **X** (yeast metabolites), are correlated to a set of binary variables, **y** (controls and Cu(II) exposed), describing the categories of X. In this work, PLS-DA analysis was performed using the data matrix **X** of the peak areas of the MCR-ALS resolved components, and the **y** data-vector where control samples were categorized as class 0 and 6 mM exposed samples were categorized as class 1. Prior to PLS-DA model calculation, the peak areas were autoscaled to give equal relevance to their possible change due to the exposure to Cu(II). PLS-DA model was assessed using leave-one-out cross-validation due to the small number of samples. PLS-DA model is assessed with the sensitivity parameter which measures the proportion of samples which are correctly identified as exposed to Cu(II), and with the specificity parameter which measures the proportion of controls which are correctly identified as not being exposed to Cu(II). In this work, PLSR was also applied to correlate the **X** –matrix (12 x 100) with the peak areas of the elution profiles (metabolites concentration) resolved by MCR-ALS and the **y**-vector of Cu(II) concentrations (0, 1, 3 and 6mM). Both, PLSR and PLS-DA methods have been described in detail elsewhere [28, 38, 39].

To further investigate the more influent variables (metabolites) in the PLSR and PLS-DA models, the variables importance in projection (VIP) method was used [24, 28]. VIP scores are defined as a weighted sum of squares of PLS weights which take into account the amount of explained **y** variance in each extracted LV. This method is frequently used as a parameter for variable selection [40, 41]. Since the average of the squared VIP scores equals 1, 'greater than one rule' is generally used as a criterion for variable selection [42]. This is not a statistically justified limit, and it can be shown that is very sensitive to the presence of non-relevant information pertaining to **X** [41, 43]. The selectivity Ratio (SR) is another variable selection method frequently used to detect the more important variables of a multivariate data set. The SR value for a variable in particular is defined as the ratio between its explained residual variances of the spectral variables on the target-projected component. In order to ascertain which variables have higher discriminatory abilities, a *F*-value

statistic is calculated for every variable to check whether exceeds the critical value from a *F*-distribution assumption [25].

2.7 Tentative identification of metabolites whose concentration change due to Cu(II) exposure

Results from the MCR-ALS optimization are given as resolved eluted profiles in $\mathbf{C}_{\mathrm{aug}}$, and pure mass spectra, $\mathbf{S}^{\mathrm{T}}$ factor matrices (Equation 3). Changes in peak areas of the resolved elution profiles in Cu(II) exposed samples compared to those of control samples were used to investigate the possible effects of Cu(II) in yeast metabolism. A first evaluation is performed from fold changes using a *t*-test. Additionally, all peak areas resolved by MCR-ALS were autoscaled, and PLS-DA was then applied to identify the most important metabolites responsible for the discrimination of Cu(II) treated and controls samples according to VIP and SR scores. Final selection of metabolites more affected by Cu(II) treatment was done taking into account the three approaches: *t*-test, VIPs and SR.

Pure mass spectra of those components having their concentrations changing considerably by Cu(II) treatment were used for their putative metabolite identification. Since the ROI *m/z* selection procedure (see above) did not lose *m/z* accuracy from raw measured data, metabolite identification was carried out directly from MCR-ALS resolved pure mass spectra in $\mathbf{S}^{\mathrm{T}}$. Since the observed signals in $\mathbf{S}^{\mathrm{T}}$ included the molecular ion and the corresponding charged molecular ion adducts of neutral metabolites, the masses of the most significative negative ion adducts were also considered in the elemental composition calculator searching engine. The calculation was conducted 8 different times for each negative ion peak (i.e., allowing for [M-H], [M-2H], [M-3H], [M-H_2O-H], [M+Na-2H, [M+HAc-H], [2M-H], [2M+HAc-H], [3M-H], where HAc = acetic acid. These metabolites were searched in two on-line databases resources, the Yeast Metabolome Database (YMDB) [44] and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [45, 46]. The final list of the tentatively identified metabolites was then used to investigate and interpret the most probable metabolic pathways and mechanisms affected by the addition of Cu(II) to yeast culture samples.

3. Results and discussion

3.1 Yeast growth

Addition of Cu up to 6mM only resulted in a decrease on growth rate below 5% (data not shown). To a lesser extent, the slow growth of *S. cerevisiae* was also observed at 3 mM of Cu(II).

3.2 Metabolomic chemometric analysis

A large number of components, such as 100, were proposed in the MCR-ALS analysis to allow for a sufficient large number of eluted metabolites plus a number of extra components describing background, solvent and other non-well defined chromatographic contributions (systematic noise sources). With this high number of components, MCR-ALS explained a total variance (R^2) higher than 99%, and a lack of fit (lof) lower than 9% (Equations 4 and 5). In Figure 1, an example of a small chromatographic region with three resolved components, out of the set of the 100 components included in the global MCR model, is given. Elution profiles of these three components for control and treated samples are shown above the figure, their corresponding pure spectra are shown left below and resolution of the three elution profiles for one of the replicates of the Cu(II) 3 mM treated sample are shown right below.

The PLS-DA discrimination model with one latent variable (LV), already explained 41.05 % of the total **X** variance (peak areas of MCR-ALS resolved elution profiles) and 96.02 % of the total class variance (**y**), with specificity and sensitivity values equal to 1 for each class. The PLSR model with one-latent variable already captured a large part of the **y** variance (around 94%). Figure 2 displays the resulting scores of this PLSR model; as it can be seen, the four sample groups (control and Cu(II) at 1mM, 2mM, and 6mM concentration) were separated according Cu(II) dose. These results suggested that the metabolic concentrations changed linearly according to exposure to increasing Cu(II) doses. Control and Cu(II) exposed samples to 1mM dose were closer, with the group of samples exposed to 3mM more separated from them. Samples exposed to 6mM Cu(II) give the group more isolated in Figure 2.

3.3 Time-course of concentration changes in specific metabolites

Changes in metabolite concentrations (fold-change) are summarized along with their metabolite assignments in Table 1. VIP and SR values (see method section 2.6), calculated from PLS-DA model revealed what variables (metabolites) had the greatest influence on the discrimination between control and 6 mM Cu(II) exposed *S. cerevisiae* metabolic samples. 41 metabolites had VIP scores higher than one. When this threshold was increased to 2, 19 metabolites were selected. A total number of 14 metabolites were coincident using three different analyses (see Table 1): Student's *t*-test ($p < 0.05$), VIP (greater than 2) and SR (*F*-test, 95%). These metabolites were identified as described in section 2.7. Table 1 also shows the tentative identification of these metabolites (compound name, molecular formula, adduct, error and KEGG number), their corresponding concentration fold change, and its concentration trend (up = increasing or down = decreasing). Among the different changes observed, glutathione (GSH) concentration was reduced to less than 1% of the levels in control samples in the highest dose group (6mM Cu(II)). Nicotinamide D- ribonucleotide and Nicotinate D-ribonucleoside concentration showed a remarkable increase with a fold-change of 26.5 and 11.6 respectively. Furthermore, concentrations of trehalose and L-Glutamic acid also increased with a fold-change higher than 6.

3.4 Biological interpretation

The 14 metabolites whose concentrations varied between control and Cu(II) exposed samples (Table 1) felt into different functional categories according to KEGG (not mutually exclusive): glutathione metabolism (glutathione and L-glutamic acid), amino acid metabolism (L-phenylalanine, 2-Isopropylmaleic acid, L-leucine, L-isoleucine and L-aspartic acid), purine metabolism (guanosine and adenosine diphosphate ribose), nicotinate and nicotinamide metabolism (nicotinate D-ribonucleoside and nicotinamide D- ribonucleotide), starch and sugar metabolism (trehalose), pyrimidine metabolism (L-dihydroorotic acid) and lysine degradation (L-pipecolate).

Exposure to sublethal Cu(II) concentrations involved a complex response that led to yeast cell to acclimate to prevented cell death. Yeast responses to Cu(II) exposure promoted three major changes in the metabolism: defense against reactive oxygen species (ROS), cell protection and DNA repair.

Glutathione was clearly downregulated when yeast was exposed to high concentration of Cu(II) (6 mM). Since free Cu(II) ions are likely to participate in the formation of reactive oxygen species

(ROS), Cu(II) and Cu(I) ions will participate actively in oxidation and reduction biochemical reactions. In the presence of reducing agents such as glutathione (GSH), Cu(II) can be reduced to Cu(I), which is then able to catalyze the formation of hydroxyl radicals (OH·) from hydrogen peroxide (H_2O_2) via the Haber-Weiss reaction [47, 48], causing the depletion of glutathione concentrations [49, 50], in agreement with results reported in Table 1. Large amounts of ROS will led to protein denaturation, membrane order alteration and damage of intracellular enzyme activity and consequent reduced metabolism [51]. In *S. cerevisiae*, exposure to low doses of H_2O_2 or oxidative stress by glutathione depletion induces apoptosis [52]. In our case, we observed a near total depletion of glutathione without significant effects on yeast growth. The proposed methodology allowed the detection of an early step of the oxidative stress process, and confirmed glutathione as the first line of defence against oxidants in the cell. The increase of concentration of L-Glutamic acid (see Table 1), a precursor of GSH, may reflect the activation of metabolic pathways leading to the restoration of physiological GSH levels.

Trehalose is capable of reducing oxidant-induced modifications of proteins during exposure of yeast cells to H_2O_2, and it is considered a general stabilizer of starving or stressed yeast cells [53]. Therefore, its increased concentration at higher doses of Cu(II) (see Table 1) may be explained by its capacity to scavenge free radicals and therefore protecting cellular constituents from oxidative damage [54].

The presence of reactive oxygen species (ROS) and of Cu(II) ions themselves in the presence of reducing agents such as glutathione may result in damage to DNA. ROS species and OH· radicals either oxidize bases, generate other radicals that result in crosslinks, or cleave the phosphoester bonds between specific nucleotides in the DNA [55]. DNA damage induces several cellular responses that enable the cell either to eliminate or cope with the damage. DNA ligase is an enzyme important for DNA repair and replication. The first step of DNA ligase-mediated DNA repair involves nucleophilic attack on the α phosphorous adenosine triphosphate (ATP) or nicotinamide adenine dinucleotide (NAD^+), resulting in the release of pyrophosphate or nicotinamide D-ribonucleotide (NMN) (see Table 1) and formation of a covalent intermediate (ligase-adenylate), in which adenosine monophosphate (AMP) is linked via phosphoamide bond to lysine [56, 57]. This or a similar mechanism may account for the increase in NMN, nicotinamide D-ribonucleoside (NAR) and adenosine diphosphate ribose (ADP-ribose) concentrations observed in Table 1 upon exposure to high Cu(II) concentrations. NAD+ is a co-enzyme of pivotal importance in the redox balance of metabolism, as it is continuously interconverted between an oxidized (NAD^+) and reduced (NADH) state. The increases of NAR and ADP-ribose concentrations (see Table 1) may be explained by the fact that NAD^+ is synthetized from NAR [58] and consumption of NAD^+ involves

release of nicotinamide and transfer of the remaining ADP-ribose moiety onto acceptor molecules [59].

4. Conclusions

In this study, a new LC-MS based metabolomics approach combined with chemometrics was used to explore Cu(II) toxicity in *S. cerevisiae* cultures . Exposure to sublethal concentrations of Cu(II) produced significant changes at the metabolic level, even at conditions where yeast growth was not significantly affected. Changes in chromatographic peak areas provided a useful insight about what potential metabolites changed their concentrations following the exposure to different Cu(II) doses. A new ROI data compression strategy allowed not losing spectral resolution nor *m/z* accuracy from LC-MS raw data and enabled the direct identification of the resolved chromatographic peaks. The intracellular metabolites best contributing to samples discrimination were selected by means of VIP-PLS and SR-PLS variable selection methodologies. Fourteen metabolites showed significant concentration changes upon Cu(II) exposure, following a dose-response effect. Observed metabolic changes were consistent with expected effects of Cu(II) intoxication and with the physiological responses to its presence.

High concentrations of Cu(II) caused increased concentrations of ROS which led to reduced yeast metabolism and DNA damage. An early step of the oxidative stress process was detected, since there was a near total depletion of glutathione without significant effects on yeast growth. Therefore, glutathione was confirmed to act as a first line of defence against oxidants in the cell. The concentration increase of L-glutamic acid reflected the activation of metabolic pathways leading to the restoration of physiological glutathione levels. Yeast cells counteracted the resulting oxidative stress by protecting their cellular constituents by increasing trehalose concentration at higher Cu(II) doses. A DNA repair mechanism activation was observed and explained by the increase of nicotinamide D- ribonucleotide, nicotinamide D-ribonucleoside and adenosine diphosphate ribose concentrations when the higher Cu(II) concentrations were used.

Acknowledgements

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. 320737.

References

1. Avery, S.V., Howlett N.G., and Radice, S. (1996). *Copper toxicity towards Saccharomyces cerevisiae: dependence on plasma membrane fatty acid composition.* Applied and Environmental Microbiology *62* (11): p. 3960-3966.
2. Komárek, M., Čadková, E., Chrastný, V., Bordas, F. and Bollinger, J.C. (2010). *Contamination of vineyard soils with fungicides: A review of environmental and toxicological aspects.* Environment International *36* (1): p. 138-151.
3. Brandolini, V., Tedeschi, P., Capece, A., Maietti, A., Mazzotta, D., Salzano, G., Paparella, A. and Romano, P. (2002). *Saccharomyces cerevisiae wine strains differing in copper resistance exhibit different capability to reduce copper content in wine.* World Journal of Microbiology and Biotechnology *18* (6): p. 499-503.
4. Flemming, C.A. and J.T. Trevors (1989). *Copper toxicity and chemistry in the environment: a review.* Water, Air, and Soil Pollution *44* (1-2): p. 143-158.
5. Cervantes, C. and F. Gutierrez-Corona (1994). *Copper resistance mechanisms in bacteria and fungi.* FEMS Microbiology Reviews*14* (2): p. 121-137.
6. Giller, K.E., E. Witter, and S.P. McGrath (1998). *Toxicity of heavy metals to microorganisms and microbial processes in agricultural soils: a review.* Soil Biology and Biochemistry *30* (10–11): p. 1389-1414.
7. Dupont, C.L., G. Grass, and Rensing C. (2011) *Copper toxicity and the origin of bacterial resistance-new insights and applications.* Metallomics *2* (11): p. 1109-1118.
8. Sun, X.y., Zhao, Y., Liu, L.l., Jia, B., Zhao, F., Huang, W.d and Zhan, J.c. (2015). *Copper Tolerance and Biosorption of Saccharomyces cerevisiae during Alcoholic Fermentation.* PLoS ONE *10* (6)
9. Azenha, M., Vasconcelos, M.T., and Moradas-Ferreira, P. (2000). *The influence of Cu concentration on ethanolic fermentation by Saccharomyces cerevisiae.* Journal of Bioscience and Bioengineering *90* (2): p. 163-167.
10. Presta, A. and Stillman, M.J. (1997). *Incorporation of copper into the yeast saccharomyces cerevisiae. Identification of Cu(I)-metallothionein in intact yeast cells.* Journal of Inorganic Biochemistry *66* (4): p. 231-240.
11. Welch, J.W., Fogel, S., Cathala, G. and Karin, M (1983), *Industrial yeasts display tandem gene iteration at the CUP1 region.* Molecular and Cellular Biology *3* (8): p. 1353-1361.
12. dos Santos, S., Piculell, L., Medronho, B., Miguel, M.G. and Lindman, B. (2012). *Phase behavior and rheological properties of DNA-cationic polysaccharide mixtures.* Journal of Colloid and Interface Science *383* (1): p. 63-74.
13. Trygg, J., Holmes, E. and Lundstedt, T. (2006). *Chemometrics in Metabonomics.* Journal of Proteome Research *6* (2): p. 469-479.
14. Aliferis, K.A. and Jabaji, S (2011). *Metabolomics – A robust bioanalytical approach for the discovery of the modes-of-action of pesticides: A review.* Pesticide Biochemistry and Physiology, *100* (2): p. 105-117.
15. Dunn, W., Erban, A., Weber, R., Creek, D., Brown, M., Breitling, R., Hankerneier, T., Goodacre, R., Neumann, S., Kopka, J. and Viant M. (2013) *Mass appeal: metabolite identification in mass spectrometry-focused untargeted metabolomics.* Metabolomics, *9* (1): p. 44-66.
16. Ortiz-Villanueva, E., Jaumot, J., Benavente, F., Piña, B., Sanz-Nebot, V. and Tauler, R. (2015). *Combination of CE-MS and advanced chemometric methods for high-throughput metabolic profiling.* Electrophoresis *36* (18): p. 2324-2335.
17. Puig-Castellví, F., Alfonso, I., Piña, B. and Tauler, R. (2015). *A quantitative 1H NMR approach for evaluating the metabolic response of Saccharomyces cerevisiae to mild heat stress.* Metabolomics 11 (6): p. 1612-1625.
18. Son, H.S., Hwang, G.S., Kim, K.M., Kim, E.Y., van den Berg, F., Park, W.M., Lee, C.H. and Hong, Y.S. (2009). *1H NMR-Based Metabolomic Approach for Understanding the Fermentation Behaviors of Wine Yeast Strains.* Analytical Chemistry *81* (3): p. 1137-1145.
19. Farrés, M., B. Piña, and R. Tauler (2015). *Chemometric evaluation of Saccharomyces cerevisiae metabolic profiles using LC–MS.* Metabolomics, *11* (1): p. 210-224.
20. Tauler, R. (1995). *Multivariate curve resolution applied to second order data.* Chemometrics and Intelligent Laboratory Systems *30* (1): p. 133-146.

21. Tauler, R., Kowalski, B., and Fleming, S. (1993). *Multivariate curve resolution applied to spectral data from multiple runs of an industrial process.* Analytical Chemistry *65* (15): p. 2040-2047.

22. Tauler, R. and Barceló, D. (1993). *Multivariate curve resolution applied to liquid chromatography—diode array detection.* TrAC Trends in Analytical Chemistry *12* (8): p. 319-327.

23. Tauler, R., Smilde, A. and B. Kowalski (1995). *Selectivity, local rank, three-way data analysis and ambiguity in multivariate curve resolution.* Journal of Chemometrics *9* (1): p. 31-58.

24. Wold, S., Johansson, A. and Cochi, M (1993). *PLS-partial least squares projections to latent structures* 3D QSAR in Drug Design, Theory, Methods, and Applications, ed. H. Kubinyi. ESCOM Science Publishers: Leiden. 523-550.

25. Rajalahti, T., Arneberg, R., Berven, F.S., Myhr, K.M., Ulvik, R.J. and Kvalheim, O.M (2009) *Discriminating Variable Test and Selectivity Ratio Plot: Quantitative Tools for Interpretation and Variable (Biomarker) Selection in Complex Spectral or Chromatographic Profiles.* Analytical Chemistry *81* (7): p. 2581-2590.

26. Rajalahti, T., Arneberg, R., Kroksveen, A.C., Berle, M., Myhr, K.M. and Kvalheim, O.M. (2009) *Biomarker discovery in mass spectral profiles by means of selectivity ratio plot.* Chemometrics and Intelligent Laboratory Systems *95* (1): p. 35-48.

27. Wold, H. (1966). *Estimation of Principal Components and Related Models by Iterative Least squares*, in *Multivariate Analysis.* Academic Press. p. 391-420.

28. Wold, S., Sjöström, M., and L. Eriksson, L. (2001). *PLS-regression: a basic tool of chemometrics.* Chemometrics and Intelligent Laboratory Systems *58* (2): p. 109-130.

29. Gorrochategui, E., Casas, J., Porte, C., Lacorte, S. and Tauler, R. (2015). *Chemometric strategy for untargeted lipidomics: Biomarker detection and identification in stressed human placental cells.* Analytica Chimica Acta *854*: p. 20-33.

30. Navarro-Reig, M., Jaumot, J., García-Reiriz, A. and Tauler, R. (2015). *Evaluation of changes induced in rice metabolome by Cd and Cu exposure using LC-MS with XCMS and MCR-ALS data analysis strategies.* Analytical and Bioanalytical Chemistry *407* (29): p. 8835-8847.

31. Tautenhahn, R., Bottcher, C. and Neumann, S. (2008). *Highly sensitive feature detection for high resolution LC/MS.* BMC Bioinformatics *9* (1): p. 504.

32. Gorrochategui, E., Jaumot, J. and Tauler, R. (2015) *A protocol for LC-MS metabolomic data processing using chemometric tools.* Protocol Exchange

33. Gonzalez, B., François, J. and Renaud, M. (1997). *A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol.* Yeast *13* (14): p. 1347-1355.

34. Lämmerhofer, M. (2010). *HILIC and mixed-mode chromatography: The rising stars in separation science.* Journal of Separation Science *33* (6-7): p. 679-680.

35. de Juan, A. and Tauler, R. (2001). *Comparison of three-way resolution methods for non-trilinear chemical data sets.* Journal of Chemometrics *15* (10): p. 749-771.

36. Golub, G.H. and Loan, C.F.V. (1996) *Matrix Computations*. The Johns Hopkins University Press (Baltimore and London).

37. de Juan, A., Jaumot, J. and Tauler, R. (2014). *Multivariate Curve Resolution (MCR). Solving the mixture analysis problem.* Analytical Methods *6* (14): p. 4964-4976.

38. Barker, M. and Rayens, W. (2003). *Partial least squares for discrimination.* Journal of Chemometrics *17* (3): p. 166-173.

39. Geladi, P. and Kowalski, B.R. (1986). *Partial least-squares regression: a tutorial.* Analytica Chimica Acta *185*: p. 1-17.

40. Andersen, C. M. and Bro, R. (2010). *Variable selection in regression—a tutorial.* Journal of Chemometrics, *24* (11-12): p. 728-737.

41. Farrés, M., Platikanov, S., Tsakovski, S. and Tauler, R. (2015). *Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation.* Journal of Chemometrics, 2015: p. 528-536.

42. Chong, I.G. and Jun, C.H. (2005). *Performance of some variable selection methods when multicollinearity is present.* Chemometrics and Intelligent Laboratory Systems *78* (1–2): p. 103-112.

43. Tran, T.N., Afanador, N.L, Buydens, L. and Blanchet, L. (2014). *Interpretation of variable importance in Partial Least Squares with Significance Multivariate Correlation (sMC)*. Chemometrics and Intelligent Laboratory Systems *138*: p. 153-160.

44. Jewison, T., Knox, C., Neveu, V., Djoumbou, Y., Guo, A.C., Lee, J., Liu, P., Mandal, R., Krishnamurthy, R., Sinelinkov, I., Wilson, M. and Wishart, D. (2012) *YMDB: the Yeast Metabolome Database.* Nucleic Acids Research, 2012. *40* (D1): p. D815-D820.

45. Kanehisa, M., et al., *KEGG for integration and interpretation of large-scale molecular data sets.* Nucleic Acids Research *40* (D1): p. D109-D114.

46. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999) *KEGG: Kyoto Encyclopedia of Genes and Genomes*. Nucleic Acids Research *27* (1): p. 29-34.

47. Bremner, I., *Manifestations of copper excess* (1998). The American Journal of Clinical Nutrition *67* (5 Suppl): p. 1069S-1073S.

48. Gaetke, L.M. and C.K. Chow, C.K. (2003). *Copper toxicity, oxidative stress, and antioxidant nutrients.* Toxicology *189* (1–2): p. 147-163.

49. Estruch, F. (2000). *Stress-controlled transcription factors, stress-induced genes and stress tolerance in budding yeast.* FEMS Microbiology Reviews *24* (4): p. 469-86.

50. Izawa, S., Y. Inoue, Y. and Kimura, A. (1995). *Oxidative stress response in yeast: effect of glutathione on adaptation to hydrogen peroxide stress in Saccharomyces cerevisiae.* FEBS Letters *368* (1): p. 73-6.

51. Berger, F. Ramírez-Hernández, M.A.H. and Ziegler, M. (2004). *The new life of a centenarian: signalling functions of NAD(P).* Trends in Biochemical Sciences. *29* (3): p. 111-118.

52. Madeo, F., Fröhlich, E., Ligr, M., Grey, M., Sigrist, S.J., Wolf, D.H. and Fröhlich K.-U. (1999) *Oxygen Stress: A Regulator of Apoptosis in Yeast.* The Journal of Cell Biology, *145*(4): p. 757-767.

53. Benaroudj, N., Lee, D.H. and Goldberg, A.L. (2001). *Trehalose Accumulation during Cellular Stress Protects Cells and Cellular Proteins from Damage by Oxygen Radicals.* Journal of Biological Chemistry *276* (26): p. 24261-24267.

54. de Jesus Pereira, E., Panek, A.D. and Eleutherio, E.C.A. (2003). *Protection against oxidation during dehydration of yeast.* Cell Stress & Chaperones *8* (2): p. 120-124.

55. Linder, M.C. (2012) *The relationship of copper to DNA damage and damage prevention in humans.* Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis *733* (1–2): p. 83-91.

56. Odell, M., Sriskanda, V., Shuman, S. and Nikolov, D.B. (2000). *Crystal Structure of Eukaryotic DNA Ligase–Adenylate Illuminates the Mechanism of Nick Sensing and Strand Joining.* Molecular Cell *6* (5): p. 1183-1193.

57. Sriskanda, V., Moyer, R.W. and Shuman, S. (2001). *NAD+-dependent DNA Ligase Encoded by a Eukaryotic Virus.* Journal of Biological Chemistry *276* (39): p. 36100-36109.

58. de Figueiredo, L.F., Gossmann, T.J, Ziegler, M. and Schuster S. (2011). *Pathway analysis of NAD+ metabolism.* Biochemical Journal *439* (2): p. 341-348.

59. Koch-Nolte, F., Fischer, S., Haag, F. and Ziegler, M. (2011). *Compartmentation of NAD+-dependent signalling.* FEBS Letters *585* (11): p. 1651-1656.

Table 1 Most relevant metabolites whose concentrations changed due to Cu(II) yeast culture exposition and their KEGG tentative identification.

Met. ID	Compound	Molecular formula	Adduct	Error (ppm)	KEGG C-number	Fold-change	Trend	*p*-value
1	Gultathione	$C_{10}H_{17}N_3O_6S$	$[M-H]^-$	20.48	C00051	105.2	DOWN	0.029
2	L-Glutamic acid	$C_5H_9NO_4$	$[M-H]^-$	0.22	C00025	5.1	UP	0.020
3	L-Dihydroorotic acid	$C_5H_6N_2O_4$	$[M-H]^-$	7.13	C00337	6.2	DOWN	0.037
4	L-Phenylalanine	$C_9H_{11}NO_2$	$[M-H]^-$	9.89	C00079	1.8	UP	0.018
5	2-Isopropylmaleic acid	$C_7H_{10}O_4$	$[3M-H]^-$	7.08	C02631	3.3	UP	0.022
6	Nicotinate D-ribonucleoside	$C_{11}H_{14}NO_6$	$[M+HAc-H]^-$	32.26	C05841	11.6	UP	0.024
7	Adenosine diphosphate ribose	$C_{15}H_{23}N_5O_{14}P_2$	$[M-H_2O-H]^-$	18.64	C00301	3.6	UP	0.009
8	L-Leucine	$C_6H_{13}NO_2$	$[M-H]^-$	10.16	C00123	3.1	UP	0.002
9	L-Isoleucine	$C_6H_{13}NO_2$	$[M-H]^-$	10.16	C00407	3.1	UP	0.013
10	L-Aspartic acid	$C_4H_7NO_4$	$[M-H]^-$	10.55	C00049	5.4	DOWN	0.006
11	Nicotinamide D-ribonucleotide	$C_{11}H_{15}N_2O_8P$	$[M-H]^-$	42.18	C00455	26.5	UP	0.010
12	Guanosine	$C_{10}H_{13}N_5O_5$	$[M-H]^-$	19.26	C00387	1.7	UP	0.037
13	L-Pipecolate	$C_6H_{11}NO_2$	$[M+HAc-H]^-$	10.21	C00408	2.8	UP	0.023
14	Trehalose	$C_{12}H_{22}O_{11}$	$[M+HAc-H]^-$	35.68	C01083	6.3	UP	0.004

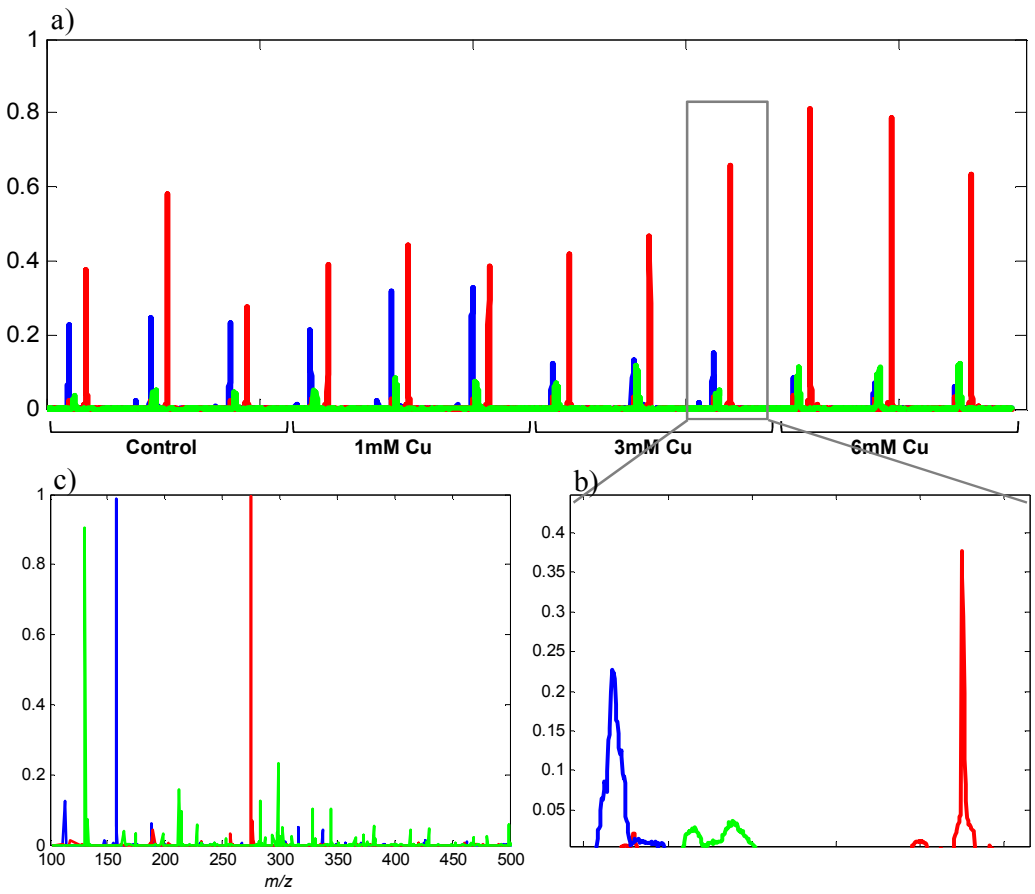


Figure 1 Example of MCR-ALS results of one one chromatographic region with three coeluted components. a) Resolved elution profiles for these three components in the individual chromatographic runs. b) Expansion of the resolution of these three components in one of the yeast samples exposed at 6 mM Cu(II), and c) pure resolved mass spectra.

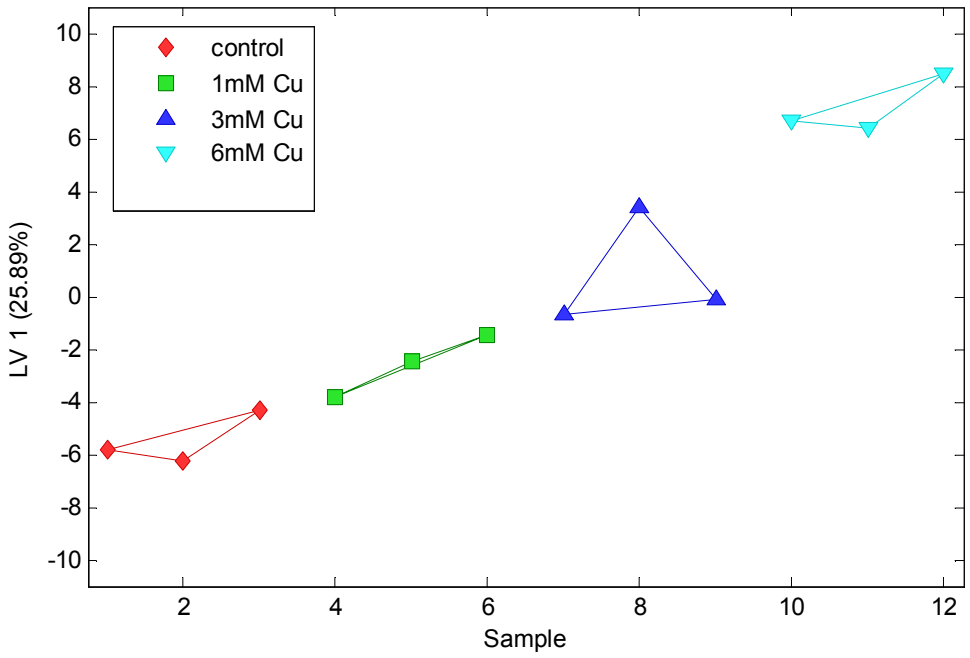


Figure 2. LV1 PLSR Scores plot in the analysis of yeast samples exposed to 0 mM (control), 1 mM, 3 mM and 6 mM of Cu(II).