Published in final edited form as:
Curr Protoc Chem Biol. 2011 ; 3(4): 181–196. doi:10.1002/9780470559277.ch110193.

Analyzing *In Vivo* Metabolite-Protein Interactions By Large-Scale Systematic Analyses

Xiyan Li and **Michael Snyder**
Department of Genetics, Stanford University, Stanford, California 94305

Abstract

Metabolites interact with proteins *in vivo* in various ways other than enzymatic reactions. Profiling of such interactions may help disclose unknown molecular mechanisms that regulate protein functions, and provide potential targets for disease treatment. Here we describe a procedure for systematic analyses of metabolite-protein interactions *in vivo*. This procedure couples protein affinity purification and mass spectrometry to identify metabolite-protein interactions. The primary effort can be completed within one day and scaled to process hundreds of samples in a batch. Originally developed in yeast, the same principle and protocol can be adapted to other organisms.

Keywords

Metabolite-protein interaction; liquid chromatography; mass spectrometry; LC-MS; metabolite; protein affinity purification; yeast

Introduction

In the cell, metabolites are a group of small compounds whose presence and levels are often dynamically regulated during the process of growth and differentiation, as well as in a diversity of biological responses. Although typically thought of primarily simple building blocks and energy sources within a cell, metabolites may also interact with proteins in various ways other than enzymatic reactions, many of which can regulate protein functions, such as those found in modulation of enzyme activity, regulation of transcription, signaling and neural transmission. Profiling of such *in vivo* interactions provides important clues to the molecular mechanisms that modulate protein functions through physical interaction. We have developed a systematic procedure to study *in vivo* metabolite-protein interactions in yeast, which may be adapted to other organisms, as long as certain criteria are met (see Commentary) (Li et al., 2010). No prior knowledge is required for the protein of interest in this assay. The whole procedure involves protein expression, protein affinity purification, metabolite extraction, liquid chromatography-coupled mass spectrometry (LC-MS), and extensive data analysis. Depending on the throughput of mass spectrometry, this procedure can be conveniently scaled up to process several hundred samples at once.

Strategic Planning

A general strategy for studying metabolite-protein interactions is shown in figure 1. The basic procedure involves purification of a protein of interest, elution of the bound small moleucles and separation and identification of small molecules using LC-MS. Several factors are crucial for successful experiments. First, an appropriate protein-expressing

Correspondence: mpsnyder@stanford.edu.

system is crucial to produce biological meaningful and reliable data in this experiment. Since there can be potential metabolomic variation between species and experimental conditions, proteins of interest should be produced in their natural hosting cells whenever possible. Second, rapid purification of the protein using either an epitope tag or capture agent is valuable. The former is typically used for high throughput analyses, and we often tag proteins using the ZZ domain of protein A which binds IgG beads with very high affinity (Lowenadler et al., 1987; Nilsson et al., 1987). The ZZ domain consists of 116 amino acid residues, and the whole tag is about 19 kD (Gelperin et al., 2005). However, cautions should be taken in immunoassays because the ZZ domain interacts strongly with most primary antibodies raised against specific proteins. A summary of additional important factors in experimental design is described in Figure 2 and also covered in details in following sections.

Another important parameter is the choice of an appropriate LC-MS method, which is the most challenging part in this experiment. Because metabolites have enormous chemical diversity, it is not possible to analyze them using one general LC-MS method and several different methods must be used to analyze as many metabolites as possible (as described in Figure 3). Two general methods are described in this protocol with a preference for the analysis of hydrophobic and hydrophilic metabolites, respectively. However, more sensitive methods are possible by focusing on analyzing a particular group of metabolites at a time (e.g. hydrophobic or hydrophilic molecules).

Basic Protocol 1

Affinity Purification of Yeast Protein and Extraction of Protein-Interacting Metabolites

Protein A (ZZ domain)-tagged protein is adsorbed on rabbit IgG-conjugated magnetic beads (Dynabeads) in a solution system that is compatible with mass spectrometry. The protein-bound metabolites are then extracted for LC-MS analyses. The protein yield is also assessed by gel staining afterwards. A graphic summary this procedure is described in Figure 1.

Materials

Cell pellet (stored at −80 °C, see basic protocol 2)

Zirconia silica beads (Biospec)

Rabbit IgG-conjugated dynabeads (see basic protocol 3)

Lysis buffer (see recipe)

Wash buffer 1 (see recipe)

Wash buffer 2 (see recipe)

Methanol, acetonitrile, ethanol and water (mass spec grade)

2× Laemmli SDS sample buffer (for SDS-PAGE)

15-well 4-12% NuPAGE Bis-Tris gel (Invitrogen)

Page Ruler Plus prestained protein ladder (Fermentas)

20× NuPAGE MOPS SDS running buffer (Invitrogen NP0001-02)

ProtoBlue Safe Colloidal Coomassie staining solution (National Diagnostics EC-722)

Gel drying solution

Eppendorf Protein Lobind tubes (2.0 ml and 1.5 ml)

Glass vials with inserts (Mass spec consumables)

FastPrep cell lyser with an adapter for 2 ml tubes

Gel drying frame

Cellophane membrane

Refrigerated microcentrifuge

Hula mixer (Invitrogen) or similar product

Magnetic stand for 1.5/2.0 ml tubes

Heat blocks maintained at 42 °C and 95 °C

Non-filter polypropylene pipette tips

Nitrile gloves

Step 1-8 should be carried out in a cold room at 4 °C.

Non-filter pipette tips should be used to avoid introducing polymers that are often found in filters.

Nitrile gloves are preferred to create a cleaner background on LC-MS.

1. Wash appropriate amount of rabbit IgG dynabeads three times in lysis buffer without DTT or protease inhibitors. To do this, each time separate beads from the supernatant on a magnetic stand, remove supernatant, add lysis buffer, and then invert-mix on the mixer for 3 min at 4 °C. Then suspend beads in lysis buffer and prepare aliquots into clean 2 ml LoBind tubes for step 3.

   We use 50 μl (~1.25 mg dynabeads) IgG dynabeads per sample, which is able to produce sufficient amount of protein for visualization by Coomassie blue staining. More dynabeads can be used to produce higher protein yield.

   Briefly spin 5-10 sec before separation on magnet to help recover beads.

2. Add equal volume of 0.5 mm Zirconia/silica beads (stored at −20 °C) to yeast cell pellet, as prepared from 150 ml culture (A600nm = 1.0). Add 950 μl Lysis buffer. Lyse on FastPrep 24, 3 × 40s min at 6.0 m/s with 2 min interval on ice.

   Cell pellets can be stored at −80 °C for 6-12 months with no apparent loss of yield.

   Glass beads (1.0 mm) or ceramic beads (1.4 mm, if milder lysis condition is desired) may be used with comparable performance.

3. Centrifuge at 20,000 ×g for 10 min at 4 °C, and transfer the supernatant (lysate) to a 2.0 ml Lobind tubes. Store at 4 °C.

   Be careful not to transfer cell debris that sometimes layers above Zirconia/silica beads.

4. Add 950 μl lysis buffer to the cell pellet and perform lysis again as step 2.

5. Repeat step 3 and combine and transfer the lysate into the 2 ml LoBind tube containing IgG dynabeads from step 1.

   Be careful not to transfer cell debris that sometimes layers above Zirconia/silica beads.

6. Invert-mix for 30-45 min at 4 °C on a Hula mixer at 20 rpm.

   The mixing speed should be fast enough to prevent dynabeads from settling to the side. A homogenate brownish suspension is expected in this step.

7. Separate and remove the lysate in the supernatant on a magnetic stand. Add 0.95 ml wash buffer 1 to the beads, mix on a mixer for 5 min at 4 °C.

   This step washes off most non-specific proteins in high molecular weight range.

   Briefly spin 5-10 sec before separation on magnet to help recover beads.

8. Separate and remove wash buffer 1 in the supernatant on a magnetic stand. Add 0.95 ml wash buffer 2, transfer bead suspension to a fresh 1.5 ml LoBind tube. Invert-mix on a mixer for 5 min at 4 °C.

   Briefly spin 5-10 sec before separation on magnet to help recover beads.

   The use of a fresh tube is to eliminate non-specific adsorption of metabolites to the old tube.

9. Separate and remove wash buffer 2 in the supernatant on a magnetic stand. Add 60 μl 75% v/v ethanol (MS grade) to the beads, pipette mix and heat at 42 °C on a heat block for 15 min.

   High organic composition in 75% ethanol is used to disrupt the metabolite-protein interaction and dissolve the protein-bound metabolites while keeping out proteins and other unwanted polymers.

   75% v/v ethanol serves as a general solvent that is suited for most metabolites. Ethanol can be replaced with pure methanol for most hydrophobic metabolites. Other solvents are possible choices, yet have not been tested.

   The volume of extraction solvent can be modified, according to the actual requirement for mass spectrometer. Each sample should have at least 3 technical replicates for each LC-MS method to allow for statistic analyses.

10. Briefly spin the tubes for 10 sec. Separate the beads on a magnetic stand, transfer the metabolite extract in the supernatant into a MS vial with glass insert. The extract is ready for LC-MS analysis.

    We use 150 μl glass inserts in 12×32mm brown glass vials with preslit PTFE/silicon septa screw caps for metabolite extract and mass spectrometry.

    LC-MS analysis of the samples should be performed, preferably, within 1-2 days after sample preparation. On some occasions, the samples can be stored at −20 °C or −80 °C for 1 month without losing certain metabolites.

11. Add 30 μl 2× Laemmli SDS sample buffer to the beads, flick the tube to bring down the beads into sample buffer, heat for 15 min on a heat block at above 95 °C.

    Here the protein of interest is stripped off the beads for abundance assessment.

    Random leakage of IgG heavy chain and light chain into the sample buffer is expected if no BS3 crosslinking is used after IgG dynabeads conjugation (see Protocol 2), because β-mercaptoethanol disrupts disulfide

bonds that are formed in between the heavy chains and light chains. Alternatively, non-reducing elution solutions or BS3-crosslinking dynabeads may be used to avoid this leakage problem if desired.

12. Briefly spin the tubes for 5 sec after cooling down the samples on bench. Separate on a magnetic stand, load 15 μl per well onto a 15-well 4-12% NuPAGE gel, and perform electrophoresis 0.5 h at 120 V and 1 h at 150 V.
13. Stain the gel with ProtoBlue Safe Colloidal Coomassie staining solution overnight at room temperature and de-station in water according to manufacturer's instruction.
14. Scan the gel with a minimum resolution of 600 dpi to assess protein yield.

    A dry gel can be prepared for long term preservation: soak the gel in gel drying solution for 15-30 min. Make a dry gel using two layers of cellophane for each gel. Scan the gel to assess protein abundance.

Reagents and solutions—Use double-deionized water in all recipes and protocol steps unless otherwise specified.

Lysis buffer

Use mass spec grade water

200 mM ammonium acetate (stock: 5 M, Sigma 73594)

5 mM EGTA (stock: 500 mM EGTA)

5 mM EDTA (stock: 500 mM EDTA)

1× nDG (1000× stock: 50 mg/ml n-dodecyl glucoside Sigma D8035)

1 mM DTT (stock: 1 M)*

Halt Protease and Phosphotase inhibitor from Pierce (1×)*

* do not use for washing IgG dynabeads.

Wash buffer 1

Use mass spec grade water

500 mM ammonium acetate (stock: 5 M, Sigma 73594)

1× nDG (1000× stock: 50 mg/ml n-dodecyl glucoside Sigma D8035)

Wash buffer 2

Use mass spec grade water

50 mM ammonium acetate (stock: 5 M, Sigma 73594)

1× nDG (1000× stock: 50 mg/ml n-dodecyl glucoside Sigma D8035)

2× Laemmli sample buffer (BIO-RAD 161-0737)

Add 50 μl β-mercaptoethanol to 950 μl 2× Laeemmli sample buffer, vortex to mix.

Gel drying solution

25% ethanol

5% glycerol

Support Protocol 1

Conjugation of Immunoglobulin to Dynabeads

Rabbit IgG is conjugated to epoxy dynabeads, via covalent interaction between the epoxy group on beads and the amine or thiol group from proteins. In some cases the conjugated beads can be further crosslinked to avoid IgG leakage into eluate (Fig. 4). The conjugated beads are later used for affinity purification of TAP-tagged proteins that has the ZZ domain of protein A. The performance of dynabeads is superior to agarose beads because of their high surface area and low background binding.

Materials

Rabbit IgG (ChromaPure rabbit IgG, 10 mg in phosphate buffer, Jackson ImmunoResearch 011-000-003)

Epoxy Dynabeads (Invitrogen, 300 mg size, #143.02D)

Dynabeads Antibody Coupling Kit (Invitrogen, #143.11D) Containing C1, C2, LB, HB, SB solutions.

Protein LoBind tubes (Eppendorf)

n-Dodecyl glucoside (nDG, Sigma #D8035-1g) stock 143 mM, eq. 50 mg/ml in methanol for downstream application in mass spectrometry.

Sodium azide

BS3 (Bis(sulfosuccinimidyl) suberate, Thermo PI-21580, 21585, or 21586), Optional

BS3 Conjugation Buffer 20 mM HEPES or 20 mM Na-PO_4, 0.15M NaCl (pH 7.5)

BS3 Quenching Buffer 1M Tris HCl (pH 7.5)

15 ml Falcon tubes

Mixer

Day 1

1. Spin the IgG solution at 20,000 ×g on a microcentrifuge for 10-15 min at 4°C. Only use the supernatant for the following steps. Determine protein concentration by Bradford assay using bovine gamma globulin (BGG) as standards.

   Note: make sure that the IgG stock solution does not contain any reagents, such as Tris, that has amine or thiol group.

   Centrifugation helps remove IgG aggregate that may leak into protein purification.

   BGG gives better estimation of the IgG concentration than BSA.

2. Resuspend the beads (300 mg) in 3 ml C1 solution. Split and transfer to two 1.5 or 2.0 ml Lobind tubes. Briefly centrifuge, then separate on a magnetic stand. Remove the supernatant.

3. Wash the beads with another 1.5 ml C1 solution in each tube for 2 min, separate on a magnetic stand, briefly centrifuge and remove the supernatant.

4. Resuspend the beads in each tube with 2 ml C1 solution and transfer the beads from each tube into a new 15ml Falcon tube.

5. To each Falcon tube, Add 450 μg rabbit IgG and appropriate volume of C1 (a total of 2 ml of IgG and C1). Put on a mixer at 20 rpm for 1 min.
6. Add 4 ml C2 to each Falcon tube, mix by inverting on a mixer at 20 rpm for 1 min.
7. Wrap the tube with aluminum foil. Incubate on a mixer (20 rpm) at 37 °C for 16-24 hours.

   The crosslink efficiency varies in between 75-85% from our experience.
8. Add 25 μl nDG OR 0.05% Tween-20 to 25 ml SB, mix on a mixer at 4 °C for Day 2.

   nDG does not suppress the ionization of mass analytes on ESI mass spectrometry, but is a weaker detergent to remove non-covalent binding in following steps, and harder to handle due to its poor solubility in water than Tween-20.

Day 2

9 Split the suspension from two Falcon tubes in 6 × 2ml LoBind tubes. Remove the supernatant on a magnetic stand.

   Note: save 1 ml supernatant in each of step 9-14 to calculate conjugation efficiency.

10 Add 1.6 ml HB to each tube. Mix for 1 min or until homogeneous on the Hula mixer, separate on a magnetic stand and remove the supernatant.

11 Add 1.6 ml LB to each tube. Mix for 1 min or until homogeneous on the Hula mixer, separate on a magnetic stand and remove the supernatant.

12 Add 1.6 ml SB with nDG to each tube. Mix for 1 min or until homogeneous on the Hula mixer, separate on a magnetic stand and remove the supernatant.

13 Add 1.6 ml SB with nDG to each tube. Mix 15 min on the Hula mixer, separate on a magnetic stand and remove the supernatant.

14 Repeat step 13.

15 Resuspend the beads in SB to a total volume of 2 ml for each tube of beads. Split to 2 tubes, each with 1 ml. Spin 3 sec to bring down the beads on the tube side. Store at 4 °C until use for up to a month.

   Sodium azide may be added to 0.02% for long-term storage up to a year.

16 Determine conjugation efficiency by measuring IgG concentration with Bradford assay, using BGG as standards. Desalting may be required for other protein assays.

   A typical dilution as following:

   Loading IgG 1000× dilution (>10 mg/ml),

   Flowthrough, HB, LB, and SB wash 5× dilution.

   Standard bovine gamma globulin range uses 1.25-20 μg/ml.

Optional: Crosslinking conjugated beads—This part is useful when IgG leakage happens during protein purification and interferes with other applications. It can be done at a

later time after bead conjugation. Sodium azide must be removed by washing or not used in previous steps.

1. Wash each 1 ml IgG-coupled Dynabeads 3 × 5 min in 1.8 ml Conjugation Buffer each.
2. Dissolve 50 mg BS3 in 18 ml freshly prepared BS3 conjugation buffer to make 5 mM solution.

   Note: prepare fresh BS3 solutions every time. Discard unused solution.
3. Add 1.5 ml 5 mM BS3 solution to the washed Dynabeads in 2 ml LoBind tubes.
4. Incubate at room temperature for 30 min with tilting/rotation.
5. Quench the cross-linking reaction by adding 75 μl Quenching Buffer.
6. Incubate at room temperature for 15 min with tilting/rotation.
7. Wash the cross-linked Dynabeads three times with 1.5 ml SB each. Add equal start volume of SB for storage at 4 °C before use.

   The beads are ready to use, and little IgG leakage into the eluate should be observed afterward.

Support Protocol 2

Yeast Growth and Cell Collection

In our study, we used MORF yeast strains, each hosting a galactose-inducible protein expression construct (Gelperin et al., 2005). Healthy cells at mid-logarithmic growth stage in raffinose medium are subject to galactose induction of target gene expression. Cells are washed and stored at −80 °C until analysis.

Materials

Yeast strain

SC-URA solid medium plates with glucose

SC-URA liquid media with glucose or raffinose

3× YP/gal

Cold sterile distilled, de-ionized water (ddH_2O)

500 ml centrifuge bottles

15 ml Falcon tubes

2 ml thick wall screw-cap tubes

1 L Flasks

Platform shaker capable of 1 L flasks

JA-10 or SLA-3000 rotor

High speed centrifuge

1. Streak yeast cells from glycerol stocks on fresh SC-URA plates, and incubate 2-3 days at 30 °C.
2. Pick fresh single colonies and inoculate into 3 ml liquid medium (SC-ura/glucose) and grow overnight (note: this part is usually done after 3PM). *If the plates have*

ever been stored in cold room, re-streak single colonies on fresh plates and wait for another 1-2 days for yeast to grow up.

3. Check the OD_{600} of yeast culture.

   Use fresh medium as blank, and make a 1×10 dilution of the starter culture.

4. Calculate the volume to be added to 200 ml SC-ura/raff (in 2 L flask):

   V (ml)= 2.75 / OD_{600}

   V is usually below 0.5 ml.

5. Shake the 200 ml culture for roughly 15.5 h at 30°C (critical for mid-log phase control).

6. Measure OD600 using 0.8 ml culture. At this point, absorbance in mid-log phase should be 0.5-0.9.

7. Add 100 ml 3× YP/gal (pre-warmed at 30 °C) to each flask, and shake for another 5.5-7 h.

8. Collect cells in 500 ml centrifuge bottle (JA-10 or SLA-3000 rotor) by centrifugation at 1,500 × g for 5 min

9. Re-suspend cells in 10 ml ice-cold sterile ddH_2O, transfer to a 15 ml Falcon tubes.

10. Pellet cells by centrifugation at 1,500 × g for 5 min.

11. Wash cells again by adding ice-cold sterile ddH_2O to a total volume of 3.8 ml, mix and transfer 1.9 ml each into two 2ml screw-cap tubes.

12. Re-pellet cells on a microcentrifuge at 4 °C. Spin at max speed for 5 min. Remove ddH_2O.

13. Quickly freeze cell pellets on dry ice and store in −80 °C freezer. The cells are good for at least a year.

Reagents and solutions—Use ddH_2O in all recipes and protocol steps unless otherwise specified.

SC-URA media

6.7 g/L Yeast nitrogen base without amino acids (DIFCO)

1.92 g/L yeast synthetic drop-out media (Sigma Y1501)

2% w/v carbon source (glucose or raffinose)

Carbon source stock solution (20% w/v) is sterilized and added after medium autoclaving.

Add 20 g/L agar before autoclaving for solid media.

3× YP/Gal

3% w/v yeast extract

6% w/v peptone

6% w/v galactose

For 1 liter, autoclave 700 ml 3× YP, then add 300 ml filter-sterilized 20% galactose.

Lc-Ms of Protein-Bound Metabolites

Due to the fast development and technical diversification of mass spectrometer models, only a general method is described here. Experimenters are advised to consult a mass spectrometry expert for the best solution available. Because of the explorative nature of this assay, a full scan of 80-1600 amu is sufficient for initial profiling. Quantitative multiple reaction monitoring (MRM) can be introduced at later stage for verification and validation.

Materials

Metabolite samples (preferred if freshly made within 1-2 days)

Reversed phase UPLC columns for low polar to very hydrophobic metabolites

UPLC C18, hexyl/phenyl, or C8 columns from Waters

Normal phase or HILIC columns for polar hydrophilic metabolites

UPLC Amide, T3, or other columns from Waters

Waters Acquity UPLC-coupled Thermo Exactive Orbitrap mass spectrometer, equipped with an electrospray ionization (ESI) probe or an atmospheric pressure chemical ionization (APCI) probe

ESI and APCI complement each other to expand the detection scope. APCI is especially suitable to detect thermostable nonpolar hydrophobic molecules with a molecular weight below 1000 Daltons. APCI is preferred when it can detect the metabolites of interest. ESI has broader detection coverage in chemical properties and molecular weight than APCI, yet it requires optimization of many parameters for sound performance.

Mobile phase solutions

Gradient elution of LC often uses 10% and 90% acetonitrile in water for ESI, and 10% to 100% methanol for APCI. Buffer reagents, such as 10 mM ammonium acetate, can be added to improve LC peak shape stability. pH can be adjusted with acetic acid or ammonium hydroxide as needed to help ionization of analytes.

Strong wash, weak wash, and needle wash solutions

Weak and strong wash solutions are the same as those which are used at start and end mobile phases in composition. For example, 10% and 90% acetonitrile in water can be used as weak and strong wash solutions for a reversed phase gradient. Needle wash is close to weak wash. Do not use buffer or salt reagents in wash solutions.

1. Calibrate the mass spectrometer with standard calibrants and prime the LC system with LC solvents.

   It may take at least 1-2 h before the performance of mass spectrometer is stable.

2. Install the UPLC column, and run at least 3-5 gradients with blank injection of weak wash solvent.

   This step stabilizes column performance.

3. Set up the sample running batch. At least 3 replicate runs should be performed for each sample.

Uninterrupted running is strongly advised for batch processing of all samples in one study.

The running time for each injection depends on the flow rate and column size.

A typical gradient usually includes a wash of 10× column bed volumes, plus 5-10 column bed volumes' weak wash to stabilize the column.

4. After acquisition, wash and store the column with solvents without buffer or salt reagents.

   Inclusion of organic solvent at 10% or higher helps prevent microbe growth.

LC-MS Data Processing Using XCMS

In addition to proprietary softwares from mass spectrometer vendors, many free softwares are available in academia with decent power. We routinely use XCMS for data processing. XCMS is a free R package that has been developed to streamline LC-MS-based metabolomics analyses (Smith et al., 2006). Raw LC-MS data are converted to centroid CDF or mzXML format before XCMS analysis. Technical replicates are placed into one folder for each sample. The following script is a script running in R, a statistic software freely available. It works under R 1.13.0 with XCMS build 1.26.0. For more information please check the XCMS page (http://metlin.scripps.edu/xcms/).

install XCMS

update.packages()

source("http://bioconductor.org/biocLite.R")

biocLite("xcms")

rm(list=ls(all=TRUE))

load xcms, set up sample list and file path

library(xcms)

change the following file path to the folder where you have all those CDF files saved.

mzXMLpath= "X:/MassSpecData"

setwd="X:/MassSpecData"

list.files("mzXML",path=mzXMLpath, recursive=TRUE, full.names=TRUE)

mzXMLfiles<-list.files("mzXML",path=mzXMLpath, recursive=TRUE, full.names=TRUE)

extract mass features in positive mode from the raw data

type "?findPeaks.centWave" in R for parameter settings for centWave

xset<-xcmsSet(mzXMLfiles, method='centWave', ppm=2.5, peakwidth=c(15,90), snthresh=3, prefilter=c(3,260), integrate=2, mzdiff=0.00005, noise=260, verbose.columns=TRUE,fitgauss=TRUE,mzCenterFun="wMean",polarity="positive", nSlaves=4)

xset<-group(xset, bw=10, minfrac=0.6)

xset

Align the retention time of mass features detected among samples

#xset2<-retcor(xset, method="obiwarp", plottype="deviation")

xset2

Fill blanks in samples without a particular mass feature

xset2<-group(xset2, bw=10)

xset3<-fillPeaks(xset2)

xset3

Write out the peak table for all sample

pt <- peakTable(xset3)

write.table(pt,file="peaktable.txt",sep="bsolt", row.names=F)

Statistic comparison of two samples, and write out the table reporting P value and fold enrichment

reporttab1<-diffreport(xset3,"name_of_control", "name_of_sample", "name_of_sample.txt")

The accurate mass reported in the "mz" column for each mass feature can be used to search for possible hits on metabolite databases such as METLIN (http://metlin.scripps.edu) and HMDB (www.hmdb.ca).

Commentary

Background Information

Metabolite-protein interaction represents another layer of complexity in biological interaction networks. Many have been reported adventitiously in the past, including some classical examples such as cofactors in enzymology, nutrient-sensing operons and sterol-sensing mechanism in regulation of transcription, second messengers in signaling, and neurotransmitter in synapses. A systematic approach was not possible until recent advances in genomics and proteomics that can predict and produce, virtually, all proteins in one organism. Several studies have since disclosed extensive and degenerate metabolite-protein interactions *in vivo* and *in vitro* (Gallego et al., 2010; Li et al., 2010; Morozov et al., 2003). These degenerate patterns may bridge gaps between other types of interaction networks and provide novel insights into how cells function as a system (Li and Snyder, 2011).

Compared to proteins, the small sizes of metabolites, in most, if not all, cases, preclude their use as baits, as immobilized metabolites would likely sterically hinder many metabolite-protein interactions. Proteins are thus used as baits in aforementioned studies. An assay to assess metabolite-protein interaction, which requires the incubation of purified protein and metabolite mixture *in vitro*, has previously been described, using protein affinity purification (Tagore et al., 2008) or protein microarray (Gallego et al., 2010; Morozov et al., 2003; Zhu et al., 2001). These studies have high throughput only when combined with protein microarray and pre-selected metabolite targets. However, it might suffer from undesired artifacts, since it can be very challenging to mimic the exact *in vivo* binding environment and physiological ranges of metabolite composition. Furthermore, unexpected interactions cannot be detected because the detection scope is much limited by the metabolites that are used. Here we couple protein purification directly with downstream mass spectrometry for fast and scalable experimental practice and unbiased metabolite profiling. We have had success in using this method to profile both hydrophobic and hydrophilic molecules.

Critical Parameters and Troubleshooting

Sufficient protein input is necessary to detect stoichiometrically protein-interacting metabolites. Because picomoles of metabolites can be comfortably detected by modern mass spectrometry, the biomaterials for each experiment should be sufficient to purify the protein of interest at within one order of 1 picomole (around 100 ng for a 100kD protein).

To minimize variation in the performance of LC-MS, it is advised that the experimenter finish the whole batch of samples with the same LC-MS method at once. Frequent switching LC-MS setups may result in unwanted migration in retention time and altered MS sensitivity.

It is important not to inadvertently introduce small molecule contamination during experiments. Handle samples with nitrile gloves and in ultra-clean tubes. Always use the highest grade reagents available. Avoid unnecessary use of polyethylglycol polymer-based detergents and membrane filters.

Analyzing LC-MS data can be daunting at present despite many software tools are available for metabolomic profiling. While different peak calling algorithms, peak alignment algorithms and statistic filters may apply, one should keep in mind that, in this assay, an all-or-none or infinite enrichment is much more desired when compared with genuine metabolome-profiling studies. The parameters in data processing have to be adjusted according to real data for best performance.

Anticipated Results

Protein-interacting metabolites can be scored by the presence and enrichment of a reproducible LC-MS feature with retention time and accurate mass measurement within several parts-per-million (ppm) of the theoretic values. The monoisotopic mass peak pattern should also match the theoretical values within a range of 20%. These true positive mass features are often not obvious to bare eyes because their abundance is much lower than most matrix mass features arising from solution ingredients.

Follow-up experiments

Several strategies can be used for validation after general profiling of protein-associating metabolites. The first two can be systematically carried out without prior knowledge, while the third requires expertise in related fields.

First, metabolite standard, if commercially available, can be analyzed under identical LC-MS conditions for matching retention time and monoisotopic and adduct patterns with the analytes of interest. This is the first and straightforward step one can take to verify the identities of protein-binding metabolites. To minimize the matrix effect on retention time (such as salt, solvent, etc.), it would be better to prepare standard in similar solutions that are used for real experiment. Additionally, fragmentation patterns of the target metabolites can be acquired by LC-MS operating in multiple reaction monitoring (MRM) in repeated biological experiment, and used to search metabolite databases for matching fragment distribution.

Second, *in vitro* binding assays between purified protein and metabolite standards can be used to characterize the binding affinity and stoichiometry. A saturable binding curve is expected for specific binding while linear binding often indicates the binding is non-specific.

An affinity constant within the physiological range of the binding metabolite and low metabolite-to-protein binding ratio often suggest that the interaction is consequential. A desalting column that removes small molecules (< 7,000 Da) can be used to remove non-

binding metabolites in this assay, as described before (Li et al., 2010; Li and Snyder, 2011). Cautions shall be taken during protein purification in order to strip off endogenously bound metabolites by extensive washing to make the binding sites available for *this in vitro* binding assay. This assay may not work if the metabolite-protein binding requires energy molecules (e.g., ATP) or other chaperones.

Third, the biological significance of metabolite-protein binding can be explored by determining changes in the function of target proteins after genetic or pharmacological alteration of target metabolite levels in cells. For example, target metabolite synthesis can be eliminated by knocking out the enzyme-encoding gene(s) or reduced by introducing specific drug inhibitors of the enzymes. Conversely, the levels of the target metabolite can be compared between cells expressing its interacting protein at normal, excessive, reduced (or none) levels. Furthermore, it is also possible to confirm the predicted metabolite-protein interactions, by introducing a chemically modified metabolite, or its precursor into the biological system (e.g., cell culture) for subsequent interaction analysis. The predicted binding event, as supported by this type of methods, could rather be a consequence of indirect association as well. Even in such case, the results could still be biological meaningful, as long as manipulation of either proteins or metabolites can be shown to have significant effects on the biology of interest.

Time Considerations

Cell collection (protocol 3) can take weeks to months for large batch of samples, which is the most time-consuming step. We have found that the cells can be stored at −80 °C up to 2 years without noticeable difference in performance. However, experimenters are encouraged to use fresh samples that are collected within 1-3 months.

The actual number of samples processed per day depends on the throughput of LC-MS. We typically run three technical replicates per sample. Each run takes 20 minutes, meaning 24 samples is the maximum one can process each day.

The protocol 1 takes ~8 hours from cell lysis to staining of the NuPAGE gels, should 24 samples be processed in one batch. The bead conjugation takes a maximum of two days. 300 mg dynabeads are enough for several hundred samples.

LC-MS data analyses can take several days to weeks of computing time for large sample size, and the manual checking raw data can be even longer.

Acknowledgments

The authors thank Dr. Peichuan Zhang for critical review of this manuscript. This research was supported by a grant from the NIH.

Literature Cited

Gallego O, Betts MJ, Gvozdenovic-Jeremic J, Maeda K, Matetzki C, Aguilar-Gurrieri C, Beltran-Alvarez P, Bonn S, Fernandez-Tornero C, Jensen LJ, et al. A systematic screen for protein-lipid interactions in Saccharomyces cerevisiae. Mol Syst Biol. 2010; 6:430. [PubMed: 21119626]

Gelperin DM, White MA, Wilkinson ML, Kon Y, Kung LA, Wise KJ, Lopez-Hoyo N, Jiang L, Piccirillo S, Yu H, et al. Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. Genes Dev. 2005; 19:2816–2826. [PubMed: 16322557]

Li X, Gianoulis TA, Yip KY, Gerstein M, Snyder M. Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. Cell. 2010; 143:639–650. [PubMed: 21035178]

Li X, Snyder M. Metabolites as global regulators: a new view of protein regulation: systematic investigation of metabolite-protein interactions may help bridge the gap between genome-wide

association studies and small molecule screening studies. Bioessays. 2011; 33:485–489. [PubMed: 21495048]

Lowenadler B, Jansson B, Paleus S, Holmgren E, Nilsson B, Moks T, Palm G, Josephson S, Philipson L, Uhlen M. A gene fusion system for generating antibodies against short peptides. Gene. 1987; 58:87–97. [PubMed: 3692175]

Morozov VN, Morozova TY, Johnson KL, Naylor S. Parallel determination of multiple protein metabolite interactions using cell extract, protein microarrays and mass spectrometric detection. Rapid Commun Mass Sp. 2003; 17:2430–2438.

Nilsson B, Moks T, Jansson B, Abrahmsen L, Elmblad A, Holmgren E, Henrichson C, Jones TA, Uhlen M. A synthetic IgG-binding domain based on staphylococcal protein A. Protein Eng. 1987; 1:107–113. [PubMed: 3507693]

Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem. 2006; 78:779–787. [PubMed: 16448051]

Tagore R, Thomas HR, Homan EA, Munawar A, Saghatelian A. A global metabolite profiling approach to identify protein-metabolite interactions. J Am Chem Soc. 2008; 130:14111–14113. [PubMed: 18831549]

Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, et al. Global analysis of protein activities using proteome chips. Science. 2001; 293:2101–2105. [PubMed: 11474067]

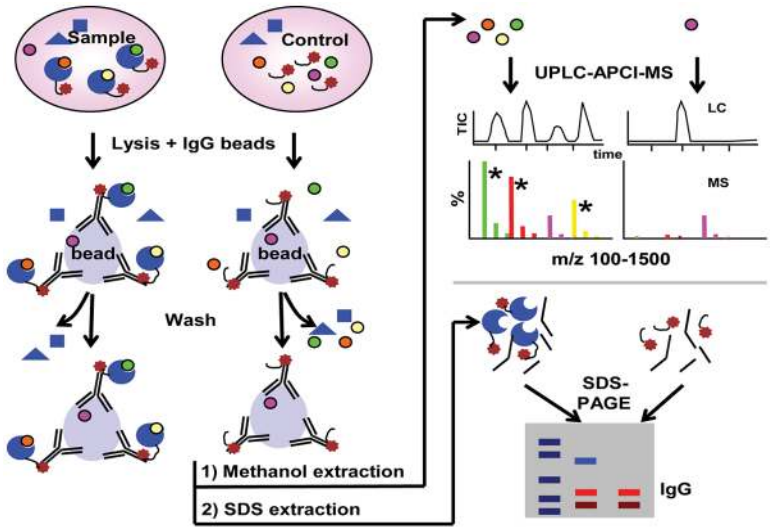

Figure 1. Flowchart for the identification of small metabolites bound to proteins

Yeast proteins tagged with an IgG-binding protein domain are isolated from lysates using magnetic beads coated with IgG. After washing, the small metabolites are then extracted in organic solvent and analyzed using liquid chromatography-coupled mass spectrometry (LC-MS). The purified proteins adsorbed to magnetic beads are later extracted with SDS sample buffer and analyzed using SDS-PAGE. A yeast strain lacking the fusion protein is used as the negative control in parallel experiments. The metabolites significantly enriched in the fusion protein sample relative to the negative control are scored as protein-bound metabolites.

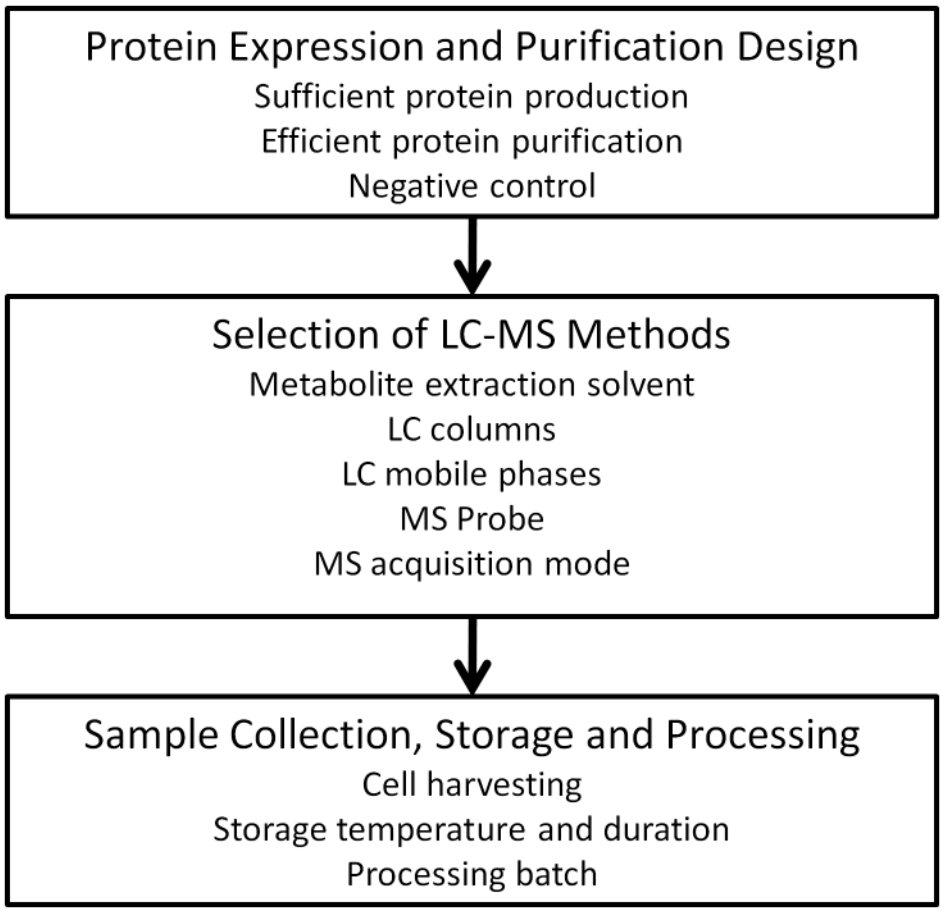

Figure 2. Strategic consideration to study metabolite-protein interactions as described in this protocol

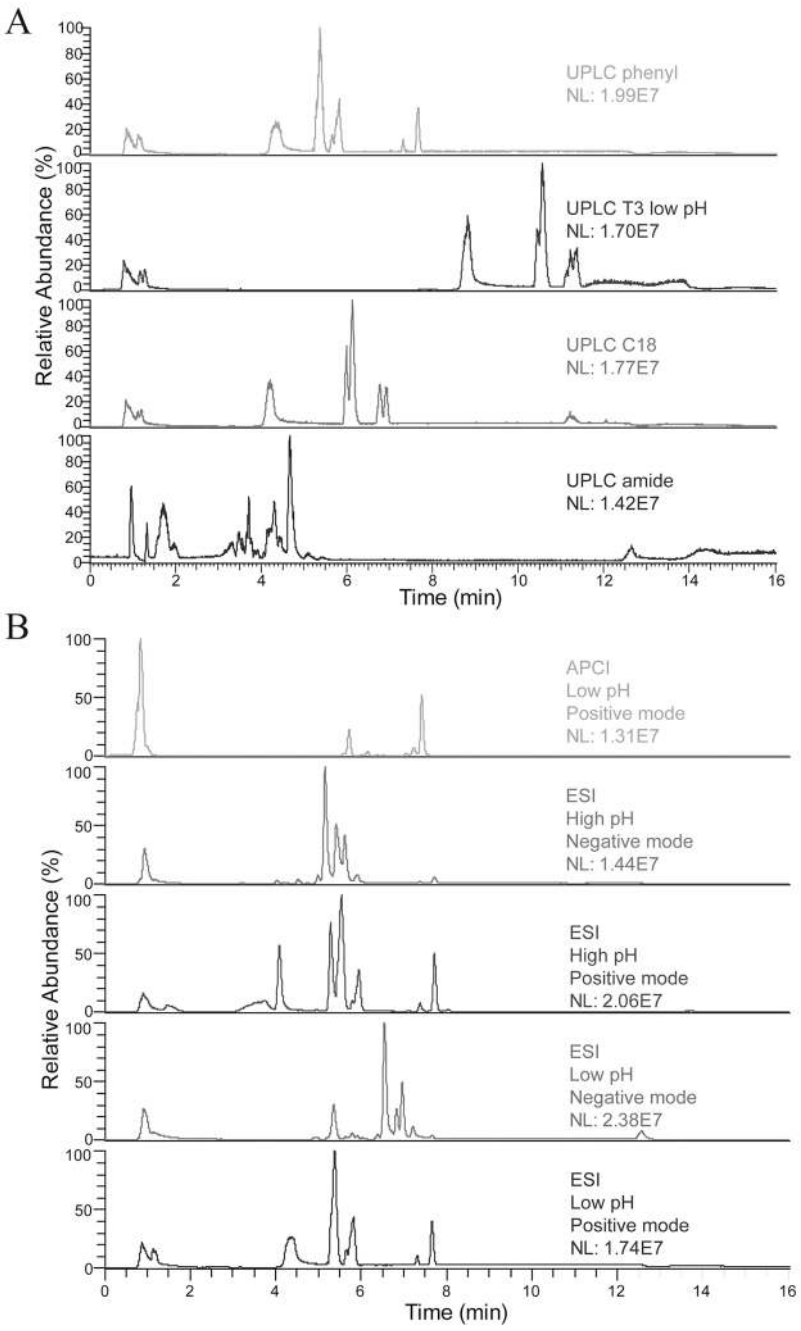

Figure 3. The influence of different LC-MS methods on mass spectral patterns

The same yeast metabolite extract is used for all conditions at a scan range of 85-1200 m/z. NL indicates the maximum intensity. All plots are base peak intensity (BPI) mass spectra.

A, The Influence of UPLC columns on mass spectra. All mass spectra are acquired in positive mode with mobile phases at pH 4.25.

B, The influence of MS probes, the pH of mobile phases, and polarity on mass spectra. APCI or ESI probes, high pH (9.75) or low pH (4.25), and polarity are indicated.

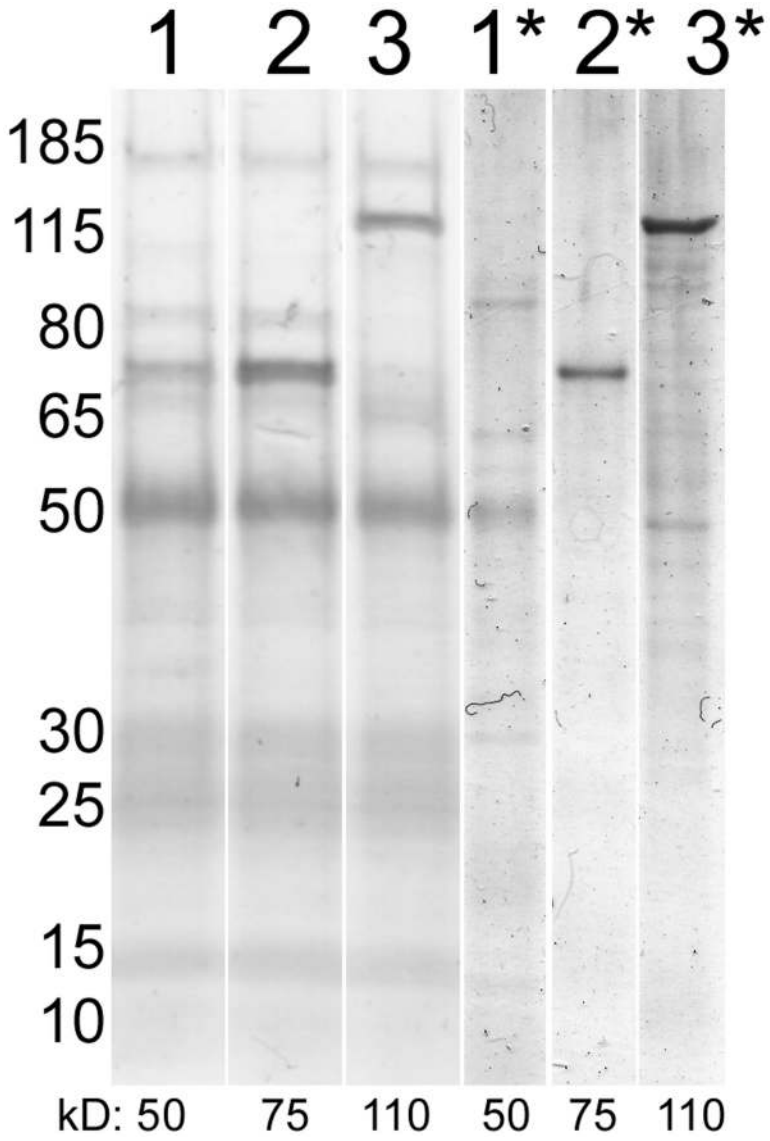

Figure 4. Protein of different sizes can be purified using rabbit IgG conjugated dynabeads

The expected size of each protein is indicated at the bottom of each lane. Molecular Markers are on the left. The gels were stained with Colloidal Coomassie Blue G250. Lanes 1-3, proteins purified on regular IgG dynabeads. Lanes 1*-3*, the same proteins purified on BS3-crosslinked IgG dynabeads, note the IgG leakage at 55 kD and 25 kD are absent in these lanes. Equal amount of yeast cells were used for each protein purification.