This item is the archived peer-reviewed author-version of:

A generic approach for shotgun analysis of the soluble proteome of plant cell suspension cultures

Reference:

de Vijlder Thomas, Valkenborg Dirk, Dewaele Debbie, Remmerie Noor, Laukens Kris, Witters Erwin.- *A generic approach for shotgun analysis of the soluble proteome of plant cell suspension cultures*

Journal of chromatography: B: analytical technologies in the biomedical and life sciences - ISSN 1570-0232 - (2014), p. 1-39

DOI: <http://dx.doi.org/doi:10.1016/j.jchromb.2014.10.024>

Accepted Manuscript

Title: A generic approach for “shotgun” analysis of the soluble proteome of plant cell suspension cultures

Author: Thomas De Vijlder Dirk Valkenborg Debbie Dewaele
Noor Remmerie Kris Laukens Erwin Witters

PII: S1570-0232(14)00665-5
DOI: <http://dx.doi.org/doi:10.1016/j.jchromb.2014.10.024>
Reference: CHROMB 19173

To appear in: *Journal of Chromatography B*

Received date: 31-7-2014
Revised date: 13-10-2014
Accepted date: 21-10-2014

Please cite this article as: T. De Vijlder, D. Valkenborg, D. Dewaele, N. Remmerie, K. Laukens, E. Witters, A generic approach for “shotgun” analysis of the soluble proteome of plant cell suspension cultures, *Journal of Chromatography B* (2014), <http://dx.doi.org/10.1016/j.jchromb.2014.10.024>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A generic approach for “shotgun” analysis of the soluble proteome of plant cell suspension cultures.

Thomas De Vijlder^{a,b,1}, Dirk Valkenburg^{a,c,d}, Debbie Dewaele^e, Noor Remmerie^{a,b,2} Kris

Laukens^{f,g} and Erwin Witters^{a,c,h}

^a Center for Proteomics (CFP), Groenenborgerlaan 171, B-2020 Antwerp, Belgium

^b Laboratory of Plant Growth and Development, Department of Biology, University of Antwerp, Groenenborgerlaan 171,

^c Vlaamse Instelling voor Technologisch Onderzoek (VITO), Boeretang 200, B-2400 Mol, Belgium

^d Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Agoralaan 1, B-3590 Diepenbeek, Belgium

^e Biomolecular & analytical mass spectrometry, Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium.

^f Biomedical Informatics Research Center Antwerp (Biomina), University of Antwerp / Antwerp University Hospital, Wilrijkstraat 10, B-2650 Edegem, Belgium

^g Advanced Database Research and Modelling, Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, B-2020 Antwerp, Belgium

^h Laboratory for Ecophysiology, Biochemistry and Toxicology, Department of Biology, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

Corresponding author: thomas.devijlder@uantwerpen.be, telephone number: +32 14 60 25 54,

Key words: *Arabidopsis thaliana*, *Nicotiana tabacum* cv “Bright Yellow-2”, proteome analysis, plant cell suspension cultures, mass spectrometry

¹Present address: Janssen Pharmaceutica NV, Turnhoutseweg 30, B-2340 Beerse, Belgium

² Present address: Multiplicom NV, Galileilaan 18, B-2845 Niel, Belgium

Abstract

Cell suspension cultures from different plant species act as important model systems for studying cellular processes in plant biology and are often used as “green factories” for the production of valuable secondary metabolites and recombinant proteins. While mass spectrometry based proteome analysis techniques are ideally suited to study plant cell metabolism and other fundamental cellular processes from a birds eye perspective, they remain underused in plant studies. We describe a comprehensive sample preparation and multidimensional ‘shotgun’ proteomics strategy that can be generically applied to plant cell suspension cultures. This strategy was optimized and tested on an *Arabidopsis thaliana* ecotype Landsberg erecta culture. Furthermore, the implementation of strong cation exchange chromatography as a peptide fractionation step is elaborately tested. Its utility in mass spectrometry based proteome analysis is discussed. Using the presented analytical platform, over 13,000 unique peptides and 2640 proteins could be identified from a single plant cell suspension sample. Finally, the experimental setup is validated using *Nicotiana tabacum* cv. “Bright Yellow-2” (BY-2) plant cell suspension cultures, thereby demonstrating that the presented analytical platform can also be valuable tool in proteome analysis of non-genomic model systems.

1. Introduction

Cell suspension cultures from a myriad of different plant species have long been valuable research tools to analyse specific and complex cellular and biochemical processes. For example, they have been used to investigate cell cycle progression [1,2], cell differentiation [3] and biochemical synthesis of secondary metabolites [4,5].

The first attempts to isolate and cultivate plant cells *in vitro* were described by Haberlandt, an Austrian botanist, over a century ago [6]. Afterwards, many scientists have continuously improved breeding and cultivation techniques. The development of the first techniques for the propagation of plant cells in liquid cultures dates back to the 1950s [7,8]. Afterwards, the feasibility of growing plant cells in suspension cultures on a large scale was swiftly established [8].

Next to their use in basic plant biology research, the utility of plant cell suspension cultures in the production of potentially valuable secondary metabolites was soon appreciated [9]. Examples are the production and extraction of shikonin (and its stereoisomer alkanin) from *Alkanna tinctoria* [10] and paclitaxel (taxol) from *Taxus sp.* [11]. Both secondary metabolites are used to combat cancer. Moreover, plant cell suspension cultures have also often been used as living 'factories' for the production of recombinant proteins, including antibodies and cytokines [12]. In many occasions, *Nicotiana tabacum* cell suspension cultures are used as host plant cell culture systems, e.g. for the production of human growth hormone [13], interleukins [14] or human erythropoietin [15].

The analysis of proteins on a proteome-wide scale has gained momentum during the last decade. Proteome analysis typically results in an overwhelming amount of valuable information describing the state of a biological entity. Despite their aforementioned importance as a basic research tool and as a vehicle for recombinant protein production, work on plant cell suspension cultures has not received a great deal of attention in proteome analysis.

Originally, most of the published reports on plant proteome analysis used a strategy based on two-dimensional gel electrophoresis (2DGE) [16]. During the last decade, chromatography based 'shotgun' approaches have been developed as well, both for plant cell suspensions of model species with a sequenced genome [19,20] as for systems like Banana, lacking a sequenced genome at the time of publication [21]. At present, gel-based and gel-free techniques coexist as complementary approaches in the analysis of proteomes of plant species with or without a sequenced genome [22,23].

In most recent shotgun proteome analysis reports, little is mentioned and no details are provided on the choice of sample preparation techniques and peptide separation strategies that were applied. However, thorough optimization of these sometimes neglected analytical steps is key in performing high-quality shotgun proteome analysis.

As stated above, plant cell suspensions are often used for the production of valuable secondary metabolites or recombinant proteins. It is worth mentioning that these secondary metabolites and proteins only account for less than 1% of a cultures dry weight, rendering their isolation and purification laborious and expensive. Therefore, thorough understanding of their metabolic

pathways and the way these react to external cues is of pivotal importance to increase the yields of the desired secondary metabolites and proteins. Holistic approaches such as proteome analysis (but also transcriptomics, metabolomics, next generation whole genome sequencing etc.) are gaining attention in this respect. For example Desgagne-Penix and colleagues describe a transcriptomic and proteomic approach to investigate the alkaloid metabolism in opium poppy cell cultures [24]. Recently, Bally and coworkers used proteome analysis to investigate changes in *Nicotiana tabacum* plastid protein abundance as a result of recombinant protein expression [25]. An excellent overview of systems biology efforts to analyse secondary metabolism is given by Schilmiller et al. [26].

Here we describe a comprehensive shotgun analysis approach based on multi-dimensional liquid chromatography (LC) and high-resolution mass spectrometry (MS) to probe the soluble proteome of plant cell suspension cultures. The *Arabidopsis thaliana* cv. Landsberg erecta cell suspension culture was used to optimize the sample preparation and analysis strategy. The approach is based on offline prefractionation of peptides by strong cation exchange chromatography (SCX-LC) prior to reversed phase chromatography (RP-LC) and mass spectrometric identification. We demonstrate the orthogonality of the SCX prefractionation and its potential to enrich acetylated and phosphorylated peptides, as described elsewhere [27]. Finally, the optimized strategy was validated using a plant cell suspension culture from a species that lacks an annotated nuclear genome, namely *Nicotiana tabacum* cv. "Bright Yellow-2" (BY-2).

2. Materials & Methods

2.1 Materials and chemicals

The *Arabidopsis thaliana* PSB-D (Plant Systems Biology-D, ecotype Landberg erecta) plant cell suspension culture was a kind gift from the department of Plant Systems Biology (Ugent, Belgium). The *Nicotiana tabacum* cv. "Bright Yellow-2" plant cell suspension cultures were obtained from dr. Toshiyuki Nagata. Urea was purchased from USB chemicals (Cleveland, OH, USA) and β -mercaptoethanol came from GE healthcare (Uppsala, Sweden). Water and acetonitrile solvents were all LC-MS grade and obtained from Acros organics (Geel, Belgium). Formic acid came from Merck chemicals (Darmstadt, Germany). Empore 3M C₈ extraction disk were a kind gift from 3M Belgium (Diegem, Belgium). All other chemicals and products were purchased from Sigma (St. Louis, MO, USA) unless mentioned otherwise.

2.2 Cell culture conditions and sampling

The *Arabidopsis thaliana* PSB-D cell suspension culture was maintained as described by Menges *et al.* [28] in 50 mL of a medium containing 4.43 g/L MSMO (Duchefa, Haarlem, The Netherlands), 30 g/L sucrose, 0.5 mg/L α -naphthaleneacetic acid and 0.05 mg/mL kinetin. The pH of the medium was adjusted to 5.7 with 0.1 M potassium hydroxide. Cell suspensions were kept in continuous dark at 25°C and under agitation at 130 rpm. Every week, cells were subcultured in fresh medium at a 1:25 dilution. A culture in exponential phase was obtained by transferring 12 mL of a 7-day old stationary culture to 50 mL of fresh medium, followed by 48h of subculture. Ten mL of Cultures in exponential phase were sampled on a filter (Whatmann nr

1) through vacuum filtration and immediately frozen in liquid N₂. Samples were stored at -80°C until protein extraction.

Nicotiana tabacum cv Bright Yellow-2 plant cell suspension cultures were maintained as described by Nagata et al. [29] in flasks containing 50 mL of culture medium consisting of 30 g/L sucrose, 4.3 g/L M&S salts (Duchefa, Haarlem, The Netherlands), 200 mg/L KH₂PO₄, 1 mg/L thiamine and 0.2 mg/L 2,4-dichlorophenoxyacetic acid. The tobacco BY-2 cell culture was maintained by weekly transfer of 0.6 mL stationary culture to 50 mL of culture medium. Cells were harvested as described for *Arabidopsis thaliana* cultures.

2.3 Protein extraction

Frozen plant cell suspension cells were homogenized by grinding in a liquid N₂ cooled mortar. The homogenate was resuspended in 1.5 mL acetone containing 10% (v/v) trichloroacetic acid (TCA) and 0.07% (v/v) β-mercaptoethanol (kept at -20°C until use) and incubated overnight at -20°C for precipitation. After centrifugation for 20 minutes (16000 g, 4°C) in an Eppendorf 5415 centrifuge (Hamburg, Germany), the supernatant was removed and 1 mL acetone (kept at -20°C until use) was added to the pellet. The mixture was then vortexed, incubated for 15 minutes at -20°C and afterwards centrifuged for 20 minutes (16000g, 20°C). This washing step was repeated twice. Afterwards the resulting pellet was air-dried under a laminar flow. The pellet was then resolubilized in 150 µL of a buffer containing 6 M urea, 2 M thiourea and 50 mM ammonium bicarbonate, pH 8.5 and kept at room temperature for 1h. Every 15 minutes, the pellet was crushed with a pipette tip, vortexed for 5 seconds and sonicated in an ice bath (Branson 3510, Danbury, CT, USA) for 5 more seconds to improve solubilisation. Protein

concentration was determined with the Bio-Rad (Hercules, CA, USA) RC-DC protein assay. Finally, the protein extract was centrifuged (16000 g, 20°C) to remove insoluble cell debris.

2.4 Proteolysis

2.4.1 Standard protein mixture

Equal amounts of serum albumin (BSA, *Bos taurus*), galactosidase (*Escherichia coli*), casein α (consisting of isoforms S1 and S2), casein β (*Bos Taurus*), ovalbumin (*Gallus gallus*) and Histone type II-A (*Bos Taurus*, containing Histone type II-B as an impurity) were dissolved in a buffer containing 6 M urea, 2 M thiourea and 50 mM ammonium bicarbonate (pH 8,5) each at a concentration of 500 ng/ μ L. The proteins were reduced by the addition of 5 mM dithiotreitol (DTT) and kept at room temperature for 30 minutes. Afterwards, 15 mM of iodoacetamide (IAM) was added and the proteins were kept in the dark for 45 minutes at room temperature for carbamidomethylation of the reduced cysteins. After observation of N-overalkylation of the peptide N-termini (see results below), the reaction of IAM was quenched by adding 25 mM DTT after alkylation. For proteolysis, the protein mixture was diluted at least 4 times in 50 mM NH_4HCO_3 and trypsin was added to a 1:60 ratio. The proteins were digested overnight at 37°C. Digests were preserved at -20°C until use.

2.4.2 Digests of plant cell suspension culture proteins

The resolubilized proteins were reduced by the addition of 5 mM DTT for 30 minutes at room temperature. Next, 15 mM IAM was added and the proteins were kept in the dark and at room temperature for 45 minutes. Afterwards, the excess of IAM was quenched by the addition of 25

mM DTT. The protein mixture was diluted 6 times in 50 mM ammonium bicarbonate and digested with trypsin (1:60) overnight at a temperature of 37°C. The digests were immediately desalted and further processed or analysed. For ITRAQ labeling (4-plex) of *Nicotiana tabacum* digests, the ammonium bicarbonate was replaced by the same concentration of triethylammonium bicarbonate. The labeling was performed according to the manufacturers' instructions.

2.5 Desalting

Peptide mixtures were desalted prior to SCX chromatography by solid phase extraction based on the "Stop and go extraction"-tip (STAGE-tip) protocol described by Rappsilber *et al.* [30] for small amounts of digest (<20 µg) or "in house" packed SPE cartridges for larger amounts. STAGE-tips were assembled by stamping out a small plug of C₈ material from a 3M Empore C₈ extraction disk using a HPLC syringe needle. Subsequently, this plug was placed in the constricted end of a GELoader tip (Eppendorf). Next, a methanolic slurry containing 500 µg of POROS Oligo R3 C₁₈ beads (Applied Biosystems) was packed on top of the plug. The STAGE-tip was washed with 100 µL of a 95% acetonitrile solution (v/v) and equilibrated with 100 µL water containing 0.5% (v/v) formic acid. The peptide samples were acidified with 5 µL of 10% (v/v) formic acid and loaded on the equilibrated STAGE-tip. The column was washed with 150 µL of the 0.5% formic acid solution and finally eluted in 100 µL of a solution comprised of 70% acetonitrile and 0.5% formic acid (v/v). For larger samples, empty fritted 1 mL cartridges (Machery-Nagel, Duren, Germany) were used and they were packed with 20 mg of POROS Oligo R3 by methanolic slurry packing. Washing of the column was done by adding 500 µL of 95%

acetonitrile, followed by equilibration in 1 mL of 0.5% formic acid. The acidified sample was loaded on the column and the column was subsequently washed with 2 mL of 0.5% formic acid. The retained peptides were eluted with 300 μ L of a solution containing 70% acetonitrile and 0.5% formic acid. All eluted samples were concentrated by vacuum centrifugation to evaporate the organic modifier used for elution.

2.6 SCX chromatography

Strong cation exchange chromatography of digested peptides was accomplished on an AKTA Basic 10 FPLC system (GE Healthcare, Uppsala, Sweden) using a 1 mL Resource S column (GE Healthcare) as described by Gruhler *et al.* [31] with modifications. The desalted peptide mixture was loaded onto a 100 μ L PEEK injection loop and separated after injection by a linear gradient from 100% buffer A (5 mM ammonium formate, pH 2.7, 30% (v/v) acetonitrile) to 100% buffer B (1 M ammonium formate, pH 3, 30% (v/v) acetonitrile) in 24 column volumes at a flowrate of 600 μ L/min. Fractions of 1.2 mL were automatically collected throughout the entire gradient (including the column flow through and washing out of unbound sample). Afterwards, the system was washed with 5 column volumes of 100% buffer B and re-equilibrated with 10 column volumes of 100% buffer A. Both UV absorption (280 nm) and buffer conductivity were monitored throughout the gradient. The collected fractions were immediately evaporated by vacuum centrifugation in an Eppendorf concentrator 5301 for 100 minutes at 40°C or stored at -20°C overnight for evaporation the next day.

2.7 LC-MS/MS analysis

A 2D nanoLC ultra system (Eksigent, Dublin, CA, USA) equipped with an Endurance 920 autosampler (Spark Holland, Emmen, the Netherlands) was connected to an LTQ-Orbitrap Velos mass spectrometer (Thermo Electron, Bremen, Germany). The trapping column was a 20 mm length x 200 μ m ID Acclaim pepmap 100 column (Dionex, Amsterdam, the Netherlands), which is coupled to a 150 mm L x 75 μ m ID Acclaim pepmap 100 analytical column (3 μ m beads) (Dionex). Peptides were injected onto the trapping column at a flowrate of 3 μ L/min. in 100 % solvent A (2% acetonitrile and 0.1% formic acid in water). A subsequent water/acetonitrile gradient to 40% solvent B (75% acetonitrile and 0.1% formic acid in water) in 1h elutes and separates the peptides at a flowrate of 350 nL/min. The LC was interfaced with a Triversa chip-based electrospray source (Advion Biosystems, Ithaca, NY, USA)

The LTQ-Orbitrap Velos mass spectrometer was operated in data dependent acquisition mode. Full scan MS spectra were acquired in the Orbitrap at 60000 resolution (measured at 400 m/z) after accumulation to a target value of 5×10^5 ions in the linear ion trap and with a maximal injection time of 500 ms. A polysiloxane ion was used as lock mass for continuous mass calibration (m/z 445.120030). In each recorded full scan spectrum, the 10 most intense ions (with z at least 2) were selected for collision induced dissociation (CID) in the LTQ Velos at a normalized collision energy of 35% after accumulation to a target value of 5000 ions and with a maximal injection time of 50 ms. Activation time was set to 10 ms. After MS/MS acquisition, the parent ion mass was dynamically excluded for MS/MS for 45 seconds and with a 5 ppm window.

2.8 Data analysis.

MS and MS/MS spectra were analysed in Xcalibur 2.1.0 (Thermo Electron) and protein identifications were acquired through the SEQUEST and Mascot search engines embedded in the Proteome Discoverer software 1.4 (Thermo Electron). The MS tolerance was set to 5 ppm and the fragment mass tolerance to 0.8 Da. One missed cleavage was allowed. The following modifications were included unless noted otherwise: carbamidomethylation of cysteine as fixed and oxidation of methionine and phosphorylation of serine, threonine and tyrosine as variable modifications. The SWISSPROT database was used to analyse spectra from the standard protein experiments and TAIR9 (version September 13th 2009) for the experiments with peptides originating from *Arabidopsis* cell suspension cultures. After database searching, the result files were loaded into the Scaffold 4.3.4 software suite (Proteome sSoftware Inc, Portland, OR, USA). The peptide false discovery rate was determined at 1% via the PeptideProphet algorithm (with Scaffold delta mass correction). For protein identification, a minimum of 2 peptides was mandatory. Protein probabilities were assigned using the ProteinProphet algorithm and a 1% protein FDR was established. Proteins that contained similar peptides and could not be discriminated based on MS/MS fragmentation were grouped into clusters to fulfill the principles of parsimony.

For the *Nicotiana tabacum* analyses, *Viridiplantae* database was used that was downloaded from the Uniprot website using following url: [http://www.uniprot.org/uniprot/?query=\(viridiplantae\)+AND+reviewed%3ano&force=yes&format=fasta&include=yes](http://www.uniprot.org/uniprot/?query=(viridiplantae)+AND+reviewed%3ano&force=yes&format=fasta&include=yes) . A target-decoy

analysis was performed with a reversed database. A false discovery rate of 1% was allowed for confident peptide identification.

Gene ontology (GO) analysis was performed with the GO tool available on the Arabidopsis TAIR website (<http://www.arabidopsis.org/tools/bulk/go/index.jsp>) and GO enrichment analysis was accomplished using the AgriGO tool for singular enrichment analysis (SEA) (<http://bioinfo.cau.edu.cn/agriGO/analysis.php>). A hypergeometric test and the Benjamini–Hochberg [32] correction were used to test the statistical significance of the enrichment of each given GO term in a list of proteins with respect to a reference list. Only functional categories with minimum 5 mapped entries are analysed. As a reference list, gene ontology annotation of *Arabidopsis thaliana* was downloaded from the GO website (www.geneontology.org). The significance threshold was set to $p < 0.05$ after p-value adjustment by the Benjamini and Hochberg procedure for multiple testing.

Data of identified spectra from the *Arabidopsis* SCX experiment were analysed by loading them into Matlab (MathWorks, Natick, MA, USA, version R2009b) after conversion to an ASCII data format by the RawXtract software [33]. Matlab was used to normalize the data and analyse the orthogonality between SCX fractionation and reversed phase chromatography and also in the calculation of the MA-plot.

The *Nicotiana tabacum* data was also analyzed with the PEAKS Studio software (Bioinformatics solutions Inc, Waterloo, Ontario, Canada, version 6). For the PEAKS *de novo* searches and the PEAKS DB database searches, the same search parameters were applied as described above.

3. Results & discussion

3.1 Experimental workflow

The sample preparation protocol was optimized based on the procedure described for 2DGE analyses by Laukens and colleagues [34]. This included a protein precipitation step with trichloroacetic acid and acetone in order to remove substances (e.g. phenolic compounds) potentially interfering with downstream analysis steps such as mass spectrometry. First, the sample preparation procedure was tested and optimized on a mixture of standard proteins to ensure reproducibility and recovery of the tryptic peptides. During optimization, a problem of alkylation of peptide amino termini upon trypsin digestion was encountered. This problem was remediated by the addition of a low concentration of DTT in the latter step (data not shown). A similar phenomenon was described before by Boja and Fales [35]. The applied workflow for sample preparation further included a STAGE-tip desalting step [30] to remove (thio)urea after proteolysis and for sample concentration. Afterwards, desalted peptide samples were analysed by nano-scale reversed phase chromatography (RPLC) coupled to mass spectrometry. An experimental workflow scheme including sample preparation, separation and analysis steps is shown in figure 1. After data acquisition, peptide identification via a database searching was applied. The PeptideProphet algorithm was used to obtain a stringent ($\leq 1\%$) peptide false discovery rate. This was combined with a protein false discovery rate of 1%, as calculated via the ProteinProphet algorithm [36], in order to provide reliable protein identifications.

3.2 Identification of *Arabidopsis thaliana* proteins

The optimized experimental workflow was further applied to an *Arabidopsis thaliana* cv. 'Landsberg erecta' plant cell suspension culture. For this one-dimensional LC-MS experiment, an amount corresponding to 50 ng of extracted protein was injected on the nanoscale C18 column and a RPLC gradient of 60 minutes (2% to 30% acetonitrile) was applied.

Many in-solution digestion protocols use rather high concentrations of urea to increase protein unfolding. In solution, a small fraction of urea will decompose to cyanic acid in a temperature-facilitated mode. Cyanic acid can react with side chains of arginine or lysine and N-terminal amine groups to form carbamylated residues. All resolubilisation and further sample preparation steps were carried out at room temperature. Also, centrifugation steps were executed at 20°C. Protein samples were vortexed only very briefly (< 5 sec.) and sonication was carried out in an ice-bath to prevent excessive temperature increase. Reduction of cysteins was accomplished at room temperature instead of 56°C, a temperature frequently used in in-gel digestion protocols [37]. Under these conditions, only 20 carbamylated peptides were identified, corresponding to about 1% of the total number of identified peptides. This is similar to the ratio reported by others [38]. The fraction of peptides that contained a carbamidomethylated N-terminus as a result of overalkylation by iodoacetamide was lower than 1%.

In this sample, 4160 unique tryptic peptides were identified (1% peptide FDR) corresponding to 737 different proteins (at 1% protein FDR). These results are in good accordance with data described by Second and colleagues, who applied a comparable nanoLC-MS setup for the analysis of *Caenorhabditis elegans* peptides [39]. These results prove that sample preparation,

chromatography and mass spectrometric analysis worked properly for *Arabidopsis thaliana* cell suspension cultures.

In order to further examine the quality of the sample preparation protocol the features of the identified proteins were assessed by their GO classifications for cellular component and molecular function and were compared to the classifications from the entire *Arabidopsis* genome. The results of this analysis are displayed in figure 2. From these pie charts it becomes clear that we were not able to identify the same proportion of proteins with “unknown” classification that make up a large portion of the *Arabidopsis* genome (31% and 30% for cellular component and molecular function respectively). To further assess whether our sample preparation protocol was biased towards specific cellular components, the dataset was further analysed using the AgriGO algorithm for singular enrichment analysis (<http://bioinfo.cau.edu.cn/agriGO/analysis.php>). The only cellular component that was significantly under-represented in our dataset was the endomembrane system (GO:0012505, p-value 2.8×10^{-16}). Moreover, only 5% of the identified proteins contained predicted transmembrane domains, whereas this was 24% for the total *Arabidopsis* proteome (see supplemental data). It should be noted that interpreting singular enrichment analysis results with the *Arabidopsis* genome as a reference to reflect methodological biases should be done with caution; the use of a genome wide reference does not take into account whether proteins are actually expressed under the given conditions. Moreover, next to protein extraction, mass spectrometric identification of peptides is known to be biased towards certain physicochemical properties as well [40]. The extraction of membrane proteins could further be facilitated by the addition of detergents during protein extraction. The use of detergents could however seriously

hamper downstream mass spectrometric analysis. Recently, novel extraction techniques for membrane proteins have been optimized to overcome this problem [37].

3.3 Prefractionation via strong cation exchange chromatography

To further investigate the comprehensiveness of the sample preparation protocol and the analysis strategy, a prefractionation step consisting of strong cation exchange chromatography of the tryptic peptides was added. Here, an offline strategy was opted for as this adds flexibility to the workflow. The addition of an SCX step under acidic conditions, based on the work by Gruhler and colleagues [31], would also allow preliminary enrichment of certain acidic classes of tryptic peptides that are of special biological interest, such as phosphorylated peptides or acetylated peptides [27].

To examine the suitability of the method and the compatibility of the SCX prefractionation step, the digest obtained from the mixture of standard proteins was applied and clearly more peptides were identified compared to reversed phase chromatography alone. Then, a peptide mixture was prepared from an *Arabidopsis thaliana* cell suspension culture and 200 µg was fractionated by offline SCX. Afterwards, the resulting fractions were injected in a randomized order on the same nanoLC-MS platform as described before.

Analysis of the fragmentation spectra with the SEQUEST search engine resulted in the identification of 19106 unique peptides with a peptide false discovery rate of 1% from 2900 different proteins with false discovery rate of 1% at the protein level. As depicted in figure 3, peptides were identified throughout the entire SCX gradient, considering the fact that fractions 1-4 comprise dead volume (including column flow through) and washing out of unbound

sample in 100% buffer A. A remarkable increase was observed in fraction 8 where more than 3500 peptides were identified. This increase is also apparent in reports using similar SCX strategies [41] and arises from the elution of doubly charged tryptic peptides. It can also be concluded from this figure that the overlap of peptides identified in consecutive SCX fractions is generally below 15%. This low overlap implies that elution peak width in SCX is generally below 2 minutes and ascertains that MS analysis time is not occupied exuberantly with peptides that were already identified in previous fractions.

As stated above, SCX chromatography is frequently used for its ability to enrich phosphorylated peptides. During this analysis, 213 phosphorylated peptides could be identified, without the implementation of specific mass spectrometric strategies such as neutral loss dependent MS³ or other fragmentation techniques such as electron transfer dissociation. Figure 4 indicates that, while the total number of identified phosphopeptides remained similar over the first 14 fractions, up to 95% of the identified peptides in that fraction were shown to be phosphorylated, without further phosphopeptide enrichment. Moreover, 54 unique n-terminal acetylated peptides were identified. Thirteen of these peptides contained at least one histidine residue and eluted during the first half of the gradient (fractions 7 to 14). Over 95% of the N-acetylated peptides without a histidine eluted in the column flow-through or subsequent washing step. These data indicate that our proteome analysis platform can be used for the enrichment of several classes of acidic peptides.

The prefractionation step by SCX chromatography was implemented to increase the total peak capacity of the analysis platform. An increased peak capacity can only be accomplished if the

two dimensions of separation (i.e. SCX and reversed phase chromatography) are orthogonal. In order to test orthogonality, the reversed phase retention time of all identified peptides was plotted against the SCX fraction in which they were identified (figure 5). This plot shows that in most of the fractions, peptides elute during the entire gradient and good orthogonality is achieved. However, true orthogonality is not accomplished and the increase in peak capacity by switching from one- to two-dimensional LC is not simply multiplicative. This observation can be explained by the role of charge density in the SCX retention mechanism, which causes larger peptides to be less retained by the SCX resin compared to smaller peptides with the same net charge [42].

3.4 Validation of the strategy on *Nicotiana tabacum* BY-2 cultures

A plant cell suspension culture of a system without a publicly available nuclear genome sequence was used to validate the analytical platform described above. In this experiment, all the sample preparation steps were kept constant. After acquisition, the data was analyzed via different approaches in order to optimize the cross species identification rate. In all cases, a compiled database containing *Viridiplantae* protein sequences was used. First, a standard database searching approach was applied using the SEQUEST algorithm as described for the *Arabidopsis thaliana* cell suspension culture. In total, 1273 unique peptides were identified. Next to this, the spectra were also interrogated via the Mascot database searching algorithm using the same search parameters and an identical database. This search yielded positive identifications for 1146 unique peptides.

Afterwards, the spectra were analyzed by the PEAKS DB and PEAKS PTM tools in the PEAKS Studio software. The PEAKS DB tool incorporates de novo sequencing results generated by the PEAKS algorithm [43] into a database sequencing approach and has been shown to improve both the number of identifications and their accuracy [44]. The PEAKS PTM tool aids in identifying post-translationally modified peptides [45]. Using these tools, significantly more unique peptides and proteins could be identified in other non-model species compared to traditional database searching. In total 2986 unique peptides were identified by the PEAKS Studios software. The result and overlap of the identification approaches is depicted in figure 6a. Remarkably, only 705 peptide sequences were identified by all approaches. Moreover, the PEAKS de novo assisted database search was able to add 1564 unique peptide identifications. In total, 1202 proteins could be identified by at least 2 unique peptides. It must be noted however that while our *Viridiplantae* database was RefSeq filtered (i.e. one single sequence for each protein per model organism), performing cross-species identification with databases containing multiple species is bound to yield some redundancy. To make the results from the different search engines comparable, the InChorus tool was used to group sequence identification conditional on the tandem MS spectrum. As a reference, Duby and colleagues recently reported the identification of 532 proteins from *Nicotiana tabacum* BY-2 cells by a 2DGE approach coupled with MALDI TOF-TOF identification [46]. The same research group later identified 670 *Nicotiana tabacum* proteins via 2DGE and 858 via LC-MS analyses [47]. These are, to our knowledge, the largest repositories of identified *Nicotiana tabacum* proteins published to date. Taken together, these data demonstrate that our shotgun proteome analysis approach is also amenable to non-genomic model species.

Moreover, the feasibility of implementing isobaric label-based quantification in our setup was assessed. Here, all sample preparation parameters were constant but the ammonium bicarbonate in the digestion buffer was replaced by triethylammonium bicarbonate and the resulting tryptic peptides were labeled with iTRAQ (isobaric tags for relative and absolute quantitation) 4-plex labels in order to enable multiplexing and the relative quantitation of different samples. In this proof-of-concept study, a tryptic digest of extracted *Nicotiana tabacum* proteins was split in half and then labeled with a different label. Afterwards, equal amounts of the differently labeled peptides were mixed to create a self-self analysis and then analyzed by one-dimensional reversed phase LC-MS. A MA-plot [48] demonstrating the efficacy and accuracy of the labeling procedure can be found in figure 6b, where the ordinate M (log ratio's) and the abscissa A (mean average of $\log_2$ intensities) for each peptide are calculated as

$$M = \log_2(114) - \log_2(117)$$

$$A = \frac{1}{2} \{ \log_2(114) + \log_2(117) \}$$

where (114) and (117) represent the intensity of the iTRAQ reporter ions at m/z 114 and m/z 117, respectively. The technical precision of the iTRAQ labelling was assessed by a replicate same-sample analysis of a standard protein mixture (data not shown). In order to analyze the iTRAQ labeling efficiency, MS/MS spectra of labeled peptides were acquired in HCD mode. In total, 2682 unique *Nicotiana tabacum* peptides were identified by the combinatorial identification approach (data not shown). This demonstrates that isobaric labeling based quantification can be implemented in our setup.

4. Conclusions

Here we provide a generic platform for analysis of the soluble proteome based on chromatography and mass spectrometry. The number of identified proteins and protein coverage was increased by the implementation of SCX as an orthogonal prefractionation step. Furthermore, it was demonstrated that this analytical platform can also be effectively used for proteome analysis of non-genomic model species. Finally the incorporation of peptide isobaric labels during sample preparation was shown to be feasible on this analytical platform. This will eventually enable us to relatively quantify protein abundance profiles between plant cell suspension cultures that are in a different physiological state.

Reference List

- [1] J. Van Leene, H. Stals, D. Eeckhout, G. Persiau, E.V. De Slijke, G. Van Isterdael, A. De Clercq, E. Bonnet, K. Laukens, N. Remmerie, K. Henderickx, T. De Vijlder, A. Abdelkrim, A. Pharazyn, H. Van Onckelen, D. Inze, E. Witters, and G. De Jaeger, A tandem affinity purification-based technology platform to study the cell cycle interactome in *Arabidopsis thaliana*, *Mol. Cell. Proteomics* 6 (2007) 1226-1238.
- [2] J. Van Leene, J. Hollunder, D. Eeckhout, G. Persiau, E. Van de Slijke, H. Stals, G. Van Isterdael, A. Verkest, S. Neiryndck, Y. Buffel, S. De Bodt, S. Maere, K. Laukens, A. Pharazyn, P.C.G. Ferreira, N. Eloy, C. Renne, C. Meyer, J.D. Faure, J. Steinbrenner, J. Beynon, J.C. Larkin, Y. Van de Peer, P. Hilson, M. Kuiper, L. De Veylder, H. Van Onckelen, D. Inze, E. Witters, and G. De Jaeger, Targeted interactomics reveals a complex core cell cycle machinery in *Arabidopsis thaliana*, *Mol. Sys. Biol.* 6 (2010).
- [3] P. Campanoni, B. Blasius, and P. Nick, Auxin transport synchronizes the pattern of cell division in a tobacco cell line, *Plant Physiol.* 133 (2003) 1251-1260.
- [4] A. Goossens, S.T. Hakkinen, I. Laakso, T. Seppanen-Laakso, S. Biondi, V. De Sutter, F. Lammertyn, A.M. Nuutila, H. Soderlund, M. Zabeau, D. Inze, and K.M. Oksman-Caldentey, A functional genomics approach toward the understanding of secondary metabolism in plant cells, *Proc. Natl Acad. Sci U.S.A.* 100 (2003) 8595-8600.

- [5] B.A. Wolucka, A. Goossens, and D. Inze, Methyl jasmonate stimulates the de novo biosynthesis of vitamin C in plant cell suspensions, *J. Exp. Bot.* 56 (2005) 2527-2538.
- [6] I.K. Vasil, A history of plant biotechnology: from the Cell Theory of Schleiden and Schwann to biotech crops, *Plant Cell Rep.* 27 (2008) 1423-1440.
- [7] G. Melchers and U. Engelmann, Die Kultur von Pflanzegevebe in flussigen Medium mit Dauerbeluftung, *Naturwissenschaften* 20 (1959) 564-565.
- [8] W. Tulecke and L.G. Nickell, Production of large amounts of plant tissue by submerged culture, *Science* 130 (1959) 863-864.
- [9] W. Steck, B.K. Bailey, J.P. Shyluk, and O.L. Gamborg, Coumarins and alkaloids from cell cultures of *Ruta graveolens*, *Phytochem.* 10 (191) 191-194.
- [10] V. Papageorgiou, A. Assimopoulou, V. Samanidou, and I. Papadoyannis, Recent advances in chemistry, biology and biotechnology of Alkannins and Shikonins, *Curr. Org. Chem.* 10 (2006) 2123-2142.
- [11] S. Malik, R.M. Cusido, M.H. Mirjalili, E. Moyano, J. Palazon, and M. Bonfill, Production of the anticancer drug taxol in *Taxus baccata* suspension cultures: A review, *Process Biochem.* 46 (2011) 23-34.
- [12] S. Hellwig, J. Drossard, R.M. Twyman, and R. Fischer, Plant cell cultures for the production of recombinant proteins, *Nature Biotechnol.* 22 (2004) 1415-1422.
- [13] J. Xu, S. Okada, L. Tan, K.J. Goodrum, J.J. Kopchick, and M.J. Kieliszewski, Human growth hormone expressed in tobacco cells as an arabinogalactan-protein fusion glycoprotein has a prolonged serum life, *Transgenic Res.* 19 (2010) 849-867.

- [14] T.H. Kwon, J.E. Seo, J. Kim, J.H. Lee, Y.S. Jang, and M.S. Yang, Expression and secretion of the heterodimeric protein interleukin-12 in plant cell suspension culture, *Biotechnol. Bioeng.* 81 (2003) 870-875.
- [15] S. Matsumoto, K. Ikura, M. Ueda, and R. Sasaki, Characterization of A Human Glycoprotein (Erythropoietin) Produced in Cultured Tobacco Cells, *Plant Mol. Biol.* 27 (1995) 1163-1172.
- [16] J.V. Jorin, A.M. Maldonado, and M.A. Castillejo, Plant proteome analysis: A 2006 update, *Proteomics* 7 (2007) 2947-2962.
- [17] K. Baerenfaller, J. Grossmann, M.A. Grobe, R. Hull, M. Hirsch-Hoffmann, S. Yalovsky, P. Zimmermann, U. Grossniklaus, W. Gruissem, and S. Baginsky, Genome-scale proteomics reveals *Arabidopsis thaliana* gene models and proteome dynamics, *Science* 320 (2008) 938-941.
- [18] C.G. Gammulla, D. Pascovici, B.J. Atwell, and P.A. Haynes, Differential metabolic response of cultured rice (*Oryza sativa*) cells exposed to high- and low-temperature stress, *Proteomics* 10 (2010) 3001-3019.
- [19] K.G. Zulak, M.F. Khan, J. Alcantara, D.C. Schriemer, and P.J. Facchini, Plant Defense Responses in Opium Poppy Cell Cultures Revealed by Liquid Chromatography-Tandem Mass Spectrometry *Proteomics, Mol. Cell. Proteomics* 8 (2009) 86-98.
- [20] Y.P. Gai, X.L. Ji, W. Lu, X.J. Han, G.D. Yang, and C.C. Zheng, A Novel Late Embryogenesis Abundant Like Protein Associated with Chilling Stress in *Nicotiana tabacum* cv. Bright Yellow-2 Cell Suspension Culture, *Mol. Cell. Proteomics* 10 (2011).

- [21] A. Vertommen, AL. Møller, JH. Cordewener, R. Swennen, B. Panis, C. Finnie, AH. America, SC. Carpentier. A workflow for peptide-based proteomics in a poorly sequenced plant: a case study on the plasma membrane proteome of banana, *J. Proteomics*, 74 (8), 2011, 1218-1229.
- [22] GK. Agrawal, A. Sarkar, PG. Righetti, R. Pedreschi, S. Carpentier, T. Wang, BJ. Barkla, A. Kohli, BK. Ndimba, NV. Bykova, C. Rampitsch, L. Zolla, MS. Rafudeen, R. Cramer, LV. Bindschedler, N. sakirpaloglou, RJ. NdimbaJ, JM. Farrant, J. Renaut, D. Job, S. Kikuchi, R. Rakwal. A decade of plant proteomics and mass spectrometry: translation of technical advancements to food security and safety issues, *Mass Spectrom. Rev.*, 32(5), 2013, 335-365.
- [23] H. Vanderschuren, E. Lentz, I. Zainuddin, W. Gruissem. Proteomics of model and crop plant species: status, current limitations and strategic advances for crop improvement. *J. Proteomics*, 93, 2013, 5-19. [24] I. Desgagne-Penix, M.F. Khan, D.C. Schriemer, D. Cram, J. Nowak, and P.J. Facchini, Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures, *Bmc Plant Biol.* 10 (2010).
- [25] J. Bally, C. Job, M. Belghazi, and D. Job, Metabolic Adaptation in Transplastomic Plants Massively Accumulating Recombinant Proteins, *Plos One* 6 (2011).
- [26] A.L. Schillmiller, F. Pichersky, and R.L. Last, Taming the hydra of specialized metabolism: how systems biology and comparative approaches are revolutionizing plant biochemistry, *Curr. Op. Plant Biol.* 15 (2012) 338-344.
- [27] S. Mohammed and A.J. Heck, Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications, *Curr. Op. Biotechnol.* 22 (2011) 9-16.

- [28] M. Menges , J.A. Murray, Synchronous Arabidopsis suspension cultures for analysis of cell-cycle gene activity, *Plant J.* 30 (2002) 203-212
- [29] T. Nagata, Y. Nemoto, and S. Hasezawa, Tobacco By-2 Cell-Line As the HeLa-Cell in the Cell Biology of Higher-Plants, *Int. Rev. Cytol.* 132 (1992) 1-30.
- [30] J. Rappsilber, Y. Ishihama, and M. Mann, Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics, *Anal. Chem.* 75 (2003) 663-670.
- [31] A. Gruhler, J.V. Olsen, S. Mohammed, P. Mortensen, N.J. Faergeman, M. Mann, and O.N. Jensen, Quantitative phosphoproteomics applied to the yeast pheromone signaling pathway, *Mol. Cell. Proteomics*, 4 (2005) 310-327.
- [32] Y. Benjamini and Y. Hochberg, Controlling the False Discovery Rate - A Practical and Powerful Approach to Multiple Testing, *J. Roy. Stat. Soc. B Met.* 57 (1995) 289-300.
- [33] W.H. McDonald, D.L. Tabb, R.G. Sadygov, M.J. MacCoss, J. Venable, J. Graumann J, J.R. Johnson, D. Cociorva, J.R. 3rd Yates, MS1, MS2, and SQT—three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications, *Rapid Comm. Mass Spectr.* 18 (2004) 2162–2168.
- [34] K. Laukens, P. Deckers, E. Esmans, H. Van Onckelen, and E. Witters, Construction of a two-dimensional gel electrophoresis protein database for the *Nicotiana tabacum* cv. Bright Yellow-2 cell suspension culture, *Proteomics* 4 (2004) 720-727.
- [35] E.S. Boja and H.M. Fales, Overalkylation of a protein digest with iodoacetamide, *Anal. Chem.* 73 (2001) 3576-3582.

- [36] Al. Nesvizhskii, A. Keller, E. Kolker and R. Aebersold. A statistical model for identifying proteins by tandem mass spectrometry, *Anal. Chem.* 75 (2003) 4646-4658.
- [37] A. Shevchenko, H. Tomas, J. Havlis, J.V. Olsen, and M. Mann, In-gel digestion for mass spectrometric characterization of proteins and proteomes, *Nature Prot.* 1 (2006) 2856-2860.
- [38] J.R. Wisniewski, A. Zougman, N. Nagaraj, and M. Mann, Universal sample preparation method for proteome analysis, *Nature Met.* 6 (2009) 359-U60.
- [39] T.P. Second, J.D. Blethrow, J.C. Schwartz, G.E. Merrihew, M.J. MacCoss, D.L. Swaney, J.D. Russell, J.J. Coon, and V. Zabrouskov, Dual-Pressure Linear Ion Trap Mass Spectrometer Improving the Analysis of Complex Protein Mixtures, *Anal. Chem.* 81 (2009) 7757-7765.
- [40] B. Louie, R. Higdon, and E. Kolker, The necessity of adjusting tests of protein category enrichment in discovery proteomics, *Bioinformatics*, 26 (2010) 3007-3011.
- [41] J.J. Benschop, S. Mohammed, M. O'Flaherty, A.J.R. Heck, M. Slijper, and F.L.H. Menke, Quantitative phosphoproteomics of early elicitor signaling in Arabidopsis. *Mol. Cell. Proteomics* 6 (2007) 1198-1214.
- [42] M. Gilar, P. Olivova, A.E. Daly, and J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, *Anal. Chem.* 77 (2005) 6426-6434.
- [43] B. Ma, K.Z. Zhang, C. Hendrie, C.Z. Liang, M. Li, A. Doherty-Kirby, and G. Lajoie, PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry, *Rapid Comm. Mass Spectr.* 17 (2003) 2337-2342.

- [44] J. Zhang, L. Xin, B. Shan, W. Chen, M. Xie, D. Yuen, W. Zhang, Z. Zhang, G.A. Lajoie, and B. Ma, PEAKS DB: De Novo Sequencing Assisted Database Search for Sensitive and Accurate Peptide Identification, *Mol. Cell. Proteomics* 11 (2012).
- [45] X. Han, L. He, L. Xin, B. Shan, and B. Ma, PeaksPTM: Mass Spectrometry-Based Identification of Peptides with Unspecified Modifications, *J. Proteome Res.* 10 (2011) 2930-2936.
- [46] G. Duby, H. Degand, A.M. Faber, and M. Boutry, The proteome complement of *Nicotiana tabacum* Bright-Yellow-2 culture cells, *Proteomics* 10 (2010) 2545-2550.
- [47] E. Van Cutsem, G. Simonart, H. Degand, A.M. Faber, P. Morsomme, and M. Boutry, Gel-based and gel-free proteomic analysis of *Nicotiana tabacum* trichomes identifies proteins involved in secondary metabolism and in the (a)biotic stress response, *Proteomics* 11 (2011) 440-454.
- [48] J.M. Bland and D.G. Altman, Statistical Methods for Assessing Agreement Between Two Methods of Clinical Measurement, *Lancet* 1 (1986) 307-310.

Figure captions:

1. Diagram of the protocol used in sample preparation and fractionation. Sample homogenization and protein extraction was followed by TCA-acetone precipitation. After resolubilisation, proteins were digested with trypsin. Tryptic peptides were desalted through SPE (C18) and directly analysed via reversed phase nanoLC-MS/MS or fractionated by strong cation exchange chromatography before nanoLC-MS/MS analysis.

2. Distribution of Gene ontology annotations for cellular components of the identified proteins.

The distribution of the identified proteins in the *Arabidopsis* 1D experiment (without fractionation, left), the *Arabidopsis* 2D experiment (with fractionation, middle) and the entire *Arabidopsis* genome (right) are shown.

3. Distribution and overlap of identified peptides between consecutive fractions in SCX chromatography.

This graph displays the total number of identified unique peptides in each fraction of the SCX chromatography (grey bars) and the number of peptides that were also identified in the previous fraction (black bars). The black line illustrates the percentage of peptides that were found in two consecutive fractions of the SCX chromatography.

4. Distribution of phosphorylated peptides across the SCX chromatography.

The black bars represent the number of unique phosphorylated peptides identified in the respective fraction while the gray line denotes the percentage of peptides identified in the fraction that contained at least one phosphate group.

5. Orthogonality plot of 2D-separated peptides from *Arabidopsis thaliana*. The SCX fraction is shown in the ordinate and the retention time of the peptide in the reversed phase chromatography is shown in the abscis. The color code displays the charge state (z) of the peptides in the mass spectrum. A correlation of -0.52 was found between SCX fraction and retention time of the reversed phase chromatography.

6. Validation of the strategy in the non-genomic model *Nicotiana tabacum*.

a. Venn diagram showing the number of identified unique peptides using different database search engines (SEQUEST, Mascot) and the *de novo* assisted PEAKS DB tool. *b.* MA-plot for a same-sample analysis of a *Nicotiana tabacum* iTRAQ-labeled peptide sample. Every blue dot represents an iTRAQ-labeled peptide. The green line is a LOWESS curve which can be used for data normalization.

Highlights:

- A generic approach for LC-MS based proteome analysis of plant cell suspension cultures was developed
- Implementation of SCX prefractionation improved identification rates and sequence coverage
- The approach was successfully tested on the non-genomic model *N. tabacum* 'BY-2'
- The platform is amenable for quantitative analyses via isobaric labeling

Cellular components

Without fractionation

With fractionation

Arabidopsis genome

Molecular functions

Without fractionation

With fractionation

Arabidopsis genome