

Whole cell proteome regulation by microRNAs captured in a pulsed SILAC Mass spectrometry approach

Olivia Ebner¹, olivia.ebner@mdc-berlin.de

Matthias Selbach¹, matthias.selbach@mdc-berlin.de

¹Max Delbrück Center for Molecular Medicine, Robert-Rössle Str. 10, D-13125 Berlin, Germany

Abstract

Since gene expression is controlled on many different levels in a cell, capturing a comprehensive snapshot of all regulatory processes is a difficult task. One possibility to monitor effective changes within a cell is to directly quantify changes in protein synthesis which reflects the accumulative impact of regulatory mechanisms on gene expression. *Pulsed stable isotope labelling by amino acids in cell culture* (pSILAC) has been shown to be a viable method to investigate *de novo* protein synthesis on a proteome-wide scale (1, 2). One application of pSILAC is to study the regulation of protein expression by microRNAs. Here, we describe how pSILAC in conjunction with shotgun mass spectrometry can assess differences in the protein profile between cells transfected with a microRNA and non-transfected cells.

Key Words: Mass spectrometry; LC-MS/MS; SILAC; microRNA; seed; transfection; in gel digestion

Introduction

Measuring how gene expression changes in response to a stimulus can provide instructive insights into biological systems. Gene expression is regulated at all stages from DNA via mRNA to the protein. Most methods exclusively quantify changes in steady state mRNA levels, neglecting posttranscriptional regulatory mechanisms. Here, we describe *pulsed stable isotope labelling by amino acids in cell culture* (pSILAC) as a method to quantify changes in protein production at a global scale. pSILAC measures the actual output of gene expression and can therefore reveal regulation at all levels. The method is particularly useful to study regulation at the level of translation. As an example, we show how the method can be used to quantify the effect of microRNAs on cellular protein production.

The technology behind pSILAC is mass spectrometry-based proteomics (3). The general workflow is that proteins in a sample are digested into peptides. The resulting peptide mixture is separated by reversed phase liquid chromatography (LC). At the end of the chromatographic column, eluting peptides are directly transferred into the orifice of a mass spectrometer by a process called electrospray ionization (ESI). The mass spectrometer performs two important tasks. First, it measures the masses and intensities of the peptides in the mixture at any given time during the LC run. Second, the device fragments individual peptides and measures masses and intensities of the fragments (tandem mass spectrometry or MS/MS). The information about the masses of the non-fragmented peptides and their fragment spectra can be used to identify the peptides and hence the proteins present in the sample. In addition to identifying peptides and proteins, it is also necessary to quantify changes in their abundance. In mass spectrometry, this is most accurately achieved by stable isotope labelling (4) : Incorporating heavy stable (i.e. non-radioactive) isotopes into peptides leads to a shift in mass. Differentially labelled samples can be combined and analyzed together so that all peptide peaks will occur in pairs. The ratio of peak intensities of such

peptide pairs accurately reflects differences in their abundance. In *stable isotope labelling by amino acids in cell culture* (SILAC) the label is introduced metabolically. Cells are cultivated in growth medium containing heavy-stable isotope versions of essential amino acids (5). After several cell generations, all proteins have incorporated the heavy label. Mixing heavy and light cells can reveal changes in steady state protein levels between both samples.

pSILAC is a variant of the SILAC approach (1). In contrast to standard SILAC, cells are first cultivated in growth medium with the normal light (L) amino acids. Concomitantly with differential treatment, cells are transferred to culture medium containing heavy (H) or medium-heavy (M) amino acids. All newly synthesized proteins will be made in the H or M form, respectively. Subsequently, both samples are combined and analyzed together. The abundance ratio of H versus M peptides reflects changes in protein production. Thus, pSILAC measures differences in protein synthesis integrated over incubation time of then pulse labelling period. In principle, any essential amino acid can be used for SILAC. We prefer lysine and arginine because the protease trypsin cleaves C-terminal of these residues. Therefore, all tryptic peptides except for the protein C-terminus contain a label and can be quantified. Heavy and medium heavy lysine (Lys8, Lys4) has a mass shift of eight and four daltons compared to the normal light form (Lys0), respectively. Similarly, we use light (Arg0), medium-heavy (Arg6) and heavy (Arg10) arginine.

pSILAC is particularly useful to quantify changes in protein production induced by microRNAs (2). The protocol described here uses overexpression of short double-stranded RNAs designed to mimic endogenous microRNAs. Note that it is also possible to knock-down endogenous microRNAs, to use cells from microRNA knock-out animals or to quantify protein production in different non-microRNA contexts. The first section describes how samples are prepared by transfecting cells with the microRNA mimic. Separation of proteins by SDS-PAGE is used as an approach to reduce sample complexity in order to achieve

deeper proteome coverage. In the second section, mass spectrometry is described for an LTQ-Orbitrap system. In the last section, we indicate how the raw data can be analyzed to identify and quantify proteins and to investigate microRNA mediated effects. See Fig.1 for an overview of the described technique.

The protocol described here relies heavily on mass spectrometry-based quantitative proteomics. While the overall procedure is straight forward, the success depends on many technical details. Therefore, experience with mass spectrometry and subsequent data analysis is generally required. Alternatively, we suggest getting in touch with an expert lab while planning the project.

Materials

In general HPLC and mass spectrometry require reagents and solvents of the highest available grade (HPLC grade or better). Reagent grades are indicated in the respective subsections.

pSILAC preparation of samples for mass spectrometry

Determination of transfection efficiency of BLOCK-iT™ Fluorescent Oligos

BLOCK-iT™ fluorescent oligo (Invitrogen)

4% paraformaldehyde (PFA)

Round cover slips (Menzel GmbH & Co. KG)

D-PBS (phosphate buffered saline) modified- without calcium chloride and magnesium chloride, sterile-filtered, liquid, cell culture tested (1X, Invitrogen).

Epifluorescence microscope (Leica DM-R)

Transfection, labelling and harvest of cells

HeLa cells or cell line of choice

D-PBS (1X, modified- without calcium chloride and magnesium chloride, sterile-filtered, cell culture tested, Invitrogen)

Trypsin-EDTA (0.05% Trypsin with EDTA 4Na, 1X, Invitrogen)

SILAC amino acids: non-labelled (for “light” medium) L-Lysine monohydrochloride and L-Arginine monohydrochloride (Sigma-Aldrich); (“Lys 4”) 4,4,5,5,-D₄-L-lysine monohydrochloride and (“Arg 6”) L-Arginine-¹³C₆ monohydrochloride (Sigma-Aldrich, cat. no. 643440) for “medium-heavy” medium; (“Lys 8”) L-lysine-¹³C₆ ¹⁵N₂ monohydrochloride (Sigma-Aldrich, cat. no. 608041) and (“Arg 10”) L-arginine-¹³C₆ ¹⁵N₄ monohydrochloride (Sigma-Aldrich, cat. no. 608033) for “heavy” medium

Amino acid stock solutions: Dissolve 84 mg/ml arginine or 146 mg/ml lysine respectively in D-PBS. Store in small aliquots at -20°C. High concentrations are necessary to avoid dilution of the culture medium. Stocks for stable isotope-labelled amino acids are prepared the same way. Sterile filtration is not necessary at this step as they will be filtered during the medium preparation.

Dulbecco’s Modified Eagle’s Medium (DMEM) High Glucose (4,5g/l) w/o L-Arg, L-Lys, L-Glu (PAA, custom preparation) supplemented with 10% sterile-filtered dialyzed fetal bovine serum (dFBS, Sigma-Aldrich) and 4mM stable Glutamine (L-alanyl-L-glutamine, PAA). Dialyzed serum is required since normal serum contains amino acids. Add dFBS and glutamine to medium, shake to mix and fill into two 250 ml to filtration devices. Add 1:3000 stock: medium (total amount of arginine 28 mg/L plus

48 mg/L of lysine/500ml medium) of non-labelled lysine and arginine for “light” (L; 0/0) SILAC medium; Lys 4 and Arg 6 for “medium heavy” (M; 4/6) SILAC medium; Lys8 and Arg 10 for “heavy” (H; 8/10) SILAC medium

Note that the given amino acid concentrations are optimized for HeLa cells. Other cell lines may require adjustments to minimize arginine to proline conversion (4, 5)

Synthetic microRNAs (Dharmacon): Prepare aliquot of 2 mM following the manufacturers protocol

DharmaFECT 1 Transfection Reagent (Dharmacon). Store at 4°C

Cell lifter (Costar)

Vacuum filtration with bottle, PES membrane 0.22µm (TPP) Techno Plastic Products AG

Cell lysis and SDS-PAGE

NanoDrop 2000 (Thermo scientific) or alternative device to measure protein concentration

RIPA buffer: 50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1% Triton-X100, 1% Sodium deoxycholate and 0.1% SDS; use a 0.22µm filter for sterile filtration

Benzonase® Nuclease-5KU (Sigma-Aldrich, optional)

NuPAGE Novex 4 to 12% gradient gels; LDS Sample buffer 4x, MES SDS Running buffer. 20x (Invitrogen)

1M DL-Dithiothreitol (DTT) in H₂O_{bidest.}; store 1ml aliquots at -20°C. Harmful, prepare with caution

Prestained molecular weight marker: SeeBlue Plus 2 Prestained Standard (Invitrogen)

Fixing solution: 20ml H₂O_{bidest.}, 25ml Methanol (MeOH), 5ml Acetic Acid (HAc, puriss. p.a., eluent additive for LC-MS, Sigma-Aldrich)

Staining Buffer A: 27.5ml H₂O_{bidest.}, 10ml MeOH, 10ml Novex StainerA (Invitrogen))

Staining Buffer B: Add 2.5ml Novex Stainer B (Invitrogen) to Stainer A buffer in tray.

In gel protein digestion

50mM Ammonium BiCarbonate (ABC, puriss. p.a., Sigma-Aldrich): Dissolve 40mg ABC in 10ml H₂O (LC-MS Chromasolv, Sigma-Aldrich). Store at RT

Ethanol (EtOH, Ethanol gradient grade LiChrosolv, Merck)

Sequence grade modified trypsin (Promega)

Trypsin solution: Dilute Trypsin (0.5µg/µl) in 50mM ABC. Demanded Protein: Enzyme ratio = 50:1. Estimated protein amount per lane = 200 µg. Prepare immediately before use and keep it always on ice to minimize autocatalysis. Undiluted stocks are stored in small aliquots at -80 °C. Trypsin is sensitive to high urea concentrations. Concentration should be below 2M urea/thiourea

Iodacetamide solution (55 mM) in 50 mM ABC: Dissolve 10.2 mg iodacetamide in 1 ml ABC. Prepare fresh or store in small aliquots at -20°C. Keep in the dark.

10 mM DL-Dithiothreitol (DTT) in 50mM ABC. To make 1 ml dilute 10 µl of a 1M DTT solution in 990 µl ABC and store in small aliquots at -20°C

Extraction solution: 3 % trifluoroacetic acid (TFA, puriss. p.a., eluent additive for LC-MS, ≥99.0% (GC), Sigma-Aldrich), 30 % acetonitrile. To make 1ml dilute 300 µl ACN (LC-MS CHROMASOLV, ≥99.9%, Sigma-Aldrich) and 30 µl TFA in 670 µl H₂O (LC-MS Chromasolv, Sigma-Aldrich). Store at RT

Buffer A: 3 % TFA, 5 % ACN in H₂O (LC-MS Chromasolv, Sigma-Aldrich). Store at RT

Desalting and purification by C₁₈ Stage Tips

C₁₈ Empore 47mm Disks (3M)

Buffer B: 0,5% HAc, 80% ACN in H₂O (LC-MS Chromasolv, Sigma-Aldrich). Store at RT

Mass spectrometry of the samples

Nano-flow HPLC system

For HPLC columns: Pack your own columns with ReproSil-Pur 120 C18-AQ, 3 µm beads HPLC bulk packing material (Dr Maisch GmbH, Germany) and a 360 mm OD, 75 mm ID pulled capillary column. Columns can also be bought ready-to-use from several companies.

High performance mass spectrometer. Ideally, this should be an instrument with high resolving power, high dynamic range, high speed and high sensitivity. We describe the protocol for an LTQ-Orbitrap system (Thermo Fisher).

Processing of mass spectrometry data

Processing of raw data

We use the freely available software package MaxQuant for peptide identification, protein assembly and quantification (6). Alternatively, MSQuant (<http://msquant.sourceforge.net/>) or other software packages capable of SILAC-based quantification can be used.

Possibilities of data analysis

Among the open or commercial (programming) tools for data analysis, we found R (Bioconductor package) and PERL extremely useful for handling big datasets. Other tools with one main focus are e.g. Sylamer, miReduce, Cytoscape, and STRING.

Methods

A common effect in microRNA overexpression experiments is the down-regulation of many (i.e. hundreds) of proteins. Yet many of these proteins are only mildly regulated (2). To be able to discern the effect and signature of the microRNA in a dataset a high transfection efficiency of the microRNA is required. Thus previous to the transfection of cells for proteomics, a test of transfection efficiency via BLOCK-iT™ Fluorescent Oligos should be performed. To accurately determine the effect of the transfected microRNA it is also necessary to treat both microRNA transfected cells and mock-transfected controls in the very same way. The cultures should have similar confluence of cells and the transfection procedure of microRNA samples and controls should be the same with the one exception; omitting the microRNA in the control. After transfection controls are usually plated into heavy (Lys8/Arg10) medium whereas microRNA samples are plated into medium-heavy SILAC medium (Lys4/Arg6). Identical treatment for mass spectrometry sample preparation is assured by combining both control and microRNA sample upon harvest (See Fig. 1). Separation of lysed proteins is achieved via SDS-PAGE and single slices of the SDS-gel are subsequently subjected to reduction, alkylation and digestion into peptides with trypsin. Prior to the digestion step, clean handling of the samples is essential to avoid contamination with contaminants such as keratin. Following the extraction of the peptides from the gel slices the samples are desalted on reverse-phase C₁₈ STAGE-tips. This step also provides further filtering and concentration of the samples. For mass spec analysis samples are eluted from the columns and further separated by HPLC. At the end of the HPLC column the peptides are ionized (electrospray ionization, ESI) and analyzed by tandem mass spectrometry (LTQ-Orbitrap). Raw data files are analyzed with the MaxQuant software package, which is able to detect peaks and SILAC labelled peptide triplets, identify peptides by database searching and

deduce proteins from the latter. For further practical and background information about SILAC and related methods also consult the available *Nature Protocols* (5-8).

pSILAC preparation of samples for mass spectrometry

Determination of transfection efficiency of BLOCK-iT™ Fluorescent Oligos

As a non-targeted dsRNA oligomer, the BLOCK-iT fluorescent oligo resembles a microRNA and the efficiency of its transfection can be measured via fluorescence microscopy. Conduct the transfection of cells (on coverslips) for oligos as well as for controls as described in point 3.2. In this case however, 6well culture plates with area-adjusted amounts of reagents are employed (*see Note 1*).

Wash the cells after 8h of transfection with D-PBS once and fix them with 4% paraformaldehyde (PFA) in D-PBS. To determine the transfection efficiency compare the fluorescence of oligo-transfected with non-transfected cells under the epifluorescence microscope.

Transfection of cells, labelling and harvest

The transfection is carried out using DharmaFECT1 according to the manufacturer's protocol. In parallel, control transfections of HeLa cells for each microRNA sample are done under the same conditions expect for serum-free DMEM replacing the synthetic microRNA. Using another short dsRNA as a control (for example with scrambled sequence) is generally not recommended because every short dsRNA will have off-target effects on protein production making result interpretation more difficult.

1. Prior to the transfection, cells are cultured in "light" SILAC medium for a week in order to give them time to adjust to these growth conditions. Split the cells every three days and cultivate 37°C with 5% CO₂.

2. Plate the cells one day before transfection on 10cm² dishes in “light” SILAC medium and incubate at 37°C with 5% CO₂ overnight. Confluence at the time of transfection should be 60-70% (*see Note2*).
3. Add 300µl microRNA (2µM; final plating concentration of 100nM) to 300µl serum-free DMEM and incubate for 5min at RT. Mix gently.
4. Add 16µl DharmaFECT1 to 584µl serum-free DMEM and incubate for 5min at RT. Mix gently.
5. Combine both, mix by slow pipetting and incubate for 20min at RT. Be careful not to disrupt the Lipid-microRNA complexes from this point on - only very slow pipetting!
6. Remove medium from cells, rinse with D-PBS once and then add 4.8ml “light” SILAC medium.
7. Slowly pipet the transfection mix (1.2ml in total) onto the cells and gently rock the plate to ensure equal distribution.
8. At 8h post transfection wash cells twice with D-PBS and change the SILAC medium of microRNA samples to “medium-heavy” (Lys4/Arg6) and of the controls to “heavy” (Lys8/Arg10).
9. Wash once again with D-PBS 24h later, scrape cells off in 6ml ice-cold D-PBS, combine the microRNA sample and control and centrifuge for 10 min at 600 x g and 4°C. Pellets can be stored at -20°C.

Lysis and SDS-PAGE

Resuspend cell pellets in 250-400µl RIPA buffer and lyse for 20min on ice with occasional vortexing. After centrifugation (14,000 rpm for 10min in a microfuge at 4°C) transfer the supernatant to a fresh microtube. If the supernatant appears to be viscous, this is due to DNA. To reduce viscosity, employ Benzonase.

Measure protein concentration of the whole-cell lysate via the NanoDrop 2000 ProteinA280 method or other suitable methods.

For SDS-Page you can use NuPAGE Novex 4 to 12% gradient gels with MES as running buffer according to the manufacturers protocol. However, other denaturing gel electrophoresis systems work as well. Note that precast gels are generally preferable because acrylamide in self-made gels can introduce protein modifications. Optimal load of protein is approximately 200µg per lane although it is possible to use as little as 50µg of sample.

Mix the desired amount of lysate with 1:20 of DTT (1M) and 1:4 of NuPage buffer and incubate at 75°C for 10min. Subsequently load the samples onto the gel together with 7.5µl of the protein marker and run 180V for approximately 50min (i.e. until the marker reaches the lower boundary of the gel).

Treat the gel first with fixing solution and then staining buffer A for 10min each. Next, incubate gel with staining buffer B for 30-60min (or until bands are visible). Wash the gel with ddH₂O once or twice until the solution of the gel appears to be clear. This step functions mainly to make sure that the gel ran smoothly and enough protein has been loaded. Loading the same sample twice on two separate lanes to have enough protein for in gel digestion is possible as well.

In gel protein digestion

All steps before the trypsin digestion step should be performed under a flow hood. Use extreme care not to contaminate samples with keratin from skin or other contaminants (*see Note3*). The samples should always be incubated using a microtube shaker. Prior to the gel extraction step, the supernatants are discarded at each step. Use enough solution to cover the gel slices. Washing can be done with an excess of solution (around 500µl) whereas extraction

solutions generally should not exceed 300µl to contain sample volume for evaporation. Unless stated otherwise all steps are done at RT.

Cut the gel into 12 slices on a clean plastic foil. Chop each slice into smaller pieces (approx. 1 x 1 mm) and place the pieces of one slice in a clean 1.5ml microtube (*see Note4*). Cutting and transferring slices is easier with a gel that is neither too dry nor too wet. When possible, it is best to cut out individually stained bands. However, for regions of low protein content where individual proteins are not clearly discernable, one can combine 2-3 slices (*see Note5*).

Wash gel pieces with 1:1 ABC/EtOH for 20 min. Gel pieces should be clear before proceeding to the trypsin digestion step. If they are clear, wash again with ABC for 20 min and then ABC/EtOH for 20 min.

Dehydrate the gel pieces by incubating for 10 min in absolute EtOH.

Dry the samples in a speed-vac for 10-15 min until the gel pieces are bouncing in the tube (*see Note6*). Samples can be stored at 4°C for several days at this point.

Rehydrate the gel pieces and reduce the proteins by incubating for 45-60 min in DTT solution at 56 °C. Discard all the liquid afterwards.

Block free sulphhydryl groups by incubating for 45 min in iodacetamide at RT in the dark.

Wash gel pieces once with ABC for 20 min at RT.

Dehydrate the gel pieces by incubating for 10 min in absolute EtOH.

Remove remaining ethanol from gel pieces by vacuum centrifugation. Samples can be stored at 4°C for several days at this point.

Add enough trypsin solution at 4 °C to cover the dehydrated gel pieces and place tubes on ice. Make sure that the gel slices are fully covered with trypsin solution (otherwise add ABC) after swelling as much as possible (approx. 20 min). Place in microtube shaker at 37 °C over night.

Add 2 µl TFA to stop the digestion and quickly finger-vortex the solution. Spin down the gel pieces at low speed in a microfuge and transfer the liquid to a fresh tube.

Extract the gel pieces by adding Extraction solution to cover the gel. Shake the mixture vigorously for 10 min at RT. Remove the liquid and combine with that from step 11.

Dehydrate gel pieces in 100 % ACN for 10 min at RT. Spin down the gel pieces, recover the supernatant and combine with supernatant from steps 11 - 12.

Dry the samples in a Speed-Vac until 10-20 % original volume to remove ACN. Adjust the samples to a low pH with buffer A by adding approx. 50 µl to the samples. The resulting pH should be < 2.5.

Desalting and purification by C₁₈ Stage Tips

1. Prepare as many desalting columns (Stage Tips) as necessary by punching out small discs of C₁₈ Empore Filter and eject the discs (three per tip) into a P200 pipette tip using a blunt-ended syringe needle (Fig.2). Ensure that the discs are securely wedged in the bottom of the tip.
2. Condition tips by adding 50µl methanol to the Empore disc. Use this step to check whether the Stage Tips are leaky. For centrifugation (5 min at 5000 rpm) a standard table-top centrifuge can be used. Punch a hole into the lid of 2 ml microtubes and place the Stage Tips into them. Discard flow-through once it

reaches the end of the tip and always make sure the Stage Tips are empty before adding the next solution.

3. Remove any remaining organic solvent in the column by adding 100µl buffer A to the disc and centrifuge for 5 min at 5000 rpm.
4. Force the acidified peptide sample (From 3.4, Nr.14) through the Stage Tip at 5000 rpm for 5 min.
5. Wash the column with 120µl buffer A for another 5 min at 5000 rpm. Samples can be stored at this point at 4°C for up to one year. When using tips that have been stored for a longer period, wash again with washing buffer.
6. Elute the peptides from the C₁₈ material using 50 µl buffer B. Elute directly into microtube the autosampler plate of your HPLC system (*see Note7*).
7. Dry the autosampler plates in the Speed-Vac at 45°C until all acetonitrile has evaporated (~3 µl remaining final volume). (Do not over dry because otherwise you will lose most of the sample) For two MS measurements of twice 6µl add sample buffer to a final volume of approx. 15µl (*see Note8*).

Mass spectrometry of the samples

While sample preparation for shotgun mass spectrometry can be done in almost every wet lab, setting up the resources for both mass spectrometry itself and subsequent data analysis requires investment in machines and computing power as well as expertise in handling both. The following protocol can thus provide the rough outlines of sample processing after in gel digestion. The analysis is performed using liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). Peptides are separated on a nano-flow HPLC system. At the end of the column they are directly ionized and transferred to the mass spectrometer. The mass spectrometer determines masses and intensities of the eluting peptides during the gradient

(MS). In addition, it isolates individual peptides for fragmentation and records their fragment spectra. Ideally, the machine should be able to fragment all peptides eluting from the column. However, due to limited dynamic range, sensitivity and speed, only a fraction of peptides are picked in each run. Measuring every sample twice alleviates undersampling.

1. Analyze the eluted samples via LC-MS/MS on a high performance mass spectrometer (LTQ-Orbitrap). Coupled to the LTQ-Orbitrap a HPLC nano-flow system (for example Eksigent) separates peptides by reversed phase chromatography on C₁₈ columns directly connected to the electrospray ion source.
2. Inject each sample twice with an injection volume of 6µl and a 150 min flow gradient ranging from 10 to 60% acetonitrile in 0.5% acetic acid at a flow rate of 250 nl/min. There are several alternative gradients that can be used at this point. Supervision by a scientist experienced in LC-MS/MS is recommended.
3. The Orbitrap performs precursor ion scan/survey MS spectra (*m/z range of 300–1700*; resolution $R = 60,000$; target value of 1×10^6 ; profile mode). For MS/MS isolate five to ten of the most intense peaks (monoisotopic precursor selection enabled; charge state ≥ 2 ; target value 5,000), fragment via collision induced dissociation (normalized collision energy 35%; wideband activation enabled; centroid mode) in the LTQ part of the LTQ-Orbitrap and exclusion selected peptides using dynamic exclusion (duration 60 seconds, exclusion list size 500). For an example of a full MS scan showing a SILAC triplet see Fig.3. .

Processing of mass spectrometry data

Pre-processing of raw data

While there are different ways to process the obtained spectra we employ the MaxQuant software package. This package consists of several modules with adjustable parameters. In most cases, the default parameter settings can be used. More information about MaxQuant can be found online (www.maxquant.org).

Settings for Quant are Arg6 and Lys4 as “medium-heavy”, Arg10 and Lys8 as “heavy” labels with maximum of three labelled amino acids per peptide. Top 6 MS/MS peaks per 100 Da are used and polymer detection is enabled. For the database search use carbamidomethylation of cysteines as fixed modification. Acetylation of the protein N-terminus and oxidation of methionine are used as variable modifications. A maximum of two missed cleavages and specificity of cleavage by trypsin should be required. Set the mass tolerance to 0.5 Da for low-resolution fragment ions and mass accuracy cut-off to 7 ppm for the high resolution parent ion. Note that this is just the initial value: MaxQuant will determine the optimal cut-off value by target-decoy-database searching.

Set-up a suitable protein database for the search. In general, an organism-specific database should be used (for example from www.uniprot.org). To this database you should add common contaminants (human/sheep keratins, trypsin, bovine serum albumin etc.).

For peptide and protein identification we normally require of at least two identified peptides per protein group in total, one of them being unique in the database. The minimum peptide length is set to six amino acids. Quantification is done only for protein groups with at least two ratio counts. A more conservative analysis uses only

proteins with three ratio counts as the median of three peptide ratios is outlier-robust. With the protocol described here it should be possible to identify at least 4000-5000 and to quantify 3000-4000 proteins per sample

Possibilities of data analysis

Besides the pre-processing via MaxQuant the data may still need further steps of filtering before being ready for downstream analysis. Recheck your list for contaminants such as keratins that might have been missed because they were not in the contaminant part of the database. Proteins with very high/low ratios should be examined carefully. As a rule of thumb, proteins quantified based on many counts are reliable despite a high variability of individual peptide ratios. For proteins represented by only a few peptide counts it may be necessary to check individual cases. For example, you can check the posterior error probability (PEP) of proteins and peptides to see how confident individual peptides/proteins were identified. If you have experience with mass spectrometry you can also look at individual fragment spectra and try to assign all peaks to expected fragments.

For data analysis it is useful to transform measured ratios into logarithmic fold changes; we use the $-\log_2$ of the heavy to medium-heavy normalized ratios of the MaxQuant output (H/M normalized) to visualise the data. The two main subsets that can be distinguished in microRNA related datasets are proteins that possess a seed site in the 3'UTRs of their respective mRNAs versus proteins without such seed sites. Even though the effect of microRNAs on many targets may not be explained by 3'UTR seed-mediated regulation, this is one of the most feasible ways to identify likely targets of a microRNA. To make sure that overexpression of the microRNA worked, monitor whether the seed signature of the microRNA is seen in the data. Sylamer is a tool to assess nucleotide sequences over- or under-represented in data sorted list of genes. When searching for enriched sequences of 6nt length, the microRNA seed sequence should be enriched mainly in the down-regulated

fraction (Fig.4). Furthermore the comparison of cumulative fractions of all proteins versus seed-containing proteins should have a bias of the latter to negative fold changes. Another way to validate data is to investigate how many known targets of the microRNA are found to be down-regulated in the mass spectrometry data. Network relationships can be obtained using tools like Cytoscape, PathVisio or STRING. To assess functional implications of the microRNA targets, especially the seed-containing ones, Gene Ontology analysis in conjunction with hierarchical clustering is a commonly used method.

Notes

The transfection protocol has been optimized for HeLa cells but can be adapted to other cell lines as well. If you choose a different cell line make sure that DharmaFECT1 has high transfection efficiency in the chosen cell line. Otherwise you might need to try different lipids for transfection. The same is true for transfection and labelling times; depending on the subject of investigation shorter or longer times until the harvest of the cells might be required.

Make sure the confluence of cells does not exceed 70% at the time of transfection; in order to sufficiently SILAC label long-lived proteins the cells need to be able to further grow after transfection in order to incorporate the labelled amino acids.

While in gel digestion is not very error prone it is crucial to avoid contaminants because they will be predominant in the mass spectrometry spectra. Therefore work under the hood and make sure all surfaces are clean. Do not work over open microtubes.

We use the inner lining of small autoclaving waste bags for cutting the gel but any clean plastic foil will do.

A good rule for cutting the gel is to keep regions with a thick distinct protein band in a single slice while pooling regions of lower protein content where you don't see clearly distinguished protein bands. In general you can cut the gel into as many pieces as you like but keep in mind that each slice prolongs measuring time. Accordingly, you may not necessarily get better results from increasing the number of gel slices.

Be careful when handling the gel cubes in the microtubes; especially after speed-vac evaporation they are prone to "jump" from the tube.

For elution from stage tips we recommend to use a 20 or 50ml syringe. Push out the sample by hand by sticking the tip on top of the syringe; do not elute too fast but rather in slow drops in order to elute most peptides on the column.

Be careful when diluting the final sample in the autosampler plate. You will need a few more μl of sample than is actually injected because of the HPLC settings. Yet, since the sample is contained in a very small volume at that time, every added μl of buffer A will lead to a strong dilution of the sample. For good MS results you should keep the sample as concentrated as possible.

Acknowledgments

The authors would like to thank Björn Schwanhäusser for providing data and templates for the shown figures.

Legends to Figures

Fig1: Experimental setup of the method described in this chapter. Cells were cultivated in “light” (L) SILAC medium and subsequently mock- (control) or microRNA transfected. Pulse labeling was conducted after 8h hours by transferring the control samples to “heavy” (H) and the microRNA transfected samples to “medium-heavy” (M) SILAC medium. 24h later, cells were harvested and combined. Sample preparation of both control and microRNA samples together ensures identical processing and comparability. Only M and H peaks represent newly synthesized proteins in the following mass spectrometry of the sample. Changes in protein production are reflected in the H/M ratio of peptide intensity peaks.

Fig3: Peptide mass spectra displaying pulsed SILAC triplets of Destrin (Sequence in parenthesis). The full MS scan of the miR-34 transfected HeLa sample shows the light (L), medium-heavy (M) and heavy (H) monoisotopic peaks of Destrin. The doubly charged peptide contains labelled arginine which leads to a shift of 3 Da for the M label and 5 Da for the H label (Arg6 and Arg10, due to the double charge the shift is cut in halve). The H/M ratio of Destrin in this sample is 1.12, indicating a reduction of Destrin in the miR-34 transfected sample of $-0.05 \log_2$ fold change.

Fig2: A, Preparation of Stage Tips. Small discs of C₁₈ Empore Filter are excelled with a blunt-ended syringe needle and inserted into the tip. Discs should be pushed to the end of the tip and wedged securely but not too tightly (impedes flow-through) to the bottom of the tip.

B, Final Stage Tip. The C₁₈ Empore Filter consists of chromatographic beads embedded in a Teflon mesh and is ready-to-use for micro-purification. To activate, wash and to load the sample, liquid is always added from the top and passes through the column by either centrifugation or pressure from an air-filled plastic syringe. Until elution the sample is retained on the beads.

Fig4: Sylamer analysis of HeLa cells transfected with miR-16. Proteins quantified in the sample are sorted in bins from most down regulated (left) to most up regulated (right) on the X axis. The nucleotide (nt) sequences of the 3'UTRs of these proteins were analyzed for enrichment of 6 nt long sequences. The hypergeometric significance for each leading bin of all detected 6nt sequences is shown on the Y axis scaled as \log_{10} (P-value). Enrichment is represented on the positive Y axis whereas depletion in negative values. The E-value threshold of 0.01 (Bonferroni-corrected) is indicated as horizontal dotted line. Significance curves for all sequence motifs are drawn along the X axis. Most of the sequences are not significantly enriched, however one sticks out of the gray shaded background. This curve represents the miR-16 seed sequence of GCTGCT. The profile shows a clear peak in the region of down regulated proteins as signature of the microRNA with lesser but significant enrichment in the whole protein fraction.

References

1. Schwanhausser, B., Gossen, M., Dittmar, G., and Selbach, M. (2009) Global analysis of cellular protein translation by pulsed SILAC, *Proteomics* **9**, 205-209.
2. Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008) Widespread changes in protein synthesis induced by microRNAs, *Nature* **455**, 58-63.
3. Aebersold, R., and Mann, M. (2003) Mass spectrometry-based proteomics, *Nature* **422**, 198-207.
4. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review, *Anal Bioanal Chem* **389**, 1017-1031.
5. Ong, S. E., and Mann, M. (2006) A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC), *Nat Protoc* **1**, 2650-2660.
6. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J. V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics, *Nat Protoc* **4**, 698-705.
7. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes, *Nat Protoc* **1**, 2856-2860.
8. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips, *Nat Protoc* **2**, 1896-1906.