SECTION I

BIOTECHNOLOGY/PROTEINS

1

TARGETED PROTEOMICS USING IMMUNOAFFINITY PURIFICATION

Karen R. Jonscher, Lei Jin, John C. Cambier, Shaikh M. Rahman, and Jacob E. Friedman

1.1 INTRODUCTION

Proteins are multimodular and multifunctional, interacting in complex networks that drive cellular function. Pathological alterations in signaling networks are thought to result in a number of diseases, particularly cancer. Understanding the roles and consequences of protein–protein interactions is therefore a fundamental goal in systems biology. The two-hybrid approach [1] emerged in the early 1990s as the first method to assay whether two proteins interact in a pair-wise fashion. A number of bait–target strategies were subsequently developed [2], including techniques exploiting affinity purifications coupled to mass spectrometry (MS) to rapidly identify potentially novel protein interactions. Initial studies were performed in yeast [3,4] and were subsequently expanded to mammalian models [5]. Since MS-based proteomics is not necessarily limited to specific sites or to specific proteins, it represents an unbiased and direct approach to studying cellular processes [6].

As recently described [6,7], immunoaffinity purification has emerged as the most frequently employed method for multiprotein complex purification. Its success is based on the principle that multiple members of a complex may be captured when one complex member is enriched, regardless of whether the complexed proteins are directly bound to the target protein. Additionally, purification of posttranslational modifications has been used extensively to globally profile modified proteins throughout cellular networks [8,9] and provides invaluable insights into signal transduction mechanisms.

A summary of typical steps employed to generate samples using an immunoaffinity-based approach is illustrated in Figure 1.1 and described in detail for two example applications below. Following purification, peptide mixtures resulting from the digestion of bands or eluates are analyzed using tandem mass spectrometry (MS/MS) and proteins are identified by database searching and spectral matching.

A gel-based approach may be useful when two conditions are being compared—bands exhibiting visual differences can be excised to yield data most likely to contrast biologically significant results (note that interesting low abundance proteins may be covered up by more abundant nonspecific proteins). Another useful method, gel-enhanced liquid chromatography–tandem mass spectrometry (GeLC-MS/MS), has also emerged for the analysis of complex protein mixtures [10] and can be applied to the separation of immunoaffinity eluates. In this approach, a protein-containing gel lane is chopped into equivalent sections, digested, and peptide mixtures analyzed. When complexed protein levels are extremely low or sample is limited, elution followed by in-solution digestion may provide a better option, as less protein is lost to sample handling. An important caveat to note is that a protein of interest may be "covered up" by comigrating background or nonspecific proteins. This is a particular concern for proteins that may comigrate with immunoglobulin (Ig)

Mass Spectrometry Handbook, First Edition. Edited by Mike S. Lee.

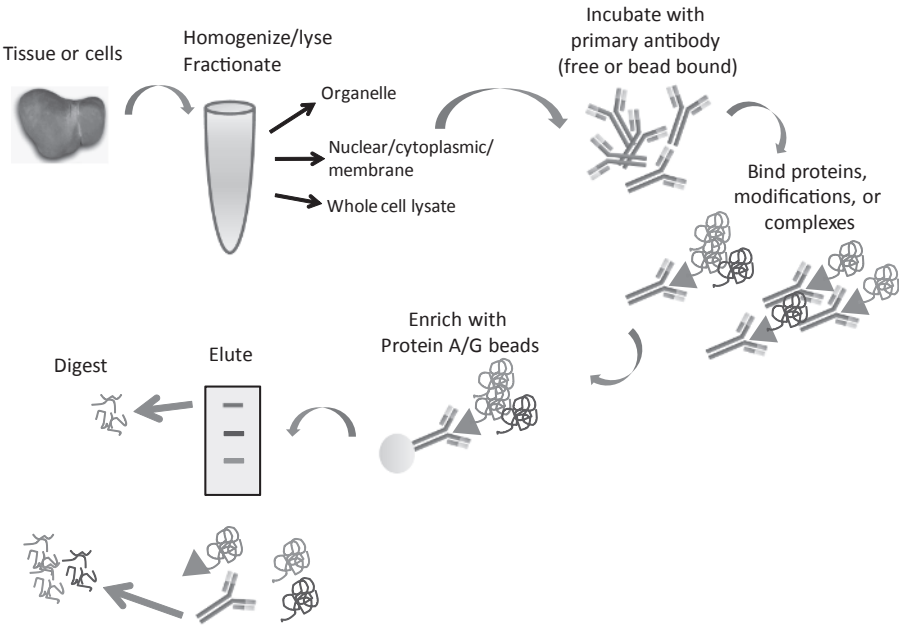

FIGURE 1.1 Flowchart of sample preparation steps for immunoaffinity purification workflows. Tissue is homogenized or cells are lysed in an appropriate buffer using protease inhibitors. Whole cell lysates may be used, or subcellular enrichment of organelles, membrane, or cytoplasmic fractions, for example, may be employed for subsequent processing. Primary antibody recognizing the protein or modification of interest is either bound to protein A- or protein G-coated beads or added to the lysate, allowing either target proteins or modifications to bind. Bead-bound protein A or protein G interacts with the primary antibody, enriching the protein complexes through a series of washes. Finally, proteins are eluted from the beads. The protein mixture, also containing the antibody, is either directly digested with an enzyme or separated by 1D gel electrophoresis, with the protein-containing bands of interest subsequently excised and digested.

heavy and light chains. Using a cross-linker such as DMP (dimethyl pimelimidate) to bind the primary antibody to Protein A will suppress elution of Ig chains when nonreducing conditions are used for elution. However, cross-linking may result in loss of affinity; an optimization workflow should ideally include both cross-linked and noncross-linked trials.

The primary challenge of immunoaffinity-based workflows lies in the difficulty of separating true low abundance interactors from nonspecifically binding proteins. Use of negative controls, such as preimmune sera or antibodies against other proteins, or, if the model allows, using a knockout or knock down of the protein of interest, can help separate out these background proteins. As described below, cross-linking the antibody, minimizing incubation times and antibody concentrations, optimizing wash buffer stringency, and other approaches may help mitigate the extent of nonspecific binding. Assessing the utility of at least a few of these parameters should be included in the optimization workflow. A good way to begin optimizing the protocol is to immunoprecipitate the protein of interest and probe for a known interactor using Western blot. Begin by titrating the primary antibody and beads to find the minimum amount required to effectively immunodeplete the sample. Then experiment with incubation times. Fewer beads and shorter incubation times will help reduce nonspecific binding. Ultimately, orthogonal techniques such as co-immunoprecipitation (IP) with Western blot should be used to validate a subset of the potential interactors, whenever possible.

In this work we present examples of workflows in which immunoaffinity-purified proteins were either separated using gel electrophoresis and bands exhibiting significant change from control were analyzed, or complexed proteins were eluted from the beads, digested in solution, and analyzed. It is important to note that these protocols provide *general guidelines* and that several optimization steps with multiple iterations of MS will likely be required for purification of a protein complex of interest.

As discussed in more detail below, data analysis and mining are critical for gleaning relevant information from proteomics studies. Careful extraction of peptide signals, determination of properly stringent search engine parameters [11], and reversed or scrambled

database searching leads to an output data set where significance of the identifications may be established with score or probability cutoffs. Although a false discovery rate of ~1% is often employed in more global approaches [12], establishing criteria for two peptide "hits" to a protein with peptide probabilities of ~95% is sufficient to provide a false discovery rate approaching 0% for immunoaffinity purification applications. Following identification, data mining is employed to obtain functional information about the proteins to begin to decipher mechanisms that may be triggered by the interaction. The tools used often include those developed for microarray analysis, where gene ontology information is used to cluster proteins with similar cellular compartments, functions, or pathways [13–16]. Downstream assays based on these results, including IP of proteins identified in the complex, allow investigators to begin to elucidate mechanisms driving cellular function.

1.2 EXPERIMENTAL PROTOCOLS

Materials and Solutions

- *Cell Lysis Buffer.* 0.33% 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS); 150 mM NaCl; 10 mM sodium pyrophosphate; 10 mM Tris-HCl pH 7.4; 1 mM phenylmethylsulfonyl fluoride (PMSF); 0.4 mM ethylenediaminetetraacetic acid **(**EDTA); 1.8 mg/mL iodoacetamide (IAA); 10 mM NaF; 2 mM Na_3VO_4; and 1 μg/mL each of aprotinin, leupeptin, and pepstatin.
- *Tissue Lysis Buffer.* Buffer A (10 mM HEPES pH 7.9, 1.5 mM KCl, 10 mM $MgCl_2$, 0.5 mM dithiothreitol (DTT), 0.1% IGEPAL CA-630 (Sigma-Aldrich, St. Louis, MO), and 0.5 mM PMSF) and Buffer B (20 mM HEPES pH 7.9, 25% glycerol, 1.5 mM $MgCl_2$, 420 mM NaCl, 0.5 mM DTT, 0.2 mM EDTA, 0.5 mM PMSF, and 4 μM leupeptin). (*Note*: These buffers were selected for the analysis of liver proteomes, as described below. Other lysis buffers may be more appropriate for the sample/proteins under investigation. A search of recent literature should provide some direction regarding appropriate buffer selection.)
- Reagents for preparation of magnetic beads as described below.
- Protein A or protein G beads. (*Note*: Either Sepharose [GE Healthcare, Piscataway, NJ] beads or magnetic beads may be used; however, magnetic beads are preferred as they tend to exhibit less nonspecific binding than Sepharose beads.)
- Pipette tips cut 5 mm from the top (will avoid damaging the beads if using Protein A/G Sepharose beads).
- Laemmli buffer.
- High performance liquid chromatography (HPLC)-grade water (Honeywell Burdick and Jackson [Morristown, NJ] or other high quality liquid chromatography–mass spectrometry [LC/MS]-grade water).
- Desalt spin columns.
- Reagents for in-gel or in-solution digestion as listed below.

Equipment

- Refrigerated centrifuge
- Rotary mixer
- Vacuum centrifuge
- Gel electrophoresis apparatus and 10% polyacrylamide gel
- Nanoscale HPLC, tandem mass spectrometer

Lysis

Note that the cell numbers and tissue amounts presented here are a guideline. These numbers should be increased if complexed proteins are of low abundance.

Cell Lysis

1. Lyse ~5×10^8 cells using 500–1000 μL cell lysis buffer at 4°C overnight. Note that if low abundance or weakly interacting proteins are to be analyzed, increase the cell numbers to as much as 10^{10} (as shown by Malovannaya and coworkers [17]).
2. In the morning, centrifuge lysates at 12,000 × *g* at 4°C for 20 min. Remove supernatants to a clean tube.

Tissue Lysis

1. Homogenize ~100 mg of tissue using a mortar and pestle over liquid nitrogen. For identification of potentially weakly binding complexes, increase tissue amount to 10–20 g (following Moresco et al. [18]) and increase lysis buffer volume to 5–10 mL.
2. Add homogenized tissue to 0.5 mL of tissue lysis buffer A (ice cold) and ultrasonicate three times, 15 s each. Place sample tubes in an ice bath for at least 1 min between sonications.
3. Incubate samples on ice for 30 min, then centrifuge at 14,000 × *g*, 4°C for 10 min. Remove the supernatant (cytoplasmic fraction) to clean tube.
4. Resuspend the membrane/organelle fraction pellet in 0.2 mL ice-cold tissue lysis buffer B and

incubate on ice for 30 min. Following centrifugation at 14,000 × *g* for 30 min at 4°C, remove the supernatant to clean tube.

Note: For all subsequent steps, be sure to keep samples on ice or at 4°C, however freezing lysates before the immunoaffinity purification should be avoided [18]. If necessary, store lysates at –80°C prior to use.

Total Protein Quantification

Use either the Bradford or bicinchoninic acid (BCA) method to quantify total protein concentrations following the manufacturer's instructions for a microwell plate assay. Make sure that the lysis buffer components are compatible with the manufacturer's stated levels. Try several dilutions to ensure the sample concentration is within the linear range of the assay.

Immunoaffinity Purification

Immunoaffinity purification may be accomplished using either soluble antibodies or antibodies cross-linked to beads. Generally the first step of the optimization should be done using cross-linked antibodies; cross-linking significantly reduces contaminating signals from Ig light and heavy chains. Procedures for both approaches are provided below.

Immunoaffinity Purification Using Magnetic Beads

Reagents for Magnetic Bead Preparation (Dynabeads Are Typically Used)

- *Citrate Phosphate Buffer, pH 5.0.* 25 mM citric acid, 50 mM sodium phosphate (Na_2HPO_4)
- *0.2 M Triethanolamine (TEA), pH 8.2.* 3.71 g triethanolamine-HCl/100 mL water
- *20 mM DMP.* 5.4 mg DMP-2HCl per milliliter of TEA buffer
- 50 mM Tris pH 7.5
- *PBS-T.* 0.01% Tween-20 (Thermo Fisher Scientific, Waltham, MA) in phosphate buffered saline
- 0.1 M glycine pH 2.5–2.7
- *Storage Solution.* PBS-T with 0.02% sodium azide

Dynabeads (Invitrogen Corp., Carlsbad, CA) are packaged as a 5% slurry. Prepare 0.5–1.0 mL of slurry to obtain 25–50 µL of packed beads. A rule of thumb is that 1 mL of slurry binds ~300 µg of antibody. Incubate with about 400 µg of the primary antibody.

Equilibrate Dynabeads

Centrifuge beads briefly, place tube in a magnetic rack, and remove the supernatant. Add 1 mL citrate phosphate buffer, vortex, spin briefly in a minifuge (1 s to remove bead solution from cap), and place tubes in a magnetic rack. Remove supernatant. Repeat two more times.

Incubate with Primary Antibody

Prepare 400 µg of primary antibody in 1 mL citrate phosphate buffer and add to beads. Reducing the volume may improve binding and may be included in subsequent optimization steps. Rotate tube end over end for 2–3 h at room temperature.

Wash

Centrifuge briefly, place tubes in a magnetic rack, remove supernatant, and add 1 mL citrate phosphate buffer and wash three times as described in the section "Equilibrate Dynabeads." Wash two times more with 1 mL 0.2 M triethanolamine-HCl.

Cross-Link

Remove final TEA wash from the beads using the magnet and add 1 mL of DMP solution. Incubate 30 min at room temperature, rotating end over end.

Clean Up

Using the magnet, remove the DMP solution and incubate beads with 50 mM Tris for 15 min to remove free cross-linking reagent.

Wash

Wash beads three times with PBS-T.

Remove Free Antibody

Incubate the beads with 0.1 M glycine for 5 min, rotating end over end at 4°C.

Wash

Wash beads three times with PBS-T.

Store Beads

Bring beads back to original packaged volume in storage solution for storage at 4°C. For a 1 mL stock, use 950 µL of storage solution.

To use, wash beads three times using 1 mL PBS, then three time with 1 mL lysis buffer. Use Western blots and bead titration to determine the minimum amount of beads to use and the minimum amount of time to incubate. A starting point for optimization may be 20 µL of packed beads and 2 h of incubation at 4°C. Elute as described below.

Immunoaffinity Purification Using Protein A/G Sepharose Beads

Incubation with Primary Antibody

Add 3–20 µg of primary antibody to the supernatant and incubate for 2 h to overnight at 4°C. This step

should be optimized to use the minimal time and generate the least amount of nonspecific binding. Increasing the incubation time generally increases background and fragmentation due to endogenous protease activity. Increased antibody concentrations are required when using more cells or when the total protein concentration is high.

Note: In order to reduce nonspecific binding of protein aggregates during incubation, ultracentrifuge samples at 100,000 × *g* for 20 min following incubation and limit incubation time to 2 h. Remove the supernatant to a clean tube. Some groups recommend that the lower 0.1 mL of lysate not be used [17]. Using magnetic beads instead of Sepharose beads may also help.

Prepare Protein A/G Sepharose Beads

Remove 100 µL of bead slurry for each sample (mix the slurry well prior to removing the beads). Centrifuge at 1500 × *g* at 4°C for 1–2 min, aspirate the supernatant, and wash with 10× bead volume of cell lysis buffer three times.

Bind to Sepharose Beads

Remove sample to the tube containing the washed beads and mix end over end at 4°C for 4 h. Centrifuge at 1500 × *g* at 4°C for 10 min and remove supernatant to clean tube.

Wash

Wash beads up to five times with 10× bead volume of cell or tissue lysis buffer by adding buffer, centrifuging at 1500 × *g* at 4°C for 1 min, and removing supernatant. (*Note*: To retain weakly interacting proteins, use a low stringency wash buffer, such as PBS or 0.5% NP-40, add 10× bead volume, and briefly invert the tube 10 times [17]. Radioimmunoprecipitation assay [RIPA] buffer may also be used to retain more strongly interacting proteins.)

Elution

Proteins may be eluted from the beads either for subsequent gel electrophoresis and in-gel digestion or for in-solution digestion.

Gels

For applications involving gel electrophoresis, add one bead volume of 2× Laemmli buffer to the bead pellet, heat at 60°C for 10–15 min, centrifuge, and remove the supernatant to clean tube. To ensure complete elution, add an additional 10 µL of Laemmli buffer to the bead pellet, repeat, and pool supernatants.

Solution

For in-solution digestion applications, elute the proteins from the antibody-coupled Sepharose beads using one bead volume of a strong acid such as 0.2 M citric acid (pH 2.0), 1% formic acid or 0.5% trifluoroacetic acid. If using citric acid, immediately add an equivalent volume of 2 M Tris to neutralize the pH prior to digestion. Alternatively, a strong and volatile base such as ammonium hydroxide can be used, or even 8 M urea. The urea will not effectively elute the protein target but will disrupt interacting proteins. Perform three similar elutions and combine the eluates. Concentrate to dryness in a vacuum centrifuge. If 8 M urea is used for elution, proceed to the protocol for in-solution digestion.

Note: The elution should be optimized by checking at least two to three different elution solvents. Eluates may be assessed using a one-dimensional (1D) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel stained with colloidal Coomassie.

Gel Electrophoresis (for Subsequent in Gel Digestion)

Separate Proteins

Using samples generated in the section "Gels," load the proteins onto a 1D 10%–14.5% SDS-PAGE gel and separate them using 120–180 V. A 13.3 × 8.7 cm × 1 mm thick gradient gel generally provides sufficient separation. If more separation is required, a larger gel can be used. A 15% gel should be used for resolving smaller proteins, while a 7%–10% gel is appropriate for investigating larger proteins. All of these types of gels may be useful for the initial stage of optimization for separating the proteins of interest. A lower separating voltage may provide more resolution.

Stain Proteins

Following electrophoresis, carefully wash the gels twice with ultrapure water. Be careful to handle the gels only by the edges. Some commercially available ready-made gels have a thick bottom edge which helps prevent the gels from ripping when they are rinsed. Use either colloidal Coomassie blue (such as EZ Blue, Sigma-Aldrich) or a fluorescent stain and follow the manufacturer's directions for staining and fixing. Oriole™ fluorescent gel stain (BioRad Corporation, Hercules, CA) is an easy-to-use, sensitive, and highly linear protein gel stain. This stain can be used without prior protein fixation or destaining. For larger protein loads (e.g., 10 mg total protein), use a standard Coomassie blue stain and follow destaining procedure below.

Visualize Proteins

Image the gel using a laser imager capable of exciting and measuring emission at wavelengths appropriate for the stain employed. For example, Oriole has an excitation maximum at 270 nm and emission at 604 nm, thus

it is appropriate for UV-based imagers. Coomassie can be imaged using a standard visible light gel scanner.

Excise Protein Bands

Excise gel bands of interest using a scalpel or a cut pipette tip. Dice the gel slice into 1 mm^3 cubes. Place into a 1.5 mL microfuge tube and remove liquid. Gels may be stored at –80°C prior to processing.

Sample Preparation for MS

In-Gel Digestion

Reagents

It must be noted, use only high quality HPLC-grade water.

- *50 mM Ammonium Bicarbonate (NH_4HCO_3).* 4 mg/mL in water.
- *Destain Reagent.* 1:1 acetonitrile (ACN) in 50 mM NH_4HCO_3.
- *Reducing Reagent.* Prepare a 1.5 mg/mL solution of DTT in 50 mM NH_4HCO_3. DTT is unstable in solution at room temperature so prepare fresh daily.
- *Alkylating Reagent.* 10 mg/mL solution of IAA in 50 mM NH_4HCO_3. This buffer should also be prepared fresh daily (while reducing) and kept in the dark.
- *Trypsin.* Use sequencing grade modified trypsin. Promega (Madison, WI) is widely used.

Destain

(This step will remove stain that might interfere with proteolytic digestion. Note that this step is not required if using Oriole stain. Follow manufacturer's directions for the stain utilized.)

1. Remove any liquid from the microfuge tube. If the gel band was not previously diced, use a sealed pipette tip to cut gel slice into small pieces.
2. Add enough destain reagent to cover gel pieces. Shake or sonicate at room temperature for 10 min. Discard liquid. If gels were stained using a fluorescent stain, go to the section "Reduction/Alkylation (for 1D Gel Bands)."
3. Repeat the second step until gel pieces are clear. Usually two to three washes are necessary for Coomassie stained gels. If gel pieces still have blue color, rehydrate by adding 50 mM NH_4HCO_3 and shake for 10 min at room temperature. Discard liquid and repeat.

Reduction/Alkylation (for 1D Gel Bands)

(This step allows for denaturing and separation of complexed proteins)

1. If not already dehydrated, dry gel pieces in vacuum centrifuge. Transfer the dried slices to clean 0.6 mL tubes.
2. Add enough of the reducing reagent to fully cover gel pieces, taking into account that the pieces will swell. Usually 20–50 µL is sufficient.
3. Incubate at 37°C for 1 h. During this time, prepare the alkylating reagent.
4. After incubating, remove excess reducing reagent and cool to room temperature (this can be done by placing samples at –20°C for 5 min). Add enough alkylating reagent to cover gel pieces. Place tubes in the dark and shake gently for 45 min.
5. Discard supernatant. Wash gel pieces with 50 mM NH_4HCO_3 for 10 min with shaking at room temperature. Discard supernatant.
6. Wash gel slices twice more using destaining reagent (10 min with shaking at room temperature and each time discarding supernatant).
7. Dehydrate gel pieces by vacuum centrifugation. Transfer the dried slices to clean 0.6 mL tubes.
8. Trypsin digestion: Add enough of a 0.2 µg/µL trypsin solution to swell the gels and incubate on ice or at 4°C for 15 min. Usually 20 µL is sufficient. The trypsin solution can be prepared and aliquoted ahead of time (20 microfuge tubes of 10 µL each) and stored in the freezer until ready.
9. Remove excess trypsin solution. Add 50 mM NH_4HCO_3/10% ACN, enough to cover the gel pieces, but not excess. Place in 37°C incubator.
10. Check pieces in 20 min, adding enough 50 mM NH_4HCO_3/10% ACN to keep pieces just covered.
11. Digest overnight (~19 h) at 37°C.

Extracting Tryptic Fragments

1. Following digestion, remove microfuge tubes from incubator. Add 1 µL of formic acid to tubes and sonicate or shake at room temperature for 15 min. Centrifuge and remove supernatant to clean tube.
2. Add enough 1.0% trifluoroacetic acid/60% ACN to cover gel pieces. Shake or sonicate for 10 min at room temperature, centrifuge, and remove supernatant and combine with supernatant from previous step.
3. Vacuum centrifuge the supernatant extract until concentrated to ~10 µL. If extract goes to dryness (not good), add 5 µL of 1.0% formic acid and vortex vigorously at room temperature.
4. Store digests at –70°C prior to analysis.

In-Solution Digestion

Note that a 100 mM Tris buffer is good to use if samples will be loaded onto a C_{18} enrichment desalting column prior to separation by nanoscale HPLC. Otherwise, use NH_4HCO_3.

Reagents

- *100 mM NH4HCO3.* 7.9 mg/mL of HPLC-grade water (or 100 mM Tris, pH 8.5).
- *8 M Urea.* Dilute 480 mg of urea in 1.0 mL of 100 mM NH_4HCO_3 solution (or 100 mM Tris buffer).
- *Reducing Reagent.* Dissolve 3 mg of DTT in 20 µL of 100 mM NH_4HCO_3 solution to make 1 M DTT (or use 100 mM Tris buffer).
- *Alkylating Reagent.* Dissolve 3.6 mg in 100 µL of NH_4HCO_3 or 100 mM Tris solution to make 200 mM IAA.
- *Trypsin Solution.* Make up a 1 mg/mL solution of trypsin in HPLC-grade water or NH_4HCO_3 (or 100 mM Tris buffer) and add 1–2 µL to each sample.

Digestion Procedure

1. *Reconstitute.* Reconstitute sample in approximately 20 µL of 8.0 M urea in a 0.5 mL microfuge tube.
2. *Reduce.* Add 1 µL of reducing reagent and mix the sample by gentle vortex. Reduce the mixture for 1 h at room temperature or in an oven at 37°C. Do not go over 37°C or the urea will react with the sample and generate carbamylated artifacts. Allow the sample to cool to room temperature.
3. *Alkylate.* Add 20 µL of alkylating reagent and alkylate for 1 h at room temperature in the dark (use aluminum foil to cover the sample). Add 4 µL of reducing reagent to consume any leftover alkylating agent (so the trypsin is not alkylated).
4. *Dilute.* Add 60 µL of NH_4HCO_3 or 100 mM Tris solution to dilute the urea before digesting it with trypsin.
5. *Digest.* Add trypsin solution in appropriate ratio (1:30) to approximate amount of protein by weight. After 1 h, add another microliter of trypsin solution. Digest 4 h to overnight at 37°C.
6. *Stop Digestion.* In the morning, or following digestion, add 1 µL of 100% formic acid to the sample. Vortex and centrifuge. Freeze sample at –70°C prior to analysis. Use a vacuum centrifuge (Speedvac) for samples with relatively low protein concentrations to maximize signal and minimize loss of sample to the tube.

Mass Spectrometric Analysis

Note that the analyses of the samples presented here were performed using a high capacity quadrupole ion trap (LC/MSD XCT Ultra [Agilent Technologies, Santa Clara, CA]). Although the quadrupole trap was historically the proteomics workhorse, hybrid quadrupole/time-of-flight (QTOF) and Orbitrap instruments are capable of rapid scanning and high resolution of fragment ions, providing a performance advantage over the quadrupole traps [6]. Parameters presented should be viewed as general guidelines that may be modified and incorporated into proteomics workflows using alternative instrumentation.

Loading Samples

Load either 5 µL of in-gel digestion sample or 10 µL of in-solution digestion into a polypropylene autosample vial.

Establishing Flow

Use a 75 µm internal diameter (ID) nanoscale column with C_{18} packing for separating peptides. Flow rate should be approximately 200–300 nL/min. Adjust following manufacturer's directions for ultra high performance liquid chromatography (UPLC) pumps.

Desalting

Samples should be desalted prior to loading on the separating column. This can be accomplished using a pipette tip desalting system (Glygen Scientific Products, Columbia, MD) or an online enrichment column with a switching valve.

Peptide Separation

Elute peptides from the separating column using a gradient from 3% to 40% buffer B (90% ACN, 0.1% formic acid). Buffer A is generally 0.1% formic acid. (*Note*: It is important to purchase high-quality formic acid and other reagents labeled as suitable for LC/MS applications.) For relatively simple mixtures, a gradient length of ~40 min is generally sufficient. Gradient times may be reduced if UPLC is employed. For more complex mixtures, extending the gradient, as well as performing steps of intact protein fraction, enables the identification of more peptides and proteins. For an in-solution digest, a 2- to 3-h gradient should be sufficient. This step will need to be optimized for different sample types.

Acquiring Spectra

Utilize data-dependent acquisition, where the most abundant peptides are analyzed a limited number of times and then placed in an exclusion list for a given amount of time (e.g., 1 min). The data shown below

were acquired by excluding the most abundant three to six peptides in a spectrum from subsequent analysis following three spectral acquisitions. Collect spectra over a mass-to-charge ratio (*m/z*) range of 350–1500 Da. (*Note*: The desired acquisition mass range will depend on the instrument employed. This mass range is appropriate when using a quadrupole ion trap.)

Database Searching

Proteomics workflows with MS/MS analysis provide extraordinarily data-rich results. A single digested band can easily result in over 10,000 fragmentation spectra. Therefore, careful analysis of the data must be performed to generate high-quality output protein lists for use in subsequent biological assays. Commonly used [19] database search algorithms include Sequest (Thermo Fisher Scientific), Spectrum Mill (Agilent Technologies), X!Tandem (The Global Proteome Machine Organization, http://www.thegpm.org/TANDEM), Mascot (Matrix Science, Inc., Boston, MA), ProteinLynx Global Server (Waters Corporation, Milford, MA), Phenyx (Geneva Bioinformatics, SA, Geneva, Switzerland), OMSSA (NCBI, http://pubchem.ncbi.nlm.nih.gov/omssa/), PEAKS (Bioinformatic Solutions, Inc., Waterloo, ON, Canada), ProteinPilot (AB Sciex, Framingham, MA), and Sequest Sorcerer (Sage-N-Research, Inc., Milpitas, CA), some of which are freely available (X!Tandem, OMSSA). Although each implementation is different, these search algorithms operate under the same general principles. These include establishment of the database and search space to be used and a statistical method of comparing experimental and theoretically generated fragmentation mass spectra, ultimately outputting a "ranked" score. Subsequently, users must determine criteria for valid score thresholds and extract potentially interesting results for follow-up.

Determining the Search Space

The search space can be limited by specifying known parameters including sample taxonomy, precursor and fragment ion mass tolerance, enzyme specificity, numbers of allowed missed enzymatic cleavages, and potential amino acid (AA) modifications. Limiting the search space significantly helps reduce false-positive results. Additionally, the selection of a more highly curated database, such as the UniProt Knowledge Base (http://www.uniprot.org/help/uniprotkb), may also help with reducing false positives. In the examples below, taxonomy was *mus musculus*, precursor ion tolerance was 1.7, and the fragment ion tolerance was 0.6. Trypsin was selected as the enzyme with two missed cleavages allowed. No AA modifications were included. The UniProtKB database was used, updated 04/2010, with 68,507 entries.

Selecting Potential Peptides to Search

First, peptide experimental mass is compared with the theoretical mass generated from an in silico digest of the protein database, and a subset of peptides with mass within the tolerance window selected in the search space are utilized for subsequent processing. Next, the fragment ion masses are compared with in silico fragmentation of the peptides (with masses within the defined mass tolerance) that passed the first criteria. Using a variety of different techniques, the algorithms assign a score or probability to the "hit" and generate a protein identification.

Interpreting the Score Data

The algorithms generate scores for all of the data within the tolerance window; however, some spectra may be falsely assigned. Generally, users select a score cutoff criterion and only consider data falling within the criterion window. Typically, a minimum of two unique peptides are required for each protein identification, although some groups have validated identifications from just one unique peptide [18]. Using reversed database searching and receiver operating curves may also help determine an appropriate cutoff. The use of data evaluation methods is not yet standardized; see Kapp and Schütz for an excellent discussion of these issues [19]. Due to the large number of spectra generated in a typical experiment it is impossible to manually inspect all of the data. However, search results from proteins that may be used in downstream biological assays should be manually validated.

Mining the Data

Proteomics data output results in long lists of proteins and scores, and it is often quite difficult to extract significant information to use for the development or testing of hypotheses. Initial insights are often obtained by searching the Gene Ontology (GO) annotations [13] for overrepresented terms. A comprehensive list of GO-based tools is available at http://www.geneontology.org/GO.tools.shtml. In the example below, we utilized The **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery (**DAVID**) v6.7 [20,21] to cluster genes with similar functions, families, or cellular locations to select a group of proteins with characteristics related to our hypothesis that were further investigated. New tools are constantly being developed and a regular review of Web sites and the literature is highly recommended.

1.3 APPLICATIONS OF THE PROTOCOLS

1.3.1 Analysis of the Acetylated Proteome of Mouse Liver in Obesity

Acetylation is now recognized as an emerging mechanism for controlling proteins mediating cellular adaptation to metabolic fuels [22] and is governed, in part, by sirtuins (SIRTs), class III NAD^+-dependent histone deacetylases (HDACs) that regulate lipid and glucose metabolism in liver during fasting and aging. However, whether acetylation or SIRTs play a pathogenic role in fuel metabolism under conditions of obesity is unknown.

In the present study, 5- to 6-week-old male C57BL/6 SVJ mice were fed a high-fat diet (45 kcal% fat) or a standard chow diet for 16 weeks, and then were sacrificed by pentobarbital overdose following treatment. Livers were harvested immediately from anesthetized mice and snap frozen at –70°C in liquid nitrogen before analysis. Utilizing the protocol provided for preparation of tissue homogenates (~200 mg liver tissue, 10 µg primary antibody, 50 µL packed beads), we prepared nuclear/mitochondrial and cytoplasmic fractions and enriched for lysine acetylated proteins using an anti-acetyllysine polyclonal antibody. Complexed proteins were eluted from the beads by boiling in Laemmli buffer, separated using a 10% polyacrylamide gel, fluorescently stained with Lava Purple and imaged.

Labeled gel bands in Figure 1.2 that differed between mice fed normal chow (control) and a high-fat diet were excised, digested, and analyzed by MS/MS using a 40-min gradient of acetonitrile. Data were analyzed by LABKEY (http://www.labkey.com), an interface utilizing the X!Tandem search algorithm and elements of the Trans-Proteomics Pipeline, including PeptideProphet and Protein Prophet (these are freely available software packages developed by the Seattle Proteome Center within the Institute for Systems Biology, Seattle, WA; http://www.proteomecenter.org/software.php). Receiver operating characteristic (ROC) curves were generated to determine peptide and protein probability cutoffs providing ~ 1% false discovery rate (typically ~0.9–1.0). Following the search, results were exported to Microsoft Excel (cytoplasmic sample, 1703 proteins; nuclear/mitochondrial sample, 307 proteins). Proteins that were observed in multiple samples were assumed to be non-specifically binding and were excluded from further analysis, as were proteins identified by one unique peptide observed only once. Potential isoforms and putative proteins were excluded as well, reducing the

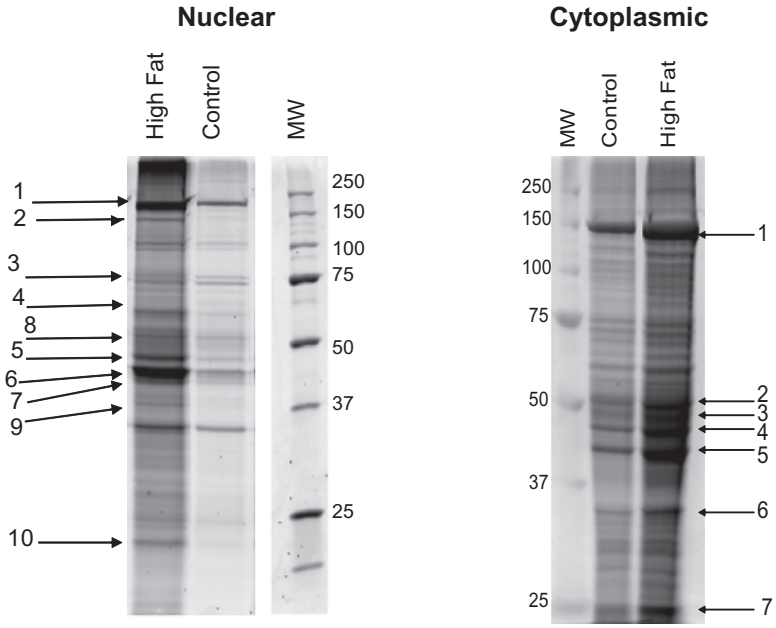

FIGURE 1.2 Immunoaffinity purification of acetylated proteins reveals increased acetylation in proteins from livers of mice fed a high-fat diet. C57BL/6 SVJ mice were fed either normal chow or a high-fat diet (45% fat) for 16 weeks and fasted overnight before sacrifice ($n = 3$ per group). Samples were extracted from liver as described in the text. Lysine-acetylated liver proteins were immunoaffinity purified from either cytoplasmic or mitochondrial/nuclear extracts, then separated by 1D gel electrophoresis using a 10% SDS-PAGE gel, visualized with Lava Purple and imaged with a Typhoon 9600 imager (GE Healthcare). Labeled bands with differential staining were excised, digested, and identified by MS/MS and database searching.

number of identifications to 227 for the cytoplasmic sample and 156 for the nuclear/mitochondrial sample. Proteins were sorted by calculated molecular weight (MW) and MWs differing from the median by more than 20% were assumed to be incorrect identifications. Finally, protein identifications with at least 5% AA coverage of the protein [18] were retained in the analysis.

As summarized in Tables 1.1 and 1.2, we identified 148 proteins from the cytoplasmic sample and 75 proteins from the nuclear/mitochondrial sample with high confidence. Each band typically contained a number of proteins that were identified. Bands excised from the cytoplasmic samples were larger to incorporate the apparently high number of closely comigrating proteins; therefore, more proteins were typically identified in those bands than in the nuclear/mitochondrial extract bands. An examination of proteins from the latter sample suggests that a high-fat diet led to hyperacetylation of proteins involved in gluconeogenesis, mitochondrial oxidative metabolism, methionine metabolism, liver injury, and the endoplasmic reticulum (ER) stress response. Not shown here, we observed that in mice lacking SIRT3, a sirtuin localized to the mitochondrion, a high-fat diet further increased the acetylation status of liver proteins compared with high-fat diet-fed wild-type mice and was associated with the disruption of mitochondrial oxidative phosphorylation complexes II, III, and V. Our results suggest that hyperacetylation of mitochondrial proteins may play a pivotal role in mechanisms regulating high-fat diet-induced mitochondrial dysfunction in livers of obese mice.

1.3.2 Identification of a Major Histocompatibility Complex Class II (MHC-II)-Complexed Death Transducer

MHC-II is primarily known to function in the presentation of antigenic peptides to T lymphocytes. However, these molecules have also been observed to transduce signals, leading to either cell activation or apoptotic death. The short, cytoplasmic tails of the two transmembrane proteins comprising MHC-II are not required for induction of apoptosis [23], therefore a protein complexing with MHC-II is likely important in mediating death signaling.

In this study, K46 cells (5×10^8) were lysed in cell lysis buffer and immunoaffinity purified, as described above. Shown in Figure 1.3 is a gel separation of proteins from a representative preparation, suggesting the presence of many potentially complexed proteins. Bound proteins were eluted from the beads with citric acid and pH was neutralized with Tris. Following in-solution digestion, ~30% of the sample was analyzed by nanoscale liquid chromatography–tandem mass spectrometry (nano-LC/MS/MS) using a 145-min gradient of acetonitrile. Data from five separate immunoaffinity purifications were analyzed by LABKEY and proteins with probabilities yielding less than 3% false discovery rates were included. Initially, a total of 2514 protein identifications were obtained from the combined five analyses. Since the data files were obtained from immunoaffinity purification of both MHC-II and a complexed protein (MPYS, subsequently termed TM173), we restricted the list to proteins that were observed in at least three runs, including one run from the MHC-II IP and one run from the MPYS IP. Identifications based on one hit were removed, as were duplicates, and putative uncharacterized proteins, resulting in a list of 237 proteins, of which 83 were isoforms of MHC-II. The list of accession numbers was then clustered using functional annotation tools in **DAVID** [21] and protein groups present are summarized in Figure 1.4. A recent review of commonly observed protein identifications [24] suggests that overrepresentation of actin, tubulin, and adenosine triphosphate (ATP) synthase isoforms may result from nonspecific binding or represent common cellular stress responses.

Since MHC-II is a transmembrane protein, we concentrated further analysis on the group of transmembrane proteins identified in this study, summarized in Table 1.3. Of these proteins, only two were potentially multispanning proteins—CD20 and TM173. Fragmentation mass spectra leading to the identification of these proteins are plotted in Figures 1.5 and 1.6, respectively. CD20 is known to associate with MHC-II but was not thought to be a likely candidate for transducing death signals [23]. As described in the UniProt Knowledge Base, the isotopes of transmembrane emp24 are single-pass transmembrane proteins with cytoplasmic domains not well described, and CDGSH iron sulfur protein, implicated in autophagy, has a short cytoplasmic domain. Ribophorin and solute carrier protein 3 are both single membrane-spanning proteins with longer cytoplasmic domains (150 and 75 AAs, respectively). MHC-II is known to associate with tetraspanins, therefore we selected TM173, a potential tetraspanin with a 204-amino-acid-long cytoplasmic domain, for subsequent study.

As we previously described [23], TM173 is a membrane protein with an immunoreceptor tyrosine-based inhibitory motif (ITIM) contained in its cytoplasmic tail. We demonstrated that TM173 becomes phosphorylated upon cross-linking with MHC-II, recruits the inhibitory signaling effectors Src homology region 2 domain-containing phosphatase-1 (SHP-1) and Phosphatidylinositol-3,4,5-trisphosphate 5-phosphatase 1 (SHIP), reduces calcium mobilization, and negatively regulates cell growth. Confirmation of the important role of TM173 in MHC-II-mediated cell death was obtained in cells with TM173 knocked down.

TABLE 1.1 Hyperacetylated Cytoplasmic Proteins from Livers of Mice Fed a High-Fat Diet

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
1	Q8CIE6	Coatomer subunit alpha	138,446	12	1.0000	6	4
1	Q9Z1Q9	Valyl-tRNA synthetase	140,084	9	1.0000	9	4
1	Q8VDN2	Sodium/potassium-transporting ATPase subunit alpha-1	112,510	9	1.0000	6	3
1	Q8VDJ3	Vigilin	141,612	6	1.0000	2	1
1	Q9JKR6	Hypoxia upregulated protein 1	107,606	6	1.0000	8	3
1	Q05920	Pyruvate carboxylase, mitochondrial	127,428	6	1.0000	3	3
2	P17563	Selenium-binding protein 1	525,14	45	1.0000	31	11
2	O35728	Cytochrome P450 4A14	58,238	19	1.0000	12	5
2	P24549	Retinal dehydrogenase 1	54,337	16	1.0000	13	5
2	Q8VCM7	Fibrinogen gamma chain	46,671	14	0.9997	4	1
3	O54734	Dolichyl-diphosphooligosaccharide–protein glycosyltransferase 48 kDa subunit	46,044	36	1.0000	27	10
3	P17182	Alpha-enolase	47,010	26	1.0000	29	7
3	Q8BWF0	Succinate-semialdehyde dehydrogenase, mitochondrial	52,012	21	1.0000	7	6
3	P53395	Lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex, mitochondrial	46,188	20	1.0000	5	4
3	P70333	Heterogeneous nuclear ribonucleoprotein H2	49,280	19	1.0000	3	2
3	O35737	Heterogeneous nuclear ribonucleoprotein H	49,199	19	1.0000	3	2
3	P12790	Cytochrome P450 2B9	55,760	17	1.0000	10	6
3	P40124	Adenylyl cyclase-associated protein 1	51,444	14	1.0000	7	4
3	Q9D2G2	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	41,470	12	1.0000	4	3
3	Q9JLJ2	4-trimethylaminobutyraldehyde dehydrogenase	53,384	9	1.0000	4	2
3	Q62465	Synaptic vesicle membrane protein VAT-1 homolog	42,965	8	0.9998	2	1
3	P50431	Serine hydroxymethyltransferase, cytosolic	52,585	7	1.0000	2	2
3	P61922	4-aminobutyrate aminotransferase, mitochondrial	53,220	6	0.9996	2	2
3	Q9Z1N5	Spliceosome RNA helicase BAT1	48,904	6	0.9999	2	2
4	Q9QZ85	Interferon-inducible GTPase 1	47,572	44	1.0000	22	7
4	P09411	Phosphoglycerate kinase 1	44,419	43	1.0000	21	8
4	Q61598	Rab GDP dissociation inhibitor beta	50,537	33	1.0000	14	7
4	P43883	Perilipin-2	46,664	30	1.0000	11	7
4	P15105	Glutamine synthetase	41,988	29	1.0000	38	9
4	Q8VCZ9	Probable proline dehydrogenase 2	50,723	28	1.0000	5	3
4	P63037	DnaJ homolog subfamily A member 1	44,552	24	1.0000	13	5

(*Continued*)

TABLE 1.1 (*Continued*)

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
4	Q91WN4	Kynurenine 3-monooxygenase	54,532	22	1.0000	9	7
4	Q91YT0	NADH dehydrogenase (ubiquinone) flavoprotein 1, mitochondrial	48,626	20	1.0000	4	3
4	P26150	3 beta-hydroxysteroid dehydrogenase/Delta 5–>4-isomerase type 3	41,900	20	1.0000	8	4
4	O88986	2-amino-3-ketobutyrate coenzyme A ligase, mitochondrial	42,850	14	1.0000	9	3
4	Q8R0F9	SEC14-like protein 4	46,053	14	1.0000	4	2
4	Q91VC3	Eukaryotic initiation factor 4A-III	46,709	13	1.0000	7	2
4	Q922E4	Ethanolamine-phosphate cytidylyltransferase	45,235	12	0.9956	2	1
4	O88844	Isocitrate dehydrogenase (NADP) cytoplasmic	46,660	11	0.9999	2	2
4	A2AKK5	Acyl-coenzyme A amino acid N-acyltransferase 1	46,070	8	1.0000	2	2
5	Q99PG0	Arylacetamide deacetylase	45,119	49	1.0000	23	11
5	Q60759	Glutaryl-CoA dehydrogenase, mitochondrial	43,737	47	1.0000	34	9
5	Q9Z2I9	Succinyl-CoA ligase (ADP-forming) subunit beta, mitochondrial	44,422	36	1.0000	27	10
5	Q8QZT1	Acetyl-CoA acetyltransferase, mitochondrial	41,414	31	1.0000	25	7
5	P08249	Malate dehydrogenase, mitochondrial	36,367	27	1.0000	17	6
5	P49429	4-hydroxyphenylpyruvate dioxygenase	44,923	27	1.0000	13	6
5	Q9QYG0	Protein NDRG2	40,658	25	1.0000	5	3
5	P51174	Long-chain specific acyl-CoA dehydrogenase, mitochondrial	44,627	21	1.0000	22	7
5	P45952	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	43,593	21	1.0000	15	5
5	Q62087	Serum paraoxonase/lactonase 3	39,351	19	1.0000	10	3
5	P06151	L-lactate dehydrogenase A chain	43,292	19	1.0000	2	2
5	Q9CZU6	Citrate synthase, mitochondrial	49,014	16	1.0000	4	3
5	P52430	Serum paraoxonase/arylesterase 1	39,434	15	1.0000	5	3
5	Q9WUZ9	Ectonucleoside triphosphate diphosphohydrolase 5	45,262	13	0.9895	2	1
5	Q91YP0	L-2-hydroxyglutarate dehydrogenase, mitochondrial	45,665	12	1.0000	2	2
5	Q924Y0	Gamma-butyrobetaine dioxygenase	44,699	12	1.0000	3	2
5	Q71RI9	Kynurenine–oxoglutarate transaminase 3	51,126	10	1.0000	2	2
5	Q99JI4	26S proteasome non-ATPase regulatory subunit 6	45,405	10	1.0000	3	3
5	P35486	Pyruvate dehydrogenase E1 component subunit alpha, somatic form, mitochondrial	40,181	10	1.0000	11	3
5	P14152	Malate dehydrogenase, cytoplasmic	38,545	8	1.0000	2	2

TABLE 1.1 (*Continued*)

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
5	P42669	Transcriptional activator protein Pur-alpha	34,884	8	0.9795	2	1
5	P50247	Adenosylhomocysteinase	47,557	6	0.9999	2	1
6	P70694	Estradiol 17 beta-dehydrogenase 5	37,048	67	1.0000	62	12
6	Q8VBT2	L-serine dehydratase/L-threonine deaminase	34,462	48	1.0000	16	7
6	P52196	Thiosulfate sulfurtransferase	32,233	45	1.0000	48	12
6	Q9CW42	MOSC domain-containing protein 1, mitochondrial	33,991	35	1.0000	18	6
6	Q8VCX1	3-oxo-5-beta-steroid 4-dehydrogenase	37,290	32	1.0000	11	6
6	Q922Q1	MOSC domain-containing protein 2, mitochondrial	34,831	31	1.0000	37	6
6	P47962	60S ribosomal protein L5	37,248	30	1.0000	13	7
6	P57776-2	Isoform 2 of elongation factor 1-delta	33,335	29	1.0000	7	3
6	P57776-3	Isoform 3 of elongation factor 1-delta	31,162	29	1.0000	7	3
6	Q9JMD3	Phosphatidylcholine transfer protein (PCTP)-like protein	32,951	26	1.0000	13	5
6	P62880	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	37,200	25	1.0000	6	4
6	P50172	Corticosteroid 11-beta-dehydrogenase isozyme 1	34,269	24	1.0000	10	3
6	Q9WUM5	Succinyl-CoA ligase (GDP-forming) subunit alpha, mitochondrial	32,098	24	1.0000	17	5
6	Q9JII6	Alcohol dehydrogenase (NADP+)	36,456	23	1.0000	5	4
6	Q9D051	Pyruvate dehydrogenase E1 component subunit beta, mitochondrial	35,768	23	1.0000	17	4
6	Q6P3A8	2-oxoisovalerate dehydrogenase subunit beta, mitochondrial	37,817	22	1.0000	3	3
6	P47911	60S ribosomal protein L6	34,216	21	1.0000	6	3
6	Q91Z53	Glyoxylate reductase/ hydroxypyruvate reductase	35,329	19	1.0000	3	2
6	Q9D5T0	ATPase family AAA domain-containing protein 1	40,744	19	1.0000	2	2
6	Q3UNZ8	Quinone oxidoreductase-like protein 2	37,809	18	1.0000	2	2
6	Q91Y97	Fructose-bisphosphate aldolase B	39,376	17	1.0000	7	4
6	Q91XE0	Glycine N-acyltransferase	34,098	16	1.0000	10	3
6	Q9CXR1	Dehydrogenase/reductase SDR family member 7	34,876	14	1.0000	2	2
6	Q8VCH0	3-ketoacyl-CoA thiolase B, peroxisomal	41,284	14	1.0000	6	3
6	Q921H8	3-ketoacyl-CoA thiolase A, peroxisomal	41,242	14	1.0000	6	3
6	Q99MZ7	Peroxisomal trans-2-enoyl-CoA reductase	32,410	14	1.0000	10	3

(*Continued*)

TABLE 1.1 (*Continued*)

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
6	Q9DBH5	Vesicular integral-membrane protein VIP36	35,455	13	0.9918	3	1
6	Q9CY50	Translocon-associated protein subunit alpha	29,689	12	1.0000	7	3
6	Q8CDN6	Thioredoxin-like protein 1	32,106	11	1.0000	3	2
6	P48758	Carbonyl reductase (NADPH) 1	33,378	11	1.0000	3	2
6	Q9D819	Inorganic pyrophosphatase	32,667	11	1.0000	3	2
6	Q9DCX8	Iodotyrosine dehalogenase 1	30,257	9	1.0000	3	2
6	P62874	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	37,246	9	1.0000	4	2
6	P51658	Estradiol 17-beta-dehydrogenase 2	30,510	8	1.0000	4	2
6	O55125	Protein NipSnap homolog 1	33,363	8	1.0000	2	1
6	Q6ZWX6	Eukaryotic translation initiation factor 2 subunit 1	35,977	8	0.9931	3	1
6	P11725	Ornithine carbamoyltransferase, mitochondrial	36,122	6	0.9999	2	2
7	Q9DCX2	ATP synthase subunit d, mitochondrial	20,893	63	1.0000	13	7
7	P35700	Peroxiredoxin-1	22,176	54	1.0000	23	7
7	Q9WVL0	Maleylacetoacetate isomerase	25,295	53	1.0000	41	9
7	P10649	Glutathione S-transferase Mu 1	25,839	48	1.0000	23	6
7	Q9DCM2	Glutathione S-transferase kappa 1	22,165	48	1.0000	20	8
7	P11352	Glutathione peroxidase 1	22,179	46	1.0000	6	3
7	Q9CQQ7	ATP synthase subunit b, mitochondrial	24,765	44	1.0000	22	9
7	Q9R257	Heme-binding protein 1	25,494	41	1.0000	2	2
7	Q9D6Y7	Peptide methionine sulfoxide reductase	22,187	40	1.0000	7	4
7	Q9DB20	ATP synthase subunit O, mitochondrial	25,988	32	1.0000	17	4
7	Q9CQC9	GTP-binding protein SAR1b	22,382	31	1.0000	9	3
7	Q64471	Glutathione S-transferase theta-1	27,245	31	1.0000	20	6
7	P35278	Ras-related protein Rab-5C	23,281	31	1.0000	6	4
7	P30115	Glutathione S-transferase A3	25,229	31	1.0000	21	4
7	Q9D1G1	Ras-related protein Rab-1B	24,015	29	1.0000	3	2
7	P46638	Ras-related protein Rab-11B	24,030	25	1.0000	7	4
7	P62492	Ras-related protein Rab-11A	23,907	25	1.0000	7	4
7	Q9CZM2	60S ribosomal protein L15	21,493	25	1.0000	4	2
7	P19157	Glutathione S-transferase P 1	23,478	23	1.0000	6	2
7	Q923D2	Flavin reductase	22,066	23	1.0000	3	3
7	P51410	60S ribosomal protein L9	21,881	23	1.0000	5	2
7	P62082	40S ribosomal protein S7	22,127	22	1.0000	8	2
7	Q9DCS9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex subunit 10	22,687	22	1.0000	5	2
7	Q80W21	Glutathione S-transferase Mu 7	25,710	22	1.0000	8	2
7	P35980	60S ribosomal protein L18	21,513	22	1.0000	7	2
7	Q91V41	Ras-related protein Rab-14	23,766	20	1.0000	6	3
7	Q9DBP5	UMP-CMP kinase	21,006	19	1.0000	3	3
7	Q9CQA3	Succinate dehydrogenase (ubiquinone) iron-sulfur subunit, mitochondrial	28,770	17	1.0000	3	2
7	Q9DCW4	Electron transfer flavoprotein subunit beta	25,564	17	0.9998	2	1

TABLE 1.1 (*Continued*)

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
7	P62827	GTP-binding nuclear protein Ran	24,292	17	1.0000	2	1
7	Q6ZWN5	40S ribosomal protein S9	22,460	16	1.0000	4	2
7	Q9WTP7	GTP:AMP (adenosine monophosphate) phosphotransferase mitochondrial	21,053	16	1.0000	3	2
7	Q9CQD1	Ras-related protein Rab-5A	23,467	16	1.0000	2	2
7	Q99PT1	Rho GDP-dissociation inhibitor 1	27,492	16	0.9996	2	1
7	Q9DCS2	UPF0585 protein C16orf13 homolog	25,573	16	1.0000	2	1
7	Q00623	Apolipoprotein A-I	27,922	16	1.0000	6	2
7	P24472	Glutathione S-transferase A4	24,275	15	0.9999	3	2
7	Q61133	Glutathione S-transferase theta-2	27,549	12	1.0000	5	2
7	Q9CPU0	Lactoylglutathione lyase	24,074	11	0.9948	4	1
7	Q02013	Aquaporin-1	28,662	11	1.0000	3	2
7	Q91VT4	Carbonyl reductase family member 4	23,333	10	0.9952	3	1
7	P20108	Thioredoxin-dependent peroxide reductase, mitochondrial	21,565	10	1.0000	3	2
7	P84099	60S ribosomal protein L19	23,466	9	1.0000	5	1
7	P10605	Cathepsin B	23,276	9	0.9957	3	1
7	Q99KF1	Transmembrane emp24 domain-containing protein 9	23,432.91	7	0.9918	3	1
7	Q922B1	MACRO domain-containing protein 1	25,415	7	0.9845	3	1
7	P19253	60S ribosomal protein L13a	27,578	6	0.9957	4	1
7	Q9CR57	60S ribosomal protein L14	22,210	6	0.9957	3	1
7	Q64105	Sepiapterin reductase	20,678	6	0.9673	2	1

[a] Excised band labeled in Figure 1.2.
[b] UniProtKB identifier number.
[c] Number of AAs identified from peptides divided by total number of AAs in the identified protein and multiplied by 100.
[d] Probability value calculated by the ProteinProphet module of the LABKEY software algorithm.
[e] Total number of fragmentation spectra identified as belonging to a protein.
[f] Number of nonredundant fragmentation spectra identified as belonging to a protein.

1.4 CONCLUSION

The use of immunoaffinity purification as a sample preparation method for downstream proteomics applications is emerging for applications involving identification of proteins in functional complexes and global identification of modified proteins. This approach allows investigators to use proteomics approaches to address hypothesis-driven research questions. Although a detailed protocol was presented here, it should be noted that each sample type and antibody will require optimization of the protocol—different buffers may work better for lysis, antibodies may bind more or less strongly—therefore, the information here is presented only as a guideline or starting place.

ACKNOWLEDGMENTS

This work was supported by NIH Grants DK59767 and P30-DK48520 (to J.E.F.), a Pilot & Feasibility Award from the University of Colorado, Center for Human Nutrition P30-DK048520-09 (K.J.), and NIH Grant AI020519-22 (to J.C.C.). K.J. thanks W.H. MacDonald at Vanderbilt University for review of the manuscript and extraordinarily helpful discussions.

TABLE 1.2 Hyperacetylated Proteins from a Nuclear/Mitochondrial Extract of Livers from Mice Fed a High-Fat Diet

Band[a]	Identifier[b]	Protein Description	MW (kDa)	AA Coverage[c] (%)	Protein Prophet Probability[d]	Total Peptides[e]	Unique Peptides[f]
1	sp|Q8C196	Carbamoyl-phosphate synthase (ammonia), mitochondrial	164,618	38	1.0000	137	37
2	sp|Q05920	Pyruvate carboxylase, mitochondrial	129,685	23	1.0000	43	17
3	sp|P20029	78 kDa glucose-regulated protein	72,422	20	1.0000	19	6
3	sp|P38647	Stress-70 protein, mitochondrial	73,528	18	1.0000	19	7
3	tr|Q7TSZ0	Heat shock protein 9	73,461	18	1.0000	19	7
3	sp|P63017	Heat shock cognate 71 kDa protein	70,871	6	1.0000	5	2
3	tr|Q504P4	Hspa8 protein	68,779	6	1.0000	5	2
4	sp|P63038	60 kDa heat shock protein, mitochondrial	60,955	17	1.0000	11	5
4	tr|Q8C6E3	Catalase	59,835	15	1.0000	13	5
4	sp|P32020	Nonspecific lipid-transfer protein	59,126	14	1.0000	12	4
4	sp|P34914	Epoxide hydrolase 2	62,515	13	1.0000	18	7
4	sp|Q63880	Liver carboxylesterase 31	63,318	7	1.0000	5	2
4	tr|B1AXW8	Aldehyde dehydrogenase 4 family, member A1	61,841	7	1.0000	5	3
4	sp|Q8VC30	Bifunctional ATP-dependent dihydroxyacetone kinase/flavin adenine dinucleotide (FAD)–AMP lyase (cyclizing)	59,691	7	0.9999	3	1
5	sp|O09173	Homogentisate 1,2-dioxygenase	49,990	18	1.0000	7	2
5	sp|P05784	Keratin, type I cytoskeletal 18	47,538	16	1.0000	7	4
5	sp|P56480	ATP synthase subunit beta, mitochondrial	56,300	16	1.0000	12	6
5	sp|P10126	Elongation factor 1-alpha 1	50,114	11	1.0000	17	4
5	sp|Q99JY0	Trifunctional enzyme subunit beta, mitochondrial	51,386	11	1.0000	5	4
5	sp|Q9D8N0	Elongation factor 1-gamma	50,061	9	1.0000	3	2
5	sp|P97807	Fumarate hydratase, mitochondrial	54,371	7	1.0000	7	2
6	sp|P16460	Argininosuccinate synthase	46,584	19	1.0000	16	6
7	tr|Q3TIT9	Acetyl-Coenzyme A acyltransferase 2 (Mitochondrial 3-oxoacyl-Coenzyme A thiolase), isoform CRA_k	41,830	6	0.9993	3	1
8	sp|Q03265	ATP synthase subunit alpha, mitochondrial	59,753	36	1.0000	62	12
8	sp|Q63836	Selenium-binding protein 2	52,610	9	1.0000	5	2
8	sp|P17563	Selenium-binding protein 1	52,514	9	1.0000	5	2
9	sp|Q91Y97	Fructose-bisphosphate aldolase B	39,507	19	1.0000	8	4
9	sp|Q8VCH0	3-ketoacyl-CoA thiolase B, peroxisomal	43,995	11	1.0000	3	2
9	sp|Q921H8	3-ketoacyl-CoA thiolase A, peroxisomal	43,953	11	1.0000	3	2
10	sp|P35700	Peroxiredoxin-1	22,176	37	1.0000	6	4
10	tr|Q5RJH8	Glutathione peroxidase	22,292	23	1.0000	8	3
10	sp|P11352	Glutathione peroxidase 1	22,179	23	1.0000	8	3
10	tr|Q58EV2	Apoa1 protein	23,022	7	0.9953	2	1

[a] Excised band labeled in Figure 1.2.
[b] UniProtKB identifier number.
[c] Number of AAs identified from peptides divided by total number of AAs in the identified protein and multiplied by 100.
[d] Probability value calculated by the ProteinProphet module of the LABKEY software algorithm.
[e] Total number of fragmentation spectra identified as belonging to a protein.
[f] Number of nonredundant fragmentation spectra identified as belonging to a protein.

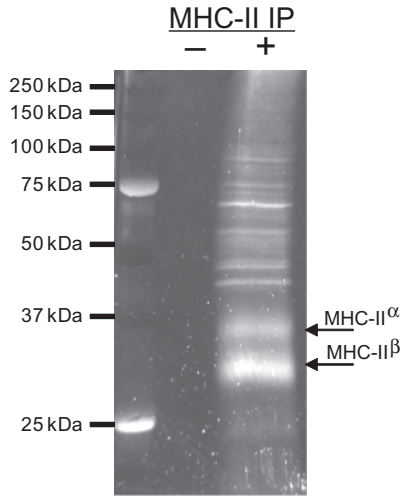

FIGURE 1.3 Immunoaffinity purification of MHC-II reveals a number of potentially complexed proteins. MHC-II-associated proteins were enriched from K46 cells lysates and analyzed using nano-LC/MS/MS. Incubation of protein A beads in the absence of primary antibody was employed as a negative control. The gel shown is a representative result from three similar experiments.

TABLE 1.3 Transmembrane Proteins Potentially Complexed with MHC-II

UniProtKB Identifier	Protein Description
4F2_MOUSE	Solute carrier family 3 (activators of dibasic and neutral AA transport), member 2
CD20_MOUSE	Membrane-spanning 4-domains, subfamily A, member 1
CISD2_MOUSE	CDGSH iron sulfur domain 2
Q3TVJ8_MOUSE	Signal sequence receptor, delta
Q6PDC2_MOUSE	Transmembrane emp24 protein transport domain containing 9
RPN1_MOUSE	Ribophorin I
TM173_MOUSE	Transmembrane protein 173
TMED4_MOUSE	Transmembrane emp24 protein transport domain containing 4
TMEDA_MOUSE	Transmembrane emp24-like trafficking protein 10 (yeast); predicted gene 4024

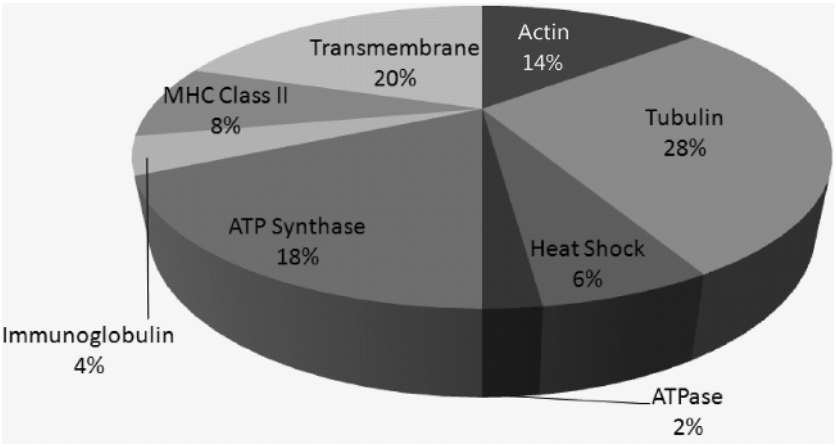

FIGURE 1.4 Functional classification of proteins identified in the MHC-II complex. The DAVID algorithm was used to assess overrepresentation of GO terms for the identified proteins. The Similarity Term Overlap was set to 3, the Similarity Threshold was 0.2, the Group Membership was 5, and the Multiple Linkage Threshold value was 0.5. Following data reduction, actins, tubulins, and heat shock proteins (25) were clustered into a group with an enrichment score of 13.7, ATPases (9) were clustered with an enrichment score of 5.7, and transmembrane proteins (17) were clustered with an enrichment score of 1.4. GAPDH was left unclustered.

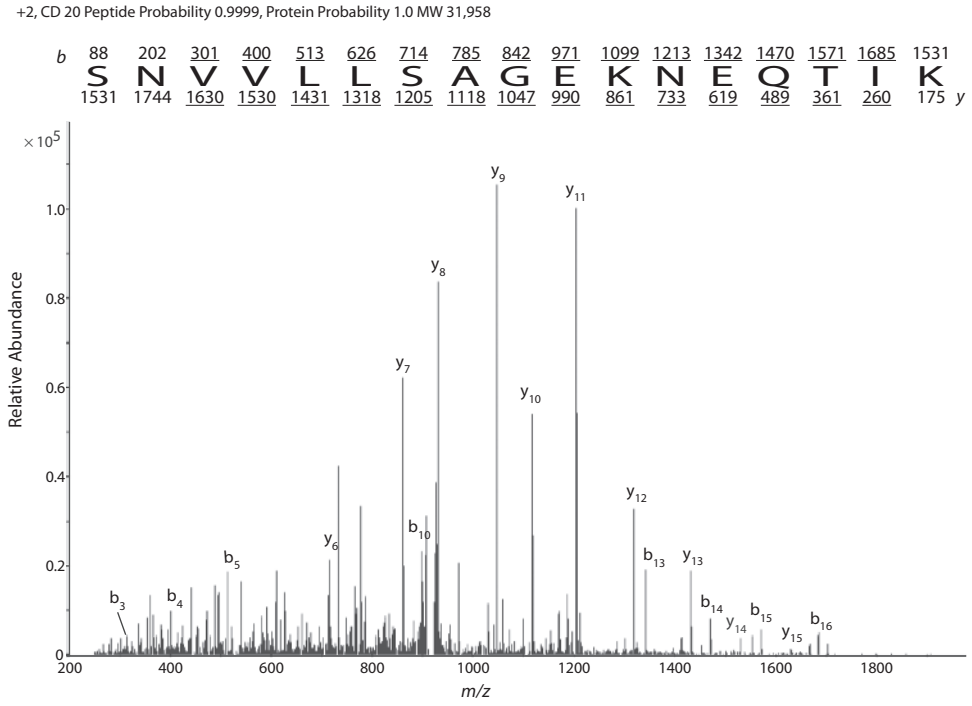

FIGURE 1.5 Fragmentation mass spectrum and sequence analysis of peptide identified from CD20. An almost complete series of singly charged b- and y-type ions were used to unambiguously identify this doubly charged peptide SNVVLLSAGEKNEQTIK from CD20 (MW 31,958 Da). The PeptideProphet probability was 0.9999 and the ProteinProphet probability was 1.000 for this identification.

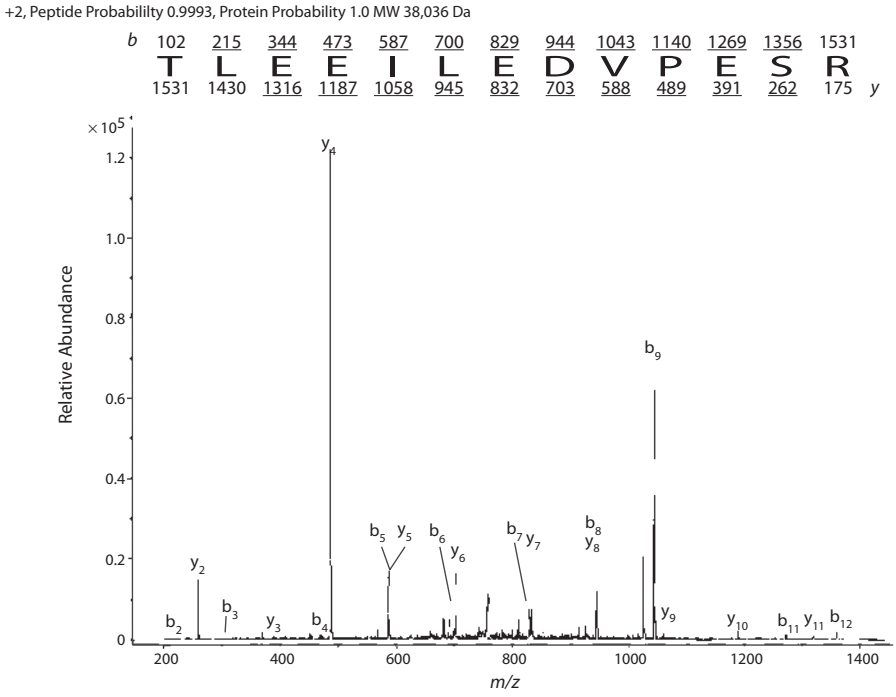

FIGURE 1.6 Fragmentation mass spectrum and sequence analysis of peptide identified from TM173. An almost complete series of singly charged b- and y-type ions were used to unambiguously identify this doubly charged peptide TLEEILEDVPESR from TM173 (MW 38,036 Da). The PeptideProphet probability was 0.9993 and the ProteinProphet probability was 1.000 for this identification.

REFERENCES

1. Chien, C.T., Bartel, P.L., Sternglanz, R., Fields, S. (1991) The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proceedings of the National Academy of Sciences of the United States of America*, *88*, 9578–9582.
2. Suter, B., Kittanakom, S., Stagljar, I. (2008) Interactive proteomics: what lies ahead? *BioTechniques*, *44*, 681–691.
3. Gavin, A.-C., Bosche, M., Krause, R., Grandi, P., et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, *415*, 141–147.
4. Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., et al. (2002) Systematic identification of protein complexes in Saccharomyces cerevisiae by mass spectrometry. *Nature*, *415*, 180–183.
5. Ewing, R.M., Chu, P., Elisma, F., Li, H., et al. (2007) Large-scale mapping of human protein-protein interactions by mass spectrometry. *Molecular Systems Biology*, *3*, 2–17.
6. Choudhary, C., Mann, M. (2010) Decoding signaling networks by mass spectrometry-based proteomics. *Nature Reviews Molecular Cell Biology*, *11*, 427–439.
7. Yang, W., Steen, H., Freeman, M.R. (2008) Proteomic approaches to the analysis of multiprotein signaling complexes. *Proteomics*, *8*, 832–851.
8. Ashman, K., López Villar, E. (2009) Phosphoproteomics and cancer research. *Clinical & Translational Oncology*, *11*, 356–362.
9. Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., et al. (2009) Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*, *325*, 834–840.
10. Schirle, M., Heurtier, M.-A., Kuster, B. (2003) Profiling core proteomes of human cell lines by one-dimensional PAGE and liquid chromatography-tandem mass spectrometry. *Molecular & Cellular Proteomics*, *2*, 1297–1305.
11. Kumar, C., Mann, M. (2009) Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Letters*, *583*, 1703–1712.
12. Hubbard, S.J., Jones, A.R., Elias, J.E., Gygi, S.P. (2010) *Proteome Bioinformatics*. Methods in Molecular Biology, Vol. *604*. London: Humana Press.
13. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., et al. (2000) Gene Ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, *25*, 25–29.
14. Okuda, S., Yamada, T., Hamajima, M., Itoh, M., et al. (2008) KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Research*, *36*, W423–W426.
15. Carvalho, P., Fischer, J., Chen, E., Domont, G., et al. (2009) GO Explorer: a gene-ontology tool to aid in the interpretation of shotgun proteomics data. *Proteome Science*, *7*, 6.
16. Carvalho, P.C., Fischer, J.S.G., Chen, E.I., Yates, J.R., et al. (2008) PatternLab for proteomics: a tool for differential shotgun proteomics. *BMC Bioinformatics*, *9*, 316.
17. Malovannaya, A., Li, Y., Bulynko, Y., Jung, S.Y., et al. (2010) Streamlined analysis schema for high-throughput identification of endogenous protein complexes. *Proceedings of the National Academy of Sciences*, *107*, 2431–2436.
18. Moresco, J.J., Carvalho, P.C., Yates, J.R. III (2010) Identifying components of protein complexes in *C. elegans* using co-immunoprecipitation and mass spectrometry. *Journal of Proteomics*, *73*, 2198–2204.
19. Kapp, E., Schütz, F. (2007) Overview of tandem mass spectrometry (MS/MS) database search algorithms. *Current Protocols in Protein Science*, *49*, 25.2.1–25.2.19.
20. Dennis, G., Sherman, B., Hosack, D., Yang, J., et al. (2003) DAVID: database for annotation, visualization, and integrated discovery. *Genome Biology*, *4*, P3.
21. Huang, D.W., Sherman, B.T., Lempicki, R.A. (2008) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, *4*, 44–57.
22. Bao, J., Sack, M. (2010) Protein deacetylation by sirtuins: delineating a post-translational regulatory program responsive to nutrient and redox stressors. *Cellular and Molecular Life Sciences*, *67*, 3073–3087.
23. Jin, L., Waterman, P.M., Jonscher, K.R., Short, C.M., et al. (2008) MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex class II and mediates transduction of apoptotic signals. *Molecular and Cellular Biology*, *28*, 5014–5026.
24. Petrak, J., Ivanek, R., Toman, O., Cmejla, R., et al. (2008) Déjà vu in proteomics. A hit parade of repeatedly identified differentially expressed proteins. *Proteomics*, *8*, 1744–1749.